



Universidad de Murcia

Departamento de Ingeniería de la Información y las Comunicaciones

DEFINICIÓN DE UN MODELO DE GESTIÓN DE LAS NOCIONES
DE CONFIANZA Y REPUTACIÓN ENTRE AGENTES. ENFOQUE
BASADO EN LA SIMILITUD ENTRE TAREAS

Tesis Doctoral

Presentada por:

Alberto Caballero Martínez

Dirigida por:

Dr. Juan A. Botía Blaya

Dr. Antonio F. Gómez Skarmeta

Febrero de 2008



Universidad de Murcia

D. Juan A. Botía Blaya, profesor Titular de Escuela Universitaria del Área de Ciencias de la Computación e Inteligencia Artificial, y D. Antonio F. Gómez Skarmeta, Profesor Titular de Universidad del Área de Ingeniería Telemática, en el Departamento de Ingeniería de la Información y las Comunicaciones de la Universidad de Murcia, AUTORIZAN:

La presentación de la Tesis Doctoral titulada "DEFINICIÓN DE UN MODELO DE GESTIÓN DE LAS NOCIONES DE CONFIANZA Y REPUTACIÓN ENTRE AGENTES. ENFOQUE BASADO EN LA SIMILITUD ENTRE TAREAS", realizada por D. Alberto Caballero Martínez, bajo nuestra inmediata dirección y supervisión, y presentada para la obtención del grado de Doctor por la Universidad de Murcia.

En Murcia, a 8 de febrero de 2008

D. Juan A. Botía Blaya

D. Antonio F. Gómez Skarmeta



Universidad de Murcia

D. Luis Daniel Hernández Molinero, Profesor Titular de Universidad del Área de Ciencias de la Computación e Inteligencia Artificial y Director del Departamento de Ingeniería de la Información y las Comunicaciones de la Universidad de Murcia, INFORMA:

Que la Tesis Doctoral titulada "DEFINICIÓN DE UN MODELO DE GESTIÓN DE LAS NOCIONES DE CONFIANZA Y REPUTACIÓN ENTRE AGENTES. ENFOQUE BASADO EN LA SIMILITUD ENTRE TAREAS", ha sido realizada por D. Alberto Caballero Martínez, bajo la inmediata dirección y supervisión de D. Juan A. Botía Blaya y D. Antonio F. Gómez Skarmeta, y que el Departamento ha dado su conformidad para que sea presentada ante la Comisión de Doctorado.

En Murcia, a 13 de febrero de 2008

D. Luis Daniel Hernández Molinero

A Diego

Agradecimientos

Ojalá que con unas cuantas líneas fuera suficiente para mencionar a quienes han ofrecido su mano y más. Sé que, aun sin escribir ninguna u omitiendo varios nombres, los que me han ayudado sabrán disfrutar de este logro y alegría, que, sin tener dudas, también es suyo. Sin embargo, explícitamente tengo que decir GRACIAS.

GRACIAS a Diego, por ser el motivo y alegría de mi vida, haciendo insignificante cualquier problema.

GRACIAS a Daymi por estar siempre, por compartir locuras, por apoyarme, por asumir un montón de mis tareas de familia, ... por quererme, ... también por revisar el manuscrito y darme buenas sugerencias.

GRACIAS a mis padres y hermana, por tanto cariño y apoyo durante toda mi vida.

Sin encontrar las palabras adecuadas, digo GRACIAS a Antonio, por su apoyo infinito y continuo desde antes de conocerme, allá por el año 2002.

GRACIAS a Juan, por tantas sesiones de trabajo, por tanto conocimiento brindado, por tanta ayuda, por tantas revisiones que hacen menos imperfecto este documento, ... por hacer jocosas muchas conversaciones sobre este tema tan serio.

GRACIAS a Francisco García Sánchez, por las sugerencias y aportaciones generadas por la revisión del manuscrito del capítulo 4, relacionado con las representaciones semánticas estudiadas.

GRACIAS a mis amigos y a los compañeros de despacho y de departamento, por compartir ideas, sugerencias, estados de ánimo, conversaciones, que han hecho este trabajo menos duro.

GRACIAS a la AECl, al programa Alban y a la Fundación Carolina por garantizarme el soporte económico para este trabajo.

Resumen

Los conceptos de confianza y reputación determinan un campo de investigación relevante para varias disciplinas científicas. Desde el punto de vista de las Ciencias de la Computación, una gran parte de los esfuerzos en esta dirección se ha dedicado a la definición de mecanismos de control adaptativos de las comunidades virtuales tomando como referencia las relaciones que se presentan en las sociedades humanas. El hecho de tener en cuenta elementos asociados a la confianza y la reputación contrarresta el elevado grado de incertidumbre bajo el cual los individuos toman sus decisiones. Así, un modelo de confianza y reputación hace frente a la incertidumbre asociada a las decisiones que toman las partes "inteligentes" de varios tipos de sistemas. En este sentido, resulta importante proporcionar algunos mecanismos de estimación de los valores de confianza y reputación que permitan ofrecer recomendaciones valiosas aun cuando no se disponga de la información suficiente para hacerlo. También destaca la necesidad de adopción de algún modelo conceptual común y eficaz para la representación de los elementos dependientes del dominio de aplicación.

Esta tesis está dedicada a la definición y estudio de un modelo de confianza y reputación que, adoptando un marco de representación ontológica para la definición de los elementos dependientes del dominio de aplicación, utiliza criterios de similitud para estimar el desempeño de las partes del sistema cuando la información relacionada se presenta incompleta. Se estudia el comportamiento del modelo para diferentes alternativas de definición de algunos de sus elementos, y frente a diferentes situaciones experimentales, determinadas por las variaciones del desempeño de los elementos del sistema.

Abstract

The concepts of trust and reputation determine an important research field for several scientific disciplines. From the point of view of Computer Sciences, a large amount of efforts in this direction has been devoted to the definition of adaptive control mechanisms for virtual communities based on the relationships that take place in human societies. Taking into account some elements associated with trust and reputation, the high degree of uncertainty, under which individuals make decisions, is reduced. Thus, a trust and reputation model could help to deal with the uncertainty associated with the decisions of "smart" parties, in some types of systems. This way, it is important to provide some mechanism to estimate the trust and reputation values in order to give valuable recommendations, even if there are no sufficient information to do so. It also stands out the necessity to adopt a common and effective conceptual model for the representation of the application domain elements.

This thesis is devoted to the definition and study of a trust and reputation model, based on a common ontological framework for the definition of domain-dependent elements, using similarity criteria for estimating the performance of parts of the system when related information is incomplete. Also it analyzes the behavior of the model to different alternatives for the definition of some of its elements, and in front of different experimental situations, determined by variations in the performance of the system elements.

Índice general

I	Introducción	1
1.	Introducción	3
1.1.	Motivación	3
1.2.	Hipótesis y objetivos de trabajo	4
1.3.	Principales contribuciones	5
1.4.	Publicaciones y trabajos relacionados	6
1.5.	Estructura de la tesis	7
II	Revisión Bibliográfica	9
2.	Modelos de confianza y reputación	11
2.1.	Aspectos clave de un modelo	13
2.1.1.	Obtención de la información de confianza y reputación	13
2.1.2.	Contexto del agente, requerimientos de los usuarios y dimensiones de la confianza y la reputación	15
2.1.3.	Visibilidad y disposición de la información	16
2.1.4.	Atributos considerados para evaluar la confianza y la reputación	17
2.1.4.1.	Integridad	17
2.1.4.2.	Previsibilidad	18
2.1.4.3.	Competencias	18
2.1.4.4.	Motivación	18
2.2.	Características particulares de algunos modelos	19
2.2.1.	Modelos independientes del contexto y de los requerimientos de los usuarios	19
2.2.1.1.	Sporas e Histos	19
2.2.1.2.	MARS	20
2.2.1.3.	AFRAS	21
2.2.1.4.	Modelo de Carter <i>et al.</i>	22
2.2.1.5.	TRAVOS	22
2.2.1.6.	FIRE	23
2.2.1.7.	Modelo de Wang y Singh	24

2.2.2.	Modelos dependientes del contexto o de los requerimientos de los usuarios	24
2.2.2.1.	Modelo de Marsh	25
2.2.2.2.	Modelo cognitivo de Castelfranchi y Falcone	25
2.2.2.3.	Modelo de Abdul-Rahman y Hailes	26
2.2.2.4.	Modelo bayesiano de Mui <i>et al.</i>	27
2.2.2.5.	ReGreT	27
2.2.2.6.	Modelo difuso de Ramchurn <i>et al.</i>	29
2.2.2.7.	TrustMail y FilmTrust	30
2.2.2.8.	Modelo bayesiano de Wang y Vassileva	32
2.2.2.9.	Modelo de Maximilien y Singh	32
2.2.2.10.	MDT-R	33
2.2.2.11.	Un modelo basado en la teoría de la información: Modelo de Sierra y Debenham	34
2.2.3.	Resumen comparativo	35
2.3.	Aspectos que requieren esfuerzos de investigación	40
2.4.	Conclusiones	41
3.	Evaluación de modelos de confianza	43
3.1.	Métricas para la evaluación y comparación de los modelos	44
3.1.1.	Métricas individuales	44
3.1.1.1.	Evolución de la confianza y la reputación	45
3.1.1.2.	Métricas relacionadas con el nivel de satisfacción	45
3.1.1.3.	Métricas relacionadas con la utilidad	46
3.1.1.4.	Métricas relacionadas con las interacciones con otros agentes	46
3.1.1.5.	Métricas relacionadas con la calidad del servicio	47
3.1.2.	Métricas sociales	47
3.1.3.	Información recolectada como fuente de otras métricas	48
3.2.	Escenarios de prueba	48
3.2.1.	Escenarios basados en el dilema del prisionero	49
3.2.1.1.	Escenario PlayGround	49
3.2.1.2.	Escenario con una fase de negociación previa	50
3.2.1.3.	Escenario donde las decisiones son guiadas por la reputación	50
3.2.2.	Recolector de alimentos	51
3.2.3.	Escenarios basados en el paradigma Consumidor - Proveedor	52
3.2.3.1.	Escenario BRS (<i>Beta Reputation System</i>)	52
3.2.3.2.	Escenario simulador de un mercado	53
3.2.3.3.	Escenario basado en servicios	53
3.2.3.4.	Experimentos de Sporas, ReGreT, AFRAS	54
3.2.4.	TOAST: <i>Trust Organisational Agent System Testbed</i>	54
3.2.5.	ART: <i>Agent Reputation and Trust Testbed</i>	56
3.3.	Conclusiones	57

4. Representaciones semánticas	59
4.1. La web semántica	61
4.1.1. Tecnologías para la web semántica	62
4.1.1.1. XML (<i>eXtensible Markup Language</i>)	63
4.1.1.2. RDF (<i>Resource Description Framework</i>) y RDFS (<i>RDF Schema</i>)	63
4.1.1.3. OWL (<i>Web Ontology Language</i>)	65
4.2. Propuestas para servicios web semánticos	66
4.2.1. OWL-S (<i>Ontology Web Language for Services</i>)	67
4.2.2. WSMO (<i>Web Service Modeling Ontology</i>)	67
4.2.2.1. WSML (<i>Web Service Modeling Language</i>)	68
4.2.2.2. Descubrimiento de servicios web según WSMO	69
4.2.3. SWSF (<i>Semantic Web Services Framework</i>)	71
4.2.4. WSDL-S (<i>Web Services Semantics</i>)	72
4.3. Medidas de similitud entre conceptos	73
4.3.1. Coeficiente de correlación de Pearson	73
4.3.2. Medida a partir de la agregación de criterios de comparación	74
4.3.3. Medida de Tversky	74
4.3.4. Medida basada en la comparación léxica	75
4.4. Relación con el enfoque de confianza y reputación	76

III Modelo de Confianza y Reputación **79**

5. El modelo propuesto: TRSIM	85
5.1. Estructura del modelo	87
5.1.1. Estructura de las bases de experiencias	90
5.2. Agrupamiento de los agentes según los criterios de confianza y reputación	90
5.2.1. Estrategias para solicitar información	92
5.3. Confianza para ofrecer soluciones	93
5.4. Confianza para ofrecer información sobre otros	94
5.5. Confianza directa	95
5.6. Reputación	96
5.7. Confianza y reputación a partir de tareas similares	98
5.7.1. Alternativas para aproximar la confianza y la reputación	99
5.8. Funciones dependientes del dominio de aplicación	101
5.8.1. Calidad de la respuesta a una tarea	102
5.8.1.1. Alternativas para determinar la calidad de una respuesta	103
5.8.2. Similitud entre tareas	105
5.8.2.1. Alternativas para determinar la similitud entre tareas	105
5.9. Actualización de las bases de experiencias	107
5.10. Conclusiones	109
5.11. Lógica de un agente para determinar la confianza en otro	111

6. Escenarios de referencia	115
6.1. Identificación de los tipos de agentes utilizando la confianza y la reputación	116
6.2. Escenario Consumidor - Proveedor de recursos	117
6.2.1. Modelo de agentes	119
6.2.1.1. Dimensión mediadora	119
6.2.1.2. Dimensión adaptativa	120
6.2.2. Modelado de las soluciones de los proveedores y los requerimientos de los consumidores	121
6.2.3. Interacciones entre los agentes	121
6.3. Escenario de tasación de pinturas ofrecido por ART	123
6.3.1. Dominio de aplicación e interacciones en ART	123
6.3.2. Estrategia de los agentes en ART	124
6.4. Conclusiones	128
7. Experimentación	129
7.1. Características generales de la simulación	130
7.2. Alternativas para la definición de algunas funciones importantes del modelo	133
7.2.1. Elección de la alternativa adecuada para la calidad	135
7.2.1.1. Evolución de la satisfacción, el error y la confianza	135
7.2.1.2. Diferenciación de los grupos de agentes según la calidad de la respuesta que ofrecen	140
7.2.2. Elección de la alternativa adecuada para la combinación de la confianza y la reputación con la similitud	145
7.2.2.1. Determinación de la mejor alternativa para combinar la confianza y la reputación	146
7.2.2.2. Independencia de la función de similitud	148
7.2.2.3. Diferenciación de los grupos de agentes según la calidad de la respuesta que ofrecen	149
7.2.3. Conclusiones del estudio de diferentes alternativas	152
7.3. Adaptación ante los cambios en el comportamiento de los agentes proveedores	153
7.3.1. Modelado de los cambios de comportamiento	154
7.3.1.1. Disminución de la calidad de la respuesta	155
7.3.1.2. Aumento de la calidad de la respuesta	156
7.3.1.3. Disminución de la calidad de la respuesta y aumento posterior hasta la calidad inicial	156
7.3.1.4. Aumento de la calidad de la respuesta y disminución posterior hasta la calidad inicial	156
7.3.2. Cambios individuales en el comportamiento de diferentes tipos de agentes	157
7.3.2.1. Un agente de calidad <i>Alta</i> empeora su comportamiento	157
7.3.2.2. Un agente de calidad <i>Baja</i> mejora su comportamiento	161
7.3.2.3. Un agente de calidad <i>Alta</i> empeora su comportamiento y vuelve a mejorarlo	163

7.3.2.4.	Un agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo	166
7.3.3.	Cambios en el comportamiento de un grupo de agentes que ofrece soluciones de calidad <i>Alta</i>	170
7.3.3.1.	Todos los agentes de calidad <i>Alta</i> empeoran su comportamiento	170
7.3.3.2.	Todos los agentes de calidad <i>Alta</i> empeoran su comportamiento y vuelven a mejorarlo	174
7.3.4.	Cambios en el comportamiento de un grupo de agentes que ofrece soluciones de calidad <i>Baja</i>	178
7.3.4.1.	Un grupo de agentes de calidad <i>Baja</i> mejoran su comportamiento	178
7.3.4.2.	Un grupo de agentes de calidad <i>Baja</i> mejoran su comportamiento y vuelven a empeorarlo	182
7.3.5.	Conclusiones del estudio de la capacidad de adaptación	185
8.	Adecuación de TRSIM a ART	187
8.1.	TRSIM desde el punto de vista de ART	188
8.2.	Estrategia de un agente ART según TRSIM	190
8.2.1.	Experimentación: Comparación de estrategias alternativas	193
8.3.	Características del agente <i>Novel</i>	194
8.4.	Conclusiones	196
8.5.	Propuesta de extensión de ART para considerar la similitud entre pinturas	196
8.5.1.	Extensión semántica del concepto de Pintura	197
8.5.2.	Similitud entre dos pinturas	198
IV	Conclusiones	199
9.	Conclusiones y trabajos futuros	201
9.1.	Conclusiones	201
9.1.1.	Conclusiones extraídas de la experimentación	203
9.1.1.1.	Comportamiento del modelo en agentes estacionarios	203
9.1.1.2.	Adaptación a los cambios de comportamiento de los agentes	204
9.2.	Trabajos Futuros	204
9.2.1.	TRSIM en otros escenarios	204
9.2.2.	Evaluación del comportamiento grupal o de sistema	205
9.2.3.	Confianza desde el punto de vista del proveedor	205
9.2.4.	Extensión de ART. Comparación con otros modelos	206
9.2.5.	Confianza en el escritorio social semántico	206
Bibliografía		209

Índice de figuras

4.1. Estructura de capas de la web semántica (tomado de: [12])	62
4.2. Ejemplo de sentencia RDF y del esquema RDF	64
5.1. Relación entre las diferentes partes del modelo	88
5.2. Coeficientes de relevancia para cada experiencia según su antigüedad, calculados para diferentes valores del parámetro δ y $ L = 10$	96
5.3. Componente de la precisión de la confianza directa (<i>DTRL</i>) que considera la cantidad de experiencias utilizadas en la obtención de la confianza directa (<i>DT</i>), para diferentes valores del parámetro <i>itm</i>	97
5.4. Algoritmo para determinar la confianza en un agente	112
6.1. Esquema general del escenario Consumidor - Proveedor de recursos	118
6.2. Ejemplo de secuencia de las interacciones entre los agentes en el escenario Consumidor - Proveedor de recursos, según los conceptos de TRSIM.	122
6.3. Ejemplo de secuencia de las interacciones entre los agentes en el escenario ART, según los conceptos de TRSIM.	127
7.1. Distribuciones de probabilidad Beta utilizadas para generar los atributos de las tareas y respuestas según su exigencia y calidad, respectivamente. a) <i>Beta</i> ($\alpha = 5, \beta = 2$) para exigencia y calidad <i>Alta</i> ; b) <i>Beta</i> ($\alpha = 5, \beta = 5$) para exigencia y calidad <i>Media</i> ; c) <i>Beta</i> ($\alpha = 2, \beta = 5$) para exigencia y calidad <i>Baja</i>	131
7.2. Algoritmo general de la simulación	133
7.3. Comparación de la satisfacción con la solución recomendada por el modelo, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	136
7.4. Comparación del error cometido por el modelo al seleccionar la solución adecuada, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	137

7.5.	Comparación del momento (número de ronda) en que el modelo alcanza la estabilidad, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, agrupadas según el nivel de exigencia de la tarea.	138
7.6.	Comparación de la confianza del agente que ofrece la solución recomendada por el modelo, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	139
7.7.	Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_a , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	141
7.8.	Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_b , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	142
7.9.	Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_c , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.	143
7.10.	Comparación del nivel de satisfacción con la solución recomendada por el modelo, para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas (la definición de las alternativas se ofrece en el apartado 5.7.1), según la exigencia de la tarea a resolver en cada ronda: a) alta, b) media.	146
7.11.	Comparación del error cometido por el modelo al seleccionar la solución adecuada, para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas (la definición de las alternativas se ofrece en el apartado 5.7.1), según la exigencia de la tarea a resolver en cada ronda: a) alta, b) media.	147
7.12.	Evolución de la confianza para cada grupo de agentes, utilizando la alternativa IT_c para estimar la confianza y la reputación a partir de tareas similares, según el tipo de tarea a resolver en cada ronda a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>	149
7.13.	Comparación del grado de satisfacción, según la exigencia de la tarea en cada ronda (a) <i>Alta</i> , b) <i>Media</i>), para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas, utilizando diferentes funciones para el cálculo de la similitud 1) Pearson, 2) Tversky, 3) Comparación léxica, 4) Diferencia entre los atributos, 5) Distancia euclídea (las alternativas se detallan en el apartado 5.8.2.1). Las gráficas correspondientes a las funciones 1) Pearson, 2) Tversky, 3) Comparación léxica se muestran en la página anterior.	151
7.14.	Tipos de variaciones del comportamiento de los agentes: a) un agente empeora su comportamiento, b) un agente mejora su comportamiento, c) un agente empeora su comportamiento y luego vuelve a mejorarlo, d) un agente mejora su comportamiento y luego vuelve a empeorarlo.	154

7.15. Comparación del nivel de satisfacción con la solución recomendada por el modelo según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento.	158
7.16. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento.	159
7.17. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento.	160
7.18. Evolución de la confianza del agente de calidad <i>Alta</i> que empeora su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	160
7.19. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento.	162
7.20. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento.	162
7.21. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento.	163
7.22. Evolución de la confianza del agente de calidad <i>Baja</i> que mejora su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>	164
7.23. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento y vuelve a mejorarlo. .	165

7.24. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento y vuelve a mejorarlo.	165
7.25. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Alta</i> empeora su comportamiento y vuelve a mejorarlo.	166
7.26. Evolución de la confianza del agente de calidad <i>Alta</i> que empeora su comportamiento y vuelve a mejorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	167
7.27. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo.	167
7.28. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo.	168
7.29. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo.	168
7.30. Evolución de la confianza del agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	169
7.31. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento.	171
7.32. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento.	172

7.33. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento.	172
7.34. Evolución del promedio de la confianza de los agentes de calidad <i>Alta</i> que empeoran su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	174
7.35. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento y vuelven a mejorarlo.	175
7.36. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento y vuelven a mejorarlo.	175
7.37. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad <i>Alta</i> empeoran su comportamiento y vuelven a mejorarlo.	176
7.38. Evolución del promedio de la confianza de los agentes de calidad <i>Alta</i> que empeora su comportamiento y vuelve a mejorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>). . . .	177
7.39. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50 % de los agentes de calidad <i>Baja</i> mejoran su comportamiento.	179
7.40. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50 % de los agentes de calidad <i>Baja</i> mejoran su comportamiento.	180
7.41. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50 % de los agentes de calidad <i>Baja</i> mejoran su comportamiento.	181

7.42. Evolución del promedio de la confianza de los agentes de calidad <i>Baja</i> que mejoran su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	181
7.43. Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad <i>Baja</i> mejoran su comportamiento y vuelven a empeorarlo.	182
7.44. Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad <i>Baja</i> mejoran su comportamiento y vuelven a empeorarlo.	183
7.45. Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad <i>Baja</i> mejoran su comportamiento y vuelven a empeorarlo.	184
7.46. Evolución de la confianza del agente de calidad <i>Baja</i> mejora su comportamiento y vuelve a empeorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia <i>Alta</i> , b) de exigencia <i>Media</i>).	185
8.1. Balance del estado de cuentas para tres tipos de agentes según diferentes estrategias para solicitar opiniones a otros agentes.	194

Índice de tablas

2.1.	Abreviaturas utilizadas en el resumen comparativo de las características más relevantes de los modelos estudiados	37
2.2.	Resumen comparativo de las características más relevantes de los modelos .	38
4.1.	Coincidencias entre las metas y las descripciones de los servicios web, según la relevancia de sus atributos y la intención especificada en el proceso de descubrimiento (tomada de [140]).	70
7.1.	Áreas entre las curvas que describen los valores de confianza del agente iniciador sobre los agentes de calidad <i>Alta</i> y <i>Media</i> para cada una de las alternativas utilizadas en la obtención de la calidad de la respuesta (desde $t = 0$ hasta $t = 400$).	144
7.2.	Áreas entre las curvas que describen los grados de satisfacción para las alternativas IT_c e IT_f , al estimar la confianza y la reputación para los tipos de tareas analizados: de exigencia <i>Alta</i> y de exigencia <i>Media</i> (desde $t = 0$ hasta $t = 400$).	147
7.3.	Efecto de cada una de las modificaciones de la calidad de la respuesta en la estabilidad del modelo. Se muestra, en su caso, el número de la ronda en que el modelo alcanza nuevamente el equilibrio tras finalizar el cambio en el comportamiento de los agentes.	186

Parte I
Introducción

Capítulo 1

Introducción

1.1. Motivación

A pesar de que los conceptos de confianza y reputación han sido ampliamente analizados desde la perspectiva de la Psicología, la Sociología y la Filosofía, fundamentalmente, aun determinan un espacio de estudio que requiere gran atención desde el punto de vista de las Ciencias de la Computación. En este sentido, las investigaciones científicas relacionadas con la confianza y la reputación en comunidades virtuales es una disciplina reciente, encaminada a aumentar la fiabilidad y desempeño de las comunidades electrónicas a partir de considerar los mecanismos de control de las sociedades humanas [111].

Una gran variedad de sistemas distribuidos (tales como el Grid, y los sistemas basados en la interacción entre pares) pueden ser vistos como sistemas multi-agente en los cuales los componentes actúan de manera autónoma y flexible intentando alcanzar sus propios objetivos [74]. En un sistema multi-agente, los agentes pueden ser consumidores y/o proveedores de información, recursos y servicios. Cada agente se muestra interesado en sus propios objetivos o metas, muchas veces sin resultar benevolentes o confiables en sus relaciones sociales. No poseen una representación exacta del entorno y de los otros agentes, que le permita predecir con exactitud su comportamiento futuro. Con frecuencia, tampoco existe una entidad central que brinde información relacionada con el desempeño de cada agente. Por tales motivos, la incertidumbre asociada a cada decisión de los agentes, que dependa de aspectos relacionados con el resto del sistema multi-agente, es muy elevada.

Para hacer frente a la falta de información o a la existencia de información incompleta o imprecisa sobre el comportamiento de otros agentes, se necesitan mecanismos que posibiliten reducir los riesgos derivados de las interacciones entre las diferentes partes. Ante una determinada situación, cada agente puede mostrar comportamientos diferentes. Por lo general, los agentes consumidores de recursos, por ejemplo, prefieren interactuar (negociar, cooperar, contratar, etc.) solo con los proveedores que exhiben los mejores comportamientos. Dado que en la mayoría de estos escenarios no existe una entidad capaz de ofrecer información acerca del comportamiento de los demás, cada agente debe crear y mantener actualizado su propio modelo de representación sobre el desempeño de sus

conocidos. Así, ante determinada situación, podrá analizar con cuáles de ellos interactuar, mediante mecanismos de cooperación, para resolver un determinado problema.

En este sentido, los modelos de confianza y reputación ofrecen una buena solución para representar y predecir el comportamiento del resto de agentes en el sistema. Los valores de confianza y reputación pueden usarse como un criterio para identificar los agentes de los que se espera un buen comportamiento, es decir, los que son capaces de ofrecer una solución de calidad para un determinado problema.

Estos modelos, por lo general, combinan información de confianza y reputación de varios tipos. Los valores pueden estimarse a partir de diferentes fuentes de información, en dependencia del punto de vista analizado, el contexto o, simplemente, la disponibilidad de conocimiento. Cuando la información de estas fuentes es insuficiente para ofrecer algún criterio de confianza o reputación, de manera intuitiva, se puede pensar en obtener una aproximación de estos valores. Para ello, una alternativa sería utilizar el conocimiento almacenado sobre el desempeño previo de otros agentes frente a requerimientos similares al deseado.

Por otra parte, la información muchas veces depende del dominio de aplicación al que se aplique el modelo de confianza y reputación. En este sentido puede resultar conveniente utilizar una representación ontológica común que facilite la manipulación y razonamiento a partir de la información recopilada por los agentes acerca del desempeño de otros. Dicha representación debe garantizar, entre otros, un modelo de los agentes en el sistema, de los requerimientos de aquellos que exhiben el rol de consumidor, las soluciones ofrecidas por los agentes desde el rol proveedor, herramientas para la comparación de los requerimientos y de las soluciones, así como la posibilidad de establecer medidas de calidad y de satisfacción de los requerimientos a partir de una determinada solución.

1.2. Hipótesis y objetivos de trabajo

Esta tesis doctoral ha sido encausada por la siguiente **hipótesis de trabajo**:

- La representación adecuada de los elementos de un modelo de confianza y reputación dependientes del dominio de aplicación, así como la estimación de la confianza y la reputación a partir de criterios de similitud entre los requerimientos (o preferencias) de los usuarios, puede elevar el desempeño del modelo y los niveles de satisfacción de los mismos.

Para demostrar o refutar dicha hipótesis se propuso el siguiente **objetivo general**:

- Definir un modelo de confianza y reputación que permita proponer las soluciones más adecuadas a los requerimientos (o preferencias) de los usuarios en un entorno de interacción entre pares, bajo condiciones de desconocimiento, adoptando un marco conceptual de representación común que permita la definición de los elementos del modelo de confianza y reputación y de los usuarios involucrados.

Los **objetivos específicos** propuestos para cumplimentar el objetivo general fueron los siguientes:

1. Ofrecer un modelo de confianza y reputación que, dados los eventuales requerimientos de los usuarios, sea capaz de brindar las soluciones más adecuadas a partir del conocimiento almacenado sobre las interacciones que han tenido lugar con anterioridad.
2. Proponer una representación adecuada para los conceptos y funciones del modelo, dependientes del dominio de aplicación, a partir de un marco conceptual, que brinde elementos útiles para el tratamiento de la confianza y la reputación. De esta forma, se pretende ofrecer un marco de representación común para todos los sistemas que utilicen el modelo.
3. Establecer mecanismos para la aproximación de los valores de confianza y reputación para aquellos casos en que el modelo carece de información almacenada sobre el desempeño del sistema en interacciones previas.
4. Brindar mecanismos adaptativos para soportar los cambios de comportamiento de los agentes y/o el entorno.

1.3. Principales contribuciones

Las principales contribuciones del presente trabajo de tesis doctoral, se encuentran enumeradas en los siguientes puntos:

- Proponer y describir el funcionamiento de un modelo de confianza y reputación en un entorno de interacción entre pares para identificar las mejores soluciones que se puedan ofrecer al usuario a partir de la especificación de sus requerimientos y preferencias.
- Utilizar medidas de similitud entre las preferencias y requerimientos de los usuarios para estimar la confianza y la reputación cuando no se dispone de conocimiento suficiente para calcular estos valores a partir de la información almacenada acerca de las interacciones previas. En los casos donde se utilizan aproximaciones de estos valores a partir de criterios de similitud se obtienen mejores desempeños del modelo propuesto.
- Tomar el marco conceptual ofrecido por la propuesta de WSMO ¹ para representar los conceptos y funciones dependientes del dominio de aplicación, a partir de la identificación y utilización de los elementos de WSMO útiles para el tratamiento de la confianza y la reputación. Se proponen y comparan diferentes alternativas de definición de los elementos dependientes del dominio de aplicación.

¹ *Web Service Modeling Ontology* <http://www.wsmo.org/>

1.4. Publicaciones y trabajos relacionados

El trabajo presentado en esta tesis ha dado lugar a las siguientes publicaciones:

- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **Adaptability of the TRSIM model to some changes in agents behaviour.** (Por aparecer.) Presentado a Autonomic and Trusted Computing, ATC 2008. LNCS xxxx, pages xxx-xxx. Springer, June 2008.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **Assesing the Effectivity of TRSIM Model for Trust and Reputation.** In R. Bordini editor, Fifth European Workshop on Multi-Agent Systems EUMAS 07, pages 38-52, December 2007.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **Adecuación del modelo de confianza y reputación TRSIM a ART.** In J. Carbó, J. Sabater-Mir, J. Molina editors, Agents in the ART testbed spanish competitions at CAEPIA 07, pages 16-21, November 2007.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **On the Behaviour of the TRSIM Model for Trust and Reputation.** In P. Petta, J. P. Müller, M. Klusch, M. Georgeff, editors, Multiagent System Technologies. 5th German Conference, MATES 2007. LNCS 4687, pages 182-193. Springer, September 2007.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **On the Convenience of Using Similarity Between Tasks in a Trust and Reputation Model for Contracting Web Services.** In R. Falcone and S. Barber and J. Sabater-Mir and M. Singh, editors, Proc. of Workshop on Trust in Agent Societies at AAMAS 07, pages 22-30, May 2007.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **A New Model for Trust and Reputation Management with an Ontology Based Approach for Similarity Between Tasks.** In K. Fischer, I. J. Timm, E. André, and N. Zhong, editors, Multiagent System Technologies, 4th German Conference, MATES 2006. LNCS 4196, pages 172-183. Springer, September 2006.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **Trust and Reputation Model based on WSMO.** In R. Falcone and S. Barber and J. Sabater-Mir and M. Singh, editors, Proc. of Ninth International Workshop on Trust in Agent Societies at AAMAS 06, pages 9-18, May 2006.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **An Approach to Achieve Intelligence into an Agent-based System to Contract Information Services.** Proc. of XII Congreso Latino Iberoamericano de Investigación de Operaciones CLAIO 04, Octubre 2004.

- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **Modelo de los datos incorporados en los agentes FIPA que conforman un sistema distribuido para la invocación automática de servicios Web**. Rev. Ingeniería Industrial, XXV(3), pages 56-61. ISPJAE. Cuba, 2004.
- A. Caballero, J. Botía, and A. Gómez-Skarmeta. **El papel de los agentes FIPA en aplicaciones basadas en Servicios Web**. Proc. of CAEPIA'2003 Conferencia de la Asociación Española de Inteligencia Artificial, pages 249-252, Noviembre, 2003.

Los resultados de esta tesis han contribuido al desarrollo de los siguientes proyectos de investigación:

- **POPEYE: Peer to Peer Collaborative Working Environment over Mobile Ad-Hoc Networking**. Financiado bajo *IST Call 5 of the EU's 6th Framework Program* (IST-2006-034241). 2006-2008. Investigadores principales: Antonio Gómez-Skarmeta, Juan A. Botía Blaya.
- **INGENIAS II**. Financiado por el Ministerio de Educación y Ciencia de España (TIN-2005-08501-C03-02). 2005-2008. Investigator principal: Juan A. Botía Blaya.

Además, utilizando el modelo de confianza propuesto, se han implementado las estrategias de comportamiento de dos agentes *MurciaAgent* y *Novel* que han participado en la *I y II Competición Nacional de Modelos de Confianza (2006 y 2007)*, respectivamente. *Novel* ha participado además en la *II Competición Internacional de Modelos de Confianza* (celebrada en el marco del AAMAS'2007).

1.5. Estructura de la tesis

El documento de tesis está organizado en cuatro partes. Esta primera parte ha estado dedicada a la introducción del trabajo realizado, comentando su motivación, hipótesis, objetivos, principales contribuciones y resultados, entre otros. También ofrece una visión general de la estructura del resto del contenido de la tesis.

En la parte II del documento de tesis se hace una revisión del estado del arte en el campo de los modelos de confianza y reputación para agentes y sistemas multi-agente, así como de las soluciones disponibles para abordar la representación semántica de los componentes de un modelo de este tipo.

En el **capítulo 2**, se comentan algunos aspectos importantes que son considerados en el desarrollo de modelos de confianza y reputación, y las soluciones más trascendentales que han aportado otros autores. Se concluye el primer capítulo puntualizando las ideas más importantes que recoge la literatura y los elementos que aun demandan gran atención en este campo.

En el **capítulo 3** se relacionan las principales formas de establecer métricas para el estudio del comportamiento de los modelos de confianza y reputación entre agentes.

También se revisan los diferentes escenarios de prueba que han sido definidos para facilitar la comparación de los desempeños de los diferentes modelos.

El **capítulo 4** describe los principales modelos semánticos que pueden ser útiles para la representación de los componentes de los modelos de confianza y reputación.

En la parte III, el documento describe la estructura, funcionalidad y comportamiento del modelo de confianza y reputación propuesto como solución a la problemática planteada.

El **capítulo 5** detalla la estructura, componentes y funcionalidad del modelo de confianza y reputación TRSIM.

En el **capítulo 6** se describen las características de dos escenarios en los cuales puede resultar factible la utilización del modelo de confianza y reputación planteado en el capítulo anterior.

En el **capítulo 7** se ofrecen resultados experimentales que permiten interpretar y ajustar el funcionamiento del modelo. Se presenta un conjunto de experimentos que muestran cómo se comporta el modelo TRSIM ante diferentes situaciones de interés. Los experimentos están encaminados fundamentalmente al estudio del comportamiento del modelo según diferentes alternativas para la definición de sus funciones, y al estudio de la capacidad de adaptación ante los cambios en el desempeño de los agentes y/o el entorno.

El **capítulo 8** muestra cómo se adecúa el modelo de confianza y reputación planteado al escenario de prueba ART ². Se destacan los puntos de contacto, diferencias y condiciones bajo las cuales tiene sentido utilizar ART para la comparación de las alternativas del modelo, o analizar su comportamiento frente a otros modelos. Además se ofrecen algunas recomendaciones al desarrollo de ART, relacionadas con el enriquecimiento semántico de los conceptos que utiliza.

Por último, en la parte IV, el **capítulo 9** se dedica a puntualizar las principales conclusiones del trabajo y algunas ideas que definen líneas de trabajo para investigaciones futuras.

²Agent Reputation and Trust Testbed. <http://www.art-testbed.net/>

Parte II
Revisión Bibliográfica

Capítulo 2

Modelos de confianza y reputación

La confianza y la reputación son conceptos muy importantes en el ámbito social, por lo que han sido abordados por muchas disciplinas científicas (por ejemplo, la Psicología, la Sociología y la Filosofía, entre otras), cada una desde perspectivas diferentes [111].

Desde el punto de vista de la inteligencia artificial distribuida [99], se han realizado importantes avances. Estos conceptos han sido incorporados en sistemas abiertos y distribuidos de manera que guíen las interacciones entre las entidades que, por la complejidad propia del entorno, manejan información incompleta, con elevado grado de incertidumbre o constantemente cambiante. Un ejemplo de este tipo de sistema son los sistemas multi-agente, donde cada componente es considerado un agente autónomo que debe cooperar de manera efectiva con otros agentes para lograr el buen desempeño del sistema.

Cada agente, bajo un control local individual, toma sus propias decisiones que determinan sus acciones a partir de sus preferencias y conocimiento. Los agentes deben incorporar herramientas adaptativas que le permitan adecuarse en cada momento a las necesidades del sistema y de su entorno, debiendo decidir entre otras cuestiones qué mecanismo utilizar, cuándo utilizarlo, con quién interactuar, etc. Estas decisiones deben tomarse bajo un elevado grado de incertidumbre e imprecisión de la información disponible. La confianza que un agente tenga sobre los otros, ayuda a reducir los riesgos de tomar decisiones bajo estas circunstancias de incertidumbre [104].

De manera general, la confianza que un agente posee en otro puede ser considerada como una medida acerca de la certeza que se tiene en que el otro agente será capaz de ejecutar eficientemente una determinada acción, teniendo en cuenta su propio conocimiento. Por otra parte, la reputación puede ser vista como la certeza que un agente tiene sobre el comportamiento de otro, compuesta a partir del conocimiento que es capaz de extraer de las relaciones con el resto de agentes, ya sea por el análisis de la red de relaciones sociales como por la información suministrada por otros. Ésta es la orientación que se da a ambos conceptos en el trabajo presentado en esta tesis.

Actualmente, el estudio y desarrollo de modelos de confianza y reputación constituye objeto de varias investigaciones [52, 59, 60, 72, 137, 151, 11, 10, 67, 102, 128, 133, 121]. En ellas el principal objetivo es brindar mecanismos auxiliares que apoyen la toma de

decisiones de los agentes en su interacción. Los modelos de confianza y reputación pueden ser utilizados para guiar las interacciones entre las diferentes partes del sistema, dando respuesta a problemas claves tales como:

- Selección de los mecanismos y protocolos para llevar a cabo las interacciones (¿cómo interactuar?).
- Determinación del momento necesario en que se necesita interactuar con otros componentes del sistema (¿cuándo interactuar?).
- Selección de los agentes con que interactuar (¿con quién interactuar?).

Se han desarrollado diferentes trabajos en cuanto a la selección de los mecanismos de interacción entre los agentes y el momento adecuado de hacerlo [14, 45]. Otros se han centrado en proponer modelos que permitan seleccionar los agentes interesantes para una determinada interacción [84, 33, 147, 116, 2, 143, 96, 28, 111, 105, 56, 57, 133, 72]. En este sentido, cada modelo trata de ofrecer una solución adecuada a aspectos clave relacionados con la confianza y reputación (por ejemplo, cómo obtener y almacenar la información, cuál utilizar, cuándo utilizarla, qué estructura utilizar para su representación e intercambio, etc.).

Por otra parte, a pesar de que los conceptos de confianza y reputación se encuentran presentes en las interacciones entre agentes a todos los niveles, Ramchurn *et al.* [104] definen dos niveles conceptuales complementarios en los que se considera la confianza:

- confianza a nivel individual, que considera la confianza como una medida de la honestidad y/o reciprocidad del resto de los agentes, y
- confianza a nivel de sistema, que considera que los agentes están forzados a confiar en las reglas, mecanismos, protocolos, etc. regulados por el sistema.

Ramchurn *et al.* [104] plantean que a pesar de que los protocolos y reglas del sistema garantizan la honradez de los agentes en el mismo, no siempre lo hacen de la manera más eficiente posible. Por tal razón, los modelos de confianza a nivel individual proporcionan información para guiar las decisiones de los agentes en cada momento. A nivel individual se tiene en cuenta la confianza como un fenómeno social inherente a las interacciones entre dos partes [93, 103, 141, 39]. La confianza es una propiedad que resulta de las interacciones entre los agentes: cada uno almacena información sobre el comportamiento de los otros para ser utilizada en futuras interacciones. Para ello, los modelos de este tipo establecen métricas que se calculan a partir de la información almacenada sobre el desempeño de los agentes en interacciones anteriores [139, 113, 21, 84, 2, 145, 105, 101, 126, 70].

Bajo este punto de vista, el concepto de reputación es un concepto similar al de confianza, solo que para su obtención se utiliza la información obtenida a partir de terceros. Se puede considerar como la agregación de las opiniones de un subconjunto de agentes de una determinada comunidad sobre el desempeño que muestra otro de ellos. La agregación de las opiniones de este subconjunto de agentes puede ser ponderada en función de la

confianza que se tiene en el agente que la ofrece, teniendo en cuenta la red de relaciones sociales que se presenta entre los agentes [146, 116, 145, 54, 56, 21].

Con el objetivo de ofrecer una visión general, el capítulo continúa comentando los aspectos más relevantes a tratar por un modelo de confianza y reputación cualquiera (sección 2.1). En la sección 2.2, se puntualizan las principales características específicas de varios de los modelos que destacan por su importancia en el desarrollo del estudio de esta temática (todos ellos centrados en el ámbito de los sistemas multi-agente), ofreciéndose un resumen comparativo a partir de cómo consideran los elementos clave tratados en la sección 2.1. En la sección 2.3, se ofrecen varias ideas sobre problemas de investigación que aun continúan abiertos en el área. Al finalizar, en la sección 2.4 se puntualizan los aspectos más importantes abordados en este capítulo.

2.1. Aspectos clave de un modelo

El tema de confianza y reputación ha sido abordado por muchos investigadores desde diferentes enfoques y con gran variedad de definiciones de estos conceptos, obteniendo soluciones para disímiles problemas, haciendo uso de la información de confianza y reputación de diferentes maneras, etc. Esto ha generado un elevado grado de ambigüedad en los conceptos utilizados en el desarrollo y aplicación de un modelo de confianza y reputación [11]. De ahí que la clasificación de los modelos de confianza existentes sea una tarea difícil [111], tanto la definición de los aspectos que la determinan como el conjunto de clases posibles.

Con el objetivo de establecer algunos elementos que faciliten el estudio de los modelos de confianza y reputación, a continuación se ofrecen algunos aspectos importantes que se consideran en su desarrollo.

2.1.1. Obtención de la información de confianza y reputación

En la mayoría de los modelos de confianza y reputación desarrollados se considera la información que se obtiene a partir del conocimiento propio del agente o del conocimiento brindado por terceros. De manera general, la diferenciación de estos dos tipos de información hace que, en la mayoría de los casos, se hable de modelos de confianza (cuando los valores de confianza son obtenidos a partir de información propia), modelos de reputación (cuando se basa en información brindada por terceros), o modelos de confianza y reputación (cuando se consideran ambos tipos).

Ésta, sin embargo, es una clasificación muy básica por cuanto existen otras fuentes de información de confianza y reputación. Por ejemplo, puede existir una entidad ajena especializada en analizar las interacciones de los agentes y brindar estadísticas acerca de su desempeño; también pudiera realizarse un análisis de la red de relaciones sociales que se dan entre los agentes; etc.

Los tipos de información de confianza utilizados dependen de las características específicas de los agentes y del entorno en el cual se ejecutan. Los agentes tienen que

ser capaces de producir y recibir la información adecuada, al mismo tiempo que el entorno debe brindar los mecanismos necesarios para ello. Por otra parte, los agentes por lo general cuentan con capacidades limitadas de almacenamiento y procesamiento de información [106]. Además, la fiabilidad de los valores de confianza aumenta en tanto mayor cantidad de tipos de información de confianza se utilicen para producirla [111]. Por tanto, existe un compromiso entre la complejidad de los tipos de información empleados, la fiabilidad de la información de confianza y las capacidades de almacenamiento y procesamiento, entre otros.

Sabater [111] destaca cuatro tipos de fuentes de confianza fundamentales: (1) experiencias directas, (2) información de testigos o de terceros, (3) redes de relaciones sociales, y (4) prejuicios. Las experiencias directas son la fuente más fiable de información de confianza y la más utilizada. Pueden almacenarse experiencias sobre el comportamiento de los agentes con los cuales se ha interactuado directamente o sobre las interacciones que tengan otros (en este caso el entorno donde se ejecuten los agentes debe permitir escuchar conversaciones ajenas). La información brindada por terceros también es comúnmente utilizada [111]. Es muy abundante aunque no tan fiable por cuanto los agentes que la ofrecen pueden ocultarla, mentir al ofrecerla o simplemente tener apreciaciones diferentes a quien la solicita.

Las redes de relaciones sociales son una fuente de información de confianza utilizada en sistemas donde existe la disponibilidad de un gran volumen de información sobre las relaciones entre los agentes de acuerdo a los diferentes roles sociales que desempeñan. Estas redes tratan de representar las complejas relaciones sociales que se dan en las sociedades humanas. Trabajos como los de Mui *et al.* [96], Yu y Singh [145], Golbeck [55] y Battiston *et al.* [10] utilizan redes de relaciones sociales para propagar la información de reputación, teniendo en cuenta los posibles caminos que se representen sobre la red, su longitud y resolviendo problemas de ambigüedades, redundancias, etc.

Los prejuicios que se tengan sobre un agente son otra fuente de información de confianza. Éstos pueden estar determinados por la pertenencia o no a un grupo social determinado. En estos modelos, se pueden asignar los valores de confianza a un agente a partir de los valores de confianza que se tengan sobre el grupo de agentes al cual pertenece. De ahí que en muchas ocasiones no solo se haga referencia al concepto de confianza en un agente, sino también al de confianza en una comunidad de agentes [134].

Becerra *et al.* [11] señalan cuatro tipos de métodos para la obtención de los valores de confianza: (1) intuición, (2) experiencia, (3) rumores o información a partir de terceros, e (4) historial. Esta clasificación se solapa en algunos casos con la dada por Sabater [111]. Sin embargo, tiene en cuenta algunos elementos no comentados arriba tales como la intuición y el manejo de un historial por parte de una autoridad del entorno.

El concepto de intuición considera todos los métodos que permiten intuir la confianza en un agente de manera general, sin considerar la información específica que se tiene sobre él. Becerra *et al.* [11] reconocen que éste es un método difícil de utilizar en un sistema multi-agente, por cuanto está matizado por un componente sentimental, de cada agente, asociado a la predisposición, favorable o desfavorable, a la interacción con otro. Sin embargo, desde el punto de vista computacional se puede pensar en un criterio

aproximado que se pueda tener sobre el comportamiento de otro agente cuando aun no se ha interactuado con él. Según un punto de vista propio, para obtener un valor de intuición acerca de cuan beneficioso puede ser interactuar con un agente se necesitaría facultades de introspección para disponer del conocimiento previo del mismo. Por ello, en muchas ocasiones no pasaría de ser un caso específico de procesamiento de las experiencias del propio agente, donde se aproximan los valores de confianza para otros agentes de los cuales se desconocen sus comportamientos.

Por otra parte, la información dada por una entidad institucional acerca del comportamiento de los agentes puede manejarse de igual manera a la información dada por testigos en la clasificación dada por Sabater [111]. La diferencia entre ambos casos es que la información brindada por una entidad institucional estaría siempre disponible y no viciada por mentiras de dicha entidad (siempre que ésta sea confiable). Sin embargo, la información que maneje la propia entidad puede ser inexacta por cuanto los agentes pueden notificar información falsa. La disponibilidad de una entidad o autoridad institucional depende del nivel de centralización de los mecanismos de coordinación que brinde el entorno.

2.1.2. Contexto del agente, requerimientos de los usuarios y dimensiones de la confianza y la reputación

Varios modelos tratan los valores de confianza y reputación como características inherentes únicamente a los agentes [146, 116, 143, 28]. Éstos no tienen en cuenta que los valores de confianza y reputación pueden variar según las características del entorno o los requerimientos de los usuarios. Tampoco consideran la presencia de múltiples dimensiones de los conceptos de confianza o reputación. Por ejemplo, un agente puede ser altamente confiable en un entorno determinado por una reunión de trabajo y no en otro dado por una subasta en la bolsa. De igual manera, si se consideran los requerimientos de los usuarios de un agente, se pueden encontrar desempeños muy diferentes de éste al resolver ecuaciones matemáticas y al trazar un camino entre dos sitios de la ciudad. También, considerando varias dimensiones, puede tenerse un agente confiable ofreciendo respuestas de alta calidad pero no cumpliendo los plazos de entrega.

En otros modelos la confianza y la reputación son dependientes del contexto [84, 33, 2, 113, 105]. En ellos se utilizan mecanismos para gestionar varios contextos al unísono. Como resultado, cada agente puede estar caracterizado por diferentes valores de confianza y reputación para cada uno de los contextos considerados.

En los modelos de confianza estudiados, cuyas principales características se muestran en apartados siguientes, el contexto aparece como un elemento que hace referencia a un conjunto de características del entorno. Desde el punto de vista de la abstracción utilizada por los modelos de confianza y reputación, no se particulariza en la definición de las características del contexto. Se considera que los contextos de los agentes, sus tipos y elementos, son definidos en un nivel de abstracción superior al tratado en los modelos de confianza y reputación. En otras palabras, para todos los modelos estudiados, la definición de los posibles contextos del agente queda fuera del alcance de estudio de los elementos de

confianza y reputación.

Si bien el tipo de tarea que debe desarrollar un agente no forma parte de su contexto, sí lo es la presencia de determinados requerimientos o preferencias de los usuarios. Por tanto, el comportamiento diferenciado de la confianza y la reputación ante diferentes requerimientos o preferencias de los usuarios puede ser considerado de manera similar al comportamiento según diferentes contextos. Es decir, en muchas situaciones, los valores de confianza y reputación dependen también de los requerimientos o preferencias de los usuarios.

Por otra parte, existe modelos que consideran que los valores de confianza y reputación resultan de la agregación de valores referidos a dimensiones específicas. Por ejemplo, Castelfranchi y Falcone [33] tienen en cuenta que la confianza es la combinación de creencias tales como la competencia, la disposición, la dependencia y el cumplimiento. De igual forma, Griffiths [57] demuestra que utilizando cuatro dimensiones de la confianza (éxito en la ejecución de la tarea, costo incurrido al ejecutar la tarea, cumplimiento del tiempo comprometido y la calidad de la respuesta) se puede mejorar significativamente la selección de agentes contrapartes para una determinada interacción.

También Gujral *et al.* [65] consideran medidas multidimensionales de la confianza para la representación de los requerimientos que imponen las metas u objetivos del agente (lo que determina sus beneficios) y para estimar el comportamiento de los agentes contrapartes a partir de sus restricciones. Las dimensiones para los requerimientos de las metas incluyen requerimientos de calidad, de tiempo y de costos fijos relacionados con el dominio de aplicación. Por otra parte, las dimensiones correspondientes a las restricciones del agente contraparte comprenden estimaciones de las restricciones de calidad, de tiempo, de disponibilidad y de los costos que se pagará a la contraparte.

2.1.3. Visibilidad y disposición de la información

La información relativa a la confianza que se tiene sobre un determinado agente, así como la reputación que éste posee entre sus vecinos, puede aparecer como una característica global de dominio público o puede permanecer en secreto, como una propiedad individual de cada agente.

Los entornos donde la información de confianza y reputación se mantiene pública, visible para todos los agentes, pueden presentar diferentes problemas. Por ejemplo, pueden existir agentes que eleven ficticiamente su nivel de confianza mediante la creación de entidades cuyo único objetivo es ofrecer evaluaciones satisfactorias acerca de su desempeño [146]. En cambio, los modelos donde la información permanece de dominio local, los agentes tienen la posibilidad de omitir información, brindar información falsa, o contradecirse mediante el envío de valores incongruentes. Por lo tanto, debe garantizarse mecanismos para el tratamiento del intercambio de información incompleta, falsa o imprecisa.

Desde otro punto de vista, la disponibilidad de la confianza y reputación puede estar determinada por el tipo de modelo de control que se utilice para la gestión de los agentes: centralizado o descentralizado. Los mecanismos de confianza y reputación usados en estos dos tipos de sistemas también son diferentes. El mantenimiento de la información de

confianza y reputación puede realizarse de manera global, o bien de manera local, en cada agente [135].

En entornos centralizados (mayormente vistos en el área del comercio electrónico), los mecanismos de confianza y reputación son relativamente simples, presentando las siguientes características comunes:

- Existe un nodo central, o alguna entidad institucional, que actúa como gestor responsable de mantener actualizada la información utilizada para obtener los indicadores de confianza.
- La información de confianza y reputación por lo general es ofrecida por el sistema. No es necesaria una representación del comportamiento del resto de componentes del sistema y de las relaciones entre éstos.
- No se genera un diálogo adicional entre las partes del sistema para obtener los valores de reputación. Las medidas de reputación se obtienen a partir de la interacción directa con la entidad gestora central.

Los modelos utilizados en entornos con control descentralizado, tal como los basados en la interacción entre pares, son más complejos. Por lo general, presentan las siguientes características:

- No existe un elemento central que ofrezca métricas sobre el desempeño de los agentes
- Cada elemento debe mantener actualizado su propio modelo sobre el comportamiento del resto de sus vecinos y de las relaciones con ellos.
- Se intercambia un gran volumen de información de confianza y reputación.
- Los valores de confianza y reputación se mantiene como información privada que en muchos casos se utiliza para obtener ventajas en las interacciones con otros.

2.1.4. Atributos considerados para evaluar la confianza y la reputación

La confianza que un agente tiene sobre otro puede ser considerada desde diferentes perspectivas. Becerra *et al.* [11] plantean que la confianza y la reputación pueden ser evaluadas a partir de cuatro atributos fundamentalmente: la integridad del agente, la previsibilidad de su comportamiento, sus competencias, y su motivación.

2.1.4.1. Integridad

El componente referido a la integridad describe cuan ético u honrado es un agente. Quizás éste sea uno de los atributos mayormente usados en los modelos de confianza y reputación. Muchas veces aparece como una medida de confiabilidad relacionada con los valores de confianza y reputación [146, 113, 28, 21, 121].

Los agentes pueden ser honestos al ofrecer tanto información de confianza como de reputación, es decir, responder a todas las preguntas y hacerlo siempre con su criterio. Por lo general, este comportamiento es empleado por agentes cuyas interacciones tienen lugar en sistemas cooperativos y que no compiten unos con otros.

Sin embargo, en ambientes competitivos es muy frecuente la presencia de agentes con comportamientos deshonestos, omitiendo la respuesta a solicitudes recibidas o respondiendo con información de confianza y reputación falsas.

Los modelos, que consideran la existencia de agentes que mienten u ocultan información de confianza o reputación, incorporan mecanismos que, de una forma u otra, detectan estos comportamientos no honestos y son capaces de representarlo de manera adecuada para considerarlo en interacciones futuras [146, 113, 21, 121, 28, 101, 126, 70, 71].

2.1.4.2. Previsibilidad

La previsibilidad del comportamiento del agente es un elemento que reduce considerablemente la incertidumbre, cuanto más predecible sea el comportamiento de un agente menor incertidumbre se tendrá para determinar la confianza en él. Mayer *et al.* [89] consideran la previsibilidad como un factor que influye en la predisposición a la cooperación entre dos agentes.

Este atributo es considerado, quizás de manera implícita, en ReGreT mediante la credibilidad en la medida de confianza directa a partir del número de experiencias utilizadas en la obtención de la medida de confianza y su dispersión [111].

2.1.4.3. Competencias

Este atributo describe la habilidad del agente para desempeñar la tarea que se le exige en dependencia de la situación en que se encuentra [11].

Varios modelos han considerado las competencias del agente para determinar la confianza que se tiene sobre él. Yu y Singh [143, 144] utilizan la calidad del servicio ofrecido por el agente en cada interacción como medida de las competencias del agente. ReGreT utiliza una estructura ontológica para evaluar las competencias del agente de acuerdo al rol que interese, a partir de la agregación de valores de reputación asociados a elementos simples [113].

2.1.4.4. Motivación

Este es un atributo menos abordado por la mayoría de los modelos de confianza existentes [11]. Se utiliza para describir el interés que puede mostrar un agente en completar la tarea asignada.

La motivación puede ser un criterio complementario a la hora de seleccionar a quién delegar una tarea. Puede representar la ambición del agente en llevar a cabo una tarea. Si la motivación de un agente aumenta, es muy probable que su integridad disminuya [11]. Por lo general, la motivación del agente no tiene necesariamente que estar relacionada con sus competencias. Así, la repercusión de la motivación en la obtención de la confianza del

agente estará determinada por las apreciaciones (creencias) e importancia que otorgue el agente que hace la evaluación.

2.2. Características particulares de algunos modelos

Actualmente existe un gran número de modelos de confianza y reputación, con diferentes objetivos, características, definición de los conceptos de confianza y reputación, funcionalidades, etc., para dar respuestas a disímiles problemas. A continuación se comentan algunas características específicas de los modelos más representativos del desarrollo alcanzando en este campo, agrupándolos de acuerdo a la dependencia del contexto y de los requerimientos de los usuarios.

2.2.1. Modelos independientes del contexto y de los requerimientos de los usuarios

Los modelos, cuyas características se comentan a continuación, consideran la confianza y la reputación como propiedades que no están relacionadas con el contexto del agente o con los requerimientos de los usuarios. Por lo general, tienen en cuenta que el agente se desenvuelve siempre en el mismo contexto para resolver el mismo requerimiento del usuario, o, que al menos, su comportamiento no está asociado a estos elementos. Así, en ninguno de ellos se ofrece una representación y gestión explícita del contexto de los agentes o de los requerimientos de los usuarios.

2.2.1.1. Sporas e Histos

Sporas e Histos son dos mecanismos de reputación complementarios para ambientes de comercio electrónico [146, 147]. Su funcionamiento no distan mucho de otros mecanismos de reputación *on-line* tales como los que se muestran en eBay ¹, Amazon Auctions ², OnSale Exchange ³, entre otros.

Por lo general, estos sistemas de comercio electrónico *on-line* mantienen información de reputación acerca de vendedores y consumidores a partir de la información recopilada al finalizar cada transacción. Los valores de reputación representan un indicador global del desempeño de los componentes del sistema, independiente del contexto o situación en que se encuentran; la información se mantiene actualizada en un registro centralizado; y no ofrecen mecanismos para la detección de información falsa.

Sporas es un mecanismo de reputación simple que puede ser implementado sin tener en cuenta el número de interacciones que se evalúen, solo se considera la evaluación de la interacción más reciente. Por su parte, Histos es un modelo más complejo que es capaz de

¹eBay. <http://eBay.com/>

²Amazon Auctions. <http://auctions.amazon.com/>

³OnSale Exchange. <http://onsale.com/>

considerar abundante cantidad de información relativa a las evaluaciones que un agente hace de otro al interactuar, considerando solo las experiencias más recientes [147].

En *Sporas*, un nuevo agente entra al sistema con reputación muy baja, que posteriormente se va actualizando con el transcurso de las interacciones entre los agentes. Los valores de reputación son actualizados después de cada interacción, teniendo en cuenta las opiniones de retroalimentación provenientes de otros agentes. *Sporas* gestiona un único valor de reputación global para cada agente e incorpora medidas de fiabilidad basadas en la desviación estándar de los valores de reputación. Establece la condición de que cada par de agentes solo pueden evaluarse una vez, así evita complicidades entre dos de ellos, no permitiendo que se sobrevaloren mutuamente. Sin embargo, pueden tener lugar complicidades entre más de dos agentes, cuestión esta que es resuelta por *Histos*.

Histos, aunque muy similar a *Sporas*, resuelve el problema de la falta de personalización de los valores de reputación. Considera la reputación como una propiedad intrínseca a la relación entre dos agentes. Para ello construye un grafo donde los nodos representan los agentes y los arcos la evaluación más reciente que un agente tiene sobre otro. La dirección de los arcos determina el sentido de la evaluación que un agente hace sobre otro: el nodo origen se corresponde con el agente que evalúa y el destino, con el evaluado. Esta representación es similar a la utilizada por Schillo et al. [116] y conocida como el nombre de *TrustNet*.

Para computar el valor de reputación entre dos agentes determinados, se recorre el grafo utilizando un algoritmo de búsqueda similar al Primero-en-Amplitud (*Breadth-First-Search*) [110] para encontrar todos los caminos entre ambos agentes, teniendo en cuenta posibles limitaciones en cuanto a su longitud. El número de caminos considerados también puede estar restringido de acuerdo a la actualidad de las evaluaciones que se tengan, por ejemplo, de todos los caminos posibles solo pueden utilizarse los que contengan las evaluaciones más recientes.

Una vez identificados los caminos, se utiliza un método recursivo para calcular el valor de reputación. Dicho valor resulta de la suma ponderada de todas las reputaciones obtenidas desde un nivel inferior. Los pesos en esta suma ponderada son los valores de los arcos del grafo (o sea, las evaluaciones directas que un agente hace de otro). En los casos donde no exista un camino entre los dos agentes en el grafo entonces se aplica el método que establece *Sporas* para el cálculo de la reputación.

Estos modelos establecen enfoques centralizados e independientes del contexto del agente para el manejo de la confianza. Aunque no tienen en cuenta mecanismos para la detección de mentiras, sí establecen medidas de confiabilidad de los valores de reputación.

Por otra parte, en el método de propagación de la reputación utilizado, no se establecen diferencias entre la reputación que tiene un agente desempeñando sus funciones y la reputación que tiene ofreciendo criterios sobre otros.

2.2.1.2. MARS

MARS (*Multiagent Referral System*) utiliza el modelo de reputación desarrollado por Yu y Singh [142, 143, 144, 145]. Este modelo se basa en almacenar un histórico de las experiencias directas de cada agente, a partir de evaluar la calidad de la respuesta ofrecida

en la última interacción. Utiliza la teoría de la evidencia de Dempster-Shafer [118] para representar y propagar explícitamente las evidencias de confianza a favor y en contra de cada agente considerado, suponiendo que la carencia de confianza no significa desconfianza. Para ello, cada agente asigna valores de certeza para representar la creencia que tiene sobre el comportamiento de otro. Estos valores de certeza no representan medidas de creencias absolutas sobre el desempeño de los agentes, sino sobre los cambios en las creencias.

Para la propagación de los valores de confianza, utiliza *TrustNet*, similar a Schillo *et al.* [116] y Zacharia [147] (esta estructura es descrita ligeramente al comentar las principales características de Histos en el apartado 2.2.1.1).

Considera dos tipos de fuentes de información de confianza: experiencias directas y opiniones de terceros. Sin embargo, no se combinan: se utiliza la información disponible sobre las experiencias directas si están disponibles, en caso contrario se utiliza la información brindada por terceros.

2.2.1.3. AFRAS

AFRAS es un modelo que utiliza conjuntos difusos [148] para representar la reputación [28, 27]. Su aplicación principal ha estado relacionada con la selección automática de mercado. Según este modelo, un agente mantiene sobre cada uno de sus conocidos una única medida difusa de la reputación, no asociada a ninguna tarea en particular, que es actualizada a partir de la suma ponderada de dos valores de reputación: el viejo valor de reputación almacenado y el nuevo valor de satisfacción dada la solución ofrecida en la interacción. Los pesos que son utilizados para combinar estos dos valores dependen de la similitud entre la interacción que se está evaluando y la anterior.

Los valores iniciales de reputación asociados a agentes cuyo desempeño se desconoce se determinan utilizando criterios de similitud entre los gustos de los agentes y mediante la demostración de la existencia de interacciones anteriores (compras en el dominio utilizado por AFRAS). Dado que en AFRAS la reputación se adquiere mediante la similitud de evaluaciones subjetivas, mostrando únicamente la similitud un agente puede recibir una valoración acorde a las preferencias de quien lo está evaluando. Así, un agente malicioso puede aprovechar esta situación y obtener valores de reputación elevados a partir de evaluaciones falsas. Por tal razón se considera la existencia de interacciones anteriores, como justificación de dichos valores de reputación.

Mediante la combinación de la demostración de los gustos, y de la existencia de compras anteriores, la aceptación de agentes desconocidos se basa en criterios justificados y no en valores prefijados arbitrarios. Sin embargo, estas dos pruebas pueden revelar los gustos y el perfil de compra del agente desconocido sin que el otro agente esté comprometido a nada. Por ello, también implementa mecanismos de seguridad para la protección de la intimidad de los gustos, tratando de que la cantidad de información que las partes comparten sea equitativa, y mantengan el anonimato y autenticidad de ambas pruebas.

Este modelo ha adoptado la representación de los valores de reputación mediante conjuntos difusos dada la propia naturaleza subjetiva e inexacta de la reputación. La idea de confiabilidad de cada valor de reputación está modelada por el área y la forma de los

propios conjuntos difusos. Un conjunto difuso de área amplia representa un valor con un elevado grado de incertidumbre, con poca confiabilidad. En cambio, los valores dados por conjuntos de área estrecha ofrecen valores de elevada confiabilidad.

Tiene en cuenta conceptos tales como la timidez, el egoísmo, la susceptibilidad y el rencor. El modelo considera que los agentes se comportan sin timidez, es decir, evitando el aislamiento, y, en consecuencia, ofrecen más opciones a los agentes desconocidos. El egoísmo decide si se ejecuta o no la intención de responder a una pregunta de otro agente en lugar de ejecutar la intención de preguntar a otro agente. La susceptibilidad representa cuánto recelo tiene un agente a la hora de tomar decisiones con un coste económico asociado. En tanto, el rencor, es definido como el recuerdo de la reputación previa del agente al calcular el nuevo valor de reputación [28].

2.2.1.4. Modelo de Carter *et al.*

Este modelo, a pesar de no considerar el contexto, ni los requerimientos de los usuarios, en la obtención de los valores de reputación, sí asocia estos valores a los posibles roles que puede desempeñar un agente. Cada valor de reputación indica el grado de cumplimiento de cada rol desempeñado. Para ello, una entidad central otorga valores positivos o negativos para indicar el cumplimiento de cada rol, previamente definido por el entorno. Un valor positivo indica un buen desempeño del rol, mientras que uno negativo, un mal desempeño.

Dicha entidad centralizada también ofrece, a todos los agentes en el sistema, un valor único de reputación a partir de la agregación de los valores para cada rol. Los pesos de dicha agregación son establecidos por la propia entidad, dependiendo de la comunidad de agentes [30].

Este modelo solo es aplicable a entornos donde el conjunto de roles de un agente es definido previamente. Es un ejemplo de modelo donde los valores de reputación están ligados a preferencias o características propias del agente, en este caso definido a partir del rol a desempeñar.

2.2.1.5. TRAVOS

TRAVOS es un modelo de confianza y reputación que considera que los agentes son egoístas y que pueden ofrecer información falsa acerca de su desempeño y el de otros agentes. Este modelo presta especial atención al tratamiento de información de reputación falsa [101, 126].

La confianza es calculada usando la teoría de probabilidades a partir de la información almacenada sobre el desempeño de los agentes en las interacciones pasadas. De esta forma, el modelo ofrece una métrica de confianza que permite establecer la comparación entre los diferentes agentes con el objetivo de identificar los que tengan mayor trascendencia en la decisión a tomar. Por otra parte, cada agente no asume que las opiniones de otros son ciertas. Los criterios de otros están modificados por los valores de reputación, basados en la confiabilidad y consistencia de quienes ofrecen las opiniones.

Mediante evaluaciones sencillas, en [126] se demuestra que mediante la utilización de TRAVOS se eleva el desempeño a la vez que los agentes se protegen contra el efecto negativo de recibir opiniones falsas.

Este modelo asume que el comportamiento de los agentes no cambia en el tiempo. En este sentido, necesitaría al menos considerar que las experiencias pierden relevancia con el tiempo.

En este modelo, las experiencias de los agentes acerca del desempeño de otros en interacciones anteriores solo indican si la interacción ha sido o no exitosa. Luego, a partir del conteo de las interacciones exitosas y las fallidas, se establecen las métricas de reputación que utiliza el modelo. Este espacio de representación binario constituye una limitante del modelo.

2.2.1.6. FIRE

FIRE es un modelo de confianza y reputación que integra varias fuentes de información para brindar valoraciones exhaustivas acerca del desempeño de los agentes. FIRE considera la confianza basada en las interacciones, la confianza basada en roles, la reputación a partir de testigos y la reputación certificada para brindar una métrica, sobre el comportamiento de los agentes, en cualquier circunstancia [69, 70].

La confianza basada en las interacciones se determina a partir de la experiencia de los agentes en sus interacciones directas. Las métricas que se obtienen bajo este tipo de confianza son similares a las utilizadas en ReGreT (sección 2.2.2.5).

La confianza basada en roles se obtiene a partir del análisis de las relaciones de roles de los agentes. Para este tipo de confianza, utiliza reglas que permiten estimar la confianza en un agente a partir del rol que desempeña. Cada agente tiene su propio conjunto de reglas que son evaluadas cada vez que necesita conocer la confianza basada en el rol que desempeña un determinado agente.

La reputación a partir de testigos se obtiene utilizando la información recopilada por otros agentes. Para identificar los agentes que pueden ofrecer información acerca de otros, utiliza un método similar a MARS [145]. Sin embargo, incorpora métodos para evaluar la credibilidad de los testigos que ofrecen este tipo de información de reputación [71].

La reputación certificada está dada por evaluaciones que cada agente mantiene sobre sí mismo a partir de las opiniones recibidas desde otros agentes en interacciones anteriores [69, 72, 73]. Este tipo de reputación permite a los agentes probar su desempeño en interacciones previas, según el criterio de los otros agentes. Se asume que los valores de reputación certificada se encuentran codificados según algún mecanismo de seguridad (por ejemplo, alguna variante de infraestructura de clave pública [124]) para evitar que cada agente sobredimensione su propio valor de reputación. Este tipo de información de reputación es útil cuando se cuenta con poca información para determinar otras medidas de confianza.

2.2.1.7. Modelo de Wang y Singh

Wang y Singh [131, 133] muestran un modelo de confianza y reputación que apoya a los usuarios en encontrar servicios ofrecidos por agentes. En el modelo se representa la confianza en términos de creencias y certezas, y mediante una distribución de certeza de probabilidades [132] se definen dos operadores para combinar los valores de confianza: el de concatenación y el de agregación. El operador de concatenación se utiliza para calcular la propagación de los valores de confianza a través de un camino sobre la red social que se establece a partir de las relaciones entre los agentes. Por otro lado, el operador de agregación se utiliza para la combinación de los valores de reputación obtenidos mediante varios caminos sobre la red.

Aunque el enfoque utilizado para representar las relaciones sociales de los agentes mediante un grafo es bastante intuitivo, el modelo ofrece una forma novedosa de definir y la propagar la confianza sobre dicho grafo. En [132] se define la confianza en términos de *probabilidades de probabilidades*. Y en [131] se comenta la definición y utilización de los operadores de concatenación y agregación sobre el grafo de relaciones sociales. Los operadores son definidos utilizando un enfoque similar al utilizado por Richardson *et al.*[108], donde la gestión de reputación es tratada como un problema bien-definido del álgebra de caminos [29].

Al igual que en otros modelos [116, 145, 146], en éste los agentes solo necesitan mantener actualizados los criterios sobre los demás y utilizar los operadores convenientemente para estimar la reputación de un determinado agente contraparte.

Estos métodos de concatenación y agregación de los caminos sobre el grafo de relaciones sociales entre los agentes, permiten estimar la confianza en proveedores de servicios sobre los cuales no se poseen experiencias. En [131] se concluye que el uso de estos operadores reduce los efectos de la decepción por parte de algunos agentes.

