UNIVERSIDAD DE MURCIA

FACULTAD DE INFORMÁTICA

Formal Framework of the Capacity of Adaptation on Information Ubiquitous Systems Based on Ambient Intelligence

Modelo Formal de la Capacidad de Adaptación en Sistemas de Información Ubicuos Basados en Inteligencia Ambiental

Dña. María Teresa García Valverde
2012

UNIVERSIDAD DE MURCIA

Facultad de Informática

FORMAL FRAMEWORK OF THE CAPACITY OF ADAPTATION ON INFORMATION UBIQUITOUS SYSTEMS BASED ON AMBIENT INTELLIGENCE

MODELO FORMAL DE LA CAPACIDAD DE ADAPTACIÓN EN SISTEMAS DE INFORMACIÓN UBICUOS BASADOS EN INTELIGENCIA AMBIENTAL

Tesis Doctoral

Presentada por:

D^a. María Teresa García Valverde

Supervisada por:

Dr. D. Antonio Fernando Gómez Skarmeta

Dr. D. Juan Antonio Botía Blaya

Murcia, Julio de 2012

UNIVERSITY OF MURCIA

Faculty of Computer Science

FORMAL FRAMEWORK OF THE CAPACITY OF ADAPTATION ON INFORMATION UBIQUITOUS SYSTEMS BASED ON AMBIENT INTELLIGENCE

MODELO FORMAL DE LA CAPACIDAD DE ADAPTACIÓN EN SISTEMAS DE INFORMACIÓN UBICUOS BASADOS EN INTELIGENCIA AMBIENTAL

PhD Thesis

Author:

D^a. María Teresa García Valverde

Thesis Advisors:

Dr. D. Antonio Fernando Gómez Skarmeta

Dr. D. Juan Antonio Botía Blaya

Murcia, July 2012

A D^a. Susana y D. Rafael,
los imprescindibles.

Agradecimientos

Es sorprendente la cantidad de gente que aprovecha este capítulo para llorar por el largo y arduo camino y agradecer a la gente que los ha acompañado en ese valle de lágrimas. Siento no poder estar más en desacuerdo. Debo haber tenido mucha suerte con los que me han acompañado en este proceso, pues para mí, han sido unos años maravillosos. Si pudiera, volvería a empezar de nuevo, una y mil veces.

Y esto se lo debo en especial a mi director Juan Botía, por ser una fuente de ideas y conocimiento que a día de hoy me sigue sorprendiendo. Gracias por haber compartido parte de eso conmigo y por transmitirme la ilusión por este mundo, por disfrutar tantísimo con este trabajo. Desde *El miedo a la página en blanco* y *El origen de las especies* supe que había tomado la decisión correcta. Has sido mucho más que mi director, por eso, a ti en particular, te pido disculpas por todo lo aguantado, por no estar a la altura, por mis cabezonerías y por no demostrarte día a día lo eternamente agradecida que estoy de que seas mi director.

Tengo que agradecer además a mi otro director, Antonio Skarmeta, que me dieran esta oportunidad y me permitieran trabajar en un grupo tan competente. La capacidad de trabajo que aquí se observa es un ejemplo a seguir.

En especial también tengo que agradecer a mi madre, incansable educadora, y a mi padre, mi mejor admirador, gracias por su confianza ciega, dedicación absoluta y apoyo incondicional durante toda mi vida. Todavía oigo a veces el *Aprovecha el tiempo*. Vosotros me habéis enseñado el amor por aprender, por superarse, por luchar por lo que uno quiere. Gracias por ese regalo, y principalmente, gracias a los dos por quererme más de lo que pudiera soñar.

A mi hermano le doy las gracias por ser un ejemplo. El brillo de tus ojos cuando hablas de la producción de hidrógeno o del nuevo microcontrolador es una fuente de energía inagotable. Eres un ejemplo a seguir de investigador apasionado, te agradezco que me hayas transmitido siempre esa ilusión y me hayas animado a continuar.

A mis amigas de Bullas, a *estas*, las de siempre. Por todo lo recorrido juntas, por aceptar sin cuestionar, por ser un refugio, el lugar dónde ir sin ser juzgado. A mis siempre compañeras de piso, especialmente a Ana Villa, por hacer del *44campanilla* un hogar y de una desconocida un apoyo.

Gracias también a mis compañeros de carrera porque todavía hoy me siguen alegrando la vida. Y tras la carrera, a mis compañeros de laboratorio por crear un buen ambiente de trabajo.

Un agradecimiento especial es para mi abuela. Una abuela diferente, adaptada a todos los tiempos y defensora de lo suyo. Gracias por renunciar tantas veces a mi sin queja y sobre todo, por *alegrarte en mi contento*.

Gracias también al resto de mi familia, tanto la de siempre como la nueva, por el apoyo y por tantos domingos ausentes.

Mil gracias por supuesto a mis compañeros de la 2.6 (los ex, los que están y los que nunca han estado). Los buenos ratos que he pasado con vosotros son inolvidables, he aprendido la teoría de cuerdas, los afluentes del Segura, y sobre todo, a cómo reírte un lunes cualquiera. En especial, tengo que agradecer a Andrés su compañerismo sincero, a Paco L. todo lo que me ha enseñado, a Paco 2 y Pablo lo mucho que me hacen reír, a Jesús por esas charlas de café, y a mi Precio, a ti Vicky tengo que agradecerte mucho este año, gracias por todo.

Gracias también a los que me acompañaron en mi estancia en Colchester. Aquí especialmente tengo que dirigirme a Hani Hagras, por enseñarme y ofrecerme tanto. Ha sido un placer poder trabajar con alguien como él tanto profesional como personalmente, y siempre le agradeceré esa oportunidad, porque, *at the end of the day*, me doy cuenta de lo mucho que he aprendido de él. En Colchester también tengo que dar las gracias a mis amigos de allí, en especial a Ays y Dario, porque me hicieron muy fácil vivir en un país desconocido.

Agradecimientos menores, aunque no menos importantes, se merece mi gato, por todas las horas que ha pasado debajo de un flexo al lado del ordenador. Y por supuesto, todos los libros y canciones que han sido banda sonora de esta tesis.

Por último, y como manda la tradición, te doy las gracias a ti. En esta tesis has sido de todo, compañero, corrector, revisor, coautor, consultor, ... pero principalmente has sido mi mayor apoyo, has aguantado altibajos, cambios de opinión, esperas, cabreos y todo tipo de tonterías. Siempre me has animado, aplacado mi inseguridad y ofrecido un plan B cuando venían las dudas. El día que me prometiste mejorarme la vida, no imaginaba que llegarías tan lejos.

Gracias.

Contents

Capítulo 1

Resumen

1.1 Motivación y Objetivos

El paradigma de la **Inteligencia Ambiental (AmI)** [1, 3, 10, 31] persigue como objetivo proporcionar entornos dotados de interfaces inteligentes e intuitivas que están integradas transparentemente en los objetos que nos rodean, y son capaces de detectar, responder y adaptarse a nuestras necesidades, hábitos y deseos.

Desde esta definición podemos detectar las tecnologías de las que deriva la AmI. En primer lugar, la capacidad de que el *entorno esté dotado de interfaces inteligentes integradas transparentemente* proviene de la **Computación Ubicua** [36, 38, 39]. El paradigma de la computación ubicua apuesta por la tecnología que desaparece, es decir, entornos donde las capacidades de computación están incluidas en objetos cotidianos de forma que son imperceptibles para sus usuarios. Decimos así que estos entornos son ubicuos, debido a la disponibilidad de tecnología en cualquier parte, y no intrusivos, ya que no alteran el entorno habitual de sus usuarios.

Por otra parte, decimos que el *entorno está dotado de interfaces inteligentes e intuitivos*. El paradigma de las **Interfaces de Usuario Inteligentes** [25,33] defiende el desarrollo de dispositivos e interfaces que interactúan con las personas sin interferir en su comportamiento cotidiano. Siguiendo la definición de AmI, un *entorno AmI es capaz de detectar usuarios e información relativa a ellos*. Esta capacidad nos ofrece la información que deseamos sobre el usuario y su hábitat en el momento adecuado y de la forma apropiada. Esto es lo que llamamos sensibilidad al contexto y se refiere al paradigma del **Context-Awareness** [2, 7].

Por último y parte clave en AmI, los *entornos son capaces de responder y adaptarse a las necesidades, hábitos y deseos de las personas*. En AmI, los servicios y aplicaciones usan la información de contexto para ofrecer beneficios a los usuarios en una forma proactiva. Así, la **Inteligencia** y la **Adaptación** se convierten aquí en cualidades clave. Estas características permiten que aplicaciones y servicios sean personalizados y se adapten a las condiciones cambiantes del usuario y el entorno, y actúen con proactividad tomando decisiones autónomamente para anticiparse al usuario cuando la situación lo requiera.

El campo de la Inteligencia Ambiental ha crecido mucho en los últimos años y se han aportado diversas propuestas y soluciones orientadas a la consecución e integración de sus beneficios en nuestras vidas. Sin embargo, actualmente ninguna de estas propuestas es capaz de ofrecer una solución que proporcione la capacidad de adaptación inherente a los sistemas AmI de una forma genérica, i.e., adaptable a diferentes tipos de escenarios y situaciones, y global, i.e. que aborde todas las características y restricciones de este tipo de sistemas desde una fase inicial de diseño hasta el desarrollo final.

Derivado de esta necesidad surge la motivación global y principal objetivo de esta tesis: **la definición de un marco genérico para proporcionar capacidades de adaptación en sistemas software complejos como los basados en Inteligencia Ambiental**. Así, en el ámbito de esta tesis, lo que pretendíamos era preguntarnos sobre la naturaleza de la adaptación en sistemas AmI (i.e. altamente cambiantes, ubicuos, no intrusivos, centrados en el usuario, con alto número de entidades en el sistema y funcionamiento a largo plazo). El principal resultado de esta actividad pretendía generar un modelo formal y de desarrollo en el cual, en cualquier tipo de despliegue de un sistema computacional, la capacidad de adaptación se ofreciera de manera ortogonal a servicios y aplicaciones para ofrecer beneficios a los usuarios.

Como vemos, en este tipo de sistemas, el usuario es la principal entidad actora. De hecho, no sólo es entidad del sistema, si no que es el beneficiario último de la adaptación. Para integrarlo de manera natural en nuestro modelo necesitábamos un modelo del usuario que considerase tanto sus características físicas (edad, altura, etc.) como sus intereses, preferencias y comportamientos. Por tanto, gran parte del esfuerzo de esta tesis ha sido dirigido hacia el modelado de usuarios en este tipo de entornos. Esta actividad de modelado nos llevó al uso de técnicas de inteligencia computacional. Estas técnicas nos permiten por una lado, el modelado de conocimiento complejo de los usuarios, y por otro, dotar a la infraestructura que pretendíamos desarrollar de una capacidad de adaptación que le permitiera interactuar de una manera cada vez más cómoda y natural con las aplicaciones y con otros usuarios del sistema.

Otra característica fundamental para proveer adaptación en AmI es la información de contexto. Era necesario ser capaces de ofrecer una caracterización del contexto de cada entidad del sistema suficientemente detallada, actualizada y funcional. Además, especialmente la información sobre la localización de los usuarios en AmI es primordial. Surgía, de este modo, la necesidad de localizar usuarios en entornos altamente cambiantes, aprendiendo y adaptándose a estos cambios de forma efectiva y perdurable en el tiempo. De hecho, mucho del trabajo de esta tesis ha estado orientado en esta dirección.

Derivado de las necesidades descritas y de las propias características de los sistemas AmI, era necesario enfocar la tesis tanto a mecanismos y técnicas de aprendizaje, como al completo desarrollo y despliegue de estas en escenarios reales, respetando las restricciones impuestas por los entornos AmI y ofreciendo una integración real y natural en ellos. Por esta razón, una de las actividades más importantes dentro de esta tesis consistió en proveer al marco general de adaptación de la posibilidad de coexistencia de elementos reales y simulados para acercar los modelos teóricos a la realidad.

Con esto, podemos resumir los objetivos específicos perseguidos para la consecución del objetivo global de esta tesis doctoral en:

1. Identificar y analizar los requisitos y necesidades de sistemas de información ubicuos basados en Inteligencia Ambiental.
2. Analizar el estado del arte actual en este tipo de sistemas, identificando aquellas deficiencias que imposibilitan el avance hacia una verdadera, sostenible y perdurable adaptación en AmI.
3. Establecer las bases y requisitos que garantizan la capacidad de adaptación en sistemas AmI y que permiten al sistema adecuarse eficientemente a cambios y nuevas situaciones que se produzcan en el entorno y sus usuarios.
4. Definir formalmente y modelar aquellas entidades necesarias en sistemas AmI considerando el alto dinamismo y heterogeneidad de usuarios, fuentes de información y escenarios.
5. Ofrecer una manera de diseñar, evaluar y desarrollar este tipo de sistemas que permita su posterior aplicación al mundo real.
6. Como parte inherente de la adaptación en sistemas AmI, proporcionar mecanismos capaces de localizar a los usuarios, que respeten las restricciones de estos entornos y provean una solución efectiva.
7. Validar estas propuestas de forma que se demuestre la viabilidad de su integración en diferentes escenarios reales.

1.2 Resultados

El trabajo realizado durante esta tesis ha dado lugar a valiosos resultados, de los cuales, gran parte se encuentran descritos en los artículos que la componen. Dados los requisitos de esta modalidad de presentación de tesis doctoral, no todo artículo fruto de esta tesis ha sido incluido en el compendio de publicaciones (la lista completa puede verse en la sección 8.2). Durante el transcurso de esta tesis la doctoranda ha participado en diferentes proyectos y colaboraciones que han dado lugar a resultados que, pese a no ser incluidos en esta tesis, han enriquecido el trabajo llevado a cabo dentro de la línea de investigación del doctorado [42,48]. A pesar de esto, los resultados más relevantes respecto al trabajo de tesis realizado, sí se encuentran descritos en los artículos presentados.

Así, y derivado del objetivo específico 1) descrito en la sección anterior, hemos analizado y definido las características propias y deseables de los sistemas AmI. Asimismo, hemos caracterizado la problemática subyacente a los sistemas de información ubicuos basados en AmI.

Tras este análisis, y planteado como objetivo 2) de esta tesis, llevamos a cabo una profunda revisión de la literatura existente en este tipo de sistemas. Esta revisión dio como resultado un estudio sobre las diferentes aportaciones, propuestas previas y trabajos realizados en adaptación en sistemas AmI.

Este resultado junto con la caracterización y problemática asociada a entornos AmI reveló la necesidad de una arquitectura capaz de considerar todas las entidades integradas y requisitos en AmI. Además, permitió detectar y establecer aquellas entidades y elementos que deben ser susceptibles de consideración en este tipo de sistemas para proveer mecanismos adaptativos eficaces y perdurables en el tiempo. Esto dio lugar a la consecución del objetivo específico 3) planteado en esta tesis.

Como parte del objetivo 4), analizamos las entidades detectadas en el objetivo 3). Para ello, definimos formalmente y modelamos estas entidades, haciendo especial hincapié en el modelado de la entidad usuario, entidad clave de un sistema AmI. De hecho, realizamos un análisis y aplicamos algoritmos de aprendizaje al modelado del comportamiento de personas reales viviendo en entornos AmI gracias a la integración de la doctoranda en varios proyectos con escenarios reales (casas particulares, hospital, residencia, oficinas) [57]. El análisis de estos patrones de comportamiento reales permitió la abstracción y desarrollo de un modelo completo de usuario que es capaz describir un usuario real en diferentes entornos [50].

Otra entidad clave en computación ubicua y AmI es el contexto. En esta tesis hemos propuesto gestionar el contexto asociado a las diferentes entidades mediante el middleware OCP (Open Context Platform). OCP es una herramienta desarrollada en la Universidad de Murcia. Dada la importancia del contexto en un entorno ubicuo, la doctorando participó en la mejora de este middleware, añadiéndole las características necesarias para su integración en un marco capaz de ofrecer adaptación: histórico de datos, razonamiento contextual y mejora del rendimiento para su ejecución online [44–46].

Finalmente, integramos los modelos de las entidades obtenidas en el objetivo 4) en el framework AmISim presentado en esta tesis [41, 58]. AmISim permite reproducir los modelos teóricos de las entidades desarrolladas sobre diferentes entornos [49]. Así, es posible validar diferentes técnicas de aprendizaje y observar su capacidad de adaptación en entornos relativamente complejos (edificios de oficinas, sistemas de evacuación de incendios, casas particulares, etc.), que suelen ser imposibles de validar en la realidad. AmISim ofrece además la posibilidad de que elementos reales y simulados coexistan, incrementando así el realismo obtenido [59, 60].

AmISim cumple el objetivo 5) de esta tesis, ofreciendo una manera de diseñar, evaluar y desarrollar sistemas AmI que, por un lado, reduce errores y esfuerzo invertido en desarrollos, y por otro, amplía la variedad de escenarios que pueden ser validados. La efectividad de AmISim ha sido probada sobre varios escenarios y de hecho, ha sido muy bien acogido por la comunidad científica de este campo, dando lugar a varias publicaciones y ponencias en congresos [43, 47].

Como hemos descrito, uno de los requisitos más importantes para los servicios en entornos AmI es que sean conscientes de la presencia del usuario y de su contexto y, por tanto, tengan la capacidad de localizarlos para satisfacer sus necesidades. En este sentido, el objetivo 6) de esta tesis estaba orientado a ofrecer servicios de localización capaces de ofrecer consciencia de la localización del usuario en entornos

AmI. Sin embargo, los servicios en un entorno AmI deben ser no intrusivos, proteger la privacidad del usuario y no requerir un despliegue muy costoso. Debido a esto, hemos afrontado el problema de la localización en AmI en *entornos controlados*: entornos donde las restricciones permiten una mínima capacidad de control sobre el despliegue e información del escenario, y en *entornos no-controlados*: entornos en los cuales tanto infraestructura, usuarios, cómo dispositivos son completamente desconocidos.

En ambos casos, el objetivo 7) que nos habíamos planteado respecto la validación de la viabilidad de las propuestas en escenarios reales ha podido ser realizado gracias al framework de AmISim, que nos ha permitido diseñar, testear y validar estar propuestas para su posterior integración en escenarios reales.

Para entornos controlados, hemos diseñado y desarrollado un servicio basado en localización para edificios inteligentes sobre los que se dispone alguna información y control. Así, el servicio de localización propuesto es capaz de predecir la localización en interiores mediante diversas técnicas de aprendizaje automático [51, 52]. El uso de AmISim junto con el servicio de localización han sido utilizados para incluir un servicio adaptativo capaz de optimizar el despliegue de la infraestructura en un entorno AmI de forma automática.

En los experimentos realizados el servicio consigue mantener y garantizar una alta precisión mientras que el coste y despliegue de la infraestructura es reducido hasta en un 50 %. Si consideramos el uso de tecnología RFID (por ejemplo, Speedway Revolution RFID Reader ® con un coste por unidad de 1300 €) en un edificio con 26 habitaciones, la solución obtenida por nuestro servicio no sólo es un 50 % menos intrusiva, sino que además supone un ahorro de casi 17.000 € en el presupuesto [53, 54].

Para el caso de entornos no-controlados, hemos desarrollado un sistema difuso de localización en interiores para entornos AmI. Este sistema no utiliza conocimiento a priori del entorno. Por tanto, cualquier entorno puede disponer de la solución de forma no intrusiva y con bajo coste, características esenciales en AmI. Este servicio es capaz de aprender automáticamente de los cambios y variaciones del entorno y adaptarse de forma online a las diferentes condiciones para proveer la localización de usuarios. Esto se realiza mediante un aprendizaje incremental a largo plazo que ajusta su comportamiento a las variaciones y cambios de las señales, entorno y usuarios. Hemos validamos el servicio sobre escenarios simulados y reales. En los experimentos sobre escenarios reales, el sistema consiguió hasta un 77 % de precisión. Los experimentos han mostrado que el servicio ofrece una solución efectiva y con coste cero, ya que no necesita despliegue de nuevo hardware [55, 56].

La tabla 1.1 resume los resultados presentados de acuerdo al objetivo planteado y la categoría ingenieril a la que pertenecen.

Desde los resultados, se observa que los objetivos planteados en la planificación de esta tesis han sido alcanzados con éxito, proporcionando resultados valiosos que ofrecen un acercamiento real hacia la integración de sistemas de Inteligencia Ambiental en nuestra vida diaria.

Objetivo	**Resultado**	**Índole**
Objetivo 1: Identificar y analizar los requisitos y necesidades de sistemas de información ubicuos basados en AmI	Análisis de los sistemas AmI y tecnologías de las que deriva: Computación Ubicua, Context-Aware, Interfaces de Usuario Inteligentes, Inteligencia y Adaptación	Estudio
Objetivo 2: Analizar el estado del arte actual en sistemas AmI e identificar deficiencias para dotarlos de adaptación	Estado del arte de las propuestas para dotar de adaptación a los sistemas AmI	Estudio
Objetivo 3: Establecer las bases y requisitos que garantizan la capacidad de adaptación en sistemas AmI	Detección de los requisitos y entidades necesarias para la adaptación en AmI	Análisis
Objetivo 4: Definir formalmente y modelar las entidades necesarias en sistemas AmI	Modelado del entorno, usuarios, contexto y adaptación	Diseño y Modelado
Objetivo 5: Ofrecer una manera de diseñar, evaluar y desarrollar sistemas AmI para su aplicación al mundo real	Framework AmISim	Implementación
Objetivo 6: Proporcionar mecanismos capaces de localizar usuarios	Algoritmos de localización y despliegue de infraestructura	Implementación
Objetivo 7: Validar las propuestas de forma que se demuestre la viabilidad de su integración en diferentes escenarios reales	Validación de la viabilidad de las propuestas presentadas	Validación

Cuadro 1.1: Objetivos y resultados

1.3 Conclusiones y Trabajos futuros

El desarrollo de sistemas complejos como los basados Inteligencia Ambiental supone un gran reto. Esto es debido a la naturaleza intrínseca de este tipo de entornos, en los cuales características como la ubicuidad, transparencia al usuario y no intrusión son requisitos fundamentales. Además, el usuario, como entidad central en este tipo de sistemas, manifiesta comportamientos complejos que pueden cambiar a lo largo del tiempo y situaciones. En este complejo contexto, los servicios en AmI deben ser capaces de aprender y adaptarse a los usuarios y sus necesidades de forma que ofrezcan un comportamiento proactivo tanto a corto como a largo plazo.

Las propuestas planteadas hasta ahora en la literatura, ofrecen soluciones para situaciones específicas y bajo determinadas condiciones; donde la mayoría no son capaces de adaptarse a las condiciones cambiantes de usuarios y entorno de una forma efectiva y duradera.

En este sentido, los resultados obtenidos en esta tesis proporcionan un acercamiento hacia una verdadera realización de entornos ubicuos, adaptativos e inteligentes en diversas situaciones. De hecho, las contribuciones de esta tesis no están sólo enfocadas hacia metodologías o algoritmos para AmI, si no que intentan proveer una solución genérica que aborde desde el diseño y desarrollo de este tipo de sistemas en su planteamiento inicial hasta el despliegue final sobre entornos reales.

Los resultados presentados proveen una solución innovadora no sólo desde una perspectiva científica, sino que además ofrecen una solución práctica que aporta beneficios desde una perspectiva industrial respetando las condiciones de los entornos AmI, de forma que puedan ser integrados en nuestra vida diaria para ofrecernos beneficios como usuarios, sin suponer grandes y costosos despliegues.