2.2.2. Modelos dependientes del contexto o de los requerimientos de los usuarios

Según se ha comentado en la sección 2.1.2, existe un grupo de modelos donde la confianza y reputación son dependientes del contexto o de los requerimientos de usuario a los cuales se les pretende dar solución [2, 33, 84, 105, 113]. En modelos de este tipo, cada agente puede caracterizarse por diferentes valores de confianza y reputación para cada uno de los contextos o preferencias de los usuarios que se consideren. Teniendo en cuenta el carácter multi-contextual que pueden mostrar la confianza y la reputación, estos modelos ofrecen mecanismos para gestionar varios contextos al unísono. Por ejemplo, un agente comprador puede tener valores de confianza y reputación diferentes en dependencia del mercado en el que se encuentre. Esta idea se extiende a la diferenciación de los valores según las preferencias de los usuarios. Por ejemplo, un agente, aun en el mismo mercado, puede tener asociados valores diferentes para comprar pescado fresco o calzado.

A continuación se comentan las principales características de modelos de este tipo. En ellos, los valores de confianza dependen del contexto de los agentes o de las necesidades de

los usuarios (la presencia de algunos requerimientos o preferencias pueden ser tratados de manera similar a las características que definen el contexto de un agente).

2.2.2.1. Modelo de Marsh

Uno de los primeros modelos de confianza fue propuesto por Marsh en 1994 al formalizar algunos términos y conceptos relacionados con la confianza en sistemas multi-agente [84]. Tiene en cuenta que los valores de confianza están determinados por las relaciones que se establecen entre dos agentes interesados en cooperar en el marco de una situación determinada. Los valores de confianza relativos a una situación dada se obtienen a partir de la agregación de funciones que tienen en cuenta la confianza general que tiene un agente en el otro (independientemente de la situación en que se encuentren), la utilidad (relación costo/beneficio) que el agente puede generar a partir de la situación, y la importancia de la situación para dicho agente.

Según el modelo, la confianza general que un agente puede tener en otro es estimada a partir de las experiencias anteriores del agente, en correspondencia con el enfoque optimista, pesimista o realista que tenga el agente. También tiene en cuenta un concepto de memoria, asociado a la vigencia de las experiencias, tomando solo las más recientes. Utiliza un mecanismo de ventana de tiempo para considerar las experiencias previas en el cálculo de la confianza. Un agente de enfoque optimista toma la mejor experiencia que tenga en su memoria. Por su parte, uno pesimista toma la peor. En cambio, uno realista toma la media de las experiencias incluidas en la ventana de tiempo considerada.

Considera que la confianza no debe ser un único elemento que guíe el proceso de toma de decisiones de las interacciones de un agente, sino un elemento que permite valorar la importancia y el riesgo de la posible interacción con sus vecinos, así como una estimación de las competencias de éstos.

2.2.2.2. Modelo cognitivo de Castelfranchi y Falcone

El modelo desarrollado por Castelfranchi y Facone [33] es un ejemplo de modelo cognitivo donde la confianza es considerada la componente mental de la delegación de tareas.

El concepto de confianza utilizado por este modelo está estrechamente ligado a la delegación de tareas. La confianza es el resultado mental de la decisión de a quién delegar una tarea, obtenido a partir de la consideración de las creencias y metas del agente. Tal como se define, *"es una actitud compleja de un agente frente a otro, en relación a su comportamiento para obtener una meta determinada"* [33].

Para representar el estado mental determinado por la confianza, se utiliza un predicado de tres argumentos que relaciona los dos agentes involucrados y el contexto que se considera. De igual forma, también utiliza un predicado de tres argumentos para representar el resultado de la delegación de una tarea.

La confianza en sí misma está compuesta por creencias asociadas al contexto del agente. El estado mental que representa la confianza se componen de diferentes tipos de creencias

definidas a partir de 1) las competencias de un agente para ejecutar determinada tarea, 2) la disposición del agente para llevarla a cabo, 3) la dependencia que se tiene del otro agente para la realización de dicha tarea, y 4) el cumplimiento de la tarea por parte del agente.

El modelo define varios grados de confianza como funciones de la fortaleza de las creencias del agente, por ejemplo, su credibilidad (que refleja tanto la probabilidad subjetiva como la confianza en una creencia dada). El grado de confianza es utilizado para formalizar el proceso de toma de decisiones relacionadas con la delegación. El grado de importancia de la confianza se obtiene mediante la multiplicación del grado de importancia de las creencias y metas involucradas en la relación de confianza, expresadas en términos de funciones de utilidad asociadas a cada alternativa. La decisión de delegar la tarea asociada, según las creencias y metas del agente, es tomada si el valor de importancia asociado a la decisión sobrepasa un determinado umbral.

2.2.2.3. Modelo de Abdul-Rahman y Hailes

El modelo desarrollado por Abdul-Rahman y Hailes [2] determina la confianza y reputación utilizando la información recopilada a partir de las experiencias directas de cada agente y las recomendaciones que se hacen sobre otros. Los valores de confianza son propiedades inherentes a un par de agentes respecto a un contexto determinado, definido previamente de acuerdo al dominio de aplicación del modelo.

Tanto para evaluar el desempeño de los demás, como la su capacidad para recomendar a otros, se establecen cuatro categorías de confianza entre los agentes: vg (muy bueno), g (bueno), b (malo) y vb (muy malo).

El modelo supone que, para cada uno de sus vecinos y cada uno de los contextos, cada agente mantiene una tupla con la cantidad de experiencias de estos cuatro tipos. De esta forma, el valor de confianza directa que al agente da a su vecino para dicho contexto es el resultante de la categoría con mayor cantidad de experiencias. Si existen categorías con igual número de experiencias, el agente considera tres posibles grados de incertidumbre a partir de las cantidades de experiencia en cada una de las cuatro categorías.

Al finalizar cada interacción, el agente actualiza la confianza directa que tiene sobre los otros agentes, así como la confianza que les otorga como recomendadores. Para la confianza directa se incrementa el contador correspondiente al valor de la nueva experiencia. Mientras que para asignar la confianza a otros agentes como recomendadores, el agente combina su criterio con la opinión ofrecida en caso de que estas no coincidan. Al combinar las opiniones, se tienen en cuenta la cantidad de niveles de confianza en que se diferencian las opiniones y se da mayor importancia a las que provengan de agentes con puntos de vistas similares en ese contexto.

Aunque el modelo da una forma de analizar la información proveniente de terceros antes de incorporarla como información útil para futuros análisis, tiene la limitante de no reconocer entre criterios mal intencionados y criterios diferentes a los del propio agente.

2.2.2.4. Modelo bayesiano de Mui *et al.*

Mui *et al.* [95, 96, 97] presentan un modelo de reputación donde las recomendaciones hechas por los agentes son ponderadas por la reputación que éstos tienen, según el contexto en cada caso. El contexto del agente está definido, de manera muy simple, a partir del criterio para el cual el agente está ofreciendo su recomendación y para el cual se está estimando su reputación (por ejemplo, precio, tiempo de espera, calidad, accesibilidad, entre otros criterios de selección de las recomendaciones). De esta forma, los valores de reputación asociados a un mismo proveedor pueden ser diferentes según el criterio seguido para ofrecer una determinada recomendación.

Para la manipulación de los valores de reputación, se utiliza un modelo probabilístico bayesiano. Bajo este enfoque, la reputación, que un agente a asigna a otro agente b para un contexto dado, es estimada a partir de la probabilidad que b sea seleccionado por a (para interactuar con él) en su próxima interacción en ese contexto. El valor de probabilidad de que sea seleccionado se estima a partir de la cantidad de interacciones anteriores con el agente b en dicho contexto.

Cuando los agentes no han tenido interacciones previas, los valores de reputación son inferidos mediante la propagación de las evaluaciones realizadas por los agentes vecinos.

Mui *et al.* [94] también estudian el comportamiento de los agentes a partir de la comparación experimental de diferentes poblaciones de agentes, donde cada tipo de agente es definido por la estrategia de cooperación utilizada. Para ello, extiende el juego clásico del problema del prisionero, incorporando nuevas estrategias. Las estrategias consideradas son las siguientes: (C) siempre cooperar, (D) siempre desertar, (TFT) cooperar al inicio y luego hacer lo que haga el oponente en la última ronda, (RTFT) cooperar al inicio en dependencia del valor de reputación asociado y luego hacer lo que haga el oponente en la última ronda. Las principales características del escenario definido para la experimentación con este modelo son comentadas en el apartado 3.2.1.3.

2.2.2.5. ReGreT

ReGreT [111, 113, 114] es un modelo de confianza y reputación modular orientado a entornos complejos de comercio electrónico, donde las relaciones sociales juegan un papel importante. La estructura de ReGreT está determinada por tres dimensiones del concepto de reputación considerado por el modelo: la individual, la social y la ontológica. Cada una ofrece una perspectiva diferente de la reputación entre los agentes.

Además de los módulos encargados de gestionar la reputación desde estas tres dimensiones, el modelo ofrece otros módulos que ofrecen indicadores de credibilidad a cerca de los valores de reputación estimados. Todos estos módulos, en su conjunto, ofrecen un modelo completo de confianza y reputación basado en el conocimiento directo y la reputación. Sin embargo, su diseño modular permite que el agente pueda decidir qué partes utilizar, adaptándose, entre otros, a la disponibilidad de las diferentes fuentes de información [111].

La dimensión individual tiene en cuenta únicamente la información que se obtiene a

partir del conocimiento relacionado con las interacciones directas con otros miembros de la sociedad de agentes. En cambio, la dimensión social hace referencia a la información que se intercambia entre los miembros de la sociedad y a la que se obtiene a partir del análisis de las relaciones que se dan entre ellos. Por otra parte, la dimensión ontológica es analizada desde el punto de vista de las diferentes facetas (dado por los intereses, criterios, etc.) que puede tener la reputación, ofreciendo la forma de combinar adecuadamente cada una de ellas.

Bajo la dimensión individual, el modelo ofrece un valor de confianza directa a partir del conocimiento que tiene sobre el desempeño de los agentes en las interacciones anteriores. Para ello, combina la información sobre las interacciones anteriores mediante un modelo descontado, donde las experiencias pierden relevancia con el transcurso del tiempo.

Los valores de confianza directa tienen asociado métricas que indican su confiabilidad. Esta medida se estima a partir de la desviación (de las experiencias en torno al valor de confianza directa calculado) y de la cantidad de experiencias que se utilizan para obtenerlo. [111, 113, 114].

Aunque la información proveniente de las interacciones directas de un agente constituyen la fuente más fiable de información, no siempre se dispone del conocimiento suficiente para garantizar una confiabilidad mínima en las medidas que se obtienen. En estos casos se recurre a analizar las relaciones con otros agentes. La dimensión social tiene en cuenta el conocimiento proveniente de la comunidad de agentes y sus relaciones.

El concepto de reputación ofrecida en la dimensión social es dividida en tres tipos especializados de reputación, en dependencia del tipo de fuente de información utilizada para obtenerlo. Si la reputación es calculada a partir de la información proveniente desde testigos, se le llama reputación brindada por testigos. Si la reputación es calculada utilizando información extraída de la red de relaciones sociales entre los agentes contrapartes, se le llama reputación brindada por el vecindario. Y, por último, si la reputación está basada en roles y en propiedades generales de los agentes se le llama reputación brindada por el sistema [111].

Los valores de confianza directa y reputación, para estas dos dimensiones (individual y social, respectivamente), se unifican en un único valor a partir de la medida de confiabilidad calculada bajo la dimensión individual.

Por otra parte, la dimensión ontológica ofrece la posibilidad de combinar, en un único valor, los valores de confianza y reputación obtenidos para cada uno de los criterios o aspectos analizados según las dimensiones individual y social. Para representar la dimensión ontológica se utilizan grafos, cuya configuración depende del dominio de aplicación tratado. Cada nodo en el grafo representa un criterio o aspecto de la reputación analizado, mientras que los arcos indican la relevancia que tiene cada uno de los criterios al ser combinados para obtener un criterio más complejo. La configuración del grafo que determina la dimensión ontológica de la reputación puede cambiar en función del estado mental del agente y del dominio de aplicación que se trate.

ReGreT ofrece dos conceptos para el cálculo de la confianza en un agente: la confianza directa y la reputación. De igual forma, el valor de reputación es la combinación de los diferentes tipos de reputación (basada en testigos, en el vecindario, y en el sistema). El

modelo propone un mecanismo de agregación de todos estos elementos para obtener un valor final de confianza. Mediante un conjunto de experimentos, definidos sobre escenarios concretos, ReGreT ofrece un ordenamiento de la importancia de la información proveniente de cada una de las fuentes de información. Sin embargo, la importancia que otorga a cada fuente depende del dominio de aplicación considerado en cada uno. El ordenamiento de la importancia de las fuentes varía en dependencia de las características de los escenarios.

Por ejemplo, en escenarios donde es fácil obtener experiencias directas es mejor no utilizar información de confianza, más aun si la posibilidad de que los agentes mientan es elevada. De igual forma, en escenarios donde la política de grupo es importante y existen fuertes vínculos entre los miembros, la reputación obtenida a partir del vecindario y del sistema adquiere gran relevancia.

De igual forma que se combinan la confianza directa y la reputación, utilizando el grado de fiabilidad de la propia confianza directa, puede pensarse en mecanismos basados en el análisis de las relaciones sociales para detectar las características del entorno y decidir la importancia relativa de cada fuente de información.

2.2.2.6. Modelo difuso de Ramchurn *et al.*

El modelo propuesto por Ramchurn *et al.* [105] utiliza un enfoque difuso para representar las creencias de un agente gestor acerca de cuan confiable es un agente contratado para garantizar los resultados esperados en determinados aspectos de un contrato. La confiabilidad del agente contratado se determina a partir de la confianza y la reputación, fundamentalmente. De manera similar a otros modelos, la confianza se obtiene utilizando las experiencias que se tienen de interacciones anteriores, mientras que la reputación se obtiene a partir del conocimiento adquirido por otros agentes. Ambas miden la misma propiedad, la creencia del agente gestor acerca de la fiabilidad de un agente contratado para cumplir un aspecto concreto de un contrato. Estas medidas son modeladas utilizando conjuntos difusos [148] para brindar a los agentes un mecanismo robusto para valorar en qué medida su agente contraparte satisface un determinado aspecto del contrato acordado.

En este modelo, un contrato es un acuerdo sobre un conjunto de elementos y los valores que estos elementos deben tomar. Por lo tanto, un contrato puede verse como un conjunto de asignaciones *elemento - valor*. En este sentido, pueden existir agentes más confiables que otros, en tanto sean capaces de garantizar los valores más adecuados para los elementos de un contrato.

Para evaluar el desempeño de un agente contratado sobre cada elemento del contrato, se utiliza un conjunto (relativamente pequeño) de etiquetas lingüísticas (que comparten todos los agentes). Estas etiquetas sirven para representar el punto de vista del gestor sobre el cumplimiento de un determinado criterio en un contrato. Así, el agente gestor representa su creencia (según el punto de vista ofrecido por la etiqueta) acerca de la posible diferencia entre el valor acordado y el valor real ofrecido por el agente contratado para un determinado elemento del contrato. En otras palabras, la confianza del agente gestor sobre el contratado para un determinado elemento del contrato, según una etiqueta lingüística

dada, está determinada por la función de pertenencia a dicha etiqueta.

El modelo plantea la forma en que el agente gestor puede obtener la confianza en el contratado para un aspecto del contrato. La confianza en un aspecto se determina a partir de la intersección de las funciones de pertenencias a cada una de las etiquetas lingüísticas para dicho aspecto del contrato.

De manera similar son representados los valores de reputación para cada par de agentes asociados a un determinado aspecto del contrato. El valor de reputación de un agente contratado, para un determinado aspecto, se obtiene a partir de la intersección de las funciones de pertenencia (de los valores de reputación) a cada una de las etiquetas lingüísticas para dicho aspecto.

Siguiendo un enfoque similar a ReGreT [111], se combinan la confianza y la reputación mediante la suma ponderada, donde los pesos establecen la importancia relativa de cada fuente de información.

2.2.2.7. TrustMail y FilmTrust

TrustMail [54, 55] y *FilmTrust* ⁴ son aplicaciones que utilizan modelos de confianza en su funcionamiento. *TrustMail* es un prototipo de cliente de mensajería electrónica que tiene en cuenta criterios de reputación para la recomendación de mensajes. Por su parte, *FilmTrust* es un sitio Web de evaluación de películas, basada en los criterios de los usuarios y las relaciones que se establecen entre ellos.

Golbeck *et al.* [52, 53, 54, 55, 56, 77, 78, 151] han realizado diferentes estudios sobre estas aplicaciones, desarrollando diferentes modelos de confianza basados en las relaciones sociales que se presentan entre los agentes. Sus trabajos integran los conceptos relativos a la confianza, su relación con la similitud entre los perfiles de los usuarios, el análisis de las redes sociales y su relación con la web semántica, la utilización de modelos de confianza para la priorización de reglas en sistemas de inferencia, entre otros.

En [53] se describe la aplicación de las redes de relaciones sociales a la web semántica para el tratamiento de la confianza. En la red de relaciones sociales utilizada, cada arco entre dos nodos representa el valor de confianza que un nodo asigna al otro. Los valores de confianza se definen a partir de nueve categorías que se corresponden con valores enteros en el intervalo [1, 9]. Los valores de reputación, entre dos nodos cualesquiera de la red, se obtienen a partir de la combinación de la reputación asociada a cada arco, teniendo en cuenta los posibles caminos entre los nodos.

Golbeck y Hendler [54, 55], mediante la aplicación *TrustMail* para el filtrado de correos, proponen un algoritmo para la inferencia de indicadores de reputación dada la red de relaciones sociales entre los individuos. También ofrecen indicadores para medir la precisión de los valores de reputación inferidos de la red.

Ziegler y Golbeck [151] han desarrollado estudios acerca de la correlación que existe entre la confianza y la similitud de los intereses de los usuarios en casos donde el empleo de modelos de confianza reemplaza o complementa técnicas utilizadas en sistemas

⁴FilmTrust. <http://trust.mindswap.org/FilmTrust/>

recomendadores tradicionales. El modelo de confianza no solo es utilizado para recomendar los vecinos más prometedores, sino también para eliminar aquellos que no resultan siquiera interesantes.

Se parte del supuesto de que la confianza en los usuarios puede variar de acuerdo a la similitud de intereses entre ellos. Por ejemplo, se otorga mayor confianza a la recomendación hecha por un agente de intereses similares. De esta forma, los intereses de los usuarios y la similitud entre ellos puede considerarse un elemento relativo al contexto de la confianza [2]. Por otra parte, y en situaciones más específicas, los intereses de los usuarios pueden estar influenciados por la confianza que se tenga en otro usuario o grupos de usuarios. Por ejemplo, el círculo de amigos es probable que comparta los mismos intereses o al menos los tolere.

Los perfiles de los usuarios son definidos a partir de los intereses de los usuarios, representados por medio de vectores; cada dimensión indica el interés por una temática específica. Las temáticas están categorizadas en una estructura arbórea, utilizada para la estimación de los intereses no conocidos. La similitud entre los intereses de dos usuarios se obtiene mediante funciones que consideran los vectores correspondientes.

Para el estudio de la correlación entre la confianza y la similitud de los perfiles de los usuarios, se utilizan técnicas de minería de datos a partir de la información almacenada en dos sistemas recomendadores, uno de libros (*All Consuming*⁵) y otro de películas (*FilmTrust*). Se crean grupos de usuarios a partir de los valores de confianza que se tiene en cada uno. (La confianza se representa en una escala entera del 1 al 10.) Así, para cada una de las 10 categorías de la confianza, se determina el promedio de las similitudes entre las evaluaciones que, cada par de usuarios hacen de un mismo elemento (libro o película). Demuestran con ello que se presenta una correlación fuerte entre la confianza y la similitud: al aumentar la confianza en los usuarios, disminuyen significativamente las diferencias entre las evaluaciones que hacen de un mismo elemento.

De este modo, recomiendan la utilización de información relativa a la similitud entre los perfiles de los usuarios cuando no se dispone de conocimiento suficiente acerca de las relaciones de confianza entre ellos. Comentan cómo para determinados escenarios, se pueden estimar los valores de confianza al utilizar criterios de similitud entre los perfiles (intereses) de los usuarios.

En trabajos siguientes [52], siguiendo una representación similar, Golbeck investiga la relación entre algunas facetas de la similitud entre los perfiles de los usuarios y la forma en que se determina la confianza. Mediante resultados experimentales precisa cuáles son las facetas de la similitud más importantes según el criterio de los usuarios y cómo deben ser utilizadas para predecir con precisión la confianza entre los usuarios de *FilmTrust* [56]. Según los datos obtenidos de *FilmTrust*, demuestra que la correlación entre la confianza y la similitud se hace más fuerte cuando se consideran diferentes facetas de la similitud que en los casos donde se tiene en cuenta una única medida de similitud.

Este análisis aunque sigue un enfoque general de los conceptos de similitud y confianza, está basado en parámetros establecidos por la aplicación utilizada *FilmTrust*, pudiendo no

⁵All Consuming. <http://www.allconsuming.net/>

ser medibles en otros entornos.

El trabajo señala la carencia de algoritmos que combinen la información de confianza proveniente de las relaciones sociales entre los agentes con la similitud de los perfiles de los usuarios. Sin embargo, es posible mejorar la precisión de los algoritmos de inferencia de la confianza utilizando toda la información disponible [52].

2.2.2.8. Modelo bayesiano de Wang y Vassileva

Wang y Vassileva [135, 136, 137] han propuesto un modelo de confianza y reputación que considera diferentes aspectos de ambos conceptos y su combinación a partir de redes bayesianas en un entorno basado en la interacción entre pares.

Consideran que la confianza es dependiente del contexto y multidimensional. Para cada una de las diferentes dimensiones (o facetas) de la confianza y reputación, los pares mantienen dos tipos de información que indican, por una parte, la capacidad de un par (agente) para ofrecer servicios, y por otra, la capacidad para ofrecer recomendaciones sobre el comportamiento de otros agentes. Un aspecto interesante en este modelo es que no solo se solicita recomendaciones a pares confiables y bien conocidos, sino que se adicionan pares cuya habilidad para ofrecer recomendaciones aún se desconoce.

Las diferentes dimensiones de la confianza y la reputación son combinadas en un único valor, que permite seleccionar el par más adecuado, utilizando una red bayesiana. Los nodos hojas representan las diferentes dimensiones de la confianza que se van combinando según la configuración de la red. Al finalizar cada interacción, cada par ajusta, según sus preferencias, la red que utiliza.

2.2.2.9. Modelo de Maximilien y Singh

Maximilien y Singh [86, 87, 88] utilizan en sus estudios un sistema multi-agente para determinar el servicio más adecuado según las preferencias del consumidor de servicios a partir de la confianza en el propio servicio y en su proveedor. El modelo de confianza se basa en compartir una ontología para la representación de las calidades de los servicios y un modelo de preferencias que tiene en cuenta las evaluaciones de los consumidores sobre dichos servicios [85]. El proceso de selección del servicio más adecuado, según las preferencias especificadas, se reduce a la selección del servicio más confiable, teniendo en cuenta el conocimiento representado en la ontología. En el marco ontológico que se ofrece, a cada servicio web se le asocia un conjunto de propiedades no funcionales, cuyos valores para cada instancia son utilizados en la determinación de la confianza asociada al servicio. Los atributos no funcionales planteados por Maximilien y Singh en [85] son: disponibilidad, interoperabilidad, desempeño, fiabilidad, robustez, escalabilidad, seguridad, integridad y estabilidad.

El modelo identifica un conjunto de propiedades no funcionales relacionadas con la calidad de los servicios (por ejemplo, capacidad, tiempo de respuesta, etc.). La confianza en un proveedor de servicio determinado está dada por la relación entre los valores de calidad de las preferencias del consumidor y por los compromisos de calidad del proveedor.

Así, el problema de seleccionar el proveedor más adecuado (más confiable) se resume a obtener un valor único de confianza acerca de la calidad que ofrece cada uno de los proveedores de servicios. Para cada proveedor se deben agregar todos los valores de calidad de las propiedades no funcionales de los servicios que ofrece, según las preferencias de los consumidores.

Dado que la calidad mostrada por un servicio varía en el tiempo, y que las opiniones recientes son un mejor indicador de la calidad actual del servicio, el modelo asigna mayor importancia a la información más reciente. De esta forma, se propone un modelo descontado en el tiempo para agregar la información relacionada con las propiedades no funcionales de los servicios.

Para establecer la relación entre las calidades ofrecidas por los servicios y las preferencias de los usuarios, el modelo define una función de utilidad de la calidad para el consumidor. Esta función indica el nivel de preferencia del consumidor por las calidades ofrecidas por el servicio.

Por otra parte, para comparar las calidades de los servicios ofrecidos, el modelo utiliza tres operadores de relación:

- *indiferencia*(Q_a, Q_b) - que indica que el consumidor es indiferente a una calidad u otra,
- *preferencia estricta*(Q_a, Q_b) - para indicar que estrictamente se prefiere la calidad Q_a a la Q_b , y
- *preferencia*(Q_a, Q_b) - que indica que se prefiere estrictamente la calidad Q_a a Q_b o bien se es indiferente.

De esta forma, se establecen las posibles relaciones entre las calidades ofrecidas por los proveedores de servicios y las requeridas por los consumidores.

En este modelo, la confianza se obtiene mediante una evaluación directa de la correspondencia entre las preferencias de los consumidores y las calidades de los servicios ofrecidos por los proveedores. El concepto de confianza tratado supone la agregación de varias dimensiones definidas a partir de cada una de las propiedades no funcionales que describen la calidad de los servicios.

2.2.2.10. MDT-R

Los primeros trabajos de Griffiths y Luke [61, 62, 63] utilizan las nociones de motivación y confianza para resolver problemas relativos a la cooperación entre los agentes. En estos casos, la confianza ofrece mecanismos para modelar y razonar acerca de la confiabilidad, honestidad, veracidad, etc. También estudian la formación de clanes de agentes, formados por agentes que confían unos en otros y que tienen objetivos similares, y su efecto en algunos problemas propios de la cooperación entre agentes. Dentro de estos clanes los agentes se favorecen unos a otros, tratando de obtener beneficios mutuos [61].

Luego, Griffiths muestra un modelo de confianza basado en experiencias para minimizar los riesgos asociados con la cooperación entre agentes mediante la delegación de tareas en

un entorno Grid. La información de confianza se combina con otros factores para apoyar el proceso de toma de decisiones en la selección de los agentes que pudieran cooperar [57, 59].

Trata el concepto de confianza multi-dimensional que permite modelar la confianza entre agentes según diferentes criterios o preferencias de los usuarios. Las preferencias son representadas por un conjunto de factores ponderados que permiten que la información de confianza sea adaptada a las prioridades de cooperación del agente. Las dimensiones consideradas para la evaluación de la confianza, en este modelo, son: el éxito (la confianza que se tiene en que el agente contraparte desempeñará la tarea delegada con éxito), el costo (la confianza en que el costo que el agente ofrece no excede el costo esperado), el tiempo (la confianza en que el agente ejecutará la tarea en el tiempo esperado) y la calidad (confianza en que la ejecución de la tarea tendrá como mínimo la calidad esperada).

En trabajos siguientes [58, 60] relacionados con el concepto de confianza multi-dimensional, se presenta un enfoque difuso para el tratamiento de la confianza. Se brindan mecanismos que permiten a los agentes razonar con información con elevado grado de imprecisión sobre la confiabilidad de otros agentes. Además el concepto de confianza, aborda mediante la lógica difusa los conceptos de desconfianza y de falta de confianza, para elevar la efectividad de las interacciones entre los agentes.

2.2.2.11. Un modelo basado en la teoría de la información: Modelo de Sierra y Debenham

Sierra y Debenham [120, 121] proponen un modelo basado en la teoría de la información, que tiene en cuenta aspectos de confianza, honor y preferencias al determinar los mecanismos para gestionar los diálogos entre los agentes.

Consideran que la negociación es un proceso de intercambio de información entre agentes, en el que cualquier información que los agentes comuniquen, o rechacen comunicar, puede ser información valiosa. Sin embargo, la información intercambiada puede ser más o menos fiable. Tal información es valiosa en el sentido que pueda ser asimilada y contrastada con el grado de confiabilidad que se tenga en ella y en quien la brinda. Por lo general, la integridad de la información decrece con el transcurso de la negociación, estos valores de confiabilidad pueden ayudar a contrarrestar esta decadencia.

Se utiliza la teoría de la información para evaluar la negociación que se está realizando. El valor de la negociación está dado por la disminución de la incertidumbre sobre el comportamiento del otro agente, a partir de la información que genera dicha negociación. La medida de la disminución de la incertidumbre utilizada por el modelo es la entropía negativa [82]. La teoría de la información puede ser utilizada también para medir las características de las relaciones entre los agentes más allá de un simple proceso de negociación.

Se considera la confianza como una medida de la desviación del comportamiento esperado del otro agente según determinado criterio, que puede ser utilizada para: 1) seleccionar la oferta de la negociación, 2) seleccionar los agentes contrapartes, y 3) simplificar los diálogos (dado que de antemano se pueden estimar los términos y condiciones de la negociación, sin tener que negociarlos una y otra vez).

Distinguen entre la confianza que un agente muestra al acordar un contrato con otro, y

el honor en mantener las condiciones acordadas. La confianza es un indicador de la relación entre los elementos acordados y los resultados percibidos del contrato. Por otra parte, el concepto de honor expresa el grado de integridad de la información y en las promesas intercambiadas en la negociación. Al igual que el de confianza, el modelo de honor se basa en la teoría de la información para estimar el comportamiento esperado, las preferencias esperadas o la certeza en el cumplimiento de las promesas [121].

Las medidas de confianza y honor se basan en la entropía negativa de la distribución de probabilidades de los posibles resultados de un contrato, teniendo en cuenta que la confianza evalúa la relación entre los acuerdos firmados y la ejecución esperada del contrato. En este sentido, define la confianza de dos maneras: una a partir de la entropía condicional y otra a partir de la entropía relativa.

La confianza obtenida a partir de la entropía condicional refleja los casos en que no se desea ningún tipo de variación en los acuerdos del contrato. La idea de este planteamiento es que se le otorgue mayor confianza a las observaciones (experiencias) menos dispersas y más cercanas a 1.

En cambio, la entropía condicional hace referencia a los casos en que se esté dispuesto a aceptar (e incluso asignar mayor confianza) los cambios, en los acuerdos del contrato, que representen un incremento de la satisfacción de los requerimientos.

2.2.3. Resumen comparativo

Debido al desarrollo que se ha venido alcanzando en el área y su aplicación a un conjunto de dominios altamente heterogéneo, existe gran diversidad de modelos de confianza y reputación. Cada uno establece su propia representación del conocimiento acerca del sistema y del comportamiento de los otros agentes; de la semántica de sus relaciones, de sus preferencias, requerimientos y tareas; de las medidas para evaluar la calidad y la satisfacción; etc. Varias de estas características se encuentran muy ligadas al dominio de aplicación del modelo y de ahí la gran variedad que presentan.

Sin embargo, todos los modelos tratan de dar respuesta a un conjunto de problemas comunes. Entre los más generales destacan:

- ¿Cómo representar las preferencias del usuario y el comportamiento del resto de los agentes? - El usuario puede establecer los criterios bajo los cuales un agente evalúa a otros. Cada usuario, según sus preferencias y determinado por el dominio de aplicación, establece la forma de evaluar la calidad de las interacciones. Algunos de los modelos de confianza y reputación vistos brindan un mecanismo adaptativo para aprender los comportamientos de los agentes según las preferencias de los usuarios
- ¿Cómo se evalúa una interacción? ¿Cómo se refleja en el modelo? - Al finalizar cada interacción se debe actualizar el modelo de confianza y reputación para interacciones futuras. Por lo general, la actualización del modelo depende de la calidad de la solución obtenida en la interacción y de los niveles de satisfacción que el usuario muestra con dicha solución.

- ¿Cómo saber con quién interactuar? - En muchos casos, éste constituye el objetivo fundamental de un modelo de confianza y reputación. Los modelos utilizan criterios de confianza, de reputación o la combinación adecuada de ambos para determinar en cada caso a quién solicitar soluciones, recomendaciones acerca de otros, etc.
- ¿Cómo tratar la información brindada por otros? - No todos los agentes son igualmente confiables, por tanto la información proveniente de cada agente debe tener un tratamiento diferenciado. Se deben incorporar mecanismos que detecten la información falsa, incompleta, incongruente; así como otros que brinden las formas de combinar la información proveniente de diferentes agentes o desde diferentes fuentes de información.
- ¿Cómo considerar la confianza y la reputación ante varios contextos? ¿Cómo combinar las diferentes dimensiones? - Por lo general, los agentes tienen comportamientos diferentes ante diferentes contextos. Aunque varios modelos no consideran que los valores de confianza y reputación son dependientes del contexto del agente, un gran número de sistemas requieren representar esta situación. De igual forma, aun cuando se considera la confianza y la reputación para un único contexto, el comportamiento de los agentes, en muchos casos, puede variar de acuerdo a diferentes facetas o dimensiones del dominio de aplicación. Por tal razón, los modelos deben ofrecer mecanismos que faciliten la gestión de la confianza y reputación para dominios de aplicación multi-contextuales y multi-dimensionales. Por ejemplo, cuando se consideran diferentes dimensiones y se requiere manejar un único valor que asista las decisiones del agente, se debe definir la forma en que deben ser agregados los valores de confianza y reputación asociados a cada dimensión.

A modo de resumen de las particularidades de los modelos estudiados, se ofrece la tabla 2.2 donde se relacionan las características de cada uno de ellos, de acuerdo a los aspectos tratados en la sección 2.1. Las abreviaturas utilizadas para cada aspecto se muestran en la tabla 2.1.

Las características de cada modelo corresponden con las descripciones que se hace de ellos en la literatura.

Este estudio amplía la clasificación dada por Sabater [111] por cuanto considera nuevos aspectos de los modelos de confianza y reputación (tales como los atributos evaluados, la dependencia de los requerimientos o la dimensionalidad), a la vez que evalúa otros modelos no existentes entonces.

A partir de la información resumida en la tabla 2.2 se pueden obtener algunas conclusiones que, de manera general, describan las principales características de los modelos estudiados:

- La mayoría de los modelos utilizan varias fuentes de información. Sin embargo, la información más utilizada proviene de las experiencias directas (quizás por ser la fuente más confiable y de fácil gestión) y de los criterios provenientes de terceros. En segundo nivel, de acuerdo al número de modelos que la utilizan, se encuentran la

Tabla 2.1: Abreviaturas utilizadas en el resumen comparativo de las características más relevantes de los modelos estudiados

Aspecto	Abreviatura	Significado
Obtención de la información	E	Experiencias directas
	T	Terceros
	R	Relaciones sociales
	P	Prejuicios
	I	Intuición
Atributos considerados para evaluar la confianza y la reputación	AI	Integridad
	AP	Previsibilidad
	AC	Competencias
	AM	Motivación
Dependencia del Contexto	IC	Independiente
	DC	Dependiente
Dependencia de los requerimientos de los usuarios	IR	Independiente
	DR	Dependiente
Dimensiones	UD	Unidimensional
	MD	Multidimensional
Visibilidad	Pv	Privada
	Pb	Publica
Centralización	Cn	Centralizado
	Ds	Descentralizado

Tabla 2.2: Resumen comparativo de las características más relevantes de los modelos

Modelo	Obtención de la información	Atributos	Dependencia del contexto	Dependencia de los requerimientos	Dimensiones	Visibilidad	Centralización
Sporas	T	AI, AP, AC	IC	IR	UD	Pb	Cn
Histos	E, T, R	AP, AC	IC	IR	UD	Pv	Cn
MARS	E, T, R	AI, AC	IC	IR	UD	Pv	Ds
AFRAS	E, T, I	AI, AP, AC	IC	IR	UD	Pv	Ds
Carter et al.	T ⁽¹⁾	AI, AC	IC ⁽²⁾	IR ⁽²⁾	UD	Pb	Cn
TRAVOS	E, T	AI, AP, AC	IC	IR	UD	Pv	Ds
FIRE	E, T ⁽³⁾ , P	AI, AP, AC	IC	IR	MD	Pv	Ds
Wang y Singh	E, R	AI	IC	IR	MD	Pv	Ds
Marsh	E	AC, AM	DC	IR	MD	Pv	Ds
Castelfranchi y Falcone	No explica ⁽⁴⁾	AI, AC, AM	DC	DR	MD	Pv	Ds
Abdul-Rahman y Hailes	E, T	AI, AP, AC	DC	IR	UD	Pv	Ds
Mui et al.	T	AI, AP, AC	DC	IR	MD	Pv	Ds
ReGreT	E, T, R, P	AI, AP, AC	DC	IR	MD	Pv	Ds
Ramchurn et al.	E, T, P	AI, AP, AC	DC	IR	MD	Pv	Ds
TrustMail y FilmTrust	E, T, R, P ⁽⁵⁾	No explica ⁽⁶⁾	IC	DR	MD	Pv	Ds
Wang y Vassileva	E, T, R, I	AC	DC	IR	MD	Pv	Ds
Maximilien y Singh	E, T ⁽⁷⁾	AC	IC	DR	MD	Pv ⁽⁷⁾	Ds
MDT-R	E, T, P	AI, AP, AC	IC	DR	MD	Pv	Ds
Sierra y Debenham	E, I	AI, AP, AC	DC	DR	MD	Pv	Ds

(1) Las opiniones sobre el desempeño de otros proviene de una entidad central .

(2) Los valores de confianza y reputación no se encuentran asociados ni al contexto ni a los requerimientos de los usuarios, sino al rol que desempeñan los agentes.

(3) Incorpora el concepto de reputación certificada.

(4) No explica cómo se obtiene la información de confianza.

(5) Agrupa los agentes según la similitud entre sus preferencias para asignar valores más altos de confianza y reputación a aquellos con preferencias similares.

(6) Considera que de alguna manera se crea la red de relaciones sociales donde cada agente evalúa al otro.

(7) La información sobre terceros queda disponible para todos los agentes.

información proveniente de prejuicios y la extraída a partir de las relaciones sociales que mantienen los agentes. Muy pocos modelos ofrecen mecanismos de intuición para obtener información de confianza y reputación.

La información proveniente de diferentes fuentes es combinada convenientemente para ofrecer un criterio unificado que apoye el proceso de toma de decisiones del agente. De alguna forma, establecen niveles de prioridades al combinar los diferentes valores de confianza. Por lo general, se atribuye mayor importancia a los valores provenientes de análisis del conocimiento sobre las interacciones directas. En algunos casos, los niveles de prioridad de la combinación de los valores están dados por los indicadores de la credibilidad de cada fuente. Otros prefieren utilizar solo la información brindada por terceros cuando no se dispone de conocimiento sobre interacciones directas previas. Sin embargo, puede resultar interesante analizar el comportamiento previo de los agentes ante requerimientos similares, para aproximar los valores de confianza y reputación en aquellos casos en que no se dispone de conocimiento suficiente sobre interacciones directas, ni se es capaz de obtener información a partir de terceros.

Por otro lado, existen aspectos que no son ampliamente abordados por los modelos. Aun cuando algunos modelos utilizan la red de relaciones sociales como una fuente más de información de confianza y reputación, se debe enriquecer la representación semántica de las relaciones y las técnicas para extraer información de ellas. Por ejemplo, muy pocos utilizan criterios de análisis basados en prejuicios a partir de la pertenencia a determinados grupos sociales.

- Un gran número de modelos utilizan varios atributos para evaluar la confianza y la reputación. La mayoría evalúan la integridad, la previsibilidad y competencias del resto de los agentes; muy pocos son los que tienen en cuenta su motivación.

Sin embargo, a pesar de evaluar ampliamente la integridad de los agentes, se requieren mecanismos eficaces para el tratamiento de mentiras o conspiraciones entre los agentes. Resulta muy difícil de medir la credibilidad de un agente, las calumnias que se generan, etc.

- Siguiendo la clasificación planteada en cuanto a la dependencia o no de los valores de confianza y reputación respecto al contexto, los modelos estudiados se diferencian en grupos de similar cantidad de elementos. En otras palabras, se han considerado cantidades similares de modelos de ambos tipos. Sin embargo, destaca el hecho que aun cuando los modelos son dependientes del contexto, muy pocos son los que consideran que los valores de confianza y reputación también están asociados a los requerimientos de los usuarios.
- La mayoría de los modelos consideran que la confianza y la reputación están referidas a múltiples dimensiones, mantienen la información con carácter privado y utilizan un enfoque descentralizado de su gestión.

2.3. Aspectos que requieren esfuerzos de investigación

A pesar del gran número de modelos desarrollados y los grandes avances alcanzados en este campo, aun existen elementos que demandan gran atención en el área de los modelos de confianza y reputación.

Dependencia del contexto y de los requerimientos - por lo general, los modelos numéricos no tienen en cuenta el hecho de que las interacciones entre los agentes se producen en el marco de un contexto ambiental específico y dinámico [104]. El no tener en cuenta las variaciones del entorno de los agentes hace que se maneje información que en determinado momento puede no ser cierta. Un agente puede presentar un comportamiento confiable bajo determinadas circunstancias, mientras que en otros casos puede no serlo. De igual forma sucede cuando varían los requerimientos, un agente puede ser altamente confiable para resolver determinada tarea mientras que para otra puede no serlo, y viceversa.

Aproximación (intuición) - cuando no se conoce el desempeño de otros agentes resulta conveniente tener una estimación de su confianza. Se deben ofrecer mecanismos que aproximen los valores de confianza y reputación cuando aun no se ha interactuado con el agente. Soto *et al.* [123] también consideran este elemento como un componente en la obtención de los valores de reputación. Otros, teniendo en cuenta el prejuicio como fuente de confianza, pueden otorgar un criterio inicial en dependencia del grupo social al que pertenece el agente [114, 44].

Confiabilidad - las medidas de confianza y reputación siempre están viciadas por errores propios de las interacciones entre los agentes, por sus buenas o malas intenciones, por la representación interna del modelo, por sus mecanismos de inferencia, etc. Los modelos de confianza deben brindar mecanismos, tanto a nivel individual como de sistema, que definan estrategias frente a la detección de mentiras y de complicidades, así como de redundancias y/o incongruencias en la información que se intercambia. Las estrategias para detectar estas situaciones deben ser capaces de modificar los valores de confianza en los agentes con los que se interactúa. A pesar, de que varios modelos han ofrecido algunas soluciones [146, 111, 28, 101, 126, 121, 70, 71, 20], este es aspecto que requiere gran atención.

Actualización del conocimiento - la mayoría de los modelos no muestran una forma clara de evaluar las interacciones entre los agentes (por ejemplo, el cumplimiento de los compromisos, la satisfacción de las expectativas del agente en cuanto a calidad, tiempo de respuesta, etc.). La manera de evaluar las interacciones entre los agentes depende en gran medida del dominio de aplicación del modelo y de las representaciones internas del conocimiento de cada agente que interactúa. En un entorno abierto, interactúan agentes de varias naturaleza, con diferentes representaciones del entorno y otros agentes, siguiendo diferentes modelos y patrones de comportamiento y utilizando diferentes criterios para evaluar sus expectativas. En este sentido ReGreT [114] establece un modelo ontológico común para evaluar la satisfacción de los agentes. Sin embargo, la ontología compartida es dependiente del dominio de aplicación en cuestión. Pudiera, además, pensarse en una

ontología común para todos los modelos, capaz de evaluar las interacciones utilizando, por una parte, atributos inherentes a los conceptos de confianza y reputación (número de errores en el tiempo, número de respuestas satisfechas en el tiempo, confiabilidad, etc.) y, por otra, los atributos propios del dominio de aplicación (por ejemplo, para evaluar un buen vendedor, buena relación calidad/precio, fecha de entrega, etc.).

Carácter social - las relaciones sociales que se presentan entre los agentes son una reflexión de las relaciones que se presentan entre los seres humanos. Sin embargo, los modelos utilizan solo una pequeña parte de esta información para calcular o mejorar el cálculo de los valores de confianza y reputación [111, 104, 11, 55, 54, 10, 145, 96]. Por lo general, solo se gestiona información relativa a la acción que se ejecutó en una interacción pasada, o el valor de confianza asociado, etc. Sin embargo, existe otro tipo de información social disponible en la red de relaciones entre los agentes que resulta muy útil (por ejemplo, la semántica de las relaciones). La definición de los diferentes tipos de relaciones que se presentan entre los agentes (por ejemplo, miembros del mismo grupo social u organización, contrapartes en la misma negociación, colaboradores, amigos, competidores por los mismos recursos, etc.) podrían ayudar a realizar una mejor agregación y evaluación de los valores de confianza y reputación.

Evaluación y comparación de los modelos - Por lo general, cada modelo de confianza y reputación presenta su propia metodología y escenario de prueba para la evaluación de su comportamiento. El escenario, muchas veces, está diseñado para hacer destacar las aspectos trascendentales del modelo asociado [111]. Esto se debe fundamentalmente a la gran variedad de características (tales como la representación del conocimiento, del comportamiento de otros agentes, de las preferencias de los usuarios, de las tareas, los métodos para evaluar la calidad de la respuesta, la satisfacción del usuario, entre otras) y a la carencia de un escenario de prueba que se adapte a todos los requerimientos de cada modelo y posibilite su evaluación bajo las mismas condiciones experimentales. En este sentido, se han desarrollado algunas propuestas entre las que destaca el escenario de prueba ART [50] y TOAST [67]. Sin embargo, consideramos que ninguno de ellos constituye una propuesta definitiva por cuanto un escenario de prueba de este tipo debe estar constantemente en evolución, incorporando los nuevos conceptos e ideas, surgidos con el propio desarrollo de los modelos de confianza y reputación.

2.4. Conclusiones

En este capítulo se han estudiado los aspectos fundamentales relacionados con el desarrollo alcanzado en el campo de los modelos de confianza y reputación. Se han destacado los elementos generales importantes de un modelo, detallando cómo se abordan en cada uno de los modelos estudiados. Al final se ha presentado una comparativa de los modelos y se han identificado las líneas de trabajo que aun requieren esfuerzos de investigación. De esta forma, en el resto de la tesis, se pretende definir las características de un nuevo modelo de confianza y reputación, de forma que aproveche las ventajas de los modelos existentes y ofrezca soluciones a algunas de las situaciones identificadas en este

capítulo.

Antes del planteamiento y evaluación del modelo desarrollado en esta tesis, los dos capítulos siguientes profundizan el estudio de algunos temas comentados en la sección 2.3. En el capítulo 3 se estudian los principales enfoques, empleados por otros modelos, para la monitorización, evaluación y validación de los modelos de confianza y reputación. Esto servirá para definir los mecanismos y medidas de evaluación del modelo propio, presentado en el capítulo 5. En tanto, el capítulo 4 se dedica al estudio de las principales representaciones semánticas disponibles para la definición de algunos elementos dependientes del dominio de aplicación del modelo a desarrollar.

Capítulo 3

Evaluación de modelos de confianza

La monitorización, evaluación y validación de los modelos de confianza y reputación es un campo que requiere esfuerzos de investigación. El desarrollo alcanzado hasta el momento en este sentido, ha estado determinado por la definición de métricas y experimentos, por cada uno de los autores de los modelos. De manera general, los planteamientos han estado encaminados a corroborar el correcto desenvolvimiento de un determinado modelo y destacar las bondades para la situación en la cual el modelo de confianza y reputación ha sido diseñado [111, 28, 143, 150, 94, 72, 75, 116, 84, 147, 32, 126, 67]. En todos estos trabajos, destaca la ausencia de un escenario de referencia común para la monitorización, evaluación y validación de los modelos, que se adapte a las características generales de todos y permita su evaluación y comparación bajo las mismas condiciones experimentales. El desarrollo de un escenario con estas características constituiría un avance considerable en el campo de trabajo.

Afortunadamente, los desarrollos más recientes, como FIRE [72], TOAST [67] y ART [50], tratan de ofrecer un marco de experimentación común para la prueba y comparación de los diferentes modelos bajo las mismas condiciones. Sin embargo, aun son tres escenarios diferentes, y en desarrollo. En los apartados 3.2.3.3, 3.2.4 y 3.2.5 se comentan las características de cada uno de los escenarios de prueba suministrados por FIRE, TOAST y ART, respectivamente.

El presente capítulo se divide en dos bloques fundamentales. El primer bloque (determinado por el contenido de la sección 3.1), donde se presenta un conjunto de métricas para la evaluación de los modelos, pretende estudiar las diferentes medidas que han sido utilizadas de manera que se puedan tomar algunas de ellas para analizar el comportamiento del modelo que se propone en esta tesis.

Luego, en el segundo bloque del capítulo (dado por la sección 3.2), el estudio de los escenarios de prueba pretende servir para seleccionar uno de ellos en el cual probar el funcionamiento del modelo planteado en esta tesis. Se espera que, utilizando un escenario de prueba adecuado, se pueda estudiar el comportamiento del modelo ante diferentes variantes de su configuración y diferentes condiciones experimentales. Por otra parte, un escenario permite la comparación del desempeño del modelo planteado frente a otros modelos ajenos. En resumen, la utilización de un escenario de prueba adecuado puede contribuir

al refinamiento del modelo en dos sentidos: 1) análisis de las diferentes alternativas de configuración del modelo, y 2) comparación del desempeño del modelo con respecto a otros.

A continuación, el capítulo continúa con la sección 3.1 donde se ofrece una clasificación de las principales métricas que han sido utilizadas para describir el funcionamiento de los modelos de confianza. Luego, la sección 3.2 comenta las principales características de los escenarios de prueba que han sido desarrollados con el fin de evaluar y comparar los modelos de confianza. El apartado 3.3 concluye los aspectos tratados más importantes.

3.1. Métricas para la evaluación y comparación de los modelos

El estudio del comportamiento de los modelos de confianza y reputación se lleva a cabo mediante el análisis de indicadores del comportamiento de los agentes y del sistema en que se implementan, fundamentalmente. Por lo general, se definen magnitudes medibles en los experimentos que permiten establecer criterios de comparación entre las diferentes configuraciones que se tratan.

Por la semejanza de su naturaleza y la forma en que se obtienen, estas métricas pueden agruparse bajo dos categorías fundamentales: las que describen el comportamiento del agente (relacionadas con indicadores individuales y locales de su desempeño) y, por otro lado, las que reflejan el comportamiento de la sociedad de agentes (relacionadas con indicadores sociales y el efecto que las interacciones entre los agentes provocan sobre el entorno).

Los modelos de confianza y reputación, así como los escenarios de prueba que se definen para su evaluación y comparación, ofrecen herramientas que facilitan la recolección, procesamiento y presentación de las medidas del comportamiento individual o social. Por lo general, las métricas presentadas se basan en el conteo y/o agregación de información almacenada durante varias interacciones. Sin embargo, también puede estar disponible otro tipo de información, que puede ser analizada más detalladamente o por técnicas diferentes al conteo o consolidación de la información generada (por ejemplo, la búsqueda de patrones de comportamiento).

A continuación, se ofrece un estudio de las métricas que han sido utilizadas por otros modelos. Éstas se presentan agrupadas, utilizando un criterio propio de clasificación, según su naturaleza y los atributos que evalúan.

3.1.1. Métricas individuales

El análisis del comportamiento de los agentes de manera individual puede resultar útil para evaluar cuan beneficioso resulta el conocimiento obtenido por un agente, a partir de la utilización de un modelo de confianza y reputación. Para ello se definen métricas cuyos valores, o variaciones de los mismos, representan el desempeño de los agentes.

Las métricas usadas como indicadores del comportamiento de los agentes de manera individual también pueden ser utilizadas por los propios agentes como un elemento más a tener en cuenta en los procesos de toma de decisiones.

A continuación se comentan las principales medidas individuales que se reportan en la literatura, clasificándolas según su propósito.

3.1.1.1. Evolución de la confianza y la reputación

De manera general, los modelos de confianza y reputación estudiados basa su análisis del comportamiento e idoneidad del modelo en el estudio de la evolución de los valores de confianza y reputación en el transcurso del tiempo [150, 117, 98, 143, 75, 147, 113, 27, 139].

Estas medidas son útiles para analizar la convergencia del modelo y realizar los ajustes necesarios para obtenerla de manera temprana. La rapidez de convergencia del modelo de confianza y reputación garantiza su adaptabilidad a cambios del entorno o del comportamiento de otros agentes.

También es muy útil el análisis de la evolución de la confianza y la reputación de agentes cuyo comportamiento se conoce de antemano. Por ejemplo, el modelo representa eficientemente el comportamiento de los agentes si los valores de confianza y reputación aumentan rápidamente para agentes cuyo desempeño es bueno; y, en cambio, para agentes malos, estos valores disminuyen.

3.1.1.2. Métricas relacionadas con el nivel de satisfacción

Varios modelos de confianza y reputación utilizados en sistemas recomendadores o de comercio electrónico, donde el principal objetivo es encontrar y evaluar previamente el conjunto de servicios que satisfacen el conjunto de los requerimientos, utilizan medidas relacionadas con la satisfacción de las soluciones que son ofrecidas por el modelo [50, 21, 136].

En este sentido, ART y el modelo de Wang y Vassileva [136] evalúan el grado de satisfacción con la solución obtenida, estudiando su evolución y posible convergencia a lo largo del tiempo. A continuación se comentan cada uno de ellos.

ART ofrece medidas acerca del error que se produce en la tasación de una pintura, a partir de la diferencia entre el valor real de la pintura y la opinión de tasación ofrecida por el agente. También calcula la desviación estándar del error de tasación que es utilizado como idea de la precisión de las opiniones del tasador. (El apartado 3.2.5 ofrece los detalles de este escenario de prueba.)

Otros, como Wang y Vassileva [134], para evaluar el nivel de satisfacción de un usuario con su modelo en un sistema de recomendación, definen métricas de satisfacción a partir del conteo de las recomendaciones exitosas. (En el dominio de aplicación tratado, una recomendación es exitosa si las necesidades de usuario son satisfechas utilizando el servidor de ficheros recomendado por el modelo.) Las principales medidas están relacionadas con:

- Porcentaje de recomendaciones exitosas (número de recomendaciones exitosas dividido por el número de las recomendaciones positivas),

- Porcentaje de interacciones exitosas basadas en recomendaciones (número de interacciones exitosas basadas en recomendaciones dividido por el número total de interacciones exitosas), y
- Porcentaje de interacciones exitosas en todas las interacciones (número de interacciones exitosas dividido por número total de interacciones).

TRAVOS [126] emplea el error de estimación de la confianza como una medida de la eficiencia del modelo. En este caso, el nivel de satisfacción queda expresado implícitamente a partir de la diferencia entre el valor real y la estimación ofrecida por el modelo.

De manera general, en este tipo de modelos, no solo puede ser interesante analizar los beneficios de la respuesta obtenida (opinión recibida, valoración estimada, etc.), sino también analizar el coste de oportunidad de no seleccionar la óptima. Esta métrica puede determinarse a partir de la diferencia entre la satisfacción con la respuesta (opinión, valoración, etc.) seleccionada y la satisfacción que reportaría la óptima.

3.1.1.3. Métricas relacionadas con la utilidad

Existe una gran cantidad de modelos aplicados a entornos competitivos donde el comportamiento de los agentes se mide en términos de funciones de utilidad. En estos entornos los agentes tienen una cantidad de recursos limitados para acceder a los servicios que necesitan. Consumen unidades de recursos para acceder a los servicios y reciben otras por los servicios brindados. Las funciones de utilidad obtienen el número de unidades de recursos que posee cada agente [50, 75, 72, 126, 67, 113, 27].

Por su similitud con las aplicaciones de comercio electrónico, la mayoría de los modelos expresan las unidades de recursos en términos de dinero. Así, un agente gasta dinero en comprar los recursos que desea y recibe ganancias por los servicios que vende. De esta manera, la utilidad de un agente como reflejo del saldo de su cuenta bancaria constituye una medida de su desempeño.

Castelfranchi *et al.* [32] definen la utilidad como la fortaleza de un agente a partir de la recolección de alimentos. El hecho de ingerir alimentos y vencer a otros agentes supone una ganancia en la fortaleza del agente, mientras que el ser derrotado supone pérdidas.

Por su parte, Gujral *et al.* [65] consideran el beneficio estimado a partir de la retribución que se tiene a partir del cumplimiento de los requerimientos de tiempo y calidad, el coste en que se incurre si el respondedor no está disponible, y el coste que tiene que pagar por la respuesta en sí.

3.1.1.4. Métricas relacionadas con las interacciones con otros agentes

Existen otras medidas del comportamiento de los agentes que hacen posible el análisis de la eficiencia de su comportamiento a partir de indicadores de las interacciones con otros agentes. Un ejemplo de este tipo de métricas es el conjunto de métricas ofrecidas por ART al considerar la cantidad y tipo de mensajes intercambiados en sus interacciones.

3.1. MÉTRICAS PARA LA EVALUACIÓN Y COMPARACIÓN DE LOS MODELOS

Con medidas de este tipo, el modelo de confianza y reputación puede ser analizado desde el punto de vista de la eficiencia del modelo al decidir a quiénes solicitar información de reputación, a quiénes comprar opiniones, a quiénes aceptar la compra de dichas informaciones, entre otras.

En este sentido, también puede analizarse el tanto por ciento de vecinos que son encuestados con cada solicitud de información. En algunos casos, puede resultar un indicador del coste asociado a las interacciones innecesarias, dado que existe un gran volumen de solicitudes que luego son desestimadas o poseen información redundante.

Los estudios realizados por Jøsang *et al.* [75] también incluyen el conteo del número de interacciones de los agentes compradores con los vendedores, como medida del funcionamiento del modelo. Esta medida es utilizada para contrastar la utilidad del modelo de reputación para la selección adecuada de los vendedores. Se apoya en la idea básica de que debe existir un número mayor de interacciones con un vendedor de elevada reputación que con uno de menor reputación.

3.1.1.5. Métricas relacionadas con la calidad del servicio

En aquellos entornos donde los agentes son intermediarios de servicios, resultan útiles las métricas que definen el comportamiento de los agentes a partir de indicadores de la calidad del servicio que representan.

Griffiths, para decidir a quién delegar una determinada tarea, no solo utiliza elementos de confianza, sino también algunos aspectos que definen la calidad del servicio: el ratio de fallos, el costo de ejecución y la calidad de ejecución [57].

3.1.2. Métricas sociales

Los escenarios, por lo general, están dedicados al estudio del comportamiento de un único agente. Esto quizás está motivado por el carácter egoísta que suelen tener las relaciones entre los agentes. Sin embargo, al igual que los agentes son egoístas, en muchos casos tienen intereses que les hace cooperar unos con otros, formando agrupaciones según sus preferencias y funcionalidades. En estos casos, pueden existir situaciones en que los agentes no alcancen su máximo desempeño individual en aras de lograr un mayor beneficio generado por las relaciones que se tiene con el resto de agentes de su grupo, o del sistema. De ahí la importancia de considerar medidas de carácter grupal que representen estas situaciones.

Al analizar el comportamiento global de los agentes en el sistema (o de un grupo específico de ellos) varios autores utilizan la agregación de los valores de las métricas definidas individualmente para cada uno [126, 136, 145]. Así, se tienen en cuenta medidas tales como:

- los valores medios de confianza y reputación,
- el valor medio del nivel de satisfacción, y

- el valor medio del error de estimación del comportamiento de los proveedores.

Las métricas de este tipo han sido menos utilizadas por los modelos y por los escenarios definidos para evaluarlos.

Por ejemplo, Wang [136] utiliza los conceptos de confianza y reputación en una comunidad de agentes. Para ello supone que los agentes se agrupan formando comunidades, cada una de las cuales puede ser caracterizada por los valores medios de confianza y reputación de los agentes que la componen. También analiza la media del nivel de satisfacción de los agentes a partir del promedio del conteo de las interacciones exitosas (una interacción exitosa es aquella en que se encuentra la solución adecuada).

De igual forma, Yu y Singh [145] utilizan el promedio de los valores de reputación de un grupo de agentes para analizar el efecto del egoísmo y la falta de honestidad en el desempeño de la comunidad de agentes.

Por otra parte, en [126] se utiliza la media del error de estimación de la confianza para comparar el comportamiento de TRAVOS [101] y de BRS [76].

3.1.3. Información recolectada como fuente de otras métricas

Los escenarios de prueba, o los propios modelos, no solo ofrecen indicadores cuyos valores describen el comportamiento de los agentes y que permiten la comparación de unos con otros. También generan información de otra índole (típicamente, elementos que caracterizan el sistema multi-agente como sistema distribuido) que, individualmente, puede resultar útil para que el desarrollador del modelo proponga sus propias métricas y realice otros tipos de análisis más personalizados.

Un ejemplo de este tipo de información es la bitácora de mensajes intercambiados entre los agentes que ofrece ART [50]. A partir de esta bitácora se puede concluir información relacionada con el comportamiento de los agentes (por ejemplo, pueden identificarse determinados patrones de comportamiento).

3.2. Escenarios de prueba

Un escenario de prueba de los modelos de confianza y reputación permite la evaluación y comparación de varios modelos a partir de la definición del entorno, las relaciones entre los agentes, los actos comunicativos, la estructura del contenido de la comunicación, entre otros. Se pretende con ello ofrecer las herramientas necesarias para la obtención y análisis, de manera sencilla y transparente, de las métricas sobre el desempeño de los agentes cuyo comportamiento está definido por el modelo.

Existen pocos escenarios de prueba y validación comunes a los modelos de confianza y reputación, debido fundamentalmente a la gran variedad de problemas y dominios de aplicación tratados. Por lo general, la mayoría de los modelos han sido probados y comparados con otros tomando como punto de partida un conjunto de experimentos definido anteriormente para una situación similar o bien definiendo los propios experimentos. En estos casos, la configuración de las pruebas se corresponde

con la problemática para la cual el modelo de confianza ha sido diseñado, de modo que los resultados destacan las características para las cuales el modelo aporta una solución eficiente. Sin embargo, muchas veces no se analiza el comportamiento integral del sistema en otras configuraciones más generales y reales.

A continuación se tratan algunas características relevantes de algunos escenarios que han sido desarrollados con el objetivo de seleccionar (de ser posible) el más adecuado para la experimentación con el modelo desarrollado en esta tesis.

3.2.1. Escenarios basados en el dilema del prisionero

Existe un grupo de escenarios de prueba basados en el dilema del prisionero, un problema clásico de la teoría de juegos [7]. Se basa en la situación en que dos prisioneros deben declarar de manera independiente sobre su culpabilidad. Cada uno tiene dos opciones: declararse inocente (mantenerse en silencio) o culpable (confesar). Si uno de ellos colabora con la policía y confiesa, queda en libertad mientras que el otro recibe la pena máxima. Si ambos confiesan la condena se rebaja a una pena intermedia.

El dilema de los prisioneros estriba en qué postura tomar: confesar o mantenerse en silencio, teniendo en cuenta los beneficios a recibir en función de la acción de su compañero. A partir de las dos alternativas posible para sus acciones, se define una matriz de utilidades donde por las filas y columnas aparecen las opciones de cada jugador. Cada jugador deberá decidir las acciones a tomar, en función de la matriz de utilidades.

La versión iterada de este juego basado en el dilema del prisionero es el fundamento de varios escenarios de prueba que consideran los movimientos (selección de alternativas) sobre una matriz de utilidades. A continuación se comentan algunos de los escenarios más representativos que han utilizado este problema como base.

3.2.1.1. Escenario PlayGround

Este escenario fue propuesto por Marsh [84] para probar su modelo. En él, el modelo determina los movimientos espaciales de los agentes sobre una tabla de 10 filas por 19 columnas donde cada celda puede ser ocupada por un único agente en cada instante de tiempo. Cada movimiento de un agentes es libre y aleatorio, aunque cuando el modelo de confianza se va ajustando, los agentes aumentan la probabilidad de moverse hacia donde se encuentran los agentes confiables. Cada interacción entre dos agentes está determinada por el dilema del prisionero que se presenta entre dos agentes cuando uno de ellos trata de ubicarse en una celda que está ocupada por el otro.

Cada matriz de utilidad, correspondiente al dilema del prisionero, define una *situación*. Todos los agentes conocen de antemano las posibles situaciones, definidas previamente. Así, cada interacción entre dos agentes está determinada por la situación que se presente entre ellos, seleccionada aleatoriamente e informada a los agentes participantes.

En cada interacción, cada agente debe decidir qué estrategia seguir: cooperar o desertar. La estrategia es adoptada utilizando su propio criterio de confianza sobre el otro para la situación en cuestión. Una vez tomada la decisión de cada agente, se actualizan los valores

de confianza de cada agente para la situación en cuestión. Las variaciones (aumento o disminución) de la confianza de cada agente dependen de las estrategias adoptadas por los dos agentes. De manera general, si el agente oponente coopera, la confianza en él aumenta, en otro caso, disminuye. Los incrementos o disminuciones que un agente realiza sobre la confianza de otro dependen de la estrategia (cooperar o desertar) que el propio agente haya tomado. Por ejemplo, el aumento en la confianza en el oponente es mayor cuando ambos agentes cooperan que en los casos en que el propio agente deserta y el oponente decide cooperar.

3.2.1.2. Escenario con una fase de negociación previa

Para evaluar el modelo desarrollado, Schillo *et al.* [116] utilizan una extensión del juego del dilema del prisionero derivado a partir de utilizar una fase previa de selección del agente oponente. La principal variante de este escenario es incluir una fase para determinar los pares de jugadores a partir de un proceso de negociación y la declaración de sus intenciones (pudiendo mentir al respecto).

En la fase de selección de pares, el proceso de negociación se desarrolla según un protocolo similar a un *contract-net* [48], que termina cuando cada jugador ha tenido la oportunidad de encontrar su compañero de juego. Luego de esta fase de selección, tiene lugar el juego basado en el dilema del prisionero, teniendo en cuenta las declaraciones de intenciones hechas en la negociación. Al finalizar, el agente puede actualizar el modelo para la selección de pares en iteraciones siguientes.

El escenario está regido por un enfoque monetario donde cada agente dispone de una cantidad de puntos e intenta incrementarlos en su interacción con el resto. Al inicio de cada interacción invierte un número de ellos en forma de apuesta. Luego, si la interacción implica un movimiento acertado, recibe bonificaciones por la inversión realizada. Si el movimiento es erróneo, los puntos invertidos se pierden. Los agentes que agoten todos sus puntos se quedan fuera del juego. Así, los agentes que más acertados sean al apostar sus puntos recibirán mayores beneficios, los más desacertados tendrán mayores pérdidas.

3.2.1.3. Escenario donde las decisiones son guiadas por la reputación

Mui *et al.* [94] definen este escenario de prueba para la comparación del comportamiento de agentes que tienen en cuenta diferentes variantes de reputación descritas en su trabajo. Este escenario utiliza una variante del juego del dilema del prisionero donde las posibles alternativas de cada agente pueden ser:

- Cooperar (C) - siempre cooperar,
- Callar (D) - siempre callar,
- Tit-for-tat (TFT) - inicialmente cooperar, y luego hacer lo que hizo el otro agente en la iteración anterior, y

- Reputation tit-for-tat (RTFT) - inicialmente cooperar dependiendo de la reputación en el otro agente y luego hacer cualquier cosa que el otro haya hecho en la última iteración.

Los participantes en una iteración del juego son seleccionados aleatoriamente de la población de agentes. Después de que se selecciona el primer participante, se selecciona el otro aleatoriamente. Al finalizar un cierto número de iteraciones (llamado generación), cada agente puede engendrar hijos de acuerdo a la utilidad alcanzada. El número de agentes en cada generación es fijo, por lo que el aumento del número de agentes de un determinado tipo, conlleva a la disminución de los agentes de otro tipo menos exitoso. Es, por tanto, un enfoque bioinspirado en el sentido en que recoge un esquema basado en el predominio de aquellos agentes más adaptados al entorno.

3.2.2. Recolector de alimentos

El escenario de prueba planteado por Castelfranchi *et al.* [32] se representa mediante una malla sobre un campo, donde en las casillas se encuentran ubicados aleatoriamente varios alimentos y los agentes funcionan como recolectores de ellos. Los agentes se mueven libremente por la red en busca de alimentos. Al moverse a una celda que contiene una pieza de comida, el agente se detiene algunos instantes de tiempo para ingerirla. En el tiempo que el agente está comiendo puede ser agredido por otros. El vencedor en el ataque es quien toma el alimento que se encontraba en la celda. Cada agente tiene una fortaleza determinada que disminuye en función de los movimientos por la red o los ataques recibidos, y aumenta según las piezas de alimento tomadas. El objetivo de cada agente en este modelo es aumentar sus fuerzas tanto como sea posible.

El escenario de prueba persigue la comparación de las estrategias de interacción (ataque) entre diferentes tipos de agentes:

- Agente ciego (B) - la agresión a otros solo está determinada por la utilidad que represente, no por la fortaleza suya ni de quien se encuentra comiendo en la celda. Los agentes atacan cuando el costo de la alternativa es alto.
- Agente estratégico (S) - agrede a otros solo si conoce que resultará vencedor, es decir, si la fortaleza de quien come no es mayor que la propia.
- Agente normado (N) - Este tipo de agente considera el concepto de norma (cuyo efecto se desea analizar mediante dicho escenario [32]). En este caso, cada explorador de comida sigue una norma (o estrategia) que le asigna previamente la región de la red y la comida que se encuentra en las celdas correspondientes. Así, otro agente no puede atacarlo cuando esté comiendo en su territorio.

De esta forma, las habilidades de un agente normado pueden estar definidas (extendidas) a partir de los conceptos de confianza y reputación. La utilización de un modelo de confianza y reputación para la definición de las estrategias de los agentes normados posibilita utilizar el conocimiento previo del agente y el que es intercambiado con otros.

3.2.3. Escenarios basados en el paradigma Consumidor - Proveedor

Muchos modelos de confianza y reputación han sido desarrollados con el propósito de seleccionar el proveedor de servicio idóneo de acuerdo al conocimiento del consumidor. Para estudiar el funcionamiento de estos modelos se han utilizado escenarios de prueba que siguen dicho enfoque. A continuación se comentan algunas características de los escenarios más relevantes.

3.2.3.1. Escenario BRS (*Beta Reputation System*)

Jøsang *et al.* [75] definen un escenario de mercado con 3 vendedores y 10 compradores que comercian con las mismas mercancías. La actividad del mercado está estructurada en sesiones de 10 rondas cada una. En cada ronda un comprador trata de comprar exactamente una mercancía. Una transacción solo tiene lugar cuando un comprador encuentra un vendedor que satisface sus criterios de compra. Al recibir la mercancía, el comprador envía un valor de retroalimentación positivo, si no la recibe envía un valor de retroalimentación negativo. Los valores de reputación de cada vendedor son actualizados por una entidad central que utiliza los valores de retroalimentación reportados.

Al inicio de cada sesión cada vendedor define el precio de venta de cada mercancía (P_{rc}) y el nivel de honestidad (H_{st}). Los productos serán enviados al comprador con una probabilidad H_{st} y no enviados con una probabilidad $(1 - H_{st})$.

Por cada transacción, el vendedor recibe una ganancia g . En los casos en que la mercancía es enviada la ganancia se determina de la siguiente manera:

$$g = P_{rc} - C_{st_{prod}}$$

donde el costo de producción $C_{st_{prod}}$ es fijo. En caso de entrega, el agente vendedor siempre obtiene ganancias por cuanto el costo de producción es fijo y conocido antes de fijar el precio de venta.

En los casos en que no se realice la entrega del producto la ganancia del agente estaría dada por el precio de venta:

$$g = P_{rc}$$

El valor intrínseco del producto para cada comprador es denotado por V_{al} . Utilizando el valor de reputación R_{ep} , que se le asigna al vendedor, los compradores deben hacer frente al riesgo de que los vendedores no entreguen el producto, considerando que la ganancia esperada g_{exp} está dada por la ecuación:

$$g_{exp} = (V_{al} - P_{rc}) \cdot R_{ep} - P_{rc} \cdot (1 - R_{ep}) = V_{al} \cdot R_{ep} - P_{rc}$$

Al inicio de cada sesión todos los compradores definen su aversión al riesgo A_{vs} ($A_{vs} \in [0, 1]$), donde $A_{vs} = 0$ significa que está dispuesto a correr cualquier riesgo, mientras que $A_{vs} = 1$ significa que no quiere correr ningún riesgo. De esta forma, el comprador

seleccionará al vendedor que le reporte mayor ganancia esperada g_{exp} y cuyo valor de reputación R_{ep} sobrepase el nivel de aversión al riesgo A_{vs} .

En caso de que el vendedor realice la entrega exitosamente, el comprador tendrá una ganancia $g = V_{al} - P_{rc}$. Si, por el contrario, la entrega no se realiza la pérdida estaría dada por $g = -P_{rc}$.

Cada comprador debe ajustar su modelo de reputación para realizar la elección oportuna en las siguientes rondas, teniendo en cuenta el comportamiento de los vendedores.

3.2.3.2. Escenario simulador de un mercado

Este escenario fue definido para validar a TRAVOS [126] a partir de su comparación con BRS [76]. Es un escenario sencillo que se basa en la simulación de un mercado donde los agentes pueden ser proveedores, consumidores o suministradores de información de reputación. Al inicio de cada experimento, se fija el comportamiento de cada proveedor y suministrador de información mediante el parámetro B_i ($B_i \in [0, 1]$). El valor de B_i indica la tendencia del agente i a cumplir sus compromisos.

Los experimentos consisten en una serie de episodios en los cuales un consumidor trata de valorar la confianza en todos los proveedores. De esta forma, se calcula la media del error de estimación del comportamiento de los proveedores, a partir de la media de las diferencias absolutas entre su comportamiento (B_i) y el valor de confianza ofrecido por el modelo (T_i):

$$\frac{1}{N} \sum_{i=1}^N abs(T_i - B_i)$$

La evolución de esta medida es un indicador de la idoneidad del modelo de confianza para estimar el comportamiento de los agentes.

3.2.3.3. Escenario basado en servicios

El escenario de prueba utilizado para la evaluación de FIRE [72] es un sistema multi-agente compuesto por agentes proveedores de servicios y agentes consumidores que utilizan dichos servicios. Asume que todos los agente ofrecen el mismo servicio; sin embargo, el desempeño de cada uno (por ejemplo, la calidad del servicio) puede ser diferente. Los agentes se ubican aleatoriamente sobre una esfera y son estacionarios durante su ciclo de vida. Cada agente tiene su propio radio de acción, que representa la capacidad del agente para interactuar con el conjunto de sus vecinos (formado por todos los agentes ubicados dentro de su alcance).

Las simulaciones se realizan en iteraciones (o rondas). En cada ronda, si un agente consumidor necesita utilizar un servicio puede localizar los agentes más cercanos, seleccionando uno de ellos. La selección del servicio adecuado se realiza a partir del conocimiento ofrecido por el modelo de confianza. El agente consumidor recibe una utilidad de la interacción, llamada UG , ($UG \in [-10, 10]$), en dependencia del desempeño del proveedor. El desempeño de cada proveedor está dado por uno de los cinco niveles prefijados

en el modelo, cada uno con valores de utilidades diferentes. Un proveedor puede servir a más de un consumidor en cada instante de tiempo.

Después de cada interacción, el agente consumidor evalúa al proveedor de acuerdo al nivel de satisfacción que recibe con su servicio. Almacena esta evaluación para análisis de decisiones futuras e informa al proveedor de su evaluación. El escenario supone que todos los agentes son honestos en el intercambio de información.

Con el fin de simular diferentes niveles de dinamismo, algunos de los parámetros del modelo cambian de una interacción a otra:

- La población de agentes: un número aleatorio de agentes abandonan el escenario y otros nuevos son ubicados.
- La localización de los agentes: los agentes cambian de localización, modificándose así el conjunto de sus vecinos. Se rompen algunas relaciones y se crean otras.
- El comportamiento de los proveedores: los agentes cambian su desempeño (pueden hacerse peores o mejores). Algunas veces estos cambios pueden estar determinados por una conducta egoísta del proveedor, al aumentar su reputación decide mejorar su desempeño y recibir mayores ganancias por el servicio brindado.

3.2.3.4. Experimentos de Sporas, ReGreT, AFRAS

Sporas [147] ha sido probado por un conjunto de experimentos que también han sido utilizados por ReGreT [111] y AFRAS[28]. Sabater y Sierra [113] los utilizan para comparar ReGreT, Sporas y los mecanismos de Amazon Auctions. Por su parte, Carbó *et al.* [27] proponen una versión extendida de dichos experimentos y los utilizan para comparar AFRAS [28] con Sporas, el modelo de Yu y Singh [142, 143] y los mecanismos de reputación utilizados por eBay y Bizrate ¹.

Aunque este conjunto de experimentos dista de ser considerado un escenario de prueba, ha resultado útil para la comparación de varios modelos bajo las mismas condiciones [27].

El conjunto de experimentos iniciales propuestos por Zacharia *et al.* [147] se centran en el estudio de la velocidad de convergencia del modelo y el efecto del desempeño anterior, analizando como los modelos reaccionan a diferentes escenarios. Luego, Carbó *et al.* [27] incluyen el estudio acerca de cómo el uso de la cooperación entre agentes mejora la convergencia de la reputación y el impacto de las coaliciones entre vendedores y compradores.

3.2.4. TOAST: *Trust Organisational Agent System Testbed*

TOAST [67] ha sido desarrollado con el fin de evaluar los modelos de confianza a partir de la influencia del modelo en la evolución de la utilidad de un agente o de una comunidad de agentes. Este escenario de prueba simula una organización virtual donde los

¹BizRate. <http://www.bizrate.com/>

agentes interactúan unos con otros para obtener sus metas. El escenario está basado en las siguientes simplificaciones:

- Solo se considera el problema de seleccionar los agentes contrapartes en cada interacción a partir de un conjunto de posibles candidatos.
- Todas las interacciones son binarias, tienen lugar entre dos agentes.
- Los agentes siempre están dispuestos a interactuar con quien se lo solicita. No se considera el problema en que el agente seleccionado puede decidir o no si la interacción es útil para él.
- Las interacciones son simuladas, los agentes no evalúan el comportamiento de sus contrapartes. Una vez finalizada la interacción, los agentes reciben las evaluaciones desde el escenario de prueba. Para evaluar el comportamiento de los agentes, el sistema utiliza el concepto de habilidad (en inglés, *Capability*). La habilidad de un agente A determina su idoneidad para desempeñar un rol R en una interacción de tipo I . La habilidad es representada mediante una distribución normal de media y varianzas constantes y es asignada a cada tupla $\langle A, R, I \rangle$. Esto supone que: 1) los agentes muestren cierta estabilidad en su comportamiento. (Sin embargo, aunque este supuesto es la base de los modelos de confianza y reputación, en muchos casos el comportamiento de los agentes varía en el tiempo), y 2) los agentes no mienten, siempre ofrecen su mejor desempeño de acuerdo a sus habilidades.

El escenario permite la definición de organizaciones virtuales compuestas, cada una, por los siguientes elementos:

- Agentes, que son capaces de desempeñar diferentes roles, en dependencia de sus habilidades, tratando de alcanzar sus metas,
- Roles, que describen la funcionalidad de la organización,
- Interacciones, que definen las acciones que la organización permite desarrollar a los agentes, y
- Metas, que tratan de alcanzar (satisfacer) los agentes al llevar a cabo un conjunto de interacciones con otros agentes.

Una vez definidos estos elementos, TOAST asocia aleatoriamente los roles a los agentes. Luego, siguiendo una distribución normal asigna los valores de habilidad para cada tupla $\langle A, R, I \rangle$.

En la ejecución de un experimento, utilizando este escenario de prueba se repiten los siguientes pasos:

- Se genera una meta para un agente. Cada meta tiene asociada una determinada interacción que el agente debe llevar a cabo para lograrla. Cada agente tiene una cola de metas por satisfacer.

- Se obtiene la lista de posibles agentes contrapartes a partir de la lista de agentes (definida para la interacción que favorece la obtención de la meta) con los que se puede interactuar para obtener la meta deseada.
- Mediante el modelo de confianza, se selecciona de la lista anterior el agente del que se espera mejor comportamiento en la interacción.
- Se envía a cada agente, la evaluación del resto de agentes que han participado en la interacción. Cada agente utiliza estos valores para actualizar su propio modelo

3.2.5. ART: *Agent Reputation and Trust Testbed*

La iniciativa que guía el desarrollo de ART [50] ha estado motivada por el objetivo de proporcionar un escenario de prueba para las tecnologías relacionadas con la confianza y la reputación. El equipo director, formado por investigadores de diferentes países, coordina el diseño del dominio de aplicación, las especificaciones del juego empleado, el desarrollo del escenario, así como la administración de las diferentes competiciones que se realizan utilizándolo. ART ha sido concebido para utilizarse de dos formas: (1) como escenario de competición donde los investigadores pueden comparar sus tecnologías a partir de métricas objetivas, y (2) como un conjunto de herramientas experimentales con parámetros flexibles que permiten a los investigadores desarrollar experimentos personalizados y fácilmente repetibles.

En algunos estudios anteriores, varios investigadores [9, 49, 112] han destacado la necesidad de estándares para justificar el funcionamiento de los modelos de confianza y reputación. Las métricas que se ofrezcan en este sentido, deben permitir la evaluación del comportamiento de los sistemas de forma transparente y estándar. ART se presenta como un escenario universal para fomentar la cohesión de los esfuerzos de los investigadores en el área de confianza y reputación, ofreciendo métodos unificados de experimentación que permitan dar solución a los problemas existentes. Mediante métricas objetivas y bien definidas, ofrece a los investigadores un conjunto de herramientas para la comparación y validación de sus enfoques. También pudiera servir como un medio objetivo para la presentación, al resto de la comunidad investigadora, de las características (ventajosas o desventajosas) de las diferentes tecnologías.

El estudio de este escenario de prueba se muestra de forma detallada en el apartado 6.3 por cuanto, será utilizado en la realización de algunos experimentos relacionados con el comportamiento del modelo que se presenta en esta tesis. El capítulo 8 comenta como el modelo ha sido adecuado a este escenario de prueba, así como las características y resultados más relevantes de dichos experimentos en los que se comparan diferentes alternativas de configuración del modelo presentado. También se comentan algunas incompatibilidades entre el modelo y el escenario de prueba, de forma que sirvan como recomendaciones para el desarrollo ulterior de ART.

3.3. Conclusiones

Los modelos de confianza y reputación necesitan la definición de modelos o escenarios de prueba y evaluación que permitan llevar a cabo el estudio de su comportamiento y del ajuste a las condiciones reales del problema que abordan. También deben facilitar las herramientas necesarias para la comparación entre modelos.

La mayoría de los modelos basan el análisis de su comportamiento en métricas individuales sobre el desempeño de los agentes. Para ello, estudian la evolución de la utilidad alcanzada por el agente y de las medidas de confianza y reputación.

Algunos modelos agregan dichas métricas individuales como indicadores del comportamiento de los agentes en el sistema (o de un grupo de ellos), fundamentalmente.

Así, dado que la evaluación y comparación de modelos es un problema que requiere gran atención, han existido una gran variedad de escenarios de prueba. Cada escenario ha surgido para dar respuesta a las propias necesidades y condiciones de los modelos para los cuales han sido propuestos. Por lo general, con grandes limitaciones para ser adaptado a otros modelos con condiciones y características diferentes.

En este sentido, ART supone un gran avance en el desarrollo de un escenario de prueba común para todos los modelos de confianza y reputación. El desarrollo de este escenario está guiado por la necesidad de ofrecer un escenario de prueba unificado para la evaluación y comparación bajo las mismas condiciones, independiente de las peculiaridades de cada modelo.

ART puede ser utilizado tanto para la experimentación, que tiene lugar en el propio desarrollo de los modelos, como para la comparación de diferentes modelos a modo de juego de competición. Para ello, ofrece un conjunto de herramientas gráficas que facilitan el análisis de la evolución de medidas relacionadas con la utilidad de los agentes, los errores cometidos en la estimación de los valores de confianza y reputación, la reputación de cada agente, las peculiaridades de los mensajes de confianza y reputación intercambiados, entre otros.

Se decide utilizar ART para la evaluación y validación experimental de algunas estrategias y configuraciones del modelo de confianza y reputación presentado en esta tesis. Con ello se pretende, (1) seleccionar las de mejor rendimiento para ser utilizadas en la implementación de las funcionalidades de un agente concreto, y (2) analizar las posibles características del modelo que no son consideradas por ART (y que dan lugar a incompatibilidades entre ambos enfoques), de forma que sirvan como recomendaciones para el desarrollo futuro de ART.

Capítulo 4

Representaciones semánticas

Tal como se ha comentado en los capítulos anteriores, existen varios elementos de los modelos de confianza y reputación que dependen directamente del dominio de aplicación tratado. Según el funcionamiento y mecanismos ofrecidos, los modelos de confianza y reputación necesitan adoptar una representación para la definición de elementos tales como los requerimientos de los usuarios, las soluciones que ofrecen los agentes, la representación del entorno, de su contexto, entre otros. Por ejemplo, en un modelo que se aplica a un escenario del tipo consumidor - proveedor de recursos, es muy importante contar con una representación adecuada para, al menos, los elementos que describen las necesidades del consumidor y las ofertas de los proveedores.

La representación adoptada por los modelos no solo debe garantizar la definición de los conceptos tratados, sino también debe ofrecer soluciones para la definición de las relaciones entre ellos. Por ejemplo, no basta con poder definir los conceptos relacionados con las necesidades de los consumidores y las ofertas de los proveedores, sino que además se debe poder definir la relación de satisfacción, que indique en qué medida una oferta es capaz de satisfacer las necesidades de un consumidor. Según los mecanismos del modelo, también podría pensarse, por ejemplo, en funciones que permitan determinar el grado de similitud entre las necesidades de dos consumidores.

Sin embargo, no es suficiente con que cada modelo adopte una representación propia para dichos elementos. Por una parte, es necesario que se utilice alguna propuesta que resulte útil para representar los atributos de los conceptos y relaciones que son de interés para el tratamiento de la confianza y la reputación. Y por otra parte, que ofrezcan en alguna medida la interoperabilidad entre los diferentes modelos de confianza y reputación y con otras tecnologías. En el caso del escenario consumidor - proveedor de recursos, dada la naturaleza del escenario, sería importante la adopción de un modelo de representación orientado a servicios. Por ello, la mayor parte de este capítulo se dedica al análisis de algunas propuestas en este sentido.

Actualmente, las arquitecturas orientadas a servicios [43], donde las funcionalidades de los sistemas son representadas como servicios distribuidos, son cada vez más utilizadas. En el mundo de los agentes, los servicios web y la web semántica favorecen la interoperabilidad entre los sistemas, haciendo uso de las tecnologías web, a la vez que ofrecen un vínculo

entre el paradigma basado en agentes y el de los sistemas basados en servicios [42].

Por tal razón, la descripción de los servicios web de manera entendible por otras entidades software constituye un aspecto importante para la cooperación dinámica y escalable entre los diferentes sistemas y organizaciones. Los servicios web ofrecidos por aplicaciones que cooperan pueden ser automáticamente descubiertos e invocados por elementos de otras aplicaciones. Los servicios pueden ser combinados para alcanzar objetivos más complejos o funcionalidades de valor añadido, gracias a la cooperación de las diferentes aplicaciones interconectadas, sin ningún acuerdo previo o modificación del código. Estos esfuerzos de integración de manera flexible y con coste mínimo son los que guían el desarrollo de lo que se conoce como servicios web semánticos.

Esta tecnología consiste en enriquecer la descripción de los servicios con una semántica entendible por todos los sistemas y dispositivos interconectados en la web, de forma que el descubrimiento de servicios, su composición e invocación se pueda realizar automáticamente por parte de entidades software (por ejemplo, agentes inteligentes) capaces de procesar la información semántica disponible [51].

Existen varios trabajos que han propuesto representaciones ontológicas que abordan los conceptos relacionados con la confianza y la reputación, para dar respuesta a la necesidad de adoptar un marco común para la descripción de los elementos tratados en los modelos [129, 36, 35, 31, 79, 137, 68, 102, 128].

Estas representaciones ofrecen, fundamentalmente, la descripción de conceptos y relaciones que garantizan, desde el punto de vista del tratamiento de la confianza y la reputación, el descubrimiento de los servicios adecuados. También ofrecen un marco de representación para algunos elementos dependientes del dominio de aplicación tales como la evaluación de la calidad de una solución (servicio propuesto), y la satisfacción del usuario, entre otros. Sin embargo, no consideran la representación de conceptos y relaciones que permitan utilizar la similitud entre los requerimientos de los usuarios como un método para la estimación de la confianza y la reputación en determinadas situaciones de incertidumbre.

Como se verá en los capítulos siguientes (capítulos 5 y 7), el modelo que se propone en esta tesis otorga singular importancia a la similitud entre los requerimientos de los usuarios como mecanismo para la estimación de valores de confianza y reputación en los casos en que no se dispone de conocimiento previo suficiente. Por tal razón, en el presente capítulo también se estudian diferentes medidas de similitud entre conceptos que permitan describir de manera factible la similitud entre los requerimientos de los consumidores.

Este capítulo, en la sección 4.1, continúa ofreciendo algunos comentarios sobre las principales propuestas de la web semántica para la representación del conocimiento y el razonamiento a partir de él. En la sección 4.2 se ofrece el estudio de diferentes aproximaciones para alcanzar un estándar en la tecnología de los servicios web semánticos. Luego, la sección 4.3 estudia varias propuestas para la comparación de instancias cuyos conceptos han sido descritos ontológicamente, a partir de las cuales se pretende definir las medidas de similitud utilizadas en el modelo propuesto. Al finalizar, a modo de conclusiones de este estudio, la sección 4.4 selecciona y destaca la utilidad de una propuesta de estándar en concreto, para la representación de los conceptos y funciones de un modelo de confianza y reputación. Este estándar se utiliza en el modelo que propone esta tesis en el capítulo 5.

4.1. La web semántica

La visión dada por la web semántica [130], propuesta por Berners-Lee [12], es una corriente de desarrollo dirigida por el consorcio W3C, cuyo último fin es lograr que las máquinas puedan entender, y por tanto utilizar, el contenido de la web. Esta nueva web estaría poblada por agentes o representantes software capaces de navegar y realizar operaciones por los humanos para ahorrar trabajo y optimizar los resultados. Para conseguir esta meta, la web semántica propone describir los recursos de la web con representaciones procesables (entendibles) no solo por personas, sino por programas que puedan asistir, representar, o reemplazar a las personas en tareas rutinarias o inabarcables para un humano. Las tecnologías de la web semántica buscan desarrollar una web más cohesionada, donde sea aún más fácil localizar, compartir e integrar información y servicios, para sacar un partido todavía mayor de los recursos disponibles [34, 37].

La web semántica propone superar las limitaciones de la web actual mediante la introducción de descripciones explícitas del significado, la estructura interna y la estructura global de los contenidos y servicios disponibles. Frente a la semántica implícita, el crecimiento descontrolado de recursos, y la ausencia de una organización clara de la web actual, la web semántica aboga por clasificar, dotar de estructura y anotar los recursos con semántica explícita procesable por las máquinas.

La web semántica mantiene los principios que han hecho un éxito de la web actual, como son los principios de descentralización, compartición, compatibilidad, máxima facilidad de acceso y la apertura al crecimiento y uso no previstos de antemano. En este contexto un problema clave es alcanzar un entendimiento entre las partes que han de intervenir en la construcción y explotación de la web: usuarios, desarrolladores y programas de muy diverso perfil. Rescata la noción de ontología del campo de la Inteligencia Artificial como vehículo para cumplir este objetivo.

Gruber [64] define el concepto de ontología como *a formal explicit specification of a shared conceptualization*. Una ontología es una jerarquía de conceptos con atributos y relaciones, que establece una terminología consensuada para definir redes semánticas de unidades de información interrelacionadas. Ofrece un vocabulario de clases y relaciones para describir un dominio, donde los elementos más importantes están relacionados con la compartición del conocimiento y el consenso en su representación.

Las ontologías son una fuente de términos definidos de manera precisa que se pueden compartir entre aplicaciones (y humanos). Básicamente, consisten en una jerarquía de conceptos relevantes para un dominio, un conjunto de propiedades de los conceptos y la enumeración de instancias de cada concepto. Si se utilizan las ontologías para marcar contenidos y funcionalidad en páginas web, se puede entonces (1) facilitar la comunicación comprador-vendedor en comercio electrónico, (2) habilitar búsquedas semánticas en el web, y (3) proporcionar descripciones de servicios mucho más ricas que redunden en un uso más ajustado a las necesidades del usuario, por parte de los agentes inteligentes, entre otras aplicaciones.

La web no solamente proporciona acceso a contenidos sino que también ofrece interacción y servicios (por ejemplo, comprar un libro, reservar una plaza en un vuelo,

hacer una transferencia bancaria, simular una hipoteca). Los servicios web semánticos son una línea importante de la web semántica, que propone describir no solo información sino definir ontologías de funcionalidad y procedimientos para describir servicios web: sus entradas y salidas, las condiciones necesarias para que se puedan ejecutar, los efectos que producen, o los pasos a seguir cuando se trata de un servicio compuesto. Estas descripciones procesables por las máquinas pretenden automatizar el descubrimiento, la composición, y la ejecución de servicios, así como la comunicación entre unos y otros.

4.1.1. Tecnologías para la web semántica

La tecnología que se ha creado para hacer posible la web semántica incluye lenguajes para la representación de ontologías, *parsers*, lenguajes de consulta, entornos de desarrollo, módulos de gestión (almacenamiento, acceso, actualización) de ontologías, módulos de visualización, conversión de ontologías, y otras herramientas y librerías [34].

Esta tecnología queda organizada según el modelo de capas que se muestra en la figura 4.1

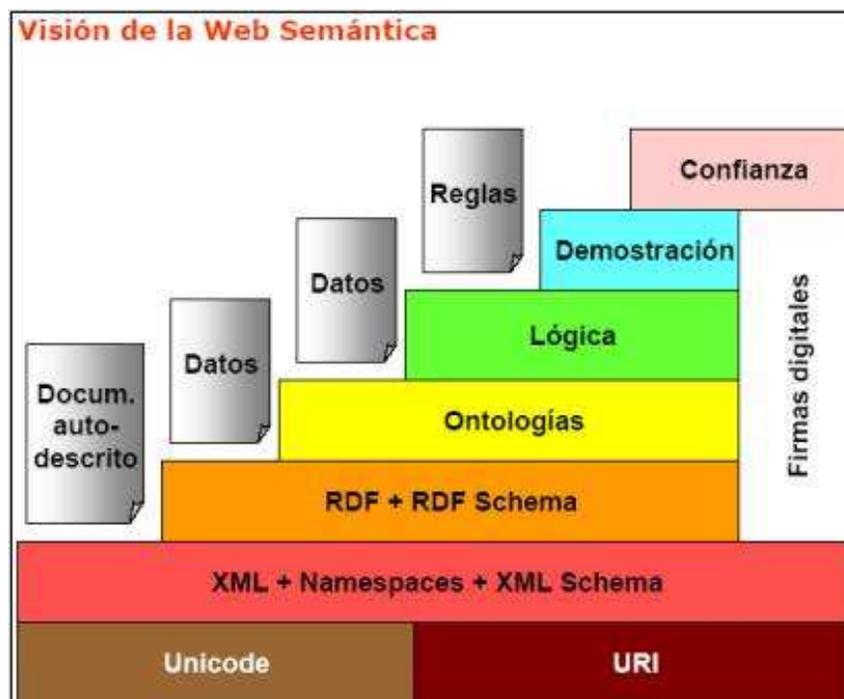


Figura 4.1: Estructura de capas de la web semántica (tomado de: [12])

Esta representación abstracta en forma de capas, relaciona los medios tecnológicos de representación con los cuales se pretenden alcanzar los objetivos de la web semántica. A continuación se comentan los aspectos más relevantes de algunos de ellos.

4.1.1.1. XML (*eXtensible Markup Language*)

El primer lenguaje para la construcción de la web semántica fue SHOE [66]. Desde entonces se han definido otros lenguajes y estándares con finalidad similar, como XML, RDF, DAML+OIL y OWL, por citar los más importantes [34].

XML representa una primera aproximación a la web semántica, y aunque no está expresamente pensado para definir ontologías, es el estándar más extendido hoy día en las aplicaciones desarrolladas con anterioridad a esta idea.

XML permite estructurar datos y documentos en forma de árboles de etiquetas con atributos. Con XML Schema (XMLS) se pueden acordar de antemano las estructuras que se van a utilizar, así como manejar tipos de datos primitivos y derivados. Con el estándar XSLT se pueden definir plantillas asociadas a las estructuras XML, que describen cómo generar código HTML para visualizar los contenidos en un navegador. *Parsers* como DOM permiten moverse por las estructuras XML desde un programa Java o C++, y existen multitud de herramientas para facilitar la compatibilidad de XML con otras tecnologías.

XML es un primer paso hacia una representación explícita de los datos y la estructura de los contenidos de la web, separada de su presentación en HTML. XML proporciona una sintaxis para hacerlo posible, pero ofrece una capacidad nula o muy limitada para expresar la semántica. El modelo de datos XML consiste en un árbol que no distingue entre objetos y relaciones, ni tiene noción de jerarquía de clases.

4.1.1.2. RDF (*Resource Description Framework*) y RDFS (*RDF Schema*)

RDF es un lenguaje para la definición de ontologías y metadatos en la web. Actualmente es el estándar más popular y extendido en la comunidad de la web semántica. El elemento de construcción básica en RDF es la sentencia, que consiste en dos nodos (sujeto y objeto) unidos por un arco (predicado), donde los nodos representan recursos, y los arcos, propiedades.

El sujeto identifica al recurso (persona, lugar o cosa) que la sentencia describe. Un recurso RDF puede ser cualquier elemento en un modelo de datos (*libro, producto, cliente, etc.*). A cada recurso se le asigna un identificador único en forma de URI.

El predicado representa una propiedad o atributo (*nombre, país, ciudad, fecha de nacimiento, etc.*) del sujeto (persona, lugar o cosa). Al igual que los sujetos, cada propiedad viene identificada por medio de una URI única.

Un objeto especifica el valor (*Gabriel García Márquez, Colombia, Aracataca, 6 de marzo de 1927, etc.*) para la propiedad (*nombre, país, ciudad, fecha de nacimiento, etc.*) del sujeto (persona, lugar o cosa).

Por ejemplo una sentencia podría expresar el hecho de que el autor (predicado) del libro *Cien años de soledad* (sujeto) fue el escritor *Gabriel García Márquez* (objeto), como se ilustra en la figura 4.2. Encadenando estas sentencias simples se construyen grafos o redes semánticas para la web [34].

RDF es un lenguaje de propósito general para representar información en la web. Ofrece una forma de expresar enunciados simples acerca de los recursos utilizando propiedades y

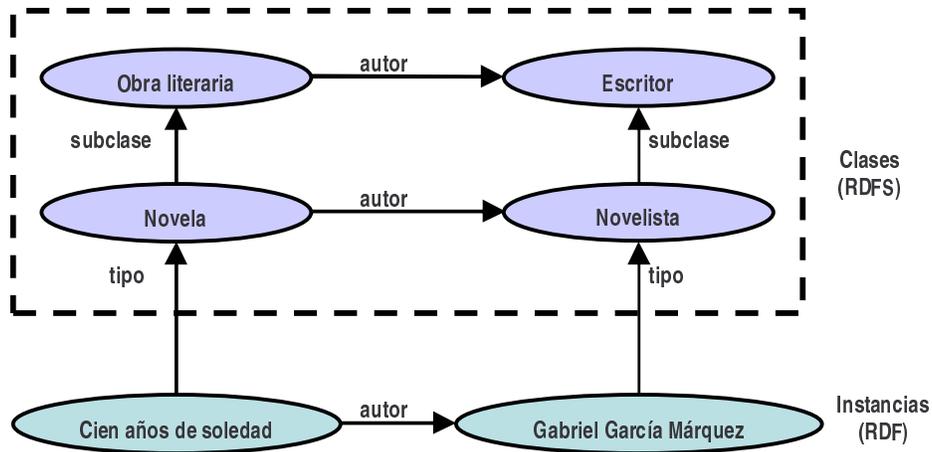


Figura 4.2: Ejemplo de sentencia RDF y del esquema RDF

valores.

RDF *Schema* (RDFS) describe cómo utilizar RDF para describir vocabularios. Con RDFS se pueden definir jerarquías de clases de recursos, especificando las propiedades y relaciones que se admiten entre ellas [16] (Figura 4.2). En RDF las clases, relaciones, y las propias sentencias son también recursos, y por lo tanto se pueden examinar y recorrer como parte del grafo, o incluso establecer sentencias sobre ellas. RDFS facilita realizar mejoras en los procesos de búsqueda y permite hacer inferencias.

Se han definido diferentes formas sintácticas para la formulación escrita de RDF, pero quizás la más extendida es la basada en XML. Es por ello que RDF se presenta a menudo como una extensión de XML.

El sistema de clases y propiedades de RDFS es similar al sistema de tipos de algunos lenguajes de programación orientados a objetos, con la diferencia de que, en vez de definir clases en términos de las propiedades que las instancias de estas clases deben tener, RDFS describe propiedades en términos de las clases de recursos a los que éstas se pueden aplicar. Para este propósito, aparecen los roles de dominio y rango. Así, por ejemplo, se podría definir la propiedad *eg:Novelista* con un dominio *eg:Novela* y un rango *eg:Escritor*. De este modo, es posible definir más propiedades que tengan como dominio *eg:Novela* o como rango *eg:Escritor* sin necesidad de modificar la descripción original de estas clases [51].

Por tanto, se puede reconocer RDFS como un lenguaje ontológico primitivo que ofrece algunas primitivas para el modelado de clases, relaciones de subclases, propiedades, relaciones de subpropiedades y restricciones de dominio y rango, con un significado fijo [5]. Desgraciadamente, RDFS es demasiado simple como lenguaje de modelado para la Web y carece de varias primitivas de modelado interesantes en relación a otros lenguajes más recientes. Entre sus limitaciones, cabe destacar las siguientes [5]:

- **Ámbito local de las propiedades:** no permite que las restricciones se apliquen solo a algunas clases (por ejemplo, *Las vacas solo comen hierba*).

- Clases disjuntas (por ejemplo, clase *Hombre* vs. clase *Mujer*).
- Combinaciones booleanas de clases: definir clases mediante unión, intersección, complementario, etc. (por ejemplo, $Persona = Hombre \cup Mujer$).
- Restricciones de cardinalidad (por ejemplo, *una persona solo tiene dos progenitores*).
- Características de propiedades: transitividad, unicidad, inversa, etc.

Sin embargo, a pesar de las limitaciones presentadas, es posible, con el uso de RDF y RDFS, realizar diversas tareas de razonamiento automáticas para inferir nuevas relaciones a partir de una base de conocimiento dada [51].

4.1.1.3. OWL (*Web Ontology Language*)

Luego de RDF y RDFS, para satisfacer algunos requerimientos que aun permanecían sin solución, se generaron paralelamente dos lenguajes: OIL (*Ontology Inference Language*), desarrollado en Europa, y DAML (*DARPA Agent Markup Language*), en EE.UU., dos lenguajes muy similares que posteriormente se fundieron en DAML+OIL. A partir de esta unión se definió el lenguaje OWL (*Web Ontology Language*), que reúne las ventajas de DAML+OIL y resuelve los problemas de este lenguaje.

OWL se puede formular en RDF, por lo que se suele considerar una extensión de éste. OWL incluye toda la capacidad expresiva de RDF(S) y la extiende con la posibilidad de utilizar expresiones lógicas. Entre las principales posibilidades adicionales que proporciona OWL sobre sus predecesores se pueden relacionar [90]:

- Definición de clases mediante restricciones sobre propiedades, valores o cardinalidad.
- Definición de clases mediante operaciones booleanas sobre otras clases: intersección, unión y complemento.
- Relaciones entre clases (por ejemplo, inclusión, disyunción, equivalencia).
- Propiedades de las relaciones (por ejemplo, inversa, simétrica, transitiva).
- Cardinalidad (por ejemplo, *únicamente una*).
- Igualdad y desigualdad de clases.
- Igualdad y desigualdad de instancias.
- Clases enumeradas.

Estas características introducidas en el lenguaje, si bien mejoran el poder expresivo del lenguaje, también limitan la capacidad de razonamiento para inferir nuevo conocimiento a partir de sentencias dadas. En particular, no se pueden garantizar (1) que en el proceso de razonamiento se encuentren todas las conclusiones válidas (completitud), ni (2) que se

obtenga la respuesta para cualquier entrada en un período de tiempo finito (*decidability*). Así se han creado tres variantes de OWL que presentan diferentes niveles de expresividad [90]:

- *OWL Lite*: es el lenguaje de menor expresividad. Permite una jerarquía de clasificación y restricciones simples (restricciones de cardinalidad, pero solo valores de cardinalidad de 0 ó 1). Entre las ventajas, se encuentran la facilidad de entender (por parte de los usuarios) y la facilidad de implementar (para constructores de herramientas). La mayor desventaja es, por supuesto, la restringida expresividad.
- *OWL DL*: de expresividad intermedia, se basa en la lógica descriptiva. Es el lenguaje indicado para aquellos usuarios que requieren el máximo de expresividad, mientras se conservan las garantías de que todas las conclusiones pueden ser obtenidas y de que se pueden obtener en tiempo finito. Incluye todos los constructores del lenguaje OWL, pero solamente se pueden utilizar bajo ciertas restricciones.
- *OWL Full*: permite la máxima expresividad y ofrece la libertad sintáctica de RDF pero carece de garantías en cuanto a la propiedad de ser computable.

Cada uno de estos sublenguajes es una extensión de su predecesor más simple, de forma que las formulaciones de un nivel de expresividad inferior son válidas también en un nivel de expresividad superior. Esto implica que toda fórmula inferida en *OWL Lite*, por ejemplo, puede ser válidamente concluida en *OWL DL* u *OWL Full*, partiendo del mismo conocimiento, y con la misma asunción.

Tal como se ha comentado, el propósito de OWL es similar al de RDFS, servir de vocabulario XML para definir clases, sus propiedades y las relaciones entre clases. En comparación con RDFS, OWL permite expresar relaciones mucho más ricas y, por tanto, disponer de capacidades de inferencia mejoradas. La semántica formal y el soporte al razonamiento en OWL se garantizan a través del mapeo de OWL en formalismos lógicos utilizando la lógica de predicados (lógica de primer orden) y la lógica descriptiva [5, 51].

4.2. Propuestas para servicios web semánticos

Por lo que respecta a los servicios web, entre los estándares ya asentados, basados en XML, podemos citar el lenguaje WSDL (*Web Service Description Language*) [13] para la descripción de servicios; el protocolo SOAP (*Simple Object Access Protocol*) [92] para el intercambio de mensajes entre servicios; y el sistema de directorio UDDI (*Universal Description, Discovery and Integration*) [1] para localizar servicios dinámicamente. BPEL4WS (*Business Process Execution Language for Web Services*) [38], desarrollado por IBM, añade a WSDL descripciones para coordinar el flujo de operaciones en procesos de negocios y transacciones.

Actualmente, W3C está examinando cuatro propuestas a estándares que abordan la problemática de los servicios web semánticos: OWL-S, WSMO, SWSF y WSDL-S. A continuación se destacan las características más importantes de cada uno.

4.2.1. OWL-S (*Ontology Web Language for Services*)

OWL-S (antes DAML-S) [100] es una propuesta de ontología para la descripción de servicios web semánticos que brinda las herramientas necesarias para permitir, con un alto grado de automatización, el descubrimiento, invocación, composición y monitorización de recursos web en forma de servicios con determinadas características.

OWL-S, similar a otras propuestas, considera la existencia de dos tipos fundamentales de servicios: atómicos y compuestos. Un servicio atómico es aquel donde una única entidad (servicio) accesible por la web es invocada, realiza su tarea y devuelve la respuesta. En cambio, un servicio compuesto es aquel que está constituido por múltiples servicios atómicos que pueden requerir una comunicación extendida entre quién lo invoca y el conjunto de servicios que lo componen [51]. Además ofrece mecanismos para garantizar:

- el descubrimiento automático de los servicios web - OWL-S permite anunciar de forma declarativa las propiedades y capacidades de los servicios.
- La invocación automática de los servicios web - OWL-S proporciona una API declarativa e interpretable que incluye la semántica de los argumentos a especificar cuando se ejecutan las llamadas, así como la semántica de los mensajes que se producen como resultado de la ejecución del servicio tanto cuando éste falla como cuando tiene éxito.
- La composición e interoperación automática de los servicios web - OWL-S proporciona especificaciones declarativas de los prerequisites y consecuencias de la aplicación de servicios individuales y un lenguaje para describir composiciones de servicio e interacciones de flujos de datos.

OWL-S permite la descripción de los servicios web en términos de un Perfil (*Profile*), que indica qué hace el servicio, un Modelo del Proceso (*Process Model*), que indica cómo funciona el servicio, y una Base (*Grounding*), que indica cómo acceder al servicio. El Perfil y el Modelo del Proceso son consideradas especificaciones abstractas en el sentido que no detallan el formato de los mensajes, de los protocolos ni de las direcciones de red que permiten instanciar un servicio web. El papel de la Base (*Grounding*) es ofrecer estos detalles concretos. WSDL (*Web Services Description Language*), desarrollado independientemente a OWL-S, brinda los medios necesarios para especificar estos elementos [100].

4.2.2. WSMO (*Web Service Modeling Ontology*)

Esta propuesta de ontología para la descripción de servicios web semánticos está basada en el marco conceptual propuesto por Fensel y Bussler [47]. WSMO refina y extiende este marco de trabajo, a la vez que desarrolla la ontología y los lenguajes formales.

Los elementos de esta representación ontológica se agrupan bajo cuatro categorías fundamentales [42]:

- Ontologías - que ofrecen los conceptos y relaciones que se presentan entre dichos conceptos, definiendo una terminología común utilizada por los otros elementos. Todos los conceptos definidos por WSMO son elementos del propio modelo. Cada uno de estos conceptos posee un conjunto de atributos no funcionales que los caracteriza.
- Descripciones de los servicios web - que representan los diferentes aspectos de las entidades computacionales que definen los servicios, ofreciendo, entre otras cuestiones, la forma de acceder a ellos. Están formadas por aspectos funcionales, atributos no funcionales y de comportamiento. Son los elementos capaces de satisfacer los objetivos o requerimientos de los usuarios.
- Metas u Objetivos - que representan los objetivos o requerimientos de los usuarios al consultar un servicio, es decir, la caracterización del problema que se desea resolver por medio de los servicios web.
- Mediadores - cuyo principal objetivo es resolver los problemas de interoperabilidad entre el resto de los elementos.

Según WSMO, los servicios web están descritos fundamentalmente por los siguientes componentes:

- un conjunto de propiedades no funcionales,
- la ontología del dominio y, posiblemente, un conjunto de mediadores que permiten la reutilización de otras ontologías definidas anteriormente,
- un conjunto de habilidades, que definen el nivel funcional del servicio web, y
- Un conjunto de interfaces del servicio para permitir la composición automática de los servicios con el objetivo de brindar funcionalidades más complejas [42].

Otro elemento central en WSMO es la representación de las metas. Las metas representan los objetivos a alcanzar con la ejecución del servicio web. Se describen de forma similar a los servicios web, de manera que determinan las características de los servicios web que satisfacen los requerimientos del usuario.

4.2.2.1. WSML (*Web Service Modeling Language*)

El lenguaje formal definido por WSMO para la descripción de los elementos que componen la ontología es *Web Services Modeling Language* (WSML). Los objetivos principales de este lenguaje son: 1) desarrollar un lenguaje de formalización propio para la especificación de los servicios web semánticos, y 2) brindar un lenguaje basado en reglas para la web semántica [40, 140].

WSML tiene diferentes variantes, cada una de ellas se corresponde con los diferentes niveles de expresividad lógica que garantiza. Todas las variantes de WSML son definidas en términos de una sintaxis entendible por los humanos a partir de palabras claves, de

manera similar a los elementos del modelo conceptual de WSMO. Además, ofrece sintaxis de intercambio XML, RDF y OWL con el objetivo de garantizar la interoperabilidad con otros modelos [40, 140].

WSML brinda las herramientas sintácticas y semánticas formales necesarias para facilitar el proceso de descubrimiento e invocación automática de los servicios web en WSMO.

4.2.2.2. Descubrimiento de servicios web según WSMO

Para el descubrimiento de los servicios, cuyas descripciones satisfacen las metas especificadas, WSMO utiliza diferentes variantes, en dependencia de los requerimientos de la aplicación [140]:

- Descubrimiento basado en palabras claves
- Descubrimiento basado en la descripción semántica simple de los servicios
- Descubrimiento basado en la descripción semántica enriquecida de los servicios

El descubrimiento de servicios a partir de palabras claves es un componente elemental en el marco de trabajo para el descubrimiento de servicios web semánticos. Por medio de este método se puede filtrar o establecer una clasificación rápida de un gran número de servicios. El centro del proceso de descubrimiento de servicios según WSMO no es precisamente esta variante. Sin embargo, se considera una técnica muy útil en un marco de trabajo completo para el descubrimiento de servicios.

La especificación de WSMO ofrece a su vez dos alternativas para efectuar el descubrimiento a partir de la descripción semántica simple de los servicios web. Una de ellas está basada en la teoría de conjuntos y la otra, en la lógica. Ambas utilizan las ontologías como representación formal, entendible por los sistemas automáticos, acerca del dominio del conocimiento tratado.

Concretamente, la alternativa que utiliza la teoría de conjuntos se basa en el hecho de que la mayoría de los modelos de conocimiento adoptados por las empresas utilizan un enfoque basado en la representación de los objetos, sus atributos y sus interrelaciones. En tales enfoques, las propiedades semánticas más importantes son las relaciones entre estos conjuntos u objetos. El establecimiento y control de estas relaciones es la principal tarea de razonamiento que permite a los agentes analizar el conocimiento formalizado en el modelo del dominio (ontología). Esta es una perspectiva de modelado muy natural y simple desde el punto de vista de los humanos.

En este sentido, una descripción de servicio debe representar los objetos que pueden ser entregados por el servicio. Mientras que una meta puede representarse por el conjunto de elementos relevantes para el cliente como resultados de la ejecución de un servicio. Así, según esta alternativa de descubrimiento, las descripciones de servicios y metas son consideradas como un conjunto de objetos. Las descripciones simples de estos conjuntos hacen referencia a las ontologías que representan el dominio de conocimiento tratado.

Siguiendo este enfoque, el proceso de descubrimiento no necesita conocer en detalles la semántica de los objetos, sino analizar las coincidencias entre los objetos correspondientes a ambos elementos. Sin embargo, al definir el conjunto de objetos existe ambigüedad al especificar la intención del servicio web o de la meta. Por ejemplo, se pudiera especificar que se desean satisfacer *todos* los atributos de una meta, o solo *algunos* de ellos. Igual puede suceder con los servicios, pueden indicar que ofrecen *todos* los objetos, o solo *algunos* de ellos.

Tabla 4.1: Coincidencias entre las metas y las descripciones de los servicios web, según la relevancia de sus atributos y la intención especificada en el proceso de descubrimiento (tomada de [140]).

Intención	$I_w = \forall$		$I_w = \exists$	
$I_g = \forall$	$R_g = R_w$	Total	$R_g = R_w$	Parcial
	$R_g \subseteq R_w$	Total	$R_g \subseteq R_w$	Nula
	$R_g \supseteq R_w$	Parcial	$R_g \supseteq R_w$	Parcial
	$R_g \cap R_w \neq 0$	Parcial	$R_g \cap R_w \neq 0$	Nula
	$R_g \cap R_w = 0$	Nula	$R_g \cap R_w = 0$	Nula
$I_g = \exists$	$R_g = R_w$	Total	$R_g = R_w$	Total
	$R_g \subseteq R_w$	Total	$R_g \subseteq R_w$	Nula
	$R_g \supseteq R_w$	Total	$R_g \supseteq R_w$	Total
	$R_g \cap R_w \neq 0$	Total	$R_g \cap R_w \neq 0$	Nula
	$R_g \cap R_w = 0$	Nula	$R_g \cap R_w = 0$	Nula

donde si U es el conjunto de todos los objetos utilizados para describir una meta o servicio web, entonces:

- $R_g \subset U$ representa el conjunto de objetos relevantes de la meta G ,
- $R_w \subset U$ representa el conjunto de objetos relevantes del servicio web W ,
- $I_g \in \{\forall, \exists\}$ representa la intención de la meta G , y
- $I_w \in \{\forall, \exists\}$ representa la intención del servicio web W .

La intención que especifica que se desean satisfacer todos los objetos de una meta o se ofrecen todos los objetos de un servicio web, se le llama universal y se denota por \forall . En cambio, la intención existencial, denotada por \exists , hace referencia a los casos que se desean satisfacer, o se ofrecen, solo algunos.

Al llevar a cabo el proceso de descubrimiento de los servicios web, se requiere especificar la intención de las descripciones de los servicios y de las metas. Muchas veces, estas intenciones quedan reflejadas implícitamente al seleccionar un conjunto de objetos propiedades en cada concepto.

Especificando el conjunto de objetos relevantes de las metas y los servicios web, así como la intención de cada uno, se pueden presentar los diferentes tipos de coincidencias en el proceso de descubrimiento. La tabla 4.1 muestra los tipos de coincidencia que se pueden presentar entre las metas y los servicios web a partir de la comparación de sus atributos relevantes y la intención especificada.

4.2.3. SWSF (*Semantic Web Services Framework*)

Este marco de trabajo pretende soportar el desarrollo de herramientas y metodologías en el entorno de los servicios web, así como facilitar el uso de procesos de razonamiento sobre dichos servicios. Al igual que el resto de aproximaciones, SWSF pretende incorporar una semántica enriquecida que de soporte a un elevado nivel de automatización de tareas como la selección e invocación de servicios, traducción del contenido de los mensajes entre servicios heterogéneos que interoperan, composición de servicios, etc. También permite llevar a cabo aproximaciones integrales para la monitorización de servicios y la recuperación de errores [125].

Esta propuesta consiste en dos componentes fundamentales: *Semantic Web Services Language (SWSL)* y *Semantic Web Services Ontology (SWSO)*.

SWSL es un lenguaje lógico utilizado para especificar las caracterizaciones formales de los conceptos de los servicios web y las descripciones individuales de los servicios. Incluye dos sublenguajes: SWSL-FOL y SWSL-Rules. SWSL-FOL se basa en la lógica de primer orden y es usado principalmente para expresar la caracterización formal (ontología) de los conceptos relativos a los servicios web. Por otra parte, SWSL-Rules está basado en el paradigma de programación lógica [15] y utilizado para dar soporte al uso de la ontología de servicios en procesos de razonamiento y en entornos de ejecución basados en ese paradigma.

SWSO define un modelo conceptual por el cual los servicios web pueden ser descritos y una representación formal o axiomatización de este modelo. La axiomatización completa es dada en lógica de primer orden utilizando SWSL-FOL, con una semántica de teoría de modelos que especifica un significado preciso de los conceptos de la ontología. Esta forma de la ontología que utiliza la lógica de primer orden se denomina *First-Order Logic Ontology for Web Services (FLOWS)*. Adicionalmente, los axiomas de FLOWS han sido traducidos al lenguaje de reglas SWSL-Rules obteniéndose de este modo una ontología basada en semántica de programación lógica denominada *Rules Ontology for Web Services (ROWS)* [125].

FLOWS, al igual que WSMO y OWL-S (sobre el cual se fundamenta y extiende), es una ontología formal que trata los conceptos relacionados con servicios, proporcionando un marco de trabajo conceptual para describirlos y razonar sobre ellos [51].

FLOWS está subdividido en tres ontologías fundamentales:

- **Descriptor de servicio:** ofrece la información básica acerca de un servicio web de forma similar a un directorio del tipo de páginas amarillas, independientes del dominio, incluyendo el concepto de Perfil de OWL-S (consultar la sección 4.2.1). Los descriptores se utilizan generalmente para el descubrimiento automatizado de

servicios web. Existe una lista inicial de descriptores que permiten identificar el nombre del servicio, el autor, la descripción textual, la versión etc., pudiendo extenderse de acuerdo a los requerimientos de los usuarios.

- Modelo de procesos: extiende la ontología de procesos genérica propuesta por PSL (*Process Specification Language*) para hacer disponibles los conceptos que son útiles en el contexto de los servicios web. Se añaden dos elementos fundamentales sobre PSL: la noción estructurada de proceso atómico tal y como se especifica en OWL-S, y una infraestructura para especificar varios tipos de flujos de datos.
- Base: de manera similar a OWL-S, esta ontología relaciona por medio de WSDL los elementos abstractos de los servicios (conceptos de SWSO para la descripción de servicios y las instanciaciones de éstos para describir un servicio particular) con detalles concretos de su implementación (por ejemplo, el formato de los mensajes, los protocolos de transporte y las direcciones de red).

4.2.4. WSDL-S (*Web Services Semantics*)

WSDL-S ha supuesto un cambio radical con respecto a las perspectivas tradicionales para la incorporación de semántica a los servicios web [51]. WSDL-S define un mecanismo para asociar anotaciones semánticas a los servicios que han sido descritos mediante WSDL. A diferencia del resto de aproximaciones, WSDL-S asume la existencia de modelos semánticos del dominio relevantes para cada servicio, mantenidos fuera del ámbito de los documentos WSDL y que pueden ser referenciados desde un documento WSDL a través de elementos de extensibilidad planteados por la propia propuesta WSDL-S [3].

Esta aproximación se basa, al igual que en los casos anteriores, en el hecho de que el estándar WSDL actual opera a un nivel sintáctico que carece de la expresividad semántica necesaria para representar los requisitos y capacidades de los servicios. WSDL-S enriquece la expresividad de WSDL con semántica mediante el empleo de conceptos similares a los de OWL-S, incluyendo las definiciones de precondiciones, entradas, salidas y efectos de las operaciones de servicios web. Los elementos y atributos de extensibilidad planteados por esta propuesta son los siguientes:

- *modelReference* - para especificar la asociación entre una entidad WSDL y un concepto en algún modelo semántico. Este atributo puede añadirse a un tipo complejo, elemento, operación, así como a los elementos de extensión *precondition* y *effect*.
- *schemaMapping* - que se le añade a los elementos XSD y a los tipos complejos, para manejar diferencias estructurales entre los elementos de diseño de un servicio web y sus correspondientes conceptos del modelo semántico.
- *precondition* y *effect* - (situados como descendientes de *operation*). Se utilizan principalmente para el descubrimiento de servicios y no son estrictamente necesarios para la invocación de un servicio dado.

- *category* - consiste en información de categorización de servicios que podría ser utilizado cuando se publica un servicio en un registro de servicios web como UDDI. Posibilita mejorar el proceso de búsqueda de servicios en los registros de directorios.

Al utilizar este enfoque los usuarios pueden describir de forma incremental todos los detalles tanto semánticos como a nivel de operaciones en WSDL, un lenguaje ampliamente conocido. Por otra parte, al dejar fuera los modelos semánticos del dominio, WSDL-S permanece independiente del lenguaje de representación de ontologías, permitiendo a los desarrolladores de servicios web anotar sus servicios con el lenguaje ontológico de su elección (e.g. UML, OWL, etc.). Esto supone una ventaja adicional, por cuanto posibilita la reutilización de modelos del dominio existentes, expresados con lenguajes de modelado como UML, al incorporar las anotaciones semánticas. Además, resulta relativamente fácil actualizar las herramientas existentes que tienen en cuenta las especificaciones de WSDL para que sean capaces de considerar este enfoque.

4.3. Medidas de similitud entre conceptos

Por lo general, y quizás por ser la manera más natural simple e intuitiva de representación del conocimiento, los conceptos de una ontología están descritos por un conjunto de atributos o propiedades. Cada combinación de valores que reciba el conjunto de los atributos define una instancia diferente de dicho concepto. De ahí que la similitud entre dos instancias determinadas se puede establecer mediante una función que considera el conjunto de propiedades que describen cada concepto al que pertenecen. En este sentido, existen varias técnicas para medir la similitud entre dos instancias dadas [115, 119, 107, 8, 127, 109, 151, 81, 46, 83]. A continuación se presentan las más destacadas.

4.3.1. Coeficiente de correlación de Pearson

El coeficiente de correlación de Pearson es utilizado para establecer medidas de similitud entre instancias de un mismo concepto [119, 107, 151]. Si s_i y s_j son las dos instancias a comparar, cada una descrita por k atributos, este coeficiente queda escrito de la siguiente manera:

$$D(s_i, s_j) = \frac{\sum_k (s_{ik} - \bar{s}_i) \cdot (s_{jk} - \bar{s}_j)}{\sqrt{\sum_k (s_{ik} - \bar{s}_i)^2 \cdot \sum_k (s_{jk} - \bar{s}_j)^2}}$$

donde s_{ik} y s_{jk} representan el k -ésimo atributo de la instancia i y j , respectivamente; y \bar{s}_i y \bar{s}_j son los valores medios de todos los atributos de las dos instancias s_i y s_j , respectivamente

Esta medida de similitud es útil cuando tiene sentido hablar del valor medio de los atributos de los conceptos. Existen casos en que cada atributo del concepto hace referencia a elementos que indican magnitudes diferentes o, además, sus valores se encuentran descritos mediante dominios y escalas diferentes.

En este sentido, existen otras variantes de funciones basadas en el análisis de los atributos según determinados criterios, y que son utilizadas en casos donde carece de sentido el valor medio de los atributos que describen el concepto.

4.3.2. Medida a partir de la agregación de criterios de comparación

Faratin *et al.* [46], para guiar las negociaciones entre agentes, utilizan una heurística que selecciona el contrato más parecido (similar) a la última propuesta que ha hecho el agente contraparte. Para estimar la similitud entre contratos usa el concepto de similitud difusa planteada por Zadeh [149]. Según estas ideas, la similitud entre dos contratos x e y es definida como una combinación ponderada de las similitudes de las características de cada contrato:

$$D(x, y) = \sum_{j \in J} w_j \cdot D_j(x_j, y_j)$$

donde J es el conjunto de características de los contratos y w_j la importancia que se le otorga a cada característica para estimar la similitud entre dos contratos, de forma que $\sum_{j \in J} w_j = 1$. $D_j(x_j, y_j)$ es el valor de similitud entre dos características x_j, y_j que se obtiene de la siguiente manera:

$$D_j(x_j, y_j) = \sum_{1 \leq i \leq m} w_i \cdot (h_i(x_j) \leftrightarrow h_i(y_j))$$

donde $\{h_1, \dots, h_m\}$ es el conjunto de criterios de comparación; \leftrightarrow es el operador de equivalencia considerado, definido como $1 - |h_i(x_j) - h_i(y_j)|$; y w_i es el conjunto de los pesos que determinan la importancia de los criterios de similitud, tal que $\sum_{1 \leq i \leq m} w_i = 1$.

Este método, siguiendo una interpretación bastante intuitiva, define para cada variable de decisión un conjunto de criterios de evaluación $\{h_1, \dots, h_m\}$. De esta forma, la medida de similitud determina cuánto un par de contratos satisfacen los criterios de comparación considerados, valor éste que es obtenido a partir de la agregación de los valores de similitud para cada uno de los criterios.

4.3.3. Medida de Tversky

La medida ofrecida por Tversky [127] es utilizada para establecer un indicador de la similitud entre instancias de conceptos pertenecientes, incluso, a diferentes ontologías [109]. Esta medida produce un valor de similitud que no solo analiza los elementos comunes entre ambos conceptos, sino que refleja también sus diferencias.

De manera general, la similitud entre dos conceptos a y b puede definirse de la siguiente forma:

$$D(a, b) = \frac{|A \cap B|}{|A \cap B| + \alpha(a, b)|A \setminus B| + (1 - \alpha(a, b))|B \setminus A|}$$

donde $0 < \alpha < 1$, y A y B son los conjuntos de propiedades de los conceptos a y b , respectivamente.

A diferencia de otros enfoques, no presupone el requerimiento de simetría de la similitud. La importancia relativa de los atributos (o propiedades) no comunes posibilita la evaluación asimétrica de la similitud. Esta medida solo es simétrica para los casos en que $\alpha(a, b) = 0,5$.

Rodríguez y Egenhofer [109] definen la función α a partir de la profundidad de los conceptos comparados, según la estructura jerárquica de la ontología a la que pertenecen.

Tal como se ha mostrado en trabajos propios [21], esta medida puede ser utilizada para estimar la similitud simétrica entre dos metas de WSMO s_i y s_j , en cuyo caso el parámetro $\alpha(s_i, s_j) = 0,5$ (ambos conceptos pertenecen a la misma clase). La comparación de ambas metas puede ser analizada a partir de la consideración de los atributos relevantes siguiendo las ideas y notación definidas en el apartado 4.2.2:

$$D(s_i, s_j) = \frac{|R_{gi} \cap R_{gj}|}{|R_{gi} \cap R_{gj}| + 0,5|R_{gi} \setminus R_{gj}| + 0,5|R_{gj} \setminus R_{gi}|}$$

donde R_{gi} y R_{gj} son los conjuntos de atributos relevantes de las metas s_i y s_j , respectivamente. La definición de estos conjuntos de atributos relevantes se ofrece en el apartado 5.8.1.

4.3.4. Medida basada en la comparación léxica

La distancia de edición definida por Levenshtein [80] es una medida utilizada para medir la diferencia entre dos cadenas de caracteres. Ella cuenta el número mínimo de inserciones, eliminaciones o sustituciones que se requieren para transformar una cadena en otra. Por ejemplo, la distancia de edición entre *Multiagente* y *Multi-agente* es igual a 1, por cuanto solo se requiere la inserción del carácter '-' para transformar la primera cadena en la segunda.

Maedche y Staab [83] definen una medida de similitud entre cadenas de caracteres a partir de la distancia de edición propuesta por Levenshtein:

$$D(L_i, L_j) = \max\left(0, \frac{\min(|L_i|, |L_j|) - ed(L_i, L_j)}{\min(|L_i|, |L_j|)}\right)$$

donde $ed(L_i, L_j)$ representa la distancia de edición entre las dos cadenas L_i y L_j . Los valores de esta medida de similitud se encuentran en el intervalo $[0, 1]$, donde 1 significa una similitud completa entre las dos cadenas y 0 una diferencia total.

Si se intenta generalizar esta medida de similitud de dos cadenas de caracteres para la definición de una medida de la similitud entre dos conceptos que están formados por un conjunto de atributos (como por ejemplo, los conceptos de metas y de servicios web definidos por WSMO), puede pensarse en el siguiente planteamiento:

$$D(s_i, s_j) = \max\left(0, \frac{\min(|R_{gi}|, |R_{gj}|) - ed(s_i, s_j)}{\min(|R_{gi}|, |R_{gj}|)}\right)$$

donde, según la notación utilizada por WSMO en el apartado 4.2.2, U representa el conjunto de las propiedades utilizadas para la definición de los conceptos, y R_g el conjunto de

atributos relevantes. Así, R_{gi} y R_{gj} son los conjuntos de atributos relevantes de las metas s_i y s_j , respectivamente. Por su parte, $ed(s_i, s_j)$ está dada por la cantidad de atributos relevantes que son diferentes en ambos conceptos, pudiendo escribirse de la siguiente manera:

$$ed(s_i, s_j) = |U \setminus (R_{gi} \cap R_{gj})|$$

4.4. Relación con el enfoque de confianza y reputación

Tal como se comenta en el apartado 2.2.3, se requiere de un modelo semántico de representación común para los requerimientos de los usuarios y las respuestas en un entorno de interacción entre pares donde los agentes desempeñan roles de proveedores y consumidores de servicios. Se necesita que los conceptos del modelo tengan en cuenta atributos que permitan la representación y manipulación de información de confianza y reputación entre los agentes.

WSMO es una propuesta de representación ontológica factible para representar las estructuras de conocimiento necesarias en un modelo de confianza y reputación entre agentes para un entorno de interacción entre pares basado en servicios web, donde los agentes pueden ser consumidores o proveedores de recursos. Utilizando WSMO, los requerimientos pueden ser representados por una especificación de tarea definida mediante el concepto de meta; mientras que las respuestas (es decir, los servicios web que satisfacen estos requerimientos) pueden ser descritos por medio del concepto de servicio web.

Esta tesis propone utilizar WSMO como marco conceptual para la representación de las especificaciones de tareas y las respuestas de los servicios web, en el modelo que se ofrece en el capítulo 5, por cuanto

1. WSMO es una nueva propuesta de estándar que puede ser utilizada para representar una gran variedad de situaciones donde los usuarios necesitan encontrar el recurso idóneo en un entorno de interacción entre pares basado en servicios web;
2. ofrece grandes facilidades para la representación y descubrimiento de servicios web de acuerdo a las requerimientos de la aplicación;
3. los conceptos de meta y servicio web están formados, entre otros, por un grupo de propiedades no funcionales que pueden ser utilizadas para gestionar la confianza y la calidad del proceso de descubrimiento de los servicios web; y
4. los procesos de descubrimiento de servicios web que ofrece WSMO pueden ser utilizados como medidas de calidad acerca de la satisfacción de una determinada tarea dado un servicio web.

De acuerdo con la razón número 3, se han identificado algunas propiedades no funcionales que pueden resultar interesantes para la gestión de la confianza y la reputación, la calidad, los costos, etc. Estos atributos no funcionales pueden ser utilizados para establecer medidas de similitud entre dos especificaciones de tareas (metas) y medidas

de satisfacción de las especificaciones de tareas a partir de la respuestas dadas por los servicios web. Estas medidas son necesarias para la gestión del modelo de confianza y reputación. Entre los atributos no funcionales que se propone considerar en la aplicación de WSMO al modelo de confianza destacan:

- precisión (*accuracy*) - número de errores generados por el servicio en un cierto intervalo de tiempo.
- calidad de servicio asociado a la red (*network-related QoS*) - representa los mecanismos de calidad de servicio que operan en la red de transporte. Puede ser medido utilizando la demora de la red, la variación de dicha demora y la pérdida de mensajes.
- rendimiento (*performance*) - representa cuan rápido la solicitud del servicio web puede ser satisfecha. Puede ser medida en términos de *throughput*, latencia, tiempo de ejecución, tiempo de transacción, tiempo de respuesta del servicio.
- fiabilidad (*reliability*) - representa la capacidad de un servicio web para desempeñar sus funciones (mantener la calidad del servicio). Puede ser medida mediante el número de fallos del servicio en un intervalo de tiempo.
- robustness (*robustness*) - representa la capacidad del servicio para funcionar en presencia de entradas incompletas o no válidas. Puede ser medida a partir del número de entradas incompletas o no válidas para las cuales el servicio web continúa funcionando correctamente.
- escalabilidad (*scalability*) - representa la capacidad del servicio web para procesar un número creciente de solicitudes. Puede ser medida mediante el número de solicitudes satisfechas en un intervalo de tiempo.
- confianza (*trust*) - representa la honradez (confiabilidad) de un servicio web.

Estos elementos de WSMO permiten adicionar nuevos atributos propios del dominio de aplicación abordado, de forma que la descripción de los requerimientos y de los servicios puede ser ampliada según atributos no funcionales propios del dominio de aplicación tratado (por ejemplo, ancho de banda requerido, velocidad de descarga, entre otros).

Por otra parte, para un modelo de confianza y reputación puede resultar muy útil considerar la similitud entre los requerimientos de los usuarios para realizar estimaciones de la confianza y la reputación cuando el conocimiento disponible no sea suficiente. El modelo, que se presenta en el capítulo 5 de esta tesis, otorga singular importancia a la similitud entre los requerimientos de los usuarios como mecanismo para la estimación de valores de confianza y reputación en estos casos. Por ello, adecúa algunas de las medidas presentadas en el apartado 4.3 de este capítulo para determinar la similitud entre dos tareas representadas mediante conceptos y atributos de WSMO. Luego, en algunos experimentos que se presentan en el capítulo 7 se pretende estudiar el efecto de varias medidas de

similitud entre los requerimientos de los usuarios para identificar, de ser posible, aquella que mejor estimación haga de los valores de confianza y reputación asociados a requerimientos desconocidos.

Parte III

Modelo de Confianza y Reputación

Notación utilizada

A continuación se relacionan los principales símbolos utilizados durante esta parte del documento de tesis.

A - conjunto de los agentes en el sistema.

S - conjunto de especificaciones de tareas que representan los requerimientos de los usuarios.

W - conjunto de soluciones de servicio web que los agentes pueden ofrecer como respuestas (soluciones) a los requerimientos de los usuarios.

$IET_i^{(t)}$ - base de experiencias del agente a_i , en la ronda t , para obtener los valores de confianza directa acerca de la capacidad del resto de agentes para solucionar los requerimientos de los usuarios.

$IER_i^{(t)}$ - base de experiencias del agente a_i , en la ronda t , para obtener los valores de confianza que se tienen en los agentes para ofrecer recomendaciones acerca del desempeño de otros.

$et_{i,j,k,l}$ - grado de satisfacción del agente a_i cuando el agente a_j ofrece una solución para la especificación de tarea s_k por l -ésima vez.

$er_{i,j,k}$ - grado de satisfacción que el agente a_i tiene cuando el agente a_j ofrece información de reputación sobre el desempeño de otros agentes al desempeñar la tarea s_k .

$T^{(t)}(a_i, a_j, s_k)$ - función que determina el valor de confianza global $f_{i,j,k}^{(t)}$ que el agente a_i otorga, en la ronda t , al agente a_j para ofrecer soluciones a una determinada especificación de tarea s_k . Se obtiene combinando $DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$, $DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$, y $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$.

$R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ - función que determina el valor de reputación asignado por el agente a_i al agente a_j , en la ronda t , para la especificación de tarea s_k de acuerdo a las experiencias de los agentes más confiables ofreciendo información de reputación,

agrupados en $CR_{sup}^{(t)}(a_i, s_k)$.

$DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ - función que determina el valor de confianza directa que el agente a_i asigna al agente a_j , en la ronda t , para la especificación de tarea s_k de acuerdo a las experiencias que posee en su propia base $IET_i^{(t)}$. Se utiliza un factor de desgaste δ para ofrecer menos relevancia a las experiencias más antiguas.

$DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ - función que determina la precisión del valor de confianza directa $DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ a partir del análisis de la cantidad de experiencias utilizadas para producir dicho valor y su desviación alrededor del valor de confianza directa. Se obtiene a partir de las funciones $D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ y $N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$.

$D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ - función que devuelve, en la ronda t , una medida de la desviación de las experiencias del agente a_i , contenidas en su base de experiencias para la confianza ($IET_i^{(t)}$), que están relacionadas con el desempeño del agente a_j para la ejecución de la tarea s_k .

$N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ - función que devuelve, en la ronda t , una medida de la cantidad de experiencias del agente a_i , contenidas en su base de experiencias para la confianza ($IET_i^{(t)}$), que están relacionadas con el desempeño del agente a_j para la ejecución de la tarea s_k . El número máximo de experiencias que mejora la medida de confianza directa está dado por el parámetro itm .

δ - parámetro modulador del tiempo, que otorga mayor importancia a las experiencias más recientes.

itm - parámetro para controlar el número máximo de experiencias consideradas para mejorar la precisión de la medida de confianza, de forma que al considerar un número de experiencias mayor a itm la precisión de la métrica de confianza no mejora.

L (o $L_{i,j,k}^{(t)}$) - subconjunto de las diferentes experiencias de $IET_i^{(t)}$, que el agente a_i tiene, en la ronda t , sobre el comportamiento del agente a_j asociado a la realización de la tarea s_k . En las fórmulas presentadas en el modelo se han eliminado los índices en aras de simplificar la escritura de las ecuaciones.

N_i - conjunto de todos los vecinos del agente a_i .

$CT_{sup}^{(t)}(a_i, s_k) \subset N_i$ - conjunto de los agentes más confiables, según el agente a_i para dar una respuesta a la especificación de tarea s_k , en la ronda t .

$CT_{inf}^{(t)}(a_i, s_k) \subset N_i$ - conjunto de los agentes menos confiables, según el agente a_i para dar una respuesta a la especificación de tarea s_k , en la ronda t .

$CT_{dud}^{(t)}(a_i, s_k) \subset N_i$ - conjunto de los agentes con un grado de confiabilidad dudosa, según el agente a_i , para dar una respuesta a la especificación de tarea s_k , en la ronda t .

$f_{i,j,k}^{(t)}$ - valor de confianza global que un agente a_i otorga, en la ronda t , al agente a_j para ofrecer soluciones a una determinada especificación de tarea s_k . Se obtiene mediante la función $T^{(t)}(a_i, a_j, s_k)$.

$CR_{sup}^{(t)}(a_i, s_k) \subset N_i$ - conjunto de los agentes más confiables ofreciendo opiniones sobre el desempeño de otros para solucionar la especificación de tarea s_k , según el criterio de a_i en la ronda t .

p_γ - percentil γ del conjunto de valores de reputación $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ de todos los agentes $a_j \in CT_{dud}^{(t)}(a_i, s_k)$ que pueden ofrecer una solución para la especificación de tarea s_k , planteada por el agente a_i .

$C^{(t)}(a_i, s_k) \subset N_i$ - conjunto de los agentes vecinos del agente a_i que, según las predicciones del modelo de confianza y reputación en la ronda t , brindarán una buena solución a la especificación de tarea s_k . En una estrategia determinada estos agentes pueden ser los más confiables para ofrecer soluciones a la especificación de tarea s_k , en la ronda t , según el criterio del agente a_i , agrupados en $CT_{sup}^{(t)}(a_i, s_k) \subset N_i$.

$C_{prom}^{(t)}(a_i, s_k) \subset N_i$ - conjunto compuesto por los agentes más confiables (según el criterio del agente a_i en la ronda t) ofreciendo soluciones a la especificación de tarea s_k y por aquellos de confianza dudosa pero con un elevado valor de reputación asociado a la realización de dicha especificación de tarea. Se ofrece en el modelo como una estrategia de selección de los agentes con quienes interactuar.

$IT^{(t)}(a_i, a_j, s_k, s_p, IET_i^{(t)})$ - función que utiliza el agente a_i para estimar, a partir de su propia base de experiencias, los valores de confianza directa en el agente a_j asociados a una tarea desconocida s_k , a partir de la combinación de la confianza directa relativa a otra tarea conocida s_p y la similitud entre ambas.

$IR^{(t)}(a_i, a_j, s_k, s_p, CR_{sup}^{(t)}(a_i, s_p))$ - función que utiliza el agente a_i para estimar los valores de reputación en el agente a_j asociados a una tarea desconocida s_k , a partir de la información de reputación ofrecida por el grupo de agentes más confiables ofreciendo opiniones sobre otros para una tarea similar $CR_{sup}^{(t)}(a_i, s_p)$.

$D(s_k, s_p)$ - función que determina el grado de similitud entre las especificaciones de tareas s_k y s_p .

$ed(s_k, s_p)$ - distancia de edición utilizada en la definición de la medida de similitud a

partir de la comparación léxica de las tareas s_k y s_p .

\oplus - operador que define la forma en que se combinan la confianza y la reputación con la similitud para estimar los valores asociados a tareas desconocidas.

$ec_{i,j,k}$ - valor de calidad acordado entre los agentes a_i y a_j para la realización de la tarea s_k .

R_u - conjunto de atributos utilizados para describir las especificaciones de tareas y las respuestas ofrecidas por los servicios web.

$R_g, R_w \subset R_u$ - conjuntos formados por los atributos más relevantes de cada especificación de tarea s_k o respuesta w_j , respectivamente.

$s_k.b_i \in R_u$ y $w_j.b_i \in R_u$ - representación de cada uno los i atributos de la especificación de tarea s_k y la respuesta w_j , respectivamente. Por cuestiones de simplicidad en la representación de algunas ecuaciones, también pueden aparecer como s_{ki} y w_{ji} , respectivamente.

$v_{s_k}(b_i)$ y $v_{w_j}(b_i)$ - funciones que determinan la conveniencia (o relevancia) del atributo b_i para la especificación de tarea s_k y la solución w_j , respectivamente.

$I^{(t)}(a_i, s_k)$ - información relativa a la interacción que el agente a_i lleva a cabo para satisfacer la tarea s_k , durante la ronda t .

$M_j^{(t)}(a_i, s_k)$ - conjunto de los agentes recomendados al agente a_i por el agente a_j para ofrecer soluciones a la especificación de tarea s_k en la ronda t .

$Q(w_j, s_k)$ - función que determina la calidad de una respuesta w_j a partir de la especificación de tarea s_k .

$P(ec_{i,j,k}, w_j, s_k)$ - función que determina el cumplimiento del compromiso de calidad $ec_{i,j,k}$ con la calidad que ofrece la respuesta w_j a la especificación de tarea s_k . Se basa principalmente en la comparación de los valores que determinan el compromiso de calidad $ec_{i,j,k}$ y el valor que ofrece la función de calidad $Q(w_j, s_k)$.

Capítulo 5

El modelo propuesto: TRSIM

Este capítulo describe la estructura y funcionalidad del modelo de confianza y reputación TRSIM que considera la similitud entre tareas. Este modelo se basa en el análisis del comportamiento de los agentes en interacciones pasadas a partir de las experiencias directas de cada agente y de la información intercambiada con otros. Analizando el conocimiento propio del agente (dado por las experiencias directas) se obtiene un valor de confianza que se combina con el valor de reputación producido a partir de la información ofrecida por terceros. En el modelo, la confianza y la reputación son vistos como un fenómeno social basado en múltiples interacciones entre los agentes y el intercambio de información entre ellos.

TRSIM está caracterizado por los siguientes supuestos generales: (1) es posible calcular los valores de confianza y reputación para indicar quién confía en quién, (2) estos valores se basan en la experiencia de cada agente (analizando el conocimiento generado en interacciones anteriores) y el intercambio de conocimiento entre los agentes, y (3) es posible ajustar dinámicamente estos valores mediante la evaluación de cada interacción y su inclusión en la base de conocimientos del agente.

Por otro lado, siguiendo la clasificación dada por Ramchurn *et al.* [104], este es un modelo adaptativo basado en el aprendizaje y la evolución de las medidas de confianza y reputación correspondiente al nivel individual (ver la introducción del capítulo 2).

TRSIM considera que los conceptos de confianza y reputación están asociados a los requerimientos de los usuarios. Para este modelo la presencia de determinados requerimientos o preferencias de los usuarios forma parte de su contexto. Por lo tanto, el comportamiento diferenciado de la confianza y la reputación ante diferentes requerimientos o preferencias puede ser considerado de manera similar al comportamiento según los diferentes contextos del agente. De esta manera, cada valor de confianza o reputación considerado en el modelo está asociado a la especificación de la tarea (requerimientos o preferencias de los usuarios) que se necesita resolver.