El trabajo realizado durante esta tesis puede ser continuado en diversas formas. En primer lugar, actualmente, se está trabajando en asegurar la robustez de las propuestas presentadas sobre diferentes dispositivos móviles, como PDAs, smartphones, etc. Al mismo tiempo, estamos trabajando sobre mecanismos que permitan la inclusión en AmISim de algunos elementos complejos no incluidos en este momento, por ejemplo condiciones ambientales o entornos exteriores. En tercer lugar, hemos abierto una nueva línea de trabajo para la integración de diferentes tecnologías para la localización (RFID, WiFi, Bluetooth, etc.) que permitan la coexistencia de diferentes señales como una unidad.

Estas tres líneas están orientadas a ofrecer robustez y flexibilidad en la instalación y desarrollo sobre entornos reales. El objetivo es que la implantación industrial de nuestros resultados sea facilitada y alcance un mayor número de ámbitos.

Otra línea de investigación muy prometedora en la que se está trabajando es en la aplicación de los modelos de usuario en entornos donde no sólo hay muchos usuarios sino que además, estos comparten recursos y cooperan entre ellos, provocando tal situación la aparición de comportamientos emergentes y grupales. Una revisión de la literatura sobre este tipo de escenarios revela que en esta dirección hay pocos avances, a pesar de ser un escenario muy frecuente. Algunos de nuestros resultados ya han sido aplicados en proyectos con este tipo de escenarios (por ejemplo, el proyecto CARONTE[1] en geriátricos o CADUCEO[2] en hospitales). Aunque los resultados preliminares están siendo satisfactorios, todavía es necesario mucho trabajo y esfuerzo en esta dirección.

Por último, otra línea de investigación derivada de esta tesis es el desarrollo de sistemas que, usando los modelos y técnicas de localización desarrolladas, puedan aprovechar su capacidad adaptativa en escenarios que presentan restricciones especiales, por ejemplo, aquellos en los que existen dispositivos de capacidades muy limitadas, escenarios con requisitos de ahorro energético, con consideraciones sociales o políticas especiales (museos, edificios históricos, etc.) o con fuertes restricciones de privacidad.

[1] `http://innovacion.grupogesfor.com/web/caronte/inicio`

[2] `http://innovacion.grupogesfor.com/web/caduceo/inicio`

Chapter 2

Abstract

2.1 Motivation and Goals

The Ambient Intelligence (AmI) [1, 3, 10, 31] paradigm pursues the building of environments in which embedded intelligent interfaces surround us, are integrated in everyday objects in a transparent way and are able to sense, response and adapt to our needs, habits and desires.

This definition presents the technologies that lead and support AmI. Firstly, the capability of *building environments in which embedded intelligent interfaces surround us and are integrated in everyday objects in a transparent way* stems from the **Ubiquitous Computing** [36, 38, 39]. Ubiquitous Computing paradigm states the concept of the disappearing computing, where the technology is integrated in our daily lives in an undetectable manner. It provides environments with the features of ubiquity, i.e. computing is everywhere, and non-intrusion, i.e. computing is indistinguishable for users.

Secondly, the AmI definition states *environments with embedded intelligent interfaces.* The **Intelligent User Interfaces** [25, 33] paradigm proposes the use of those devices and interfaces that are able to interact with users without hindering other daily activities of the users. Thirdly, according to the AmI definition *the environment is able to sense our needs, habits and desires.* This ability provides us with the information about users and environment in the appropriated moment and manner. That refers to the awareness of the context, the **Context-Awareness** paradigm [2, 7].

Finally, *environments are able to response and adapt to our needs, habits and desires.* This is essential part in AmI. AmI services and applications use contextual information to offer benefits to users. In fact, **Intelligence** and **Adaptation** are key in AmI. Services and applications with these features are able to personalise and adapt to the users and the changing conditions of the environments. They act proactively making decisions autonomously to anticipate to the users according to the situation.

The Ambient Intelligence field has grown considerably in the last years and several initiatives have been proposed in order to achieve the integration of its benefits in our lives. Nevertheless, no approach is able to offer the capability of adaptation of the AmI systems as a generic solution, i.e. a solution for different scenarios and situations, and complete, i.e. a solution that covers all the features and requirements of this kind of systems and accomplishes its development from the design in an initial phase to the final deployment.

Under these circumstances, the main motivation and objective in this thesis is raised: **definition of a formal framework of the capacity of adaptation on information ubiquitous systems based on Ambient Intelligence**. Thus, in this thesis, we pursued to analyse the nature of the adaptation in AmI systems (i.e. highly dynamic, ubiquitous, non-intrusive environments, user-centered, with several heterogeneous entities and working in a long term). The main result from this activity involved to generate a formal model able to offer the capacity of adaptation to services and applications in any kind of computing system to benefit users.

As described above, users are the main entity in this kind of systems. What is more, they are the final beneficiary of the adaptation. With the aim of integrating users in our model in a natural way, we needed a user model which covers all the features related to users, like physical features (e.g. age, height, etc.), interests, preferences and behaviours. Therefore, a big part of the effort of this thesis has been focused on the user model. This activity led to the use of artificial intelligence techniques. These techniques support the complex model of knowledge of the users and are able to deal with the capacities of adaptation required to interact with users in an easy and natural way.

In order to provide systems with these capacities of adaptation, contextual information is essential. There is a need of a complete characterisation of the context of each entity in the system, which must be detailed, updated and functional. In that sense, it becomes especially important in AmI the location information of the users. It involves the ability to localise the whereabouts of the user in AmI environments to learn and adapt to its changing conditions and face unknown situations in a long term. A large amount of the work done in this thesis has been focused on this direction.

According to these premises and the features of AmI systems, we found the need to focus this thesis on mechanisms and techniques of learning, as well as on the development and deployment of these systems over real environments; preserving the AmI requirements and supporting a real and natural integration. In that context, one of the most important activities accomplished was to make possible in the general framework the coexistence of real and simulated elements in order to bring closer the theoretical and the real world.

Thus, the specific goals that we pursue in this thesis can be summarised as follows:

1. Identify and analyse the requirements and needs of the information ubiquitous systems based on Ambient Intelligence.

2. Analyse the current state of the art of this kind of systems and identify possible deficiencies in the studied proposals which prevent a progress towards a real, sustainable and long term adaptation in AmI.

3. Define the keys and requirements that enable the capacity of adaptation in AmI and allow the system to fit in with the changes and unknown situations of the environment and users.

4. Define and model formally those entities which are needed in AmI systems with high dynamism and heterogeneity of users, information sources and scenarios.

5. Suggest a method able to design, validate and develop this kind of systems to integrate them in the real world.

6. As essential part of the adaptation in AmI, develop mechanisms that are able to localise users effectively, preserving the AmI requirements.

7. Validate our proposals in order to show its suitability to be integrated in different real scenarios.

2.2 Results

The work done in this PhD thesis has obtained valuable results. Most of them are described in the articles that compose it. Nonetheless, due to the requirements of this kind of thesis presentation, not every article or work that derives from this thesis has been included in the publications compilation (the full list can be found in section 8.2). The PhD candidate has participated in different projects and collaborations during this thesis period. This work has obtained also meaningful results, which, despite not being included as part of the publications compilation, have enriched the PhD thesis [42, 48]. In spite of that, the most relevant results of this work are included in the presented publications as part of this thesis.

Thus, as consequence of the specific goal 1), described in the previous section, we have analysed and defined the basic and desirable features of AmI systems. Likewise, we have characterised the problems related to the information ubiquitous systems based on AmI.

Then, as we proposed in goal 2), we accomplished a deep analysis of the current literature in this field. This revision resulted in a study about the different contributions, approaches and works about adaptation in AmI systems.

This study and the characterisation and the problems related to AmI environments showed the need of an architecture able to consider all the entities and requirements of AmI. Furthermore, it allowed us to detect the entities and elements which are needed in this kind of systems in order to provide efficient long term adaptive mechanisms. It satisfied the specific goal 3) proposed in this thesis.

According to goal 4), we analysed the detected entities from goal 3). We formally defined and modelled these entities. A special effort was focused on the model of the entity user because it is a fundamental part in AmI. In this sense, we have done an analysis and learning of the knowledge of real people living in AmI environments. That was possible thanks to the participation of the PhD candidate in projects with real scenarios (homes, hospital, old people's home, office buildings) [57]. With the analysis over real pattern behaviours, the abstraction and deployment of a complete user model was possible. This model is able to describe a real user over different scenarios [50].

Another important key in ubiquitous computing and AmI is context. In this thesis, the OCP (Open Context Platform) middleware has been proposed to manage the contextual information of the AmI entities. OCP is a tool developed at the University of Murcia. Because of the importance of the contextual information in ubiquitous computing, the PhD candidate worked in the improvement of this middleware in order to provide it with the required features for its integration in a general framework able to offer adaptation: historical data, contextual reasoning and online performance improvement [44–46].

Finally, we integrated these models of the entities in the framework AmISim presented in this thesis [41, 58]. AmISim is able to reproduce theoretical models of the defined entities over different environments [49]. AmISim allows us to validate learning techniques and to observe their capacities of adaptation over complex environments (office buildings, emergency systems, homes, etc.), which are usually impossible to validate in the reality. AmISim also offers the possibility of real and simulated elements coexist, increasing the obtained realism [59, 60].

AmISim accomplished the goal 5) of this thesis. It suggests a method to design, validate and develop AmI systems. AmISim minimizes the errors and risks in the final deployments and enables the validation over more scenarios. The effectiveness of AmISim has been shown over different scenarios. In fact, it has received considerable attention by the researches of this field and have led several publications [43, 47].

As described above, one of the most important requirements in AmI is being aware of the context and presence of the user. Thus, there is a need to localise users in order to provide them with the nearest features of interest. Consequently, goal 6) defined in this thesis was focused on localisation services to offer location-awareness to the AmI systems. These services should protect the privacy of the users, be non-intrusive and not require a costly deployment. With these premises, we faced the problem of the localisation in AmI *controlled and uncontrolled environments*. Controlled environments are those where we have some control of the deployment and some available information about the scenario. Uncontrolled environments are those where the information about the infrastructure, the users o the devices is completely unknown.

In the two kinds of environments, the defined goal of validating the suitability of our proposals, goal 7), was achieved by the use of the AmISim framework. AmISim allowed us to design, test and validate our approaches before the final integration in real scenarios.

In the case of controlled environments, we have designed and developed a location based service for intelligent buildings where some information and control are available. The service can predict the location of user in indoor environments by automatic learning techniques [51, 52]. This service and AmISim have been used to include an adaptive service able to automatically optimise the deployment of the infrastructure of an AmI environment.

Our experiments have shown that the service gets a high accuracy while reduces the cost and the size of the infrastructure up to a 50%. In fact, if we consider RFID technology (e.g. Speedway Revolution RFID Reader ®, with a cost per unit of 1300 €) and a building with 26 rooms, the solution obtained for our proposal is a 50% less intrusive, and saves almost 17000 € [53, 54].

In the case of uncontrolled environments, we have developed a fuzzy logic based system of indoor localisation in AmI environments. This system does not need prior information of the environment. Therefore, it can be used in any environment in a non-intrusion and low-cost way, as the AmI paradigm pursues. Additionally, the service is able to automatically learn and adapt to the changing and unknown situations of the environment in order to localise users. This is possible by an incremental lifelong learning which adapts its behaviour according to the changes of signals, environment and users. The service has been validated over simulated and real experiments. Our experiments show that the service can achieve up to 77% of accuracy over real scenarios. The experiments have probed that this service offers an effective zero-cost solution, since no new hardware is required [55, 56].

Table 2.1 summarises the presented results linked to the goal that accomplish and their engineering classification.

From the results, it is clear that the goals pursued in this thesis have been successfully achieved. Our contributions offer valuable results which make closer a real integration of Ambient Intelligence systems to our daily lives.

2.3 Conclusions and Future work

The development of complex systems based on Ambient Intelligence is a hard task. This is caused by the own nature of these systems, where features like ubiquity, transparency and non-intrusion should be assured. Furthermore, the user, as main entity in AmI systems, shows complex behaviours that can change according to the moment and the situation. Under such complex context, AmI services need to learn and adapt to users and their needs in order to response proactively to them in a short and long term.

The existing approaches in the literature offer solutions for specific situations and under certain conditions. Thus, the majority of them are not able to adapt to changing conditions of users and environments in an effective and lifelong manner. In this sense, the results presented in this PhD thesis achieve an approachment towards a truly realization of ubiquitous, adaptive, intelligent environments over different situations. In fact, our contributions can be integrated over different scenarios in an easy way and preserving the desirable conditions of the AmI environments.

The presented results offer novel solutions for the integration of AmI in our daily lives in order to offer users benefits without annoying or costly deployments. We think that they are suitable to be put into practice and will have a considerable impact even from an industrial perspective.

The work done in this thesis can be continued in several directions. Firstly, we are currently working in order to guarantee the reliability of the proposals over different devices, like PDAS, smartphones, etc. At the same time, we are studying mechanisms

Goal	Result	Classification
Goal 1: Identify and analyse the requirements and needs of the information ubiquitous systems based on Ambient Intelligence	Study of the AmI systems and the technologies that lead and support it: Ubiquitous Computing, Context-Aware, Intelligent User Interfaces, Intelligence and Adaptation	Study
Goal 2: Analyse the current state of the art of AmI systems and identify possible deficiencies to provide them with capacities of adaptation	State of art of the current proposals to enable adaptation in AmI systems	Study
Goal 3: Define the keys and requirements that enable the capacity of adaptation in AmI	Identification of the requirements and entities in adaptive AmI systems	Analysis
Goal 4: Formally define and model those entities which are needed in AmI systems	Models of the environment, users, context and adaptation	Design and modelling
Goal 5: Suggest a method able to design, validate and develop this kind of systems to integrate them in the real world	AmISim framework	Implementation
Goal 6: Develop mechanisms that are able to localise users	Algorithms of localisation and deployment of the infrastructure	Implementation
Goal 7: Validate our proposals in order to show its suitability to be integrated in different real scenarios	Validation of the suitability of the presented approaches	Validation

Table 2.1: Goals and results

that allow us to include some complex elements not considered yet in AmISim, for example, weather conditions or outdoor environments. Thirdly, we are interested in a new direction focused on the integration of different localisation technologies (RFID, WiFi, Bluetooth, etc.) working as a unity.

Consequently, these three lines of work are oriented towards increasing the reliability and flexibility of our proposals in the development and deployment over real scenarios. The aim is to facilitate the industrial establishment of our results and to reach more variety of scenarios.

Another new line of promising research, in which we are involved, is the extension of the user model for environments with lots of users, where these users cooperate and share resources, causing emergent group behaviours. Despite this kind of scenarios is quite common, a review of the existing literature in this field shows that little consideration has been given to them. Some of our results have been applied to projects which imply such scenarios (e.g. the CARONTE project[1] in old people's home or the CADUCEO project[2] in hospitals). Our first results are promising, but a lot of work is still needed in this direction.

Finally, another line that arises from this thesis consists in the development of systems that, using the proposed techniques and models, will be able to adapt in scenarios with special constraints, for example, scenarios with resource-limited devices, energy saving impositions, ethical or political considerations (museums, cultural heritage, etc.) or with strong privacy protection needs.

[1] `http://innovacion.grupogesfor.com/web/caronte/inicio`

[2] `http://innovacion.grupogesfor.com/web/caduceo/inicio`

Chapter 3

Readers' Guide

This chapter pretends to be a guide for helping readers to easily find the right material or content in this PhD thesis and to understand some aspects of the structure of this text. Note that this chapter is not an introduction, a summary or a presentation of the articles included as part of the thesis. There are sections specifically intended for that aim.

This thesis is presented under the scheme of publications compilation and, therefore, some requirements have been taken into account in order to accomplish with the regulation of this kind of presentation. In consequence, in chapters 1 and 2 a brief summary of the presented PhD thesis is included in Spanish and English respectively. Both languages are tackled because this thesis is written in English and, according to the regulation, a thesis written in a language different from Spanish and with the international PhD mention request must contain the summary in both languages. Likewise, the regulation requires that the summary involves a brief description of the goals of the PhD and the final conclusions, in which, the results of the papers presented are unified.

According to these criterions, we have included in the summary the following sections 1) *Motivation and Goals* (section 1.1), 2) *Results* (section 1.2) and 3) *Conclusions and Future work* (section 1.3). The first two sections match the requirements of the regulation. The last one is included because we consider that a description of the conclusions obtained is an essential part in a PhD thesis. Similarly, we found that the lines in which the work can be continued are also interesting for any reader of this text.

Following the regulation, chapter 4 describes an introduction to this thesis. In accordance with the regulation, the introduction has to present the articles included as part of the thesis and justify the unity of the works. Under these premises, the introduction presents some fundamental principles and definitions of Ambient Intelligence and the technologies which compound it. Then, the contributions of the thesis are presented explaining the context and the goal that they try to solve as part of the global goal of this thesis. Additionally, the relevance and feasibility of the results are presented, for the individual target of the specific work and for the global goal of the thesis.

The complete contributions can be found in chapter 5. This chapter includes the articles that conform this thesis presented under the scheme of publications compilation. Moreover, as the regulation requires, a full reference of each article, the personal details of the authors and the journals, in which the articles have been published, are included in the section.

Specifically, in the work included in section 5.1, *"Combining the real world with simulations for a robust testing of Ambient Intelligence services"*, is proposed a general architecture for testing, validating and verifying AmI environments: AmISim. This architecture is the first one that is able to cover all the features involved in AmI, i.e. environment model, user model, context model and adaptation model. The work shows that AmISim is a framework to test AmI services and applications in scenarios that would be impossible to enable in real environments with an easier and less costly deployment. This work takes into consideration the goals 1-5 of this thesis (see table 2.1).

According to the focus of this thesis, the adaptation in AmI, our main contributions are given in the fields of the modelled of user and the intelligent and adaptive services to provide users with benefits. These fields are covered under the user and adaptation models in AmISim. Therefore, the rest of the articles presented as part of this thesis contribute to these models.

Thus, given the importance of the user in AmI, the user model of AmISim is tackled in the article presented in section 5.2, entitled *"Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach"*. In this work the problem about how to simulate humans with realistic behaviours is faced. A proposal to achieve realistic behaviours based on social simulation is shown. As the realistic modelling of users as its validation are key elements in this work. A real application example which shows the level of reality reached in the users' models and the benefits of the proposed techniques is presented. This paper accomplishes with part of the goal 4 of this thesis (see table 2.1).

As essential part of the adaptation in AmI, and specifically as part of the adaptation model of AmISim, we proposed the goal 6 of this thesis (table 2.1). This goal was focused on the development of mechanisms able to localise users in controlled and uncontrolled smart environments in order to adapt to users according to their nearest features of interest. Controlled and uncontrolled environments are faced in this PhD thesis. Note that, although different technologies are used in each article, the proposals are general for any kind of technology based on RSSI. In both kinds of environments, the validation of the suitability of our proposals, goal 7 of this thesis (see table 2.1), was achieved by the use of AmISim.

Thus, for controlled environments, the work *"Improving RFID's Location Based Services by means of Hidden Markov Models"*, included in section 5.3, builds a Location Based Service (LBS), using Hidden Markov Models (HMMs), that is able to locate users in an intelligent building. The work shows that, by the use of AmISim, the infrastructure of location can be significantly optimized regarding the number and location of the antennas of the infrastructure.

This work leads us to a multiobjective optimization problem, in which, the best configuration of antennas that minimizes the set of antennas but maximizes the precision of the prediction should be found. With this premise in the article entitled *"Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach"*, in section 5.4, we present a Memetic approach for the multiobjective problem. The presented memetic algorithm uses a combination between metaheuristics and domain knowledge to deal with this problem. The work shows that our proposal is able to improve the LBS in AmI systems while reduces the cost of the infrastructure.

Finally, in order to accomplish the location over uncontrolled environments, the section 5.5, includes the work *"An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments"*. In this work a fuzzy logic based indoor localisation system is proposed at zero-cost deployment with high accuracy. The proposed system is able to adapt online incrementally in a lifelong learning mode to deal with the uncertainties and changing conditions facing unknown indoor structures and without prior knowledge. The system was tested in simulated and real environments and it gave high accuracy to locate users in the different AmI scenarios. In addition, it was able to adapt its behaviour to changes in the environment.

After the articles which compound this thesis, the acceptance letters of the described articles are included in chapter 6. This section is also compulsory according to the regulation. It should contain the acceptance letters of the articles that have been accepted but they are not published yet. We have also introduced the full reference of those articles already published because it can help to better understand the structure of the text.

As part of the information of the journals required for the regulation, chapter 7 defines the publications relevance of the articles included in this thesis. With this aim, we have included for each journal, in which an article has been published, its reference in the ISI Web of Knowledge (`http://www.accesowok.fecyt.es/`). The ISI Web of Knowledge is today's premier academic citation indexing and search service. Here, for each journal the relevance in its field and the impact factor are shown. In the case of the conferences, their description and their raking in the ERA Conference Ranking are provided. ERA Conference Ranking is a well-known ranking of conferences in Computer Science.

Finally, chapter 8 includes the bibliography of this thesis where section 8.1 lists works used for the introduction, and section 8.2 presents a full list of all the publications obtained by the PhD candidate as result of this thesis. Note that not all work resulted of this PhD is included as part of the publications compilation. For this reason, every work obtained in the PhD thesis is explained and linked in section 1.2, where the results of this thesis are presented. Thus, the reader can match each result in that section with the corresponding reference in this part of the bibliography, although the article is not part of the publications compilation.

Chapter 4

Introduction

Ambient Intelligence (AmI) refers to a seamless and invisible computing environment that is "aware" of our presence and context, and that is sensitive, adaptive and responsive to our needs [10]. AmI aims to create digital environments through the usage of embedded and non-intrusive devices which are particularly designed to interact with people in an intelligent and smooth manner. Hence, users are the main entity in AmI systems and every service is user-centered, i.e. services in AmI work for the users, in a proactive and adaptive way.

AmI stems from the convergence of important paradigms: Ubiquitous Computing, Context-Awareness and Intelligent User Interfaces. *Ubiquitous Computing* was introduced by Mark Weiser in his work The Computer for the 21st Century [38]. Weiser envisioned that "the most profound technologies are those that disappear". Thus, ubiquitous computing defines environments where computational capabilities are included into everyday objects in a non-intrusive and invisible way. These smart environments support interaction between environment and users making possible to acquire context knowledge from them, i.e. *Context-Awareness*. *Intelligent User Interfaces* are responsible for this intelligent and smooth interaction in a natural way. This interaction between environment and users makes the services become aware of the users' context with the aim to being able to personalise and adapt to them.

In the literature is usual that these concepts are wrongly used as the same concept, this specially occurs with AmI and ubiquitous computing. As well as ubiquitous computing, AmI also supports the idea of many non-intrusive and invisible devices communicated and integrated into the environment. These devices generate information about the environment, users and the changes in both of them. Nevertheless, in AmI this information is always used to benefit users. In AmI, the human user is the central element, always in control, playing multiple roles in society [31]. The other important point in the definition of AmI is *the need for a "sensible" system, and this means a system with intelligence* [3]. Correspondingly, an AmI environment is characterised by its ubiquity, transparency and finally, its intelligence; i.e. an intelligent environment.