Así, por ejemplo, si un agente necesita conocer la confianza de otro, para asignarle la realización de una determinada tarea, solo inspecciona la información de su base de experiencias relacionada con la especificación de dicha tarea. Sin embargo, es posible que no disponga de ninguna información sobre el comportamiento previo del otro para dicha tarea,

pero sí para alguna tarea similar. (El apartado 5.7 muestra los detalles de cómo se utiliza el conocimiento relacionado con una tarea similar a la especificada en la interacción.) De igual forma sucede con las solicitudes de información de reputación. Es posible que, al solicitar información sobre el comportamiento de otros, no se disponga de referencias para la tarea especificada y sí para alguna tarea similar. En estos casos, TRSIM propone estimar los valores de confianza y reputación utilizando el conocimiento disponible (propio u ofrecido por otros) sobre alguna tarea similar. En el algoritmo que se presenta en el apartado 5.11 se ilustra, entre otras cuestiones, cómo se lleva a cabo este proceso de decisión.

De manera general, además de la intuición que supone la aproximación de los valores de confianza y reputación, el modelo utiliza las experiencias directas, la información proveniente de terceros y algunos conceptos relacionados con prejuicios como otras fuentes de información (las generalidades de cada fuente de confianza se tratan en el apartado 2.1.1).

La información de confianza y/o reputación proveniente de cada una de estas fuentes se complementa mutuamente. Las experiencias directas son una fuente muy fiable de información de confianza y reputación, por cuanto reflejan la evaluación que hace el propio agente. Sin embargo, pueden existir situaciones (dadas, por ejemplo, por interacciones en instantes de tiempo tempranos, por requerimientos de nuevas especificaciones de tareas a resolver, por la presencia de vecinos nuevos o de los que se desconoce su desempeño, entre otras) en que no se disponga de experiencias para emitir un criterio de confianza directa. En estos casos, se necesita utilizar otras fuentes de información de confianza y reputación, que aunque sean menos fiables pueden estar más disponibles (por ejemplo, la información proveniente de terceros, la intuición a partir de comportamientos en situaciones similares, y los prejuicios dado la pertenencia a determinado grupo).

Teniendo en cuenta estas cuatro fuentes de información de la confianza y la reputación, TRSIM realiza una evaluación multidimensional de atributos tales como el de competencias, previsibilidad e integridad de los agentes.

Otro punto importante de este modelo es la adopción de un marco de representación base para la definición de los conceptos y funciones dependientes del dominio de aplicación. Propone la utilización de WSMO para la definición de las estructuras de conocimiento necesarias en el modelo. Las solicitudes de servicios de información se representan por medio del concepto de meta de WSMO, mientras que la respuesta a dicha solicitud es representada utilizando el concepto de servicio web dado por WSMO. Basado en la representación dada por WSMO, el modelo define, entre otras, métricas para la evaluación del grado de similitud entre dos especificaciones de tareas dadas y la calidad de una respuesta teniendo en cuenta la especificación de tarea a la que debe satisfacer.

Una vez mencionadas las características generales del modelo, el presente capítulo continúa con la descripción detallada de sus elementos. La sección 5.1 describe la estructura del modelo y la relación entre sus diferentes partes. El apartado 5.2 comenta cómo se utilizan los criterios de confianza para identificar diferentes grupos de agentes, de acuerdo a la calidad de la respuesta que se espera de ellos.

El concepto de confianza considerado en el modelo está determinado a partir de dos dimensiones fundamentalmente: (1) la confianza de un agente para ofrecer soluciones y (2)

la confianza de un agente para ofrecer información sobre el comportamiento de otros. En la sección 5.3 se define cómo determinar la confiabilidad en un agente para ofrecer soluciones a partir de la combinación de la confianza directa y la reputación. Por su parte, la sección 5.4 establece la forma de obtener la confiabilidad en un agente para ofrecer información sobre el comportamiento de otros. La forma de calcular los valores de la confianza directa y la reputación se muestra en las secciones 5.5 y 5.6, respectivamente.

TRSIM también ofrece un mecanismo para la estimación de los valores de confianza y reputación asociados a especificaciones de tareas desconocidas para los agentes. La propuesta para la aproximación de dichos valores utilizando criterios de similitud entre tareas es presentada en la sección 5.7. En este sentido, el modelo requiere la utilización de una medida de similitud entre dos tareas determinadas. La definición de la función de similitud entre tareas se ofrece en la sección 5.8. Esta es una función dependiente de las características del dominio de aplicación. En la misma sección donde se describe la similitud entre tareas, se define otra función dependiente del dominio de aplicación: la calidad de la respuesta a partir de los requerimientos de los usuarios. Esta función es necesaria en el modelo por cuanto indica el nivel de satisfacción de la tarea con la respuesta recomendada por el modelo. La sección 5.9 describe como la función de calidad, entre otras funciones del modelo, es utilizada en la actualización de la información de confianza y reputación que se mantiene en las bases de experiencias de cada agente. La sección 5.10 puntualiza las características más importantes asociadas al desarrollo de TRSIM. Al finalizar, la sección 5.11 presenta el algoritmo utilizado por un agente al determinar los valores de confianza y reputación en sus vecinos, para la solución de una determinada tarea.

5.1. Estructura del modelo

El modelo está compuesto por un conjunto de bases de información, donde cada agente almacena información necesaria para representar el comportamiento de otros, y un conjunto de funciones para operar convenientemente con estas bases. Siguiendo el enfoque distribuido, impuesto por un entorno de interacción entre pares, cada agente mantiene dos bases de experiencias. Las funciones que operan con la información almacenada en dichas bases, producen valores que ayudan fundamentalmente a guiar las interacciones entre los agentes, teniendo en cuenta criterios de confianza y reputación. Básicamente, el modelo se estructura y opera según el esquema ofrecido en la figura 5.1.

A continuación se relacionan los principales elementos (bases de experiencias y funciones) del modelo:

- *IET* (acrónimo en inglés de *Interactions Experiences of Trust*): es la base que gestiona las experiencias para la producir los valores de confianza directa acerca de la capacidad del resto de agentes para desempeñar una determinada tarea.
- *IER* (acrónimo de, en inglés, *Interactions Experiences of Reputation*): es la base que gestiona las experiencias relacionadas con la confianza que se tienen en los agentes para ofrecer recomendaciones acerca del desempeño de otros.

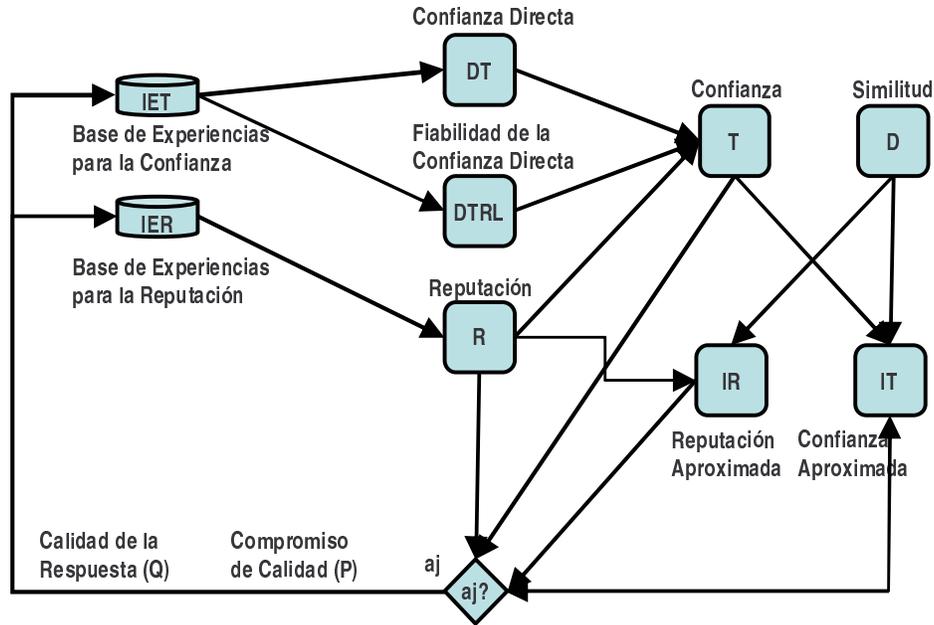


Figura 5.1: Relación entre las diferentes partes del modelo

- *DT* (acrónimo de, en inglés, *Direct Trust*): es la función que encuesta la base de experiencias para la confianza *IET* y devuelve un valor de confianza a partir de las experiencias directa del agente.
- *DTRL* (acrónimo de, en inglés, *Direct Trust Reliability*): es una función que analiza la precisión de la medida de confianza directa que se obtiene mediante la función *DT*. También encuesta la base de experiencias para la confianza *IET* analizando la cantidad de experiencias utilizadas para obtener el valor de *DT* y la desviación de las experiencias utilizadas con respecto al valor de *DT*.
- *R* (acrónimo de, en inglés, *Reputation*): es la función que devuelve el valor de reputación de un agente, a partir del criterio ofrecido por otros.
- *T* (acrónimo de, en inglés, *Trust*): es la función que combina adecuadamente los valores de confianza directa *DT* y reputación *R*, teniendo en cuenta la precisión ofrecida por *DTRL*.
- *D* (Similitud): función que determina el grado de similitud entre dos especificaciones de tareas. Se emplea por las funciones *IT* e *IR* en los casos en que el modelo no posee conocimiento suficiente asociado a una determinada especificación de tarea. En estas situaciones, se estiman los valores de confianza y reputación a partir de otra tarea similar sobre la que sí se disponga de conocimiento asociado a ella.
- *IT* (Confianza Directa Aproximada): es la función que se utiliza para estimar los valores de confianza directa asociados a una tarea desconocida, a partir de la

combinación de la confianza directa relativa a otra tarea conocida y la similitud entre ambas.

- *IR* (Reputación Aproximada): es la función que se utiliza para estimar los valores de reputación asociados a una especificación de tarea desconocida, a partir de la combinación de la reputación relativa a otra tarea conocida y la similitud entre ambas.
- *Q* (acrónimo de, en inglés, *Quality*): es la función que determina la calidad de una respuesta a partir de la especificación de los requerimientos de usuarios que pretende satisfacer. Se utiliza para evaluar la respuesta que recomienda el modelo de confianza como solución a una especificación de tarea dada.
- *P* (Cumplimiento del compromiso de calidad): es la función que evalúa el cumplimiento del compromiso de calidad que ha hecho un agente, al pretender dar respuesta a una determinada especificación de tarea.

Existen dos bases de experiencias para obtener los valores de confianza y reputación para una tarea dada: base de experiencias para la confianza *IET* y para la reputación *IER*. Estos valores se obtienen mediante la combinación adecuada del conjunto de funciones ofrecidas por el modelo. Primeramente, por la introspección de las bases de experiencias, el modelo calcula la confianza directa *DT*, la reputación *R* e indicadores de la confiabilidad de la confianza directa *DRTL*. Estos tres valores son combinados para producir un único valor de confianza global, utilizando la función *T*. (Las funciones *DT*, *DTRL* y *T* se definen de manera similar a ReGreT [111], por cuanto ofrece una manera eficiente y sencilla. Sin embargo, el descuento de la relevancia de las experiencias es diferente.) El valor, que se obtiene de agregar la confianza directa *DT* y la reputación *R*, es utilizado para seleccionar los agentes contrapartes en una determinada interacción, con el fin de identificar aquellos cuyo valor de confianza global haga pensar que ofrecerán buenas soluciones.

Si las bases de experiencias no poseen información sobre una determinada tarea, el modelo puede obtener los valores de confianza (utilizando la función *DT*) y reputación (utilizando la función *R*) para una tarea similar a la deseada y combinar estos valores con el grado de similitud existente entre las dos tareas, dado por la función *D*. Así, el modelo utiliza las funciones *IT* e *IR* para seleccionar los agentes contrapartes en una interacción en estos casos.

Luego de seleccionar el agente contraparte, e interactuar con él, el agente que inicia la interacción tiene como respuesta una solución a su tarea, pudiendo evaluar la interacción con este agente y ajustar su modelo de confianza. El ajuste del modelo tiene lugar mediante la actualización de la información contenida en las bases de experiencias según las estrategias de actualización utilizadas y las dos funciones siguientes: *P*, para representar el cumplimiento del grado de satisfacción prometida, y *Q*, que indica la calidad de la respuesta obtenida.

5.1.1. Estructura de las bases de experiencias

La base de experiencias IET está formada por un conjunto de tuplas almacenadas por el agente iniciador, donde se describe el desempeño de otros agentes al ser contratados para dar solución a una tarea determinada:

$$IET_i^{(t)} = \{(a_j, s_k, et_{i,j,k,l}) | a_j \in A, s_k \in S, et_{i,j,k,l} \in [0, 1]\}$$

donde $IET_i^{(t)}$ es la base de experiencias para el agente iniciador a_i en el instante de tiempo t , A es el conjunto de todos los agentes en el sistema, S es el conjunto de todas las especificaciones posibles de tareas que el agente pudiera contratar, $et_{i,j,k,l}$ es el grado de satisfacción del agente a_i cuando el agente a_j ofrece una solución para la tarea s_k por l -ésima vez.

También, el agente iniciador almacena información, en la base de experiencias IER , acerca de la confiabilidad de otros agentes para ofrecer información de reputación:

$$IER_i^{(t)} = \{(a_j, s_k, er_{i,j,k}) | a_j \in A, s_k \in S, er_{i,j,k} \in [0, 1]\}$$

donde $er_{i,j,k}$ es el grado de satisfacción que el agente a_i tiene cuando el agente a_j ofrece información de reputación sobre el desempeño de otros agentes al desempeñar la tarea s_k .

Estas bases de experiencias son actualizadas al finalizar cada interacción. El apartado 5.9 explica como son actualizadas ambas bases de experiencias.

5.2. Agrupamiento de los agentes según los criterios de confianza y reputación

De acuerdo a la concepción de TRSIM, el concepto de grupo de agentes se deja abierto a la consideración de los propios agentes que utilicen el modelo. Claramente, el criterio de agrupamiento que se siga, así como las estrategias para interactuar con cada grupo, son decisiones que corresponde tomar a cada agente, o, en última instancia, al desarrollador del sistema multi-agente en su diseño. Sin embargo, en el planteamiento de TRSIM se ofrece la identificación de varios grupos de agentes que pueden resultar de interés para el establecimiento de criterios o estrategias más complejas y reales, de acuerdo al dominio de aplicación del modelo.

De manera general, en una comunidad de agentes, cada uno mantiene una lista con sus vecinos con el objetivo de interactuar con ellos, independientemente del análisis que pueda hacer sobre su desempeño y/o funcionalidad. Los vecinos de un agente a_i pueden ser representados por el conjunto:

$$N_i = \{a_j | a_j \in A, neighbor(a_i, a_j) = true\},$$

donde la función booleana $neighbor()$ tiene una definición dependiente del dominio de aplicación tratado y de la representación utilizada.

5.2. AGRUPAMIENTO DE LOS AGENTES SEGÚN LOS CRITERIOS DE CONFIANZA Y REPUTACIÓN

El modelo considera diferentes particionamientos del conjunto de agentes vecinos según el grado de confiabilidad para ofrecer soluciones a una tarea dada, así como para ofrecer opiniones sobre el desempeño de otros frente a dicha tarea. En este sentido, se pueden formar los siguientes conjuntos de agentes vecinos, de acuerdo al grado de confianza que se tiene sobre cada uno para ofrecer soluciones a una determinada tarea:

- $CT_{sup}^{(t)}(a_i, s_k) \subset N_i$: Conjunto de los agentes más confiables para dar una respuesta,
- $CT_{inf}^{(t)}(a_i, s_k) \subset N_i$: Conjunto de los agentes menos confiables para dar una respuesta, y
- $CT_{dud}^{(t)}(a_i, s_k) = N_i \setminus (CT_{sup}^{(t)}(a_i, s_k) \cup CT_{inf}^{(t)}(a_i, s_k))$: Conjunto de los agentes con un grado de confiabilidad dudoso para dar una respuesta.

La selección de los agentes que componen los grupos $CT_{sup}^{(t)}(a_i, s_k)$ y $CT_{inf}^{(t)}(a_i, s_k)$ se realiza utilizando el método de la ruleta [91]. Para seleccionar los agentes pertenecientes al conjunto $CT_{sup}^{(t)}(a_i, s_k)$, se ubican todos los agentes de N_i sobre una ruleta, de forma que cada porción de la ruleta corresponda con la probabilidad de selección de un agente determinado. La amplitud de cada porción de la ruleta está dada por un valor proporcional a la confianza, que se tenga en el agente correspondiente, para ofrecer soluciones a la tarea dada. Al hacer girar la ruleta varias veces se pueden seleccionar, de forma probabilística, los agentes más confiables en este sentido.

Por su parte, para la obtención de $CT_{inf}^{(t)}(a_i, s_k)$ se considera que la amplitud de cada porción de la ruleta es inversamente proporcional a la confianza que se tenga en los agentes. Así, al girar la ruleta varias veces, se obtienen con mayor probabilidad los agentes de menor confianza ofreciendo soluciones.

La confianza que un agente a_i otorga a otro agente a_j para ofrecer soluciones a una determinada tarea s_k está dada por el valor $f_{i,j,k}^{(t)}$. La obtención de este valor, a partir de las bases de experiencias y de la información intercambiada por los agentes, se describe en la sección 5.3.

Siguiendo un razonamiento análogo al de particionar el conjunto de agentes según los criterios de confianza para ofrecer soluciones, se pueden identificar varios grupos de agentes teniendo en cuenta su habilidad para ofrecer opiniones acerca de otros. A pesar de que de manera sencilla pueden ser identificados tres grupos, como en el caso de la confianza, solo resulta interesante para TRSIM la identificación de los agentes más confiables ofreciendo opiniones sobre el desempeño de otros agentes para una tarea específica:

$$CR_{sup}^{(t)}(a_i, s_k) \subset N_i$$

En este grupo se encuentran los agentes a los cuales se sugiere solicitar información de reputación. Carece de interés gestionar la información de este tipo que es suministrada por agentes poco o medianamente fiables ofreciendo opiniones sobre el desempeño de otros.

Para la selección de los agentes más fiables ofreciendo opiniones sobre los demás, se utiliza también el método de la ruleta. En este caso, la amplitud de cada porción de la

ruleta está determinada por la confianza que se tenga en cada agente para ofrecer opiniones sobre el desempeño de otros en la realización de una tarea dada.

La confianza que un agente a_i otorga a otro agente a_j para ofrecer información de reputación relacionados con la tarea s_k está dada por el valor $er_{i,j,k}^{(t)}$. La obtención de este valor se describe en la sección 5.4.

Mediante este particionamiento del conjunto de vecinos, el modelo considera el prejuicio como una fuente de información más. La forma de interactuar con cada agente está en dependencia del grupo al que pertenece. En este caso, todos los agentes del mismo grupo reciben el mismo tratamiento en cuanto a la decisión de interactuar o no con cada uno de ellos. Cada agrupamiento se define a partir de las características y habilidades de los agentes. Por ejemplo, es factible pensar que el agente a_i solo solicitará información de reputación a sus vecinos cuando los considere altamente confiables para ofrecer criterios sobre otros agentes, asociados a la realización de la tarea s_k (agrupados en el conjunto $CR_{sup}^{(t)}(a_i, s_k)$).

El carácter probabilista del método de la ruleta es de gran utilidad para el modelo. Su empleo ofrece flexibilidad al proceso de selección de los agentes de cada grupo. Por ejemplo, al conformar el conjunto $CT_{sup}^{(t)}(a_i, s_k)$, no selecciona de manera exacta los de mayor confiabilidad, sino que posibilita la selección de otros menos confiables. Esto lejos de ser una imprecisión del modelo es una ventaja, por cuanto ofrece la posibilidad de seleccionar agentes de confiabilidad diferente al propósito del grupo, dando cobertura a los casos en que el comportamiento de los agentes no es estacionario.

5.2.1. Estrategias para solicitar información

Es lógico suponer que solo se solicitará soluciones para la tarea s_k a los agentes del grupo $CT_{sup}^{(t)}(a_i, s_k)$, dado que está formado por los agentes más confiables ofreciendo soluciones para dicha tarea. Sin embargo, es posible que otros agentes, por ejemplo, algunos pertenecientes al conjunto $CT_{dud}^{(t)}(a_i, s_k)$, también sean capaces de ofrecer alguna solución de valor, e incluso mejor. Esto se debe al hecho de que los valores de confianza y reputación constituyen indicadores aproximados del comportamiento de los agentes. De esta manera, destacan dos estrategias entre todas las que pueden definirse para seleccionar el conjunto $C^{(t)}(a_i, s_k)$ compuesto por los agentes que, según las predicciones del modelo de confianza y reputación, brindarán una buena solución a la tarea en cuestión:

- Agentes con mayor grado de confiabilidad asociado:

$$C^{(t)}(a_i, s_k) = CT_{sup}^{(t)}(a_i, s_k)$$

- Agentes con mayor grado de confiabilidad asociado más aquellos sobre los cuales se tiene dudas y poseen un elevado valor de reputación:

$$C^{(t)}(a_i, s_k) = CT_{sup}^{(t)}(a_i, s_k) \cup C_{prom}^{(t)}(a_i, s_k)$$

con

$$C_{prom}^{(t)}(a_i, s_k) = \{a_j \mid a_j \in CT_{dud}^{(t)}(a_i, s_k), R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k)) \geq p_\gamma\}$$

donde la función $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ devuelve el valor de reputación asignado por el agente a_i al agente a_j para la tarea s_k de acuerdo a la experiencia de los agentes más confiables ofreciendo información de reputación, agrupados en $CR_{sup}^{(t)}(a_i, s_k)$. La sección 5.6 ofrece la definición de dicha función. Por su parte, p_γ es el percentil γ del conjunto de valores de reputación $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ de todos los agentes $a_j \in CT_{dud}^{(t)}(a_i, s_k)$ que pueden ofrecer una solución para la especificación de tarea s_k , planteada por el agente a_i .

Utilizando una de estas dos alternativas, o cualquiera que el usuario del modelo pueda definir a partir de estos conjuntos, TRSIM forma las listas de agentes a los cuales se recomienda solicitar una respuesta para la especificación de tarea en cuestión.

En este sentido, es válido destacar que la selección de los agentes confiables (tanto para ofrecer soluciones como información acerca de otros) es una estrategia cuya definición queda fuera del planteamiento del modelo, correspondiendo su implementación a los usuarios que lo utilicen. El modelo solo define algunos conjuntos de agentes básicos que pueden ser útiles en la definición de otras estrategias (más complejas) del agente que utilice TRSIM. En este apartado se ofrecen dos estrategias para solicitar información (soluciones) por cuanto serán las utilizadas en algunas pruebas experimentales que se realicen con el modelo. Por ejemplo, en los experimentos del capítulo 7 se utiliza la primera estrategia donde únicamente se solicita información a los agentes más confiables.

En la sección 8.2.1, al analizar la adaptación de TRSIM al escenario de prueba ART, se comparan diferentes estrategias para crear la lista de agentes confiables a quienes solicitar soluciones (en el caso de ART, opiniones de tasación).

5.3. Confianza para ofrecer soluciones

El concepto de confianza utilizado por el modelo no solo tiene en cuenta al posible agente contraparte en la interacción, sino que los valores de confianza están asociados a la tarea que se pretende resolver. La función de confianza global, denotada por T , está definida en el rango $[0, 1]$, cuya definición algebraica es la siguiente:

$$T : A \times A \times S \rightarrow [0, 1]$$

Se recuerda que A denota el conjunto de agentes en el sistema y S el conjunto de todas las especificaciones posibles de tareas que el agente pudiera contratar. Los valores $f_{i,j,k}^{(t)}$, usados en la definición de los conjuntos $CT_{sup}^{(t)}(a_i, s_k)$ y $CT_{dud}^{(t)}(a_i, s_k)$ (ver la sección 5.2), se refieren a los valores de confianza global obtenidos utilizando esta función. Dichos valores representan la confiabilidad que para el agente a_i tiene el agente a_j para ofrecer soluciones para la tarea s_k . Así:

$$f_{i,j,k}^{(t)} \equiv T^{(t)}(a_i, a_j, s_k)$$

se ha definido de manera similar a ReGreT [111]. La formulación ofrecida por ReGreT ofrece una forma sencilla y eficiente de combinar las informaciones de confianza y reputación, obtenidas a partir de las experiencias de los agentes. Se considera que el valor de confianza global de un agente está compuesta por la confianza directa DT y la reputación R , otorgando dinámicamente más o menos importancia a la confianza directa de acuerdo a un indicador de su confiabilidad $DTRL$:

$$T^{(t)}(a_i, a_j, s_k) = DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) + \\ (1 - DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)})) R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$$

donde $DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ representa el valor de confianza directa que el agente a_i asigna al agente a_j para la tarea s_k de acuerdo a las experiencias que posee en su propia base $IET_i^{(t)}$; $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ es el valor de reputación que el agente a_i asigna al agente a_j para la tarea s_k de acuerdo a la confianza y criterios de los agentes más confiables ofreciendo información de reputación; y $0 \leq DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) \leq 1$ indica la precisión (confiabilidad) de la medida de confianza $DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$.

Siguiendo las ideas de ReGreT, el valor de $DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$, que puede ser diferente en cada interacción, es calculado a partir de la desviación ($D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$) y la multiplicidad ($N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$) de las experiencias que son utilizadas para obtener la medida de confianza directa. Las experiencias $et_{i,j,k,l}$ se recuperan mediante la introspección de la base de experiencias $IET_i^{(t)}$, combinando convenientemente las experiencias asociadas a los agentes a_j para la realización de la tarea s_k . El apartado 5.5 ofrece los detalles de la obtención de estos valores.

5.4. Confianza para ofrecer información sobre otros

La credibilidad (certeza de la veracidad) que un agente otorga a las opiniones de otros también es un elemento tratado en el modelo. De manera muy elemental, la confiabilidad de las opiniones se corresponden con la confiabilidad que se le otorgue al agente que las ofrece al emitir criterios sobre el comportamiento de otros.

La base de experiencias $IER_i^{(t)}$ almacena la confiabilidad para ofrecer información de reputación, dado por el valor $er_{i,j,k}^{(t)}$. Este valor es utilizado, entre otros lugares, en la definición del conjunto $CR_{sup}^{(t)}(a_i, s_k)$ (ver la sección 5.2) y para en la función de reputación $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ (ver la sección 5.6).

Cada agente a_i almacena en la base de experiencias $IER_i^{(t)}$ un único valor de confiabilidad $er_{i,j,k}^{(t)}$ acerca del agente a_j al ofrecer opiniones sobre otros, para cada tarea s_k . Al igual que la base de experiencias IET , esta base actualiza este valor tras concluir cada interacción (ver la sección 5.9).

5.5. Confianza directa

La confianza directa (DT) es un término que se utiliza en el modelo para referirse a la confianza que se obtiene a partir de las experiencias directas del agente. Asociado a este concepto se define el de precisión de la confianza directa ($DTRL$), para indicar el valor informativo que ofrece la confianza directa en dependencia de la cantidad y calidad de la información utilizada en su obtención. Por ejemplo, $DTRL$ debe indicar una mayor precisión de la medida de confianza directa DT en los casos cuando se ha utilizado para obtenerla un número relativamente grande de experiencias y la variabilidad entre ellas es poca.

El valor de confianza directa DT y su precisión $DTRL$ son obtenidos mediante funciones que encuestan la base de experiencias $IET_i^{(t)}$. Para la combinación de las experiencias directas, TRSIM utiliza un enfoque descontado, teniendo en cuenta que las experiencias pierden relevancia con el tiempo. Si $0 \leq \delta \leq 1$ es un parámetro modulador del tiempo, que otorga mayor importancia a las experiencias más cercanas a t , la confianza directa puede ser calculada de la siguiente manera:

$$DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) = (1 - \delta)^{|L|-1} et_{i,j,k,0} + \sum_{l_p \in L} \delta(1 - \delta)^{|L|-p-1} et_{i,j,k,p}$$

donde L es el subconjunto de las diferentes experiencias que el agente a_i tiene sobre el comportamiento del agente a_j asociado a la realización de la tarea s_k ($L \subset IET_i^{(t)}$, $|L| \leq t$). El subíndice p en el nuevo conjunto L , $0 \leq p < |L|$, indica cuan vieja es la experiencia $et_{i,j,k,p}$: l_{p_2} es más reciente que la experiencia l_{p_1} solo si $p_2 > p_1$. Por su parte, $et_{i,j,k,0}$ representa la experiencia más antigua que el agente a_i tiene sobre el desempeño de a_j para la tarea s_k .

La figura 5.2 muestra los valores de los factores de relevancia, según la antigüedad de las experiencias, para diferentes valores del parámetro δ , considerando el tamaño del subconjunto de experiencias $|L| = 10$.

Para conocer la precisión de la medida de confianza directa, debido su sencillez y efectividad, se siguen las ideas de Sporas [146] y ReGreT [111]. El valor de precisión se obtiene a partir del número de experiencias usadas para calcular la confianza y la variabilidad de estas experiencias:

$$DTRL^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) = N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) \cdot (1 - D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)}))$$

donde

$$N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) = \begin{cases} \sin\left(\frac{\pi \cdot |L|}{2 \cdot itm}\right) & : |L| \leq itm \\ 1 & : \text{en otro caso} \end{cases}$$

y

$$D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)}) = (1 - \delta)^{|L|-1} (|et_{i,j,k,0} - DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})|) + \sum_{l_p \in L} \delta(1 - \delta)^{|L|-p-1} (|et_{i,j,k,p} - DT^{(t)}(a_i, a_j, s_k, IET_i^{(t)})|)$$

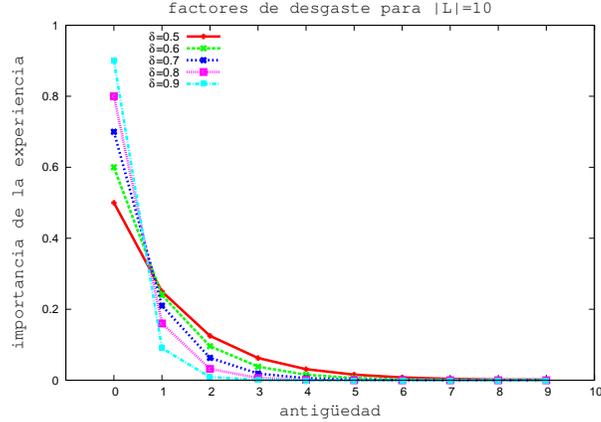


Figura 5.2: Coeficientes de relevancia para cada experiencia según su antigüedad, calculados para diferentes valores del parámetro δ y $|L| = 10$.

De esta forma, los valores más altos de precisión de la confianza directa se obtienen cuando la confianza directa se calcula a partir de un número elevado de experiencias directas con poca desviación de la media, en cada instante de tiempo. El componente $N_o^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ (en adelante N_o) es indicador de la cantidad de experiencias utilizadas para calcular el valor de la confianza directa. El componente $D_v^{(t)}(a_i, a_j, s_k, IET_i^{(t)})$ (en adelante D_v) refleja las desviaciones de cada experiencia respecto a la confianza directa DT que se ha calculado, según el modelo descontado, a partir de ellas.

El componente N_o tiene en cuenta que la precisión de la confianza directa es mayor en tanto se utilice un mayor número de experiencias directas para calcular DT . Sin embargo, considera que a partir de una cantidad de experiencias igual a itm no se mejora la precisión de DT . itm es un parámetro dependiente del dominio de aplicación para controlar el número máximo de experiencias consideradas para mejorar la precisión de la medida de confianza, de forma que al considerar un número de experiencias mayor a itm la precisión de la métrica de confianza no mejora.

La figura 5.3 muestra la función N_o para diferentes valores de $itm = 10$.

El otro componente, D_v , se obtiene mediante el mismo enfoque descontado utilizado para calcular la confianza directa. Analiza la desviación de cada experiencia respecto a la medida de confianza directa calculada, teniendo en cuenta que las diferencias entre el valor de la experiencia y la confianza directa pierden relevancia con el tiempo. El coeficiente de desgaste utilizado para calcular D_v es el mismo que el que se emplea para calcular DT (Figura 5.2).

5.6. Reputación

Al igual que los de confianza, los valores de reputación están asociados a una especificación de tarea en particular. Así, el modelo propone una función de reputación

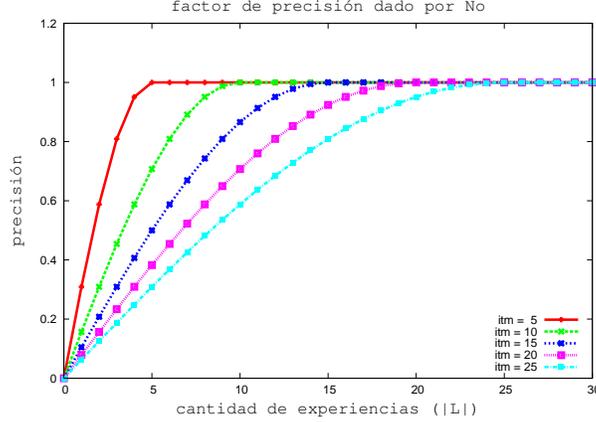


Figura 5.3: Componente de la precisión de la confianza directa ($DTRL$) que considera la cantidad de experiencias utilizadas en la obtención de la confianza directa (DT), para diferentes valores del parámetro itm .

basada en la propagación de la información de reputación a partir de las opiniones brindadas por los agentes más confiables ofreciendo información de este tipo, $CR_{sup}^{(t)}(a_i, s_k)$.

Esta función se define teniendo en cuenta algunas ideas de otros modelos de confianza y reputación que utilizan *TrustNet* [116]. Las propuestas de Golbeck y Hendler [55, 56], Zacharia [146] y Schillo [116] siguen un enfoque de propagación de la confianza y la reputación a través de la red de relaciones que se presentan entre los agentes. De esta forma, en TRSIM cuando un agente a_i desea conocer el valor de reputación del que goza un agente a_j frente a un grupo de agentes $CR_{sup}^{(t)}(a_i, s_k)$, se combinan los valores de confianza directa que tengan dichos agentes sobre el agente a_j con los valores de confianza que el agente a_i les otorga como recomendadores de otros.

La función de reputación utilizada por el modelo puede escribirse de la siguiente manera:

$$R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k)) = \frac{\sum_{a_q \in CR_{sup}^{(t)}(a_i, s_k)} DT^{(t)}(a_q, a_j, s_k, IET_q^{(t)}) \cdot er_{i,q,k}^{(t)}}{\sum_{a_q \in CR_{sup}^{(t)}(a_i, s_k)} er_{i,q,k}^{(t)}}$$

donde a_i es el agente interesado en obtener la información de reputación que al agente a_j tiene, asociado a la realización de la tarea s_k , según los criterios de los agentes a_q (agrupados en el conjunto CR , donde se encuentran los agentes más confiables ofreciendo información acerca del comportamiento de los demás).

Al solicitar información de reputación, el agente a_i encuesta a los agentes $a_q \in CR_{sup}^{(t)}(a_i, s_k)$, de quienes recibe su criterio $DT^{(t)}(a_q, a_j, s_k, IET_q^{(t)})$. Las opiniones de todos los agentes a_q son agregadas en un único valor teniendo en cuenta la confianza que a_i tiene sobre cada uno de ellos para ofrecer información acerca del comportamiento de otros.

En este sentido, las opiniones ofrecidas por los agentes encuestados a_q se corresponden con los valores de la confianza directa $DT^{(t)}(a_q, a_j, s_k, IET_q^{(t)})$ que cada agente a_q otorga

al agente a_j , objeto de la consulta. TRSIM supone que los agentes siempre responden con la verdad.

5.7. Confianza y reputación a partir de tareas similares

Tal como se ha tratado en los apartados anteriores, la reputación, la confianza directa y la precisión de ésta son medidas obtenidas a partir de la información intercambiada por los agentes y de la almacenada en las bases de experiencias $IET_i^{(t)}$ y $IER_i^{(t)}$ para la s_k a resolver en la interacción t .

Sin embargo, es posible que un agente no posea en sus bases de experiencias información sobre el comportamiento de otros para la especificación de tarea de la interacción, ni sea capaz de obtener información sobre su comportamiento a partir de las referencias de otros agentes. En este caso, TRSIM hace una propuesta original, planteando una solución a la necesidad de estimar la confianza y la reputación. Para ello utiliza la información relacionada con alguna especificación de tarea similar. El modelo obtiene esta aproximación utilizando el grado de similitud entre la especificación de tarea bien conocida y la desconocida. Por tal razón, necesita incorporar una función para obtener el grado de similitud entre dos especificaciones de tareas. El valor de esta función depende de la definición del concepto de tarea y de las métricas utilizadas para ello. El apartado 5.8.2 detalla las características de esta función y las diferentes alternativas para su definición, consideradas en el modelo.

Para estimar la confianza y la reputación, asociada a una tarea desconocida s_k , resulta interesante tener en cuenta la especificación de tarea s_p *más parecida* de la que se tenga información sobre el desempeño del agente.

Así, la especificación de tarea s_p *más parecida* a la desconocida s_k ($s_p \neq s_k$), que puede utilizar el agente a_i para estimar la confianza en el agente a_j , queda definida formalmente de la siguiente manera:

$$s_p \in S \mid \left(\exists et_{ijp}(a_j, s_p, et_{ijp}) \in IET_i^{(t)} \right) \wedge$$

$$\nexists s_r \mid D(s_k, s_r) \geq D(s_k, s_p), \forall s_r, s_k \in S, s_r \neq s_k, s_r \neq s_p.$$

De igual forma, para estimar la reputación del agente a_j se puede utilizar la información suministrada por los agentes más confiables ofreciendo recomendaciones para la tarea s_p *más parecida* a la desconocida s_k ($s_p \neq s_k$):

$$s_p \in S \mid \left(\exists er_{ijp}(a_r, s_p, er_{ijp}) \in IER_i^{(t)}, a_r \in A \right) \wedge$$

$$\nexists s_r \mid D(s_k, s_r) \geq D(s_k, s_p), \forall s_r, s_k \in S, s_r \neq s_k, s_r \neq s_p.$$

donde, se recuerda que, $IET_i^{(t)}$ y $IER_i^{(t)}$ son las bases de experiencias del agente a_i en el instante de tiempo t para la confianza y la reputación, respectivamente.

Por medio de la función de similitud entre tareas (D) y utilizando las definiciones de confianza directa (DT) y reputación (R), se definen las funciones de confianza directa aproximada y reputación aproximada, cuando no se posee información suficiente relacionada con la especificación de tarea:

$$IT : A \times A \times S \times S \rightarrow [0, 1]$$

para aproximar la confianza directa, y

$$IR : A \times A \times S \times S \rightarrow [0, 1]$$

para aproximar la reputación.

De esta forma, cuando un agente a_i no dispone de información suficiente, en su base de experiencias $IET_i^{(t)}$, sobre el desempeño de a_j para la realización de la tarea s_k , la confianza aproximada puede obtenerse según la siguiente ecuación, utilizando la información asociada a la especificación de tarea conocida *más parecida* s_p :

$$IT^{(t)}(a_i, a_j, s_k, s_p, IET_i^{(t)}) = DT^{(t)}(a_i, a_j, s_p, IET_i^{(t)}) \oplus D(s_k, s_p)$$

donde operador \oplus define la forma en que se combinan ambos valores (en el apartado 5.7.1 se ofrecen algunas alternativas para su definición).

De manera similar, cuando el agente a_i no dispone de información suficiente para determinar el grupo $CR_{sup}^{(t)}(a_i, s_k)$ de agentes recomendadores para la realización de la tarea s_k , la reputación aproximada puede obtenerse encuestando a otro grupo de agentes $CR_{sup}^{(t)}(a_i, s_p)$, confiables como recomendadores para la especificación de tarea más parecida s_p . La reputación aproximada puede obtenerse a partir de la siguiente ecuación:

$$IR^{(t)}(a_i, a_j, s_k, s_p, CR_{sup}^{(t)}(a_i, s_p)) = R^{(t)}(a_i, a_j, s_p, CR_{sup}^{(t)}(a_i, s_p)) \oplus D(s_k, s_p)$$

5.7.1. Alternativas para aproximar la confianza y la reputación

Existen varias alternativas para definir el operador \oplus para combinar los valores de confianza y reputación con el grado de similitud entre las tareas. Con el objetivo de simplificar la notación, al escribir las diferentes variantes para el operador \oplus , se utiliza:

- IT para referirse a $IT^{(t)}(a_i, a_j, s_k, s_p, IET_i^{(t)})$,
- IR para referirse a $IR^{(t)}(a_i, a_j, s_k, s_p, CR_{sup}^{(t)}(a_i, s_k))$,
- DT para referirse a $DT^{(t)}(a_i, a_j, s_p, IET_i^{(t)})$,
- R para referirse a $R^{(t)}(a_i, a_j, s_p, CR_{sup}^{(t)}(a_i, s_p))$, y
- D para referirse a $D(s_k, s_p)$

De esta forma, se pueden definir las siguientes alternativas:

$$IT_a: IT = DT \cdot D$$

$$IR = R \cdot D$$

$$IT_b: IT = \sin\left(\frac{\pi}{2}DT \cdot D\right)$$

$$IR = \sin\left(\frac{\pi}{2}R \cdot D\right)$$

$$IT_c: IT = DT \cdot \sin\left(\frac{\pi}{2}D\right)$$

$$IR = R \cdot \sin\left(\frac{\pi}{2}D\right)$$

$$IT_d: IT = \sin\left(\frac{\pi}{2}DT\right) \cdot \sin\left(\frac{\pi}{2}D\right)$$

$$IR = \sin\left(\frac{\pi}{2}R\right) \cdot \sin\left(\frac{\pi}{2}D\right)$$

$$IT_e: IT = \frac{DT+D}{2}$$

$$IR = \frac{R+D}{2}$$

IT_f : no considera ninguna estimación

Estas parecen ser las alternativas más triviales de acuerdo a las consideraciones y características del modelo.

Las alternativas IT_a y IT_e combinan la confianza/reputación con la similitud de acuerdo a la multiplicación y media aritmética de dichos valores, respectivamente. A pesar de ser considerados, se piensa de antemano que estas variantes no mejoran el nivel de satisfacción producido por la alternativa IT_f , que no considera ninguna aproximación. IT_a produce valores de confianza y reputación aproximadas mucho menores que los valores que combina. Para valores altos de confianza/reputación y similitud, puede obtener valores aproximados promedio. Por ello, se ha pensado en una transformación previa a los valores que se combinan, de acuerdo a las intenciones del modelo.

Por otro lado, la alternativa IT_e no parece ser la más idónea para aproximar la confianza y la reputación: no tiene sentido considerar el punto medio entre dos valores que representan magnitudes diferentes.

El resto de las alternativas incorporan la función *sin* para amplificar el valor de la similitud, de la confianza/reputación, o de su combinación. En la mayoría de los casos, el valor de la estimación debe ser lo más cercano posible al valor original de confianza o reputación, por cuanto la tarea conocida utilizada para la aproximación de estos valores posee un alto grado de similitud con la tarea desconocida. En este sentido, la variante más adecuada parece ser IT_c .

5.8. Funciones dependientes del dominio de aplicación

Existen dos funciones en este modelo cuya definición depende directamente de la representación de las especificaciones de las tareas y las respuestas de los servicios: la calidad de una respuesta, dada la tarea que satisface, y la similitud entre dos tareas.

Tal como se propone en el apartado 4.4, al analizar la relación entre la representación que brinda WSMO y el enfoque de confianza y reputación, TRSIM utiliza los conceptos de WSMO para representar algunos elementos dependientes del dominio de aplicación tales como las especificaciones de tareas (que representan los requerimientos de los usuarios) y las respuestas de los servicios (que describen las soluciones a dichas tareas). Las especificaciones de tareas son representadas por medio del concepto de *meta* y las respuestas de los servicios por el de *servicio web*, de manera que cada elemento es descrito por un conjunto de atributos no funcionales definidos por WSMO. Sin embargo, se pueden incorporar otras propiedades definidas por usuario de acuerdo al dominio de aplicación tratado, gracias a la simplicidad de su representación.

Siguiendo las ideas de la medida de Zadeh [149] (descrita en el apartado 4.3.2), para cada atributo de las especificaciones de tareas o las soluciones del servicio, el modelo define una función de normalización que hace independiente los valores manejados por el modelo, de los del dominio de las propiedades en el mundo real. El modelo utiliza los valores de esta función de normalización para representar la relevancia o conveniencia de un determinado atributo, independientemente de su dominio original. Tanto los atributos no funcionales definidos por WSMO, como los específicos del dominio de aplicación, son transformados utilizando funciones de normalización que representan su relevancia o conveniencia. Para cada atributo b_i se definen las funciones de conveniencia $v_{s_k}(b_i)$ y $v_{w_j}(b_i)$ para representar la conveniencia del atributo b_i para la especificación de tarea s_k y la solución w_j , respectivamente. Los valores de estas funciones son definidos en el intervalo $[0,1]$, de forma que un valor cercano a 0 indica un valor no relevante en la propiedad, mientras que los valores cercanos a 1 indican valores altamente relevantes (consultar los apartados 4.2.2.2 y 4.4).

Tratando de ofrecer un ejemplo que permita ilustrar lo comentado anteriormente, se puede considerar que, en un entorno proveedor de recursos multimedia, los servicios utilizan en su descripción propiedades tales como la velocidad de descarga (representado por el atributo *download speed*, definido por el usuario) y el número de errores generados en un intervalo de tiempo (representado por el atributo *accuracy* definido por WSMO). De esta forma, cuando la velocidad de descarga que ofrece un servicio es muy alta, el valor normalizado del atributo *download speed* es cercano a 1. En cambio, si el número de errores generados en un intervalo de tiempo es elevado, el valor de la atributo *accuracy* está cercano a 0.

De esta forma, la definición de estas dos funciones dependientes del dominio de aplicación (calidad y similitud) se realiza sobre la base de que los conceptos que utilizan son descritos mediante los elementos que ofrece la representación ontológica de WSMO. Por ello, ambas funciones pueden definirse de manera general como la comparación semántica

de elementos de WSMO, sin considerar los detalles del dominio de aplicación al que hacen referencia las especificaciones de tareas y las respuestas de los servicios. Tal como se comenta anteriormente, cada elemento que los representa, *meta* o *servicioweb*, está descrito por un conjunto de atributos no funcionales (definido por WSMO o por el usuario) que indican la relevancia de los atributos para el concepto. De esta forma, las funciones de calidad y similitud pueden quedar definidas sin tener que analizar el significado semántico de cada atributo, sino únicamente los valores de conveniencia de los atributos de los conceptos que se comparan en cada función. Este es uno de las características distintivas de TRSIM.

A continuación se comentan las principales características de cada una de estas funciones, a la vez que se ofrecen diferentes alternativas para su definición.

5.8.1. Calidad de la respuesta a una tarea

El modelo utiliza dos funciones para evaluar la satisfacción del agente iniciador de la interacción mediante el cumplimiento de los acuerdos de calidad prometidos y la calidad real de la solución obtenida a partir de la especificación de la tarea a resolver.

Se considera que en el proceso de interacción, el agente respondedor a_j se compromete, ante el agente iniciador a_i , a garantizar una calidad mínima al ofrecer la respuesta w_j a la tarea s_k . El valor de calidad acordado entre los agentes a_i y a_j para la realización de la tarea s_k es denotado por $ec_{i,j,k}$.

Básicamente, el valor de esta función es el resultado de la comparación entre el valor del acuerdo de calidad $ec_{i,j,k}$, para la solución w_j a la tarea s_k , y el valor real de la calidad de la solución, denotado por $Q(w_j, s_k)$. Para determinar el cumplimiento del acuerdo de calidad, se define la función P :

$$P(ec_{i,j,k}, w_j, s_k) = \begin{cases} 1 & : Q(w_j, s_k) \geq ec_{i,j,k} \\ 1 - (ec_{i,j,k} - Q(w_j, s_k)) & : Q(w_j, s_k) < ec_{i,j,k} \end{cases}$$

El valor de la función $P(ec_{i,j,k}, w_j, s_k)$ representa el grado de cumplimiento del acuerdo de calidad ($ec_{i,j,k}$), comparando este valor con la calidad de la respuesta w_j como solución a la tarea s_k . Si el valor de calidad real sobrepasa el valor prometido, la función devuelve 1, en caso contrario, es un indicador de la diferencia entre el valor prometido y el real.

La calidad de la solución, denotada por $Q(w_j, s_k)$, indica en que medida la respuesta w_j satisface los requerimientos especificados en la tarea s_k . El cálculo de este valor esta basado en la comparación de ambos conceptos.

De acuerdo a la conveniencia de los valores para cada propiedad o atributo, y siguiendo la recomendación WSMO, se pueden definir los atributos más relevantes de un determinado concepto [140]. Para cada tarea (s_k) o respuesta (w_j), se pueden agrupar los atributos relevantes en los conjuntos R_g y R_w , respectivamente (consultar el apartado 4.2.2.2) .

Si R_u es el conjunto de propiedades utilizadas para definir una tarea y su respuesta, entonces $R_g, R_w \subset R_u$ están formados, de acuerdo al valor de cada propiedad, por las propiedades más relevantes de la tarea y su respuesta, respectivamente. Para construir

estos conjuntos, se considera que cada atributo b_i de s_k (que puede escribirse $s_k.b_i$) es un atributo relevante y por lo tanto $b_i \in R_g$ si $s_k.b_i \geq \lambda_i$ (donde λ_i es un valor de umbral dependiente del dominio de aplicación). De igual forma, un atributo b_i de w_j es un atributo relevante y $b_i \in R_w$ si $w_j.b_i \geq \lambda_i$.

5.8.1.1. Alternativas para determinar la calidad de una respuesta

Mediante la identificación de estos tres conjuntos R_g , R_w y R_u pueden definirse varias alternativas para obtener la calidad de la respuesta w_j a partir de la tarea s_k . Entre ellas destacan las tres siguientes:

Q_a : siguiendo el enfoque ofrecido por el proceso de descubrimiento de servicios web dado por WSMO

$$Q(w_j, s_k) = \begin{cases} 1 & : R_g = R_w \\ q_3 & : R_g \subseteq R_w \\ q_2 & : R_g \supseteq R_w \\ q_1 & : R_g \cap R_w \neq \emptyset \\ 0 & : R_g \cap R_w = \emptyset \end{cases}$$

donde $0 \leq q_1 \leq q_2 \leq q_3 \leq 1$. De acuerdo a la definición de la función de satisfacción $P(ec_{i,j,k}, w_j, s_k)$, el máximo grado de satisfacción se obtiene cuando todos los atributos importantes de la tarea (agrupados en R_g) coinciden con todos los atributos importantes en la respuesta w_j (agrupados en R_w). Por el contrario, la peor satisfacción se produce cuando ninguno de los atributos importantes de la tarea s_k son satisfechos por los atributos importantes de la respuesta w_j . Esta función también considera algunos casos intermedios.

En esta alternativa, el proceso de descubrimiento de WSMO, basado en la descripción semántica simple de los servicios web, actúa como una función que indica el grado de coincidencia entre el servicio web (respuesta w_j) y la meta deseada (tarea s_k), teniendo en cuenta la coincidencia entre los atributos relevantes de ambos conceptos.

Q_b : considerando cuántos atributos de la especificación de tarea son satisfechos por la respuesta

$$Q(w_j, s_k) = \sin \left(\frac{\pi}{2} \cdot \frac{|Q'_{w_j, s_k}|}{|R_u|} \right)$$

donde R_u es el conjunto de todos los atributos de las tareas o respuestas, $Q'_{w_j, s_k} \subset R_u$ es el conjunto de estas propiedades tales que sus valores en la tarea s_k son menos restrictivos que los valores en la respuesta w_j :

$$Q'_{w_j, s_k} = \{b_i | b_i \in R_u, v_{s_k}(b_i) \leq v_{w_j}(b_i)\}$$

Se recuerda que $v_{s_k}(b_i)$ y $v_{w_j}(b_i)$ representan la conveniencia del atributo b_i para la especificación de tarea s_k y la solución w_j , respectivamente.

Esta función de satisfacción funciona como una razón entre los atributos satisfechos de la tarea y el número total de atributos de cualquier tarea o respuesta, agrupados en el conjunto R_u . El máximo grado de satisfacción es obtenido cuando todos los atributos deseados en la tarea s_k (no solo los más relevantes) son satisfechos por la respuesta w_j . Por el contrario, el peor grado de satisfacción se obtiene cuando ningún atributo de la especificación de tarea s_k es satisfecho por los atributos de la respuesta w_j .

Q_c : considerando cuántos atributos relevantes de la especificación de tarea son satisfechos por la respuesta

$$Q(w_j, s_k) = \sin \left(\frac{\pi}{2} \cdot \frac{|Q''_{w_j, s_k}|}{|R_g|} \right)$$

donde R_g es el conjunto de los atributos relevantes de la especificación de tarea s_k , $Q''_{w_j, s_k} \subset R_g$ es el conjunto de aquellas propiedades tales que sus valores en la especificación de tarea s_k son menos restrictivos que en la respuesta w_j , si la propiedad es considerada relevante en la tarea s_k :

$$Q''_{w_j, s_k} = \{b_i | b_i \in R_g, v_{s_k}(b_i) \leq v_{w_j}(b_i)\}$$

Esta alternativa es similar a la anterior. Pero, en este caso, la satisfacción es una razón entre los atributos importantes de la especificación de tarea s_k que son satisfechos y el número de atributos importantes de la especificación de tarea s_k , contenidos en R_g . El valor máximo de satisfacción se obtiene cuando todos los atributos relevantes, deseados en la especificación de tarea s_k , son satisfechos por la respuesta w_j . Por el contrario, la peor satisfacción se produce cuando ningún atributo relevante de la tarea s_k es satisfecho por los atributos de la respuesta w_j . Esta alternativa es un caso intermedio entre las dos anteriores Q_a y Q_b .

A priori, puede pensarse que Q_a no produce los mejores valores de satisfacción dado que solo considera un escaso rango de valores de satisfacción $(0, q_1, q_2, q_3, 1)$, haciendo muy limitado el poder de expresividad de esta función. La carencia de expresividad de esta alternativa se hace más acentuada en tanto mayor sea el número total de atributos b_i que describen a la especificación de tareas s_k y la respuesta del servicio web w_j . Los resultados dependen del tipo de coincidencia entre la especificación de tarea y la respuesta, siguiendo las ideas del proceso de descubrimiento de servicios web dado por WSMO. Considera la satisfacción de grupos de atributos importantes como un conjunto, no tiene en cuenta la satisfacción de los atributos de manera independiente.

Frente a las limitaciones del enfoque seguido por la alternativa Q_a , se propone la alternativa Q_b que produce un amplio espectro de valores, con mayor expresividad. Q_b considera la cantidad de atributos de la tarea que son satisfechos por la respuesta, sin tener en cuenta la relevancia de los atributos. Esta alternativa debe ofrecer mejores valores de calidad que la anterior, por cuanto es capaz de representar de mejor manera la gran variedad de casos que tienen lugar en la comparación de tareas y respuestas.

Por otra parte, Q_c es un caso intermedio entre Q_a y Q_b que tiene en cuenta la satisfacción de cada atributo de la tarea (tal como Q_b), pero solo de los atributos relevantes (tal como Q_a).

El apartado 7.2.1 ofrece un conjunto de evidencias experimentales que permiten la comparación del comportamiento del modelo para cada una de las tres alternativas consideradas.

5.8.2. Similitud entre tareas

La determinación del grado de similitud entre dos especificaciones de tarea es una función cuya definición depende de la representación de las tareas, de acuerdo al dominio de aplicación. Tal como se muestra en la figura 5.1 esta función es utilizada fundamentalmente para estimar la confianza y la reputación asociadas a un agente del que se desconoce su desempeño dando solución a una tarea determinada. En estos casos, se recurre a la información almacenada sobre el desempeño del agente realizando tareas similares.

La función que propone TRSIM para obtener el grado de similitud entre dos tareas tiene la siguiente definición algebraica:

$$D : S \times S \rightarrow [0, 1]$$

donde S es el conjunto de todas las especificaciones de tareas posibles.

A continuación se plantean algunas alternativas para la definición de la función de similitud D .

5.8.2.1. Alternativas para determinar la similitud entre tareas

Las diferentes alternativas para definir la función D se basan fundamentalmente en la comparación de los atributos de cada tarea, representada por el concepto de *meta* de WSMO. Estas alternativas están relacionadas con las ideas recopiladas de la literatura, algunas de ellas tratadas en el apartado 4.3. Entre las más importantes destacan las siguientes:

D_a : Basada en el Coeficiente de Pearson

El coeficiente de correlación de Pearson se usa para obtener una medida de la similitud entre dos elementos de un mismo concepto, descrito por un conjunto de atributos (consultar el apartado 4.3.1). Esta medida puede ser utilizada para calcular el grado de similitud entre dos tareas, ambas representadas mediante el concepto de *meta* de WSMO. Al instanciar el coeficiente de correlación de Pearson, para establecer una medida de la similitud entre dos tareas s_k y s_p (cada una descrita por i atributos: s_{ki} y s_{pi} , respectivamente), se puede plantear la función de similitud D de la siguiente manera:

$$D(s_k, s_p) = \frac{\sum_i (s_{ki} - \bar{s}_k) \cdot (s_{pi} - \bar{s}_p)}{\sqrt{\sum_i (s_{ki} - \bar{s}_k)^2 \cdot \sum_i (s_{pi} - \bar{s}_p)^2}}$$

donde \overline{s}_k y \overline{s}_p son los valores medios de los atributos de las tareas s_k y s_p , respectivamente.

D_b : Basada en la medida de Tversky

La medida de Tversky [127], basada en la teoría de conjuntos, es un indicador de la similitud semántica entre entidades que pueden estar descritas utilizando, incluso, diferentes ontologías. En el modelo, ambas tareas, cuyo grado de similitud se desea conocer, están representadas mediante el concepto de *meta* de WSMO.

Considerando que para TRSIM la medida de similitud debe ser simétrica, y que ambas tareas están descritas mediante el mismo concepto de la misma ontología ofrecida por WSMO, el parámetro α queda fijado en el valor $\alpha = 0,5$ (consultar el apartado 4.3.3).

Por otra parte, teniendo en cuenta que la similitud se obtiene a partir de la relación entre los conjuntos de atributos relevantes de las dos tareas a comparar s_k y s_p , la expresión de la medida Tversky puede quedar definida de la siguiente manera:

$$D(s_k, s_p) = \frac{|R_{gk} \cap R_{gp}|}{|R_{gk} \cap R_{gp}| + 0,5|R_{gk} \setminus R_{gp}| + 0,5|R_{gp} \setminus R_{gk}|}$$

donde los conjuntos R_{gk} y R_{gp} contienen los atributos relevantes de las tareas s_k y s_p , respectivamente.

D_c : Basada en la comparación léxica

La similitud entre dos tareas puede obtenerse a partir de la comparación de las representaciones de ambas, utilizando el enfoque basado en el comparación léxica de cadenas de caracteres, descrito en el apartado 4.3.4. Generalizando este enfoque, se puede pensar que cada cadena corresponde a cada tarea a comparar, y que los caracteres están representados por los atributos de la tarea. Así, puede plantearse la medida de similitud de la siguiente manera:

$$D(s_k, s_p) = \max \left(0, \frac{\min(|R_{gk}|, |R_{gp}|) - ed(s_k, s_p)}{\min(|R_{gk}|, |R_{gp}|)} \right)$$

donde R_{gk} y R_{gp} son los conjuntos de atributos relevantes de las tareas s_k y s_p , respectivamente. Por su parte, $ed(s_k, s_p)$ está dada por la cantidad de atributos relevantes que son diferentes en ambos conceptos, pudiendo escribirse de la siguiente manera:

$$ed(s_k, s_p) = |R_u \setminus (R_{gk} \cap R_{gp})|$$

donde, según la notación utilizada por WSMO en el apartado 4.2.2, R_u representa el conjunto de las propiedades utilizadas para la definición de los conceptos.

D_d : Basada en la media de las diferencias entre los atributos

La similitud entre dos tareas s_k y s_p puede ser obtenida a partir de la media de la diferencia absoluta entre los atributos de cada tarea:

$$D(s_k, s_p) = 1 - \frac{1}{n} \cdot \sum_{i=1}^n |s_{k_i} - s_{p_i}|$$

donde s_{k_i} es el i -ésimo atributo de la tarea s_k , y s_{p_i} es el i -ésimo atributo de la tarea s_p ; y $n = |R_u|$ es el número de atributos del concepto de tarea considerado en el modelo a partir de los atributos no funcionales descritos por WSMO y los definidos por el usuario, específicos del dominio de aplicación tratado.

D_e : Basada en la distancia euclídea

De manera análoga a la anterior, la similitud puede ser definida a partir de la distancia euclídea entre los dos conceptos:

$$D(s_k, s_p) = 1 - \sqrt{\sum_{i=1}^n (s_{k_i} - s_{p_i})^2}$$

La mayoría de los experimentos que se muestran a lo largo del capítulo 7 utilizan la alternativa D_d para calcular la similitud entre dos tareas. Sin embargo, en el apartado 7.2.2.2 se ofrecen evidencias experimentales que permiten comparar algunos elementos del comportamiento del modelo para diferentes formas de calcular la similitud.

5.9. Actualización de las bases de experiencias

Una vez que cada agente obtiene y evalúa la respuesta que ofrece el agente recomendado por el modelo de confianza y reputación, se debe actualizar el conocimiento de forma que pueda ser utilizado eficazmente en futuras interacciones con sus vecinos. El modelo supone que las interacciones entre los agentes se repiten varias veces, estableciendo el concepto de ronda como la unidad de tiempo que fija un orden secuencial en el diálogo de los agentes en una determinada interacción. Para TRSIM, una ronda comienza cuando un agente consumidor debe satisfacer sus requerimientos de usuario (descritas en la especificación de tarea s_k) y utiliza conceptos de confianza y reputación para seleccionar el vecino que ofrece la respuesta más adecuada. Cada ronda finaliza cuando el agente consumidor actualiza el conocimiento almacenado en las bases de experiencias a partir de la evaluación de la respuesta obtenida w_j .

El modelo ofrece los mecanismos para que al finalizar cada ronda, se puedan actualizar las dos bases de experiencias a partir de la información generada. Para actualizar las bases de experiencias, el agente iniciador a_i evalúa la interacción de la ronda t (utilizando, entre otras, la función de calidad Q , descrita en el apartado 5.8.1), teniendo en cuenta la solución w_j ofrecida por el agente respondedor a_j como respuesta a la especificación de tarea s_k .

La información relativa a cada interacción t , que el agente a_i lleva a cabo para satisfacer la tarea s_k , puede ser agrupada en el conjunto:

$$I^{(t)}(a_i, s_k) = \{(a_j, w_j) | a_j \in C^{(t)}(a_i, s_k), w_j \in W\},$$

donde w_j es la respuesta ofrecida por el agente a_j en dicha interacción; W es el conjunto de todas las posibles respuestas; y $C^{(t)}(a_i, s_k)$ denota el conjunto de los agentes a los cuales se les ha solicitado una solución para la tarea s_k en la interacción t . El apartado 5.2.1 comenta algunas estrategias para la formación de este grupo de agentes.

Para actualizar la base de experiencias $IET_i^{(t)}$, para cada agente a_j , que ofrece la solución w_j a la tarea s_k , el agente a_i tiene una nueva experiencia:

$$edt_{i,j,k} = (a_i, a_j, s_k, et_{i,j,k})$$

con

$$(a_j, w_j) \in I^{(t)}(a_i, s_k)$$

donde el valor $et_{i,j,k}$ es una medida obtenida a partir de la calidad de la solución (obtenido mediante la función de calidad Q) y el cumplimiento del compromiso de satisfacción acordado $ec_{i,j,k}$ (obtenido mediante la función de satisfacción P):

$$et_{i,j,k} = Q(w_j, s_k) \cdot P(ec_{i,j,k}, w_j, s_k)$$

De esta forma, el modelo evita que un agente a_j , con un bajo valor de satisfacción comprometida $ec_{i,j,k}$ y con un valor medio de calidad de la solución para la tarea s_k , pueda obtener un elevado valor de satisfacción $et_{i,j,k}$. El grado de satisfacción debe ser la combinación de la calidad de la solución dada la especificación de la tarea (Q) y el cumplimiento de los acuerdos de calidad acordados (P). La definición de las funciones P y Q se ofrece en el apartado 5.8.1.

Teóricamente, el modelo puede incorporar todas las nuevas experiencias $edt_{i,j,k}$ que se generen, sin tener que analizar cuantas experiencias existen en la base $IET_i^{(t)}$ asociadas al agente a_j para la tarea s_k . Sin embargo, pensando en la implementación y eficiencia computacional del modelo se propone limitar el número total de experiencias asociadas a un agente a_j en la realización de la tarea s_k a un valor igual al parámetro itm , por cuanto un número mayor de experiencias no mejoraría la precisión de las medidas de confianza (consultar el apartado 5.5).

Por otra parte, la base de experiencias para la reputación $IER_i^{(t)}$ está formada por valores de reputación únicos $er_{i,j,k}$ para indicar, de acuerdo a la experiencia del agente a_i , la confiabilidad del agente a_j para ofrecer información sobre el desempeño de otros agentes ante la tarea s_k . Para actualizar este valor, el modelo propone un enfoque propio, que tiene en cuenta la variación, que se produce durante la interacción, en la confianza sobre los agentes que son recomendados por el agente a_j .

El conjunto de los agentes recomendados al agente a_i por el agente a_j para la tarea s_k en cada ronda t es denotado por $M_j^{(t)}(a_i, s_k)$. Así, para cada agente recomendado $a_r \in$

$M_j^{(t)}(a_i, s_k)$, el modelo obtiene el valor de confianza que a_i tenía sobre a_r antes de la interacción, denotado por $f_{i,r,k}^{(t-1)}$, y el nuevo valor de confianza al finalizar la interacción, denotado por $f_{i,r,k}^{(t)}$. El valor de confiabilidad, del agente a_j para ofrecer opiniones sobre otros, al inicio de la interacción $er_{i,j,k}^{(t-1)}$ es modificado combinándolo adecuadamente con la media de las diferencias entre los valores de confianza inicial y final de cada agente a_r :

$$er_{i,j,k}^{(t)} = \begin{cases} er_{i,j,k}^{(t-1)} + (1 - er_{i,j,k}^{(t-1)}) \cdot \Delta & : \Delta \geq 0 \\ er_{i,j,k}^{(t-1)} \cdot (1 + \Delta) & : \Delta < 0 \end{cases}$$

donde Δ es la media de todas las diferencias entre los valores de confianza final e inicial:

$$\Delta = \frac{\sum_{a_r \in M_j^{(t)}(a_i, s_k)} f_{i,r,k}^{(t)} - f_{i,r,k}^{(t-1)}}{|M_j^{(t)}(a_i, s_k)|}.$$

De acuerdo al comportamiento de esta función, el valor final de reputación $er_{i,j,k}^{(t)}$ es mejor que el valor inicial $er_{i,j,k}^{(0)}$ cuando la confianza en los agentes recomendados por a_j , reunidos en el conjunto $M_j^{(t)}(a_i, s_k)$, es mejorada, de manera global, durante la interacción t . Es decir, la confiabilidad de un agente como recomendador de otros es incrementada o decrementada en el mismo sentido que la variación de la confianza en los agentes que recomienda. La confiabilidad de un agente ofreciendo información de reputación aumenta cuando la media de las diferencias de los valores de confianza es positiva. Si la confianza en los agentes que se recomiendan disminuye, la confianza del agente como recomendador es disminuida también.

5.10. Conclusiones

Se ha propuesto, en este capítulo, un modelo para el tratamiento de la confianza y la reputación para ser utilizado en un entorno basado en la interacción entre pares, donde los agentes son capaces de encontrar el servicio cuyo comportamiento es el adecuado para resolver los requerimientos planteados por los usuarios. El modelo emplea la representación ontológica ofrecida por la recomendación WSMO para describir, en términos de atributos funcionales y no funcionales, las especificación de las tareas (que reflejan los requerimientos de los usuarios) y sus respuestas (que determinan el servicio web que los satisface).

El modelo considera que la confianza y la reputación son conceptos referidos a las preferencias y requerimientos de los usuarios, dados en cada momento por la especificación de tarea que se desea resolver. Así, ofrece los mecanismos para obtener los valores relacionados con estos conceptos a partir del propio conocimiento del agente o el intercambio de información. Destaca el hecho que si para alguna especificación de preferencias o requerimientos se desconocen los valores de confianza y reputación asociados, el sistema ofrece mecanismos para estimar estos valores a partir de una especificación similar bien conocida.

Para la descripción de los servicios y los requerimientos de los usuarios, el modelo utiliza los conceptos dados por WSMO: *servicio web* y *meta*. Propone utilizar, para la descripción de cada concepto, un grupo de atributos no funcionales definidos por la propuesta de estándar. Sin embargo, se pueden incorporar otros atributos definidos por el usuario propios del dominio de aplicación. Además, mediante el uso de WSMO, se facilita la definición de algunas características y funciones del modelo que son dependientes del dominio de aplicación, tales como la calidad de una respuesta dada la tarea a resolver, o la similitud entre dos tareas.

Es válido destacar que existe un conjunto de elementos que deben ser establecidos por cada agente antes de que el modelo pueda ser utilizado. Entre los más importantes se encuentran:

- la función Q para determinar la calidad de la solución a partir de los requerimientos de usuario que satisface,
- la función D para determinar la similitud entre dos tareas que representan los requerimientos de usuario,
- el operador \oplus que determina la forma en que se combinan la confianza y la reputación con la similitud entre dos tareas al estimar los valores asociados a tareas que representan requerimientos desconocidos (funciones IT e IR),
- varios parámetros de configuración, tales como:
 - γ , tal que p_γ es el percentil utilizado para seleccionar del conjunto de valores de reputación $R^{(t)}(a_i, a_j, s_k, CR_{sup}^{(t)}(a_i, s_k))$ los agentes $a_j \in CT_{dud}^{(t)}(a_i, s_k)$ que puedan ofrecer una solución para la especificación de tarea s_k , planteada por el agente a_i .
 - δ , utilizado para reflejar la pérdida de relevancia de las experiencias con el transcurso del tiempo,
 - itm , utilizado para definir la cantidad de experiencias en la base de experiencias para la confianza (IET),
 - los valores de umbrales λ_i para analizar la relevancia de los i atributos de las especificaciones de tareas y respuestas. Se utilizan para seleccionar los atributos relevantes de cada especificación de tarea s_k o respuesta w_j , tomando aquellos tales que $s_k.b_i \geq \lambda_i$ o $w_j.b_i \geq \lambda_i$, respectivamente.

En el capítulo 7 se han fijado los valores de algunos de estos parámetros (por ejemplo, γ , δ , itm) para analizar el comportamiento del modelo ante diferentes alternativas para la selección de la función de calidad Q , el operador \oplus utilizado en las funciones de estimación de la confianza IT y la reputación IR para tareas desconocidas, y la forma de calcular la similitud D entre dos tareas. El establecimiento de los valores de los umbrales λ_i no han sido considerados por cuanto no se han utilizado los valores reales de cada atributos, sino se ha usado directamente su relevancia: $s_k.b_i$, para la tarea s_k , o $w_j.b_i$ para una respuesta w_j .

También se realizan algunos experimentos que estudian la capacidad de adaptación de TRSIM a algunos cambios concretos en el comportamiento de los agentes, de manera individual o grupal.

5.11. Lógica de un agente para determinar la confianza en otro

A continuación, en la figura 5.4, se presenta el algoritmo, seguido por el agente a_i , que desea resolver una tarea s_k , para determinar la confianza en otro agente a_j que ofrece respuestas.

Para determinar la confianza que se tiene en un agente a_j , para ofrecer soluciones a la especificación de tarea s_k , se deben obtener los valores de confianza directa DT , de reputación R y de precisión de la medida de confianza directa $DTRL$.

Para ello, primeramente, el agente inspecciona la base de experiencias para la confianza IET , buscando experiencias relacionadas con el desempeño de a_j asociadas a la realización de la tarea s_k . Si las experiencias están disponibles, se agregan según la propuesta del modelo, y se obtienen los valores de confianza directa DT y precisión de la confianza directa $DTRL$. Si las experiencias, asociadas a s_k , no son suficientes se analiza el comportamiento del agente para la realización de la especificación de tarea s_p más parecida a s_k . Así, se obtiene un valor aproximado de la confianza directa para la realización de s_k , a partir de la combinación de la confianza directa en la especificación de tarea similar s_p y el grado de similitud entre ambas tareas. Si tampoco se tienen experiencias para otra especificación de tarea similar (es decir, no se conoce el desempeño del agente para ninguna especificación de tarea) se le asigna, por defecto, valores de confianza directa y precisión previamente definidos por el agente que implementa el modelo. Estos valores representan la confianza directa que un agente otorga a quienes le son desconocidos por completo y la relevancia que se le da a la confianza directa como fuente de información.

Luego, el agente a_i selecciona el grupo de agentes *recomendadores* a los cuales solicita información sobre el comportamiento del agente a_j para la realización de la tarea s_k . Para determinar los valores de reputación R se combinan las opiniones de cada agente recomendador con la confianza que a_i le otorga a cada uno para ofrecer opiniones acerca del comportamiento de otros. Sin embargo, si la información contenida en la base de experiencias para la reputación IER no es suficiente para seleccionar los agentes recomendadores, el agente utiliza información relacionada con otra especificación de tarea s_p similar a la especificación de tarea s_k . De esta manera, selecciona los agentes recomendadores del comportamiento del agente a_j para la realización de la especificación de tarea similar s_p , combinando las opiniones ofrecidas con el grado de similitud entre ambas especificaciones de tareas. Por otra parte, si la información de la base IER tampoco contienen experiencias para otra especificación de tarea similar (es decir, no se conoce la calidad de los agentes como recomendadores de otros para ninguna especificación de tarea) se le asigna, por defecto, un valor de reputación previamente definido por el agente que

- Procedimiento *ObtenerConfianza*
 1. *ObtenerConfianzaDirectaYPrecisión(dt, dtrl)*
 2. *ObtenerReputación(r)*

$$Confianza = dtrl \cdot dt + (1 - dtrl) \cdot r$$
 3. Devolver *Confianza*

- Procedimiento *ObtenerConfianzaDirectaYPrecisión(dt, dtrl)*
 1. Si la información de $IET(a_i, s_k)$ es suficiente
 - Obtener la confianza directa y su precisión para s_k :

$$dt = DT(a_i, a_j, s_k)$$

$$dtrl = DTRL(a_i, a_j, s_k)$$
 2. si no es suficiente
 - Buscar la tarea s_p más similar a s_k
 - Si existe la tarea s_p
 - * Obtener la confianza directa para s_p y combinarla con la similitud:

$$dt = DT(a_i, a_j, s_p) \oplus D(s_k, s_p)$$
 - * Obtener la precisión de la confianza directa para s_p :

$$dtrl = DTRL(a_i, a_j, s_p)$$
 - Si no
 - * Asignar a dt y $dtrl$ valores predeterminados
 3. Devolver $dt, dtrl$

- Procedimiento *ObtenerReputación(r)*
 1. Si la información de $IER(a_i, s_k)$ es suficiente
 - Obtener la lista de agentes recomendadores:

$$recomendadores = CR_{sup}(a_i, s_k)$$
 - Obtener la reputación ofrecida a partir de los recomendadores

$$r = R(a_i, a_j, s_k, recomendadores)$$
 2. si no es suficiente
 - Buscar la tarea s_p más similar a s_k
 - Si existe la tarea s_p
 - * Obtener la lista de recomendadores para la tarea similar s_p :

$$recomendadores = CR_{sup}(a_i, s_p)$$
 - * Obtener la reputación para la tarea s_p y combinarla con la similitud:

$$r = R(a_i, a_j, s_p, recomendadores) \oplus D(s_k, s_p)$$
 - Si no
 - * Asignar a r un valor predeterminado
 3. Devolver r

Figura 5.4: Algoritmo para determinar la confianza en un agente

5.11. LÓGICA DE UN AGENTE PARA DETERMINAR LA CONFIANZA EN OTRO

implementa el modelo. Este valor representa el valor de reputación que un agente otorga a los agentes de quienes no puede obtener referencias.

Finalmente, los valores de confianza directa DT y reputación R (independientemente de la forma en que se obtienen) son combinados teniendo en cuenta la precisión de la medida de confianza directa $DTRL$.

Capítulo 6

Escenarios de referencia

Los modelos de confianza y reputación son útiles al asistir en las decisiones de los agentes sobre sus interacciones con otros. Mediante dichos modelos, se puede estimar o predecir el comportamiento de los agentes, por ejemplo, ante una determinada situación o consecución de determinados objetivos. Su funcionamiento se basa en la gestión eficiente de las propias experiencias del agente, almacenadas durante interacciones anteriores, la información suministrada por otros, la información que se pueda obtener a partir de las relaciones que se mantienen entre todos los agentes, las características del entorno, entre otros. (El apartado 2.1.1 detalla la forma de utilización de cada una de las fuentes de información de confianza y reputación.)

La información suministrada por las diferentes fuentes puede estar disponible y estructurada de disímiles maneras. Las características específicas de cada fuente de información dependen en gran medida del dominio de aplicación tratado y de los modelos de representación de los agentes vecinos, de las interacciones entre ellos, del entorno, de los requerimientos de los usuarios, etc. Esta línea de razonamiento sugiere que la aplicabilidad de un modelo a un determinado problema práctico depende, entre otros, de la compatibilidad entre las características del modelo y del escenario determinado por el problema, así como de la utilidad que pueda aportar el modelo a su solución. En este capítulo se describen las características de dos escenarios en los cuales puede resultar factible la utilización del modelo de confianza y reputación presentado en el capítulo anterior.

Además, se ofrecen algunas recomendaciones de uso, específicas de cada escenario, para facilitar la aplicación de este paradigma adaptativo. La definición del modelo, por sí sola no es trivial.

El resto del capítulo se estructura de la siguiente manera: la sección 6.1 comenta algunas consideraciones generales, comunes a ambos escenarios, relacionadas con la utilidad de un modelo de confianza y reputación al identificar los agentes contrapartes en una determinada interacción. En la sección 6.2 se definen las características de un escenario del tipo Consumidor - Proveedor de recursos que reúne todos los elementos generales para los cuales un modelo como TRSIM puede resultar útil. En este apartado se comenta el modelo de agentes adoptado, cómo se utiliza un marco ontológico para la representación

de los elementos dependientes del dominio de aplicación, cómo se seleccionan los agentes proveedores, entre otros. Luego, en la sección 6.3 se describe el escenario de prueba ART (presentado en la sección 3.2.5). Por último, en la sección 6.4 se comentan las principales características comunes entre ambos escenarios y se dan algunas ideas para realizar experimentos que permitan analizar la efectividad y comportamiento del modelo en ellos.

6.1. Identificación de los tipos de agentes utilizando la confianza y la reputación

Ante una determinada situación, cada agente puede mostrar comportamientos diferentes. Por ejemplo, de manera general, los agentes consumidores de recursos prefieren interactuar (negociar, cooperar, contratar, etc.) solo con los proveedores que exhiben los mejores comportamientos. Dado que en la mayoría de estos ambientes no existe una entidad capaz de ofrecer información acerca del comportamiento de los demás, cada agente debe crear y mantener actualizado su propio modelo de representación sobre el desempeño de sus vecinos. Así, ante determinada situación, podrá analizar con cuáles de ellos interactuar para resolver un determinado problema.

En este sentido, los modelos de confianza y reputación ofrecen una buena solución para representar y predecir el comportamiento del resto de agentes en el sistema. Los valores de confianza y reputación pueden ser un criterio para identificar los agentes de los que se espera un buen comportamiento, es decir, los que son capaces de ofrecer una solución de calidad para un determinado problema.

De esta forma, un modelo de confianza y reputación puede brindar mecanismos adaptativos para guiar las interacciones entre los agentes. De manera abstracta, un modelo de confianza y reputación en un sistema multi-agente puede ser visto como una capa de inteligencia auxiliar que incorporan todos los agentes y ofrece criterios que apoyan la toma de decisiones de los agentes en la selección de su contraparte en una determinada interacción.

Estos modelos, por lo general, combinan información de confianza y reputación de varios tipos. Los valores pueden estimarse a partir de diferentes fuentes de información, en dependencia del punto de vista analizado, el contexto o, simplemente, la disponibilidad de conocimiento. Cuando la información de estas fuentes es insuficiente para ofrecer algún criterio de confianza o reputación, de manera intuitiva se puede pensar en obtener una aproximación de estos valores. Para ello, se puede utilizar el conocimiento almacenado sobre el desempeño de otros agentes frente a requerimientos similares al deseado.

Por otra parte, la información muchas veces depende del dominio de aplicación al que se aplique el modelo de confianza y reputación. Se requiere un esquema ontológico común que facilite la representación y razonamiento a partir de la información recopilada por los agentes acerca del desempeño de otros. Dicho esquema debe garantizar (1) la representación de un modelo de los agentes en el sistema, (2) la representación de los requerimientos que desean satisfacer los agentes, (3) las soluciones que ofrecen, (4) la representación de las

herramientas para la comparación de los requerimientos, y (5) la posibilidad de establecer medidas de calidad y de satisfacción de los requerimientos a partir de una determinada solución, entre otros.

6.2. Escenario Consumidor - Proveedor de recursos

Quizás, los escenarios del tipo Consumidor - Proveedor de recursos (o servicios) son la representación más natural e intuitiva de una gran variedad de situaciones prácticas donde existe un conjunto de consumidores que necesitan satisfacer sus requerimientos a partir de las soluciones (productos) que ofrece un conjunto de proveedores. La sencillez y familiaridad de este escenario están determinadas por la existencia de diferentes relaciones de este tipo, tanto las que tienen lugar en las sociedades humanas como en los modelos de negocio implementados por muchos sistemas.

Desde el punto de vista de la inteligencia artificial distribuida, dicho escenario resume de manera sencilla algunas características comunes de una gran variedad de sistemas que pueden ser vistos como sistemas multi-agente en los cuales los componentes actúan de manera autónoma y flexible intentando alcanzar sus propios objetivos [74]. Estos sistemas, donde los agentes pueden ser consumidores y/o proveedores de información, recursos y servicios, se caracterizan, entre otros, por los siguientes aspectos:

- Cada agente se muestra interesado en sus propios objetivos o metas, muchas veces sin resultar benevolentes o confiables en sus relaciones sociales.
- Los agentes no poseen una representación exacta del entorno y de sus vecinos que le permita predecir con exactitud su comportamiento futuro.
- No existe una entidad central que brinde información relacionada con el desempeño de cada agente.
- La incertidumbre asociada a cada decisión de los agentes es muy elevada.

Por tales razones, para hacer frente a la falta de información o a la existencia de información incompleta o imprecisa sobre el comportamiento de otros agentes se necesitan mecanismos que posibiliten reducir los riesgos de las interacciones entre las diferentes partes. En este sentido, el uso de los modelos de confianza y reputación persigue aumentar la fiabilidad y desempeño de las comunidades electrónicas a partir de imitar algunos mecanismos de control presentes en las sociedades humanas [111].

En este escenario, el modelo de confianza y reputación puede ser considerado como el conjunto de herramientas de representación y decisión que cada agente mantiene y comparte convenientemente, para resolver el problema de encontrar la solución (o servicio) que mejor satisface los requerimientos de un consumidor determinado. Un esquema general del modelo planteado en el escenario puede ser el ofrecido en la figura 6.1.

Este esquema resume las características generales que se han comentado arriba. Existe un conjunto de agentes proveedores y consumidores de recursos donde cada proveedor

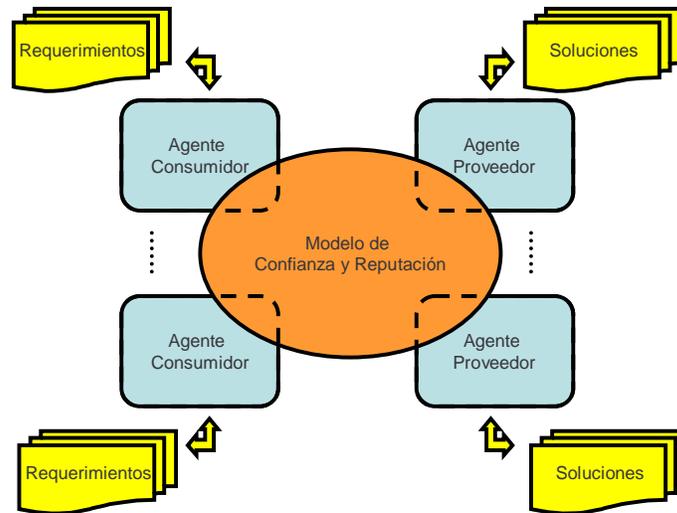


Figura 6.1: Esquema general del escenario Consumidor - Proveedor de recursos

posee un grupo de soluciones a los requerimientos que tienen los consumidores. El modelo de confianza y reputación ofrece a cada agente los mecanismos de representación de las capacidades del resto, así como garantiza el intercambio de información de confianza y reputación entre los agentes del sistema.

Las características del escenario imponen un conjunto de características al modelo de confianza y reputación que se utilice en él. Entre las más relevantes se pueden relacionar:

- los valores de confianza y reputación dependen del dominio de los requerimientos de los agentes proveedores,
- el conocimiento previo sobre el desempeño de un agente conocido, almacenado a partir de interacciones anteriores, es la fuente de confianza y reputación más fiable,
- la información de reputación se utiliza para intercambiar las apreciaciones que los agentes tienen sobre el desempeño de otros,
- se utilizan mecanismos de intuición que permiten la estimación de la confianza y la reputación asociada a agentes desconocidos,
- se distingue entre la confianza que ofrece un agente para solucionar un determinado requerimiento de la confianza que ofrece para brindar información acerca del desempeño de otros agentes,
- la información se mantiene privada y no existe una entidad centralizada que mantenga métricas sobre el comportamiento de los agentes.

Estas son precisamente las características más importantes de TRSIM, por lo que pudiera resultar factible su utilización en este escenario.

Esta sección continúa realizando algunas puntualizaciones en cuanto al modelado de los diferentes elementos que componen el escenario:

- los agentes (en el apartado 6.2.1 se comentan las características del modelo de agente considerado),
- las soluciones que ofrecen los agentes proveedores y los requerimientos de los agentes consumidores (el apartado 6.2.2 comenta la representación, mediante WSMO, de las respuestas que ofrecen de los agentes proveedores y los requerimientos de los consumidores),
- las interacciones entre los agentes (el apartado 6.2.3 comenta las principales características de las interacciones entre los agentes y la utilización del modelo de confianza en ellas).

6.2.1. Modelo de agentes

En este escenario, tanto los proveedores como los consumidores se representan mediante agentes. El modelo de agente que se propone contempla la existencia de dos dimensiones del agente fundamentalmente: la mediadora y la adaptativa. Estas dimensiones corresponden a requerimientos del escenario de referencia o a características propias de la arquitectura que se propone como solución. La dimensión mediadora está referida a las capacidades de intercambio de información de confianza y reputación entre los agentes del sistema. Mientras, la dimensión adaptativa hace referencia a las capacidades de aprendizaje del agente a partir de la evaluación de las interacciones y actualización del modelo de confianza y reputación incorporado.

A continuación se comentan algunos elementos importantes relacionados con cada una de las dos dimensiones de los agentes de este escenario.

6.2.1.1. Dimensión mediadora

La dimensión mediadora de los agentes en este escenario está dada por la ayuda que brindan los agentes a sus vecinos en la selección de los proveedores de servicios. Esta dimensión se modela a partir de las capacidades de representación e intercambio de la información de confianza y reputación entre los agentes.

Los agentes que, desde esta perspectiva, funcionan como intermediarios, gestionan las capacidades de los proveedores de servicios, adoptando un modelo de representación para ellos. Uno de los problemas que se encuentran bajo esta perspectiva del agente, es la heterogeneidad sintáctica y semántica de la información que se trata. Por tanto, para solucionar este problema se debe utilizar conocimiento ontológico ente los agentes, de manera que gestionen ontologías comunes de dominio general y específico [4].

También se exige que los agentes garanticen una determinada calidad del servicio a los clientes que requieren de sus prestaciones como mediador, dada fundamentalmente por criterios tales como la accesibilidad, disponibilidad de los servicios finales; la corrección,

completitud y consistencia de los datos; así como otros relacionados con su eficiencia y funcionalidad. Existen representaciones ontológicas, como la ofrecida por WSMO, que brindan elementos que, entre otras funciones, permiten la representación de estos criterios.

Un elemento muy importante en el uso de agentes mediadores es la confianza que estos inspiren a los agentes que utilizan sus servicios, así como la reputación que goce entre sus vecinos respecto al correcto desempeño de sus funciones. De manera general, y de forma bien intencionada, el agente debe garantizar el cumplimiento de sus compromisos y la veracidad del conocimiento que intercambia con otros, así como implementar mecanismos que lo hagan robusto, inmune a ataques malintencionados de otros.

6.2.1.2. Dimensión adaptativa

Los agentes deben ser capaces de aprender a partir de la información generada en el sistema con el objetivo de utilizar, posteriormente, el conocimiento aprendido. Para ello puede utilizar la información resultante de las interacciones directas con otros agentes, la información brindada por el entorno o por otros agentes y los que pueda obtener mediante procesos de intuición a partir de información o conocimiento similares.

Desde esta dimensión se deben garantizar los mecanismos internos de almacenamiento y gestión eficiente del conocimiento propio del agente, sobre su comportamiento, el de otros agentes o las características de su entorno. Si se almacena el conocimiento que genera una determinada interacción con otro agente, la eficiencia del agente puede mejorarse cuando se presenten situaciones similares.

Desde el punto de vista adaptativo, el agente tiene que saber analizar cuándo necesita información que otros agentes poseen y a quién pedirla. Debe ser capaz de intercambiar información con otros agentes con el fin de completar su conocimiento acerca de determinada situación. En este sentido, debe establecer mecanismos para evaluar la veracidad y calidad de la información obtenida, rechazando la información falsa, y gestionando aquella que resulte incompleta e imprecisa.

Cada agente tiene que ser capaz de estimar el conocimiento relacionado con la calidad de la solución que le ofrece otro, y con el valor de la información que recibe en sus interacciones. Dicho conocimiento debe ser incorporado en un modelo de representación interna que facilite los procesos de razonamiento y la obtención de medidas que contribuyan a la toma de decisiones.

Los agentes deben aprender las normas de conducta de los otros agentes y del sistema, de forma que sean capaces de predecir el comportamiento de éstos ante determinadas situaciones. Es necesario detectar los comportamientos malintencionados, las conspiraciones entre agentes, etc. También deben poseer mecanismos para advertir los cambios en el comportamiento de sus vecinos y tomar las acciones oportunas con el objetivo de que dichas variaciones no afecten sus intereses.

Los agentes, por lo general, no muestran las mismas habilidades ante diferentes situaciones ni ante diferentes problemas. Por lo que el conocimiento que describe su comportamiento debe estar asociado al contexto y a las características de sus requerimientos.

6.2.2. Modelado de las soluciones de los proveedores y los requerimientos de los consumidores

En el escenario, tanto los proveedores como consumidores están representados por agentes, por lo que se definen según el modelo de agente descrito en el apartado 6.2.1, incorporando todas las características mencionadas. Cada agente proveedor tiene asociado un conjunto de soluciones que ofrece como respuesta para satisfacer los requerimientos de los consumidores, en su interacción con ellos. Por su parte, cada agente consumidor tiene asociado un conjunto de requerimientos que debe satisfacer a partir de su interacción con los proveedores.

El objetivo fundamental de los agentes en este escenario está determinado por la necesidad de encontrar la respuesta idónea a una especificación de usuario dada. Éste es el mismo objetivo de los procesos de descubrimiento de servicios que incorporan las tecnologías relacionadas con la web semántica. (Algunas de las cuales se comentan en el capítulo 4.) Luego, resulta factible pensar en utilizar algún modelo ontológico relacionado con la web semántica para representar las soluciones brindadas por los proveedores y los requerimientos de los consumidores. En [17, 19] se ofrecen algunas consideraciones iniciales para la representación ontológica de los agentes, los requerimientos de los usuarios y los servicios.

Por otra parte, como se comentó en el apartado 4.4, WSMO ofrece varias ventajas para la representación de los elementos de un modelo de confianza y la reputación. También en el apartado 5.8 se explica cómo se utiliza WSMO para representar los conceptos y funciones dependientes del dominio de aplicación. Además, es interesante utilizar el escenario definido por alguna tecnología relacionada con la web semántica para evaluar el modelo de confianza y reputación planteado.

Por todas estas razones, mencionadas anteriormente, se decide utilizar WSMO en este escenario para la definición de las soluciones brindadas por los proveedores y los requerimientos de los consumidores. De esta forma, los requerimientos se representan mediante tareas y las soluciones mediante respuestas a dichas tareas (ver las secciones 4.4 y 4.2.2).

6.2.3. Interacciones entre los agentes

En este apartado se comentan las características de las principales interacciones que tienen lugar entre los agentes desde el punto de vista de la utilización del modelo de confianza. La decisión de utilizar o no la información suministrada por el modelo de confianza y reputación, así como el resto de estrategias a seguir, queda fuera del alcance de la definición de este escenario de referencia. Se supone que el modelo de confianza se utiliza cuando el agente consumidor de recursos necesita valerse de algún criterio para seleccionar, entre sus vecinos, los que puedan ofrecer el servicio de mejor desempeño de acuerdo a sus necesidades. En [18] se identifican las decisiones de los agentes que pueden ser asistidas a partir de la incorporación de un mecanismo adaptativo de este tipo.

De manera general, la interacción con los agentes vecinos sigue un protocolo del tipo

Contract-Net [122]. Mediante mensajes del tipo *cfp*, el agente proveedor lanza una solicitud de propuestas acerca del compromiso de calidad que puedan ofrecer sus vecinos al dar solución a la tarea s_k . Por su parte, los agentes a_j que juegan el rol proveedor, pueden hacer su propuesta (mediante un mensaje del tipo *proposal*), comprometiéndose a garantizar un valor de calidad ec_j . Una vez recibidos todos los compromisos de calidad, el agente consumidor, iniciador de esta interacción, acepta (mediante un mensaje *accept-proposal*) las propuestas de los agentes que considere oportunos. Estos agentes responden (mediante un mensaje *response*) con la especificación de la solución (servicio) w_j que satisface los requerimientos de la tarea s_k . Un ejemplo de conversación entre el consumidor y el proveedor se muestra en la figura 6.2, ésta sigue la especificación del protocolo estándar que propone FIPA [48].

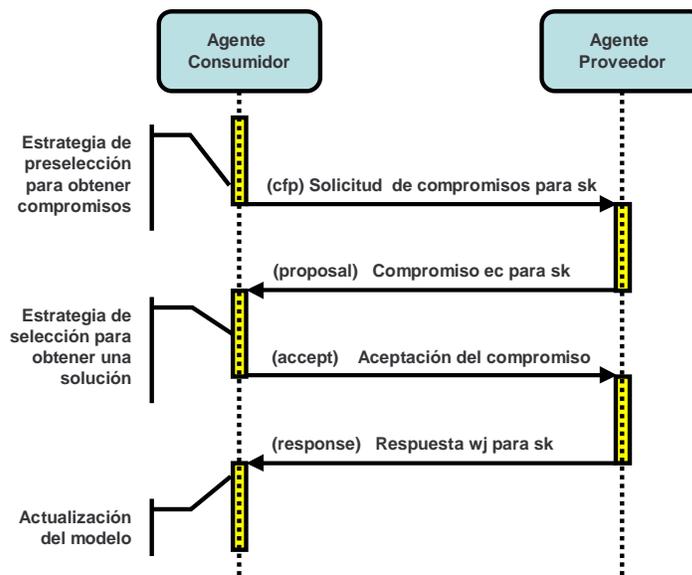


Figura 6.2: Ejemplo de secuencia de las interacciones entre los agentes en el escenario Consumidor - Proveedor de recursos, según los conceptos de TRSIM.

La utilización de un modelo de confianza y reputación adecuado puede aumentar la efectividad de las negociaciones de este tipo. De manera general, se pueden emplear criterios de confianza y reputación para definir las estrategias de identificación de los conjuntos de agentes a_j a los cuales se les pide propuestas o se les acepta para ofrecer la solución a una determinada tarea. Estas dos estrategias pueden requerir la solicitud de información de reputación, relacionada con el comportamiento de otros agentes (este diálogo no se encuentra representado en el diagrama de la figura 6.2). El modelo de confianza y reputación que se utilice debe actualizarse al recibir la respuesta w_j a dicha tarea s_k . De manera intuitiva, la actualización del modelo debe considerar en qué medida la respuesta w_j satisface los requerimientos de la tarea s_k y, por otra parte, ser capaz de evaluar en qué medida el agente a_j ha cumplido sus compromisos de calidad ec_j .

Las interacciones entre los agentes en este escenario de referencia se repiten varias veces, cada una de las iteraciones de negociaciones entre los agentes se les llama ronda. Así, una ronda comienza cuando se selecciona el agente consumidor que debe satisfacer sus requerimientos de usuario (especificados en la tarea s_k) y envía los mensaje *cfp* a los agentes proveedores que considere oportuno. Cada ronda finaliza cuando el agente consumidor actualiza el modelo de confianza a partir de la evaluación de la respuesta obtenida w_j .

6.3. Escenario de tasación de pinturas ofrecido por ART

Como se ha comentado en la sección 3.2.5, ART es un escenario de prueba de modelos de confianza y reputación que ofrece un marco común para la experimentación y competición. En su evolución, se pretende ir reflejando los esfuerzos de investigación que aporten avances significativos en el desarrollo de este campo. A diferencia del escenario anterior, definido a partir de una situación práctica concreta, las ideas que se comentan a continuación se corresponden con el desarrollo natural de un escenario de prueba que se encuentra en evolución.

6.3.1. Dominio de aplicación e interacciones en ART

En ART se define un dominio de aplicación en forma de juego. Los agentes que participan representan las solicitudes y respuestas de servicios de información sobre un dominio de aplicación relativo a la tasación de obras de arte. Las respuestas a las solicitudes de tasaciones pueden ser dadas a partir del propio conocimiento del agente o a partir de la compra de opiniones a otros. Cada agente recibe beneficios por las respuestas acertadas que ofrezca a quien las solicita, a la vez que invierte recursos en comprar opiniones a otros. Cada agente tiene como principal objetivo hacer crecer sus beneficios, por lo que la decisión de a quién comprar opiniones es una razón suficiente para considerar elementos de confianza y reputación con el fin de tomar las decisiones adecuadas en el momento de comprar o vender información.

Los agentes en ART siguen un modelo de interacciones definidos por el escenario de prueba. Corresponde al diseñador del agente la implementación del comportamiento del agente de acuerdo a las estrategias a seguir, según el modelo de confianza y reputación, para la compra y venta de informaciones acerca del valor de las obras de arte y de la reputación de los vecinos.

En ART el dominio de aplicación es la tasación de pinturas. Los agentes se encargan de tasar pinturas de diferentes eras para los clientes, quienes pagan una determinada cantidad de dinero por cada tasación. Cuanto más acertada sea la tasación realizada por el agente, mayores beneficios recibirá por su trabajo [50].

Ante una tasación, un agente puede dar su opinión que será más o menos acertada en función de su habilidad. Si no se cree capaz para tasar una determinada pintura puede

pedir (comprar) opiniones a otros agentes. Además de solicitar opiniones acerca del valor de las pinturas, un agente puede solicitar información acerca de la habilidad de otros agentes como tasadores. Las opiniones que se obtengan sobre la habilidad de otros agentes puede ser utilizada para saber a quiénes comprar información acerca del valor de una determina pintura. El valor de la tasación que hace un agente cuando se basa en la compra de opiniones a otros agentes, se obtiene a partir de la combinación de las opiniones compradas y la confianza que el agente otorga a sus vecinos.

El juego en ART es un proceso iterativo donde cada ronda de tasaciones está definida a partir de las interacciones entre los agentes. Es un proceso fijo y secuencial en cuanto al orden de las operaciones de los agentes. Primeramente, se compra y vende información de reputación (necesaria para estimar la habilidad de otros agentes ofreciendo tasaciones acertadas) y luego las opiniones acerca del valor de las pinturas.

Antes de comenzar una nueva iteración o ronda de tasaciones, cada agente puede ajustar la confianza en el resto a partir del análisis de información que ofrece el escenario de prueba y de la estrategia de actualización que defina el modelo de confianza en concreto.

Durante todo el juego el escenario de prueba ofrece un conjunto de parámetros que son conocidos por todos los agentes:

- *Client Fee* (f): cuánto paga el cliente por la tasación de la pintura.
- *Opinion Cost* (c_p): el costo de comprar una opinión a otro agente.
- *Reputation Cost* (c_r): el costo de comprar información acerca de la habilidad de otros ofreciendo opiniones acertadas (información de reputación).
- *Expertise Values*: valores de la habilidad del propio agente para dar opiniones acertadas para cada una de las eras.
- *Timestep*: número de la iteración actual

ART presupone que los costos de obtención de información acerca de la habilidad de otros agentes es mucho menor que el costo de la información de tasación, de forma que se favorezca el intercambio de valores de reputación entre los agentes.

6.3.2. Estrategia de los agentes en ART

En términos de implementación, cada comportamiento del agente y su comunicación con el resto están dados por la definición de métodos estratégicos en la clase abstracta *Agent* suministrada por el escenario de prueba. Por lo tanto, al implementar un agente para ser probado o competir en ART lo que se debe hacer es implementar las estrategias en cada uno de los métodos correspondientes. Los métodos estratégicos propuestos por ART son los siguientes:

1. *prepareReputationRequests()*: En este método se pueden obtener las respuestas de opiniones y resultados de tasaciones de la interacción anterior. Es aquí donde se

pueden efectuar los procesos de actualización de los valores de confianza y reputación a partir de los resultados obtenidos en la unidad de tiempo anterior.

También es donde se envían las solicitudes de información de reputación. Al solicitar información de reputación el agente está dispuesto a pagar un precio c_r por la información de reputación que obtenga, sin embargo, este valor es mucho menor que el precio que debe pagar por una opinión (c_p).

2. *prepareReputationAcceptsAndDeclines()*: Se aceptan o se rechazan las solicitudes de información de reputación que se han recibido.
3. *prepareReputationReplies()*: Se envían los valores de reputación de aquellas solicitudes que han sido aceptadas previamente
4. *prepareOpinionRequests()*: Se reciben las respuestas con información de reputación y se envían las solicitudes de opiniones. También se puede analizar a quién enviar las solicitudes y sobre qué pinturas se desea solicitar opiniones.
5. *prepareOpinionCertainties()*: En este método se envían los valores de certeza con los que se responderá cada una de las solicitudes de opiniones de tasación que han sido recibidas.
6. *prepareOpinionRequestConfirmations()*: Desde este método se confirma la solicitud de opiniones luego de analizar las certezas recibidas. Al responder con una confirmación afirmativa se está aceptando el pago que conlleva esta solicitud de información.

Es un buen lugar para tener en cuenta la confianza que se tiene sobre los agentes que han enviado su certeza de respuesta a la solicitud.

7. *prepareOpinionCreationOrders()*: Se dan las opiniones a partir de los precios c_g que se pagan por ellas, tanto para las opiniones que otros agentes han solicitado como para las tasaciones de las pinturas asignadas.

El costo que se ofrece por cada tasación representa el tiempo que el agente emplea en analizar la pintura. Por lo que al pagar un valor elevado de c_g , el tasador incrementará la exactitud de su opinión.

El grado de habilidad, asignado al agente para cada era, está descrito por una distribución normal del error entre la opinión del agente y el valor real de la pintura. La simulación crea opiniones de acuerdo a esta distribución que tiene media cero y desviación estándar s :

$$s = (s^* + \frac{\alpha}{c_g})t$$

donde

s^* es único para cada era y es asignado por el simulador de ART a cada agente de acuerdo a una distribución uniforme,

t es el valor real de la pintura, y

α es un parámetro de la competición fijo para todos los agentes.

Por tanto, la opinión a enviar estaría determinada por el costo c_g que se decida pagar por emitir una opinión.

8. *prepareOpinionProviderWeights()*: En este método es donde se deben enviar los pesos w_i al simulador para después poder calcular el valor de la tasación global p^* , sin que los agentes puedan variar sus pesos una vez conocida la tasación de cada pintura.

Estos pesos serán usados para conocer el valor de la opinión, calculada a partir de la suma ponderada de todas las opiniones solicitadas a otros agentes

$$p^* = \frac{\sum_i (p_i \cdot w_i)}{\sum_i (w_i)}$$

9. *prepareOpinionReplies()* : En este método se envía la opinión al agente que la ha solicitado, las solicitudes ya se han generado al establecer el costo de la opinión y enviarlas al simulador.

Las estrategias de interacción que cada agente utiliza para interactuar con sus vecinos quedan definida a partir de la implementación de estos nueve métodos estratégicos. En una ronda de interacciones, determinada por el parámetro *timestep*, los nueve métodos estratégicos se ejecutan según el protocolo de interacciones definido por ART [50]. La figura 6.3 muestra un ejemplo de conversación entre un agente solicitante y un tasador, durante una ronda determinada. Mediante ella se pretende ejemplificar cómo pueden ser utilizados algunos conceptos de TRSIM en la definición de las estrategias de los agentes según ART.

Las interacciones entre los agentes en ART, durante una ronda, se dividen en dos bloques fundamentales que tienen lugar de manera secuencial: las relacionadas con las transacciones de reputación y las relacionadas con las transacciones de confianza [50]. Mediante las primeras, el agente solicitante pide, a los tasadores, información de reputación sobre el comportamiento de otros para tasar una determinada pintura, debiendo pagar un costo por ellas. En este sentido, el agente solicitante debe establecer estrategias para (1) la preselección de los vecinos a quienes solicitar información de reputación, y (2) la compra de la información de reputación que puedan ofrecer los tasadores. El segundo bloque de transacciones están encaminadas a obtener las opiniones de los tasadores para una determinada pintura. En este sentido, el agente solicitante debe establecer estrategias para (1) determinar a qué vecinos solicitar compromisos (certeza de ofrecer una opinión de calidad, en términos de ART), y (2) comprar las opiniones adecuadas.

De manera general, en todas las estrategias del agente solicitante se pueden emplear criterios de confianza y reputación para la selección de los conjuntos de agentes. Por ejemplo, en la preselección de los vecinos a quienes solicitar información de reputación

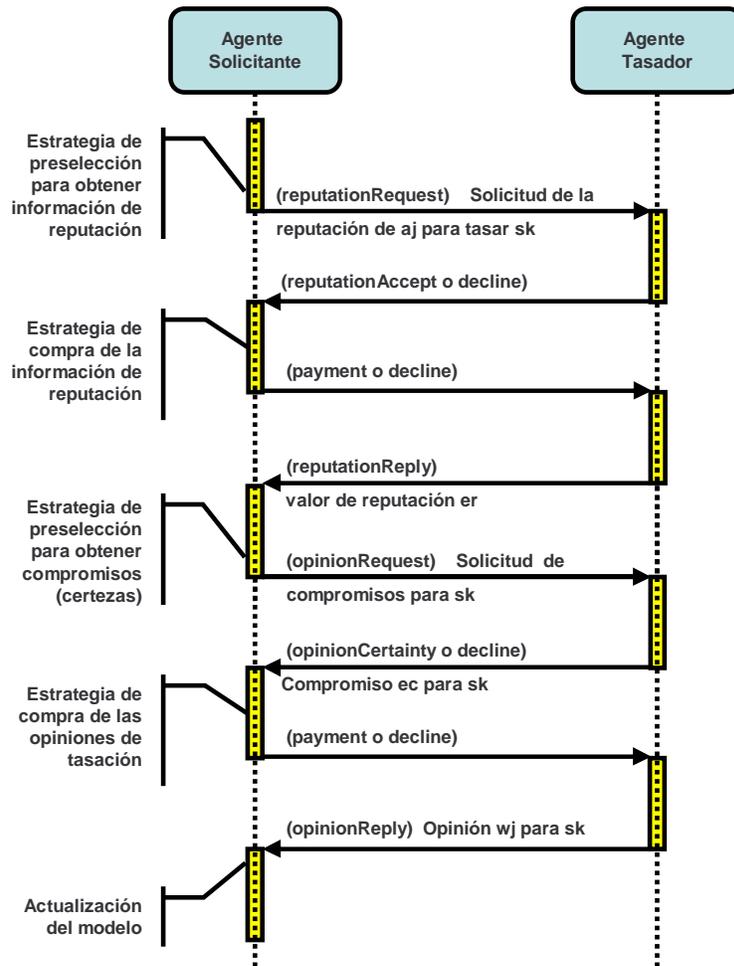


Figura 6.3: Ejemplo de secuencia de las interacciones entre los agentes en el escenario ART, según los conceptos de TRSIM.

se puede utilizar la estrategia comentada en el apartado 5.2. Ésta propone que el agente a_i solicite información a los agentes más confiables ofreciendo referencias ($CR_{sup}^{(t)}(a_i, s_k)$), sobre el comportamiento de los agentes de reputación dudosa ($CT_{dud}^{(t)}(a_i, s_k)$) para ofrecer opiniones de tasación relacionadas con una determinada pintura s_k .

El apartado 8.2 ofrece la definición detallada de las estrategias de un agente que se ejecuta en ART, según el modelo de confianza y reputación TRSIM (cada una de ellas, implementadas en el método correspondiente).

6.4. Conclusiones

En este capítulo se han mostrado dos escenarios de referencia donde puede resultar provechosa la utilización de TRSIM. El primero de ellos es un escenario práctico que reúne un grupo de características comunes a varios problemas distribuidos desde un enfoque en que los agentes desempeñan los roles de consumidores y proveedores de recursos. El segundo, ART, es un escenario de prueba en el que se ofrecen las herramientas necesarias para la evaluación y competición de agentes que implementan un determinado modelo de confianza y reputación.

A continuación, el capítulo 7 ofrece el análisis un conjunto de pruebas realizadas a TRSIM sobre el escenario Consumidor - Proveedor. Por su parte, el capítulo 8 aborda la adecuación de TRSIM al escenario ART, comentando los puntos de convergencia, de divergencia y las posibles extensiones que se pueden realizar a ART para que considere las características más importantes del diseño de TRSIM (la representación ontológica de algunos conceptos del modelo y la posibilidad de utilizar la similitud entre requerimientos de usuarios como método de intuición).

Capítulo 7

Experimentación

En este capítulo se presenta un conjunto de experimentos que muestran cómo se comporta el modelo TRSIM ante diferentes situaciones. Los experimentos se agrupan en dos bloques fundamentales de acuerdo a su finalidad:

1. Estudio del comportamiento del modelo según diferentes alternativas para la definición de algunas funciones importantes del modelo. (Los principales resultados relacionados con el análisis de las diferentes alternativas se muestran en la sección 7.2.)
 - a) Selección de la alternativa más adecuada para obtener la calidad de la respuesta a partir de la especificación de tarea que satisface (Q). La alternativa se selecciona a partir de analizar:
 - Para cada alternativa, la evolución de métricas tales como el nivel de satisfacción con la solución ofrecida por el modelo, el error cometido al ofrecerla, y la confianza en el agente que la recomienda.
 - Frente a cada alternativa para la obtención de la calidad, la capacidad del modelo para diferenciar, utilizando los valores de confianza y reputación, cada grupo de agentes, según la calidad de la respuesta que ofrecen.
 - b) Selección de la alternativa más adecuada para estimar los valores de confianza y reputación a partir de tareas similares (IT e IR). En estos casos se analiza:
 - Para cada alternativa, la evolución de métricas tales como el nivel de satisfacción con la solución ofrecida por el modelo, el error cometido al ofrecerla, y la confianza en el agente que la recomienda.
 - La independencia del comportamiento del modelo respecto a la función de similitud empleada. El comportamiento del modelo mejora cuando se estiman adecuadamente los valores de confianza y reputación a partir de la información sobre especificaciones de tareas similares.
 - Frente a cada alternativa para estimar la confianza y la reputación a partir de especificaciones de tareas similares, la capacidad del modelo para

diferenciar, utilizando los valores de confianza y reputación, cada grupo de agentes, según la calidad de la respuesta que ofrecen.

2. Estudio de la capacidad de adaptación del modelo ante los cambios en el comportamiento de los agentes proveedores. (Los resultados más relevantes en este sentido se ofrecen en la sección 7.3.)
 - a) Estudio de la capacidad de adaptación del modelo en diferentes situaciones, dadas por cambios en el comportamiento de un único agente. Las principales situaciones están dadas porque:
 - un agente bueno empeora su comportamiento,
 - un agente malo mejora su comportamiento,
 - un agente bueno empeora su comportamiento y vuelve a mejorarlo, o
 - un agente malo mejora su comportamiento y vuelve a empeorarlo.
 - b) Estudio de la capacidad de adaptación del modelo en diferentes situaciones, dadas por cambios en el comportamiento de un grupo de agentes. Las principales situaciones están dadas porque:
 - un grupo de agentes buenos empeoran su comportamiento,
 - un grupo de agentes buenos empeoran su comportamiento y vuelven a mejorarlo,
 - un grupo de agentes malos mejoran su comportamiento, o
 - un grupo de agentes malos mejoran su comportamiento y vuelven a empeorarlo.

Todos los experimentos se han desarrollado considerando el mismo escenario para la simulación a partir de la definición de un conjunto unificado de condiciones experimentales. El marco de los experimentos está determinado por las características del escenario Consumidor - Proveedor definido en el apartado 6.2.

A continuación, la sección 7.1 comenta las principales características de las simulaciones realizadas sobre este escenario, y las condiciones comunes de experimentación más importantes. Luego, en las secciones relacionadas con cada grupo de experimentos (7.2 y 7.3), se explican las cuestiones específicas de cada uno.

Los experimentos del bloque 1 no se presentan en un orden casual, sino que constituyen pasos progresivos en el proceso de refinamiento del modelo. Los resultados alcanzados en cada experimento son considerados en el siguiente. De esta forma, en cada simulación, se estudia el comportamiento del modelo a partir de la configuración que hasta el momento garantiza los mejores desempeños de los agentes.

7.1. Características generales de la simulación

El marco de los experimentos se define a partir de las interacciones entre agentes que desempeñan los roles de proveedores y consumidores de recursos, en un escenario

del tipo Consumidor - Proveedor, descrito en el apartado 6.2. En la simulación que se realiza se considera que todos los agentes exhiben el rol de proveedor de recursos, mientras que por cada ronda de la simulación solo uno de ellos juega el rol de consumidor, según los requerimientos de usuario (representados por las especificaciones de tareas) que se le asignan aleatoriamente en cada ronda.

Al inicio de la simulación se definen los conjuntos de especificaciones de tareas (requerimientos) S y de respuestas (soluciones) W que serán utilizadas, compuestos por 6 especificaciones de tareas y 10 respuestas, respectivamente. Siguiendo las características del escenario, en la simulación se considera que cada especificación de tarea o respuesta está descrita por un conjunto de siete propiedades, tal como se comenta en el apartado 5.8.

Un tercio del conjunto de especificaciones de tareas S corresponden a tareas de exigencia *Alta*, otro tercio a tareas de exigencia *Media* y otro tercio a tareas de exigencia *Baja*.

El nivel de exigencia de una tarea indica cuan difícil es de satisfacer dicha tarea a partir del conjunto de soluciones que se ofrecen en el sistema. En la definición del escenario de simulación, se ha considerado que la exigencia de la tarea está determinada por los valores de sus atributos de forma que cuanto mayores son los valores de dichos atributos mayor es el grado de exigencia de la tarea.

Para generar los atributos de las tareas de exigencia *Alta* se utiliza una distribución de probabilidad Beta con parámetros $\alpha = 5$ y $\beta = 2$. Es decir, para una tarea s_k de exigencia *Alta* se definen todos sus atributos s_{k_i} mediante la distribución $Beta(\alpha = 5, \beta = 2)$. La función de distribución de probabilidades que se utiliza para generar los atributos de las tareas de calidad *Alta* se muestra en la figura 7.1.a.

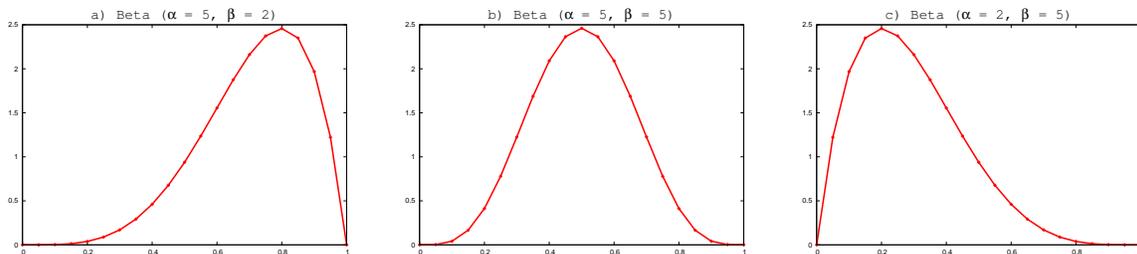


Figura 7.1: Distribuciones de probabilidad Beta utilizadas para generar los atributos de las tareas y respuestas según su exigencia y calidad, respectivamente. a) $Beta(\alpha = 5, \beta = 2)$ para exigencia y calidad *Alta*; b) $Beta(\alpha = 5, \beta = 5)$ para exigencia y calidad *Media*; c) $Beta(\alpha = 2, \beta = 5)$ para exigencia y calidad *Baja*.

Para obtener los atributos de tareas de exigencia *Media* se utilizan los parámetros $\alpha = 5$ y $\beta = 5$. La función de distribución de probabilidades que se utiliza para generar los atributos de las tareas de calidad *Media* se muestra en la figura 7.1.b.

Por otra parte, para obtener los atributos de tareas de exigencia *Baja* se utilizan los parámetros $\alpha = 2$ y $\beta = 5$. La función de distribución de probabilidades que se utiliza para generar los atributos de las tareas de calidad *Baja* se muestra en la figura 7.1.c.

De igual forma, el conjunto de soluciones W también se obtiene mediante distribuciones de probabilidad Beta, identificándose de igual forma tres tipos de respuestas según su

calidad. Para generar soluciones de calidad *Alta* se utiliza la distribución $Beta(\alpha = 5, \beta = 2)$, para las de calidad *Media* $Beta(\alpha = 5, \beta = 5)$, y para soluciones de calidad *Baja* $Beta(\alpha = 2, \beta = 5)$.

A cada agente, al desempeñar el rol de proveedor, se le asigna al inicio de la simulación una solución del conjunto W , definiendo a su vez tres tipos de agentes. Al 30% de los proveedores se les asignan respuestas de calidad *Alta*, al 30% de calidad *Baja* y al 40% de calidad *Media*. Al inicio de una réplica de la simulación, se designa un único agente consumidor para representar los requerimientos de usuarios cuya descripción se obtiene aleatoriamente del conjunto de especificaciones de tarea S .

Una ronda de la simulación es la mínima unidad de tiempo considerada en el modelo. Se inicia cuando se establecen los requerimientos del agente iniciador mediante la especificación de tarea seleccionada, y finaliza cuando se actualizan las bases de experiencias según la evaluación que el agente hace de la solución ofrecida por el agente más confiable, según el criterio del modelo de confianza y reputación. En cada experimento se llevan a cabo 400 rondas de la simulación donde en cada una el agente iniciador utiliza los valores de confianza y reputación que ofrece el modelo para seleccionar el agente de mejor desempeño. Al finalizar cada ronda, el agente iniciador evalúa las respuestas y retroalimenta el modelo para futuras iteraciones. Así, con el transcurso del número de rondas, existe un mayor volumen de información, lo que debe traducirse en la estabilización de los indicadores del funcionamiento del modelo (por ejemplo, el nivel de satisfacción con la respuesta recomendada por el modelo, el error cometido al seleccionar la respuesta adecuada, la confianza en el agente que ofrece la respuesta, entre otros).

De manera general, el algoritmo de la simulación, que se repite durante 20 réplicas se ofrece en la figura 7.2. Sin embargo, la forma de asignar a cada agente de A una respuesta de W supone situaciones experimentales diferentes, por ello, en cada experimento se analizan los siguientes casos:

- la tarea asignada al agente iniciador se selecciona aleatoriamente del conjunto de tareas W ;
- siempre se le asigna, al agente iniciador, una tarea de exigencia *Alta*;
- siempre se le asigna, al agente iniciador, una tarea de exigencia *Media*;
- siempre se le asigna, al agente iniciador, una tarea de exigencia *Baja*;

Aunque TRSIM ha sido planteado para ser considerado en entornos negociadores entre ambas partes: proveedor y consumidor, las simulaciones que se presentan han simplificado este hecho. Tal como se trata en el apartado 5.8.1, el modelo evalúa la interacción con cada agente teniendo en cuenta la calidad de la respuesta dada la tarea que satisface, así como el cumplimiento del compromiso de satisfacción hecho por el agente proveedor. En los experimentos realizados, se considera que el compromiso de calidad que cada proveedor hace frente al consumidor siempre es el máximo posible $ec = 1$. Esto supone que el agente proveedor se compromete a ofrecer la mejor de sus soluciones (en estas simulaciones, cada

1. crear el conjunto de agentes A
2. crear el conjunto de respuestas W
3. crear el conjunto de tareas S
4. asignar a cada agente de A una respuesta de W
5. seleccionar un agente a_i de A
6. Para cada ronda t hacer:
 - seleccionar una tarea s_k de S como requerimiento de a_i
 - seleccionar el(los) agente(s) a_j a los cuales se les pide solución
 - evaluar la interacción utilizando la solución w_j dada por cada agente a_j :
 $InteractionResults := (t, a_i, a_j, s_k, q = Q(w_j, s_k))$
 - actualizar las bases de experiencias a partir de la evaluación de la interacción:
 $updateBases(InteractionResults)$

Figura 7.2: Algoritmo general de la simulación

proveedor ofrece una única solución) y que considera que su oferta es la mejor posible para satisfacer los requerimientos de usuarios que se plantean.

Los datos experimentales que se muestran en los apartados siguientes son el resultado de la simulación del funcionamiento del modelo TRSIM a partir de una implementación propia utilizando Java. Cada dato es el promedio de los valores obtenidos durante las 20 réplicas de cada simulación.

7.2. Alternativas para la definición de algunas funciones importantes del modelo

El primer grupo de experimentos está encaminado a seleccionar las alternativas que muestren el mejor desempeño para la definición de funciones importantes en el modelo de confianza y reputación.

En estos experimentos se estudia la evolución del nivel de satisfacción que se tiene con la solución recomendada, así como la del error cometido por el modelo al recomendarla según la especificación de tarea dada. El error cometido se define como la diferencia entre la calidad de la respuesta recomendada por el modelo y el valor óptimo de calidad (es decir, el valor de calidad de la mejor de todas las respuestas). Tal como se comenta en el apartado 3.1.1.2, las métricas relacionadas con el nivel de satisfacción del usuario (en este

caso representado por el agente iniciador) resultan útiles en sistemas recomendadores de este tipo. Por lo general, el estudio de la eficiencia del modelo se basa en la evolución y convergencia de estos indicadores a lo largo de la utilización del modelo.

Otra medida de comparación es el momento en que el modelo alcanza un estado estable para las distintas alternativas. En estos casos se considera que se alcanza la estabilidad del modelo cuando el error cuadrático medio cometido en la selección de la respuesta adecuada no supera 0,01. Carbó [28] utiliza este mismo criterio, teniendo en cuenta que Zacharia y Maes [147] estimaron el momento de equilibrio cuando el error cuadrático medio fue inferior a $0,01 \cdot Dom^2$, donde Dom es la amplitud del dominio de en el que se define el error. En el caso de TRSIM, este valor está dado por $Dom = 1$.

Por otra parte, el objetivo principal del modelo de confianza y reputación es asistir el proceso de selección de la mejor respuesta para satisfacer la especificación de tarea que representa las necesidades de los usuarios. El modelo recomienda la selección de la respuesta que ofrece el agente más confiable (con mayor valor de confianza asociado a la realización de la tarea en cuestión), no considerando como criterio de selección la calidad de la respuesta en sí, por cuanto es desconocida.

Por lo tanto, en estos experimentos también se estudia la evolución de la confianza media para tres tipos de agentes según la calidad de las respuestas que tienen asignadas. (El apartado 3.1.1.1 comenta la utilidad de estas medidas para analizar la convergencia del modelo y realizar los ajustes necesarios para obtenerla de manera temprana.) La clasificación de los agentes se ha realizado tomando como criterio la calidad de la respuesta que ofrecen: *Alta*, *Media* o *Baja*. Cada respuesta tiene asociado una etiqueta de calidad que define los valores de sus propiedades, según la distribución probabilística que los genera al inicio de la simulación (consultar la sección 7.1).

Todos los experimentos de este bloque se desarrollan teniendo en cuenta que el comportamiento de los agentes es estacionario a lo largo de toda la simulación. (En el siguiente bloque de experimentos, se estudia la adaptabilidad del modelo a cambios de comportamientos no estacionarios.)

Por otra parte, al obtener los valores de confianza directa y su desviación a partir de la información de la base de experiencias para la confianza se ha utilizado $\delta = 0,6$ como parámetro modulador del tiempo. Este parámetro modela el hecho de que las experiencias pierdan relevancia con el tiempo (consultar la sección 5.5).

Se ha utilizado el valor $\gamma = 0,80$ de forma que se toman los agentes cuyo valor de reputación, según el criterio los agentes más confiables ofreciendo información sobre el comportamiento de otros, supera el percentil $p_{\gamma=0,80}$ en el conjunto de agentes CT_{dud} .

Se ha establecido el valor $itm = 10$ como límite de experiencias consideradas para mejorar la precisión de la medida de confianza directa. En la simulación, este valor limita el número máximo de experiencias de un agente en su base de experiencias para la confianza IET .

7.2.1. Elección de la alternativa adecuada para la calidad

Este experimento compara las diferentes medidas comentadas para tres alternativas utilizadas en la obtención del nivel de satisfacción de una tarea a partir de la respuesta ofrecida por el modelo. Las tres alternativas para obtener la calidad (Q_a, Q_b, Q_c) son descritas en el apartado 5.8.1.1:

Q_a : siguiendo el enfoque ofrecido por el proceso de descubrimiento de servicios web dado por WSMO

$$Q(w_j, s_k) = \begin{cases} 1 & : R_g = R_w \\ 0,75 & : R_g \subseteq R_w \\ 0,5 & : R_g \supseteq R_w \\ 0,5 & : R_g \cap R_w \neq \emptyset \\ 0 & : R_g \cap R_w = \emptyset \end{cases}$$

donde R_g es el conjunto de los atributos relevantes de la tarea s_k y R_w es el conjunto de los atributos relevantes de la respuesta w_j .

Q_b : considerando cuántos atributos de la tarea son satisfechos por la respuesta

$$Q(w_j, s_k) = \sin \left(\frac{\pi}{2} \cdot \frac{|Q'_{w_j, s_k}|}{|R_u|} \right)$$

donde R_u es el conjunto de todos los atributos de las tareas o respuestas, y

$$Q'_{w_j, s_k} = \{b_i | b_i \in R_u, v_{s_k}(b_i) \leq v_{w_j}(b_i)\}$$

Q_c : considerando cuántos atributos relevantes de la tarea son satisfechos por la respuesta

$$Q(w_j, s_k) = \sin \left(\frac{\pi}{2} \cdot \frac{|Q''_{w_j, s_k}|}{|R_g|} \right)$$

donde

$$Q''_{w_j, s_k} = \{b_i | b_i \in R_g, v_{s_k}(b_i) \leq v_{w_j}(b_i)\}$$

En este experimento no se utilizan los criterios de similitud entre tareas para estimar los valores de confianza y reputación de un agente para tareas sobre las cuales se desconoce su desempeño.

7.2.1.1. Evolución de la satisfacción, el error y la confianza

En la figura 7.3 se muestra la evolución del nivel de satisfacción con la solución propuesta para cada alternativa, de acuerdo a la forma en que se selecciona la tarea asignada al agente iniciador. En esta figura se observa que, de manera general, no existen diferencias significativas entre los valores de satisfacción alcanzados con cada alternativa para ningún

tipo de selección de la tarea. Sin embargo, el modelo se comportan de manera diferente de un tipo de tarea a otro. Cuando se seleccionan tareas de exigencia *Baja* los valores de satisfacción que se alcanzan son más altos que en los casos en que se selecciona otro tipo de tareas. Evidentemente, cualquier respuesta genera mayores niveles de satisfacción en tanto menor sea el grado de exigencia de la tarea.

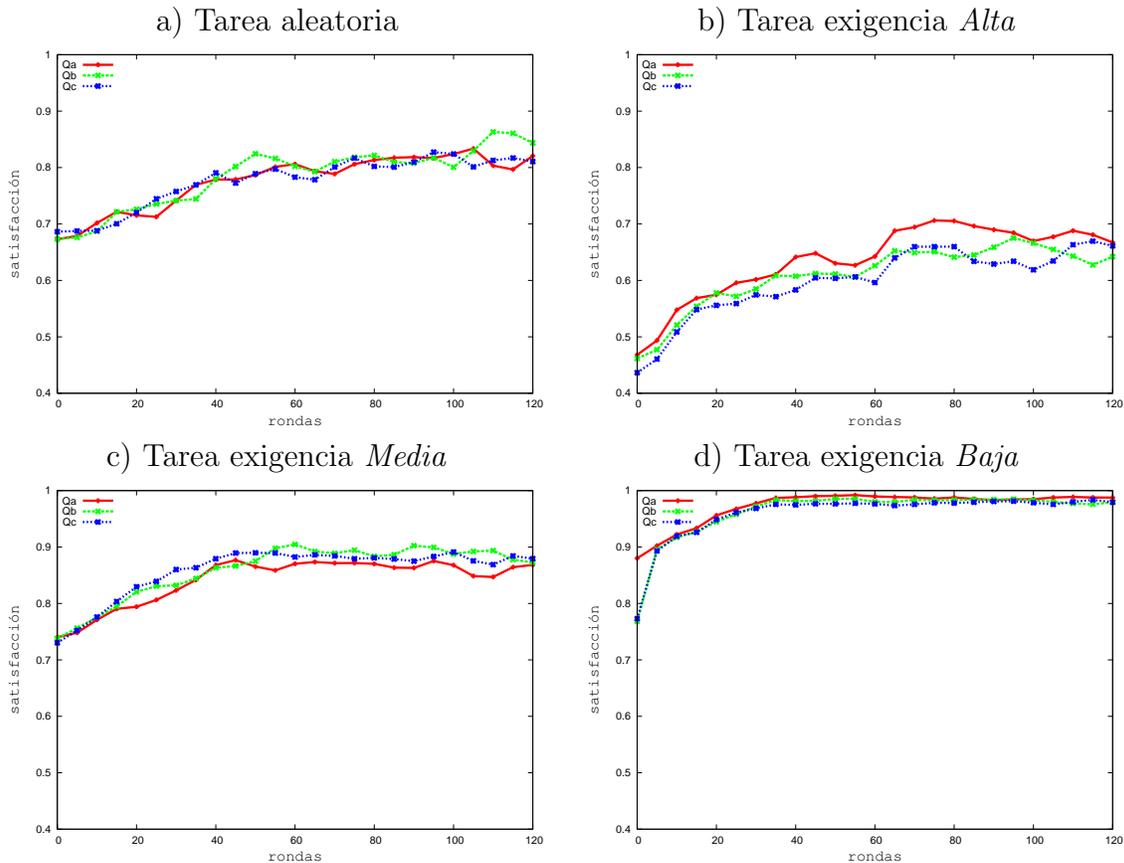


Figura 7.3: Comparación de la satisfacción con la solución recomendada por el modelo, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

Adicionalmente a la figura anterior se ofrece la figura 7.4 donde se muestra la evolución del error cometido por el modelo al seleccionar la solución adecuada según la especificación de tarea dada. El comportamiento del modelo en las tres alternativas se estabiliza luego de un número relativamente pequeño de rondas, para todos los tipos de tarea. Sin embargo, el comportamiento del modelo es diferente para diferentes formas de seleccionar la especificación de tarea. Con cada alternativa, los errores obtenidos cuando se seleccionan tareas de exigencia *Baja* son menores que cuando se selecciona otro tipo de tarea. Esto se debe a que existe un mayor número de agentes cuyas soluciones pueden ser adecuadas para resolver una tarea de exigencia *Baja*, y por tanto, le será más fácil al modelo recomendar

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

una respuesta adecuada. La figura 7.5 compara los instantes (número de ronda) en los cuales se estabiliza el modelo según cada variante, de acuerdo al criterio de estabilización considerado.

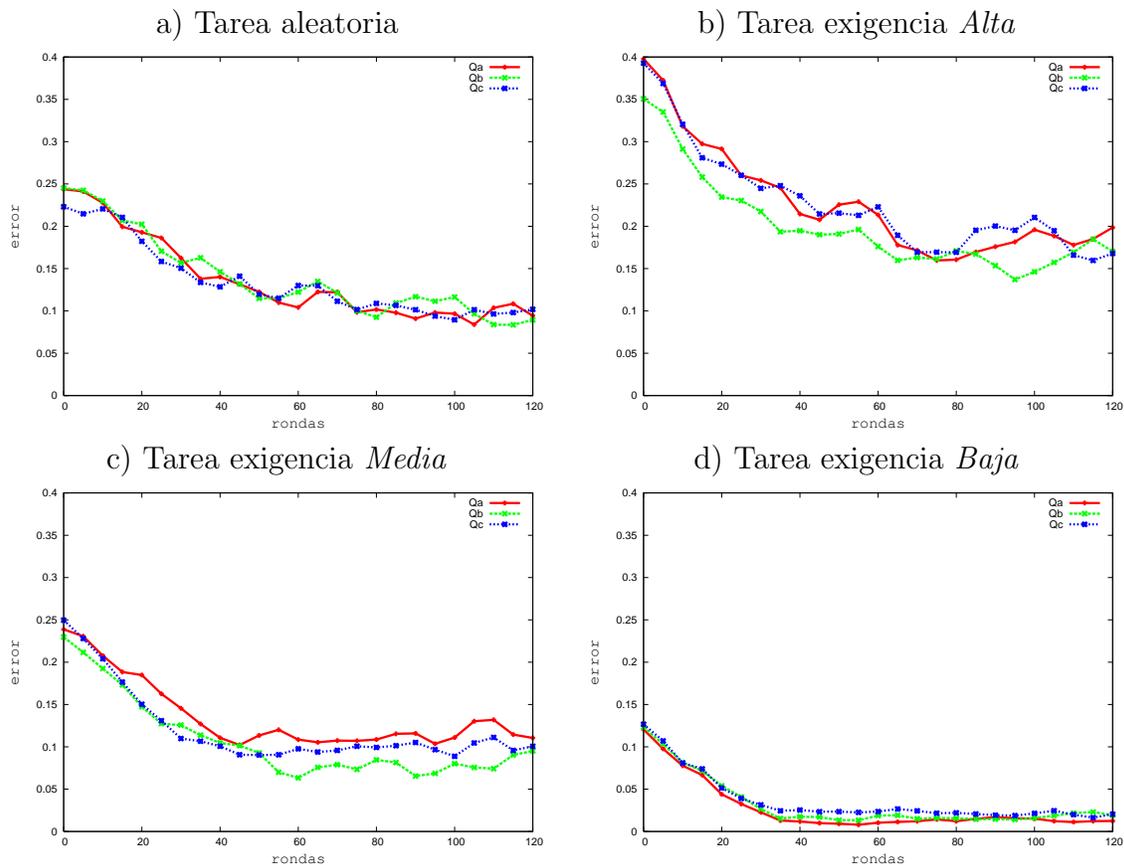


Figura 7.4: Comparación del error cometido por el modelo al seleccionar la solución adecuada, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

De acuerdo a los resultados mostrados en la figura 7.5, la alternativa Q_b se estabiliza antes que Q_a y Q_c cuando la exigencia de la tarea se selecciona aleatoriamente o es de exigencia *Baja*. En los casos en que se seleccionan tareas de exigencia *Alta* o media, las alternativas Q_a y Q_b garantizan la estabilidad del modelo en el mismo instante de tiempo. En estos dos casos, la alternativa Q_c se muestra de manera diferente: frente a tareas de exigencia *Alta* ofrece el mejor tiempo de estabilización, mientras que para tareas de exigencia *Media*, el peor.

En otro sentido, en la figura 7.5 (al igual que en las figuras 7.3 y 7.4) se observa que la estabilidad del modelo se alcanza más rápidamente en tanto menor sea la exigencia de la tarea que se pretende resolver. Los tiempos más bajos de estabilización se alcanzan cuando la tarea es de exigencia *Baja*, mientras que los más altos corresponden a tareas de exigencia

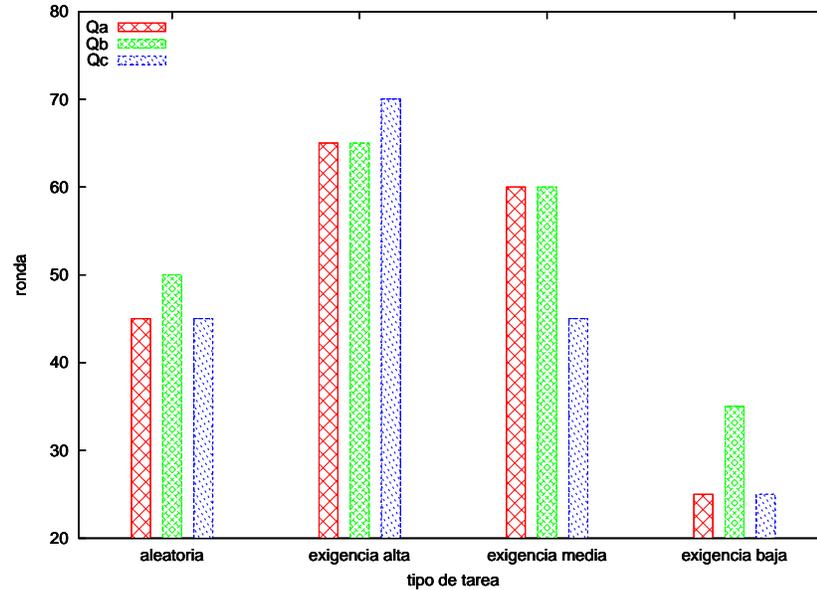


Figura 7.5: Comparación del momento (número de ronda) en que el modelo alcanza la estabilidad, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, agrupadas según el nivel de exigencia de la tarea.

Alta. Esto es un comportamiento esperado en el modelo ya que cuanto más fácil sea de satisfacer una tarea dada, más fácil será de proponer, por parte del modelo, la respuesta que garantice los mayores valores de satisfacción.

En todas las alternativas analizadas (Q_a , Q_b , Q_c), la satisfacción con la respuesta recomendada por el modelo se estabiliza, independientemente del método utilizado para su selección.

No se observan diferencias significativas entre los errores cometidos por cada alternativa al recomendar la solución que ofrece mayor nivel de satisfacción para la tarea a resolver.

La disminución del error depende considerablemente de la exigencia de la tarea a resolver en cada instante, de forma que cuanto menor es el nivel de exigencia de la tarea, más pequeños son los valores alcanzados.

Por otra parte, la figura 7.6 muestra la evolución de la confianza del agente iniciador en los agentes que se seleccionan, en cada alternativa, para ofrecer la respuesta a la tarea a resolver en cada ronda. Es decir, los puntos de cada curva se corresponden con los valores de confianza que tiene el agente iniciador en el agente que se recomienda en cada ronda, para ofrecer la solución a la especificación de tarea planteada. Dado que el modelo recomienda al agente de mayor confianza, el valor de cada curva, en cada instante de tiempo, puede entenderse como el máximo valor de confianza que el agente iniciador tiene en sus vecinos en dicho instante.

En esta figura se observa que la confianza en el agente que ofrece la respuesta en cada instante de tiempo aumenta a lo largo del tiempo hasta alcanzar un nivel máximo al que converge, aun cuando la tarea a resolver varía aleatoriamente (Figura 7.6.a). Por

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

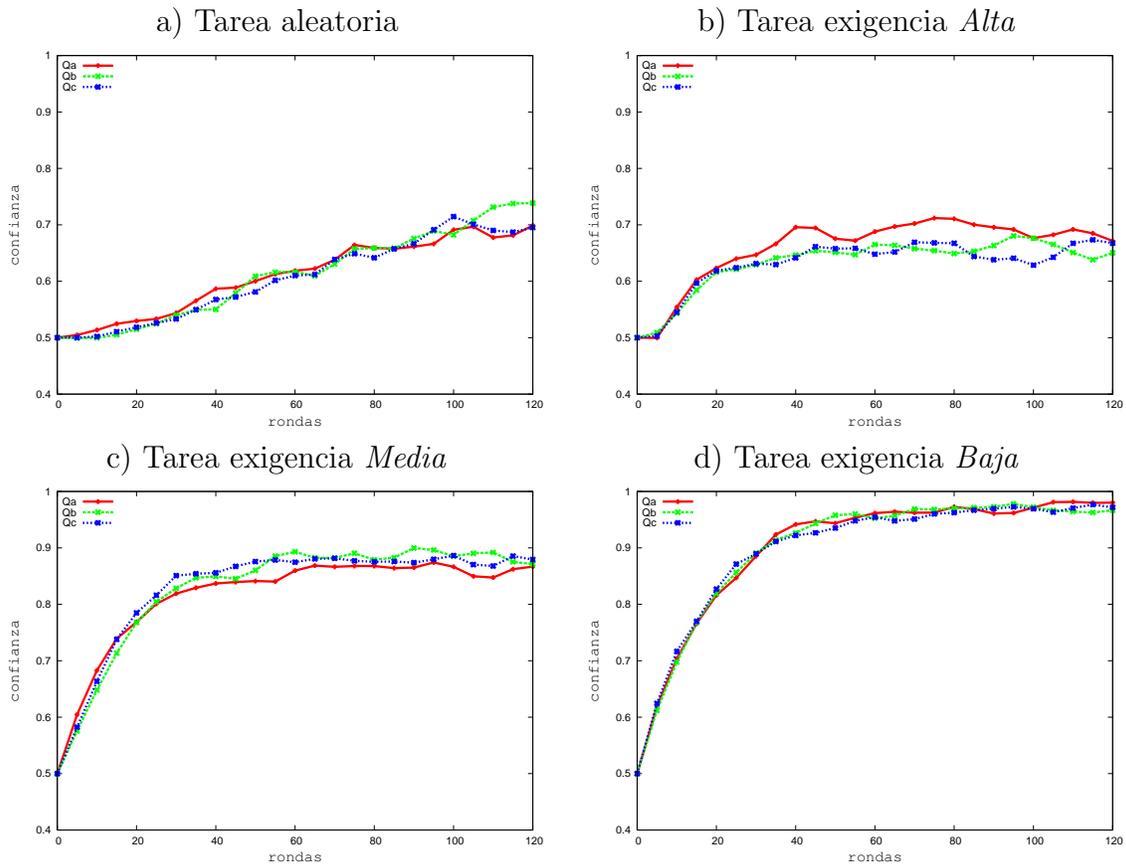


Figura 7.6: Comparación de la confianza del agente que ofrece la solución recomendada por el modelo, para cada alternativa utilizada en la obtención de la calidad a partir de la tarea que satisface, según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

tanto, se puede concluir que el modelo, para estas funciones de calidad de la respuesta, alcanza el nivel de equilibrio para cualquier tipo de tareas, incluso cuando se selecciona aleatoriamente. El nivel de equilibrio del modelo se alcanza con mayor rapidez en tanto la exigencia de la tarea sea menor.

De igual manera, los valores de confianza dependen considerablemente de la tarea a resolver en cada instante, de forma que se alcanzan valores de confianza más elevados cuanto menor es el nivel de exigencia de la tarea. Esto es lógico ya que la confianza es resultado, entre otros indicadores, del nivel de satisfacción que se tiene con un agente para la realización de una tarea dada. Así, cuanto menos exigente sea la tarea, se tendrán niveles de satisfacción más elevados y por consiguiente, el agente que pretende solucionar la tarea otorgará valores de confianza más altos.

Podría pensarse que es injusto asignar mayor confianza a agentes que satisfacen tareas de exigencia *Baja*. Sin embargo, en este caso la confianza debe verse como un indicador de cuan seguro está el agente consumidor del comportamiento de un agente concreto para una tarea determinada.

De manera general, no se observan diferencias significativas en la evolución de la satisfacción, el error y la confianza entre las tres alternativas utilizadas (Q_a , Q_b , Q_c) en la obtención de la calidad de la respuesta a partir de la tarea que satisface. Aunque el experimento no ha servido para discriminar sobre la conveniencia de usar una u otra alternativa, sí ha resultado esclarecedor en el análisis de cómo se alcanza la estabilidad del modelo, mostrando su dependencia del tipo de tarea que se requiere satisfacer en cada instante de tiempo.

En este sentido, para experimentos siguientes resultaría igualmente interesante la utilización de cualquiera de las alternativas para la obtención de la calidad de la respuesta a partir de la tarea que satisface. Sin embargo, la diferenciación de cada tipo de agentes que se produce con cada alternativa, es un criterio que puede servir para seleccionar la más adecuada según los intereses del modelo planteado.

7.2.1.2. Diferenciación de los grupos de agentes según la calidad de la respuesta que ofrecen

En este apartado se analiza la evolución de la confianza del agente iniciador en el resto de agentes, según la calidad de la respuesta que ofrecen. Se pretende con ello, encontrar la alternativa más conveniente para la definición de la calidad de la respuesta (Q_a, Q_b, Q_c) que mejor diferenciación haga de los grupos de agentes, utilizando criterios de confianza.

En las figuras 7.7, 7.8 y 7.9 se muestra la evolución de la confianza del agente iniciador que desea resolver la tarea en cada ronda, agrupadas según la alternativa utilizada en la obtención de la calidad de la respuesta a partir de la tarea que satisface. La figura 7.7 muestra la evolución de la confianza en cada tipo de agentes de acuerdo a la forma en que se selecciona la tarea asignada al agente iniciador cuando se utiliza la alternativa Q_a . De igual forma, las figuras 7.8 y 7.9 muestran la evolución de la confianza cuando se utiliza Q_b y Q_c , respectivamente.