From these requirements, the first step in this PhD thesis was to find all the features which should be considered in an AmI system. With this aim, we considered, in order

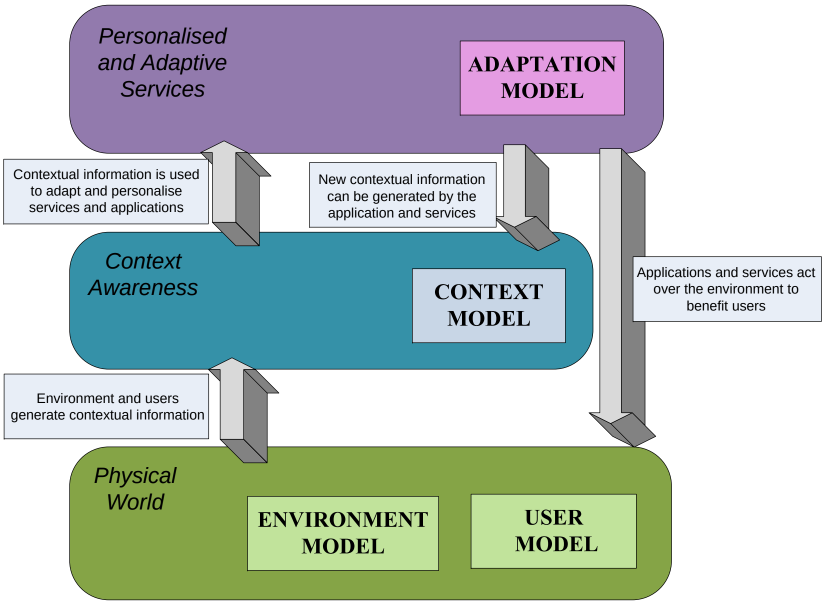


Figure 4.1: Required elements in Ambient Intelligence

to design a complete AmI system, the following elements (see Figure 4.1):

- **Environment**. In an AmI system the environment is *the physical world, every object which surrounds the user and the communications between them.* For example, if the AmI system is built in an office building, the building, offices, desks, even the temperature in the building could be part of the AmI environment. Similarly, devices, sensors, actuators and so on are part of the AmI environment. Sensors and actuators are especially important in AmI because they are in charge of the interaction between the system and the physical world. Sensors allow the AmI system to analyse and measure the information from the physical environment, while actuators allow the AmI system to influence the physical world.

- **User**. Users are the people who use the system and/or those that receive some benefit from it. As described above, *user is the main entity in an AmI system.* That is the most important characteristic of an AmI system. Every object that exists in an AmI system exists because is useful for the user in some way. Users generate information to the AmI system transparently with their behaviours and interactions with the physical world. Users' properties like age, role in an organization, usual daily routines or the state of mind are considered for the AmI system if they are useful for providing benefits to the user. In such a manner, every service and application is user-centered, directly or indirectly.

- **Context**. Several definitions of "context" have been given in literature [2, 7]. Nevertheless, all of them point that context-awareness is key for AmI systems. We define context as *every relevant information from the entities of the environment which allow the system to adapt to the demands and changes of the user and the environment*. Context-Aware systems should gather and facilitate contextual information. Note that the focus of this thesis is not how the context is modelled because a lot of effort and proposals have been done in that field (in [19] and [35] can be found detailed surveys of approaches about modelling of context). In this thesis, we are interested in how this contextual information is gathered and facilitated, i.e. in the context management. In that sense, some context-aware systems support a middleware for context management. A middleware for context-aware is an abstraction layer between an operating system and the applications running in the environment. The middleware informs services and applications about contextual information and, besides, gathers new contextual information.

- **Adaptation**. The theory of adaptation has grown out from a variety of disciplines, including biology, psychology and artificial intelligence. Nonetheless, the question about what adaptation means and what constitutes an adaptive system have been widely debated throughout history. In living systems, adaptation is linked to the learning and evolution of the species according to the Darwin's theory [8]. In that sense, adaptation is any alteration in the structure or function of an organism or any of its parts that insure species survival [40]. From the point of view of the social psychology, adaptation is the process in which the group or the individual modifies his/her behavioural patterns to fit the rules of his/her social environment based on his/her already existing social abilities [21].

  The notion of adaptation in artificial systems emerged in the seventies. In this context, adaptation is usually understood as *a mean which generates better-performing solutions through interaction with the environment* [40]. Then, an adaptive system is a system that is able to adapt its behaviour according to changes in the environment or in its users.

  In AmI, adaptation is based on the those adaptive control mechanisms which are able to adapt, in a lifelong mode, the learnt behaviours to satisfy the different users and system objectives [16]. Adaptation in AmI implies that services and applications benefit users and adapt to their needs without prying into their lives. With this aim, services and applications use the contextual information to infer and learn about users and environment. But furthermore, they are able to produce new contextual information from these inference and learning processes. For example, a location service can use the GPS coordinates available in the context to infer the street in which the user is located. The inferred street is new information which should be available in the context for other services or applications which need this kind of information. In such a way, services and applications are always user-centered because in a direct or indirect way they finally benefit users.

The four aforementioned elements cover the components of a complete AmI system.

Despite that, to the best of our knowledge no architecture for AmI systems in the literature was able to consider and cover all the components of this kind of systems (see the related work of the article presented in section 5.1 [60]). Moreover, the high variety of AmI scenarios, heterogeneous devices and sources of information, changing contexts, the requirements of transparency, intelligent user interfaces and the rest of constraints caused by AmI become the development of this kind of system in a hard process.

One of the most usual solution for the problem of the difficult design and development in AmI systems consists in the use of Living Labs, i.e. controlled imitations of real environments which can be used to test, verify and validate systems [13]. The main disadvantage of the use of Living Labs is their inability to be used in some situations, like environments with lots of users, with privacy concerns or emergency systems. The other common solution is the use of simulators. But, as described above, no solution or simulator in the literature covers all the requirements and necessities of an AmI system.

For this reason, we found that **it was necessary a general framework which was able to test, verify and validate AmI systems, and it was able to cover all the features presented in AmI environments**. According to this need, we present an architecture for AmI systems, AmISim, that builds a model for each one of the four described AmI elements: *environment model, user model, context model* and *adaptation model*. The main strengths of this architecture can be summarised as follows:

1. AmISim has a multi-layered structure which covers completely the four models to accomplish with all the components and requirements of AmI systems and abilities to change easily a model without interfering with the rest of the system.

2. AmISim enables developers to design and integrate real and simulated elements. The integration of real and simulated elements gives more realistic models due to the injection of real features in the simulation.

3. AmISim is based on a modular decoupled architecture. This kind of architecture presents several advantages in the AmI context: modularity, portability, flexibility, easiness of deployment, heterogeneity among devices and online restarting of services.

In the proposed architecture, models are encapsulated in different layers (see Figure 4.1). The lowest level of the architecture contains the environment and user models. In the middle level, the context model is found. In this level, a context manager is used to model and handle the contextual information. Finally, the upper level contains services and applications which are able to use contextual information to adapt to the environment and the users, i.e. the adaptation model. The environment and user models generate contextual information which is gathered by the context model. At that point, the information is available to be used by the adaptation model in order to adapt and personalise applications and services that act over the environment for satisfying the needs and wishes of the users. Furthermore, applications and services can generate new contextual information which will be managed by the context model.

Hence, AmISim becomes a general AmI framework in which all the AmI requirements and constraints are considered while an easy deployment is still possible. The four models are designed according to the elements that cover. Since the focus of this thesis is adaptation in AmI, the main contributions are given in the field of the modelled of user and the personalised and adaptive services. These fields are covered in the user model and the adaptation model. Nevertheless, the rest of the models are needed in order to get a complete definition and correctness of a general AmI system. Following subsections describe each model and the proposed solutions in each one.

4.1 Environment model

We introduced that the AmI environment involves the physical world. That means that the environment model describes the real world which includes buildings, complex objects, sensors, actuators, communications between devices, etc. Because of the heterogeneity of AmI environments, this model needs to be flexible, configurable and extensible.

The physical environment and the users constitute the real world. The use of real environments and real users is sometimes unfeasible (emergency scenarios, large and tedious testing, too big scenarios like a city, etc.). In such cases, the use of simulators lets a way to validate AmI services or applications, i.e. to test and validate the system under test (SUT). This PhD thesis rests on the hypothesis that **it is possible to test, verify and validate the SUT by simulating users and the physical environment**. In such a way, we use the presented architecture AmISim to integrate the real SUT in a previous simulation process in order to help in its development and minimize risks of errors in a final deployment phase.

Such hypothesis has been successfully validated over the real world in this thesis. Note that the simulation is not an overall solution. When the final deployment has been done, more tests would be needed over the real environment. Despite that, it is possible to achieve more realistic simulations using injection of real elements on the simulation, combining real and simulated elements. Reality injection means to inject or connect real elements in the simulation in order to obtain more reliable simulations. These real elements can refers to some parts of the environment, like all sensors, only a subset of them, actuators, etc. or to the users, i.e. users interacting with a simulated environment. The realism increases the reliability of the AmI services that are deployed in simulated environments.

In this thesis, the simulation of the environment and users models is based on Multi-Agent Based Simulation (MABS). MABS can build powerful models to represent real world environments and complex users' behaviours that have a degree of dynamics [24].

With this aim, AmISim uses UbikSim, a tool of the University of Murcia [34]. UbikSim is a MABS for multi-agent models of social systems. We utilise the UbikSim simulator for practical experiments as a previous phase before the final deployment over the real world. Even though the design of tools for simulation is not the goal of this thesis, we need a tool to test the experiments and UbikSim provides us with an easy and low cost solution to test and validate AmI systems. UbikSim includes tools to simulate the physical environment (building, sensors, actuators, furniture, etc.) and humans able to interact with the environment. It allows AmISim to design flexible AmI environments like complex systems. Finally, note that the context and the adaptation models are not simulated because they are real. Therefore, they are not designed using UbikSim.

A more detailed description about the environment model can be found in the article presented in section 5.1. This article presents the four models of the AmISim architecture and, specifically, it describes the environment model, their components, requirements and the interactions with the rest of the models. Additionally, some examples related to the construction of this model in different scenarios are shown. More examples of environments models for AmI systems, built in this thesis, can be found in the articles included in sections 5.2, 5.3 and 5.4.

4.2 User model

This model is in charge of describing the features related to the users of the system, their interactions and their individual or group behaviours. Because of the importance of the user in AmI systems, this model is one of the most important. Nevertheless, we found that in the literature is usual to forget or simplify the user model. But AmI is focused on users and, therefore, this model has to be an essential key in a system of this kind. With this premise, the user modelling was a fundamental goal in this thesis (see goal 4) in Table 2.1).

There is a need of designing and modelling users' behaviours in order to 1) understand the users and their actions and interactions, 2) being able to simulate them to test services and applications and 3) having a computational model of the users and their behaviours in the system. This last point is especially important in this thesis focused on adaptation. With the aim to get services able to adapt to the users, they need to know how the users are acting and how they will act to decide in consequence. Under these circumstances, a model of the user in the system is needed for AmI services even in the case where users are not simulated. Therefore, a big part of the effort of this thesis has been made in that point.

As described above, in this model, we proposed MABS, which is a branch of the social simulation area, to simulate the users of the system. The use of MABS technology in AmISim results in a user model well defined, flexible and more reliable than previous approaches in the literature. Using MABS, we introduced a probabilistical modelling for the users' behaviours. We use probabilistic models because a perfect prediction of the behaviours and the reactions of users to different stimulus is not possible. Hence, a probabilistic approach gives more realism to the model.

A hierarchical automaton is used as model. The states define the situations that the user plays in each moment and the transitions between behaviours are probabilistic.

The hierarchy defines a structure of behaviours where the highest level is a number of complex behaviours and each state of the highest level has a subordinate automaton with simpler behaviours. With this kind of structure, each behaviour can be treated separately and the modeller is abstracted of unnecessary details. Moreover, the complexity of the hierarchy of the automaton depends on the complexity of the specific behaviour to be modelled and/or the information available of the behaviour in the AmI system (e.g. if there is no information about users' location in the system, users' behaviour of their movements will not be modelled). Likewise, if the behaviour to be modelled is a simple behaviour or the available information about the user is simple, then the automaton will be a simple automaton without subordinated automata.

In this model, when the user is in a specific state, i.e. playing a behaviour, a change of state can happen, i.e. a new behaviour is activated. That means a transition between states. The transitions are modelled probabilistically according to a probability distribution function (pdf). A pdf is a function that describes the probability of a random variable to take a certain value. That means that can be used to model the frequencies and probabilities of that an event occurs over the time. In this case, a transition between two users' behaviours.

This thesis proposes the use of different pdfs according to the kind of behaviour. Three kinds of behaviours were found. The first one is *monotonous behaviours*. These behaviours are manifested by the user approximately at specific hours or time intervals, for example, having lunch or sleeping. This kind of behaviours is similar to a waiting time model because the probability of their occurrence increases with the time.

The second type is *non monotonous behaviours*. This type of behaviour is manifested by the user repeatedly, but it is not bounded to a specific time or interval, for example, going to the toilet. Non monotonous behaviours are also like waiting time models, but they are defined in specific periods (like for example, awake periods).

The last type of behaviours are the *any time behaviours*. These behaviours represent the basic behaviour, the usual one, for example, working in an office environment or resting in a home environment. This type of behaviours is continuously interrupted by the others which causes that the behaviour is played at small times.

The proposed user model has been tested in order to show the reliable of the realism achieved. Our works contrasted with empirical data from real environment and real users by statistical tests that it is possible to use them to model the users' behaviours. Nevertheless, users are not completely predictable and more tests are need over the real environment. In spite of that, user simulation is a tool able to minimize the error in the final deployment and in some situations, as described above, is the only way to test the system.

Moreover, one of the most important features of AmISim is the ability to merge simulated and real components. This ability is possible thanks to the flexibility of the architecture which enables the coexistence of real and simulated elements. In the user and environment models this capability is essential. In AmISim is possible to achieve more realistic simulations using injection of real elements within the simulated world, for example, merging some real information from the users but keeping the simulated user model or keeping the simulated environment but with real users.

As a result, despite a completely predictable user model is not possible, it is still needed. The more realistic the user model is, the higher the probability of success will be in the deployment over the real world.

The described proposals and a whole description of the user model is presented in the article included in section 5.2. Although a presentation of this model and its interaction with the other models is shown in the article of the section 5.1, the work of the section 5.2 includes a complete definition, design and validation of the proposals for this model.

4.3 Context model

Contextual information is key for AmI systems in order to discover and take advantage of meaningful information from the users and the environment. This information must be represented in a specific model to be used. Therefore, we found that the context model should be able to store, gather, merge, interpret and reason over all the relevant information in the system.

In context-awareness the representation of the context is a fundamental factor. This representation will determine other aspects such as the reasoning techniques to apply. The need for reasoning in context aware systems derives from the basic characteristics of context data: unknown, ambiguous, imprecise, and erroneous [17]. In fact, context-sensitive processing is needed by AmI services to behave proactively.

Several platforms and middlewares for context representation with diverse approaches were studied in this thesis [4, 14, 22]. They are based on different principles, like object-oriented models, ontology based models or key-valued models. We proposed the use of semantic technologies to represent the context model. The usefulness of semantic technologies in context-aware is widely known [15, 37]. They ensure a common framework for adding contextual information easily, reusing it, interpreting it, and reasoning over contextual information [28]. The result is that a more entire flexible context model is attained.

Furthermore, we found that the model of context needs some components of memory and communication. Firstly, the memory component is needed to know the history of context, i.e. the past information. Historical context reports on the necessary knowledge for learning from the experience, reasoning about past and current facts, predicting future behaviours, detecting abnormal situations, etc.

Secondly, context-aware systems should be able acquire and use the context information from different heterogeneous sources and in some cases, devices with limited resources which are not able to manage or process large sizes of data. In order to solve this problem, we propose the use of a management middleware which is in charge of managing and storing the context. This middleware should be able to make available the context to services and applications and to gather new context from the environment. With this aim, we proposed the use of the Open Context Platform (OCP) [28] in this model. However, since the flexibility of the architecture of AmISim, other context management middlewares can be used.

OCP is a middleware for management of contextual information. It has been developed at the University of Murcia. This middleware represents the information in an ontological model and offers automatic reasoning using ontologies, rules (defined by the user with SWRL [20]) as well as higher order processing using the so-called *Context Information Processing Blocks* (CIPBs) [44]. Besides, OCP was modified in this thesis to add the necessities caused by the components of memory and communication. The new version of OCP [44] facilitates an API with several temporal reasoning functions as well as a query language (Sparql [29]) for semantic databases in order to cover the requirements of the component of memory. Furthermore, its performance was substantially improved for working online, feature needed for adaptation in AmI. The information and details about this new version of OCP can be found in the co-authored articles by the PhD candidate [44], [45] and [46].

The context model is presented in the general architecture AmISim, which is presented in the article included in section 5.1. In this work, the context model is defined, setting its requirements and constraints and the necessary components. The interaction with the rest of the models is also explained, concretely, how the exchange of contextual information is produced between the different layers of the architecture. Finally, how OCP is integrated in the architecture is shown with examples to get a better understanding.

4.4 Adaptation model

AmI services and applications should be intelligent, adaptive and responsive to users. Therefore, this model needs to be able of supporting applications and services that, using the contextual information, give an adaptive and personalised help or benefit to the users. This model is the most flexible model in AmISim because is highly dependant on the kind of scenario and their requirements.

In order to create really adaptive AmI services and applications, we found that there is a need for the system to be aware of the user presence and context which entails the need to localise the whereabouts of the user. With information about location, systems can locate nearest features of interest of the users and adapt to their requirements. As a consequence, location information is one of the most essential and important requirements for AmI systems.

There are various localisation means available for outdoor spaces such as those which rely on triangulation and decoding timing signals from the satellites. Nonetheless, these outdoor localisation means cannot be used in indoor environments where their effectiveness is hindered by the obstruction of walls or objects or is limited by the reflections. What is more, indoor localisation need a higher accuracy than outdoor localisation due to the relatively narrower spaces available in indoor environments where few meters could make a difference and mean a different room or even a different floor of the building.

There are a lot of projects focused on the development of adaptive control of indoor environments based on location [1, 27, 32]. These projects are focused on home environments with usually one or a few inhabitants. Because of that, the models do not fit well in a multi-inhabitant environment. Likewise, these approaches usually need prior knowledge of the environment and users, which is not feasible in the most of environments (especially private buildings).

Another typical problem in the existing works is that they are not able to work in the long term. In AmI environments, where conditions and users are continuously changing is needed that services and applications work in the long term. That means that they should support a continuous adaptation to changes and learn and adapt to novel situations that can happen in the future.

A final constraint should be considered from the AmI requirements. In order to keep the transparency, invisibility, and non-intrusion that an AmI environment should keep, the use of a big and visible infrastructure is not desirable. Hence, there is a need to rely on non-intrusive and cheap sensors trying to keep a minimal infrastructure and to avoid annoying users.

As described, most indoor localisation proposals in the literature have poor ability to adapt to changes, involve high cost and usually require the installation and deployment of a big new hardware infrastructure, which cause them to be less ubiquitous.

Thus, every localisation proposal in AmI should respect the invisibility of the infrastructures and the non-intrusion. Regarding these constraints, two different types of environments can be considered according to the available information and the possibility of controlling them: 1) environments where a minimal infrastructure and a little prior knowledge can be controlled preserving the AmI requirements, and 2) environments completely unknown, i.e. no information about users, infrastructures or similar is available. We call them *controlled and uncontrolled environments*, respectively.

Controlled environments are those in where we have some information about the infrastructure of the building or map and the users. Usually, in this kind of environments we are able to deploy new hardware and devices (preserving the AmI constraints). Examples of these scenarios are outdoor environments with GPS [11] or the Galileo Positioning system [5], where the knowledge about the satellites is available. For indoor environments, good examples of controlled environments are hospitals or office buildings. In these scenarios, users (or their roles) are known (e.g. doctors, nurses or workers) and also the structure of the building. This reports on some information about their daily routines. In this kind of environments is usual that the effort and budget invested in new location infrastructure compensates the acquired benefits. Imagine for example a hospital, where some services, like the localisation of a doctor in an emergency situation, requires more precision than for example a marketing service based on localisation in a shopping center.

On the contrary, uncontrolled environments are those in which we can not suppose information about the map or buildings, users or infrastructure. Good examples of these environments are scenarios where the quantity and identity of the users is unknown, like a highway where vehicles are pretended to be tracking or a big shopping center where users are continuously changing and no information about them is available. Systems in these environments need to be more flexible than in controlled environment since no prior knowledge can be used.

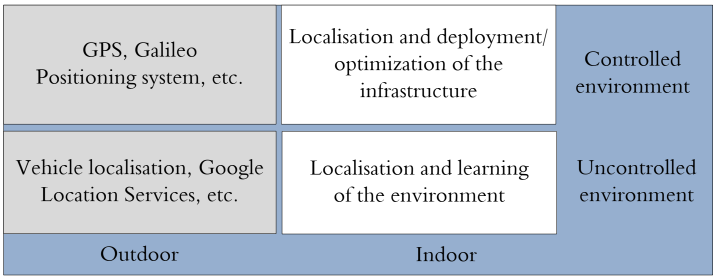


Figure 4.2: Localisation in AmI

Both kinds of environments keep the features of AmI. From the AmI paradigm point of view, uncontrolled environments have stricter restrictions because no control can be assumed. Figure 4.2 summarizes the kinds of localisation in AmI environments according to the type of environment (controlled or uncontrolled) and its outdoor/indoor nature.

In this thesis the problem of localisation in AmI indoor environments is faced, as for controlled as for uncontrolled environments. As described, we are focused on indoor environments, which is the doctoral line of this thesis. The white squares of Figure 4.2 represent the problems faced in this thesis. The following subsections present the works performed to deal with the presented problem.

4.4.1 Localisation in controlled environments

A Location Based Service (LBS) delivers personalised services based on users' locations. In this context, the system can locate nearest features of interest of the users and adapt to their requirements using the LBS and, optionally, other contextual information about them.

Localisation in AmI environments implies dealing with two groups of requirements which are often contradictory. On the one hand, AmI requires non-intrusive and transparent environments. Consequently, a minimal infrastructure and interaction with the user should be maintained. On the other hand, in order to accomplish with high quality of location, the system must obey the following requirements: high accuracy on location information, complete coverage of the zone, availability of the service and moderate cost [6].

Note that a LBS requires special hardware to detect user location. Usually, the hardware is bought specifically for the location service, only in a few cases is already available before the development of the service, for example, if Wi-Fi is used as the wireless infrastructure (like in the second approach presented in this thesis). Whether the hardware is available or not, the people in charge of deploying the service have to decide the number of devices to be installed and their allocations. Here, getting the highest coverage and accuracy implies, most of the times, a considerable deployment of antennas, which is expensive. In addition, this kind of infrastructures usually does not go unnoticed, being more intrusive. High cost and intrusion are undesired properties in AmI.

For these reasons, a trade-off between coverage/accuracy and intrusion/cost is needed. Providing a configuration (number and position of the location antennas) which guarantees the maximum coverage and accuracy is a difficult task. The best configuration depends on the physical space in which the service is deployed. For example, in a small room, it is possible for a device to cover all the room, but in the case of a bigger room, like a showroom, only one device could be insufficient.

With such premises, in this thesis, a probabilistic framework based on Hidden Markov Models (HMM) [30] was used to predict user location, at the same time that an appropriated deployment of the whole infrastructure was designed. The use of location prediction makes possible that the building is not fully covered because the system can predict the uncovered areas (i.e. without antenna and out of reception coverage).

Our process for engineering the LBS uses UbikSim in order to simulate the environment and user models. An intelligent office building and their workers are artificially reproduced in simulations. The LBS is a real implementation integrated in the simulation. Because of this process is previous to the deployment over the real world, only a little location information about the environment and users is available.

Using the simulated environment and user models, expensive and annoying tests with users can be avoided at a first phase. Then, when the LBS has been completely tested and an appropriated configuration of the infrastructure has been achieved, it can be deployed over a real scenario. This process attains an easy and low-cost solution to configure and deploy LBS in AmI.