En estas figuras se observa que independientemente del tipo de tarea, el modelo es capaz

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

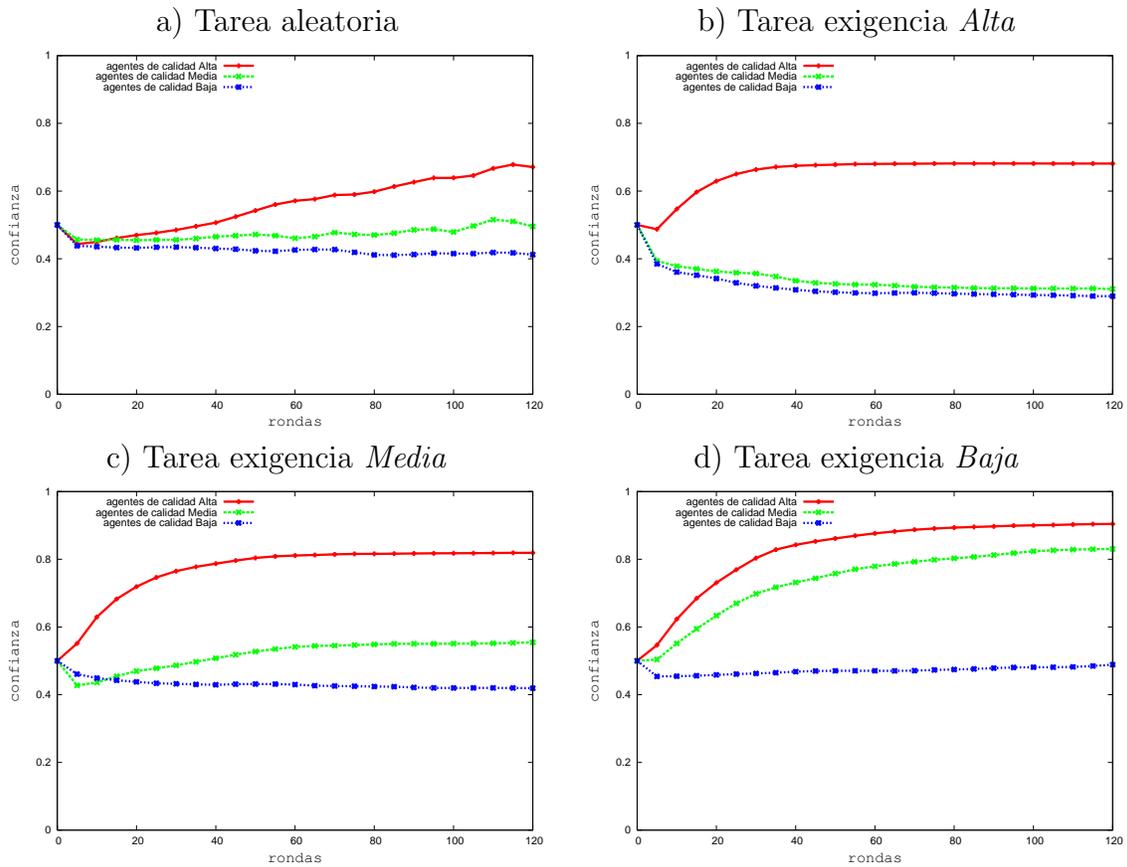


Figura 7.7: Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_a , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

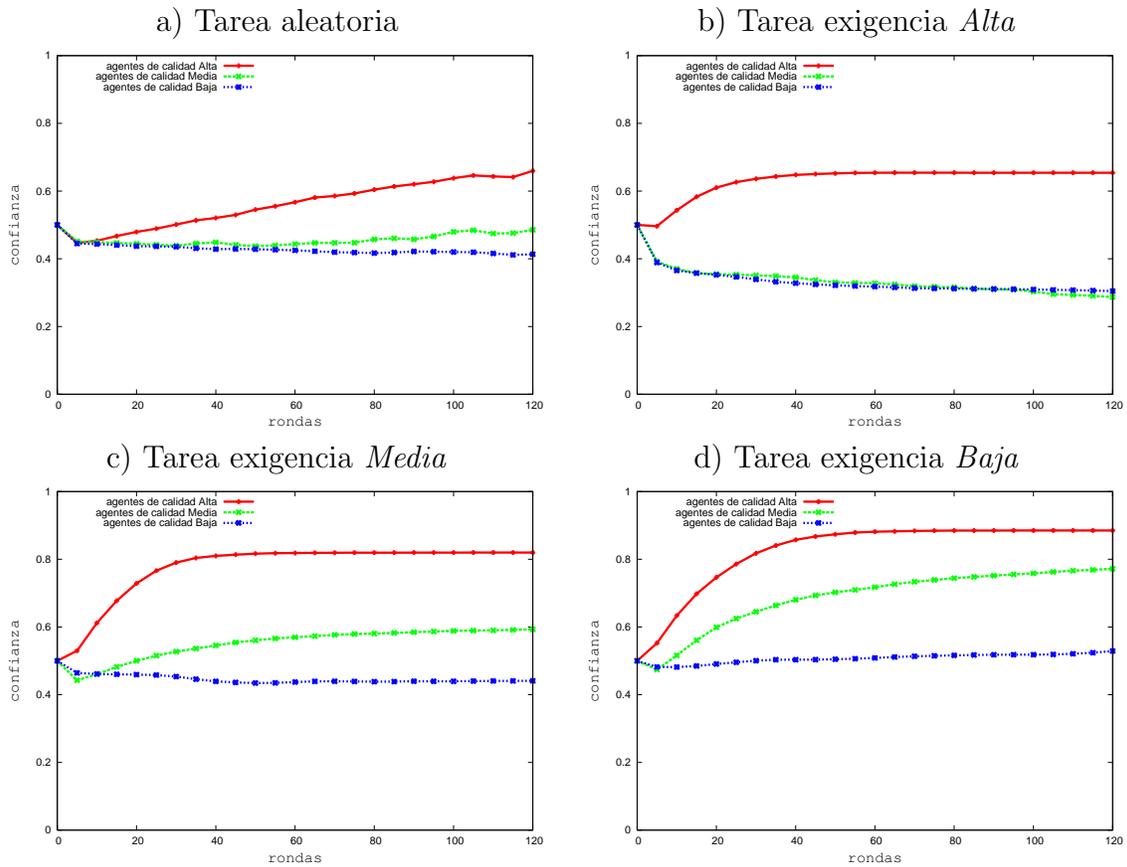


Figura 7.8: Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_b , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

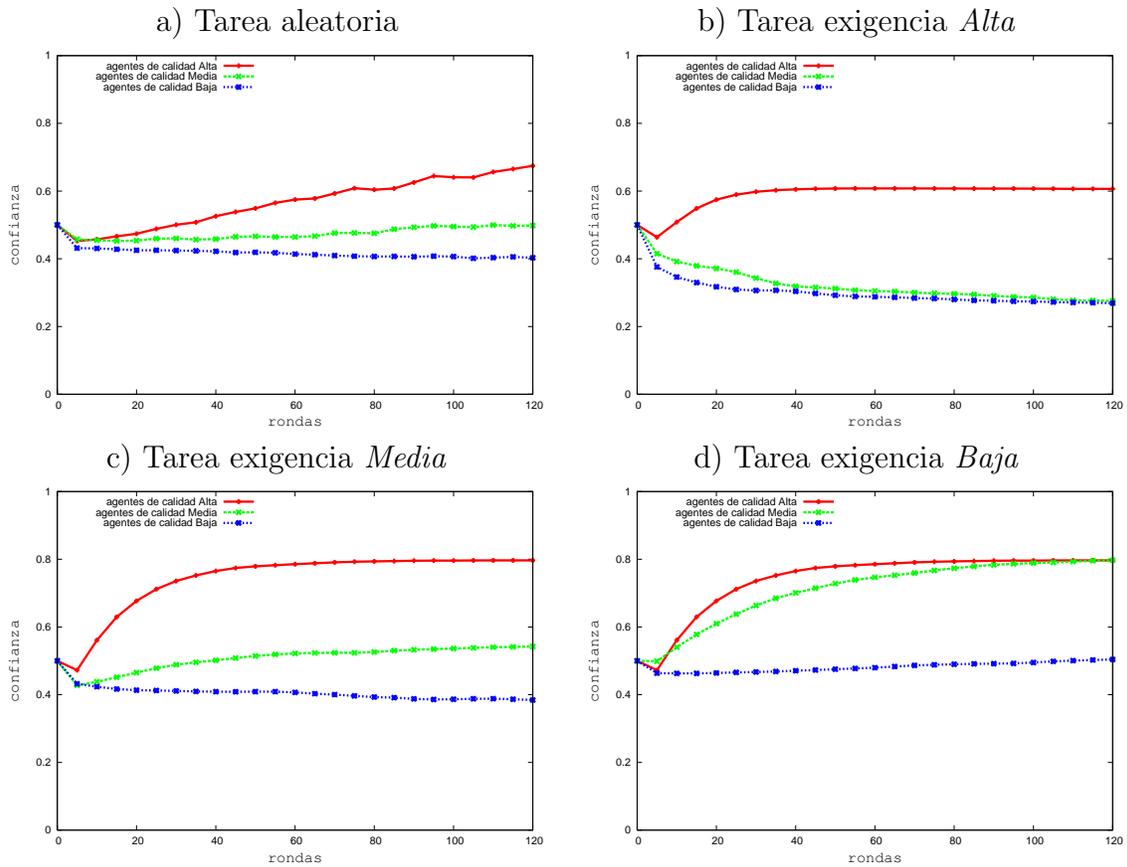


Figura 7.9: Comparación de la evolución de la confianza para cada grupo de agentes para la alternativa Q_c , según la exigencia de la tarea a resolver en cada ronda: a) aleatoria, b) alta, c) media, d) baja.

de reconocer la calidad de las soluciones que ofrecen los agentes utilizando los valores de confianza en ellos. De manera general, los que ofrecen soluciones de calidad *Alta* aumentan su confianza, los de calidad *Media* se mantienen en valores intermedios y los de calidad *Baja* muestran un ligero decrecimiento. Por ejemplo, al considerar la alternativa Q_a se observa que cuando la tarea se selecciona aleatoriamente (Figura 7.7.a) los agentes de calidad *Alta* aumentan su confianza en mayor medida que los agentes de calidad *Media*; a la vez que los de calidad *Baja* muestran una ligera disminución. Este comportamiento es similar para los casos en que la tarea es de exigencia *Alta* (Figura 7.7.b), de exigencia *Media* (Figura 7.7.c) o de exigencia *Baja* (Figura 7.7.d)

Comportamientos similares se observan para las otras dos alternativas Q_b y Q_c , lo que indica que todas son capaces de diferenciar los grupos de agentes según la calidad de su respuesta. Sin embargo, ninguna es capaz de diferenciar los agentes de calidad *Media* y los de calidad *Baja* para ofrecer respuestas a tareas de exigencia *Alta* (Figuras 7.7.b, 7.8.b y 7.9.b). Esto se debe fundamentalmente a que dada la alta exigencia de las tareas de este tipo, el modelo solo incrementa los valores de confianza en los agentes de calidad *Alta*; ni los de calidad *Media* ni los de calidad *Baja* son adecuados para resolverlas, y por tanto, en estas situaciones su confianza disminuye de igual forma. Este es un comportamiento deseado del modelo. En otros casos, donde la exigencia de la tarea no es *Alta*, el modelo muestra evoluciones diferentes de la confianza en agentes de calidad *Media* y en agentes de calidad *Baja*.

Por otra parte, cuando la tarea es de exigencia *Alta* o *Media*, la diferenciación entre los grupos de agentes de calidad *Media* y calidad *Baja* es menos pronunciada por cuanto se favorece la interacción con agentes de calidad *Alta*, con cuya evaluación se retroalimenta el modelo. Esto se debe a que el modelo posee más experiencias relacionadas con agentes buenos que con el resto.

Desde otro punto de vista, al comparar las tres alternativas para los casos en que las tareas se seleccionan aleatoriamente (Figuras 7.7.a, 7.8.a y 7.9.a) se observa que la alternativa Q_b es capaz de diferenciar mejor los agentes de calidad *Alta* de los agentes de calidad *Media*. Este resultado se ha obtenido comparando las áreas entre las curvas que representan los agentes de calidad *Alta* y los de calidad *Media* para cada alternativa, desde la ronda inicial $t = 0$ hasta la última ronda de la simulación $t = 400$. (Nótese que las figuras 7.7, 7.8 y 7.9 solo muestran la evolución de la confianza hasta la ronda $t = 120$.) Los valores de las áreas entre las curvas de los agentes de calidad *Alta* y *Media* para cada alternativa se muestra en la tabla 7.1.

Tabla 7.1: Áreas entre las curvas que describen los valores de confianza del agente iniciador sobre los agentes de calidad *Alta* y *Media* para cada una de las alternativas utilizadas en la obtención de la calidad de la respuesta (desde $t = 0$ hasta $t = 400$)

	Q_a	Q_b	Q_c
Diferencia entre Altos y Medios	20.69	25.25	21.80

La alternativa Q_b será utilizada en el resto de experimentos para determinar el nivel de

satisfacción de la tarea a partir de la respuesta brindada por el modelo. Se ha elegido esta alternativa por cuanto es la que mejor diferencia los grupos de agentes según la calidad de la respuesta que ofrecen, utilizando los valores de confianza. Según los resultados mostrados en las figuras 7.7.a, 7.8.a y 7.9.a), la diferencia entre la evolución de la confianza en los agentes de calidad *Alta* y la evolución de la confianza en los agentes de calidad *Media*, cuando se selecciona la tarea aleatoriamente, es mayor para la alternativa Q_b .

En los siguientes experimentos encaminados a estudiar el comportamiento del modelo, no se considerarán los casos en que se presentan tareas de exigencia *Baja*. Las tareas de exigencia *Baja* pueden ser satisfechas eficientemente por cualquier tipo de respuesta. Con la selección aleatoria de la respuesta se pueden obtener valores elevados de satisfacción (muy cercanos a 1), similares a los obtenidos con las soluciones propuestas por cada una de las alternativas del modelo de confianza y reputación.

7.2.2. Elección de la alternativa adecuada para la combinación de la confianza y la reputación con la similitud

Cuando los agentes no tienen ninguna experiencia relacionada con el comportamiento de otros para una tarea específica, necesitan estimar los valores de confianza y reputación. En este sentido, el modelo propone combinar los valores de confianza y reputación asociados a una especificación de tarea bien conocida con el grado de similitud entre ésta y la desconocida. Se utiliza la función D para determinar el grado de similitud entre dos tareas (consultar la sección 5.8.2.1).

Los experimentos que se muestran en este apartado analizan el comportamiento del modelo ante diferentes alternativas para llevar a cabo la combinación de los valores de confianza y reputación con la similitud entre tareas. Los experimentos se agrupan fundamentalmente en los siguientes grupos:

- Comparación de varias alternativas para la combinación de los valores de confianza y reputación con la similitud entre tareas, con el objetivo de determinar la mejor forma para estimar la confianza y la reputación a partir del conocimiento disponible sobre tareas similares. Las alternativas consideradas se relacionan en el apartado 5.7.1.
- Estudio del comportamiento del modelo utilizando diferentes funciones para obtener la similitud entre tareas.
- Estudio de la capacidad del modelo para diferenciar los tipos de agentes (según la calidad de la respuesta que ofrecen) utilizando los valores de confianza al estimarlos a partir de tareas similares.

La función utilizada para determinar la calidad de la respuesta a partir de la tarea es la dada por la alternativa Q_b , propuesta como resultado del experimento anterior. Esta alternativa tiene en cuenta la cantidad de atributos de la tarea que son satisfechos por la respuesta.

Los resultados experimentales que se presentan en este grupo de experimentos se agrupan según la exigencia (*Alta* o *Media*) de la tarea que especifica los requerimientos del agente iniciador en cada instante de tiempo. Se han utilizado solo estos dos tipos de tareas por cuanto en los experimentos del apartado 7.2.1 se evidencia que en estos casos es donde el modelo de confianza y reputación resulta realmente útil (es decir, es muy fácil seleccionar agentes que ofrezcan soluciones con valores de satisfacción elevados para resolver tareas de exigencia *Baja*).

7.2.2.1. Determinación de la mejor alternativa para combinar la confianza y la reputación

Este experimento pretende comparar las diferentes alternativas para la estimación de la confianza y la reputación cuando el modelo carece de información suficiente relacionada con la tarea a resolver. Las alternativas utilizadas para la combinación de la similitud entre tareas con la confianza y la reputación son descritas en el apartado 5.7.1.

La figura 7.10 muestra la calidad (grado de satisfacción) de la solución recomendada por el agente más confiable, para cada alternativa utilizada en la estimación de la confianza y la reputación a partir del conocimiento asociado a tareas similares. Se observan diferencias entre la evolución de los valores de satisfacción según cada alternativa de estimación. Estas diferencias son menores para las tareas de exigencia *Media*. Por otra parte, algunas alternativas de estimación no son capaces de ofrecer valores de satisfacción más altos que los ofrecidos cuando no se considera estimación alguna.

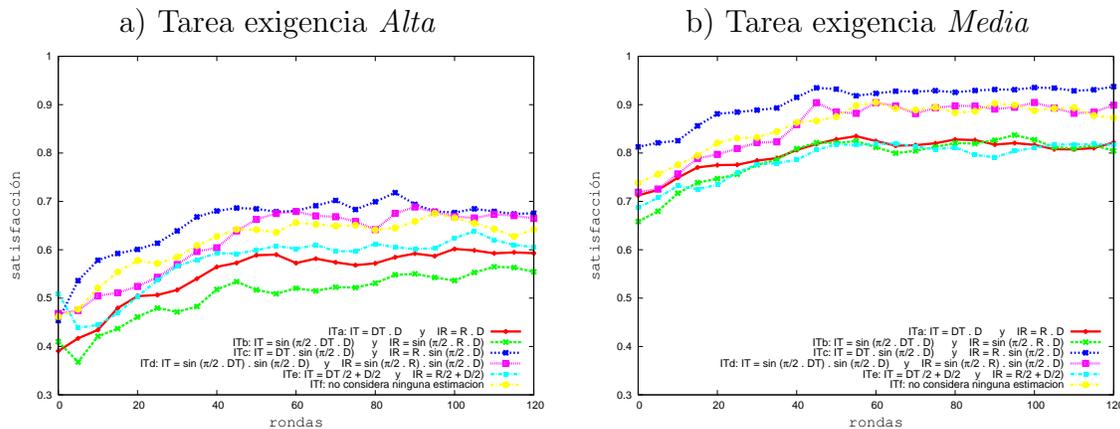


Figura 7.10: Comparación del nivel de satisfacción con la solución recomendada por el modelo, para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas (la definición de las alternativas se ofrece en el apartado 5.7.1), según la exigencia de la tarea a resolver en cada ronda: a) alta, b) media.

Por su parte, la figura 7.11 muestra la evolución del error cometido por el modelo al seleccionar la solución adecuada según la especificación de tarea dada.

Para ambos tipos de tareas (de exigencia *Alta* y de exigencia *Media*), los mejores resultados se obtienen utilizando la alternativa IT_c , donde la confianza y la reputación

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

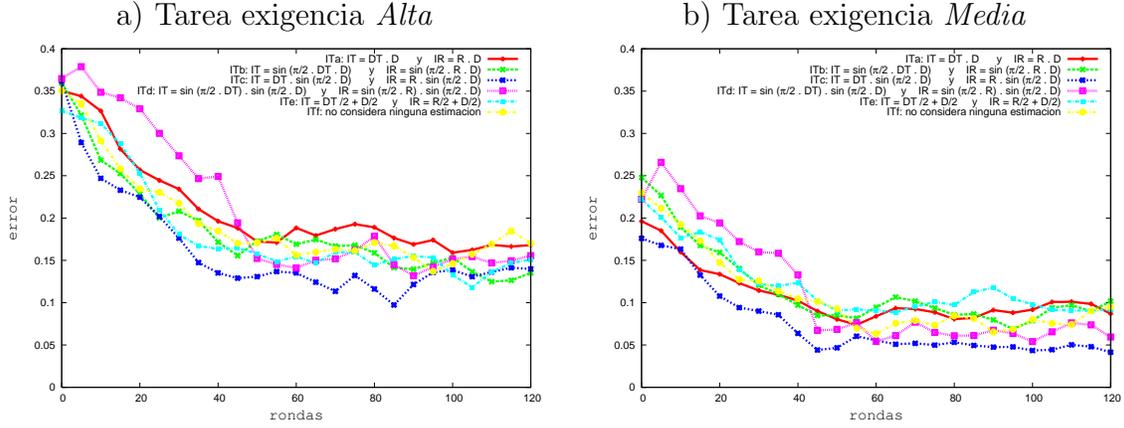


Figura 7.11: Comparación del error cometido por el modelo al seleccionar la solución adecuada, para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas (la definición de las alternativas se ofrece en el apartado 5.7.1), según la exigencia de la tarea a resolver en cada ronda: a) alta, b) media.

se estiman utilizando las ecuaciones siguientes:

$$IT^{(t)}(a_i, a_j, s_k, s_p, IET_i^{(t)}) = DT^{(t)}(a_i, a_j, s_p, IET_i^{(t)}) \cdot \sin\left(\frac{\pi}{2}D(s_k, s_p)\right)$$

y

$$IR^{(t)}(a_i, a_j, s_k, s_p, CR_{sup}^{(t)}(a_i, s_p)) = R^{(t)}(a_i, a_j, s_p, CR_{sup}^{(t)}(a_i, s_p)) \cdot \sin\left(\frac{\pi}{2}D(s_k, s_p)\right).$$

Con esta variante, el nivel de satisfacción con la solución recomendada es mayor que los ofrecidos cuando no se realiza aproximación alguna. Este resultado se ha obtenido gráficamente al comparar las áreas entre ambas curvas para los casos en que se solucionan tareas de exigencia *Alta* y de exigencia *Media* para esta alternativa (IT_c), y cuando no se considera ninguna aproximación para tareas desconocidas (alternativa IT_f). El área de la diferencia entre las curvas para cada caso se muestra en la tabla 7.2. Estos valores han sido calculados teniendo en cuenta la evolución de la satisfacción desde la ronda inicial $t = 0$ hasta la última ronda de la simulación $t = 400$. (Nótese que la figura 7.10 solo muestran la evolución de la satisfacción hasta la ronda $t = 120$.)

Tabla 7.2: Áreas entre las curvas que describen los grados de satisfacción para las alternativas IT_c e IT_f , al estimar la confianza y la reputación para los tipos de tareas analizados: de exigencia *Alta* y de exigencia *Media* (desde $t = 0$ hasta $t = 400$).

	Tarea exigencia <i>Alta</i>	Tarea exigencia <i>Media</i>
Diferencia entre IT_c e IT_f	2.69	10.76

La satisfacción con la calidad de la solución recomendada utilizando la alternativa IT_c es mayor que la que se obtiene cuando no se considera ninguna aproximación para

tareas desconocidas (alternativa IT_f) para ambos tipos de tareas. Las diferencias son más acentuadas cuando la tarea a resolver es de exigencia *Media*.

A partir de estos resultados, se demuestra que la estimación adecuada de la confianza y la reputación asociada a tareas desconocidas, utilizando el conocimiento disponible sobre otras tareas, mejora el proceso de selección de la solución más apropiada.

7.2.2.2. Independencia de la función de similitud

Los experimentos anteriores evidencian que, cuando se carece de información relacionada con la especificación de tarea a resolver, resulta provechoso utilizar convenientemente el conocimiento relacionado con alguna tarea similar. Se demuestra que, utilizando la alternativa IT_c (descrita en el apartado 5.7.1), se aumenta el nivel de satisfacción con la respuesta propuesta y se reduce el error cometido por el modelo cuando se combina la confianza y la reputación asociados a una tarea similar bien conocida con el grado de similitud entre ésta y la tarea desconocida a la que se quiere dar respuesta.

Los resultados experimentales, que comparan las alternativas para la combinación de la confianza y la reputación con la similitud, consideran que la similitud entre las tareas se calcula a partir de las diferencias entre los atributos de ambas tareas. La definición de esta medida se ofrece en la alternativa D_d , descrita en el apartado 5.8.2.1. Sin embargo, resulta factible analizar si el comportamiento del modelo, ante diferentes alternativas para la combinación de la confianza y la reputación con la similitud, depende de la forma en que se obtenga la similitud entre las tareas.

Así, el objetivo principal de este experimento es analizar el comportamiento del modelo al combinar la confianza y la reputación con la similitud, utilizando diferentes variantes para la obtención de la similitud entre tareas. Las alternativas que se plantean para el cálculo de la similitud se relacionan en el apartado 5.8.2.1.

De igual manera que en el experimento anterior, se analizan los casos en que la tarea a resolver es de exigencia *Alta* o *Media*.

Las gráficas de la figura 7.13 reflejan el comportamiento del modelo según las diferentes alternativas para combinar la confianza y la reputación con la similitud, para algunas de las variantes utilizadas para calcular la similitud entre tareas al resolver los dos tipos de tareas analizados.

Este experimento muestra que el modelo se comporta de manera similar al utilizar diferentes formas para calcular la similitud, para ambos niveles de exigencia de la tarea a resolver.

En todos los casos, la mejor alternativa para la combinación de la confianza y la reputación con la similitud entre tareas sigue siendo IT_c . Los niveles de satisfacción con la respuesta recomendada por el modelo son superiores a los obtenidos cuando se utilizan otras alternativas o cuando no se estima la confianza y la reputación a partir de tareas similares. Aunque IT_c produce mayores valores de satisfacción en todas las configuraciones analizadas, las diferencias con otras alternativas es mayor para los casos en que se pretende solucionar una tarea de exigencia *Media*, al utilizar cualquier función para el cálculo de la similitud entre tareas.

Desde otro punto de vista se observa que, para todas las funciones, que se han empleado para calcular la similitud entre tareas, el modelo se comporta de manera similar para todas las alternativas. En todos los casos, la alternativa IT_c es la que mayor satisfacción produce, IT_d muestra un comportamiento muy similar a IT_f (cuando no se realiza estimación alguna) y el resto de alternativas (IT_a , IT_b e IT_e) producen valores de satisfacción más bajos que IT_f .

Se puede concluir con este experimento, que el comportamiento del modelo ante diferentes alternativas para estimar la confianza y la reputación a partir de tareas similares es independiente de la definición de la función de similitud.

7.2.2.3. Diferenciación de los grupos de agentes según la calidad de la respuesta que ofrecen

A continuación se estudia la evolución de la confianza de los agentes según la mejor de las alternativas consideradas en los experimentos anteriores (IT_c). De esta forma, se pretende analizar la capacidad del modelo para distinguir la calidad de la respuesta de los agentes, utilizando los valores de confianza, cuando se utiliza la alternativa IT_c para estimar la confianza y la reputación a partir de tareas similares.

Para ambos niveles de exigencia de la tarea, se analiza la evolución de la confianza en cada tipo de agente, según la alternativa IT_c .

La figura 7.12 muestra la evolución de la confianza en los tres grupos de agentes de acuerdo a la calidad de la respuesta que ofrecen, usando la alternativa IT_c para estimar la confianza y la reputación a partir de tareas similares, según el nivel de exigencia de la tarea a resolver.

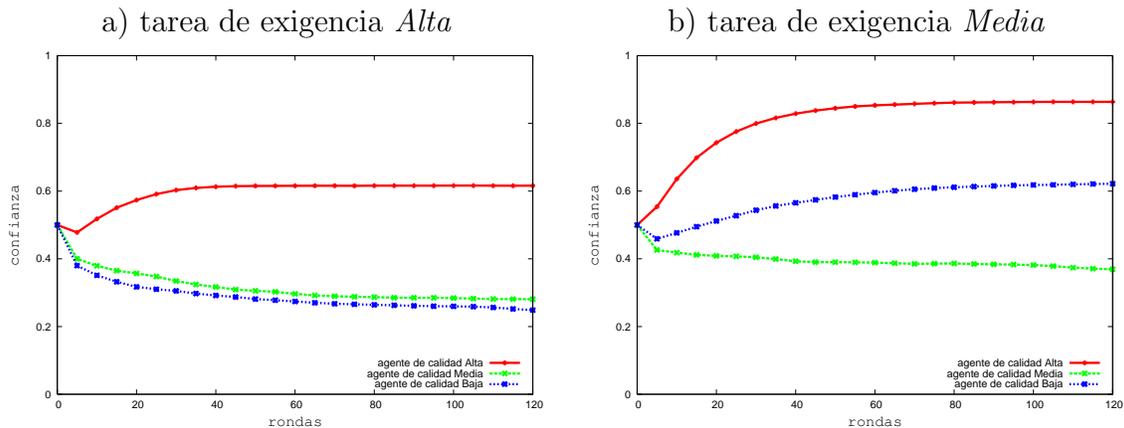
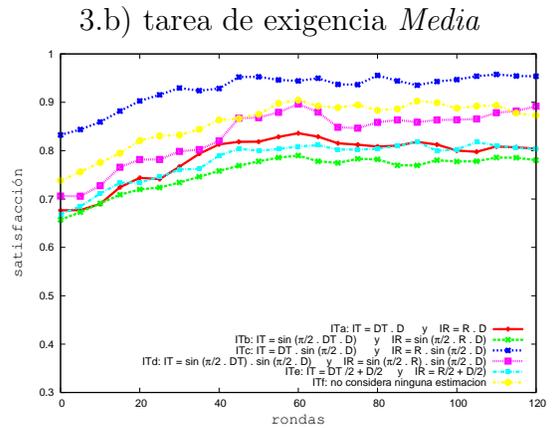
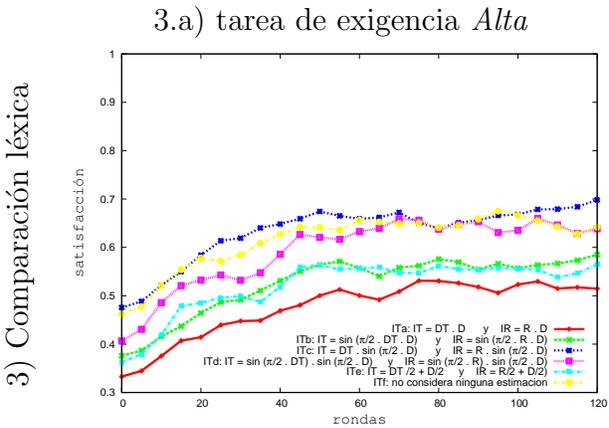
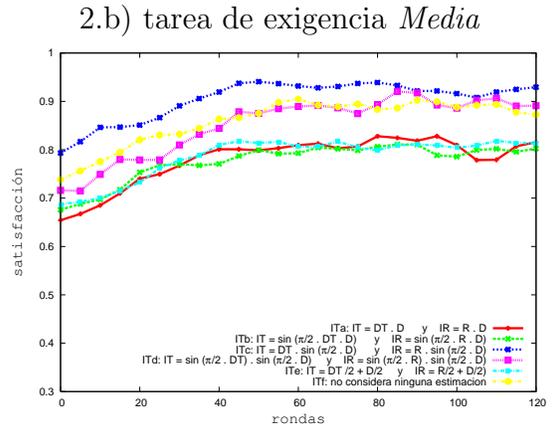
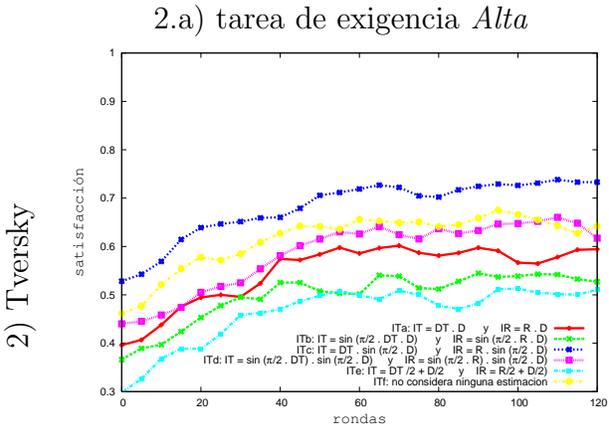
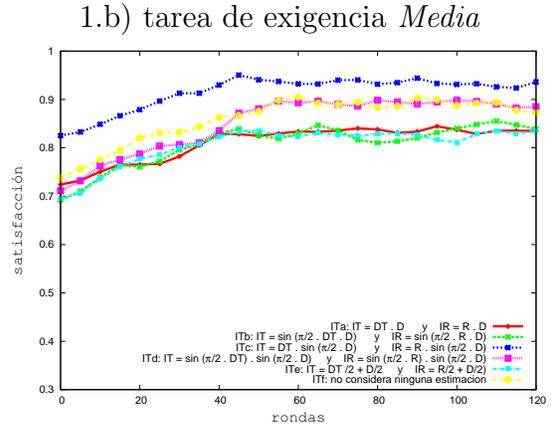
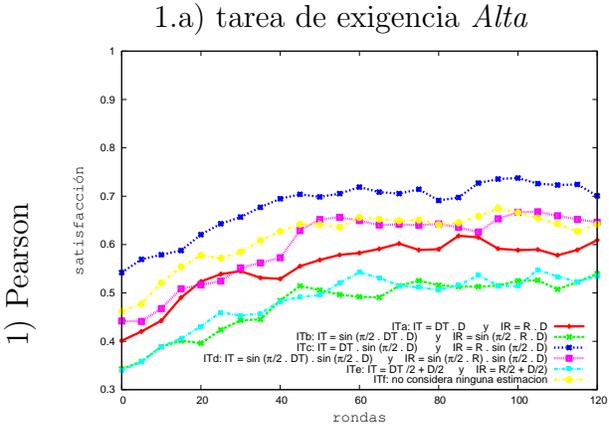


Figura 7.12: Evolución de la confianza para cada grupo de agentes, utilizando la alternativa IT_c para estimar la confianza y la reputación a partir de tareas similares, según el tipo de tarea a resolver en cada ronda a) de exigencia *Alta*, b) de exigencia *Media*.

Para ambos tipos de tareas, la confianza en los agentes que ofrecen solución de calidad *Alta* aumenta significativamente; y la confianza en los que ofrecen soluciones de calidad



... continúa en la página siguiente ...

7.2. ALTERNATIVAS PARA LA DEFINICIÓN DE ALGUNAS FUNCIONES IMPORTANTES DEL MODELO

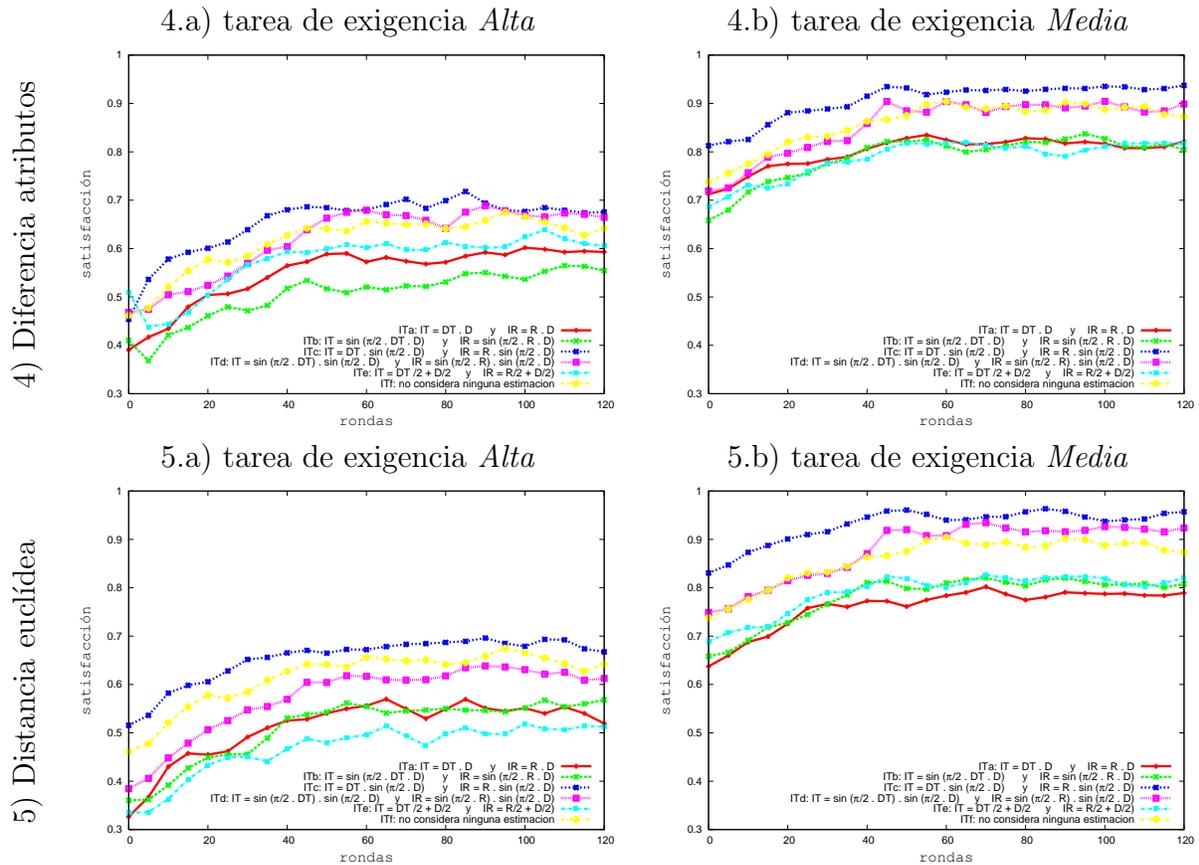


Figura 7.13: Comparación del grado de satisfacción, según la exigencia de la tarea en cada ronda (a) *Alta*, b) *Media*), para cada alternativa al combinar la confianza y la reputación con la similitud entre tareas, utilizando diferentes funciones para el cálculo de la similitud 1) Pearson, 2) Tversky, 3) Comparación léxica, 4) Diferencia entre los atributos, 5) Distancia euclídea (las alternativas se detallan en el apartado 5.8.2.1). Las gráficas correspondientes a las funciones 1) Pearson, 2) Tversky, 3) Comparación léxica se muestran en la página anterior.

Baja decrece. Por su parte, la confianza en los que ofrecen soluciones de calidad *Media* se comporta de manera diferente según la exigencia de la tarea a resolver. Frente a tareas de exigencia *Media*, los valores aumentan ligeramente; y frente a las de exigencia *Alta* se muestra un decrecimiento. Sin embargo, frente a ambos tipos de tareas, los valores de confianza en agentes que ofrecen soluciones de calidad *Media* se mantiene entre los valores de confianza correspondientes a los que ofrecen soluciones de calidad *Alta* y los que ofrecen soluciones de calidad *Baja*, para ambos tipos de tarea.

Para los agentes que ofrecen soluciones de calidad *Alta*, la confianza evoluciona desde un valor inicial (establecido por el modelo al inicio de la simulación) hasta un valor más cercano a 1. Esto significa que el modelo de confianza y reputación reconoce que los agentes de este tipo ofrecen buenas soluciones y, en consecuencia, la confianza en ellos aumenta.

Por otra parte, el modelo mantiene los valores de confianza alrededor del valor inicial para aquellos agentes que ofrecen soluciones de calidad *Media*. En cambio, la confianza en agentes cuyas soluciones son de calidad *Baja* disminuye ligeramente. El modelo penaliza las malas soluciones con la disminución de la confianza en los agentes que las ofrecen.

Los juicios anteriores indican que al estimar la confianza y la reputación para tareas desconocidas, utilizando la alternativa IT_c , no se afecta la evolución de la confianza de los agentes, independientemente del nivel de exigencia de la tarea. En otras palabras, las variaciones en la confianza de un agente están determinadas por la calidad de la solución que cada agente ofrece, no por la forma en que se combina la confianza y la reputación con la similitud entre las tareas. Sin embargo, al igual que en los casos donde no se realiza estimación alguna, la diferenciación de los valores de confianza para cada grupo de agentes es más destacada en aquellos casos donde la exigencia de la tarea es menor.

La figura 7.12 corrobora la capacidad del modelo (aun estimando la confianza y la reputación a partir de tareas similares) para representar la calidad de la solución que ofrecen los agentes utilizando los valores de confianza, mostrando un comportamiento diferenciado de acuerdo a cada tipo de agente.

7.2.3. Conclusiones del estudio de diferentes alternativas

A pesar de que todas las variantes analizadas para obtener la calidad de una respuesta, dada la tarea que satisface, muestran un comportamiento similar, la alternativa Q_b es la que garantiza la mejor diferenciación de los valores de confianza en los agentes de calidad *Alta* del resto de los agentes.

El modelo incrementa el grado de satisfacción con la respuesta recomendada cuando se estiman, de forma adecuada, los valores de confianza y reputación para especificaciones de tareas desconocidas, con respecto a los casos donde no se realiza ninguna estimación. IT_c constituye la mejor alternativa para aproximar estos valores a partir del conocimiento disponible sobre alguna tarea similar.

El problema de aproximar la confianza y reputación para tareas desconocidas es independiente de la forma de obtención de la similitud entre tareas y sí del procedimiento utilizado para combinar la similitud y los valores de confianza y reputación sobre alguna tarea conocida.

El modelo es capaz de mostrar comportamientos diferentes de acuerdo a la calidad de la respuesta que ofrecen los diferentes tipos de agentes. En todas las alternativas analizadas, la confianza en los agentes que ofrecen soluciones de calidad *Alta* tiende a crecer, la confianza en los que ofrecen soluciones de calidad *Baja* tiende a disminuir, y la confianza en los que ofrecen soluciones de calidad *Media* se mantiene en torno a un valor medio.

Algunos resultados relacionados con el contenido de bloque de experimentos se encuentran descritos en [25, 24, 23].

7.3. Adaptación ante los cambios en el comportamiento de los agentes proveedores

El grupo de experimentos, que se presenta en esta sección, está encaminado a estudiar la capacidad de adaptación del modelo TRSIM ante cambios en el comportamiento de algunos tipos de agentes. Este tipo de experimento es interesante por cuanto permite saber cómo y cuándo el modelo es capaz de recuperar su estado de equilibrio tras variaciones concretas en el comportamiento de algunos agentes. Para ello, tras la estabilización del modelo se imponen algunos cambios en la calidad de la respuesta que ofrecen los agentes y se analiza cómo reacciona el modelo ante dichos cambios. Los principales experimentos están relacionados con las siguientes situaciones:

- Un agente bueno empeora su comportamiento.
- Un agente malo mejora su comportamiento.
- Un agente bueno empeora su comportamiento y vuelve a mejorarlo.
- Un agente malo mejora su comportamiento y vuelve a empeorarlo.
- Un grupo de agentes buenos empeoran su comportamiento.
- Un grupo de agentes buenos empeoran su comportamiento y vuelven a mejorarlo.
- Un grupo de agentes malos mejoran su comportamiento.
- Un grupo de agentes malos mejoran su comportamiento y vuelven a empeorarlo.

En cada experimento se estudia la evolución del nivel de satisfacción con la solución recomendada por el modelo según la tarea a resolver en cada ronda, comparándola con los casos en que no se provoca cambio alguno en el comportamiento de los agentes. Como en otros experimentos, se muestra la evolución del error cometido por el modelo al seleccionar la solución adecuada. A partir de la evolución de estas medidas de calidad se analiza el momento en que el modelo vuelve a alcanzar el equilibrio, es decir, se adapta, tras un cambio de comportamiento de los agentes.

Por otra parte, se utilizan los valores de confianza del agente iniciador sobre el agente que cambia el comportamiento para estudiar la capacidad del modelo para reflejar los cambios de comportamiento de los agentes.

7.3.1. Modelado de los cambios de comportamiento

Los cambios de comportamiento en los agentes están determinados por el aumento o disminución de los valores de los atributos de las respuestas que los agentes ofrecen. Para cada uno de los cambios propuestos se definen funciones que modifican los valores de los atributos de la respuesta del agente.

Se considera que los cambios de comportamiento de los agentes se producen gradualmente, a lo largo de 10 ó 20 rondas (en dependencia del tipo de cambio) y que tienen lugar después de que el modelo ha alcanzado un estado de equilibrio, a partir de la ronda 200. Todas las curvas, que se ofrecen para analizar la capacidad de adaptación del modelo a cambios en el comportamiento de los agentes, grafican los indicadores del desempeño del modelo a partir de un instante de tiempo anterior al que se aplica el cambio de comportamiento en los agentes, desde la ronda $t = 190$, hasta la ronda $t = 250$.

La figura 7.14 representa las variaciones de los atributos de las respuestas de un agente que modifica su comportamiento a lo largo del cambio. Cada gráfica se corresponde con las variaciones de comportamiento estudiadas en los experimentos que se tratarán a continuación.

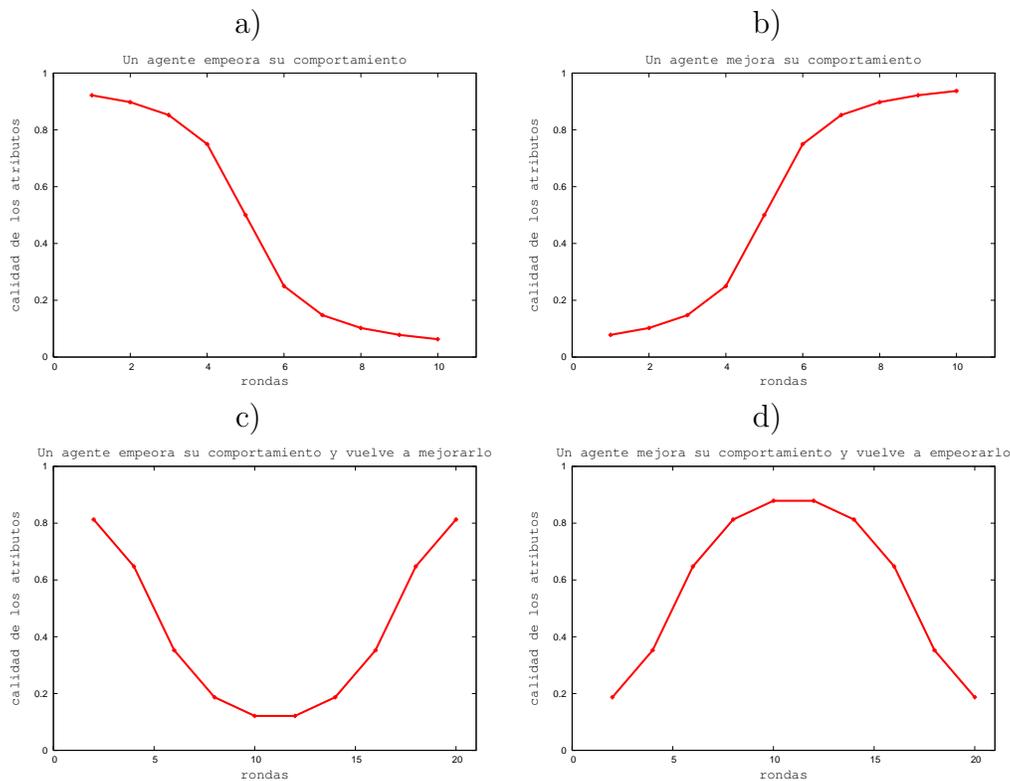


Figura 7.14: Tipos de variaciones del comportamiento de los agentes: a) un agente empeora su comportamiento, b) un agente mejora su comportamiento, c) un agente empeora su comportamiento y luego vuelve a mejorarlo, d) un agente mejora su comportamiento y luego vuelve a empeorarlo.

La figura 7.14.a muestra cómo empeora en el tiempo la calidad de una respuesta: los valores de sus atributos van disminuyendo gradualmente. En cambio, la figura 7.14.b muestra el mejoramiento de los atributos de una respuesta: los valores de sus atributos van aumentando gradualmente. Estos dos tipos de variaciones tienen lugar durante 10 rondas.

La figura 7.14.c muestra el empeoramiento y recuperación de los atributos de una respuesta: los valores de sus atributos van disminuyendo gradualmente y luego vuelven al valor que tenían inicialmente. En cambio, la figura 7.14.d muestra el mejoramiento y empeoramiento siguiente de los atributos de una respuesta: los valores de sus atributos van aumentando gradualmente y luego vuelven al valor que tenían inicialmente. Estos dos tipos de variaciones en los atributos transcurren durante 20 rondas. Son el resultado de realizar dos variaciones consecutivas a partir de la combinación de las variaciones más simples, definidas en las figuras 7.14.a y 7.14.b.

Luego, los cambios del comportamiento que experimentan los proveedores quedan descritos por alguno de los cuatro cambios definidos a continuación:

- Disminución de la calidad de la respuesta.
- Aumento de la calidad de la respuesta.
- Disminución de la calidad de la respuesta y aumento posterior hasta la calidad inicial.
- Aumento de la calidad de la respuesta y disminución posterior hasta la calidad inicial.

La calidad de la respuesta está dada por los valores de sus atributos, así, la variación de la calidad de la respuesta se provoca modificando cada uno de sus atributos. Para modelar las variaciones en cada uno de los atributos se han propuesto las funciones que se presentan en los apartados siguientes, para cada tipo de cambio.

7.3.1.1. Disminución de la calidad de la respuesta

Si se considera que $s_{k_i}^{(0)}$ es el valor del atributo s_{k_i} en el momento en que se inicia el cambio de comportamiento, se puede obtener el valor de dicho atributo en un instante t de la siguiente manera:

$$s_{k_i}^{(t)} = h_a^{(t)} \cdot s_{k_i}^{(t-1)}$$

donde t es el instante de tiempo de simulación transcurrido desde el momento en que se comienza el cambio en la calidad del proveedor; y $h_a^{(t)}$ es la función que modula la variación de los valores de los atributos de una respuesta para este tipo de cambio de comportamiento. Esta función es la representada en la figura 7.14.a, definida de la siguiente manera:

$$h_a^{(t)} = \frac{\arctan\left(\frac{M}{2} - t\right)}{\pi} + \frac{1}{2}$$

donde M es la duración del cambio de comportamiento. (En este caso, $M = 10$ rondas.)

7.3.1.2. Aumento de la calidad de la respuesta

Al aumentar la calidad de la respuesta, el valor del atributo s_{k_i} en el instante t después de que se inicia el cambio de comportamiento, se puede obtener de la siguiente manera:

$$s_{k_i}^{(t)} = s_{k_i}^{(t-1)} + h_b^{(t)} \cdot (1 - s_{k_i}^{(t-1)})$$

donde t es el instante de tiempo de simulación transcurrido desde el momento en que se comienza el cambio en la calidad del proveedor; y $h_b^{(t)}$ es la función que modula la variación de los valores de los atributos de una respuesta para este tipo de cambio de comportamiento. Esta función es la representada en la figura 7.14.b, definida de la siguiente manera:

$$h_b^{(t)} = 1 - h_a^{(t)} = \frac{\arctan(t - \frac{M}{2})}{\pi} + \frac{1}{2}$$

donde M es la duración del cambio de comportamiento. (En este caso, $M = 10$ rondas.)

7.3.1.3. Disminución de la calidad de la respuesta y aumento posterior hasta la calidad inicial

En este tipo de cambios, los valores de los atributos disminuyen durante la primera mitad de la duración total del cambio ($M/2$). A partir de este momento, los valores aumentan hasta volver a alcanzar, al final del cambio, el valor inicial $s_{k_i}^{(0)}$. En estos casos la variación de los atributos se describe de la siguiente manera:

$$s_{k_i}^{(t)} = \begin{cases} h_c^{(t)} \cdot s_{k_i}^{(t-1)} & : t \leq \frac{M}{2} \\ s_{k_i}^{(M-t+1)} & : t > \frac{M}{2} \end{cases}$$

donde t es el instante de tiempo de simulación transcurrido desde el momento en que se comienza el cambio en la calidad del proveedor; y $h_c^{(t)}$ es la función que modula la variación de los valores de los atributos de una respuesta para este tipo de cambio de comportamiento. Esta función es la representada en la figura 7.14.c, definida de la siguiente manera:

$$h_c^{(t)} = \frac{\arctan(\frac{M}{4} - t)}{\pi} + \frac{1}{2}$$

donde M es la duración del cambio de comportamiento. (En este caso, $M = 20$ rondas.)

7.3.1.4. Aumento de la calidad de la respuesta y disminución posterior hasta la calidad inicial

En este tipo de cambios, los valores de los atributos aumentan durante la primera mitad de la duración total del cambio ($M/2$). A partir de este momento, los valores disminuyen hasta volver a alcanzar, al final del cambio, el valor inicial $s_{k_i}^{(0)}$. En estos casos la variación de los atributos se describe de la siguiente manera:

$$s_{k_i}^{(t)} = \begin{cases} s_{k_i}^{(t-1)} + h_d^{(t)} \cdot (1 - s_{k_i}^{(t-1)}) & : t \leq \frac{M}{2} \\ s_{k_i}^{(M-t+1)} & : t > \frac{M}{2} \end{cases}$$

donde t es el instante de tiempo de simulación transcurrido desde el momento en que se comienza el cambio en la calidad del proveedor; y $h_d^{(t)}$ es la función que modula la variación de los valores de los atributos de una respuesta para este tipo de cambio de comportamiento. Esta función es la representada en la figura 7.14.d, definida de la siguiente manera:

$$h_d^{(t)} = 1 - h_c^{(t)} = \frac{\arctan(t - \frac{M}{4})}{\pi} + \frac{1}{2}$$

donde M es la duración del cambio de comportamiento. (En este caso, $M = 20$ rondas).

7.3.2. Cambios individuales en el comportamiento de diferentes tipos de agentes

A continuación se ofrece un grupo de experimentos que persigue estudiar el desempeño del modelo cuando cambia el comportamiento de un agente de manera individual. Se pretende comprobar que el modelo, tras un cambio en el comportamiento de un agente de forma individual, es capaz de ajustar los valores de confianza en dicho agente y continuar ofreciendo en cada momento la mejor respuesta posible. Para ello, se considera que los cambios de comportamiento de los agentes se producen en un instante de tiempo donde el modelo ha alcanzado la estabilidad. Se evalúan los mismos indicadores de calidad que los experimentos mostrados anteriormente, analizando a partir de ellos la capacidad del modelo para alcanzar nuevamente un estado estable.

7.3.2.1. Un agente de calidad *Alta* empeora su comportamiento

La figura 7.15 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en un agente que ofrece soluciones de calidad *Alta*, comparándolo con el caso donde el agente mantiene su comportamiento constante. Se comparan los casos en que, en cada ronda, se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*. En esta figura se observa cómo disminuye el nivel de satisfacción con la solución que recomienda el sistema cuando se produce un cambio en el comportamiento de un agente de calidad *Alta*. Esta disminución es más acentuada para las tareas de exigencia *Alta*.

Para ambos tipos de tareas, el nivel de satisfacción disminuye hasta alcanzar los valores más bajos, aproximadamente en el instante en que finaliza el cambio de comportamiento. A partir de este momento, en que alcanza los valores mínimos, el nivel de satisfacción aumenta ligeramente. En ambos casos, este aumento no alcanza los valores de satisfacción que se tenían antes del cambio. Esto indica que los ajustes de los valores de confianza (utilizados como criterio para recomendar la respuesta en cada ronda) no se producen instantáneamente: el modelo debe interactuar con el agente que cambia durante un número de rondas para reconocer que su calidad ha cambiado. Durante el cambio, la disminución de la satisfacción es mayor por cuanto el agente que cambia el comportamiento goza de buena confianza (por lo que el modelo lo sigue recomendando) aun cuando su calidad va cambiando. De ahí que la satisfacción recomendada diste de la óptima. Sin embargo, tras

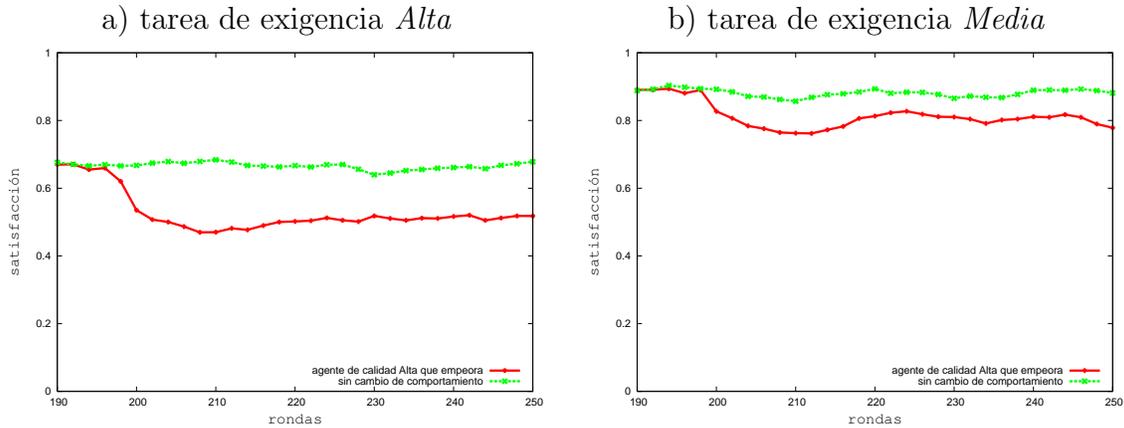


Figura 7.15: Comparación del nivel de satisfacción con la solución recomendada por el modelo según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento.

algunas interacciones, el modelo es capaz de reconocer que el agente ha cambiado y por tanto, seleccionar otro que, aunque no ofrezca tan buena satisfacción, sí garantiza un valor más alto que la del agente que ha modificado su calidad.

Este comportamiento del modelo, que se ha comentado anteriormente, se debe fundamentalmente a dos razones: (1) disminuye la cantidad de agentes de calidad *Alta* en el sistema, y por tanto la cantidad de soluciones que producen valores elevados de satisfacción, y (2) los cambios en el comportamiento de los agentes no se refleja inmediatamente en los valores de confianza, se necesita interactuar con los agentes para actualizar el conocimiento del modelo.

Por su parte, la figura 7.16 muestra la evolución del error cometido por el modelo al seleccionar la respuesta adecuada para la tarea a resolver en cada ronda, agrupadas según el nivel de exigencia de ésta. Para ambos tipos de tareas, el error cometido por el modelo aumenta cuando cambia el comportamiento de un agente de calidad *Alta*. Los valores del error aumentan, alcanzado su valor más alto en el instante en que finaliza el cambio de comportamiento del agente. Posteriormente, se estabilizan en un valor más bajo, aunque superior al error que se tenía antes de que el agente modificara la calidad de su respuesta.

Las variaciones del error cometido por el modelo se debe a las mismas razones comentadas anteriormente. El conjunto de agentes de calidad *Alta* en el sistema se reduce y por tanto, pueden recomendarse soluciones por parte de agentes que no ofrecen una solución tan buena como la que el agente que ha empeorado ofrecía antes del cambio. Por ejemplo, si antes de cambiar su comportamiento el agente que empeora es quien ofrece las soluciones de mayor calidad, luego del cambio, se recomiendan soluciones (por parte de otros agentes) de menor calidad. En estos casos, el error, cometido por el modelo al recomendar la respuesta, es mayor que antes del cambio de comportamiento del agente que ofrecía las mejores soluciones.

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

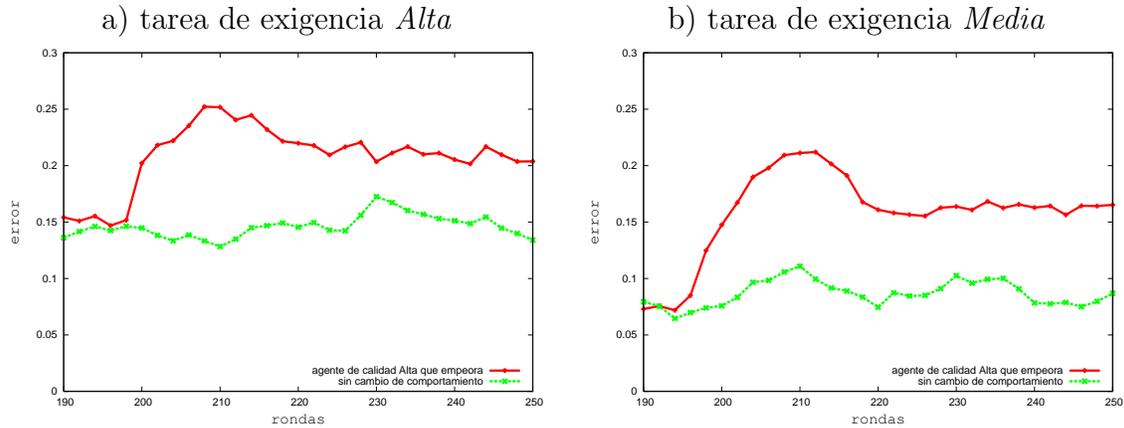


Figura 7.16: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento.

También se ofrece la figura 7.17 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que un agente de calidad *Alta* empeora su comportamiento.

Tanto para las tareas de exigencia *Alta* como para las tareas de exigencia *Media*, el nivel de satisfacción con la respuesta ofrecida por el modelo, el error cometido en la recomendación y la confianza en el agente que recomienda una solución se ven afectados tras empeorar el comportamiento de un agente de calidad *Alta*. El nivel de satisfacción con la respuesta (que se muestra en la figura 7.15) y la confianza promedio que se tiene en el agente recomendado por el modelo (que se muestra en la figura 7.17) disminuyen tras el empeoramiento de la calidad de un agente de calidad *Alta*. De igual forma, se observa un aumento del error de recomendación tras el cambio del comportamiento de este tipo de agente (Figura 7.16).

También se ofrece la figura 7.18 que muestra la variación de la confianza en el agente de calidad *Alta* que empeora su comportamiento, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. En este caso, el modelo disminuye la confianza que se tiene en el agente que empeora la calidad de la respuesta. La disminución de la confianza, que el agente iniciador tiene en el agente que cambia su comportamiento, es mayor en los casos en que se soluciona una especificación de tarea de exigencia *Alta* por cuanto se requiere de soluciones de mayor calidad para satisfacer este tipo de tareas.

Durante el cambio, e incluso algunas rondas siguientes, el modelo se desestabiliza,

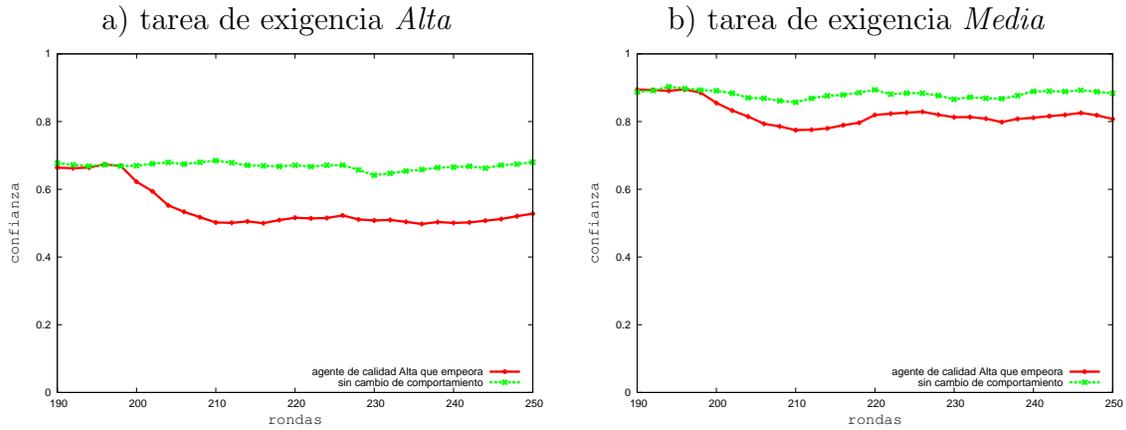


Figura 7.17: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento.

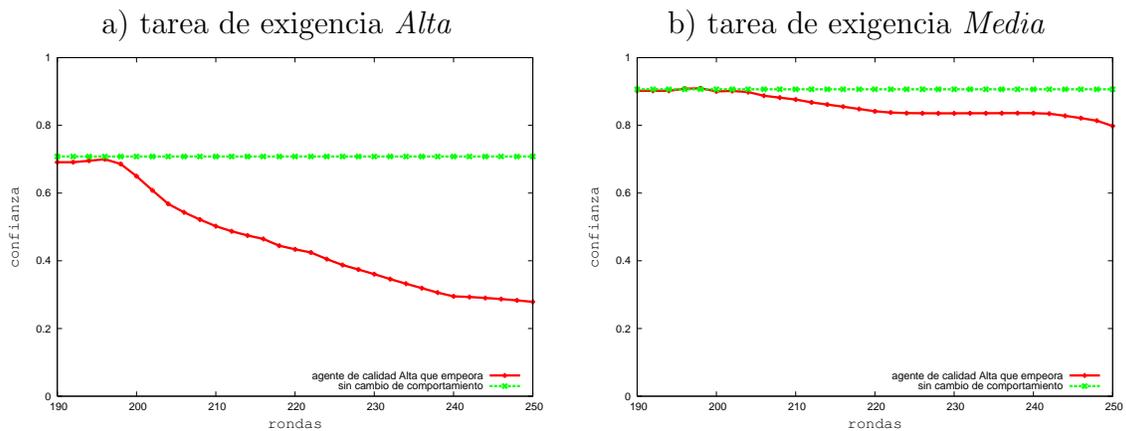


Figura 7.18: Evolución de la confianza del agente de calidad *Alta* que empeora su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

alcanzando nuevamente el equilibrio aun cuando el nuevo valor del error de estimación es mayor que el que garantizaba el modelo antes del cambio. Para ambos tipos de tareas, el modelo alcanza la estabilidad en la ronda $t = 218$, 8 rondas después de finalizado el cambio de comportamiento del agente. Estos resultados evidencian que el modelo es robusto a este tipo de cambios en el comportamiento de los agentes, volviendo a alcanzar un punto de equilibrio.

7.3.2.2. Un agente de calidad *Baja* mejora su comportamiento

Se espera que, ante este tipo de cambios en el comportamiento de los agentes, los indicadores utilizados se mantengan sin variaciones significativas en sus valores. Al mejorar el comportamiento de un agente de calidad *Baja* no se esperan aumentos significativos del nivel de satisfacción con la respuesta recomendada y de la confianza en el agente que la ofrece, ni una disminución del error cometido al recomendarla. Luego del cambio, el agente que varía su comportamiento ofrece una solución de calidad *Alta*, lo que no significa que los agentes que ya ofrecían soluciones de calidad *Alta* dejen de hacerlo. Por lo tanto, de manera general, el modelo seguirá recomendando soluciones de tan alta calidad como las ofrecidas por agentes confiables hasta el momento.

A pesar de que no se esperan variaciones en los indicadores de la satisfacción y del error cometido por el modelo, así como de la confianza en el agente que la recomienda, sí se esperan en la confianza que el agente iniciador otorga al agente que ha cambiado su comportamiento. El agente de calidad *Baja* que mejora su comportamiento debe aumentar su confianza frente a los agentes que necesitan resolver una tarea. Esto se debe a que, a partir del cambio, dicho agente puede ofrecer una solución que garantiza un elevado nivel de satisfacción para varios tipos de tareas.

La figura 7.19 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en un agente que ofrece soluciones de calidad *Baja*, comparándolo con el caso donde el agente mantiene su comportamiento constante. Por su parte, la figura 7.20 muestra la evolución del error cometido por el modelo al seleccionar la respuesta adecuada para la tarea a resolver en cada ronda. Ambas figuras ofrecen la evolución de las métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

También se ofrece la figura 7.21 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que un agente de calidad *Baja* mejora su comportamiento.

Estas tres figuras (7.19, 7.20 y 7.21) muestran que el modelo responde correctamente al mejoramiento del comportamiento de un agente de calidad *Baja*, manteniendo su estabilidad. Se mantienen los niveles de confianza en el agente que recomienda la solución y los del error cometido por el modelo al recomendar una solución. Por su parte, el nivel de satisfacción ofrecido por el modelo muestra un ligero aumento, para las tareas de exigencia *Media*, dado que al mejorar el comportamiento de un agente de calidad *Baja* aumenta el

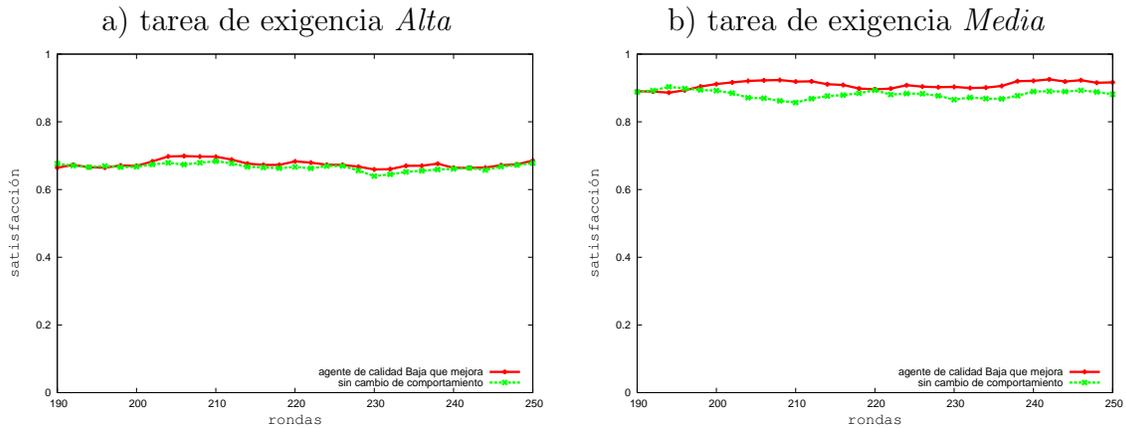


Figura 7.19: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento.

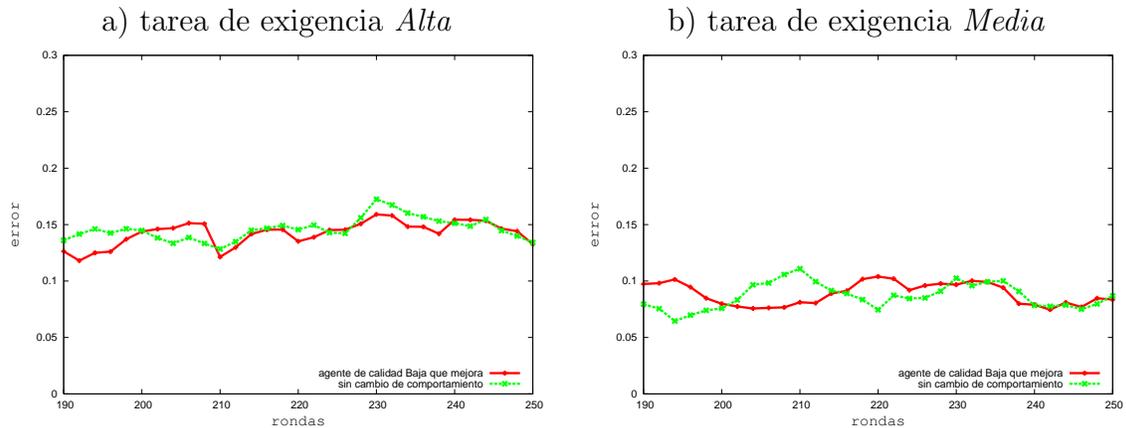


Figura 7.20: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento.

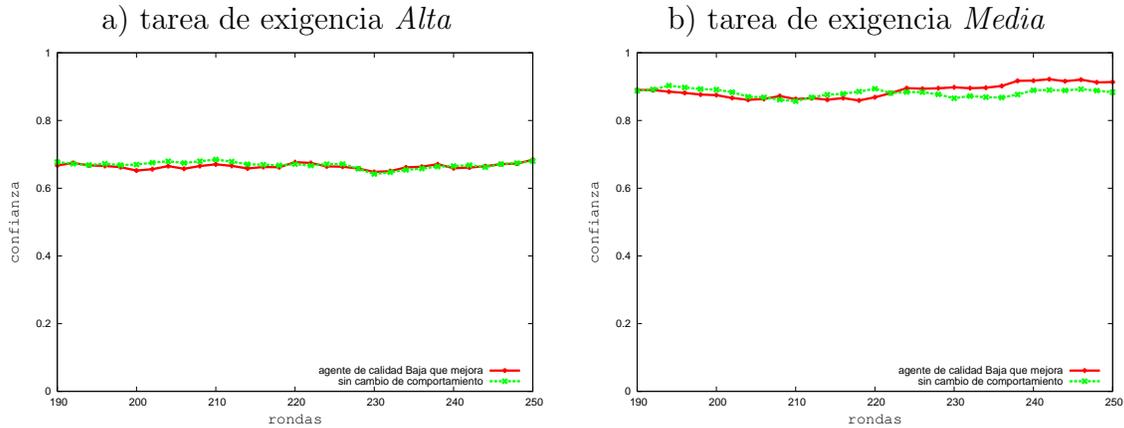


Figura 7.21: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento.

conjunto de agentes con buenas soluciones y con ello la posibilidad de que la respuesta de dicho agente, tras el cambio, sea de mejor calidad que las que se ofrecían hasta el momento por el resto de agentes buenos. El error y la confianza en el agente recomendado no varían, por cuanto el modelo es capaz de seguir seleccionando la solución más adecuada a partir de los valores de confianza en los agentes.

También se ofrece la figura 7.22 que muestra la variación de la confianza que el agente iniciador tiene en el agente de calidad *Baja* que mejora su comportamiento, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. Frente a ambos tipos de tareas, el modelo aumenta la confianza que tiene en el agente que mejora la calidad de la respuesta, hasta alcanzar valores similares a los que poseen otros agentes de calidad *Alta*. El aumento de los valores de confianza depende significativamente del tipo de tarea a resolver. Cuando la tarea es de exigencia *Media* los agentes alcanzan valores de confianza más elevados. Evidentemente, de manera general, un mismo agente es más confiable para la ejecución de una tarea de exigencia *Media* que para una tarea de exigencia *Alta*.