Some experiments were made over a simulated office building with different configurations of the number and position of the antennas. The experiments showed that the coverage is not the most important issue in a location service. In fact, a good prediction accuracy of the used technique resulted more important than the coverage. This is due to the fact that a system with high accuracy can predict the uncovered areas and remove noisy. In this sense, the system based on HMMs got better precision rates and a better adaptation to changes in users' behaviours than other studied techniques. In that context was also discovered that it is possible to remove antennas representing usual starting points for data traces from users. Another important discovery from the experiments was that antennas in corridors are more important for prediction than in other places. Under these considerations, a reduction of the 23% in number of antennas was made without a significant reduction in prediction quality.

In order to find such a good configuration of the number and the position of the antennas, which reduces the number of antennas and keeps the prediction quality, these experiments were based on trial and error. This particular search was handcrafted. As a consequence of that, we established that there was a need of finding the optimal configuration of antennas in the environment automatically. Additionally, we found that the learnt knowledge about the structure of the building and the patterns of behaviours could be useful in this search.

Accordingly, two features should be considered. Firstly, finding the best configuration of antennas implies obtaining the minimum number of antennas while high quality of the accuracy of the prediction is maintained. This problem can be stated as a multiobjective optimization problem (MOP) with two subobjectives: maximizing the accuracy of the prediction by the HMMs and minimizing the number of allocated antennas. Secondly, from our previous experiments is clear that an optimization algorithm could get profit from the adquired knowledge. The consideration of domain knowledge in optimization processes has shown to improve the results.

In MOPS no solution causes an optimum in all objectives. Algorithms for MOPs try to find a solution in which the objectives are satisfied in an acceptable factor. This solution can not be improved on any objective without degrading the other. Our first attempt to solve this problem was using Genetic Algorithms (GAs) [18]. GAs are usually a suitable technique in the resolution of MOP because they maintain a population of solutions where finding solutions in parallel is possible [12]. This mechanism improved the results accomplished by the handcrafted process.

Furthermore, since some knowledge from the domain was available, we designed a improved solution which uses this advantage using Memetic Algorithms (MAs) [23,26]. Memetic algorithms are an optimization paradigm based on the concept of "meme". This concept was coined by Dawkins in 1976 and is analogous to the gen in genetic, but the meme in cultural evolution. A meme can be defined as a unit of cultural transmission or imitation [9]. Memetic algorithms describe an optimization paradigm which combines domain knowledge with different metaheuristics. Several works have shown that MAs are able to get better solutions than its precursor, the GAs, in NP-hard optimization problems, like the presented one (see the related work of the article presented in section 5.4 [53]).

Our experimentation showed that this proposal helps to find an optimal configuration of the location infrastructure more effectively and efficiently than our previous work, by the handcrafted process and a classic genetic algorithm, while the quality of the location prediction is kept. This contribution gives an optimal infrastructure which reduces cost and achieves a less intrusive environment, avoiding annoying or costly tests over real environments and with only a bit of information available.

The proposals, techniques and results described in this subsection can be found in the articles included in sections 5.3 and 5.4. The first article, in section 5.3, shows that it is possible to use AmISim to design and test a LBS which predicts the users' locations. In the same work, it is shown that AmISim can help in the deployment of a LBS reducing costs and infrastructure, but keeping a good accuracy of the prediction. As a continuation of this work, the article presented in section 5.4 proposes the use of a memetic approach to optimize the configuration of the LBS. This approach uses some available information to optimize the configuration and improve the accuracy of the prediction.

4.4.2 Localisation in uncontrolled environments

The previous subsection describes our proposal in AmI environments where a minimum control of the environment and the infrastructure is available. Such approach leads to a complete solution of the problem of localisation in controlled AmI environments.

The solution covers the deployment of the infrastructure and the mechanisms for the prediction of location of the users.

Nevertheless, this part of the thesis is focused on those environments where no information is available. We assume that no information about users, structure of the buildings or the antennas of the infrastructure is available. In such a way, our solution for this problem does not require the installation of new hardware, preserving, in this manner, the requirements of AmI environments, the privacy of the users and at zero-cost easy deployment of the infrastructure.

From these assumptions, an adaptive fuzzy logic based system for indoor localisation in AmI environments was presented. The system uses the Received Signal Strength Indicator (RSSI). RSSI measures the power present in a received radio signal. In the article included as part of this thesis, WiFi is used as signal because almost every modern building has installed or accessible nearby WiFi Access Points (APs). Accordingly, almost every building can use this solution without the need of deploying new hardware and at zero-cost. Nonetheless, any radio signal, like Bluetooth or RFID (or even a combination of them), can be used in our approach because the solution only requires signals with RSSI.

In an environment as the described one, where no prior information is available, the conditions can change continuously and the signals often present noise and uncertainty, a mechanism capable of learning online and adapt to the highly dynamic environment is a must. To our knowledge no previous work in the literature is able to deal with these conditions. Our system is the first able to operate online and adapt its behaviour in an incremental and lifelong way in order to attain high success rates of users' location.

The system is based on fuzzy logic in order to face the short and long term uncertainties and noises from the environment in an understandable way by users. Additionally, some improvements have been included to allow the system to work if the conditions change or even if a new situation is observed. We have also designed a mechanism which is able to learn and adapt itself unobtrusively and incrementally, in order to extend the basic mechanism with a lifelong localisation solution.

Thanks to the flexibility of AmISim, the system was tested in a first phase over two simulated scenarios. The system was probed over three real environments. Note that in the work presented as part of this thesis only one scenario is shown. The other two scenarios are currently under publication process [56]. The experiments showed that presented solution obtains high accuracy, low cost and easy deployment in very different AmI scenarios with only a few days of training and no prior knowledge or new hardware. In addition, our solution outperforms other studied techniques, being the only one which is able to operate and adapt online in a long term mode.

This proposal, their experiments and results are described in detail in the article presented in section 5.5. This work presents the problem, the approaches in the literature to solve it and a description of the reasons which caused the necessity of improving the existing works. Then, the techniques applied to solve the problem are described deeply and the experiments that show the suitability of the proposal are presented.

As described in this introduction, the contributions of this thesis are not only focused on the methodologies or algorithms for AmI, but we have tried to provide developers with a whole solution for the deployment of this kind of systems preserving their features and requirements. These contributions can be easily integrated over real current environments as part of our lives for the realization of truly ubiquitous and intelligent environments. In fact, we think that our results are not only interesting in an academic or scientific domain, they are suitable and useful in our daily lives.

Chapter 5

Publications composing the PhD Thesis

5.1 Combining the real world with simulations for a robust testing of Ambient Intelligence services

Title	Combining the real world with simulations for a robust testing of Ambient Intelligence services
Authors	Garcia-Valverde, T. and Serrano, E. and Botia, J.A.
Type	Journal
Journal	Artificial Intelligence Review
Impact factor (2010)	0.429
Publisher	Springer Netherlands
Pages	1-24
Year	2010
Month	April
ISSN	0269-2821
DOI	`http://dx.doi.org/10.1007/s10462-012-9340-4`
URL	`http://www.springerlink.com/content/qp645u6v0x731l12/`
State	Published

Journal details: Artificial Intelligence Review
Editor-in-Chief: David Robertson
ISSN: 0269-2821 (print version)
ISSN: 1573-7462 (electronic version)
Publisher: Springer Netherlands
Impact factor (2010): 0.429
Website: `http://www.springer.com/computer/ai/journal/10462`

Authors – Personal details	
Name	**María Teresa Garcia Valverde**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Emilio Serrano Fernández**
Position	Member of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Juan Antonio Botía Blaya**
Position	Lecturer of the Department of Information and Communications Engineering
University/ Organization	University of Murcia

Contribution of the PhD student
The PhD student, María Teresa García Valverde, declares to be the main author and the major contributor of the paper *Combining the real world with simulations for a robust testing of Ambient Intelligence services*

Artificial Intelligence Review manuscript No.
(will be inserted by the editor)

Combining the real world with simulations for a robust testing of Ambient Intelligence services

Teresa Garcia-Valverde · Emilio Serrano · Juan A. Botia

Received: date / Accepted: date

Abstract This paper proposes a general architecture for testing, validating and verifying *Ambient Intelligence (AmI)* environments: *AmISim*. The development of AmI is a very complex task because this technology must often adapt to contextual information as well as unpredictable behaviours and environmental features. The architecture presented deals with AmI applications in order to cover the different components of these kinds of systems: environment, users, context and adaptation. This architecture is the first one that is able to cover all these features, which are needed in a full AmI system. The paper shows that AmISim is able to cover a complete AmI system and to provide a framework which can test scenarios that would be impossible to test in real environments or even with previous simulation approaches. Simulated and real elements coexist in AmISim for a robust testing, validation and verification of the AmI systems, which provide an easier and less costly deployment.

Keywords Ambient Intelligence · Multi-agent based simulations · Testing · Ubiquitous Computing

1 Introduction

Mark Weiser envisioned in his work The Computer for the 21st Century (Weiser, 1995) that "the most profound technologies are those that disappear". He based on these arguments to introduce his vision of Ubiquitous Computing. The last years have seen substantial progress in Computing systems. Some of their previous constraints (like memory capacity, processor speed, communication bandwidth or even the cost) have been noticeably improved (Garlan et al, 2002). This has caused that computer systems are now widely spread. Furthermore, there have been great advances in some technologies such as wearable computers and devices, wireless networks, sensors devices and control appliances. The improvements regarding to their connectivity, battery life, weight and size make it feasible to integrate computer systems in our daily lives, making ubiquitous environments a reality.

University of Murcia
E-mail: {mtgarcia,emilioserra,juanbot}@um.es

The Ubiquitous Computing provides a smart environment, where a set of software and hardware elements supports an invisible and nonintrusive interaction between environment and users. This interaction makes it possible to acquire context knowledge from the environment, i.e. Context-Aware, and to provide services that support daily user tasks in a smooth and adaptive way. The convergence of these points is called Ambient Intelligence (AmI). AmI refers to a seamless and invisible computing environment, which is able to provide users with proactive and adaptative services.

AmI aims to create digital environments that are aware of our presence and context through the usage of embedded and nonintrusive devices. These devices generate information about the environment, users and the changes in both of them. This information allows services and applications in AmI systems to adapt to changes. Therefore, key features of AmI systems are *embedded*, *context-aware*, *personalized*, *adaptive* and *anticipatory* (Aarts, 2004).

In AmI, access information, communication and services must be ubiquitous, wireless and transparent to users. Furthermore, the interaction between the system and the user must be unobtrusive and natural. Hence, embedded devices to access to the information in a natural and ubiquitous way are needed in an AmI environment. These devices are able to recognize users, the environment and the contextual information. This meaningful information about users and their environments can be used by the applications to adapt to them. In a nutshell, the system is context-aware. Contextual information allows applications to personalize, adapt and anticipate to the environment and to the users' necessities, desires and behaviours.

Considering all the described requirements (embedded, context-aware, personalized, adaptive and anticipatory), it is found that the following models must be considered in AmI:

- *Environment model*. A model that describes the physical world, i.e., buildings, complex objects, sensors, actuators, communications between devices, etc. This model must be flexible and configurable.
- *User model*. A complex multi-agent model that simulates physical and human features, interactions between them, and individual and group behaviours.
- *Context model*. This model gathers, merges, interprets, reasons and stores the contextual information, i.e., all relevant information that surrounds the environment and the users.
- *Adaptation model*. A model capable of supporting applications and services that use contextual information. This contextual information is used to adapt, personalize and anticipate to the environment and users in a reactive or proactive way.

There are a large number of developments and projects that apply the issues of AmI over different contexts. Specifically, most of them are focused on indoor environments. Popular examples of these projects are: Project Aura (Garlan et al, 2002), Oxygen Project (Rudolph, 2001), House_n (Intille et al, 2000) or Aware Home (Kidd et al, 1999a). These proposals include controlled imitation of real environments (called Living Labs (Gabel et al, 2005)) to test, validate and verify AmI systems.

Testing an AmI system includes detecting, locating and repairing errors on the system. Verification includes checking that the initial requirements have been achieved correctly. Validation checks that the functionality is achieved as expected. Testing, verifying and validating an AmI system by real environments is rather difficult. Moreover, the application of these labs is impractical in several situations, like an environment with thousands of individuals or an emergency scenario.

In order to address these problems, this paper proposes a general architecture for testing, validating and verifying AmI environments: AmISim. AmISim is a layered architecture which considers and covers the four aforementioned models of an AmI system. To the best of our knowledge, AmISim is the first architecture that is able to do this. Furthermore, AmISim enables developers to integrate real and simulated elements from these four models. AmISim provides a user model (essential in AmI system) without the need of interaction with the user. Therefore the number of individuals is not limited. This facilitates to test, validate and verify applications for hundreds of users or more. Finally, the contextual information in AmISim is formally represented using semantic technologies in an easy way.

The paper setup is presented as follows. Section 2 describes other works related with simulation in AmI or specific parts of AmI. Section 3 presents AmISim and the proposed architecture, which covers the requirements for the four models explained above. Section 4 illustrates the use of the AmISim architecture for an advanced AmI service into a scenario of an office building. Finally, Section 5 presents the conclusions and future work.

2 Related works

Most of the approaches for the testing, validation and verification of AmI systems are only focused on one or some parts of an AmI system. Nevertheless, to our knowledge no work covers all the requirements and necessities of an AmI system, i.e., the four models explained in the introduction. Therefore, the related works presented here about testing, validation and verification of AmI systems have been separated into four different topics according to their topics. (1) The first topic is Living Labs, which allow us to perform feasibility and usability studies by real life environments. The following three categories involve simulators, which deal with a specific part of an AmI system or emphasize such part. Thus, (2) the second topic is simulators that are focused on the use and analysis of the environment model of AmI systems. (3) The third category describes simulators that explicitly use a context-aware model. (4) Finally, the fourth topic presents the use of multi-agent approaches to simulate AmI systems. These last works usually emphasize the user model.

2.1 Living Labs

The most intuitive approach for the testing, validation and verification of an AmI environment is the use of real users receiving feedback from them. A *Living Lab* is a laboratory that consists of real life environments where users and researchers look together for new AmI services. The first serious works to build laboratories that could be used to conduct feasibility and usability studies in AmI started in 2000 (Friedewald et al, 2005). As a result, the HomeLab of Philips was opened on

April 24, 2002. Another well-known living lab was developed within the Aware-Home Project (Kidd et al, 1999b). This living lab has been used to research on context awareness and ubiquitous sensing, individual interaction with the home, person identification and location, and finding lost objects. Another example is the Essex intelligent apartment (iSpace), a test bed for ubiquitous-computing environments developed by Hagras et al. (Hagras et al, 2004). The iSpace is fitted with several sensors (temperature, occupancy, humidity, etc.) and actuators (such as door actuators, heaters and blinds). The target is to response to user's necessities. The *European Network of Living Labs* is a grown up initiative coming from the own European Living Lab and sponsored by the European Community. The Open Living Labs web[1] shows contact information for Living Labs in about 30 countries.

Developers in a Living Lab can make a compilation of relevant data about the execution of the application interacting with users. Then, these data can be analyzed to evaluate batteries of tests executed over the AmI application. This approach of real-world testing in laboratories has proliferated since the results are very reliable because the application is tested with the final users. However, the main disadvantage of the use of Living Labs is that it is not feasible in applications with hundreds or thousands of users. For example, the study of AmI services in office buildings, such as the location service used as case study in this paper, would be impractical with this approach. Furthermore, a huge economic investment would be needed. In such cases alternative approaches like simulations are useful even when the use of Living Labs is feasible. This is because simulations enable us to test services before their deployment, therefore, the cost of the faults is lower than if they appear in the development.

2.2 Simulators of environments

Building a Living Lab can be slow and costly. Furthermore, it can not be used with large-scale problems. Software developers should be able to develop AmI applications even if they do not have all hardware available yet. With the aim of solving this problem, AmI simulators have appeared in the last years. There are very few simulators designed specifically for this domain (Reynolds et al, 2006). They allow researchers to investigate AmI usability aspects and also functionality issues without physical limitations imposed from using real settings. This section deals with simulators modelling mainly the AmI environment.

Several 3D first-person-shooter (FPS) games, which have appeared for PCs since the late 1990's, have released software development kits (SDKs) allowing programmers to modify these games. Examples of these games are Half-Life, Quake III, and Unreal Tournament (O'Neill et al, 2005). Several proposals for simulating AmI system components try to exploit the 3D graphics engine of these games to model a realistic environment. For example, Quake III Arena is a FPS that models the physical environment in a 3D view. TATUS (O'Neill et al, 2005) is also an example of simulator with this philosophy. It also uses a 3D FPS network game, Half-Life (Coporation, 1998), that allows up to 32 players competing in a single game. TATUS introduces an interesting feature; it supports research and develop-

[1] Open Living Labs website: `http://www.openlivinglabs.eu/`

ment of adaptive software. The features of sensors and actuators are modelled by a SDK for Half-Life. This SDK includes map objects. A map contains information about objects, their names, types and coordinates on the environment. Defining physical objects, invisible and intangible entities like sound or lights, and triggers for modelling events is also possible with this SDK. An additional message XML-based definition tool is used to allow the information to pass between the simulator and the SUT, i.e. the system under test, during the experiment. Other popular cases of AmI environment simulations employing FPS games are UbiWise (Barton and Vijayaraghavan, 2002) and QuakeSim (Bylund and Espinoza, 2002). They are explained in next section under the topic of simulators that include a context-aware model since this model is their main focus.

The main advantage of these proposals is obtained by working with simulations. This is, quick and cheap prototyping without hardware or environment limitations. The major problem using these simulators is that they do not simulate the user. In these cases, the user is a player of the game (i.e. a person). Consequently, running a suite of experiments is slow and costly because the simulator requires interaction with the user (a player of the game) in every simulation. Therefore, these games limit the number of individuals in the environment, their relations and behaviours. Moreover, these proposals do not utilize a context-aware model. AmISim, presented in this paper, employs an environment modelling tool based on SweetHome3D[2], a free indoor design application, in order to model realistic environments.

2.3 Simulators with a context-aware model

There are many definitions in the literature about *context*. Several authors had tried to characterize the term context-awareness since the first time that appeared, about 1994. Brown (Brown, 1996) defines context as those elements of the user's environment that are known by the computer. Other authors like Abowd et al. (Abowd et al, 1999), define context as any relevant information for the applications about the entities in the environment. Nevertheless, all definitions concur in stating that the use of context-awareness is key for AmI systems in order to adapt to the demands of the users and their environments.

There are several works that explicitly use a context-aware model. An example of simulator with this philosophy is UbiWise (Ubiquitous Wireless Infrastructure Simulation Environment)(Barton and Vijayaraghavan, 2002), which uses Quake III Arena. UbiWise is focused on the simulation of computing and communication devices. This simulator emerged from two existing simulators, UbiSim and WISE. UbiSim generates contextual information from raw simulated data in Quake III Arena and this information is processed in the Context Toolkit (Salber et al, 1999), which produces meaningful context to applications. The second simulator, WISE, offers a 2D view for simulating and setting devices, their connectivity, protocols, scenarios and the interaction between users and the environment.

Another example is QuakeSim (Bylund and Espinoza, 2002), this simulator is a tool that interactively manages context information in real time. A 3D world and different kinds of context information are simulated in this tool. QuakeSim

[2] Sweet Home 3D: `http://www.sweethome3d.eu/es/index.jsp`

modifies Quake III Arena, like UbiSim does, to add simulation of sensors and actuators and the Context Toolkit manages all this information.

Morla et. al (Morla and Davies, 2004) present another approach with a context-aware model. The approach deals with testing and evaluation of network related issues on location-based applications. This proposal includes an environment model and uses the NS network simulator[3]. The authors illustrate their approach employing it to a mobile remote heart and to a lung health monitoring application.

Context-aware software is an emerging kind of application. These applications operate in a highly dynamic environment, where the testing is more complicated. These systems register parts of their context-aware logic in a middleware. However, most of the conventional testing techniques, such as unit tests, do not consider this logic (Lu H. and Tse, 2006). Consequently, some approaches to test context-aware applications have been proposed in recent years. Tse et al. (Tse et al, 2004) use case generation based on metamorphic testing, a property-based testing strategy. Given multiple executions of the software under test, a metamorphic relation is an expected relation over a set of input data and their corresponding output values. The metamorphic testing is based on checking if a group of test cases satisfy these metamorphic relations. In the same group, Lu et al. (Lu H. and Tse, 2006) have proposed a novel family of testing criteria that considers some contextual events and their associated actions. Starting from a context-aware data flow, the evolution of contexts is studied in order to obtain context-aware data flow associations and testing criteria. The authors illustrate their approach with the construction of adequate test sets and the evaluation of test results for an RFID-based location-sensing system.

These approaches, as the ones presented in the previous section, do not provide a user model, which is essential for testing AmI systems. AmISim, presented in this paper, utilizes semantic technologies to represent the context model. This ensures a common framework for: adding contextual information easily, reusing it, interpreting it, and reasoning over the context in order to integrate entities with a low level of information into other more abstract entities (Nieto et al, 2006). The result is that a more entire flexible context model is provided.

2.4 Simulators based on multi-agent technologies

One reason that justifies the necessity of new solutions for the testing, validation and verification of AmI systems is that one of the most complex parts of simulating these systems is the simulation of users. Despite this, there are few simulators that really manage complex behaviours of users and their relations, the contextual information and its use to adapt to the users. The use of social simulation in general and Multi-Agent Based Simulation (MABS) in particular provide the possibility of easily integrating sociological studies about user behaviours in the simulations. This is because many studies of this nature are carried out through MABS (Qiu and Hu, 2010).

There are some works that have followed multi-agent based approaches for some components of AmI environments although they do not cover complete AmI systems. For example, Reynolds et. al (Reynolds et al, 2006) simulate sensors,

[3] NS network simulator website: `http://www.isi.edu/nsnam/ns/`

actuators, and the environment, but without simulation of context or adaptative services facilities. The work of Liu et. al (Liu et al, 2006) proposes a scalable framework for prototyping and testing mobile context-aware applications. They use a distributed intelligent multi-agent system to model complex and dynamic user behaviours. For example, the user movement may be simulated based on policies specified by the software tester such as speed or fine grained trajectories. Another ubiquitous computing simulator in which agents are used to model users is presented in the work of Martin and Nurmi (Martin and Nurmi, 2006). They simulate separately agents, environment and context. The context model is defined with variables of the context and maps. AmISim, presented in this paper, claims to be more complete than the approaches mentioned because it covers environment, context-aware, users and adaptation models.

The field of *Multi-Agent Systems* (MAS) is complementary in several aspects to MABS (Drogoul et al, 2002) and its borders are increasingly blurred. The MAS paradigm has been used widely for the development of AmI applications. This is demonstrated by numerous papers about what has become known as *Agent-based Ubiquitous Computing* (Mangina et al, 2009). The approach presented in this paper is not developing AmI using MAS, but using MABS to model some of the main parts in AmI systems, such as environment and users, where real tests are too costly or impractical.

3 The AmISim architecture

AmISim is a generic multi-layered architecture that enables to simulate complete AmI systems. On the one hand, the generic architecture enables developers to test, validate and deploy any kind of AmI systems. On the other hand, the different layers cover the four presented models of an AmI system: environment model, user model, context model and adaptation model. The layered architecture allows developers to change easily a model without interfering with the rest of the system.

The AmISim architecture is showed in Fig. 1. The multi-layered architecture is based on the Open Services Gateway initiative (OSGi)(Marples and Kriens, 2001). OSGi is a framework for Java that supports the development and deployment of modular applications, called bundles. It manages the life-cycle of bundles and provides tools for the discovery, publishing and cooperation of devices and services. Thus, OSGi presents a large amount of advantages according to the features of the AmI systems, like easiness of deployment, portability, heterogeneity among devices and other specific features from the requirements of the Ubiquitous Computing, like the possibility of restarting a service without stopping the system, which supports the non intrusion requirement of a ubiquitous system.