Frente a este tipo de cambio en el comportamiento de un agente, el modelo se mantiene estable en todo momento, siendo capaz de ajustar los valores de confianza en el agente que varía su desempeño.

7.3.2.3. Un agente de calidad *Alta* empeora su comportamiento y vuelve a mejorarlo

A continuación se ofrece el estudio del comportamiento del modelo ante situaciones en que la calidad de la respuesta, de un agente de buen desempeño, varía durante un período

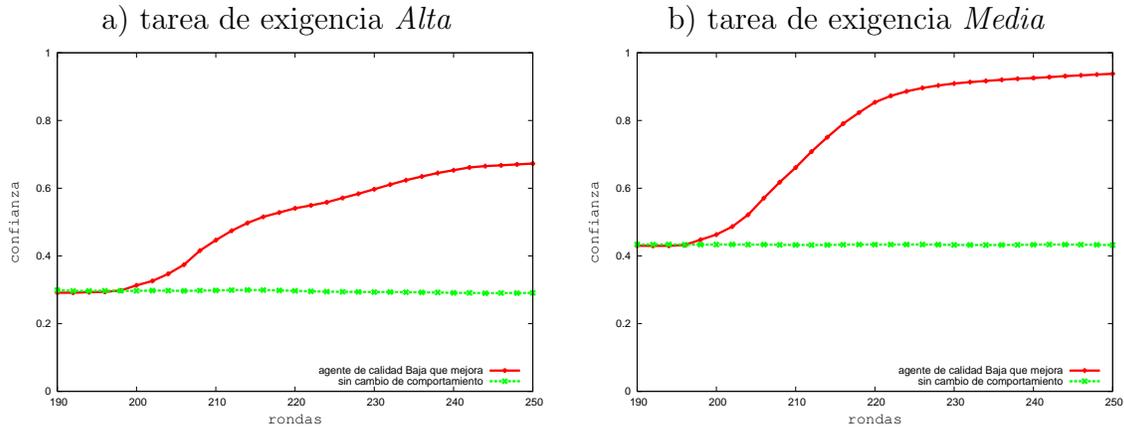


Figura 7.22: Evolución de la confianza del agente de calidad *Baja* que mejora su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*.

de tiempo determinado. Los cambios, en este caso, están dados por el empeoramiento, y posterior recuperación, de la respuesta que ofrece un agente de calidad *Alta*. Se pretende, con ello, analizar el comportamiento del modelo de confianza y reputación cuando un agente deja de ofrecer soluciones de buena calidad durante un tiempo determinado. Por ejemplo, en algunas situaciones prácticas, este cambio de calidad puede estar dado por un juicio equivocado del agente, que lo lleva a generar respuestas erradas (de baja calidad) durante un período de tiempo dado, tras el que es capaz de reajustar su comportamiento y volver a ofrecer respuestas de la misma calidad que ofrecía antes.

La figura 7.23 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un empeoramiento y recuperación de la calidad de un agente que ofrece soluciones de calidad *Alta*, comparándolo con el caso donde el agente mantiene su comportamiento constante. Por su parte, la figura 7.24 muestra la evolución del error cometido por el modelo al seleccionar la respuesta adecuada para la tarea a resolver en cada ronda. Las figuras ofrecen la evolución de ambas métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

También se ofrece la figura 7.25 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que un agente de calidad *Alta* empeora su comportamiento y vuelve a mejorarlo.

Para ambos tipos de tareas, las figuras 7.23 7.24 y 7.25 muestran como el modelo se adapta a este tipo de cambios en el comportamiento de un agente de calidad *Alta*. Durante el cambio, la satisfacción con la respuesta ofrecida por el modelo y la confianza en el agente que la recomienda disminuyen y se recuperan en el mismo sentido en que varía la calidad de la solución del agente que cambia. El error de recomendación de la solución adecuada aumenta tras empeorar la calidad del agente y va disminuyendo paulatinamente en tanto

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

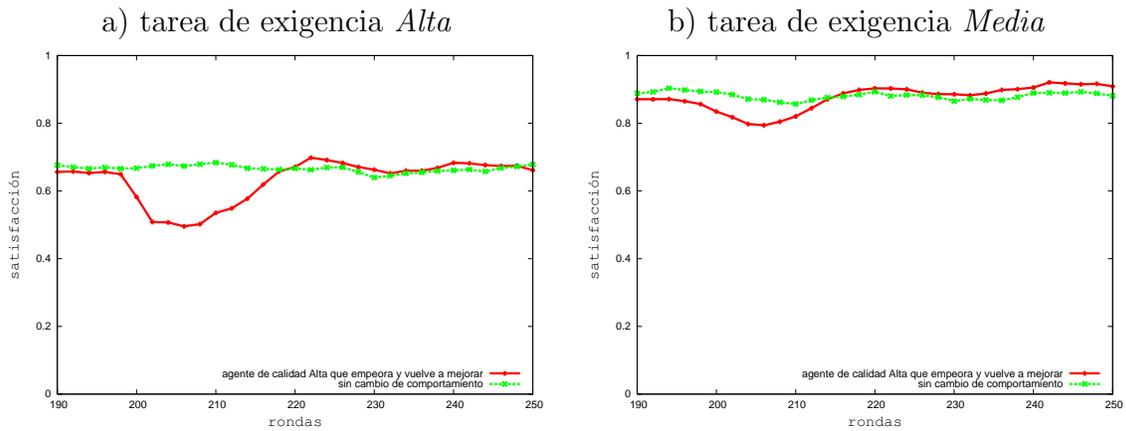


Figura 7.23: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento y vuelve a mejorarlo.

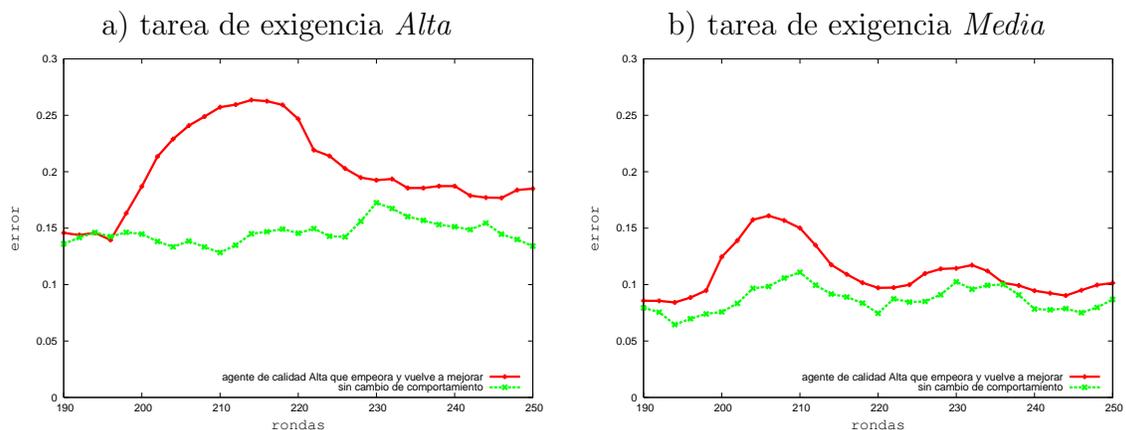


Figura 7.24: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento y vuelve a mejorarlo.

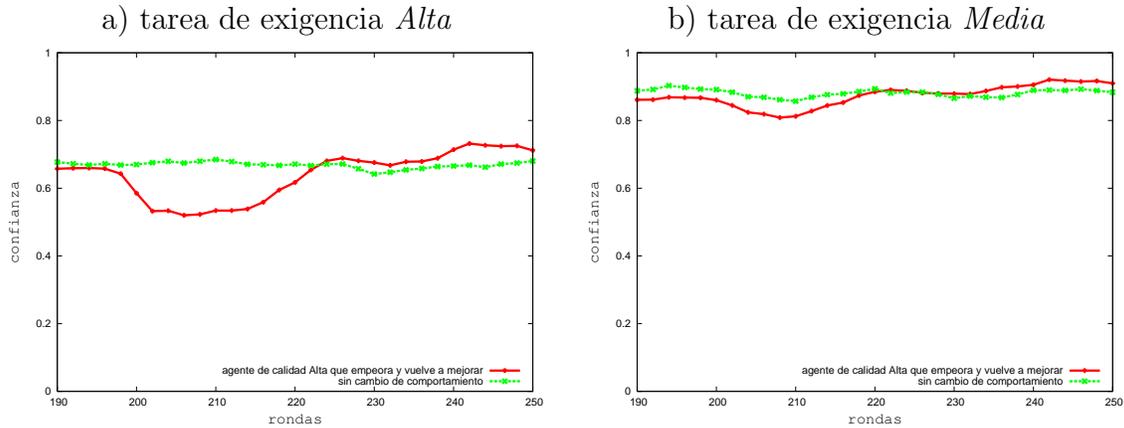


Figura 7.25: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Alta* empeora su comportamiento y vuelve a mejorarlo.

la calidad del agente se recupera. Las variaciones de estas medidas en el caso en que la tarea a resolver es de exigencia *Alta* son mayores que cuando se pretende resolver una tarea de exigencia *Media*.

La figura 7.26 muestra la variación de la confianza en el agente de calidad *Alta* que empeora su comportamiento y vuelve a mejorarlo, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. En este caso, durante la primera parte del cambio de comportamiento del agente (cuando éste empeora) el modelo disminuye la confianza que se tiene en él, pero en tanto comienza a recuperarse, la confianza asciende hasta alcanzar valores similares a los que tenía inicialmente.

Al igual que sucede con el resto de los indicadores graficados, la variación de la confianza en el agente que modifica su comportamiento es menor cuanto menor es la exigencia de la tarea a resolver en cada ronda.

Luego del cambio, tras un número de rondas relativamente pequeño, el modelo vuelve a alcanzar la estabilidad. Para las tareas de exigencia *Alta* el equilibrio se alcanza en el instante $t = 228$, mientras que para las tareas de exigencia *Media*, en $t = 224$.

7.3.2.4. Un agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo

La figura 7.27 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en un agente que ofrece soluciones de calidad *Baja*, comparándolo con el caso donde el agente mantiene su comportamiento constante. Por su parte, la figura 7.28 muestra la evolución del error

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

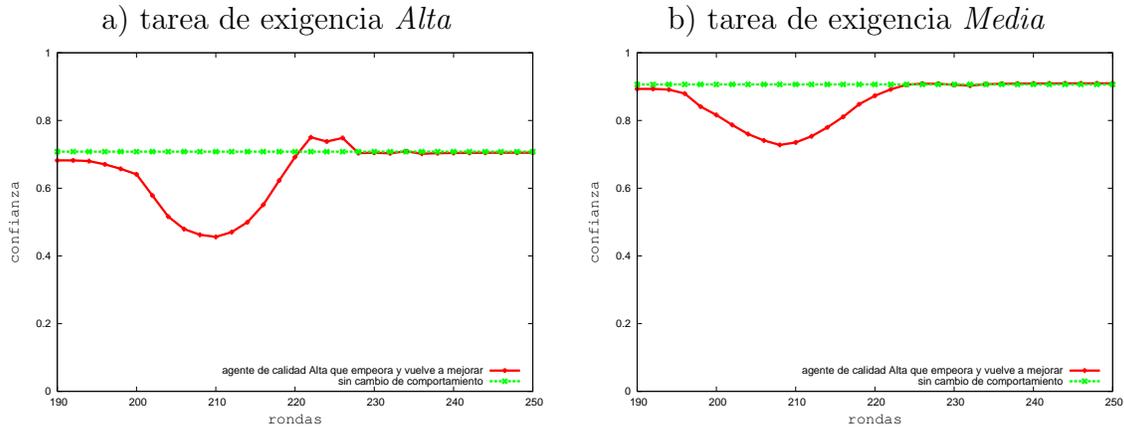


Figura 7.26: Evolución de la confianza del agente de calidad *Alta* que empeora su comportamiento y vuelve a mejorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

cometido por el modelo al seleccionar la respuesta a la tarea de cada ronda. Las figuras ofrecen la evolución de ambas métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

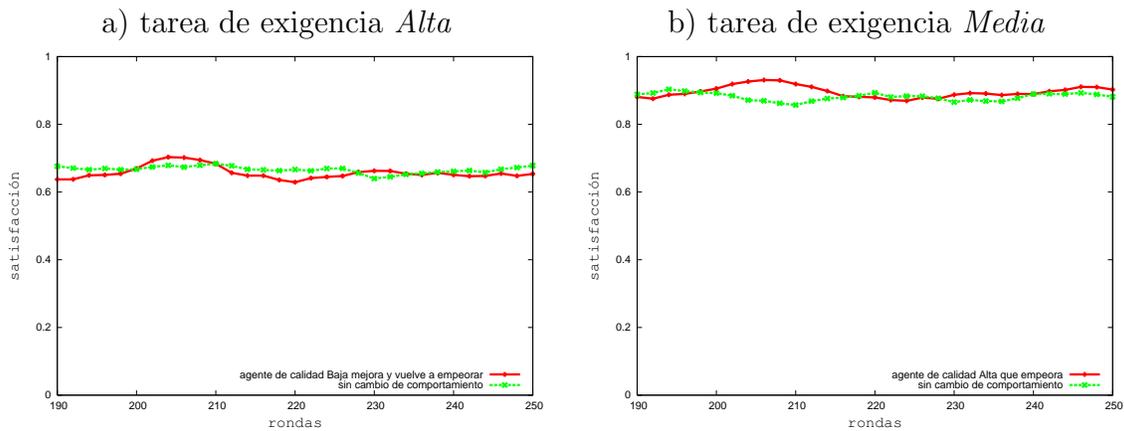


Figura 7.27: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo.

También se ofrece la figura 7.29 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que un agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo.

Estas tres figuras (7.27, 7.28 y 7.29) muestran que el modelo responde correctamente a

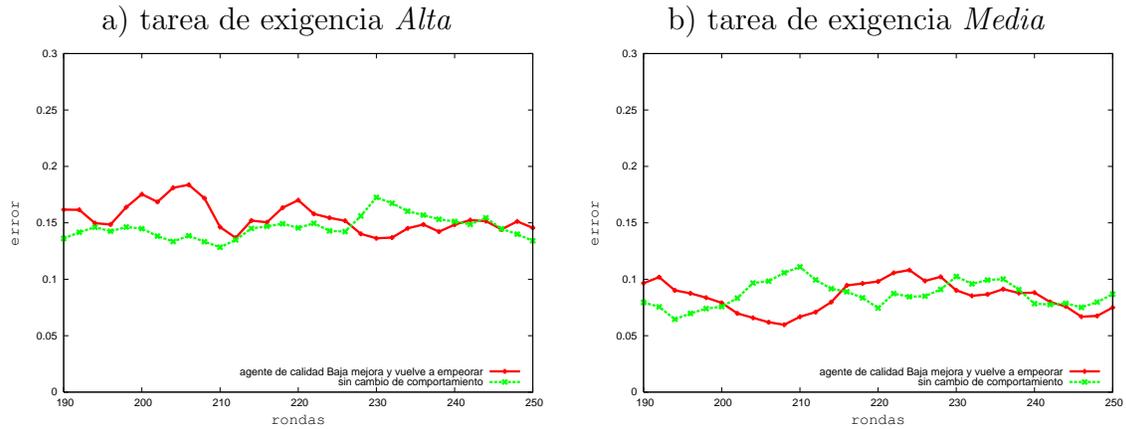


Figura 7.28: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo.

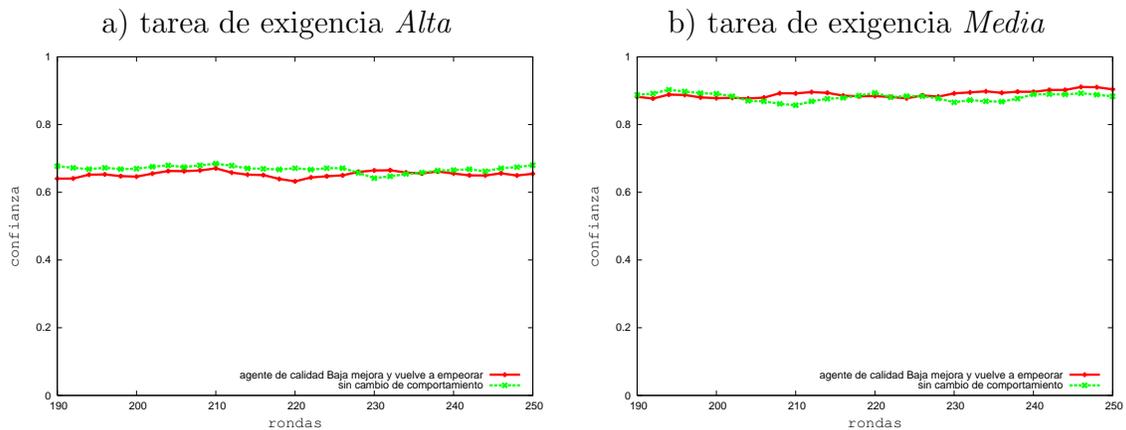


Figura 7.29: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando un agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo.

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

este tipo de cambio en el comportamiento de un agente de calidad *Baja*, manteniendo su estabilidad.

De manera general, los valores de las medidas del desempeño del modelo (satisfacción con la respuesta recomendada, error cometido y confianza en al agente respondedor) se mantienen sin grandes variaciones. Esto se debe en lo fundamental a que únicamente varía la calidad de la respuesta de un agente de calidad *Baja*. En el conjunto de agentes confiables que ofrecen respuestas de calidad elevada permanecen los mismos agentes que si no se hubiera producido cambio alguno.

Sin embargo, cuando la tarea es de exigencia *Media*, durante la primera parte del cambio de comportamiento, los niveles de confianza en la solución propuesta aumentan. Esto se debe a que la solución del agente que cambia puede llegar a garantizar mejores niveles de satisfacción que el resto de agentes. Luego de comenzar el descenso de la calidad del agente que cambia, se restablecen los valores iniciales de satisfacción. Un comportamiento similar muestra la evolución del error cometido. Cuando la calidad de la solución asciende el error cometido por el modelo sufre ligeras variaciones para ambos tipos de tarea.

Por otra parte, también se ofrece la figura 7.30 que muestra la variación de la confianza en el agente de calidad *Baja* que mejora su comportamiento y vuelve a empeorarlo, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. En este caso, durante la primera parte del cambio de comportamiento del agente (cuando éste mejora) el modelo aumenta la confianza que se tiene en él. Sin embargo, en tanto se comienza a degradar la calidad de la respuesta hasta su estado inicial, la confianza disminuye hasta alcanzar valores similares a los que tenía antes del cambio.

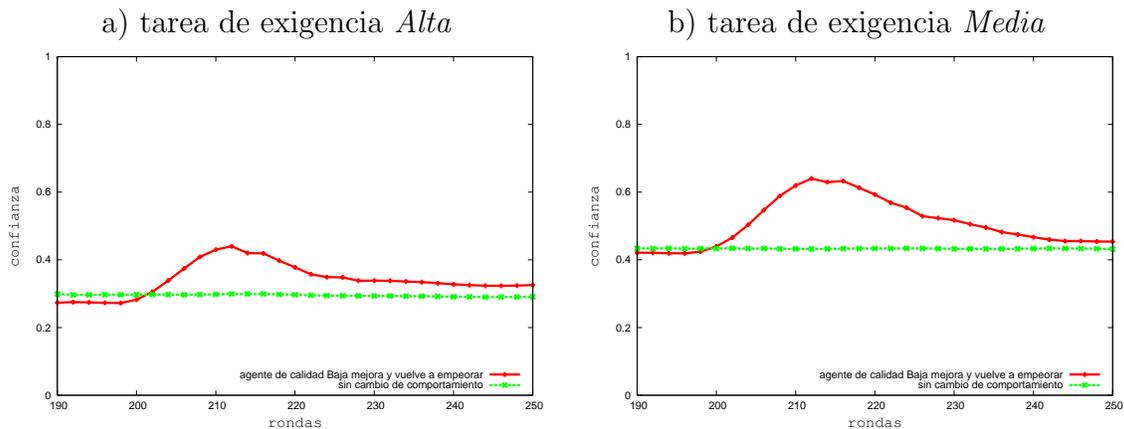


Figura 7.30: Evolución de la confianza del agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

Frente a este tipo de cambio en el comportamiento de un agente, el modelo se mantiene estable en todo momento, siendo capaz de ajustar los valores de confianza en el agente que

varía su desempeño.

La variación de la confianza en el agente que modifica su comportamiento es menor cuanto menor es la exigencia de la tarea a resolver en cada ronda.

7.3.3. Cambios en el comportamiento de un grupo de agentes que ofrece soluciones de calidad *Alta*

Muchas veces, en un sistema multi-agente los agentes se agrupan formando coaliciones con el objetivo de reunir los recursos de cada agente miembro de la coalición, y llevar a cabo sus actividades en conjunto. Las coaliciones se establecen de forma que dentro de cada grupo, los agentes coordinan sus actividades de manera diferente a como lo hacen con el resto. Por ejemplo, los agentes dentro de una misma coalición pueden coordinar sus actividades pero no hacerlo con agentes de otras coaliciones [4, 138].

El tratamiento diferenciado de la coordinación entre los miembros de una coalición, con respecto al resto del sistema, puede estar referido a la importancia que cada uno ofrece a las opiniones de los otros agentes dentro de la coalición. Así, si uno de los agentes tiene una apreciación (creencia) *no acertada* acerca del entorno, de sus vecinos, de sus acciones, etc., puede repercutir negativamente en la apreciación que hagan otros miembros del grupo.

Por ejemplo, en términos relacionados con el escenario utilizado para la experimentación, un miembro de la coalición podría ejecutar alguna acción que provoque, como perjuicio, la disminución de la calidad de todos los elementos de la coalición a la que pertenece. Dada la importancia y relativa frecuencia de este tipo de situaciones, donde se produce un cambio masivo en el comportamiento de los agentes, el presente grupo de experimentos trata de analizar la capacidad de TRSIM para adaptarse a ellas.

Los experimentos, cuyos resultados se muestran a continuación, consideran la composición de la población de agentes tal como se comentan en el apartado 7.1: el 30 % de los proveedores ofrecen respuestas de calidad *Alta*, el 30 % de calidad *Baja* y el 40 % de calidad *Media*.

Los experimentos suponen que todos los agentes que ofrecen soluciones de calidad *Alta* forman una coalición, y que, por una causa determinada, la calidad de las respuestas que ofrecen cada uno de ellos se ve afectada. Las situaciones que se estudian, a partir del cambio de comportamiento de agentes de calidad *Alta*, son las siguientes:

- todos los agentes de calidad *Alta* empeoran su comportamiento, y
- todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo.

Al igual que en experimentos anteriores, se muestran por separado las métricas sobre el comportamiento del modelo para los casos en que se resuelven tareas de exigencia *Alta* y *Media*.

7.3.3.1. Todos los agentes de calidad *Alta* empeoran su comportamiento

La figura 7.31 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento de todos los agentes que

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

ofrecen soluciones de calidad *Alta*. Se comparan los casos en que cambia el comportamiento de los agentes de calidad *Alta* y en que todos los agentes mantienen su comportamiento constante. Por su parte, la figura 7.32 muestra la evolución del error cometido por el modelo al seleccionar la respuesta adecuada para la tarea a resolver en cada ronda, comparando las mismas situaciones. Ambas figuras ofrecen la evolución de las métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

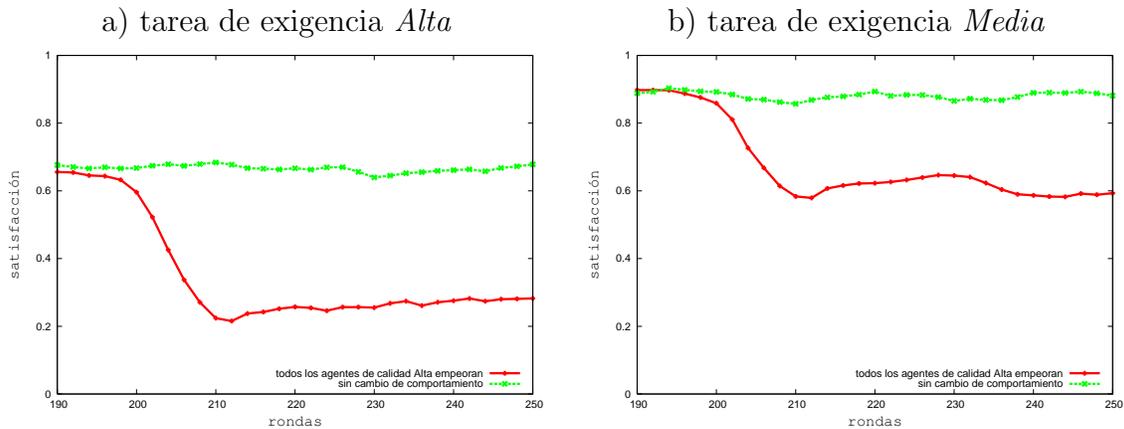


Figura 7.31: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento.

También se ofrece la figura 7.33 que muestra la evolución de la confianza en el agente recomendado por el modelo en cada ronda, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que todos los agentes de calidad *Alta* empeoran su comportamiento.

Al igual que sucede con los experimentos mostrados en la sección 7.3.2.1, donde la situación experimental está dada por el empeoramiento de un agente de calidad *Alta*, se observan variaciones significativas en la evolución de los indicadores representados en las figuras 7.31, 7.32 y 7.33. Para todos los indicadores, las variaciones son mayores en el caso en que cambia el comportamiento de todos los agentes de calidad *Alta*, que en el caso donde solo cambia el comportamiento de uno de ellos.

Tanto para tareas de exigencia *Alta* como para las de exigencia *Media*, el nivel de satisfacción ofrecido por el modelo, el error cometido en la recomendación y la confianza en el agente que recomienda una solución se ven afectados tras empeorar el comportamiento de todos los agentes de calidad *Alta*. El nivel de satisfacción (que se muestra en la figura 7.31) y la confianza promedio que se tiene en el agente recomendado por el modelo (que se muestra en la figura 7.33) disminuyen tras el empeoramiento del desempeño de todos los agentes que inicialmente eran de calidad *Alta*. De igual forma, el error de recomendación

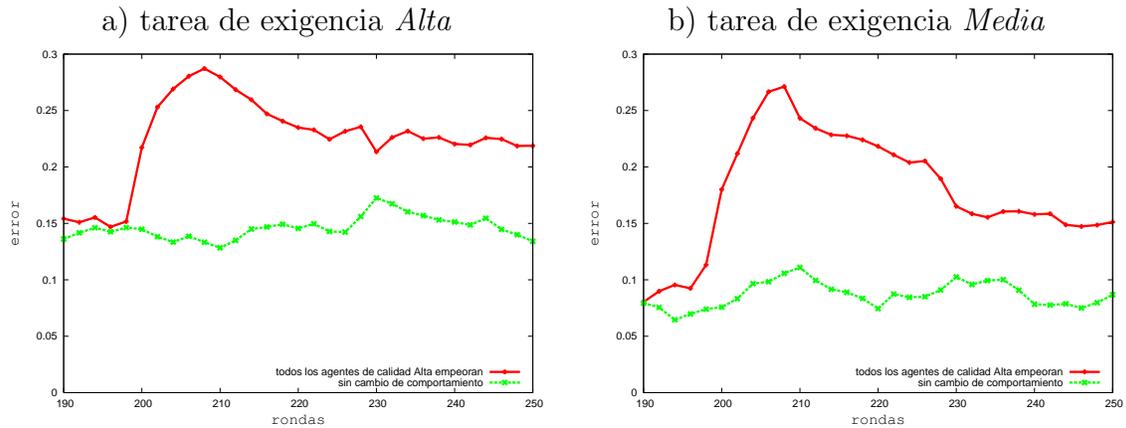


Figura 7.32: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento.

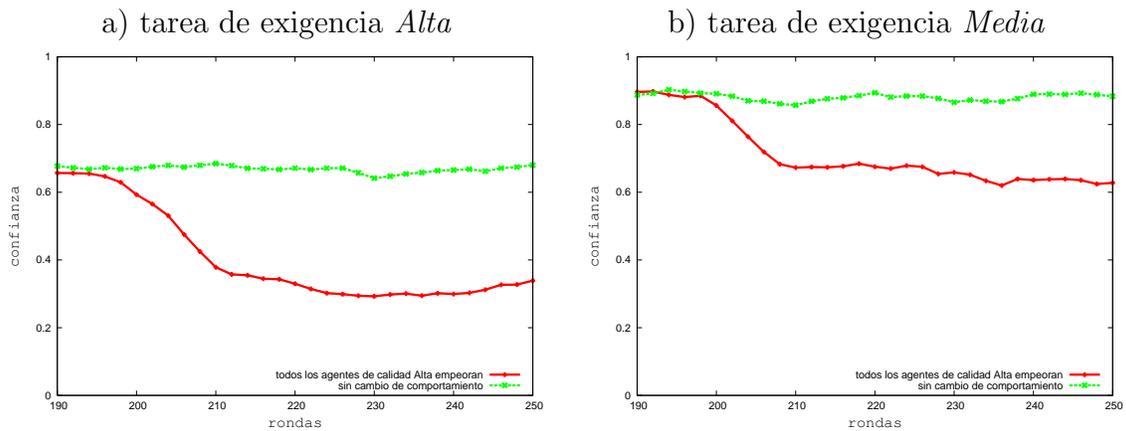


Figura 7.33: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento.

aumenta, para cada tipo de tareas, tras el cambio del comportamiento de todos los agentes que ofrecen soluciones de calidad *Alta* (Figura 7.32). El error alcanzan los valores más altos cuando finaliza el cambio de comportamiento de los agentes. Posteriormente, disminuyen y se estabilizan en un valor más bajo, aunque superior al que se tenía antes de que los agentes modificaran la calidad de las respuestas.

La confianza en el agente recomendado en cada ronda varía en la misma medida en que lo hace la calidad de las respuestas que ofrecen los agentes. En todo momento, el agente que recomienda el modelo es aquel que mejor respuesta ofrece en cada ronda. Así, la confianza en el agente que recomienda la respuesta decrece paulatinamente desde un nivel alto (correspondiente al momento en que existen agentes que ofrecen soluciones de calidad *Alta*) hasta niveles más bajos. Los valores de confianza alcanzados, al final del cambio, se corresponden con la confianza que merece un agente que ofrece soluciones de calidad *Media*, por cuanto sus soluciones son las mejores en el sistema cuando todos los agentes de calidad *Alta* han empeorado. El modelo es capaz de reflejar, utilizando los valores de confianza, la disminución de la satisfacción que se produce al empeorar la calidad de la respuesta que ofrece el agente recomendado en cada ronda.

Al comparar este experimento con el del apartado 7.3.2.1, donde cambia el comportamiento de un único agente de calidad *Alta*, se observa que en él las variaciones de la satisfacción son menores que cuando cambian todos los agentes. Esto se debe a que, en esos casos, después de la modificación aun permanecen agentes en el sistema que ofrecen soluciones de calidad *Alta*. En cambio, en la situación de este experimento, al empeorar la calidad de las soluciones de todos los agentes, que al inicio son de calidad *Alta*, ya no existen soluciones de buena calidad y por tanto, las mejores soluciones disponibles, después del cambio, son las ofrecidas por los agentes de calidad *Media*. Esta también es la causa de que la confianza en el agente que recomienda la solución en cada ronda alcance valores más bajos, y el error de recomendación valores más altos, cuando empeora la calidad de todos los agentes que al inicio ofrecen respuestas de calidad *Alta*.

Estas tres figuras (7.31, 7.32 y 7.33) muestran que el modelo responde correctamente al empeoramiento del comportamiento de un grupo de agentes de calidad *Alta*.

También se ofrece la figura 7.34 que muestra el promedio de la variación de la confianza que el agente iniciador tiene en los agentes de calidad *Alta* que empeoran, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes.

Frente a ambos tipos de tareas, el modelo disminuye la confianza que tiene en los agentes que empeoran la calidad de su respuesta. La disminución de los valores de confianza dependen significativamente del tipo de tarea a resolver. Cuando la tarea es de exigencia *Alta* los agentes alcanzan valores de confianza más pequeños. Evidentemente, de manera general, un mismo agente es menos confiable para la ejecución de una tarea de exigencia *Alta* que para una tarea de exigencia *Media*.

Durante el cambio, e incluso algunas rondas siguientes, el modelo se desestabiliza, alcanzando nuevamente un estado estable, aun cuando el nuevo valor del error de estimación es mayor que el que garantizaba el modelo antes del cambio. Para los casos en que se

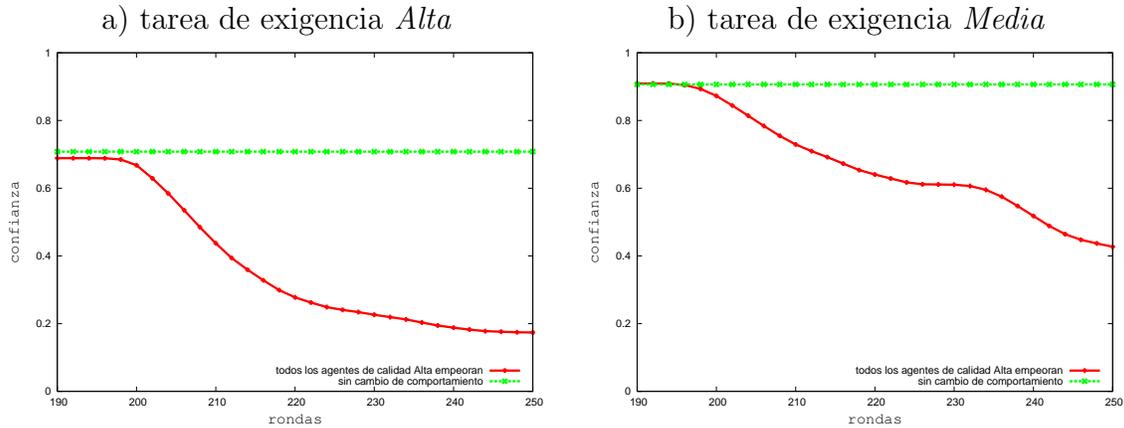


Figura 7.34: Evolución del promedio de la confianza de los agentes de calidad *Alta* que empeoran su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

solucionan tareas de exigencia *Alta*, el modelo alcanza la estabilidad en la ronda $t = 234$, mientras que para las de exigencia *Media* lo hace en la ronda $t = 232$. Estos resultados evidencian que el modelo es robusto a este tipo de cambios en el comportamiento de los agentes, volviendo a alcanzar un estado estable.

7.3.3.2. Todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo

La figura 7.35 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en un grupo de agentes que ofrece soluciones de calidad *Alta*. Se comparan los casos en que cambia el comportamiento de los agentes de calidad *Alta*, y en que todos los agentes mantienen su comportamiento constante. Por su parte, la figura 7.36 muestra la evolución del error cometido por el modelo al seleccionar la respuesta a la tarea de cada ronda, comparando las mismas situaciones. Ambas figuras ofrecen la evolución de las métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

También se ofrece la figura 7.37 que muestra la evolución de la confianza en el agente recomendado por el modelo en cada ronda, según el tipo de tarea a resolver, comparando los casos en que el comportamiento de los agentes permanece constante y en que todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo.

Al igual que sucede con los experimentos mostrados en la sección 7.3.2.3, donde la situación experimental está dada por el empeoramiento y recuperación de un agente de calidad *Alta*, se observan variaciones significativas en la evolución de los indicadores representados en las figuras 7.35, 7.36 y 7.37 durante el período en que varía la calidad de las respuestas. Para todos los indicadores, las variaciones son mayores en el caso en que cambia el comportamiento de todos los agentes de calidad *Alta*, que en el caso donde solo cambia el comportamiento de uno de ellos.

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

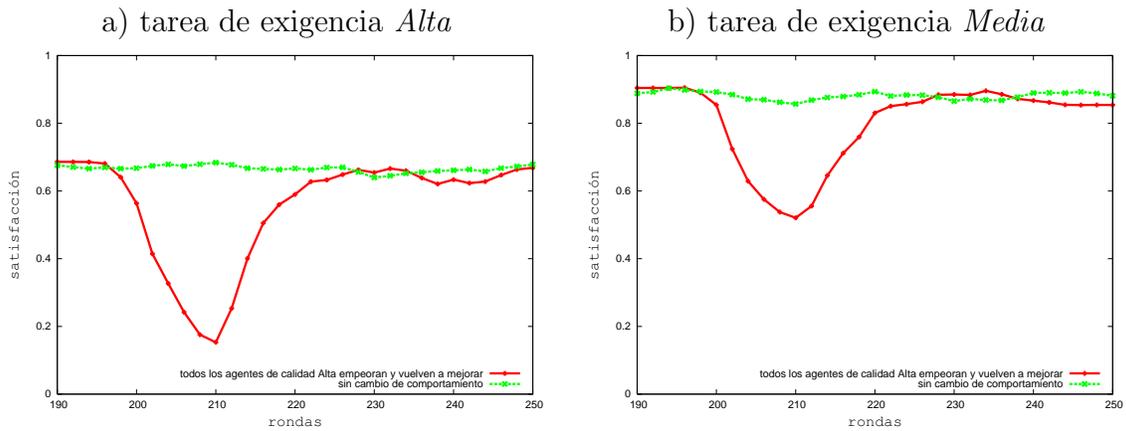


Figura 7.35: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo.

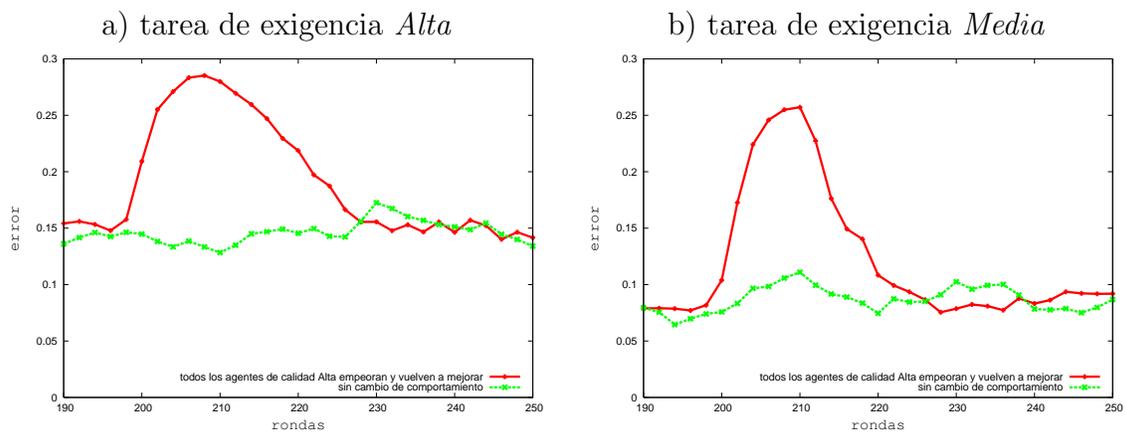


Figura 7.36: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo.

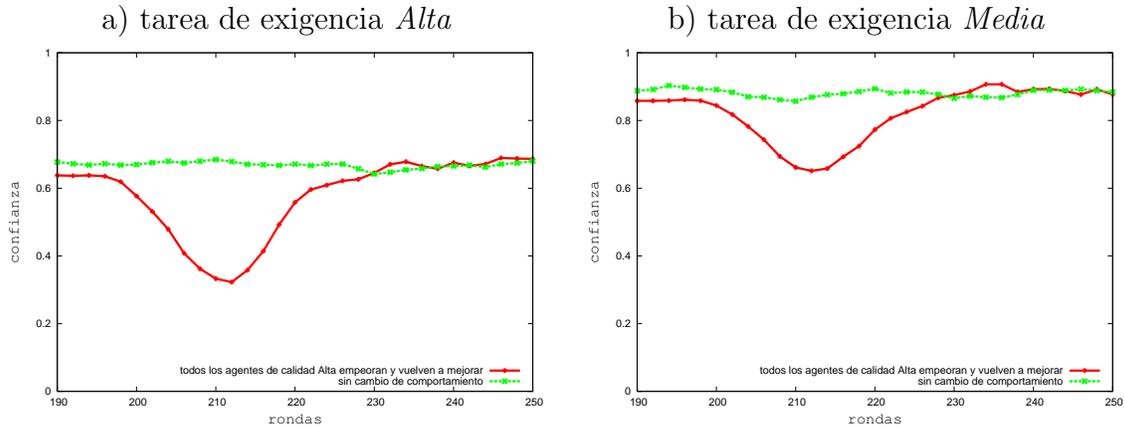


Figura 7.37: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando todos los agentes de calidad *Alta* empeoran su comportamiento y vuelven a mejorarlo.

Tanto para las tareas de exigencia *Alta* como para las de exigencia *Media*, el nivel de satisfacción ofrecido por el modelo, el error cometido en la recomendación y la confianza en el agente que recomienda una solución se ven afectados durante el período en que se modifica el comportamiento de todos los agentes de calidad *Alta*.

El nivel de satisfacción (que se muestra en la figura 7.31) y la confianza promedio que se tiene en el agente recomendado por el modelo (que se muestra en la figura 7.33) disminuyen durante la primera mitad del cambio, cuando disminuye la calidad de las respuestas de todos los agentes que inicialmente eran de calidad *Alta*. Luego, durante el período en que se recupera la calidad de la respuesta, el nivel de satisfacción y la confianza en el agente recomendado alcanzan valores similares a los que tenían antes del cambio. De la misma manera, el error de recomendación aumenta durante la primera mitad del cambio del comportamiento de los agentes; después, disminuye hasta alcanzar valores similares a los iniciales (Figura 7.32).

Al comparar este experimento con el del apartado 7.3.2.3, donde cambia el comportamiento de un único agente de calidad *Alta*, se observa que, en él, las variaciones de la satisfacción son menores que cuando cambian todos los agentes. Esto se debe a que, en esos casos, después de la modificación aun permanecen agentes en el sistema que ofrecen soluciones de calidad *Alta*. En cambio, en la situación de este experimento, al empeorar la calidad de las soluciones de todos los agentes que al inicio son de calidad *Alta*, ya no existen soluciones de alta calidad. Es posible que, en muchos instantes de tiempo durante el cambio, las mejores soluciones disponibles sean las ofrecidas por los agentes de calidad *Media*. Esta también es la causa de que la confianza en el agente que recomienda la solución en cada ronda alcance valores mínimos más bajos, y el error de recomendación valores máximos más altos, cuando empeora, y luego mejora, la calidad de todos los agentes de calidad *Alta*.

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

Estas tres figuras (7.35, 7.36 y 7.37) muestran que el modelo responde correctamente al empeoramiento y posterior recuperación del comportamiento de un grupo de agentes de calidad *Alta*.

Por otra parte, también se ofrece la figura 7.38 que muestra el promedio de la variación de la confianza en el grupo de agentes de calidad *Alta* que empeoran su comportamiento y vuelven a mejorarlo, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes.

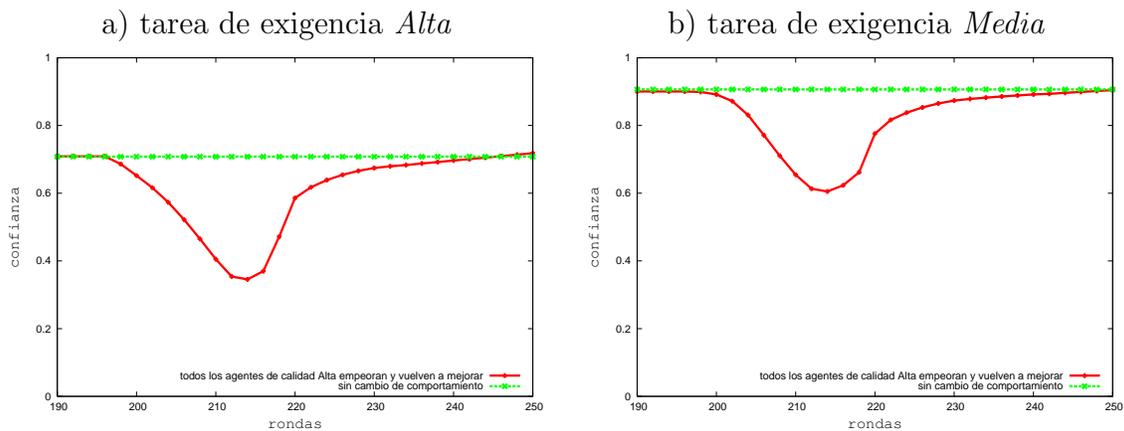


Figura 7.38: Evolución del promedio de la confianza de los agentes de calidad *Alta* que empeora su comportamiento y vuelve a mejorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

Frente a ambos tipos de tareas, durante la primera parte del cambio de comportamiento de los agentes (cuando éstos empeoran) el modelo disminuye la confianza que se tiene en ellos. Sin embargo, en tanto se comienza a recuperar la calidad de la respuesta hasta su estado inicial, la confianza aumenta hasta alcanzar valores similares a los que tenía antes del cambio.

La variación de los valores de confianza dependen significativamente del tipo de tarea a resolver. Cuando la tarea es de exigencia *Alta* las variaciones de la confianza son mayores. Evidentemente, de manera general, un mismo agente es menos confiable para la ejecución de una tarea de exigencia *Alta* que para una tarea de exigencia *Media*.

Durante el cambio, e incluso algunas rondas siguientes, el modelo se desestabiliza, alcanzando nuevamente un estado estable. Para los casos en que se solucionan tareas de exigencia *Alta*, el modelo alcanza la estabilidad en la ronda $t = 230$, mientras que para las de exigencia *Media* lo hace en la ronda $t = 228$. Estos resultados evidencian que el modelo es robusto a este tipo de cambios en el comportamiento de los agentes, recuperando su estabilidad.

7.3.4. Cambios en el comportamiento de un grupo de agentes que ofrece soluciones de calidad *Baja*

A continuación se ofrece otro grupo de experimentos que analiza el comportamiento del modelo en situaciones donde varía la calidad de las respuestas de los agentes de calidad *Baja*. En los experimentos presentados en los apartados 7.3.2.2 y 7.3.2.4, donde cambia el comportamiento de los agentes de calidad *Baja*, no se muestran variaciones significativas en los parámetros de calidad del modelo (satisfacción con la respuesta recomendada por el modelo, error cometido al recomendar dicha respuesta, y la confianza en el agente recomendado). Esto se debe fundamentalmente a que en la población de agentes existen varios agentes de calidad *Alta* que, en el momento del cambio, gozan de buena confianza y reputación y por lo tanto no hacen significativas las variaciones de calidad de un agente de calidad *Baja*.

El grupo de experimentos, cuyos resultados se muestran a continuación, consideran que en la población de agentes no existen agentes de calidad *Alta*. En ellos, se analiza si el modelo es capaz de reconocer los cambios de comportamiento en los agentes de calidad *Baja*, por cuanto, cuando existen agentes de calidad *Alta* no se favorece la selección de otros tipos de agentes y por tanto los parámetros del modelo se mantienen invariables. Es decir, se estudia el comportamiento de los parámetros del modelo cuando un grupo de agentes de calidad *Baja* mejoran la calidad de las soluciones que ofrecen, en ausencia de agentes de calidad *Alta*. Partiendo de una población inicial de agentes donde el 60 % son agentes de calidad *Baja*, el 40 % de calidad *Media* y no existen agentes de calidad *Alta*, se estudian varias situaciones a partir del cambio de comportamiento de agentes de calidad *Baja*:

- el 50 % de los agentes de calidad *Baja* mejoran su comportamiento, y
- el 50 % de los agentes de calidad *Baja* mejoran su comportamiento y vuelve a empeorarlo.

A continuación se ofrecen los resultados experimentales más relevantes relacionadas con estas situaciones. El resto de condiciones experimentales consideradas son las comentadas en el apartado 7.1.

Al igual que en experimentos anteriores se muestran, por separado, los parámetros del modelo para los casos en que se resuelven tareas de exigencia *Alta* y *Media*.

7.3.4.1. Un grupo de agentes de calidad *Baja* mejoran su comportamiento

La figura 7.39 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en varios agentes que ofrecen soluciones de calidad *Baja*. Se comparan los casos en que cambia el comportamiento del 50 % de los agentes de calidad *Baja* y en que todos los agentes mantienen su comportamiento constante. Por su parte, la figura 7.40 muestra la evolución del error cometido por el modelo al seleccionar la respuesta adecuada para la tarea a resolver en

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

cada ronda, comparando las mismas situaciones. Ambas figuras ofrecen la evolución de las métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

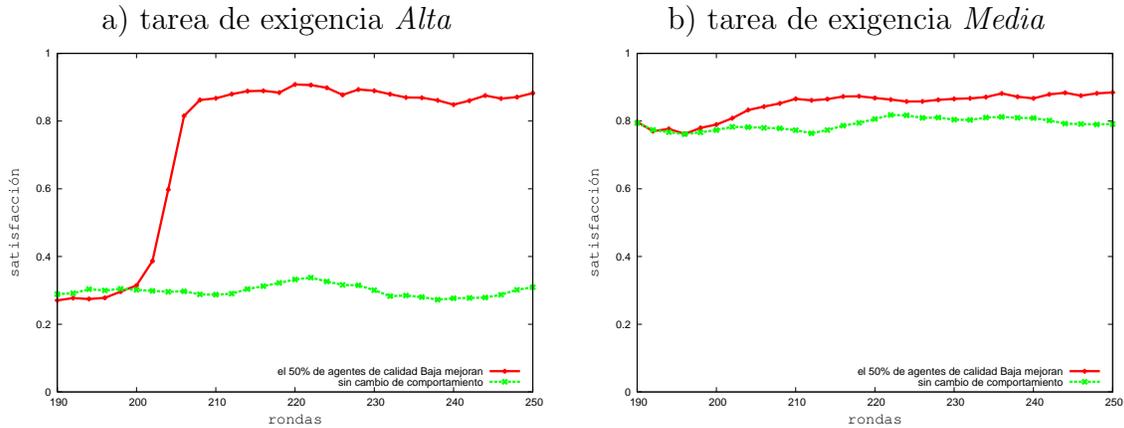


Figura 7.39: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad *Baja* mejoran su comportamiento.

Tal como muestra la figura 7.39.a, el nivel de satisfacción con la respuesta recomendada por el modelo aumenta considerablemente cuando aumenta la calidad de la respuesta que ofrecen los agentes. Esta diferencia entre los niveles de satisfacción, obtenida en cada caso analizado, se deben fundamentalmente a que, luego del cambio de comportamiento de los agentes, existen soluciones de mayor calidad que antes de tener lugar la modificación del comportamiento de los agentes de calidad *Baja*. Antes del cambio, el nivel de satisfacción de las tareas de exigencia *Alta* se mantiene en niveles bajos por cuanto solo se ofrecen respuestas de calidad *Media* o *Baja*. A medida que el comportamiento de los agentes comienza a mejorar, el 50% de las soluciones de calidad *Baja* pasan paulatinamente a ser de calidad *Alta* y por tanto, producen mayores niveles de satisfacción para la tarea de exigencia *Alta*.

Un comportamiento similar se muestra para las tareas de exigencia *Media* (Figura 7.39.b). Se obtienen los mayores niveles de satisfacción cuando varía el comportamiento del 50% de los agentes que al inicio ofrecen respuestas de calidad *Baja*. Sin embargo, las diferencias entre los niveles de satisfacción son menores. Esto se debe fundamentalmente a que, además de las respuestas de calidad *Alta* que paulatinamente van ofreciendo los agentes que cambian su comportamiento, existen agentes de calidad *Media* cuyas respuestas también producen elevados niveles de satisfacción para las tareas de exigencia *Media*.

A pesar del comportamiento diferenciado que experimenta la satisfacción con la respuesta recomendada (Figura 7.39), el modelo no muestra grandes variaciones en el error cometido al seleccionar la respuesta que mayor satisfacción produce en cada momento. La

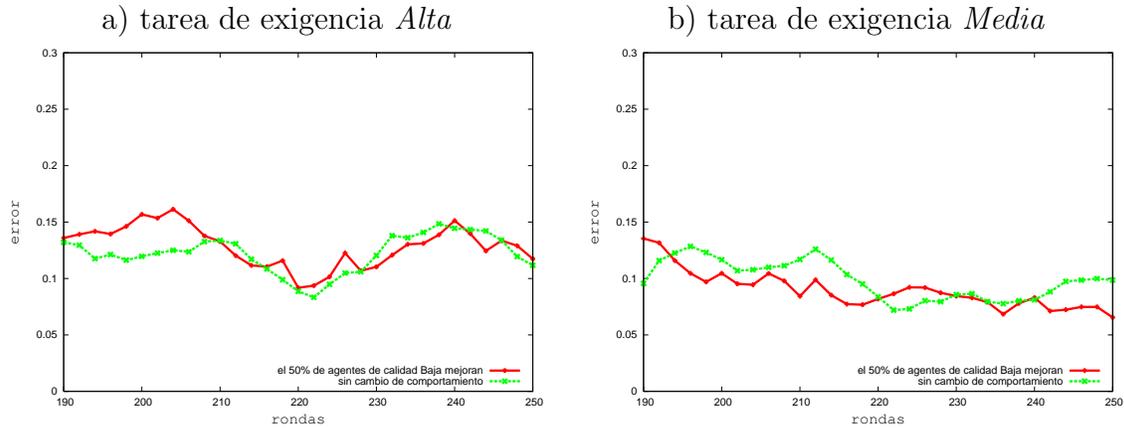


Figura 7.40: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad *Baja* mejoran su comportamiento.

figura 7.40 muestra cómo, para ambos tipos de tarea, el modelo reconoce en cada momento el agente que ofrece la respuesta adecuada, utilizando los valores de confianza. O sea, en todo momento se selecciona la respuesta que mejor satisface a la tarea. Por lo tanto, los valores del error cometido por el modelo al recomendar una respuesta se mantiene sin diferencias significativas en cada caso.

También se ofrece la figura 7.41 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los mismos casos que las dos figuras anteriores.

La confianza en el agente recomendado varía en la misma medida en que lo hace la calidad de las respuestas que ofrecen los agentes. En todo momento, el agente que recomienda el modelo es aquel que mejor respuesta ofrece en cada ronda. Así la confianza en el agente que recomienda la respuesta crece desde un nivel bajo (correspondiente al momento en que los agentes ofrecían soluciones de calidad *Baja*) hasta niveles más cercanos a 1 (que corresponden con los niveles de confianza que merece un agente que ofrece soluciones de calidad *Alta*).

Estas tres figuras (7.39, 7.40 y 7.41) muestran que el modelo responde correctamente al mejoramiento del comportamiento de un grupo de agentes de calidad *Baja*. A la vez que mantiene la estabilidad, el modelo es capaz de reflejar, utilizando los valores de confianza, el incremento de la satisfacción que se produce al mejorar la calidad de la respuesta que ofrece el agente recomendado en cada ronda.

También se ofrece la figura 7.42 que muestra la evolución del promedio de la confianza que el agente iniciador tiene en los agentes de calidad *Baja* que mejoran su comportamiento, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar,

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

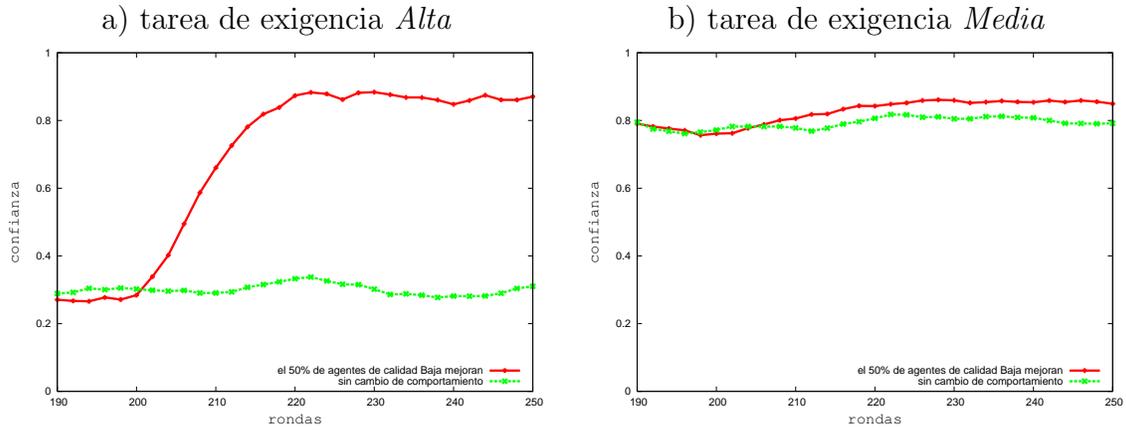


Figura 7.41: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50 % de los agentes de calidad *Baja* mejoran su comportamiento.

utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. Frente a ambos tipos de tareas, el modelo aumenta la confianza que tiene en los agente que mejoran la calidad de su respuesta. El aumento de los valores de confianza dependen significativamente del tipo de tarea a resolver. Cuando la tarea es de exigencia *Media* los agentes alcanzan valores de confianza ligeramente más elevados. Evidentemente, de manera general, un mismo agente es más confiable para la ejecución de una tarea de exigencia *Media* que para una tarea de exigencia *Alta*.

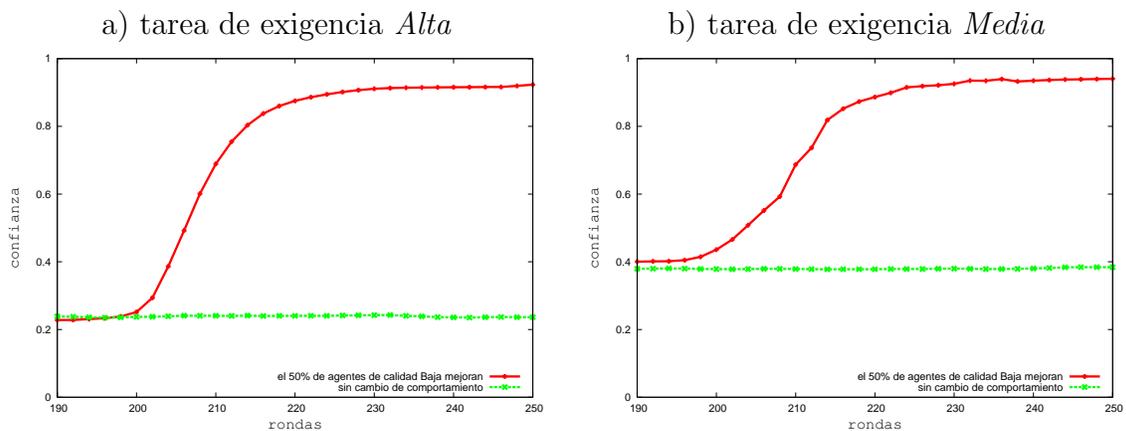


Figura 7.42: Evolución del promedio de la confianza de los agentes de calidad *Baja* que mejoran su comportamiento, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

7.3.4.2. Un grupo de agentes de calidad *Baja* mejoran su comportamiento y vuelven a empeorarlo

La figura 7.43 muestra la evolución del nivel de satisfacción con la solución recomendada por el modelo cuando se produce un cambio de comportamiento en un grupo de agentes que ofrece soluciones de calidad *Baja*. Se comparan los casos en que cambia el comportamiento del 50% de los agentes de calidad *Baja*, y en que todos los agentes mantienen su comportamiento constante. Por su parte, la figura 7.44 muestra la evolución del error cometido por el modelo al seleccionar la respuesta a la tarea de cada ronda, comparando las mismas situaciones. Ambas figuras ofrecen la evolución de las métricas para los dos casos tratados: cuando se selecciona una tarea de exigencia *Alta* y cuando se selecciona una tarea de exigencia *Media*.

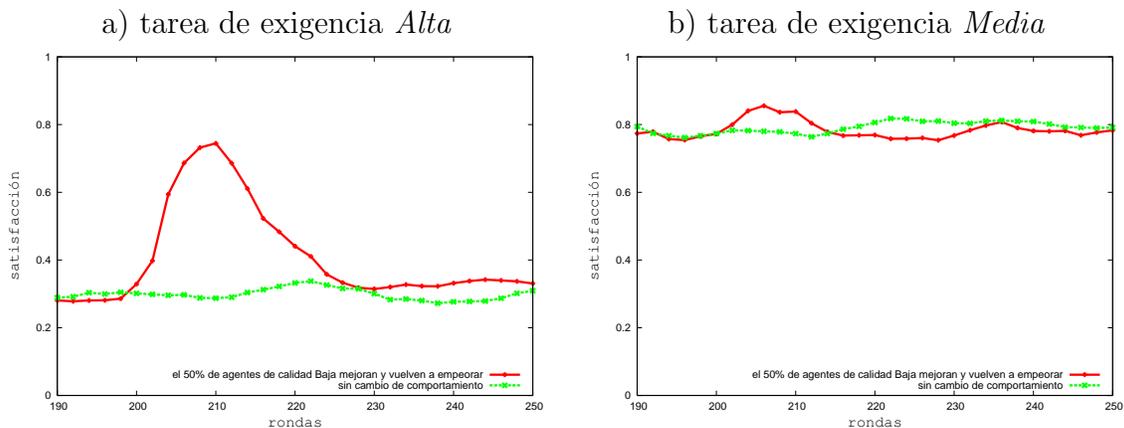


Figura 7.43: Comparación del nivel de satisfacción con la solución recomendada por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad *Baja* mejoran su comportamiento y vuelven a empeorarlo.

Tal como muestra la figura 7.43.a, para tareas de exigencia *Alta*, el nivel de satisfacción con la respuesta recomendada por el modelo aumenta durante la primera parte del cambio (cuando aumenta la calidad de la respuesta que ofrecen los agentes). Esto sucede debido a que la solución que va ofreciendo el agente (y que mejora paulatinamente) llega a garantizar mejores niveles de satisfacción que el resto de agentes que no varían su comportamiento. O sea, la calidad de la respuesta de los agentes que mejoran su comportamiento llega a ser superior que la ofrecida por los agentes de calidad *Media*, que antes del cambio eran los que ofrecían las mejores soluciones. A medida que el comportamiento de los agentes mejora, el 50% de las soluciones de calidad *Baja* pasan paulatinamente a ser de calidad *Alta* y, por lo tanto, producen mayores niveles de satisfacción para la tarea de exigencia *Alta*. Luego de comenzar a disminuir la calidad de la respuesta de los agentes que cambian, se restablecen los valores iniciales de satisfacción.

7.3. ADAPTACIÓN ANTE LOS CAMBIOS EN EL COMPORTAMIENTO DE LOS AGENTES PROVEEDORES

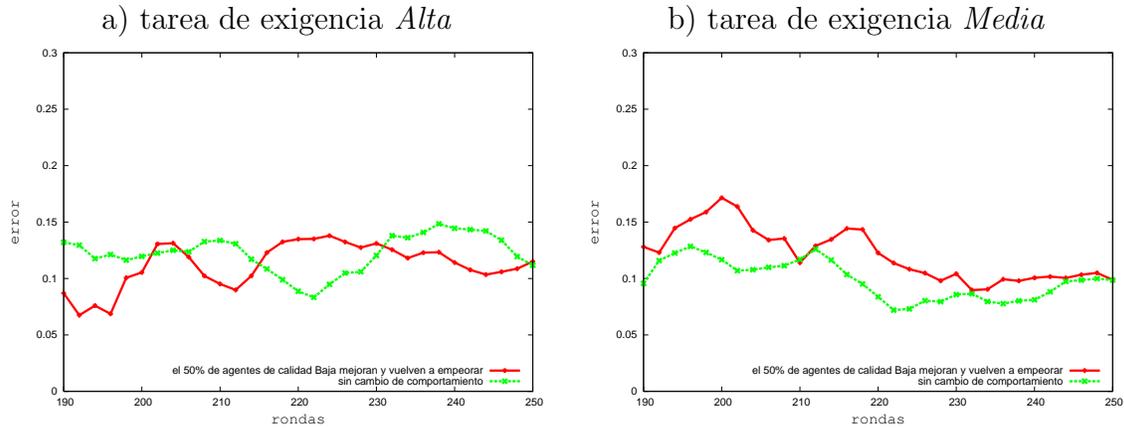


Figura 7.44: Comparación del error cometido por el modelo al seleccionar la solución adecuada, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad *Baja* mejoran su comportamiento y vuelven a empeorarlo.

Un comportamiento similar se muestra para las tareas de exigencia *Media* (Figura 7.43.b). Aumentan ligeramente los niveles de satisfacción mientras los agentes de calidad *Baja* van mejorando su comportamiento; luego disminuyen hasta valores similares a los obtenidos antes del cambio. Sin embargo, las diferencias entre los niveles de satisfacción son menores. Esto se debe, en lo fundamental, a que además de las respuestas de calidad *Alta*, que paulatinamente van ofreciendo los agentes que cambian, existen agentes de calidad *Media* cuyas respuestas también pueden ofrecer niveles de satisfacción elevados para las tareas de exigencia *Media*.

A pesar del comportamiento diferenciado que experimenta la satisfacción con la respuesta recomendada (Figura 7.43), el modelo no muestra grandes variaciones en el error cometido al seleccionar la respuesta que mayor satisfacción produce en cada momento. La figura 7.44 muestra cómo, para ambos tipos de tarea, el modelo reconoce en cada momento el agente que ofrece la respuesta adecuada, utilizando los valores de confianza. En otras palabras, en todo momento se selecciona la respuesta que mejor satisface a la tarea. Por lo tanto, los valores del error cometido por el modelo al recomendar una respuesta se mantiene sin diferencias significativas en cada caso.

También se ofrece la figura 7.45 que muestra la evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver, comparando los mismos casos que las dos figuras anteriores.

La confianza en el agente recomendado varía en la misma medida en que lo hace la calidad de las respuestas que ofrecen los agentes. En todo momento, el agente que recomienda el modelo es aquel que mejor respuesta ofrece en cada ronda. Así, en la primera mitad del cambio de comportamiento (mientras la calidad de la respuesta mejora) la confianza en el agente que recomienda la respuesta crece desde un nivel bajo

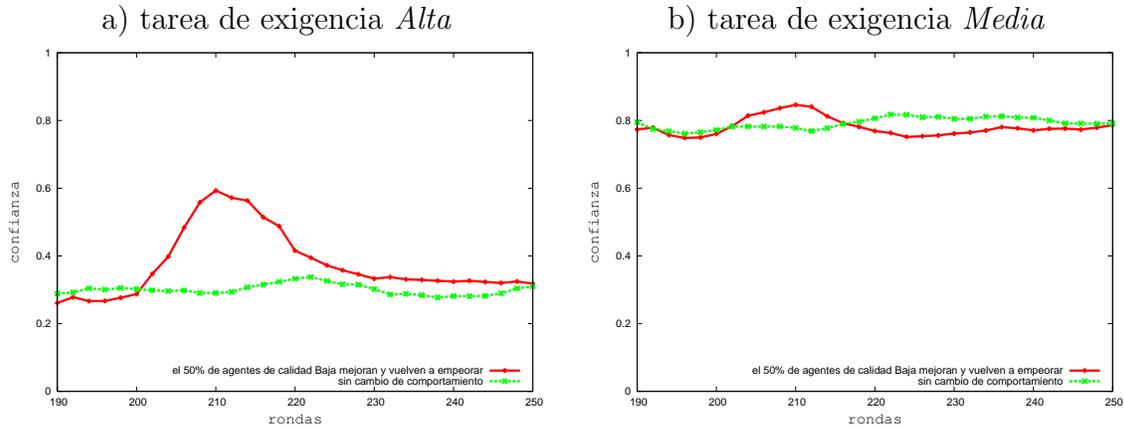


Figura 7.45: Evolución de la confianza en el agente recomendado por el modelo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*), para los casos en que el comportamiento de los agentes permanece constante a lo largo de la simulación y cuando el 50% de los agentes de calidad *Baja* mejoran su comportamiento y vuelven a empeorarlo.

(correspondiente al momento en que los agentes ofrecían soluciones de calidad *Baja*) hasta niveles más altos (que corresponden con los niveles de confianza que merece un agente que ofrece soluciones de calidad *Alta*). De igual forma, en la segunda mitad del cambio (mientras la calidad de la respuesta empeora) la confianza en el agente que recomienda la respuesta decrece nuevamente hasta un nivel similar al que tenía antes de iniciar el cambio.

Estas tres figuras (7.43, 7.44 y 7.45) muestran que el modelo responde correctamente a este tipo de cambio en el comportamiento de un grupo de agentes de calidad *Baja*.

Por otra parte, también se ofrece la figura 7.46 que muestra el promedio de la variación de la confianza en el grupo de agentes de calidad *Baja* que mejoran su comportamiento y vuelven a empeorarlo, según el tipo de tarea a resolver, comparándola con los casos en que el comportamiento de los agentes permanece constante. En ella se muestra cómo el modelo es capaz de representar, utilizando los valores de confianza, los cambios que se producen en el comportamiento de los agentes. En este caso, durante la primera parte del cambio de comportamiento de los agentes (cuando éstos mejoran) el modelo aumenta la confianza que se tiene en ellos. Sin embargo, en tanto se comienza a degradar la calidad de la respuesta hasta su estado inicial, la confianza disminuye hasta alcanzar valores similares a los que tenía antes del cambio.

La magnitud de la variación de la confianza depende del tipo de tarea a resolver. Cuando la tarea es de exigencia *Media* los agentes, que modifican su comportamiento, alcanzan valores de confianza más elevados. Evidentemente, un mismo agente es más confiable para la ejecución de una tarea de exigencia *Media* que para una tarea de exigencia *Alta*.

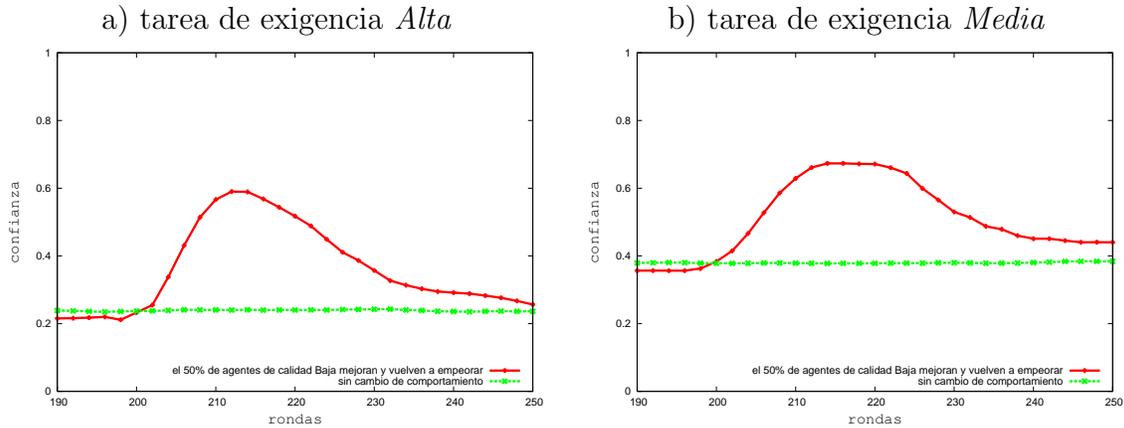


Figura 7.46: Evolución de la confianza del agente de calidad *Baja* mejora su comportamiento y vuelve a empeorarlo, según el tipo de tarea a resolver en cada ronda (a) de exigencia *Alta*, b) de exigencia *Media*).

7.3.5. Conclusiones del estudio de la capacidad de adaptación

En esta sección, se han realizado un conjunto de experimentos que han permitido estudiar la capacidad de adaptación del modelo frente a algunos cambios en el comportamiento de varios tipos de agentes. Tanto cuando varía el comportamiento de un único agente como cuando lo hace un grupo de ellos, el modelo es capaz de ajustar, por medio de los valores de confianza, el criterio que se tiene en los agentes que varían la calidad de las soluciones que ofrecen.

En los casos en que se empeora la calidad de un agente de calidad *Alta*, el modelo se desestabiliza. Sin embargo, tras un número de rondas relativamente pequeño, el modelo vuelve a alcanzar la estabilidad.

Cuando el cambio de comportamiento se produce en agentes de calidad *Baja*, el modelo no se desestabiliza. Es capaz de reflejar el cambio de comportamiento de estos agentes, aumentando la confianza en ellos y la satisfacción con la respuesta recomendada por el modelo.

La tabla 7.3 muestra, según el tipo de tarea, un resumen del efecto de cada una de las modificaciones de la calidad de la respuesta en la estabilidad del modelo. En los casos en que el modelo se desestabiliza se muestra el instante de tiempo (número de la ronda) en que el modelo vuelve nuevamente a una situación de equilibrio tras finalizar un cambio en el comportamiento de los agentes.

Algunos resultados relacionados con el contenido de bloque de experimentos se encuentran descritos en [26].

Tabla 7.3: Efecto de cada una de las modificaciones de la calidad de la respuesta en la estabilidad del modelo. Se muestra, en su caso, el número de la ronda en que el modelo alcanza nuevamente el equilibrio tras finalizar el cambio en el comportamiento de los agentes.