In the proposed architecture, every component from the different layers is encapsulated in the form of a bundle within the OSGi environment. The lowest level of the architecture contains the bundles that configure the environment and user models. Both models are based on MABS. In the middle level, the bundles of the context model are found. In this level, a context manager bundle and a domain specific ontology are used to model the contextual information. Finally, the upper level contains located bundles which are able to use contextual information to adapt to the environment and the users.

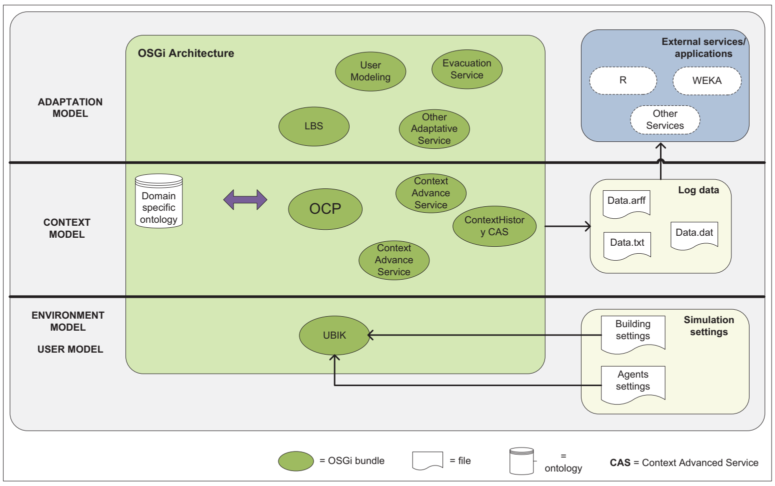


Fig. 1 AmISim architecture

OSGi enables the inter-bundle communication among all the bundles. Nevertheless, AmISim architecture divides the bundles into levels according to the model they define. This division makes a natural separation between applications or services, their contexts and the technical details about the physical devices (i.e., each model that composes an AmI system). Therefore, the adaptation model knows the context model, but it does not need to know the features of the physical devices, their locations and the rest of low level details.

The use of independent bundles provides AmISim with an important feature: mixing simulated bundles with real bundles. Thus, it is possible, for example, using a real context middleware in the context model to test real services (i.e. real SUTs) whereas a simulated environment and user model are used. Another interesting example of this feature is the possibility of simulating some sensors over a simulated environment but using one or more real sensors in order to test their performance. Merging simulated and real components is one of the most important features in AmISim, since the final deployment over the real world will be easier.

The flow of inter-bundle information is performed by levels. In the environment model, there are simulated sensors and actuators that perceive the changes of the environment and users. These sensors and actuators are defined by bundles. When a change in the environment or user is produced, these bundles send these data to the context manager bundle, i.e., they produce context. The context manager bundle, using these data, generates contextual information and stores it in the historical context information database. Hence, the contextual information is available for the bundles in the adaptation model, i.e., the consumers of context. Using this meaningful information, these bundles can offer services adapted to the environment and the users.

This is the ascending flow of information in the system. In a similar way, there is a descending flow of information. The bundles in the adaptation model can generate new contextual information (for example, a service of machine learning that infers new information from the contextual information) or send to the actuators commands to order them some actions (for example, turn on the lights). In the first case, the context manager bundle stores properly the new contextual information. In the second case, it sends the orders to the actuators abstracting to the applications from the low level.

The next sections describe deeply each model that composes the system in depth. For each model, this work presents a proposal to cover it according its requirements and components. Therefore, the environment and user models are instantiated by Ubik (Serrano et al, 2009) , the context model by the Open Context Platform (OCP) (Nieto et al, 2006) and the adaptation model by intelligent services and external tools. Nevertheless, as described above, the architecture of AmISim allows the developer to change one or more of these instantiations without changes in the others.

3.1 Environment model and user model

MABS can build powerful models to represent real world environments that have a degree of complexity and dynamic (Luck et al, 2005). In this work, the use of MABS for the environment and user models is proposed. Both models represent the real world, i.e., the scenario and the people regarding to this scenario. Therefore, they are suitable for representing with MASB. On the contrary, the other two models can not be represented using this technology because they do not simulate parts of the real world.

The environment model can be simulated in a large number of scenarios thanks to the flexibility of MABS. The simulated physical scenario can also be configurable in order to generate different case studies. Each physical space has particular properties derived from its own nature, for example, a person can walk in a room, but not in a window.

Another important factor in a simulator for AmI environments, as mentioned above, is the model of sensors and actuators since they enable sensing context and acting on it. In AmISim, sensors and actuators can be simulated. Thus, it is possible to simulate a sensor or an actuator and its behaviour, i.e., how the sensor receives the information from the environment, what information it receives, how the actuator sends the information to the environment and what information it sends.

With this aim, sensors and actuators in AmISim can have a wide range of features and properties. In this work, the approach for modelling these devices only describes their most fundamental features. A set of features is defined for setting several different behaviours of sensors and actuators. However, other features can be easily added thanks to the design philosophy of MABS.

Both sensors and actuators can be wearable or static, and they can generate data in periodical intervals or just only sporadically. All these features can be configured for the simulation setting. More examples of the features that can be configured for sensors and actuators are: their usual range of data, the probability or frequency of their occurrences, the probability that unusual data are produced (noise) or the range of data of these unusual occurrences.

Finally, other interesting properties can be defined for the simulation, such as the scope of the sensors and actuators. Out of this scope, sensors and actuators have no influence, i.e., sensors can not capture any data and actuators can not act in the environment. In AmISim, actuators are considered in a similar way that sensors, but they have some particular properties, for example, the effect of an action in the environment.

Regarding the user model, as described in section 2, simulators of AmI or similar environments usually forget or simplify the user model. Nevertheless, AmI is focused on users, their experiences and interactions. Thus, the user model is a fundamental goal in a simulator of this kind. The MABS technology, which is a branch of the social simulation, allows us to integrate sociological studies about the users' behaviours into the user model of AmISim. The result is a user model well defined, flexible and more reliable than previous approaches in simulation. Note that a simulated user model is not an overall solution because users' behaviours are not completely predictable. Thus, more tests would be conducted after the deployment. However, it offers a tool to minimize the error before the deployment. Therefore, the more realistic the user model is, the higher the probability of success will be in the deployment over the real world.

Thus, the behaviour of simulated people is modelled probabilistically in this work. Behaviours are defined as situations that the agent should play in each moment and transitions between behaviours are probabilistic. The underlying model is a hierarchical automaton (i.e. in the highest level there is a number of complex behaviours that the agent may play and, when it is in a specific state, there is subordinate automaton with simpler behaviours that defines that state). Hence, the modelling of each behaviour is treated separately and the modeller is abstracted of unnecessary details. In the lowest level (basic actions), each state is atomic. An agent never carries out two behaviours of the same level simultaneously.

In this approach, the behaviours of a user are divided into three types:

- Monotonous behaviours: the kind of behaviour that a person manifests always approximately in the same time slot, and on a daily basis (e.g. meeting or having coffee).
- Non monotonous behaviours: the kind of behaviour that a person usually manifests, not bounded to a concrete time slot, and repeated within a non constant period (e.g. going to the toilet).
- Any time behaviours: such behaviours are the usual behaviours and they are often interrupted by the others (e.g. working if the simulated person is a worker).

When the user is in a state, a change of state can happen, i.e., a new behaviour is arisen. As described above, these transitions between behaviours are modelled probabilistically. Then, probability distribution functions (pdf) are used according to the described type of behaviours. For example, the behaviour going to the toilet, is a kind of non monotonous behaviour and it can be fitted with an exponential distribution, which represents waiting time models (Garcia-Valverde et al, 2010).

As described above, one of the most important features of AmISim is the ability to merge simulated and real components. This ability is possible thanks to the flexibility of the architecture which enables the coexistence of real and simulated bundles. In the user and environment models this capability is a fundamental key. Recall that simulation is not an overall and final solution. Therefore, the possibility of having simulations composed of real and simulated elements increases the reliability of validating AmI services which are deployed by simulated environments.

Thus, in AmISim is possible to achieve more realistic simulations using injection of real elements within the simulated world.

On the one hand, it is possible to simulate some sensors over a simulated environment but using one or more real sensors in order to test their performance as described at the beginning of this section. On the other hand, having real sensors also enables to gather information from real users. In that sense, merging real users with simulated users makes more realistic the behaviour of the user model. Thus, using AmISim is possible to have real users or some real users' features over real environments coexisting with simulated users. This ability is caused by the real context middleware.

The sensors installed in the real environment capture raw data regarding the real users, their movements, their habits and so on. This information is sent to the context middleware bundle which generates contextual information and stores it in the historical context information database. Therefore, as in the case of simulated users, the information of the users is available for the bundles in the adaptation model.

In the previous case, the user model is not simulated, but real like in the final deployment. Note that this case is not the final deployment since some sensors/actuators can be still simulated. Nevertheless, as the introduction described, having real users is unfeasible in several situations (remember the case of an emergency scenario or a scenario with thousands of users) or it is not possible due to some budget requirements, ethical problems, etc. or just because the development is in an initial phase which does not require real users yet.

Even in these cases, AmISim offers some possibilities to merging some real and simulated information, i.e., merging some real information from the users but keeping the simulated user model. As described, according to the type of behaviour a pdf is used. However, the parameters of theses pdfs can be slightly different for each person. For example, the real workers in an office use to have coffee at 9 a.m. while the simulated workers have it at 11 a.m. In both cases the pdf used is the same but the time slot is different.

This information is detected by the sensors and sent to the context middleware. Then, the adaptation model can use it to modify the parameters of the pdf in the user model. In a similar way, it is also possible to modify other parameters of the user model such as the common destinations, schedules, paths followed across the building (for example, a worker who avoids his boss' office), etc.

In both situations, the use of a full real user model or only some real features from it, merging simulated and real components, reduces the distance between the simulated models and the real world. This makes user model more realistic and facilitates the final deployment over the real world.

3.1.1 Ubik

This paper proposes a specific MABS called *Ubik* to simulate the environment and user models in AmISim. Ubik is developed in MASON[4]. There are numerous frameworks for the development of MABS and stands out from them MASON. MASON was chosen rather than other platforms because it is open source (very useful to understand deeply the implemented models), is fast (speed is necessary

[4] MASON website: `http://cs.gmu.edu/eclab/projects/mason/`

for models involving thousands of agents), replicable (quality needed to repeat exactly experiments of interest), and especially for being self-contained (Luke et al, 2004). Being self-contained is important because allows MASON to be easily a sub element of another software. In this case, the simulations developed in MASON are a component of the AmISim architecture. Another good option to develop MABS is Repast[5], which has a large community of users and, according to some opinions (Railsback et al, 2006), is the most complete platform for social simulations in Java. The web of the Open Agent Based Modelling Consortium[6] nowadays lists 21 of these frameworks and comparatives among them.

Using Ubik is possible to simulate different buildings. These buildings, as environment model of an AmI system, can be configured with different number of rooms, several kind of these, windows, doors, stairs, etc.

Each user of an AmI service is simulated in Ubik as an agent. Agents in Ubik have several features, both physical and related to the behaviours. For example, it is possible to define the age, role in the organization, speed of movement, its duties, its leadership skills, its happiness, etc. Furthermore, Ubik enables to configure the different agent's states, for example, working or running, and the actions that can be performed on the environment. All these features compose the behaviour of an agent in Ubik.

Like in the model of sensors and actuators, other behavioural features can be added to provide a wide set of configurations. In this case, this is possible thanks to the design philosophy of MABS in general and MASON in particular. This simulation platform was designed to allow a JAVA programmer to add new features easily (Luke et al, 2004).

Ubik uses a graphical 3D tool, Ubik3D, a Java3D editor to model the physical elements and the people that composes the environment and the user models in an easy way. As mentioned above, different structures of building with rooms, stairs, corridors, several floors, etc. can be simulated. The editor also contains a wide range of objects available to be incorporated to the simulated model, like different domotic devices, diverse furniture or different people. Any indoor space can be modelled using Ubik in an intuitive way without previous experience.

When the environment and user models are modelled, Ubik reads their configuration from Ubik3D and registers the elements into the context model bundle. Furthermore, Ubik defines the behaviour of each element and sets the step for each behaviour in the simulation using MASON.

Ubik is included in the AmISim architecture as a bundle of OSGi. When the elements from the environment and user models are registered in the context model, a bundle for each element is created for updating the changes of this element. This bundle will be a context producer or consumer bundle registered in OSGi (the following section gives further details).

Fig. 2 shows two figures of a floor in an office building modelled by Ubik. Both figures show the same floor with different views. There are several rooms connected by corridors. The rooms have one or more desks, chairs and some special rooms have a specific furniture. Finally, different users are simulated in this scenario.

---

[5] Repast website: `http://repast.sourceforge.net/`
[6] OpenABM Consortium website: `http://www.openabm.org/`

a) b)

Fig. 2 a) Office floor in 2D. b) Office floor in 3D

3.2 Context model

Context-awareness is key for AmI systems in order to discover and take advantage of meaningful information from the users and the environment, e.g. user location, weather, devices, user activity, etc. Contextual information must be represented in a specific model to be used. In this work, the context model is built by a context manager (real or simulated). This context manager must model the contextual information for classifying situations of interest and triggering off the relevant actions at each moment.

The context manager represents the contextual information in an ontological model. Ontologies are essential for building AmI systems because they provide a common framework to share knowledge, interoperating between users and devices and reasoning about contextual information. The usefulness of ontologies in context-aware is widely known (Gu et al, 2004; Wang et al, 2004).

Thus, the context manager is a middleware in which the applications can consume or produce information. It captures the raw information from the devices in the environment model and provides contextual information thanks to the ontology. Furthermore, it obtains new information by inference and merging context. Finally, the applications in the adaptation model can also generate new contextual information (for example, using rule-based inference or machine learning) and send it to the context manager.

Another important component in the context model is the set of context advanced services (CAS). These services are specialized services, which offer elaborated information from the contextual information in the ontology or from the historic of the ontology using the context manager. A more efficient and scalable access to applications and external services is enabled by these services. An example of these services could be offering the number of individuals in a specific place.

There is a special CAS, the *ContextHistory Context Advance Service* (see Fig.1). This CAS gives a visual interface between the users and the contextual information. Thus, the users can navigate and search visually the historic contextual information.

As well as the other context advanced services, this CAS has the semantic advantages of the semantic web technology given by the context manager. Then, SWRL rules or query languages for semantic data (like SPARQL (Pérez et al, 2009)) can be used in order to search specific contextual information. Thus, the developers are able to do complex queries over the contextual information (e.g. the produced information by the sensor with ID=7 in the last two days at intervals of 30 minutes). Finally, they can extract the required information to files in the desired format, e.g. plain text or Weka (Hall et al, 2009) format (arff). The processed contextual information is ready to be used for external services like Weka, R (Team, 2008) or external services defined by developers.

The context manager and the context advanced services are integrated into AmISim by OSGi bundles. Therefore, the context manager is a bundle that offers context management. Other bundles (in any model) use the context information as producer or consumer bundles of this context manager bundle. The ability of activate/deactivate bundles, adding or removing the context advance services according to the necessities of the system, is enabled by OSGi.

The OSGi framework maintains a central Service Registry. The Service Registry makes use of the WhiteBoard pattern. With this pattern multiple data sources as well as multiple viewers are allowed. It also has the effect of completely decoupling producers and consumers of context. Therefore, the services can be dynamically added and removed from the registry in an easy and unobtrusive way.

3.2.1 OCP

This work integrates the use of *OCP* (*Open Context Platform*) as an instantiation of the context model. OCP is a middleware that provides support for management of contextual information. This middleware represents the information in an ontological model and interprets it using SWRL rules (Horrocks et al, 2004).

OCP is integrated into the system as a bundle. Therefore, the integration and use of OCP is immediate. A context producer needs to implement: the interface *ContextProducer* to send contextual information to OCP, the interface *BundleActivator* to initialize the bundle, and the methods activate/deactivate to indicate to the producer that the service ContextService is active. Finally, the producer has to register the service in OSGi following the WhiteBoard pattern:

```
# Access to other methods: the bundle can interact with OSGi
private BundleContext bc;
...

#Register the service context producer
bc.registerService(ContextProducer.class.getName(), this , idService);
```

When the context producer is registered in OSGi, it is able to send contextual information to OCP as follows:

```
private ContextService cs;
...

#Send context information: the new location
cs.setContextItemRelation("Person", idPerson, "located", "Room", idRoom);
```

As described in the example above, the context producer sends to OCP the new location (*idRoom*) of the person *idPerson*. In the same way, when a context consumer is registered in OSGi (in this case, it needs to implement the interface *ContextConsumer*), it can receive contextual information. The context consumer registers the entities in which it has interest. The interfaces to access as a context consumer are:

```
private ContextService cs;
...

#Register for listening the changes of Person
cs.register("Person", idPerson, this);
...
```

Then, the consumer is registered to receive the changes of context of the entity *Person*. When there is a change in the context of this entity, OCP notifies it to the consumer calling to the method *notifyContextChange*, which is implemented in the consumer:

```
public void notifyContextChange(ContextEntityItems cs){

#Send context information: the new location
idRoom = cs.getContextItemString( "Person" , "412", "Room");

}
```

Therefore, the context consumer obtains the ID of the room where the person 412 is located.

In OCP there is a generic predefined OWL-based ontology. This ontology should be refined with the domain specific ontology. The ontology contains a historic of contextual information (e.g. it may contain the values of a temperature sensor for the last three days). Some applications or services require this history of the context, for example: temporal reasoning, predicting future actions or locations, learning behaviour of the users, etc. Nevertheless, there are applications that can not store this information (e.g., a mobile device with memory constraints). In such cases, the middleware provides them with this service. The generic OWL-based ontology can be accessed from: `http://darwin.inf.um.es/ocp`

The use of OCP as a context manager in the context model has an important advantage for the final deployment in the real world: it can be used directly as a real context manager of the AmI system. This means that no change is needed in the context model for the deployment over the real world .

3.3 Adaptation model

The adaptation model is the most flexible model in AmISim. It consists of all applications and services that are offered to the environment and the users. Both services and applications use contextual information from the context model to adapt to the environment and offer personalized and proactive services to the environment and users.

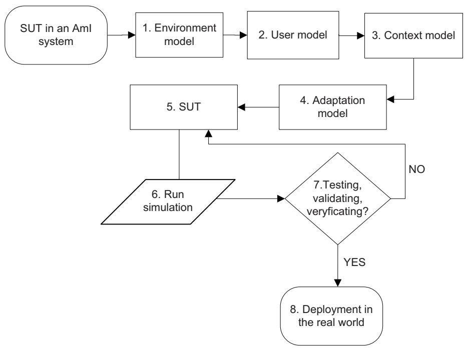


Fig. 3 Overall process

Services and applications in the adaptation model are packaged into OSGi bundles. Then, on the one hand, the service has a public API that can be accessed by the other services in the adaptation model. On the other hand, it is possible to activate and deactivate the service according to the requirements of the system.

The interaction among services and applications is implemented by OSGi. Hence, they can be discovered and used by other bundles. Furthermore, the services and applications of the adaptation model can access to the context model as context producers or consumers (as described the previous section). Thus, a service can extract the needed contextual information (from the context model) and convert it to its own format. In the same way, the service can introduce the contextual information that has been generated.

OSGi provides AmISim with a great flexibility making possible to introduce any adaptation mechanism into the bundles. Hence, it is feasible to design the adaptation logic by semantic SWRL rules, neural networks, reinforcement learning techniques, mining data stream techniques, machine learning methods using the historical contextual information, etc.

This design in the adaptation model supports other interesting features, for example, the services and applications of AmISim can be remotely controlled, managed and diagnosed without stopping the system. This characteristic is really useful for services that need some kind of offline learning. For example: obtaining offline data from the ContextHistory CAS, using this data for learning tasks, updating the information, and finally, activating the updated service.

3.4 Overall process

This section describes the overall process to test, verify and validate a SUT in an AmI system using the architecture presented.

In the first step, the environment model is defined. This process involves defining the physical world and the domotic devices (sensors and actuators), which compose the real world where the SUT is deployed. As described in section 3.1.1, Ubik is used in this work for this process. Ubik contains a Java3D editor to model the environment and users easily. Therefore, different structures of building can be simulated. Furthermore, the editor contains a wide range of objects available to incorporate to the simulated model.

The next step, step 2, includes the definition of typical user features and behaviours according to the real world. Therefore, if the real world is an office building, users will be workers of the office. However, if the real world is a hospital, the users will be patients, doctors or nurses. Again, for this process, this work uses Ubik. Ubik defines some default behaviours, which can be used; for example, movements in a building, answering a call, having a coffee, etc.

After the environment and user models are simulated, the context is defined in the step 3. This step depends on the context manager employed. Using OCP, the developer has to define the domain ontology and which components in the system are producers and consumer of context. This step is straightforward in OCP because only a few instructions are needed (see section 3.2.1).

In step 4, the adaptation model is defined. This step is highly dependent on the kind of scenario and the requirements of the SUT. Therefore, according to the SUT to be tested, some services will be needed or not. For example, if the SUT is a system to control the air conditioning depending on the location of users, an intelligent service of location will be needed in this model. Finally, when the four models are defined, in the step 5 is defined the SUT and their initial conditions.

The step 6 starts the simulation. In this step, the simulator uses the configuration defined in the previous steps to obtain the simulations. This simulation provides the needed data and views to test, verify and validate the system. Then, the suitability of the SUT can be analyzed and compared with other alternatives in the step 7. If the results obtained from the simulation find faults in the SUT, the developer ought to go back to the step 5 in order to modify the SUT and repeat the next steps. Finally, in the step 8, the obtained SUT, free of faults in its interaction with the four models, can be deployed in a real-world scenario.

4 Example of AmI service in indoor environments: location service

In order to illustrate the effectiveness of the approach presented, an example of an indoor location service is built with AmISim. Location information is one of the most essential and used contextual information in context-aware systems (Weiser, 1993). With information about location, systems can provide users with nearest features of interest and adapt to their requirements, i.e., Location-Based Services(LBSs). Users usually have some degree of regularity in their motion. Therefore, the system can predict their destination and acts proactively (Garcia-Valverde et al, 2009).

The LBS presented here can predict the location of workers in an intelligent office building. AmISim is able to simulate the building and sensors and actuators. Furthermore, the behaviour and movement of the workers in the building are also modelled and simulated. These components represent the complete environment and user models of this scenario to simulate and test the LBS.

Nevertheless, note that the context and the adaptation models are real. The context model stores the information about workers' positions and this information is used in the adaptation model to learn behaviours about users' motion and location prediction. Different algorithms and strategies can be built from the simulation to validate the service. When the LBS is tested, it can be deployed directly over the real-world environment. Hence, the use of AmISim reduces errors, problems, effort and, finally, because of that, it reduces costs in the final deployment.

The following sections describe this process and the four models of AmISim in the location service example.

4.1 Environment model

The environment and user models are simulated in this indoor location approach. Ubik bundle is used to simulate these models, as described in section 3.1.1.

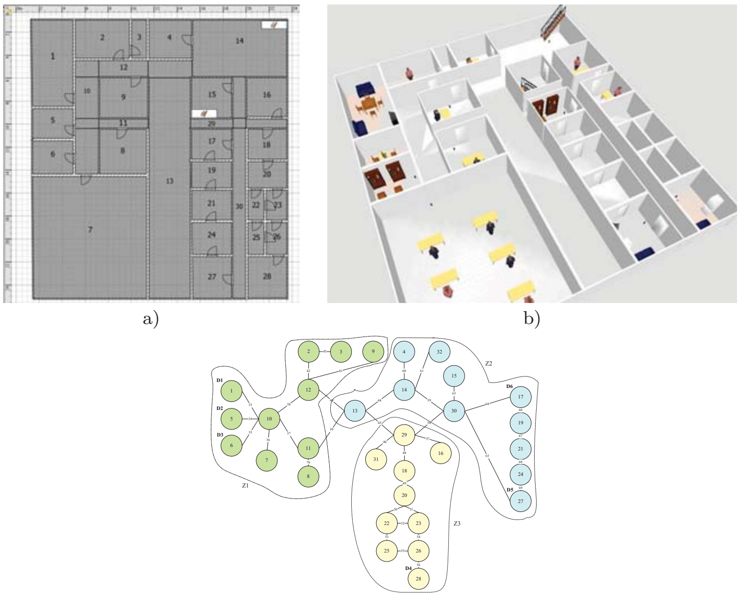
The environment model provides an office building that consists of a configurable number of floors, stairs, rooms, corridors, offices, etc. In each office, several workers agents have their desks with their own phones. Furthermore, there are some special rooms: a room to relax with a coffee maker, a warehouse and a meeting room.

There are plenty of technologies which are able to automatically identify and inform about the location of people and objects. One of the most widespread is RFID (Radio Frequency Identification). A RFID system consists of tags, which are usually passive wireless elements, attached to persons or objects. Then, fixed wireless readers, collocated within the environment, retrieve information from those tags, identifying people and objects as they are perceptible by them. Thus, when a person or an object passes near a RFID antenna, the location of this person or object is associated with that antenna.

RFID is used in this scenario to build a location service. Nevertheless, not all places in a building are usually sensorized (for reasons of budget, privacy, etc.). In such situations, the information generated by the RFID antennas can be used to perform machine learning and predict user location more effectively. With this aim, this service predicts the location of the workers in an indoor environment. Each worker wears a RFID tag, which identifies him, and the building is equipped with RFID antennas in some places where locating people is interesting. When a worker passes near an antenna, a message is generated reporting her current position to the context model. Therefore, a list of the antennas that the worker traverses is generated, i.e., a trace.

The office building and their elements are designed using the graphical tool of Ubik, Ubik3D. The figure 4 shows some views of the environment model. Figure 4 a) shows a 2D view and figure 4 b) displays the same view but in 3D. The RFID antennas are represented in blue. The building is divided into three zones $\{Z1, Z2, Z3\}$. This division is made in order to show the fact that, in a real environment, the workers of the zone Zi usually will go to destinations in this zone, the nearest destinations. For instance, if the user has a desk in zone 1, she usually will go to the relaxation room located in zone 1.

A destination is every location where a worker needs to be located in the building and there is not any antenna. In this example, the destinations, $\{D1, D2, D3, D4, D5, D6\}$, are the special rooms and they are spread out over the three zones. D1, D4 y D5 are the relaxation rooms of the zones 1, 2 and 3, respectively (the relaxation rooms are areas with coffee-maker, rest-room and a terrace for smokers). In the destination D2, there is a meeting room and the other destinations, D3 and D6, are warehouses.



a) b) c)

Fig. 4 a) Office floor in 2D. b) Office floor in 3D. c) Graph of the floor

When the Ubik bundle is registered in OSGi, the elements of the environment model are registered in the context model. Then, the context model builds an abstract schema of the building as a graph (the construction of the graph is described below). The abstract schema represents the static context about the environment. This static contextual information is used to know the structure of the building and to calculate paths through the building. The figure 4 c) shows the graph of the building. The three zones of the building are represented in the graph. Z1 is the green zone, Z2 the blue zone and Z3 the yellow zone.

4.2 User model

As described above, the user model is a fundamental goal in this kind of simulator (focussed on users). We propose MABS, by Ubik, to achieve this goal. With the use of Ubik, the behaviours of the users are modelled and simulated by using probabilistic automata.

Since users are workers is possible to assume a basic state *go to work*. The worker remains in the basic state most of the time. This kind of behaviour is an Any time behaviours. The action of this basic state can be interrupted in order to go to other states and change the behaviour. These behaviours are of two types: Monotonous behaviours and Non monotonous behaviours. Monotonous behaviours, like *having a coffee* are usually in the same time slot, whereas Non

Monotonous are not bounded to a concrete time slot, for example, *answering a phone call* or *going to the bathroom*.

Monotonous behaviours must be activated at specific time hours or within specific time intervals. For example, in the case of *having a coffee*, a new necessity of changing to such state will be generated at the moment of having coffee. In the case of *answering a phone call* or *going to the bathroom*, the necessity is not bounded to a specific time interval. The distribution that models these transitions is bounded in this time slot.

In order to generate transitions in real time, pdfs are used according to the type of the behaviour and its features. These distributions are member of the exponential family. It has been shown that these distributions are suitable to model these behaviours (Garcia-Valverde et al, 2010).

Each agent always starts its behaviour in the same point (its desk) and moves to different places in the building: relaxation room, meeting room and warehouse. These behaviours are made in a specific proportion and they depend on the zone where the agent has its desk. Users usually go to nearest destinations to their desks.

These behaviours imply movement through the office building. In order to solve the simulated movement of the users, the agents have between their beliefs the abstract schema of the map of the building (see Fig. 4 c)). Then, agents can access to the structure of the building and use it to calculate the best path for their movements using the Dijkstra algorithm (Dijkstra, 1959). When the agent knows the best path to a particular destination, this path is added to its beliefs, in a similar way as in a real scenario. This service of paths is provided from the context model by two context advanced services (see following section).

As described below, thanks to the flexibility of the architecture of AmISim, changing the strategy of the calculation of the paths using other algorithms is simple and intuitive. However, note that these behaviours can not be completely realistic because it is impossible to predict entirely the real users' behaviours. Despite that, simulations help to test and reduce risks of errors in the final deployment.

4.3 Context model

The context model is managed by the OCP bundle, as described in section 3.2.1. Moreover, in this example, OCP includes two context advanced services, the special ContextHistory CAS and a domain specific ontology. The context advances services are integrated by OSGi bundles and can communicate with the OCP bundle as context consumer/producer. The generic OWL-based ontology is refined with a domain specific ontology, which includes information about the building, the RFID devices and the workers. The ontology defines all the information regarding the environment and user model. Thus, it defines the elements of the building, the RFID antennas, their ranges and their locations, the workers and the RFID tags that identifies them.

As described above, the ontology defines the whole structure of the building. When the Ubik bundle is registered in OSGi, the elements of the building are registered in OCP and stored in the ontology. This includes the information about the floors, rooms, doors, the location of the RFID antennas, and the rest of the elements of the building (desks, phones, etc.).

There is a context advanced service bundle in the context model, which uses this contextual information to build the abstract schema of the building (see Fig. 4 c)). The abstract schema of the map of the building defines the structure of the building with areas big enough, as room, corridor, etc., but smaller areas like a cell of a grid are not considered because this would be very inefficient. This bundle builds the schema by an undirected graph with weights that models the structure of the building. Hence, the topology of the building is based on a graph $G = (V, E)$, where V is the set of rooms, corridors and stairs, E is the set of doors and crossings and $v, v' \in V$ are connected by the edge $e \in E$ if they are connected spaces in the building. Each edge has a weight X that defines the distance between the nodes that connects.

When the bundle has built the graph, the ontology is updated for setting the connections between the areas. For example, if there is an edge between $Room1$ and $Room2$ connected by $Door1$, the bundle will create a new concept $ConnectionAmongDwellings$ with the following properties: $ConnectionWith_1 = Room1$, $ConnectionWith_2 = Room2$, $WayOfLink = Door1$. Thus, the graph is stored in the ontology of OCP.

The other CAS bundle in the context model uses the stored graph to calculate the possible paths between all the rooms (nodes of the graph). The paths are calculated according to the shortest path between an origin and a destination using the Dijkstra algorithm. Hence, as described in the user model, the agents know the structure of the building and they are able to move from their desks to other destinations.

The construction of the abstract schema and the paths is divided in two bundles in order to provide flexibility to the system. Thus, it is possible to define other algorithms to calculate the paths in the building. For example, instead of the use of the shortest path, it is possible to define other strategies: the shortest path avoiding the $Room3$ (boss' desk), the path that never uses the stairs and so on. The modeller only needs to integrate a bundle with the desired strategy and this bundle will use the abstract schema for building the paths according to its strategy.

Finally, the ContextHistory CAS gives sets of data for learning tasks from the contextual information. When the environment and user models are simulated, the workers move in order to do their daily routines. Meanwhile, the activated antennas in the paths of the workers register their positions, i.e., the traces of the users in the building. Thus, the ContextHistory CAS is able to provide the set of traces of workers in order to train and validate offline techniques for location prediction in the adaptation model.

4.4 Adaptation model

In the adaptation model, three external services are proposed to location prediction: Hidden Markov Models (HMM) (Rabiner, 1989), PART (Frank and Witten, 1998) and C4.5 (Quinlan, 2003) algorithms. The three methods model the users' movement patterns. Ubik is used to generate the synthetic data set that allows the model to be trained and validated. For this reason, the three external services need to use the ContextHistory CAS bundle in order to obtain the synthetic data from the historical context information, as described in the previous section.

Here, Ubik simulates workers at work. Thus, it simulates also movements of workers at the building. By using a priori knowledge about the location of their offices, appropriate trajectories between their offices and other places can be generated and transformed them into traces. This simulation of the movements is made according to the described user model (see section 4.2). The frequency of the behaviour patterns for a user in a month are described in table 1. Finally, users are uniformly spread between the zones.

Zones \ Destinations	D1	D2	D3	D4	D5	D6
Z1	20	10	5	1	1	1
Z2	1	10	5	1	20	1
Z3	1	10	1	20	1	5

Table 1 Behaviour patterns of the users

As described in the environment model, there are special rooms: D1, D4 and D5 are relaxation rooms of the zones Z1, Z2 and Z3, respectively; D2 is a meeting room in zone Z1; and finally, D3 and D6 are warehouses of the zones Z1 and Z2. Therefore, according to the table 1, the relaxation room D1 is visited 20 times by the users in the zone 1, the meeting room D2 is visited 10 times by these users and the warehouse D3 is visited 5 times. They visit only once the other destinations. In a similar way, the relaxation room of the zone 2 (D5) is visited 20 times by the users of the zone 2 and they goes to the warehouse D3 5 times, to the meeting room 10 times and to remaining destinations once. Finally, the behaviour in the zone 3 is similar: the users of zone 3 go to the meeting room 10 times, to the relaxation room (D4) 20 times, to the warehouse D6 5 times and once to other destinations.

Note that all places are not sensorized and, therefore, it can not be assured at any precise moment t the location of an user. In these lack of information situations, the only available data are the ones generated by the antennas when the workers go past near them. Data are compounded by traces. Traces are a list of antenna locations that the agent passes through until its final destination is reached. Each time the worker leaves the desk, she is supposed to go somewhere. The list of antennas that the worker traverses, generates a trace. Then, the set of traces of a simulation is represented as

$$\begin{array}{l} \{ \ \langle (t_{1,0}, a_{1,0}), (t_{1,1}, a_{1,1}), \ldots, (t_{1,n}, a_{1,n}, l_1) \rangle, \\ \quad \langle (t_{2,0}, a_{2,0}), (t_{2,1}, a_{2,1}), \ldots, (t_{2,m}, a_{1,m-1}, l_2) \rangle, \\ \quad \ldots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \}, \end{array}$$

where $(t_{i,j}, a_{i,j})$ indicates that a user passed near $a_{i,j}$ antenna, at time $t_{i,j}$. Furthermore, l_k will be a concrete final location (e.g. a warehouse or a meeting room) if $t_{k,r} < \tau$, where τ is the threshold for a given location and $t_{k,r}$ the time that the user spends in a specific location. Such traces of antennas are analyzed to detect the most frequent destinations of users.

The ContextHistory CAS bundle classifies groups of traces to learn models according to the most frequent destination. The set of all users' traces in a month according to the above pattern makes the set of synthetic data. 2/3 of the set is used for training and the other 1/3 of data is used for validation. The ContextHistory CAS bundle is in charge of providing the data files with the 2/3 of the set for training services.

The implementation of the HMM is carried out with JAHMM(Francois, 2006), a Java implementation of HMMs related algorithms. An HMM is a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden). The basic idea is to define a set of hidden states, where each state has an associate probability distribution over a set of observables states. In each hidden state, observable information is shown according to these probability distributions. The hidden states are interconnected by probabilistic transitions between them. The HMM parameters are estimated by using the Baum-Welch algorithm. Thus, learning the HMMs consists on finding the maximum-likelihood estimator of the parameters of the HMM given a set of observed feature vectors (group of training samples for a single destination).

The other two external services to predict location use the C4.5 and the PART algorithms by the Weka data mining tool. C4.5 algorithm has been implemented using the J48 implementation in Weka. C4.5 builds decision trees using the information entropy. Decision trees are a classic way to represent information and offer a fast and powerful way to express structures in data. The trees are built using the synthetic training sequences classified by the destination. C4.5 determines in each node the attribute that best divides the data set according to the information gain.

As well as the service that uses C4.5, the service that uses PART also is built using the implementation of PART in Weka. The PART algorithm is a mixing between the C4.5 and the RIPPER algorithms. PART builds a rule and removes the instances that this rule covers. Then, the algorithm continues recursively adding rules until there are not any instance.

For the three techniques, the models of users' movements are learnt by using the data set provided by the ContextHistory CAS bundle, i.e., the preprocessed groups of traces ordered by their destinations. The final learnt models by the different techniques are built in bundles in order to offer the models of the users' behaviours to the rest of the system.

A last bundle is built in the adaptation model in order to predict the destinations. This bundle uses the learnt models and the validation data set (1/3 of data set described above) to predict the destination in each trace and validate such prediction. Besides, in the context model, the bundles that model the users' behaviours are separated from this bundle for flexibility reasons. Therefore, it is possible the use of the users' models for other purposes, not only for location prediction.

Finally, in order to show the effectiveness of the proposed framework, the three learnt models are evaluated over the synthetic validation data set. The error rate and precision are shown in tables 2 and 3. The error rate is the number of false negatives divided by true positives plus false negatives. The precision is measured as true positives divided by the sum of true positives and false positives.

Note that the HMM approach is robust with respect to an unbalanced training data distribution. For example, the HMM for D_3 has a small training set in comparison with the others. However, the learning process is not affected by this fact. Actually, comparing HMMs approach with the other machine learning ap-

Table 2 Error rates

	Destinations						
Error rate	D1	D2	D3	D4	D5	D6	Global
HMMs	20.09	65.664	47.26	1.58	25.17	24.84	30.77
PART	5.5	36.9	100	7.14	8.219	30.13	27.25
C4.5	6.78	27.45	100	0	7.30	28.76	23.65

Table 3 Precision

	Destinations						
Precision	D1	D2	D3	D4	D5	D6	Global
HMMs	0.565	0.573	0.318	0.997	0.926	0.756	0.689
PART	0.585	0.697	0	0.97	0.744	0.81	0.634
C4.5	0.604	0.683	0	1	0.906	0.765	0.659

proaches, this particular issue can be observed. For example, when PART and C4.5 algorithms are used, the prediction of D_3 and D_6 gets worse. Nevertheless, other destinations clearly benefit from this (better error rates are obtained for them) because they never predict D_3 as destination. Such destinations, D_1 and D_2, directly compete against D_3.

Additional tests were performed with all destinations equally probable. In such case, error rates were similar for the three techniques. Nevertheless, the technique based on HMMs still gets a better precision. This gives to the system based on the models from HMMs a better adaptation to changes in the workers' pattern behaviours, which is common in a dynamic scenario such as the proposed here. Thus, the example presented illustrates the flexibility and effectiveness of this approach since using the architecture is easy to choose the algorithm most suitable according to the requirements of the specific system.

5 Conclusions and future work

This work presents the AmISim architecture. AmISim integrates the basic elements of AmI systems in order to provide a framework for testing, verifying and validating this kind of systems. Four models are basic in AmI systems: environment, user, context and adaptation model. The environment includes the physical surroundings of people interacting with the AmI system. This model also includes sensors and actuators. The user model represents the user in the simulation. In AmI systems, users are the key element and all the system is designed around them. The context model is needed to provide users with services that adapt to them in the adaptation model.

The general approach proposed in this work integrates simulation and real models to test, validate and verify services under development. The proposal attempts, on the one hand, to reduce the errors and effort usually involved in the final AmI deployments over the real world. Therefore, the cost of the deployment is also reduced. On the other hand, AmISim makes feasible to test scenarios whose testing would be impossible in real environments, like an environment with thousands of individuals or an emergency scenario. We have illustrated the use and advantages of the approach by its application to a predictive Location Based Service, which is able to predict the user locations to provide them nearest features and services adapted to their requirements.

Future work will focus on improving the simulated features in order to include more complex elements, like weather conditions, advanced devices and more complex user behaviours. Another promising future work is to adapt AmISim to implement an infrastructure for mobile environments. In this line, we are trying to integrate real mobiles devices in the simulation. This inclusion will offer more realistic models due to the injection of real features from the environment and the users. The flexibility of the AmISim architecture will allow us to carry out this future work to reduce the distance between the simulated and real worlds.

Acknowledgments

This research work is supported by the Spanish Ministry of Science and Innovation under the grants AP2007-04269 and AP2007-04080 of the FPU program and in the scope of the Research Projects DIA++ (TRA2009_0141) and TIN2011-28335-C02-02, and through the Fundación Séneca within the Program "Generación del Conocimiento Científico de Excelencia" (04552/GERM/06).

References

Aarts E (2004) Ambient Intelligence: A multimedia perspective. IEEE MultiMedia 11(1):12–19

Abowd GD, Dey AK, Brown PJ, Davies N, Smith M, Steggles P (1999) Towards a better understanding of context and context-awareness. In: HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing, Springer-Verlag, London, UK, vol 1707, pp 304–307

Barton JJ, Vijayaraghavan V (2002) Ubiwise: A ubiquitous wireless infrastructure simulation environment, tech. report hpl2002-303. HP Labs 2002

Brown PJ (1996) The Stick-e Document: a Framework for Creating Context-aware Applications. In: Proceedings of EP'96, pp 259–272

Bylund M, Espinoza F (2002) Testing and demonstrating context-aware services with Quake III Arena. Communications of the ACM 45(1):46–48

Coporation V (1998) Half life. Half Life URL http://en.wikipedia.org/wiki/Half-Life_(video_game)

Dijkstra E (1959) A note on two problems in connection with graphs. Numerische Mathematik 1:269–271

Drogoul A, Vanbergue D, Meurisse T (2002) Multi-agent based simulation: Where are the agents? In: MABS, pp 1–15

Francois JM (2006) Jahmm-hidden markov model (hmm): An implementation in java. URL: http://www run montefiore ulg ac be/francois/software/jahmm

Frank E, Witten IH (1998) Generating accurate rule sets without global optimization. In: Proceedings of the Fifteenth International Conference on Machine Learning, University of Waikato, Department of Computer Science, pp 144–151

Friedewald M, Da Costa O, Punie Y, Alahuhta P, Heinonen S (2005) Perspectives of ambient intelligence in the home environment. Telematics and Informatics 22(3):221–238

Gabel O, Litz L, Reif M (2005) NCS testbed for Ambient Intelligence. In: 2005 IEEE International Conference on Systems, Man and Cybernetics, IEEE, vol 1, pp 115–120

Garcia-Valverde T, Garcia-Sola A, Lopez-Marmol F, Botia JA (2009) Engineering Ambient Intelligence services by means of MABS. International Conference on Practical Applications of Agents and Multi-Agent Systems

Garcia-Valverde T, Campuzano F, Serrano E, Botia J (2010) Human behaviours simulation in ubiquitous computing environments. In: Workshop on Multi-Agent Systems and Simulation at MALLOW

Garlan D, Siewiorek DP, Smailagic A, Steenkiste P (2002) Project Aura: Toward distraction-free pervasive computing. Pervasive Comput 1(2):22–31

Gu T, Pung HK, Zhang DQ (2004) Toward an OSGi-Based Infrastructure for Context-Aware Applications. IEEE Pervasive Computing 3(4):66–74

Hagras H, Callaghan V, Colley M, Clarke G, Pounds-Cornish A, Duman H (2004) Creating an ambient-intelligence environment using embedded agents. IEEE Trans Intell Syst 19(6):12–20

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1):10–18

Horrocks I, Patel-Schneider P, Boley H, Tabet S, Grosof B, Dean M (2004) SWRL: A Semantic Web Rule Language Combining OWL and RuleML. Tech. rep., National Research Council of Canada, Network Inference, and Stanford University

Intille S, Larson K, Kukla C (2000) House_n

Kidd CD, Orr R, Abowd GD, Atkeson CG, Essa IA, MacIntyre B, Mynatt E, Starner TE, Newstetter W (1999a) The aware home: A living laboratory for ubiquitous computing research. Lecture notes in computer science pp 191–198

Kidd CD, Orr R, Abowd G, Atkeson CG, Essa IA, MacIntyre B, Mynatt E, Starner TE, Newstetter W (1999b) The Aware Home: A Living Laboratory for Ubiquitous Computing Research. Lecture Notes in Computer Science pp 191–198

Liu Y, OGrady MJ, OHare GMP (2006) Scalable Context Simulation for Mobile Applications, Lecture Notes in Computer Science, vol 4278, Springer Berlin Heidelberg, chap chapter 42, pp 1391–1400

Luck M, McBurney P, Shehory O, Willmott S (2005) Agent technology: computing as interaction (a roadmap for agent based computing). University of Southampton

Luke S, Cioffi-Revilla C, Panait L, Sullivan K (2004) MASON: A New Multi-Agent Simulation Toolkit. In: Proceedings of the 2004 Swarmfest Workshop

Lu H WK Chan, Tse TH (2006) Testing context-aware middleware-centric programs: a data flow approach and an RFID-based experimentation. In: SIGSOFT '06/FSE-14: Proceedings of the 14th ACM SIGSOFT international symposium on Foundations of software engineering, ACM Press, pp 242–252

Mangina E, Carbó J, Molina JM (eds) (2009) Agent-based Ubiquitous Computing, Atlantis Ambient and Pervasive Intelligence, vol 1. Atlantis Press

Marples D, Kriens P (2001) The Open Services Gateway Initiative: An introductory overview. IEEE Communications Magazine 39(12):110–114

Martin M, Nurmi P (2006) A generic large scale simulator for ubiquitous computing. Mobile and Ubiquitous Systems, Annual International Conference on 0:1–3

Morla R, Davies N (2004) Evaluating a location-based application: A hybrid test and simulation environment. IEEE Pervasive Computing 3(3):48–56

Nieto I, Botía JA, Gómez-Skarmeta AF (2006) Information and hybrid architecture model of the OCP contextual information management system. Journal of Universal Computer Science 12(3):357–366

O'Neill E, Klepal M, Lewis D, O'Donnell T, O'Sullivan D, Pesch D (2005) A testbed for evaluating human interaction with ubiquitous computing environments. In: Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005. Tridentcom 2005. First International Conference on, IEEE, pp 60–69

Pérez J, Arenas M, Gutierrez C (2009) Semantics and Complexity of SPARQL. ACM Transactions on Database Systems (TODS) 34(3):1–45

Qiu F, Hu X (2010) Modeling group structures in pedestrian crowd simulation. Simulation Modelling Practice and Theory 18(2):190–205

Quinlan JR (2003) C4.5: programs for machine learning. Morgan Kaufmann

Rabiner LR (1989) A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2):257–286

Railsback SF, Lytinen SL, Jackson SK (2006) Agent-based simulation platforms: review and development recommendations. SIMULATION 82(9):609–623