Situación estudiada	Tarea exigencia <i>Alta</i>	Tarea exigencia <i>Media</i>
Un agente bueno empeora	218	218
Un agente malo mejora	no se desestabiliza	no se desestabiliza
Un agente bueno empeora y vuelve a mejorar	228	224
Un agente malo mejora y vuelve a empeorar	no se desestabiliza	no se desestabiliza
Todos los agentes buenos empeoran	234	232
Todos los agentes buenos empeoran y vuelven a mejorar	230	228
El 50 % de los agentes malos mejoran	no se desestabiliza	no se desestabiliza
El 50 % de los agentes malos mejoran y vuelven a empeorar	no se desestabiliza	no se desestabiliza

Capítulo 8

Adecuación de TRSIM a ART

En este capítulo se describe la adaptación del modelo TRSIM al escenario de prueba ART. Los principales objetivos de esta integración están relacionados con el hecho de utilizar ART para el estudio del comportamiento de algunas variantes de TRSIM y la selección de las configuraciones de mejor rendimiento, para ser empleadas en la implementación de las funcionalidades del agente *Novel*. El agente *Novel* ha participado en la *II Competición Nacional de Modelos de Confianza y Reputación* celebrada en el marco de la *III Escuela sobre Agentes (Valencia, 2007)*, y en la *II Competición Internacional de Modelos de Confianza y Reputación* celebrada en el marco del AAMAS 2007.

Aunque, a partir de la participación en dichas competiciones, el desempeño del modelo de confianza y reputación implementado por *Novel* ha podido ser contrastado frente a otros modelos, los resultados obtenidos no son representativos de la utilidad y funcionalidad de TRSIM. Para poder ser utilizado en ART, TRSIM ha sesgado sus principales características distintivas que realzan su comportamiento: (1) ha debido adoptar una función específica para la evaluación de la calidad de la respuesta, sin tener en cuenta la función de calidad propuesta a partir de los elementos del modelo descritos ontológicamente, y (2) no ha podido utilizar la intuición, como fuente de información de confianza y reputación, dado que ART no ofrece una representación semántica suficiente para que dichos mecanismos sean implementados.

Por ello, tanto los experimentos que se muestran en el capítulo anterior como los de éste, se limitan a analizar el comportamiento del modelo ante diferentes configuraciones y condiciones experimentales dada por cambios de comportamiento en los agentes. No se ofrecen experimentos que comparan TRSIM con otros modelos. La comparación del TRSIM con otros modelos de confianza y reputación es una tarea que se propone para estudios siguientes a esta tesis, por cuanto, se requiere de un escenario de prueba que haga factible la utilización de las principales ventajas que ofrece la propuesta dada por TRSIM: (1) la representación ontológica de requerimientos y respuestas, y (2) la estimación de la confianza y la reputación a partir del conocimiento asociado a requerimientos similares.

El presente capítulo se estructura de la forma siguiente: en la sección 8.1, se relacionan los puntos de contacto y divergencia entre los dos enfoques. La sección 8.2 describe las funcionalidades de los métodos estratégicos de un agente que utilice TRSIM en el

escenario de prueba ART, ofreciendo variantes para algunos de ellos. En este sentido, se ofrecen algunos resultados experimentales que permiten seleccionar las variantes de mejor rendimiento para su implementación en el agente *Novel*. Las características de este agente son comentadas en el apartado 8.3. Por último, se ofrecen algunas conclusiones de este estudio (sección 8.4) y se hacen varias recomendaciones para la extensión de ART a partir de considerar la similitud entre tareas (sección 8.5).

8.1. TRSIM desde el punto de vista de ART

TRSIM es un modelo de confianza y reputación entre agentes, en un entorno basado en la interacción entre pares, utilizado para proponer la respuesta idónea según los requerimientos de un cliente [21, 20]. En él, la confianza y la reputación son fenómenos sociales basados en las múltiples interacciones entre los agentes y las relaciones que se presentan entre ellos. Los valores de confianza indican la creencia que se tiene en la satisfacción que genera la respuesta ofrecida por el agente contraparte en una interacción determinada, según el tipo de tarea especificado. En cambio, los de reputación indican, de alguna manera, la credibilidad que se le otorga a los criterios de un agente al recomendar a otros para dar solución a un determinado tipo de tarea.

Por su parte, para ART los conceptos de confianza y reputación también resultan del análisis de las interacciones entre los agentes para tasar obras de arte de una era determinada. Los valores de confianza son indicadores de la destreza de un agente para hacer tasaciones de pinturas de una era específica. Mientras que la reputación esta referida a la capacidad de un agente para ofrecer criterios sobre la habilidad de otros tasando pinturas de una era determinada.

A priori, la adecuación de TRSIM a ART parece una tarea factible. Al mismo tiempo, teniendo en cuenta las peculiaridades de TRSIM y de ART se advierte que se deben realizar algunos ajustes al modelo para poder ser probado en el escenario. A continuación se relacionan algunas características de TRSIM, a la vez que se analiza su representación utilizando las conceptos de ART:

- La confianza y reputación son valores asociados a la tarea que el agente necesita resolver en un momento dado. En algunos casos, estos valores no pueden ser calculados por falta de experiencias relacionadas con la tarea a resolver. Para ello, TRSIM propone estimarlos a partir de la combinación de la confianza y la reputación en alguna tarea similar y el grado de similitud entre esta tarea conocida y la tarea cuyo comportamiento se desconoce.

La similitud entre tareas es un elemento clave para la estimación de la confianza y la reputación cuando no se dispone de experiencias previas [25].

En ART, la habilidad de un agente como tasador (s^*) y el criterio, que un agente tiene sobre otros, están referidos a una era determinada. Sin embargo, la representación del concepto de era no permite establecer criterios de similitud para realizar algún tipo de estimación como la comentada anteriormente.

- Utiliza WSMO [140] como base para la definición de sus principales conceptos. Se emplean algunas propiedades no funcionales de los conceptos ofrecidos por WSMO para la descripción de las tareas, respuestas y funciones del modelo tales como la similitud entre tareas y la satisfacción de una tarea dada una respuesta.
- Cada agente mantiene dos bases de experiencias, una que almacena información sobre la calidad de los agentes ofreciendo soluciones (*IET*) y otra que almacena información sobre la calidad de los agentes ofreciendo información sobre el comportamiento de otros agentes (*IER*). Todas las experiencias de ambas bases hacen referencia a la calidad de un agente según la especificación de una determinada tarea.
- El modelo posee un conjunto de funciones útiles para calcular, en cualquier instante de tiempo, los valores de confianza y reputación asociados a una determinada tarea, a partir de las experiencias de ambas bases. Estos valores hacen posible identificar los agentes más prometedores para pedir soluciones u opiniones acerca de otros agentes. Así, el modelo de confianza y reputación guía el proceso de selección de los agentes contraparte en una determinada interacción para resolver una tarea dada.

En ART tanto los grados de habilidad para dar opiniones como los pesos que se utilizan para calcular el valor de la tasación cuando se compran opiniones están referidos a una era determinada (consultar la funcionalidad del método estratégico *prepareOpinionProviderWeights()* en el apartado 6.3). En TRSIM, los valores de confianza, de reputación y las experiencias de ambos tipos dependen también del tipo de tarea a la que hacen referencia.

- Se considera la similitud entre tareas para obtener información de confianza y reputación sobre agentes de los cuales se desconoce su desempeño en una determinada tarea.

En el caso de ART sería muy útil poder considerar alguna medida de similitud entre las eras o las pinturas para tener una aproximación del desempeño de un tasador cuando no se tenga información sobre una determinada era.

- Al obtener una solución para la tarea dada, el modelo actualiza cada una de las bases de experiencias según los valores de satisfacción de la tarea a partir de la solución obtenida, el compromiso de satisfacción mínima (para el caso de la confianza) y incremento del valor de confianza (para la reputación).

A priori, la utilización del modelo en ART supone cambios en la forma de obtener la satisfacción por cuanto ésta se obtiene a partir de la comparación de dos conceptos complejos que representan la tarea y la respuesta. Al utilizar el modelo en ART, se tendría que establecer una nueva función de satisfacción a partir de las tasaciones (opiniones) de cada agente y del valor real de la pintura, fundamentalmente.

- Dado que el modelo ha sido concebido para un escenario basado en la interacción entre pares, donde los agentes representan consumidores y proveedores de recursos,

se tiene en cuenta que siempre los agentes proveedores responden con información verdadera acerca del compromiso de calidad de la solución que ofrecen y de la calidad de otros proveedores de servicios. Las decisiones en los agentes que muestran el rol proveedor de recursos son simples.

Por tal razón, en dicho escenario, carece de sentido la prueba del modelo frente a otros que no respondan con la verdad. La utilización de ART, para la realización de algunos experimentos, pretende fundamentalmente ajustar los parámetros del modelo para aumentar la satisfacción de los requerimientos del usuario dadas las posibles respuestas del sistema.

De manera general, las características de TRSIM posibilitan adecuarlo a ART para analizar el desempeño del modelo en los casos donde no se considera la similitud entre tareas, los proveedores de recursos no mienten, y se hace una abstracción de la evaluación de la satisfacción de los agentes iniciadores a partir de la solución obtenida y de los valores ofrecidos por ART.

8.2. Estrategia de un agente ART según TRSIM

A continuación se describe la funcionalidad de cada método de la clase abstracta *Agent* cuando son implementados en el agente *Novel* que tiene en cuenta el modelo TRSIM. (En el capítulo 5 se detallan los conceptos y funciones del modelo.)

1. **prepareReputationRequests()**: En este método se actualizan las bases de experiencias y se solicita información de reputación a otros agentes.

Para actualizar las bases de experiencias $IER_i^{(t)}$ y $IET_i^{(t)}$ se sigue la alternativa propuesta en [25]. La actualización de las bases de experiencias requiere la evaluación de la ronda anterior, para ello TRSIM utiliza las funciones P (cumplimiento del compromiso de satisfacción) y Q (satisfacción de la tarea a partir de la respuesta obtenida). En TRSIM, Q se define a partir de las propiedades de los conceptos definidos por WSMO, por lo tanto hay que establecer una nueva forma para obtener la satisfacción del agente a partir de los conceptos de ART. En este sentido se proponen dos variantes:

$$Q_1 = \begin{cases} 1 - \frac{|t-p^*|}{t} & : |t - p^*| < t \\ 0 & : |t - p^*| \geq t \end{cases}$$

$$Q_2 = \frac{p^*}{t}$$

donde p^* es la opinión dada y t es el valor real de la pintura.

Por otra parte, para solicitar información de reputación se pueden seguir diferentes estrategias:

- RR_1 : para las eras con valores bajos de habilidad, se solicita información de reputación a todos los agentes sobre los demás

- RR_2 : para las eras con valores bajos de habilidad, se solicita información de reputación a los agentes confiables para dar información de reputación (CR_{sup}) sobre los demás
- RR_3 : para las eras con valores bajos de habilidad, se solicita información de reputación a los agentes confiables para dar información de reputación (CR_{sup}) sobre los agentes de confianza dudosa para dar opiniones (CT_{dud}) sobre dichas eras

Solo se consideran las eras con valores bajos de habilidad por cuanto para las eras donde la habilidad para dar opiniones sea alta, se optará por emitir una opinión, nunca por comprar opiniones a otros agentes.

Como la solicitud de información de reputación lleva asociado un costo c_r , a priori, la mejor alternativa en este punto parece ser la estrategia RR_3 .

2. **prepareReputationAcceptsAndDeclines()**:

Aquí se debe analizar si resulta provechoso enviar información de reputación a todas las solicitudes recibidas.

TRSIM supone que siempre los agentes proveedores están dispuestos a responder a las solicitudes y siempre utilizando información verdadera, por tanto, en este método se propone responder a todas las solicitudes de información sobre el desempeño de otros agentes.

Sin embargo, pudiera ser interesante contrastar dos comportamientos diferentes:

- RA_1 : aceptar todas las solicitudes de información de reputación recibidas.
- RA_2 : aceptar todas las solicitudes de información reputación, excepto aquellas sobre agentes en los cuales se tiene una elevada confianza (CT_{sup}).

Similar al caso anterior, como la solicitud de información de reputación lleva asociado un costo c_r , a priori, la mejor alternativa en este punto parece ser la estrategia RA_2 .

3. **prepareReputationReplies()**: Se responde a las solicitudes, aceptadas en el método anterior, con el valor de confianza directa DT que se tiene sobre el agente objeto de consulta para la era especificada.

4. **prepareOpinionRequests()**: En este método se reciben las respuestas a las solicitudes de información de reputación y se solicitan opiniones de tasación de aquellas pinturas que se crea oportuno.

Se solicita opiniones de tasación a otros agentes cuando la habilidad para tasar una pintura de una era determinada no sea suficiente para ofrecer una tasación adecuada. Para las eras donde la habilidad para dar opiniones sea alta, se optará por emitir una opinión, nunca por comprar opiniones a otros agentes.

Para solicitar tasaciones a otros agentes se pueden seguir diferentes estrategias:

- OR_1 : se solicita opiniones a todos los agentes.
- OR_2 : se solicita opiniones solo a los agentes más confiables (CT_{sup}).
- OR_3 : se solicita opiniones a los agentes más confiables (CT_{sup}) y a aquellos de confianza dudosa con un elevado valor de reputación según los agentes encuestados (C_{prom}).

A priori, las alternativas OR_2 y OR_3 deben ofrecer mejor desempeño que OR_1 por cuanto la compra de opiniones lleva asociado un costo c_p .

5. **prepareOpinionsCertainties()**: En este método, el agente que recibe solicitudes de opiniones debe enviar el grado de certeza del valor de tasación que más tarde emitirá. En TRSIM, este valor de certeza puede ser considerado como el compromiso de satisfacción $ec_{i,j,k}$ que hace el agente encuestado a_j al agente a_i para la opinión que da sobre la pintura de la era s_k .

El compromiso también debe ser una valoración previa del tasador sobre si le conviene o no interactuar con el solicitante. Por ejemplo, para opiniones que interese responder se puede enviar un valor de certeza alto, para las que no sean interesantes, un valor bajo. El compromiso puede obtenerse en función de la habilidad que tenga dicho agente para dar opiniones sobre pinturas de la era en cuestión. Responderá siempre con la verdad, de acuerdo a la habilidad que tenga para tasar pinturas de esa era:

$$ec = 1,1 - s^*$$

donde s^* es la habilidad que tiene el agente para emitir tasaciones de pinturas de una era.

6. **prepareOpinionsRequestConfirmations()**: En este método se decide cuáles opiniones se tendrán en cuenta o no, analizando su costo, su certeza, la confianza en el agente que la da, etc.

TRSIM solo confirma aquellas solicitudes que tengan un elevado valor de certeza y que provengan de agentes con un elevado valor de confianza y reputación para la pintura en cuestión.

7. **prepareOpinionCreationOrders()**: En este método se responde a las solicitudes de tasación de otros agentes y a las asignaciones de solicitudes de tasación que ha recibido de los clientes. La exactitud de la opinión está determinada por el costo que esté dispuesto a pagar el agente por emitir dicha opinión (c_g), o sea, cuántos recursos se invierten en determinar el valor de la pintura.

Pueden plantearse formas diferentes de estimar el valor de las pinturas asignadas y las solicitudes recibidas por otros agentes.

Para determinar el costo de opinión para solicitudes recibidas:

$$c_g = \begin{cases} 0,8 \cdot \frac{f}{c_p} & : \text{timestep} = 0 \\ 0,8 \cdot \frac{|CT_{sup}|}{|N|} \cdot \frac{f}{c_p} & : \text{timestep} \geq 1 \end{cases}$$

donde N representa el conjunto de todos los agentes participantes en el juego ART. Para determinar el costo de la tasación para las pinturas asignadas:

$$c_g = \left(\frac{s^* + 1,6}{2} \right) \cdot \frac{f}{c_p}$$

Cuando se ofrecen valores de tasaciones a las solicitudes recibidas no se analiza la habilidad para dar opiniones, por cuanto solo se han aceptado solicitudes cuya habilidad es alta. Sin embargo, se considera que el costo de la opinión debe ser dada en función de la cantidad de agentes que también pudieran ofrecer una opinión para dicha pintura.

Por otro lado, cuando se tasan pinturas asignadas al propio agente, el valor es estimado a partir de la habilidad que se tiene. A priori, se puede pensar que el costo de la opinión debe ser tal que cuánto más hábil se sea en tasar pinturas de una determinada era, menor debería ser el costo de la opinión, por cuanto, el valor de tasación propuesto (calcula por ART a partir del costo dado por el agente) debe estar más cerca del valor real de la pintura.

8. **prepareOpinionsProviderWeights()**: En este método se envía al simulador, en forma de pesos, el criterio que se tiene sobre el resto de los agentes.

Los pesos que se envían corresponden al valor de confianza global que el agente asigna a cada uno del resto de los agentes para tasar pinturas de cada era.

9. **prepareOpinionReplies ()**: En este método se toman todas las solicitudes de opiniones que han sido confirmadas y son enviadas.

8.2.1. Experimentación: Comparación de estrategias alternativas

Utilizando ART se han desarrollado varios experimentos donde se compara el comportamiento de varios tipos de agentes. Cada tipo de agente se ha configurado a partir de las variantes definidas para los métodos estratégicos *prepareReputationRequests* y *prepareOpinionRequests* (ver el apartado anterior).

Todos los experimentos se han ejecutado bajo las mismas condiciones generales, similares a las definidas para la II Competición Internacional en el marco de AAMAS'07 [6], limitando el número de iteraciones a $Time\ steps = 25$. Cada experimento se ejecutó 3 veces con 10 agentes de cada tipo, obteniendo el promedio de las cuentas bancarias para cada tipo de agente.

Según los resultados de los experimentos se configuró el agente Novel que participó en la competición. Se decidió por utilizar la alternativa Q_1 para calcular la calidad; la estrategia RR_3 para solicitar información acerca del comportamiento de otros agentes (sin diferencias significativas con RR_2 , debido quizás a que las solicitudes de reputación son poco costosas); y la estrategia OR_3 para solicitar opiniones.

A modo de ejemplo, la figura 8.1 muestra la evolución del estado de cuentas para los tres tipos de agentes, cada uno de los cuales sigue una estrategia diferente al decidir a quiénes solicitar opiniones de tasación.

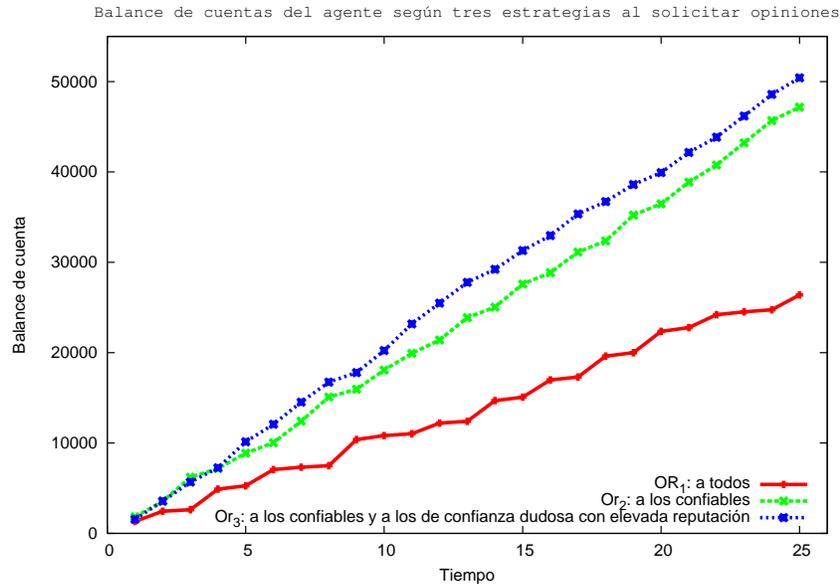


Figura 8.1: Balance del estado de cuentas para tres tipos de agentes según diferentes estrategias para solicitar opiniones a otros agentes.

Los mayores beneficios se generan cuando se solicita opiniones a los agentes confiables y a los que, teniendo asociado una confianza dudosa, gozan de buena reputación entre los agentes confiables para dar información de reputación.

8.3. Características del agente *Novel*

Teniendo en cuenta los resultados experimentales que permiten comparar algunas alternativas para la definición de las estrategias de un agente en ART según el modelo TRSIM, se ha definido el comportamiento del agente *Novel*. Las principales características de este agente se resumen a continuación:

- Responde con valores verídicos acerca del criterio que tiene sobre otros y el compromiso de satisfacción de la respuesta que ofrece.
- Para actualizar las bases de experiencias, utiliza la función de satisfacción:

$$Q = \begin{cases} 1 - \frac{|t-p^*|}{t} & : |t - p^*| < t \\ 0 & : |t - p^*| \geq t \end{cases}$$

- Al solicitar información de reputación, para las eras con valores bajos de habilidad, solicita información de reputación a los agentes confiables para dar información de reputación (CR_{sup}) sobre los agentes de confianza dudosa para dar opiniones (CT_{dud}) sobre dichas eras.
- Acepta todas las solicitudes de información de reputación recibidas.
- Al ser encuestado, responde a las solicitudes con el valor de confianza directa DT que se tiene sobre el agente objeto de consulta para la era especificada.
- Al solicitar tasaciones, solo envían solicitudes a los agentes más confiables (CT_{sup}), relacionadas con pinturas pertenecientes a eras donde la habilidad para dar opiniones no sea alta. Cuando la habilidad sea alta, se opta por ofrecer una opinión de tasación.
- Al ofrecer su compromiso (certeza) para emitir opiniones, responderá siempre con la verdad, de acuerdo a la habilidad que tenga para tasar pinturas de esa era:

$$ec = 1,1 - s^*$$

donde s^* es la habilidad que tiene el agente para emitir tasaciones de pinturas de una era.

- Confirman aquellas solicitudes cuya certeza (compromiso) sobrepasa la habilidad que se tiene tasando esa pintura, o sea, confirma la solicitud cuando:

$$ec \geq 1,1 - s^*$$

- Para determinar el costo de opinión para solicitudes recibidas, utiliza:

$$c_g = \begin{cases} 0,8 \cdot \frac{f}{c_p} & : \text{timestep} = 0 \\ 0,8 \cdot \frac{|CT_{sup}|}{|N|} \cdot \frac{f}{c_p} & : \text{timestep} \geq 1 \end{cases}$$

donde N representa el conjunto de todos los agentes ejecutándose en el juego de ART.

- Para determinar el costo de la tasación para las pinturas asignadas, utiliza:

$$c_g = \left(\frac{s^* + 1,6}{2} \right) \cdot \frac{f}{c_p}$$

- Los pesos que envía corresponden al valor de confianza global que el propio agente asigna a cada uno del resto de los agentes para tasar pinturas de cada era.

8.4. Conclusiones

Se ha adaptado el modelo de confianza TRSIM al escenario de prueba ART con el fin de comparar diferentes alternativas estratégicas de los agentes. Las variantes de TRSIM que se han comparado mediante ART no consideran en ningún caso la estimación de los valores de confianza y reputación cuando no se dispone de experiencias relacionadas con la era (o tarea) en cuestión, por cuanto ART aun no tiene en cuenta la similitud entre eras.

Es indudable que en aquellos casos en que no se dispone de conocimiento acerca del desempeño de un agente para una era determinada, la estimación de la confianza y la reputación, a partir de eras similares, puede resultar provechoso. Por ello, sugerimos que el concepto de pintura sea enriquecido semánticamente, incorporando nuevos elementos a su descripción. De esta manera, se podrán establecer criterios de similitud entre pinturas con el fin de estimar los valores de confianza y reputación en situaciones de desconocimiento. El atributo que determina la era de una pintura constituyen un criterio de agrupamiento de pinturas. Resulta factible pensar que todas las pinturas pudieran estar descritas por otros atributos comunes, precisamente aquellos que permitan establecer criterios de similitud.

Por otra parte, la función de calidad establecida a partir de la representación mediante WSMO no puede ser utilizada, por cuanto no existen atributos comunes entre las solicitudes de tasación y las opiniones recibidas. Las nuevas alternativas para la definición de las funciones de calidad se han definido a partir de los conceptos propios de ART: opinión dada por el agente tasador y el valor real de la pintura.

El principal elemento definido por ART que queda fuera de la definición de TRSIM es el establecimiento del costo de la opinión de tasación, teniendo en cuenta que es este parámetro el que determina la desviación de la opinión respecto al valor real de la pintura. La forma de estimar el costo de las opiniones se obtiene a partir de la relación existente entre el beneficio a obtener por una tasación acertada (f) y el costo que supone emitir una opinión (c_p). También tiene en cuenta la cantidad de agentes que pudieran ofrecer una opinión acertada.

Además, al ajustar TRSIM a ART, hay que refinar el procedimiento para limitar la lista de agentes a los cuales se le solicita información. Este análisis no solo debe considerar la confianza, sino también realizar un análisis de costos.

Algunos resultados relacionados con el contenido de este capítulo se encuentran descritos en [22].

8.5. Propuesta de extensión de ART para considerar la similitud entre pinturas

Presentadas, en la sección anterior, las principales incongruencias entre TRSIM y ART, en esta sección se comentan algunas ideas sobre la extensión del escenario teniendo en cuenta las propias características de TRSIM, definido en el capítulo 5. Este modelo destaca por dos elementos distintivos: (1) evaluación de un componente de intuición de la confianza y la reputación de los agentes, a partir de considerar la similitud entre tareas, y (2) la

utilización de un modelo ontológico sencillo y extensible para la definición de los elementos dependientes del dominio de aplicación.

Las principales propuestas de extensión de ART están relacionadas con el enriquecimiento semántico del concepto de *Pintura* y la consideración de medidas de similitud. Se propone así la base para establecer y evaluar mecanismos de intuición (o estimación) de los valores de confianza y reputación en determinadas situaciones. Se pretende, de esta forma, proponer algunas recomendaciones para el desarrollo de ART, dada su importancia como escenario de referencia adoptado para la validación de los modelos de confianza y reputación.

En esta sección no se comentan los detalles de funcionamiento y estructura del escenario de prueba, por cuanto son descritos en la sección 6.3. Solo se abordan los temas relacionados con su posible extensión.

8.5.1. Extensión semántica del concepto de Pintura

ART en su versión actual supone que la confianza y la reputación son dependientes de los requerimientos de los usuarios. (Los requerimientos de usuario son definidos por la solicitud de tasación de una pintura de una era, perteneciente a un conjunto de eras predefinido por el simulador.) Cada solicitud de tasación de pintura está asociado a una era concreta. El único atributo de una pintura (conocido por el agente al recibir la solicitud de tasación) es la era a que pertenece y por tanto es el único que define los requerimientos del usuario.

Es válido decir, que cada pintura también tiene asociado un atributo que define su valor real. Al solicitar la tasación de la pintura, este atributo es de dominio del simulador. Al final del proceso de tasación, el simulador da a conocer el valor real de la pintura para que cada tasador realice las evaluaciones pertinentes y actualice su modelo.

En la versión actual de ART, la única información útil para el modelo de confianza y reputación es la pertenencia de la pintura a una era determinada. Sin embargo, sería conveniente que al tasar una pintura, perteneciente a una era sobre la que no se dispone de suficiente información, se pudiera tener en cuenta el conocimiento asociado a la tasación de pinturas de otra era, quizás similar.

De manera sencilla y sin falta de generalidad, se puede proponer a ART adoptar un modelo ontológico para la representación de las pinturas, que, entre otras cuestiones, permita la definición de cada pintura a partir de un conjunto de atributos que la describan. De esta manera, pueden adoptarse, por ejemplo, los conceptos de *Goal* y *Web Service* de WSMO (consultar las secciones 4.2.2 y 4.4) para representar las solicitudes de tasación de una pintura y la valoración real de la misma, respectivamente.

Siguiendo este enfoque, la representación de una Pintura, definida mediante el concepto de *Goal* de WSMO, estaría determinada por un conjunto de atributos. Además de la referencia a la era a que pertenece, se pueden incorporar atributos tales como: pintor, estilo, técnica empleada, etc.

De este modo, el concepto de Pintura puede ser utilizado por los modelos para establecer mecanismos de intuición de los valores de confianza y reputación. Los valores

correspondientes a pinturas de eras sobre las que no se disponga de información podrían ser estimados a partir del conocimiento sobre otras pinturas que se han tasado anteriormente, teniendo en cuenta el grado de similitud entre ellas.

A continuación se comentan algunas consideraciones relacionadas con el hecho de tener en cuenta la similitud entre pinturas.

8.5.2. Similitud entre dos pinturas

El considerar o no la similitud entre pinturas en las estrategias de los agentes que utilizan ART, es un hecho que permanece fuera del alcance del escenario. Corresponde al modelo de confianza y reputación implementarlas. Como se ha comentado anteriormente, la definición de medidas de similitud entre las pinturas requiere una representación semántica adecuada del concepto. Tal como se describen las pinturas en la versión actual de ART (únicamente a partir de la era a que pertenece) no es posible establecer medidas de similitud que permitan intuir los valores de confianza y reputación cuando el conocimiento no es suficiente.

Una vez adoptado el modelo de representación ofrecido por WSMO donde la pintura queda definida por un conjunto de atributos, se puede definir la similitud entre dos pinturas. Por la naturaleza del dominio de aplicación tratado, donde los atributos propuestos para la representación de la pintura toman valores cualitativos, pudiera pensarse en emplear la medida de similitud de Zadeh [149], estudiada en la sección 4.3.2. Esta medida se basa en la agregación de un conjunto de criterios de comparación de la similitud entre los atributos o grupos de atributos de los conceptos comparados.

Siguiendo este enfoque, la confianza en un tasador para una pintura perteneciente a una era desconocida (y por consiguiente el valor de la pintura) puede estimarse a partir del conocimiento sobre otra pintura similar. Por ejemplo, pudieran agregarse dos criterios: 1) pinturas del mismo autor y 2) pinturas de la misma técnica y estilo. Así de manera conveniente la confianza será intuida a partir del conocimiento que se tenga sobre otras pinturas del mismo autor, o de otros autores, pero que compartan técnica y estilo.

Por supuesto, al igual que con cualquier otra variante para la definición de la similitud entre pinturas, los criterios de similitud utilizados y la forma de agregarlos depende del criterio seguido por el modelo implementado en cada agente.

Parte IV

Conclusiones

Capítulo 9

Conclusiones y trabajos futuros

Este capítulo concluye los trabajos presentados en la tesis y ofrece nuevas ideas para el desarrollo de futuras investigaciones. La sección 9.1 resume las principales contribuciones y resultados alcanzados, mientras que la sección 9.2 relaciona algunos aspectos que aun requieren esfuerzos de investigación.

9.1. Conclusiones

En los últimos años se han desarrollado varios modelos de confianza y reputación como solución a diferentes problemas prácticos. De manera general, cada modelo asume un marco de representación propio, dependiente del dominio de aplicación tratado. Esto ha dado lugar a la existencia de una gran diversidad de propuestas de representación de la información de confianza y reputación, así como de la forma de utilizarla en las decisiones de los agentes, fundamentalmente. Así, se advierte la necesidad de adoptar un marco conceptual común para la representación de los elementos que dependen del dominio de aplicación en un modelo de este tipo. Éste debe brindar los conceptos y herramientas necesarias para la representación y manipulación eficiente de los elementos más importantes de un modelo de confianza y reputación.

Por otra parte, Becerra *et al.* [11] han propuesto un estudio que pretende unificar los atributos y métodos utilizados en los modelos de confianza y reputación. En él destacan la evaluación de la intuición como un método de gran utilidad para la obtención de información de confianza y reputación. La consideración de este método puede resultar muy complejo, sin embargo, pueden diseñarse algunos mecanismos simples que permitan intuir la confianza o reputación que no tengan en cuenta la información específica relacionada con el comportamiento de un determinado agente.

La definición y evaluación del modelo TRSIM, propuesto en esta tesis, se basa en ambas consideraciones: (1) emplea el marco conceptual ofrecido por WSMO para la representación de los elementos del modelo dependientes del dominio de aplicación, y (2) implementa mecanismos de intuición de la confianza y la reputación a partir de la combinación del conocimiento de confianza y reputación con la similitud entre los requerimientos de los

usuarios. De esta forma, se ha contrastado la hipótesis de trabajo planteada mediante la consecución de los objetivos trazados, obteniendo los siguientes resultados:

- Se ha definido y evaluado un modelo de confianza y reputación que permite proponer las soluciones más adecuadas a los requerimientos (o preferencias) de los usuarios en un entorno de interacción entre pares, bajo condiciones de desconocimiento. A partir de los eventuales requerimientos de los usuarios, el modelo es capaz de brindar las soluciones idóneas a partir del conocimiento almacenado sobre las interacciones que han tenido lugar con anterioridad.
- Se ha adoptado el marco conceptual ofrecido por WSMO para representar los conceptos y funciones del modelo, dependientes del dominio de aplicación, tras identificar, en esta tesis, los elementos útiles para el tratamiento de la confianza y la reputación.
- Se ha propuesto la utilización de criterios de intuición para estimar los valores de confianza y reputación cuando no es suficiente el conocimiento que se tiene sobre el desempeño del sistema en interacciones previas. Demostrando que, en dichas situaciones, la consideración de mecanismos de intuición mejora el desempeño del sistema al encontrar la solución adecuada a un determinado requerimiento de usuario.
- Se ha demostrado que el modelo propuesto garantiza mecanismos adaptativos para soportar los cambios de comportamiento de los agentes en el sistema.

Además, en esta tesis se han descrito dos escenarios de referencia en los cuales el modelo resulta útil, considerando que se aprovechan sus funcionalidades. El primero de ellos, al que se ha llamado Consumidor - Proveedor de recursos, hace referencia al problema práctico de encontrar el proveedor que garantice los mayores niveles de satisfacción para las necesidades de un consumidor. Teniendo en cuenta este escenario, se ha mostrado cómo los agentes pueden representar y utilizar, en sus interacciones, los elementos del modelo TRSIM.

El otro escenario abordado, constituye una recomendación de extensión al escenario de prueba ART a partir del enriquecimiento semántico del concepto de pintura, y de la posibilidad de considerar la similitud entre pinturas.

A pesar de que ART, en su versión actual, no permite representar en su totalidad las funcionales de TRSIM, se ha adaptado el modelo de confianza y reputación a dicho escenario de prueba. Se ha mostrado cómo el escenario ha sido utilizado para el estudio del comportamiento del modelo de confianza. Se han realizado algunas evaluaciones que permiten la comparación de diferentes estrategias de los agentes que utilizan TRSIM. Seleccionando las estrategias adecuadas para las decisiones del agente que utiliza TRSIM, se han implementado dos agentes que han participado en sendas competiciones entre modelos de confianza y reputación que utilizan ART.

9.1.1. Conclusiones extraídas de la experimentación

En la tesis se presentan un conjunto de evaluaciones experimentales relacionadas con el estudio del comportamiento del modelo ante diferentes alternativas de configuración, teniendo en cuenta que los agentes son estacionarios, no cambian la calidad de la respuesta. Se ofrece otro grupo de experimentos que estudian la capacidad de adaptación de TRSIM a determinados cambios en el comportamiento de los agentes en el sistema. A continuación se comentan las principales conclusiones que se pueden extraer de los experimentos realizados desde cada punto de vista.

9.1.1.1. Comportamiento del modelo en agentes estacionarios

Se han realizado un grupo de experimentos que consideran que los agentes no cambian su comportamiento a lo largo de toda la simulación. Con ellos, se ha analizado el desempeño del modelo frente a diferentes alternativas para la consideración de elementos importantes en su configuración, tales como: la forma de obtener la calidad de una respuesta a partir de la tarea que satisface, la forma de combinar la confianza y la reputación con la similitud entre tareas, la función utilizada para calcular la similitud entre dos tareas, entre otras. A continuación se puntualizan algunas de las conclusiones más importantes:

- A pesar de que todas las variantes analizadas para obtener la calidad de una respuesta, dada la tarea que satisface, muestran un comportamiento similar, la alternativa Q_b es la que garantiza la mejor diferenciación de los valores de confianza en los agentes de calidad *Alta* del resto de los agentes.
- El modelo incrementa el grado de satisfacción con la respuesta recomendada cuando se estiman, de una forma adecuada, los valores de confianza y reputación para tareas desconocidas, con respecto a los casos donde no se realiza ninguna estimación. IT_c constituye la mejor alternativa para aproximar estos valores a partir del conocimiento disponible sobre alguna tarea similar.
- El problema de aproximar la confianza y reputación para tareas desconocidas es independiente de la forma de obtención de la similitud entre tareas y sí del procedimiento utilizado para combinar la similitud y los valores de la confianza y la reputación sobre alguna tarea conocida.
- El modelo es capaz de mostrar comportamientos diferentes de acuerdo a la calidad de la respuesta que ofrecen los diferentes tipos de agentes. En todas las alternativas analizadas, la confianza en los agentes que ofrecen soluciones de calidad *Alta* tiende a crecer, la confianza en los que ofrecen soluciones de calidad *Baja* tiende a disminuir, y la confianza en los que ofrecen soluciones de calidad *Media* se mantiene en torno a un valor medio.

9.1.1.2. Adaptación a los cambios de comportamiento de los agentes

Se han realizado un conjunto de experimentos que han permitido estudiar la capacidad de adaptación del modelo frente a algunos cambios en el comportamiento de varios tipos de agentes. Tanto cuando varía el comportamiento de un único agente como cuando lo hace un grupo de ellos, el modelo es capaz de ajustar, por medio de los valores de confianza, el criterio que se tiene en los agentes que varían la calidad de las soluciones que ofrecen.

En los casos en que se empeora la calidad de uno o varios agentes de calidad *Alta*, el modelo se desestabiliza. Sin embargo, tras un número de rondas relativamente pequeño, el modelo vuelve a alcanzar la estabilidad.

Cuando el cambio de comportamiento se produce en agentes de calidad *Baja*, el modelo no se desestabiliza. Es capaz de reflejar el cambio de comportamiento de estos agentes, aumentando la confianza en ellos. Cuando los cambios se producen en entornos donde no existen agentes de calidad *Alta*, aumentan la satisfacción con la respuesta recomendada por el modelo y la confianza en el agente que la recomienda.

9.2. Trabajos Futuros

9.2.1. TRSIM en otros escenarios

TRSIM ha sido pensado para ser utilizado en un escenario del tipo Consumidor - Proveedor de recursos en que las interacciones entre los agentes por lo general no son mal intencionados (no mienten, no cambian su identidad para mejorar su imagen frente a otros, etc.). Por lo tanto, los mecanismos que implementa en este sentido son muy simples, pudiendo ser adecuados en algunos casos e insuficientes en otros. Por ejemplo:

- La obtención de la confianza incluye medidas de confiabilidad a partir de la desviación de las experiencias que se utilizan, así, un agente de comportamiento muy variable puede tener menos confianza que uno de comportamiento estable. Sin embargo, el agente que mantenga un comportamiento estable en sus mentiras puede no ser detectado por el modelo.
- El modelo es capaz de representar los cambios de comportamiento de los agentes proveedores, utilizando los valores de confianza. Sin embargo, resulta vulnerable cuando los agentes toman una nueva identidad.
- El modelo tras cada solicitud de información de reputación evalúa al agente que ofrece la recomendación. Sin embargo, éste es un mecanismo insuficiente para detectar y manipular posibles coaliciones entre varios agentes.

Por ello, la posibilidad de aplicar TRSIM a otros escenarios supone la implementación de mecanismos adecuados para el tratamiento de las distintas condiciones que se presenten en el dominio de aplicación.

9.2.2. Evaluación del comportamiento grupal o de sistema

El estudio del desempeño del modelo ha estado centrado en el análisis del comportamiento de un conjunto de métricas individuales y grupales. Se han tomado métricas que han sido utilizadas para el análisis de otros modelos o en la definición de algunos escenarios de prueba. Por lo general, las métricas recomendadas para el análisis del comportamiento grupal han sido obtenidas a partir de la agregación de las medidas individuales de cada uno de los agentes componentes del grupo. A partir de este estudio se ha notado que la definición y utilización de medidas grupales, que sirvan para evaluar el comportamiento de un grupo de agentes desde el punto de vista de la confianza y la reputación, ha sido un tema poco abordado y que requiere esfuerzos de investigación.

En este sentido, se hace necesario definir métricas grupales (o relacionadas con el sistema en su totalidad) que reflejen de alguna medida la sinergia que se obtiene al emplear sistemas multi-agente en la solución de determinados problemas. Como ideas preliminares, pudiera pensarse en métricas asociadas a evaluar el valor de las interacciones entre los agentes de un mismo grupo y/o con agentes externos al grupo, el grado de confiabilidad, cohesión, confianza, reputación entre los agentes del grupo, entre otras.

9.2.3. Confianza desde el punto de vista del proveedor

TRSIM, al igual que otros modelos de confianza y reputación, se utilizan para identificar los agentes adecuados en una determinada interacción. Por lo general, el planteamiento y desarrollo de los modelos es enfocado desde el punto de vista del agente iniciador de la interacción, como criterio para seleccionar con quién interactuar.

TRSIM ha sido desarrollado como herramienta para que los agentes consumidores (que desempeñan el rol de iniciador de la interacción) seleccionen los proveedores adecuados. Cada consumidor es capaz de representar el conocimiento de confianza y reputación sobre el comportamiento de los proveedores, utilizarlo en sus decisiones y luego actualizarlo a partir de las soluciones obtenidas como respuesta a sus requerimientos. Para la actualización del modelo se utilizan criterios que consideran (1) la satisfacción de los requerimientos de los usuarios a partir de la solución brindada, y (2) el cumplimiento del compromiso de calidad que ha hecho el agente proveedor durante la interacción.

En los experimentos presentados en esta tesis se ha considerado que los agentes proveedores siempre se comprometen a ofrecer su mejor respuesta (garantizar la mayor calidad posible). Sin embargo, esta simplificación no es válida en otros escenarios o situaciones. Desde el punto de vista del proveedor, el agente debe poder decidir su compromiso de calidad en dependencia del consumidor, sus requerimientos, las propias soluciones que puede ofrecer, etc. En otras palabras, el modelo de confianza debe hacerse extensible al rol de proveedor (respondedor) sirviendo como elemento para la toma de decisiones de este tipo de agentes.

9.2.4. Extensión de ART. Comparación con otros modelos

Durante la tesis se ha analizado ART como escenario para la evaluación de TRSIM y para la comparación con otros modelos. Sin embargo, los elementos fundamentales de TRSIM (la representación ontológica de varios elementos del modelo y la consideración de la similitud como mecanismo de intuición) no pueden ser reflejados en ART. En este sentido, la evaluación de TRSIM ha estado relacionada con el análisis del comportamiento del modelo ante determinadas situaciones, la comparación de diferentes alternativas de configuración y la evaluación de la capacidad de adaptación ante cambios en el comportamiento de los agentes, fundamentalmente.

Mediante los experimentos desarrollados se demuestra que utilizando dicha representación y dicho mecanismo de intuición el modelo se comporta de manera deseada, según el conjunto de parámetros estudiados en el escenario de prueba Consumidor - Proveedor. Por lo tanto, tal como se discute en la sección 6.3, se recomienda la inclusión de estas ventajas de TRSIM en el escenario ART, contribuyendo así al desarrollo de este escenario.

Una vez obtenido un escenario donde se consideren todos los elementos de TRSIM, se podrá comparar su comportamiento con el de otros modelos.

9.2.5. Confianza en el escritorio social semántico

TRSIM constituye una integración entre la representación utilizada por una tecnología de la web semántica, WSMO, y los modelos de confianza y reputación. WSMO es una propuesta de ontología para la descripción de servicios web semánticos y por lo tanto factible para ser utilizada en escenarios del tipo Consumidor - Proveedor de recursos.

Por otra parte, relacionadas con la web semántica se desarrollan otras tecnologías que abordan escenarios diferentes. Un ejemplo lo constituye la infraestructura de colaboración en grupo conocida como Escritorio Social Semántico (en inglés: *Social Semantic Desktop*) [41]. Esta tecnología ha surgido a partir del desarrollo de puntos comunes de la web semántica, las redes P2P (basadas en la interacción entre pares) y las redes sociales en línea, fundamentalmente. El escritorio social semántico es un nuevo entorno de colaboración que permite la creación, distribución y desarrollo de datos y metadatos.

NEPOMUK ¹ es uno de los proyectos pioneros en el desarrollo de esta tecnología. Tiene como objetivo potenciar el conocimiento (y acciones) de cada individuo a partir de una explotación eficiente de sus datos personales y de la comunicación e intercambio fructíferos con sus colegas, dentro de los límites de la red de relaciones sociales de una organización.

En NEPOMUK se aúnan los esfuerzos de los investigadores, desarrolladores de software industrial y representantes de los usuarios, para extender el concepto de *escritorio personal* dentro de un entorno colaborativo. El objetivo fundamental del proyecto es garantizar la gestión eficiente de la información personal y potenciar la comunicación a partir del intercambio a través de las relaciones sociales y organizacionales.

¹NEPOMUK - the Social Semantic Desktop. <http://nepomuk.semanticdesktop.org/>

Como se evidencia claramente, el proyecto NEPOMUK y el concepto del escritorio social semántico definen un amplio campo de investigación donde resulta interesante la integración de técnicas y modelos de confianza y reputación.

Bibliografía

- [1] *Introduction to UDDI: Important Features and Functional Concepts*. <http://uddi.org/pubs/uddi-tech-wp.pdf>, 2004.
- [2] A. Abdul-Rahman and S. Hailes. Supporting Trust in Virtual Communities. In *Proc. of the Hawaii's International Conference on Systems Sciences*. Maui, Hawaii, 2000.
- [3] R. Akkiraju, J. Farrell, J. Miller, M. Nagarajan, A. Sheth, and K. Verma. *Web Service Semantics - WSDL-S*. W3C, <http://www.w3.org/2005/04/FSWS/Submissions/17/WSDL-S.htm>, 2005.
- [4] Ana Mas. *Agentes Software y Sistemas Multiagente. Conceptos, Arquitecturas y Aplicaciones*. Pearson, 2005.
- [5] G. Antoniou, , and F. van Harmelen. *A Semantic Web Primer*. The MIT Press, 2004.
- [6] ART Testbed Team. *Agent Reputation and Trust Testbed*. <http://www.art-testbed.net/>, 2007.
- [7] R. Axelrod. *The Evolution of Cooperation*. Basic Books, 1985.
- [8] R. A. Baeza-Yates and B. A. Ribeiro-Neto. *Modern Information Retrieval*. ACM Press / Addison-Wesley, 1999.
- [9] K. S. Barber, K. Fullam, and J. Kim. Challenges for Trust, Fraud and Deception Research in Multi-agent Systems. In R. Falcone, K. S. Barber, L. Korba, and M. P. Singh, editors, *Trust, Reputation, and Security: Theories and Practice*, LNCS 2631, pages 167–174. Springer, 2002.
- [10] S. Battiston, F. E. Walter, and F. Schweitzer. Impact of Trust on the Performance of a Recommendation System in a Social Network. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. of Ninth Workshop on Trust in Agent Societies at AAMAS*, pages 1 – 8, May 2006.
- [11] G. Becerra, J. Heard, R. Kremer, and J. Denzinger. Trust attributes, methods, and uses. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. of the Tenth Workshop on Trust in Agent Societies*, pages 1 – 6, May 2007.

-
- [12] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, May 2001.
- [13] D. Booth and C. K. Liu. *Web Services Description Language (WSDL) 1.1*. <http://www.w3.org/TR/2007/REC-wsdl20-primer-20070626/>, March 2001.
- [14] R. A. Bourne, C. B. Excelente-Toledo, and N. R. Jennings. Run-Time Selection of Coordination Mechanisms in Multi-Agent Systems. In W. Horn, editor, *Proc. of 14th European Conference on Artificial Intelligence*, pages 348–352, 2000.
- [15] I. Bratko. *Prolog Programming For Artificial Intelligence*. Addison Wesley, 2000.
- [16] D. Brickley and R. Guha. *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C, <http://www.w3.org/TR/rdf-schema/>, February 2004.
- [17] A. Caballero, J. Botía, and A. Skarmeta. El papel de los agentes FIPA en aplicaciones basadas en Servicios Web. In *Proc. of CAEPIA2003 Conferencia de la Asociación Española de Inteligencia Artificial*, pages 249–252, Noviembre 2003.
- [18] A. Caballero, J. Botía, and A. Skarmeta. An approach to achieve intelligence into an agent-based system to contract information services. In *Proc. of XII Congreso Latino Iberoamericano de Investigación de Operaciones CLAIO*, 2004.
- [19] A. Caballero, J. Botía, and A. Skarmeta. Modelo de los datos incorporados en los agentes fipa que conforman un sistema distribuido para la invocación automática de servicios web. *Rev. Ingeniería Industrial*, XXV(3):56–61, 2004.
- [20] A. Caballero, J. Botía, and A. Skarmeta. A New Model for Trust and Reputation Management with an Ontology Based Approach for Similarity Between Tasks. In E. A. K. Fischer, I. J. Timm and N. Zhong, editors, *Multiagent System Technologies*, LNCS 4196, pages 172–183. Springer, September 2006.
- [21] A. Caballero, J. Botía, and A. Skarmeta. Trust and Reputation Model based on WSMO. In J. S.-M. R. Falcone, S. Barber and M. Singh, editors, *Proc. of Ninth Workshop on Trust in Agent Societies at AAMAS*, pages 9–18, May 2006.
- [22] A. Caballero, J. Botía, and A. Skarmeta. Adecuación del modelo de confianza y reputación TRSIM a ART. In J. Carbó, J. Sabater, and J. M. Molina, editors, *Proc. of Agents in ART testbed Spanish competitions Workshop at CAEPIA2007*, pages 16–21, November 2007.
- [23] A. Caballero, J. Botía, and A. Skarmeta. Assessing the Effectivity of TRSIM Model for Trust and Reputation. In R. Bordini, editor, *Proc. of Fifth European Workshop on Multi-Agent Systems*, pages 38–52, December 2007.

- [24] A. Caballero, J. Botía, and A. Skarmeta. On the behaviour of TRSIM model for trust and reputation. In M. K. P. Petta, J. Muller and M. Georgeff, editors, *Multiagent System Technologies*, LNCS 4687, pages 182–193. Springer, September 2007.
- [25] A. Caballero, J. Botía, and A. Skarmeta. On the Convenience of Using Similarity Between Tasks in a Trust and Reputation Model for Contracting Web Services. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. of Tenth Workshop on Trust in Agent Societies at AAMAS*, pages 22–30, May 2007.
- [26] A. Caballero, J. Botía, and A. Skarmeta. Adaptability of the trsim model to some changes in agents behaviour. In *Enviado a: Autonomic and Trusted Computing*, LNCS xxxx. Springer, 2008.
- [27] J. Carbó, J. M. Molina, and J. Dávila. Trust Management Through Fuzzy Reputation. *International Journal of Cooperative Information Systems*, 12(1):135–155, 2003.
- [28] J. Carbo. *Reputation through Agent Mediated Electronic Commerce*. PhD thesis, Universidad Carlos III de Madrid, 2002.
- [29] B. Carre. *Graphs and networks*. Oxford: Clarendon Press., 1978.
- [30] J. Carter, E. Bitting, and A. Ghorbani. Reputation formalization for an information-sharing multi-agent system. *Computational Intelligence*, 18(2):515–534, 2002.
- [31] S. Casare and J. Sichman. Towards a Functional Ontology of Reputation. In F. Dignum, V. Dignum, S. Koenig, S. Kraus, M. P. Singh, and M. Wooldridge, editors, *4rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005)*, pages 505–511, New York, NY, USA, 2005. ACM Press.
- [32] C. Castelfranchi, R. Conte, and M. Paolucci. Normative reputation and the costs of compliance. *Journal of Artificial Societies and Social Simulation*, 1(3), 1998.
- [33] C. Castelfranchi and R. Falcone. Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification. In *In Proceedings of the International Conference on Multi-Agent Systems*, pages 72 – 79, 1998.
- [34] P. Castells. La web semántica. In C. Bravo and M. A. Redondo, editors, *Sistemas Interactivos y Colaborativos en la Web*, pages 195–212, 2003.
- [35] E. Chang, F. K. Hussain, and T. S. Dillon. Reputation Ontology for Reputation Systems. In R. Meersman et al., editor, *On the Move to Meaningful Internet Systems: OTM Workshops*, LNCS 3762, pages 957–966. Springer, 2005.

-
- [36] E. Chang, F. K. Hussain, and T. S. Dillon. Towards defining an ontology for reputation. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2601–2607, October 2005.
- [37] L. Codina and C. Rovira. La Web Semántica. In *Tendencias en documentación digital*, pages 9–54. Trea, 2006.
- [38] F. Curbera, Y. Golland, J. Klein, F. Leymann, D. Roller, S. Thatte, and S. Weerawarana. *Business Process Execution Language for Web Services*. <http://download.boulder.ibm.com/ibmdl/pub/software/dw/specs/ws-bpel/ws-bpel1.pdf>, July 2002.
- [39] P. Dasgupta. Trust as a commodity. *Trust: Making and Breaking Cooperative Relations*, D. Gambetta (ed.), pages 49–72, 1998.
- [40] J. de Bruijn, H. Lausen, R. Krummenacher, A. Polleres, L. Predoiu, M. Kifer, and D. Fensel. *D16.1v0.2 The Web Service Modeling Language WSML. WSML Final Draft 20 March 2005*. W3C, <http://www.wsmo.org/TR/d16/d16.1/v0.2>, 2005.
- [41] S. Decker and M. Frank. The Social Semantic Desktop. Technical Report 2004-05-02, DERI - Digital Enterprise Research Institute, May 2004.
- [42] J. Domingue, D. Fensel, and D. Roman. Semantic Web Services with the Web Service Modeling Ontology (WSMO). *Agentlinks News*, (19):7–9, November 2005.
- [43] T. Erl. *Service-Oriented Architecture (SOA): Concepts, Technology, and Design*. Service-Oriented Computing. Prentice Hall, 2005.
- [44] B. Esfandiari and S. Chandrasekharan. On how agentes make friends: Mechanisms for trust acquisition. In *Proc. of the Fourth International Workshop on Deception, Fraud and Trust in Agent Societies*, pages 27–34, 2001.
- [45] C. B. Excelente-Toledo and N. R. Jennings. Learning when and how to coordinate. *Int J. of Web Intelligence and Agent Systems*, 1(3-4):203–218, 2003.
- [46] P. Faratin, C. Sierra, and N. Jennings. Using similarity criteria to make issue tradeoffs in automated negotiations. *Artificial Intelligence*, 142(2):205–237, 2002.
- [47] D. Fensel and C. Bussler. The Web Service Modeling Framework WSMF. *Electronic Commerce Research and Applications*, 1(2), 2002.
- [48] Foundation for Intelligent Physical Agents, <http://www.fipa.org/>. *FIPA Contract Net Interaction Protocol Specification SC00029*, 2002.
- [49] K. Fullam and K. S. Barber. Evaluating Approaches for Trust and Reputation Research: Exploring a Competition Testbed. In M. Paolucci, J. Sabater, R. Conte, and C. Sierra, editors, *Proc. of IAT-2004 Workshop on Reputation in Agents Societies*, pages 20–23, 2004.

- [50] K. K. Fullam, T. B. Klos, G. Muller, J. Sabater, A. Schlosser, Z. Topol, S. barber, J. S. Rosenschein, L. Vercouter, and M. Voss. A Specification on the agent Reputation and Trust (ART) Testbed: Experimentation and Competition for Trust in Agent Societies. In *Proc. of the 4th International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 512–518, July 2005.
- [51] F. García. *Sistema Basado en Tecnologías del Conocimiento para Entornos de Servicios Web Semánticos*. PhD thesis, Facultad de Informática. Universidad de Murcia, 2007.
- [52] J. Golbeck. Trust and Nuanced Profile Similarity in Online Social Networks. *Submitted to ACM Transactions on the Web. MINDSWAP Tech Report TR-MS1291*, 2007.
- [53] J. Golbeck and J. Hendler. Trust Networks on the Semantic Web. In *Proc. of Cooperative Information Agents*, August 2003.
- [54] J. Golbeck and J. Hendler. Accuracy of Metrics for Inferring Trust and Reputation in Semantic Web-based Social Networks. In *Proc. of EKAW'04*, 2004.
- [55] J. Golbeck and J. Hendler. Reputation Network Analysis for Email Filtering. In *Proc. of the Conference on Email and Anti-Spam (CEAS)*, July 2004.
- [56] J. Golbeck and J. Hendler. FilmTrust: Movie Recommendations Using Trust in Web-based Social Networks. In *Proc. of the IEEE Consumer Communications and Networking Conference*, January 2005.
- [57] N. Griffiths. Task Delegation using Experience-Based Multi-Dimensional Trust. In C. Castelfranchi, S. Barber, J. Sabater, and M. Singh, editors, *Proc. of Eighth Workshop on Trust in Agent Societies at AAMAS*, 2005.
- [58] N. Griffiths. A Fuzzy Approach to Reasoning with Trust, Distrust and Insufficient Trust. In M. Klusch, M. Rovastos, and T. Payne, editors, *Cooperative Information Agents X*, LNCS 4149, pages 360–374. Springer, 2006.
- [59] N. Griffiths. Enhancing peer-to-peer collaboration using trust. *Expert Systems with Applications*, 31(4):849–858, 2006.
- [60] N. Griffiths, K.-M. Chao, and M. Younas. Fuzzy Trust for Peer-to-Peer Systems. In *Proc. of the 26th IEEE International Conference Workshops on Distributed Computing Systems*, page 73, July 2006.
- [61] N. Griffiths and M. Luck. Coalition formation through motivation and trust. In *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 17–24, 2003.

-
- [62] N. Griffiths, M. Luck, and M. d’Inverno. Cooperative Plan Selection Through Trust. In F. Garijo and M. Boman, editors, *Proc. of the Ninth European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW’99)*, pages 137–155, 1999.
- [63] N. Griffiths, M. Luck, and M. ”d’Inverno”. Annotating Cooperative Plans with Trusted Agents. In S. Poslad, K. S. Barber, and J. Kim, editors, *Proc. of the Fifth International Workshop on Deception, Fraud and Trust in Agent Societies*, pages 137–155, 2002.
- [64] T. R. Gruber. A Translation Approach to Portable Ontology. *Knowledge Acquisition*, 5(2):199–220, 1993.
- [65] N. Gujral, D. DeAngelis, K. K. Fullam, and K. S. Barber. Modelling Multi-Dimensional Trust. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. Ninth Workshop on Trust in Agent Societies at AAMAS*, pages 35 – 41, May 2006.
- [66] J. Hefin, J. Hendler, and S. Luke. SHOE: A Knowledge Representation Language for Internet Applications. Technical Report CS-TR-4078, 1999.
- [67] R. Hermoso, H. Billhardt, and S. Ossowski. Effective Use of Organisational Abstractions for Confidence Models. In G. O’Hare, A. Ricci, M. O’Grady, and O. Dikenelli, editors, *Engineering Societies in the Agents World VII*, LNCS 4457. Springer, September 2007.
- [68] F. K. Hussain, E. Chang, and T. S. Dillon. Defining Reputation in Service Oriented Environment. In *Proc. of International Conference on Internet and Web Applications and Services*, 2006.
- [69] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. Developing an Integrated Trust and Reputation Model for Open Multi-Agent Systems. In R. Falcone, S. Barber, J. Sabater, and M. Singh, editors, *AAMAS-04 Workshop on Trust in Agent Societies*, 2004.
- [70] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. FIRE: An Integrated Trust and Reputation Model for Open Multi-Agent Systems. In *Proc. of the 16th European Conference on Artificial Intelligence (ECAI)*, 2004.
- [71] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. On Handling Inaccurate Witness Reports. In C. Castelfranchi, S. Barber, and M. S. J. Sabater, editors, *Proc. of Eighth Workshop on Trust in Agent Societies at AAMAS*, pages 63–77, 2005.
- [72] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. An Integrated Trust and Reputation Model for Open Multi-Agent Systems. *Journal of Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, 2006.

- [73] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. Certified Reputation: How an Agent Can Trust a Stranger. In *AAMAS '06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1217–1224, 2006.
- [74] N. R. Jennings. An agent-based approach for building complex software systems. *Communications of the ACM*, 44(4):35–41, 2001.
- [75] A. Josang, S. Hird, and E. Faccer. Simulating the Effect of Reputation Systems on e-Markets. In *Proc. of the First International Conference on Trust Management (iTrust'03)*, May 2003.
- [76] A. Josang and R. Ismail. The Beta Reputation System. In *Proc. of the 15th Bled Electronic Commerce Conference*, June 2002.
- [77] Y. Katz and J. Golbeck. Social Network-based Trust in Prioritized Default Logic. In *Proc. of the Twenty-First National Conference on Artificial Intelligence (AAAI 06)*, 2006.
- [78] Y. Katz and J. Golbeck. Using Social Network-based Trust For Default Reasoning On The Web. *Submitted to Journal of Web Semantics*, 2007.
- [79] T. Kerkiri, A. Manitsaris, and I. Mavridis. E-Reputation Ontology for Adaptation in E-Learning. In *Proc. of International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL)*, 2006.
- [80] I. V. Levenshtein. Bynary Codes capable of correcting deteleations, insertions, and reversals. *Cybernetics and Control Theory*, 10(8):707–710, 1966.
- [81] P. W. Lord, R. D. Stevens, A. Brass, and C. A. Goble. Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. *Bioinformatics*, 19(10):1275–1283, 2003.
- [82] D. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.
- [83] A. Maedche and S. Staab. Measuring Similarity between Ontologies. In *Proc. of the European Conference on Knowledge Acquisition and Management - EKAW-2002*, LNCS 2473, pages 251–263. Springer, October 2002.
- [84] S. Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, Department of Mathematics and Computer Science, University of Stirling, 1994.
- [85] E. M. Maximilien and M. P. Singh. A Framework and Ontology for Dynamic Web Services Selection. *IEEE Internet Computing*, 08(5):84–93, 2004.

-
- [86] E. M. Maximilien and M. P. Singh. Toward autonomic web services trust and selection. In *ICSOC '04: Proceedings of the 2nd international conference on Service oriented computing*, pages 212–221, 2004.
- [87] E. M. Maximilien and M. P. Singh. Agent-based trust model involving multiple qualities. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 519–526, 2005.
- [88] E. M. Maximilien and M. P. Singh. Self-adjusting trust and selection for web services. *icac*, 00:385–386, 2005.
- [89] R. C. Mayer, J. H. Davis, and D. F. Schoorman. An integration model of organizational trust. *The Academy of Management Review*, 20(3):709734, July 1995.
- [90] D. L. McGuinness and F. van Harmelen. *OWL Web Ontology Language Overview*. W3C, <http://www.w3.org/TR/owl-features/>, February 2004.
- [91] T. M. Mitchell. *Machine Learning*. Computer Science Series. McGraw Hill International Editions, 1997.
- [92] N. Mitra and Y. Lafon. *SOAP Version 1.2 Part 0: Primer*. <http://www.w3.org/TR/2007/REC-soap12-part0-20070427/>, June 2002.
- [93] L. Molm, N. Takahashi, and G. Peterson. Risk and trust in social exchange: An experimental test of a classical proposition. *American Journal of Sociology*, 105(5):1396–1427, 2000.
- [94] L. Mui, A. Halberstadt, and M. Mohtashemi. Notions of Reputation in Multi-Agents Systems: A Review. In C. Castelfranchi and L. Johnson, editors, *Proc. of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'2002)*, pages 280 – 287, 2002.
- [95] L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt. Ratings in Distributed Systems: A Bayesian Approach. In *Workshop on Information Technologies and Systems (WITS'2001)*, 2001.
- [96] L. Mui, M. Mohtashemi, and A. Halberstadt. A Computational Model of Trust and Reputation. In *Proc. of 35th Hawaii International Conference on System Science*. IEEE Computer society, 2002.
- [97] L. Mui, P. Szolovits, and C. Ang. Collaborative sanctioning: applications in restaurant recommendations based on reputation. In *AGENTS '01: Proceedings of the fifth international conference on Autonomous agents*, pages 118–119, New York, NY, USA, 2001. ACM Press.

- [98] G. Muller, L. Vercouter, and O. Boissier. A Trust Model for the Reliability of Agent Communications. In C. Castelfranchi, S. Barber, J. Sabater, and M. Singh, editors, *Proc. of Eighth Workshop on Trust in Agent Societies at AAMAS*, pages 106–118, 2005.
- [99] G. O’Hare and N. Jennings, editors. *Foundations of Distributed Artificial Intelligence*. Sixth-Generation Computer Technology Series. John Wiley and Sons, 1996.
- [100] OWL-S Team. *OWL-S 1.2*. DAML Services, <http://www.daml.org/services/owl-s/>.
- [101] J. Patel, W. T. Teacy, N. R. Jennings, and M. Luck. A Probabilistic Trust Model for Handling Inaccurate Reputation Sources. In *Proc. of the 16th European Conference on Artificial Intelligence (ECAI)*, pages 193–209, 2005.
- [102] I. Pinyol, J. Sabater, and G. Cuni. How to Talk About Reputation Using a Common Ontology: From Definition to Implementation. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. of Tenth Workshop on Trust in Agent Societies at AAMAS*, pages 90–101, 2007.
- [103] M. Prietula. Advice, trust, and gossip among artificial agents. *Dynamics of Organizations*, A. Lomi and E. R. Larson (eds.), pages 141–177, 2000.
- [104] S. Ramchurn, D. Huynh, and N. Jennings. Trust in Multi-Agent Systems. *Knowledge Engineering Review*, 1(19):1–25, 2004.
- [105] S. D. Ramchurn, N. R. Jennings, C. Sierra, and L. Godo. A computational Trust Model for Multi-Agent Interactions based on Confidence and Reputation. In R. Falcone, S. Barber, L. Korba, and M. Singh, editors, *Proc. of the 6th International Workshop on Trust, Privacy, Deception and Fraud in Agent Systems (AAMAS’03) July 2003 Melbourne, Australia*, volume Workshop, pages 69–75. ACM Press, 2003.
- [106] A. Rao and M. Wooldridge. *Foundations of Rational Agency*, volume 14 of *Applied Logic Series*, chapter Foundations of Rational Agency, pages 1–10. Springer, 1999.
- [107] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Proc. of ACM 1994 Conference on Computer Supported Cooperative Work*, pages 175–186, Chapel Hill, North Carolina, 1994. ACM.
- [108] M. Richardson, R. Agrawal, and P. Domingos. Trust management for the semantic web. In *Proc. of Second International Semantic Web Conference*, 2003.
- [109] M. A. Rodriguez and M. J. Egenhofer. Determining Semantic similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, 15(2):442–456, 2003.

-
- [110] S. Russell and P. Norving. *Inteligencia Artificial: Un enfoque moderno*. Prentice Hall, 1995.
- [111] J. Sabater. *Trust and reputation for agent societies*. PhD thesis, Universidad Autónoma de Barcelona, 2002.
- [112] J. Sabater. Toward a test-bed for trust and reputation models. In R. Falcone, S. Barber, J. Sabater, and M. Singh, editors, *Proc of. AAMAS-04 Workshop on Trust in Agent Societies*, pages 101–105, 2004.
- [113] J. Sabater and C. Sierra. REGRET: a Reputation Model for Gregarious Societies. In C. Castelfranchi and L. Johnson, editors, *Proc. of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS '2002)*, pages 475–482, 2002.
- [114] J. Sabater and C. Sierra. Social ReGreT, a reputation model based on social relations. *ACM SIGecom Exchanges*, 3(1):44–56, 2002.
- [115] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Application of Dimensionality Reduction in Recommender Systems. In *Proc. of ACM WebKDD Workshop*. Boston, USA, August 2000.
- [116] M. Schillo, P. Funk, and M. Rovatsos. Using Trust for Detecting Deceitful Agents in Artificial Societies. *Applied Artificial Intelligence, Special Issue on Trust, Deception, and Fraud in Agent Societies*, 14(8):825–848, 2000.
- [117] A. Schlosser, M. Voss, and L. Brückner. On the Simulation of Global Reputation Systems. *Journal of Artificial Societies and Social Simulation*, 9(1):4–19, 2005.
- [118] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [119] U. Shardanand and P. Maes. Social Information Filtering: Algorithms for Automating "Word of Mouth". In *Proc of ACM CHI'95 Conference on Human Factors in Computing Systems*, volume 1, pages 210–217, 1995.
- [120] C. Sierra and J. Debenham. An information-based model for trust. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 497–504, 2005.
- [121] C. Sierra and J. Debenham. Trust and honour in information-based agency. In *AAMAS '06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1225–1232, 2006.
- [122] R. G. Smith. The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. *IEEE TRANSACTIONS ON COMPUTERS*, 29(12):1104–1113, 1980.

- [123] J. P. Soto, A. Viscaino, J. Portillo-Rodríguez, and M. Piattini. Managing Trustworthy Knowledge with Intelligent Agents. In *Memorias del Taller DESMA'2007 en el Congreso Español de Informática (CEDI'2007)*.
- [124] W. Stallings. *Cryptography and Network Security*. Prentice Hall, 4 edition edition, 2005.
- [125] SWSF Team. *Semantic Web Services Framework (SWSF) Overview*. W3C, <http://www.w3.org/Submission/SWSF/>, 2005.
- [126] W. T. Teacy, J. Patel, N. R. Jennings, and M. Luck. "TRAVOS": Trust and Reputation in the Context of Inaccurate Information Sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.
- [127] A. Tversky. Features of Similarity. *Psychological review*, 84(4):327–352, 1977.
- [128] L. Vercouter, S. J. Casare, J. S. Sichman, and A. Brandão. An Experience on Reputation Models Interoperability Based on a Functional Ontology. In M. M. Veloso, editor, *Proc. of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*, pages 617–622, January 2007.
- [129] L. Viljanen. Towards an Ontology of Trust. In H. Federrath and D. H. M. Casassa, editors, *Trust, Privacy and Security in Digital Business*, pages 175–184, August 2005.
- [130] W3C, <http://www.w3.org/2001/sw/>. *Semantic Web*, 2007.
- [131] Y. Wang and M. P. Singh. Trust Representation and Aggregation in a Distributed Agent System. In *Proc. of the 21st National Conference on Artificial Intelligence (AAAI)*, pages 1425–1430, 2006.
- [132] Y. Wang and M. P. Singh. *Trust via evidence combination: A mathematical approach based on certainty*. TR 2006-11, North Carolina State University, Raleigh, 2006.
- [133] Y. Wang and M. P. Singh. Formal Trust Model for Multiagent Systems. In M. M. Veloso, editor, *Proc. of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*, pages 1551–1556, 2007.
- [134] Y. Wang and J. Vassileva. Trust and Reputation Model in Peer-to-Peer Networks. In *Proc. of IEEE Conference on P2P Computing*. Linköping, Sweden, September 2003.
- [135] Y. Wang and J. Vassileva. Trust and reputation model in peer-to-peer networks. In *Peer-to-Peer Computing*, pages 150–157, 2003.
- [136] Y. Wang and J. Vassileva. Trust-Based Community Formation in Peer-to-Peer File Sharing Networks. In *Proc. of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 341–348, 2004.

-
- [137] Y. Wang and J. Vassileva. A Review on Trust and Reputation for Web Service Selection. In *Proc. of the 27th International Conference on Distributed Computing Systems Workshops*, page 25, 2007.
- [138] G. Weiss, editor. *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, 2001.
- [139] M. Witkowski, A. Artikis, and J. Pitt. Experiments in Building Experiential Trust in a Society of Objective-trust Based Agents. In *Trust in Cyber-societies, Integrating the Human and Artificial Perspectives*, LNCS 2246, pages 111–132, 2001.
- [140] WSMO Team. *WSMO Web Service Discovery. WSML Working Draft*. W3C, <http://www.wsmo.org/2004/d5/d5.1/v0.1/>, 2005.
- [141] T. Yamagishi, K. Cook, and M. Watabe. Uncertainty, trust, and commitment formation in the United States and Japan. *American Journal of Sociology*, 104(1):165–194, 1998.
- [142] B. Yu and M. P. Singh. A Social Mechanism of Reputation Management in Electronic Communities. In M. Klusch and L. Kerschberg, editors, *Cooperative Information Agents IV - The Future of Information Agents in Cyberspace*, LNCS 1860, pages 154–165, 2000.
- [143] B. Yu and M. P. Singh. An Evidential Model of Distributed Reputation Management. In C. Castelfranchi and L. Johnson, editors, *Proc. of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS '2002)*, pages 294–301, 2002.
- [144] B. Yu and M. P. Singh. Distributed reputation management for electronic commerce. *Computational Intelligence*, 18(4):535–549, 2002.
- [145] B. Yu and M. P. Singh. Searching Social Networks. In J. S. Rosenschein, T. Sandholm, M. Wooldridge, and M. Yokoo, editors, *AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, pages 65–72, New York, NY, USA, 2003. ACM Press.
- [146] G. Zacharia. *Collaborative reputation mechanisms for communities*. PhD thesis, Massachusetts Institute of Technology, 1999.
- [147] G. Zacharia and P. Maes. Trust management through reputation mechanisms. *Applied Artificial Intelligence*, 14(9):881–907, October 2000.
- [148] L. A. Zadeh. Fuzzy Sets. *Information and Control*, (8):338–353, 1965.
- [149] L. A. Zadeh. Similarity relations and fuzzy orderings. *Information Sciences*, 3:177–200, 1971.

- [150] J. Zhang and R. Cohen. A Personalized Approach to Address Unfair Ratings in Multiagent Systems. In R. Falcone, S. Barber, J. Sabater-Mir, and M. Singh, editors, *Proc. of Ninth Workshop on Trust in Agent Societies at AAMAS*, pages 89–98, May 2006.
- [151] C.-N. Ziegler and J. Golbeck. Investigating interactions of trust and interest similarity. *Decision Support Systems*, 43:460 – 475, 2007.