Reynolds V, Cahill V, Senart A (2006) Requirements for an ubiquitous computing simulation and emulation environment. In: InterSense '06, ACM Press, p 1

Rudolph L (2001) Project Oxygen: pervasive, human-centric computing - an initial experience. In: CAiSE '01: Proceedings of the 13th International Conference on Advanced Information Systems Engineering, Springer-Verlag, London, UK, pp 1–12

Salber D, Dey AK, Abowd DG (1999) The context toolkit: aiding the development of context-enabled applications. In: CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, New York, NY, USA, pp 434–441

Serrano E, Botia JA, Cadenas JM (2009) Ubik: a multi-agent based simulator for ubiquitous computing applications. Journal of Physical Agents 3(2):39

Team R (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria ISBN 3(10)

Tse TH, Yau SS, Chan WK, Lu H, Chen TY (2004) Testing context-sensitive middleware-based software applications. In: Computer Software and Applications Conference, 2004. COMPSAC 2004. Proceedings of the 28th Annual International, IEEE, pp 458–466

Wang XH, Zhang DQ, Gu T, Pung HK (2004) Ontology based context modeling and reasoning using OWL. Pervasive Computing and Communications Workshops, 2004 Proceedings of the Second IEEE Annual Conference on pp 18–22

Weiser M (1993) Hot topics-ubiquitous computing. Computer 26(10):71–72

Weiser M (1995) The computer for the 21st century. Scientific American 272(3):78–89

5.2 Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach

Title	Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach
Authors	Garcia-Valverde, T. and Campuzano, F. and Serrano, E. and Villa, A. and Botia, J. A.
Type	Journal
Journal	Journal of Ambient Intelligence and Smart Environments
Impact factor (2010)	1.5
Year	2012
State	In press

Journal details: Journal of Ambient Intelligence and Smart Environments
Editors-in-Chief: Hamid Aghajan and Juan Carlos Augusto
ISSN: 1876-1364 (print version)
ISSN: 1876-1372 (electronic version)
Publisher: IOS Press
Impact factor (2010): 1.5
Website: `http://www.jaise-journal.org/`

Authors – Personal details	
Name	**María Teresa Garcia Valverde**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Francisco José Campuzano Adán**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Emilio Serrano Fernández**
Position	Member of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Ana Villa Caballero**
Position	MSc in Computer Science and Researcher in Ami2
University/ Organization	AmI2, Ambiental Intelligence and Interaction (Murcia)
Name	**Dr. Juan Antonio Botía Blaya**
Position	Lecturer of the Department of Information and Communications Engineering
University/ Organization	University of Murcia

Contribution of the PhD student
The PhD student, María Teresa García Valverde, declares to be the main author and the major contributor of the paper *Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach*

Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach

Teresa Garcia-Valverde [a], Francisco Campuzano [a] Emilio Serrano [a] Ana Villa [b] and Juan A. Botia [a,*]

[a] *Universidad de Murcia, Campus Espinardo, Murcia, Spain*
E-mail: {mtgarcia, fjcampuzano, emilioserra, juanbot}@um.es
[b] *AmI2, Ambiental Intelligence and Interaction, CEEIM, Campus Espinardo, Murcia, Spain*
E-mail: {ana.villa}@ami2.net

Abstract. An Ambient Intelligence (AmI) system is a pervasive system in which services have some intelligence in order to smoothly interact with users immersed in the environment. Users are the main entity in AmI systems. Thus, the services are user-centred with an adaptive interaction that provides users with more personalized facilities in a non-intrusive way. The deployment of AmI services is a hard task, especially when the system includes complex users' behaviours or complex deployments in terms of the devices and software participating at the system. Since one of the intrinsic requirements in these services is smooth interaction with users, the user, or at least a model of the user, should be incorporated in the development process. This makes the process of verification and validation of such services quite difficult. This paper proposes the use of Agent Based Social Simulation for a quick and feasible validation of AmI systems. Simulation in the first stages enables easy validation of the services before real environment deployment. The main challenge in this approach is how to simulate humans with realistic behaviours. In this paper, Social Simulation (SS) is used to tackle this problem. Specifically, this paper presents a methodology for the validation of AmI systems by SS. The realistic modelling of users and its validation are key elements of this process and involve challenges that the techniques proposed here are able to deal with. A real application example which shows the level of reality reached in the users' models and the benefits of the methodology is presented in this paper.

Keywords: Ambient Intelligent, Ambient Assisted Living, Behaviour Simulation, User Modelling

1. Introduction

An Ambient Intelligence [2] system is a set of appliances, services and applications which silently surrounds and interacts with the user in an intelligent manner. Some of the commonest examples of these projects are: MavHome [14], iSpace [19], Aware Home [31] or the House_n [29]. In these projects the environment has intelligence and makes decisions regarding the current context and interactions with users.

The last years have seen substantial progress in AmI [1]. There have been lots of developments and projects which apply the issues of ambient intelligence in different contexts: smart homes [51], health monitoring and assistance [11], hospitals [38], transportation [48], emergency services [16], education [55], workplaces [24], etc. A state-of-the-art in-depth view of AmI applications in different fields can be found in [13]. In such scenarios, the user is the central entity. Services and applications

*Corresponding author. E-mail: juanbot@um.es.

are built using the user as the central element of product creation.

One of the central issues of AmI services development is verification and validation of such services by testing. Testing software is the process of executing a program in order to find errors in the code [41]. Such errors must then be debugged. According to the IEEE Standard Glossary of Software Engineering Terminology [46]: “Validation is the process of evaluating a system or component during or at the end of the development process to determine whether it satisfies specified requirements”. Software validation is, accordingly, the process of checking that what it does it what it is supposed to do. A failure is found when it is detected that the software does something different to what it should do by its design. Testing and validation are concepts that are tightly related to software quality [3].

Considering an AmI service or application as the system under test (SUT), some of the errors may be found by using a Unit approach (e.g. out-of-range variable values, shoddy checking of return values from methods and so on). Robustness of the code is crucial here. However, in this particular kind of SUT, and provided that robustness is already guaranteed (e.g. by previous Unit testing and debugging), the correct functionality of the SUT must be verified.

There is a variety of AmI based ubiquitous computing systems. In terms of the number of devices and software subsystems implied, a simple example is a continuous monitoring system for elderly people who live independently in their own houses. A more complex system would be an intelligent building in charge of controlling the whole building. All of them have elements in common: they include users and devices, mainly sensors and actuators. How do we design and perform a validation process of the SUT when users and devices are central elements, without real users and devices? To put the question in another words, how can we attain sufficient SUT quality in a stage prior to deployment of the real system, and so minimize the risk of finding errors when the real system is installed?

This work rests on the following hypothesis: it is possible to validate the SUT by recreating users and devices in a simulated physical environment (e.g. a house or a building). This statement will be validated below. It is important to note that simulation is not an overall solution. Probably, more validation tests would be needed after deploying it in a real environment. But integrating the real SUT in a previous simulation process may help to validate the SUT, or at least minimize the risks of finding errors when the AmI system is deployed in production mode (i.e. as a product). Moreover, it is possible to achieve more realistic simulations using injection of real elements within the simulated world. Thus, simulations are composed of both real and simulated elements. This realism increases the reliability of validating AmI services which are deployed in simulated environments.

Our particular approach for simulation is based on social simulation (SS). SS is a cross disciplinary research and application field which combines computer simulation and social science [18]. Human societies are often complex adaptive systems with a high number of complex interactions between their members. SS allows us to deal with these complex social processes since traditional computational and mathematical models have difficulties doing it [35]. To that end, SS uses an extremely simple version of the agent metaphor to specify single components and interactions among them. But using SS models in simulations implies that such models should also be validated. Validating the SUT with guarantees of success requires that simulation models reflect reality. So, such models can be seen as social models under test (SMUT) and some validation process should be applied to them before they are used for validating the SUT. The process of validating SMUTs and use them to validate SUTs is not trivial because it implies a lot of complex components. Thus, the paper shows a methodology intended to guide the developer in the process. Note that, although the methodology copes with both validation of SMUTs and SUTs, the paper only illustrates the application of the methodology for the validation of a SMUT. This illustration is based on a real domain. The application domain is AAL (Ambient Assisted Living). An AAL system is an ICT (Information Communication and Technology) based solution which is devoted to augmenting the quality of life for elderly people. In this case, the interest is focused on a system called Necesity [10]. The paper shows how the methodology is applied to validate models for the elderly. These models can then be used to validate any AmI service to be built for such domain by following an approach based on simulation.

The paper is structured as follows. Section 2 introduces the methodology. Section 3 details the AmI system used to illustrate the proposal. Section 4 introduces the computational models employed to artificially reproduce the behaviour of the user. Section 5 presents the application of the methodology to validate such computational models. Section 6 discusses some related works. Finally, the conclusions and future works are presented.

2. Proposal of the Methodology

Suppose an AAL system capable of monitoring elderly people who live alone at home [10]. It continuously monitors the elderly person and takes decisions about their state by a fixed mechanism (i.e. it behaves in the same manner with any elderly person and uses no adaptation). Such mechanism detects when something is wrong, by using data coming mainly from activity sensors deployed at the house. Suppose that the requirements of the problem slightly change so that a new service is needed. New adaptive capabilities are needed: i.e. the mechanism in charge of detecting abnormal situations should now adapt to the elderly people's patterns of behaviour in order to produce a more effective response (i.e. adaption to particular users implies that the system can manifest a quicker reaction). How can such a new service be developed? Moreover, how do we assess the effectiveness of such a service and demonstrate that it behaves better than the earlier, rigid version? A first engineering choice would be to program a new service in the lab, make a controlled test of its adaptive capabilities and then test it during some weeks with real users. Some drawbacks are in the way: what if we do not have such real users available for test purposes? What if real users are available but no abnormal situations are produced during the test phase? An approach based on simulation would be equally effective to validate the adaptive service, provided that the house, the sensors and appropriate elderly people can be simulated. Reality should be preferred to simulation, but when a real set up is not affordable, simulation may help. Moreover, simulation models can be reused when new AmI services are needed within the same or similar domains.

This section proposes a methodology to address this problem: verification and validation of AmI services by using simulation. Thus, the methodology also addresses the development of simulation models (including their validation as SMUTs) as they are needed to validate the SUT. It is an extension of the methodology promulgated by Gilbert and Troitzsch [27] for the development and use of SS models.

Gilbert and Troitzsch's methodology [27] is interesting in this context for a number of reasons. The first is that it pays the necessary attention to model building (not only is a good design of a model important, but also a good implementation). Other similar methodologies, such as the one proposed by Fishwick [22,23], do not do this. The second is that it is the most popular methodology for developing social simulations [20].

We have extended Gilbert and Troitzsch's methodology with two new and necessary elements: (1) the existence of a step to generate simpler simulation models to cope with high complexity and (2) the consideration of including real elements in the simulation to get a more effective validation of SUTs. The methodology is expressed as a flowchart in Figure 1[1]. It is called AVA, which stands for Ambient intelligence services VAlidation.

Gilbert and Troitzsch [27] define the target of a social simulation as some real world phenomenon in which the researcher is interested. In this case, the whole phenomenon refers to an ambient intelligence system deployed at a physical space in which users evolve. Thus, the phenomenon includes a physical environment, persons acting within the environment and the devices and software of the AmI system. Let us suppose that the AmI software is made up by services, applications and a middleware in charge of supporting services and applications. Let us also suppose that the middleware and some services and applications have been validated previously (i.e. they all are already mature). Then, the methodology is proposed to vali-

[1]The flowchart uses standard elements of the classic flowcharts [28] as flow of control (represented as arrows), processes (represented as rectangles), decisions (rhombuses), input/output (parallelograms), start and end symbols (ovals) and predefined processes (rectangles with vertical lines at the sides).

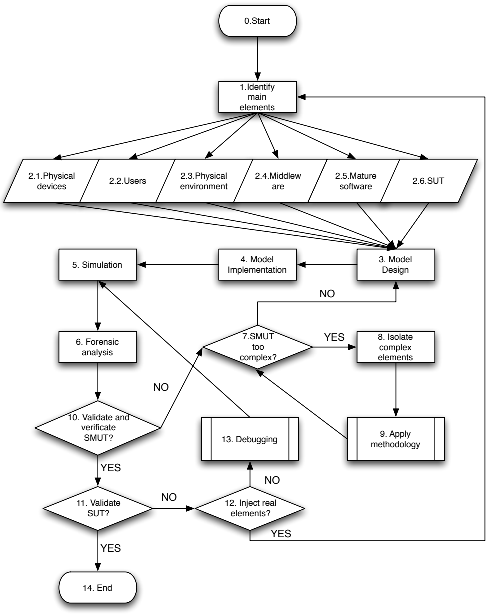


Fig. 1. Proposal of a methodology for the validation of AmI services

date those services and applications which are already developed but not totally mature (i.e. the SUTs) by means of creating, validating and using SS models (i.e. the SMUTs) for physical environment, users and devices.

The application of the methodology is based on the concept of cycle and iteration. A cycle comprehends the application of the methodology to validate a single SUT (e.g. a location service based on Bluetooth for an intelligent building) and it is compound of a number of iterations. An iteration inside a cycle refers to the application of the flow diagram process of Figure 1 in the context of validating the SUT. Within a cycle, a number of iterations will be executed until the SMUTs are validated and, after that, the SUT is validated also. Thus, given n SUTs to validate, the methodology should be applied by means of n cycles, one for each SUT.

Step 1 of the methodology, labelled with *Identify main elements* includes the identification and accounting of physical devices (element 2.1), kind and number of users (element 2.2), a description of the physical environment (element 2.3), the middleware (element 2.4) and the separation of services and applications into two subsets: those which are already mature (element 2.5) and those which are chosen to be validated (element 2.6). A third subset is left out and includes these services which need be validated but will not be in this application cycle of the methodology. This identification of an accounting task is a necessary familiarization with the domain required by the simulation models developer. Now, the developer of the models is ready to design the models [27].

Step 3 of the methodology refers to the design of the SS models for elements in 2.1, 2.2 and 2.3 (physical devices, users and physical environment) and how they are integrated in the simulation with the real software (i.e. middleware (2.4), mature software (2.5) and SUT (2.6)). Note that an effective validation of the SUT needs realistic models for users, mainly.

Step 4 deals with the necessary activities of implementation. On the one, the model design outcome must be coded into the language of the specific simulation platform used. On the other hand, elements 2.4, 2.5 and 2.6 (middleware, mature software and SUT) must be integrated with the specifications of the simulation models. The outcome from step 4 should be something ready for simulation. To pass from a simulation model design to a model implementation ready for simulation in a concrete SS platform is not a trivial task [20]. A general programming language is not a feasible option because some elements physical devices (2.1), users (2.2) and physical environment (2.3) usually have considerable complexity. Furthermore SS platforms can help to alleviate the complexity of coding the design of models (i.e. tools for several of the typical tasks in the construction of SS have been included in this kind of software packages [27]). The web of the Open Agent Based Modelling Consortium[2] nowadays lists over twenty of these SS platforms.

Step 5 is in charge of simulation. The first iterations within a methodology application cycle are devoted to validating the SMUTs. In subsequent iterations within the same cycle, once the SMUTs have been validated, the focus of the simulations is put on validating the SUT. SS platforms will offer the possibility of easily performing different experiments of interest and generate simulation data

[2] OpenABM Consortium website: http://www.openabm.org/

for analysis in a convenient way. It is up to the developer how to design the experiments, depending on the particular scenario and also if the target is a SMUT or a SUT. Note that these simulation experiments are actually software tests. Simulation models are used to make tests on the SUT.

Once simulations are run, it is time to analyze data. Forensic analysis (Step 6) is an offline simulation data analysis to be conducted on the data generated in the previous step. The analysis should consider three important questions that should guide the process. The first has to do with the complexity of SMUTs. If they are affordable, the second one addresses whether the SMUT of the corresponding methodology iteration is correct in order to validate it. Finally, the third question, once all the SMUTs have been validated, is about the correctness of the SUT in the corresponding methodology cycle.

One of the innovative points in the methodology is covered in Step 7. It consists of including the possibility of dividing complex simulation models into simpler ones, i.e. creating and validating a set of $SMUT_1, \ldots, SMUT_m$ from a complex SMUT (Step 8). Hence, an iteration of the methodology is applied to each $SMUT_1, \ldots, SMUT_m$ (Step 9). Once all the $SMUT_1, \ldots, SMUT_m$ are validated, it is easier to validate the SMUT. Note that a SMUT is considered as too complex when it is difficult to assess whether the behaviour of that SMUT is the expected one. For example, some behaviours of users can be so complex that they need to be evaluated in isolation. An example of this type of behaviour would be the resolution of collisions on the motion of a large number of dynamic entities. This behaviour is a problem to be studied by itself and its validation as a SMUT would be much more complicated with additional elements in the simulation (additional users' behaviours, an environment model, the SUT and the rest of the software).

Step 10 is a branch which indicates that if the SMUT of the corresponding iteration is not mature, the model design must be revised. Only when all the SMUTs have been validated, is it time to try validate the SUT (Step 11). In other words, when the simulation models are validated, then the whole simulation is ready and the focus of attention can be directed to the SUT validation.

Step 12 is another innovative aspect of the methodology: the integration of the simulation with real elements. Think, for example, of a very simple simulation in which a real activity sensor is connected with the simulation through the simulation platform. This sensor is associated to a concrete room in the simulated environment of the methodology's element 2.3. In the rest of rooms, the corresponding activity sensors are simulated. This kind of set up would be useful for validating aspects like the integration of real sensors with the middleware in charge of reading, storing and redistributing sensor data to others by comparing readings generated by the real sensor with those generated by simulated ones. Thus, the benefit of this innovation is that the simulation gains in realism, in some sense, when real devices interact with simulated elements (i.e., injection of real elements). If no injection is necessary in this methodology cycle, and the SUT is still not ready, then debugging the SUT code (Step 13) is necessary (i.e. it is supposed that in Step 5, the SUT was tested and Step 6 generated evidences about possible errors). As debugging is locating and fixing errors in the SUT, it does not imply changing the simulation model produced in Step 4. Thus, from Step 13, the process goes again to simulation, Step 5. A cycle on the methodology, i.e. the validation of a concrete SUT, ends when its behaviour is correctly verified.

2.1. Model Design

The design of a simulation model is performed in Step 3 of the methodology depicted in Figure 1. This step is in charge of producing designs for simulation models of physical devices (2.1), users (2.2) and physical environment (2.3). Such elements, in the form of simulation models, are the so called SMUTs. Step 3 can be further refined as a flow chart. It appears in Figure 2.

The first point at the design of a SS model is the definition of classes [27]. Recall from the last section that models have to be designed for physical devices (2.1), users (2.2) and physical environments (2.3). Thus, classes referring to such categories are needed. Additional classes can be specializations of these, depending on the domain and scenario. Such classes must be separated into two main categories: agent-based classes and the rest. A class must be modelled as an agent if objects of that class must evolve over time. A typical complex element which evolves over time is a user. But

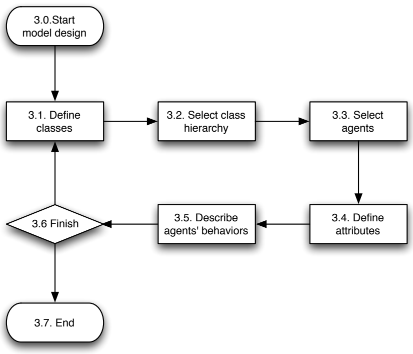


Fig. 2. Flowchart for the model design process of the methodology

simple elements could also be modelled as agents. For example, regarding classes for the physical environment, a wall in a building does not change, it is fixed over time. However, a door can be opened or closed, but there is no evolution over time if the change is caused exclusively by a user opening or closing it. Nevertheless, if a door must react autonomously according to external events (e.g., be closed in some periods or after some alarms), it could be modelled as an agent.

For each concept modelled as a class, the corresponding attributes describing their properties must be specified. The next step is to describe the behaviour of those elements modelled as agents. These behaviours must include interactions between elements (i.e. interactions between individual agents and the environment or among agents themselves, when appropriate [27]). Model dynamics may be modelled with any abstract machine, e.g. finite and deterministic automata, petri nets, etc.

It is important to stop modelling at the right level of complexity (i.e. the level of complexity required for a correct validation of the SUT). For example, if the SUT is a location service which depends on the user's trajectory in a house, the user behaviour should perform interesting trajectories in order to generate the necessary tests to evaluate the location service's accuracy. If the SUT is an adaptive service which depends on the particular user habits at home, the behaviour should be able to artificially reproduce such habits. Moreover, it should be customizable so as to reproduce different users. For all of this, it is necessary to clearly identify to what extent behaviours are realistic. In other words, when does the developer stop modelling? (finish condition in Figure 2).

The condition that must be checked to decide when to stop is whether the model is *minimalist* (i.e. it includes only the relevant elements for the problem at hand) in order to simplify the analysis phase as much as possible [40,27]. But at the same time, the model must describe reality in some sense. This trade off introduces the confrontation between the KISS approach ("Keep it Simple Stupid") [7] and the KIDS approach ("Keep it Descriptive Stupid") [21] for social simulation based modelling. The KIDS approach leads to as descriptive as possible design models, but often these models are quite complex. Later, the model can be simplified if it is sufficiently justified. On the other hand, the KISS approach aims to make models as simple as possible. These models only become more complex if they are unable to reproduce possible and interesting phenomena that may occur in reality. The KISS approach proposed by Axelrod [7] is the mainstream approach in social simulation. The goals of KISS and KIDS approaches are different. KISS simulations study the macro-micro links in the artificial society while KIDS simulations are used to produce realistic simulations [30]. To correctly validate a SUT, whose correct behaviour depends on users, realistic simulations rather than simple simulations are better, in principle. Thus, the proposed methodology focuses on KIDS models.

2.2. Forensic Analysis

The forensic analysis step (Step 6 in Figure 1) refers to the study of the results from the simulation runs. Such study covers data selection, data transformation and data representation and analysis tasks. However, no conclusions are drawn at this stage. They are generated afterwards, in steps 7, 10 or 11 depending whether the SMUT is too complex, or when it is ready to be validated or when the SUT is ready to be validated, respectively. Note that conclusions drawn in steps 7, 10 and 11, are used to decide if the object under evaluation should be validated. But, when a pilot study follows the application of the methodology, such conclusions should again be checked with the testbed [20].

Steps 10 and 11 work similarly: they both address verification first and then validation. In verification, bugs are searched for in the simulation data analysis results. When no bugs are detected, the object (either a SMUT or a SUT) is verified. Then comes validation. Validation checks, again by using simulation data analysis results, that the behaviour of the object is the desired one by proving that requirements are accomplished. One of the most popular ways for verifying SS models is the Unit approach. It can detect implementation bugs such as values of variables out of range or incorrect return values from methods. It works by writing tests, simultaneously with the coding activity, and performing these tests automatically each time the code is modified. Note that, if the simulation engine incorporates displays of some interesting aspects of the simulation, such displays can play a relevant role in the verification process. Through them, a developer can quickly detect bugs in the simulation. In this sense, displays are equivalent to monitoring the evolution of a large number of variables. SS models validation is a more abstract task. Most authors perform the validation by using numerical evidence tests [40].

For example, SMUT simulation results may be compared with real data obtained from the real phenomenon. This data can be obtained from the sensors in particular, the context in general and information external to the system, such as the users' profiles. Naturally, to have real data available is not always possible. In the experiments of this paper, this type of approach is used (see section 5). Another effective but costly option for validation of simulation models is to implement the model twice by using two different target simulation platforms and two different programmers [27]. Then, the behaviour of the two simulations are compared. Note that, the verification of a SUT is similar to that applied to a SMUT. The Unit approach is usually addressed. In fact, verification of the SMUT is by the application of the methodology. It uses the Unit approach as each simulation run implies that the SMUT is used many times with different input values. Such input values are usually conditioned by sensor data, which at the same time, are conditioned by the evolution of the user is the environment and how sensors perceive it. The same discussion on validation of the SMUT may be applied to the SUT.

Step 7 detects if the SMUT is too complex when the simulation data analysis results give evidences indicating that the SMUT can not be validated. Note that when there is no possible validation, it can be due for two reasons. The first one is that there are bugs in the model design. In this case, the flow of control should return to Step 3. The second one is that requirements are not accomplished but not for a bug. This is due to handling a extremely complex model.

2.3. Software Tools for Application of the Methodology

The methodology presented above depends on the availability of software tools for a successful application when a real problem is faced. For example, in order to validate a concrete SUT (e.g. a location service), and following the methodology, a simulation engine and the corresponding simulation models for elements 2.1 (physical devices), 2.2 (users) and 2.3 (physical environment) are needed. Besides, the necessary middleware and additional services needed by the SUT to correctly run must also be available. Furthermore, additional tools to perform forensic analysis should be also available in order to avoid performing simulation data analysis from scratch. For this purpose, the UbikSim[3] tool [26] has been created together with the methodology described. It has the following main features:

- It is based on the MASON social simulation engine; thus, implementation of the models must adjust to this engine.
- Basic SS models for physical environments (e.g. offices building floors, hospital floors, etc.) are already developed and validated.
- Basic SS models for humans are already included in the framework and validated for specific environments like, for example professors in universities, caregivers in hospitals, etc. They were created to be used in former projects and can be reused and/or tailored in similar situations.
- The same goes for simple SS models for sensors (e.g. presence, pressure, open door sen-

[3] UbikSim website: `http://ubiksim.sourceforge.net`. The website offers several videos to view the operation and evolution of UbikSim.

sors and Bluetooth and RFID tags and antennas).
- All the releases of UbikSim at the moment are integrated with a specific ubiquitous computing middleware (i.e. the 2.4 element in Figure 1).
- Basic ubiquitous computing and AmI services (i.e. 2.5 element in Figure 1) have been developed, verified and tested as SUTs and can be used as mature software.

The architecture of UbikSim is structured in software components. It is based on an OSGi service oriented architecture [4]. OSGi brings a very important benefit to the UbikSim architecture: flexibility and capacity of evolving through time. For example, UbikSim's current configuration is based on MASON but this could be easily changed by integrating any simulation platform (e.g. Repast or Netlogo) into a bundle with the same API (Application Program Interface). OSGi is responsible for coordinating components during runtime. Figure 3 shows the interrelations among components by means of a collaboration diagram.

OSGi and MASON are 100% Java. Therefore, SMUTs (which are simulated by MASON) and SUTs (which must be integrated into an OSGi bundle) should also be based in Java. OSGi components are called bundles. Bundles in Figure 3 are labelled with Bn, where n is the index of the bundle. For example, $B1$ is an environment modelling tool based on Java3D for modelling people and their environment. It is possible to create buildings on a 2D plane and it also offers a 3D navigable view (Figure 4). $B2$ is a 3D viewer that shows how the simulation evolves while it is running. The use of this display is also very useful for verification (see section 5.1).

$B3$ is the MASON simulator engine of UbikSim. MASON is the platform of choice because it is one of the most popular and powerful opensource SS platforms. $B4$ is a middleware for context information management. Recall that to deploy an ubiquitous computing system, a middleware is necessary to allow deploying context-awareness and flexible management of entities. The middleware used in UbikSim is Open Context Platform (OCP), developed in our lab [43]. OCP was OSGi based before UbikSim existed, so integrating it was extremely easy. As the bundle has a clear API, other middlewares (e.g. JCAF or the Context Toolkit to name a few) might be alternatives to OCP. Note from the Figure that OCP works by using ontologies for the representation and storage of context. This means any other bundle can use past sensor lectures or other user context for its own purposes. B5 is the SUT. Note that the SUT interacts mainly with $B4$. It could use user context (e.g. location information) or user's past history to develop an adaptive service by machine learning.

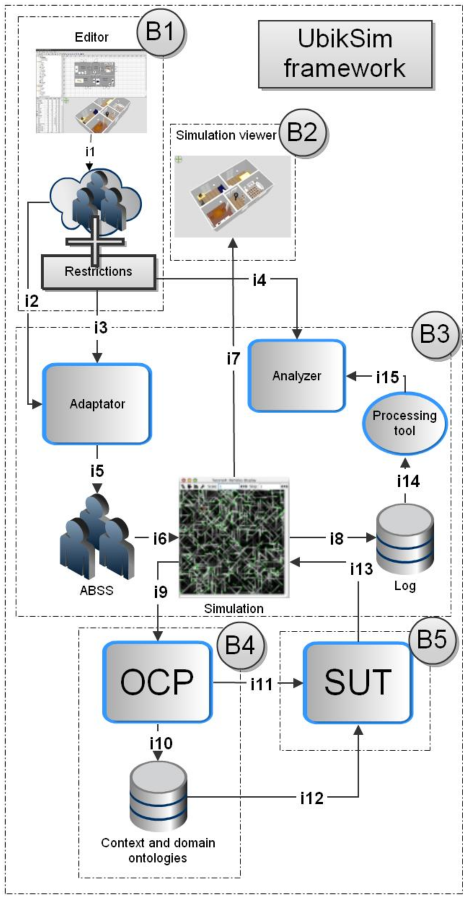


Fig. 3. Collaboration diagram for the UbikSim simulator framework

3. Introducing the Running Example

This section is the first part of three devoted to giving an example of how to apply the methodology. This first part introduces the Necesity system, verified and validated by using the methodology and how the physical environment and devices were created using UbikSim. The second part, in section 4, explains the kind of model we use to simulate the user of this example. The third part, in section 5, explains how to verify and validate such model. Thus, a concrete iteration of the methodology is illustrated, with the user simulation model as a SMUT.

Necesity[4] is an Ambient Assisted Living (AAL) system. An AAL system is any ICT system which contributes to enlarge the amount of time that elderly people living independently can maintain such independent living. The Necesity system is designed to detect falls or other domestic incidents involving elderly people who live alone in their own houses. The system aims to reduce the long waiting times before a person is finally attended after suffering a fall or a faint. Additionally, it avoids this incident passing unnoticed. It is able to detect automatically and non-intrusively the situations while preserving subject's privacy.

The Necesity system is in charge of continuously monitoring elderly people who live alone and healthy in their own houses. Such monitoring is based on a wireless sensor network deployed at home. Sensors connected to the network capture presence, pressure in armchairs and open door events. All the events are sent to a miniPC (i.e. the size of a small laptop) which takes decisions about the state of the subject (i.e. the elderly person being monitored). When a problem is detected (e.g. long periods of inactivity in a context in which this should not happen according to patterns of behaviour learnt from the subject) an emergency actuation protocol is raised (usually an alarm is sent to an emergency response team through 3G communication). Such patterns of behaviour are key here. One of the most important elements of Necesity is the adaptation algorithm used to learn from user habits. This algorithm is used in a service in charge of quickly detecting when an unexpected situation is happening at the same time that false positives (i.e. false alarms) are minimized and false negatives (i.e. the algorithm does not detect an undesired situation which is occurring) are avoided. This service is the SUT considered in this paper. In order to test this algorithm

[4]Necesity website: http://www.necesity.net/

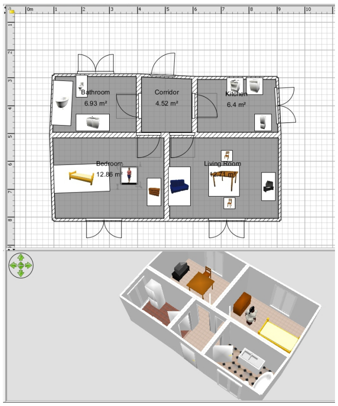


Fig. 4. A plane model in UbikSim's editor and its 3D representation

over the long term, some mechanism to generate artificial sensor data is needed (i.e. real data is scarce and it usually takes a long time to get a data series which is long enough). Thus, the idea is to create a SS model for general subjects. A model which could be tailored to manifest different habits to differentiate between subjects. This SS model is the SMUT that will be validated in the running example by applying the methodology.

Let us start illustrating the application of the methodology by explaining how to create SS models for elements 2.1 and 2.3 of the methodology. After the creation and validation of these elements, they are incorporated as validated SMUTs that are used, in subsequent iterations of the methodology, to validate the SMUT regarding subject's behaviour. The creation of such models is based on using the UbikSim editor (see bundle B1 in Figure 3) [53]. A snapshot of the editor, with a simple apartment already created appears in Figure 4 (it is based on SweetHome3D editor and adapted for the particular domain of ubiquitous computing and AmI).

The editor includes basic furniture (i.e. passive objects) and active objects (i.e. sensors in this case, which include presence, open door and pressure among others). The Figure shows the physical model of a simple house with a kitchen, a bathroom, a bedroom and a living room. A 2D view edited by the user appears in the upper part, and the corresponding 3D view, which is created on the fly, in the lower part. In the model, presence sensors are included in every room of the house. Furthermore, a sensor for open door (it is necessary to know when the elderly person leaves the home) is also included in the front door. Pressure sensors are installed in the bed and the subject's favourite armchair. They allow the system to know when the subject is resting. Simulated sensors are connected to the middleware (i.e. see bundle B4 of the architecture of UbikSim) by actually using the bundle's API. Thus, the process of sending events from sensors, in simulation time, is exactly the same as when the system actually runs in a real house.

4. Realistic Users' Behaviour Modelling

The kind of monitored subject under the Necesity system is a typical aged person, who lives independently and alone in his own house. In this scenario, a probable situation is that the person suffers some sudden health problem and stays immobilized on the floor, as a consequence, for a long time before anybody comes and notices that something is wrong. Detecting such situations and acting in consequence is the main task of the Necesity system. Moreover, this section introduces the application of the methodology to validate a SS model of the subject. This SMUT is used afterwards to validate, among other things, the mechanism in charge of detecting that something is wrong in the house by using the UbikSim tool.

Basic assumptions for the subjects are: (1) the simulated elderly person should be necessarily simulated 24 hours a day, repeatedly for a determined number of weeks; (2) the day is divided into time slots (i.e. morning, noon, afternoon and night); (3) the simulated subject behaves depending on the time slot the simulation is in.

The development of this SMUT follows the process defined in Figure 2. The *MonitoredSubject* class is created and added to the model's class hierarchy. The *MonitoredSubject* class must be declared as an agent: it represents an object that is going to evolve along the simulation. Let us describe the SMUT's dynamics. In the approach followed in this work, the SMUT for a subject is modelled probabilistically. The basic idea is to identify simple behaviours (e.g. sleeping, eating, having a shower, cleaning hands, and so on). The simulation model should then decide, depending on the time of the day, what is the basic behaviour displayed by the subject at any time. Simple behaviours can be aggregated into higher abstraction level ones (e.g. having a shower and cleaning hands could be aggregated into a *having a wash*). Thus, a hierarchical automaton is a natural solution to represent possible transitions from one behaviour to others and hiding the unnecessary complexity depending on the granularity level needed. Figure 5 represents the highest abstraction level automaton in (a) with the corresponding Q set of states and transitions represented as arrows at the graph. At the (b) part, all the automata which are encapsulated into states in Q are represented. For example, when at state $a_2 \in Q$, the subject is having a meal. So Q_2 and the corresponding automaton represent that while eating, the subject can go to the toilet and come back to finish eating. A further refinement of the a_{20} state is represented by the automaton in (c). It represents a sequence of activities done to prepare meal before eating (this granularity level is not used in this paper as explained above).

Depending on how they occur, there are three different types of behaviours:

- *Monotonous behaviours*: the kind of behaviour the subject manifest always approximately in the same time slot, and on a daily basis (e.g. sleeping, having meals, medication and so on).
- *Non monotonous behaviours*: the kind of behaviour the subject usually manifests, not bounded by a concrete time slot, and repeated within a non constant period (e.g. going to the toilet, having a shower, cleaning the house and so on).
- *Anytime behaviours*: these behaviours are interrupted by others and represent the state of the subject when she is not doing anything of interest for the system. The subject displays this behaviour when not displaying any other, either monotonous or non monotonous (i.e. anytime behaviours represent spare time).

We have just introduced the interrelation of the subject's behaviours in terms of the switching between them. Such switching has the form of a hierarchical automaton. But a mechanism is needed to simulate when to switch and to what behaviour. For this purpose, a combination of an interpreter and probability distribution (pdf) functions is used. These pdfs are members of the exponential family [15,33].

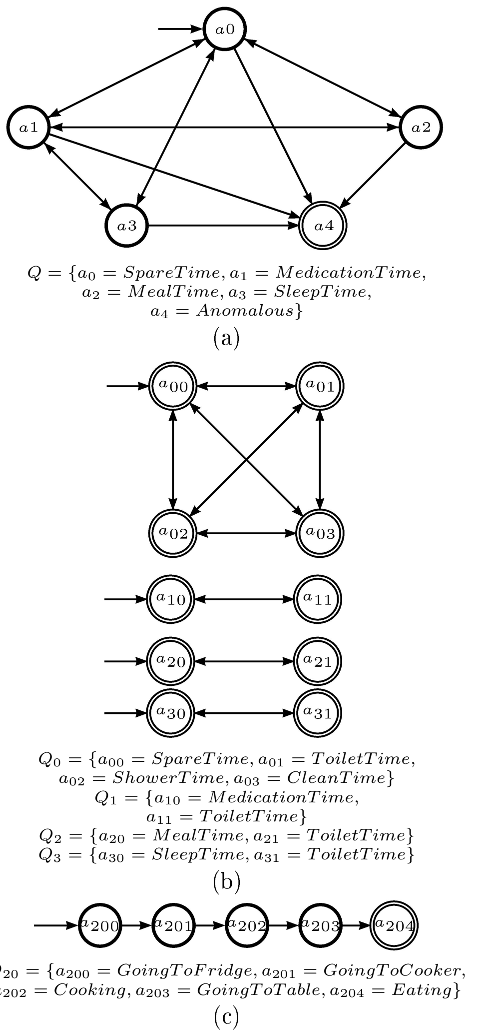


Fig. 5. (a) Level 0 automaton, (b) Level 1 automata, (c) Level 2 automaton for state a_{20}

Monotonous behaviours must be activated at specific time hours or within specific time intervals. For example, in the case of *MedicationTime* (see Figure 5(a)), a transition to this state will be generated exactly at the time of taking medicines. In the case of *MealTime* and *SleepTime*, the activation is generated within a time interval. Hence, the pdf which models these transitions is bounded to this time slot. Also, it must be assured that the simulated subject eats and sleeps every day. Because of that, if no transition is probabilistically generated within the time slot, it is necessarily generated at the end of the slot. Such time slots are similar for all subjects (i.e. a person usually goes to sleep or has dinner at the same hours). The gamma pdf is usually used as a probability model for waiting times, in this case, the time the interpreter has to wait to launch the behaviour. Thus, a gamma distribution may be used for modelling when monotonous behaviours occur. With the same reasoning, non monotonous behaviour's occurrence is simulated with an exponential pdf. Non monotonous behaviours are also waiting time models, but they are defined in awake periods. Actually, this is a special case of the gamma distribution. Finally, the group of anytime behaviours are behaviours which will be often interrupted by the other kind of behaviours. The time that it takes for such a behaviour to be displayed is small, due to interruptions. This causes a characteristic heavy-tailed pdf. A well known example of a pdf showing this appearance is the Pareto II or Lomax pdf. All these assumptions will be empirically demonstrated in section 5.

The relation between behaviours and the basic mechanisms used to simulate when to make the transition to a new behaviour are explained above. Now, the interpreter, which simulates the subject's dynamics, is introduced. Figure 6 presents a pseudocode for it. For this SMUT, only levels 0 and 1 of the automaton introduced in Figure 5 are used. Only such granularity of reality level is needed. First of all, behaviours of level 0 and 1 are initialized. For example, line number 15 initializes, for all anytime behaviour, the instants of time, from when the simulation starts until it ends, when they should be activated. The same goes for monotonous and non monotonous behaviours at lines 20 and 22, respectively. The case of monotonous behaviour, as mentioned above, is special: the pdf used to generate times for activation is bounded to the interval in which the behaviour should occur. At line 35, the automaton starts to execute. At each iteration, simulation time advances (line 61) and when it is equal to the time the active behaviour must end, the automaton returns to the initial behaviour (lines 37-39). At each iteration the priority of the first behaviour in the `pt` list (i.e. list of pending tasks ordered by priority) is compared with the priority of the active behaviour (i.e. `currentState`). If the priority is higher for that behaviour on the list, the active behaviour is interrupted and is put at the `pt` list. The behaviour with higher priority is then activated (lines 43-50). The next lines, from 51 to 64 are devoted to maintaining `pt`, adding new behaviours that should be activated now.

The interpreter can be customized using some configuration parameters. The first is the time limit in each room before entering an anomalous state (t_{max}), i.e. if the subject stays more than two hours in the bath, an alarm should be raised, thus, the automata should go to *Anomalous*. The second parameter is, obviously, the pdf used for each kind of behaviour. The third and last one is the set of time slots of monotonous behaviours.

5. Forensic Analysis of Users' Behaviour

At this point, the SMUT is designed and implemented. Following the methodology, the model is already designed in Step 3, and implemented in Step 4. This section is devoted to verification and subsequent validation of the SMUT.

5.1. Verification

The verification of SMUT ensures that the implementation of these models is correct. The SMUT has been verified by using the unit test package JUnit[5]. JUnit is open source and intensively used in the open-source community. This package allows developers to create unit tests which, once written, are performed automatically, giving an error report in response.

An example of unit test is the verification of the transitions in the automata. Figure 7 shows the JUnit code for the automaton of level 2 shown in Figure 5.

Furthermore, UbikSim contributes to verification by means of the displays. The simulation displays are very useful to build the model and observe that the artificial society behaviour is appropriate and realistic. At the same time, it can help to detect some simple simulation errors. UbikSim works on MASON and can use its features, for example inspectors. They are a means of graphically visualizing the evolution of variables of interest for the simulation. A large number of inspectors for various simulation variables can be used and monitored dynamically as the simulation evolves. They can be used to check that such variables always take reasonable values. Additionally, UbikSim shows a 3D display of the physical environment, devices installed, furniture and other physical elements and simulated humans. This visualization allows designers to evaluate some simple features of the system. Examples of such features are the coherence of the environment (a door must be in a wall) or some behaviour of the users (a user only can move through free areas, i.e. without obstacles).

[5] JUnit website: `http://www.junit.org/`

5.2. Validation

The SMUT in the running example is a model of the elderly person. This model refers to a realistic behaviour in terms of where the user is now, towards where he will move and when he will move. Eventually, he can sit or lie down on the bed. Why this level of abstraction? Why not model the subject cooking or reading a book? The reason is simple: the Necesity system can not perceive such granularity of activities. It only perceives presence, pressure on beds or armchairs and open or closed doors. As a consequence we can say that the granularity of behaviours to simulate will be given by the activities the system can perceive. Thus, the SMUT should be validated at the same granularity level.

There are many techniques for validating simulations [34,58] and more specifically, for validating simulated models based on agents [50]. In this work, we start with fragments of real data obtained from Necesity logs, thanks to a pilot experiment with an early release of the Necesity system. This pilot experience involved about 25 elderly people living independently in their own houses. The basic idea consists of comparing real behaviour data with synthetic data generated by simulations of the SMUT. The rest of the section explains how this idea was enabled by two phases. The first is in charge of processing real data to make it ready for SMUT validation. The second phase, model diagnosis, assesses whether that the proposed SMUT resembles real data.

5.2.1. Data Preprocessing

For validation of the SMUT, we selected data from three subjects of the pilot experiment mentioned above. These data refer to monitoring the

```
Let pt be a list of pending tasks ordered by priority
Let states be a list with all posible automata's states
Let times be a list of time instants when transitions
are going to be generated
Distribution Functions to model the ocurrence
of monotonous behaviours:
Let mb be the function to model monotonous behaviours
Let nb be the function to model non monotonous behaviours
Let ab be the function to model anytime behaviours

//Instants of time initialization

for all s in states do
  if isAnytime(s) then
      times(s)<-ab()
  else if isMonotonous(s) then
    //Initial and final instants of bounded time slot
    i <- ini(s)
    e <- end(s)
    times(s)<-mb(i, e)
  else if isNonMonotonous(s) then
    times(s)<-nb()
  endif
endfor

currentTime <- 0
//current state is defined by 3 numbers, one per level
level0 <- 0
level1 <- 0
level2 <- 0
currentState<-newState(level0,level1,level2)
//Automata begin to move when the times are initialized

//If an anomalous state is reached, the execution stops
while(level0(currentState)<>4)
  //If current task ends, initial state is activated
  if timeLeft(currentState)=0 then
    currentState <- newState(0,0,0)
  endif
  //If next state has higher priority than current state,
  //current state becomes a pending task and it is stored
  //in pt
  if (size(pt))>0
    nextState <- first(pt)
    if priority(nextState)>priority(currentState) then
      add(pt,currentState)
      currentState<-nextState
      remove(pt,nextState)
    endif
  endif
  for all s in states do
    time <- first(times(s))
    //If current time matches first time instant in
    //times, the task associated with this time instant
    //is added to pending tasks
    if currentTime=time then
      remove(times(s),time)
      add(pt,s,priority(s))
    endif
  endfor
  currentTime <- currentTime + 1
  //Decrement remaining time for finishing current task
  time <- timeLeft(currentState) - 1
  setTimeLeft(currentState, time)
endwhile
```

Fig. 6. Pseudocode for the abstract machine interpreter which simulates subject's behaviour

```
@Test
public void mealTimeLevel2Transitions() {
  actualState <- newState(2,0,0); //first state of meal time in level 2
  for(int i=1;i<=4;i++){
    currentState.nexState();//generating transition
    //for a201 the index in level 0 is 2, index in level 1 is 0, and so on.
    int j = currentState.getStateIndexForLevel(2);
    String alertMessage="The transition from" + currentState.toString() + "should be to state a20" + i;
    assertEquals(alertMessage,i,j); //check if transition is right
  }
}
```

Fig. 7. JUnit code for the automaton of level 2

elderly people over two months. Source data is in the form of a log file. This log shows in which room of the house the subject is located (including also if the subject is leaving the house, seated or sleeping).

In this approach, validating the SMUT means assuring that the right probability distribution function is used to reproduce the transition between the different states of the automation (i.e. behaviours). Thus, preprocessing log data is transforming it into a data series which consists of the instants of time in which transition between behaviours take place. This data series is then compared with a similar data series obtained by means of SMUT simulations in the same conditions. Comparison is made by a goodness of fit statistical test.

In the case of the non monotonous behaviours and anytime behaviours, preprocessing the log data involves the extraction of the time series of the moments in which each event is produced. These time series only include values within the typical awake period of the corresponding user. Obviously, while the user is sleeping, his daytime routines change. The log treatment for monotonous behaviours is slightly different. This kind of behaviour is usually produced in a bounded time slots. For example, having dinner, having lunch or sleeping are behaviours which occur during specific daily time periods. So, the preprocessing task involves extracting behaviour events inside such time slots. Note that time slots can be slightly different for each elderly person, but it is possible to define an approximation of them which will be valid for all (section 4 shows the configuration parameters). Finally, time intervals between each event are measured. The extracted time series are composed of these time intervals.

5.2.2. Model Diagnosis

SS models which describe social processes, like the model proposed here, are generally hard to validate. In this approach, the behaviour is probabilistically modelled. A good approach to assess that the model explains real data with a reasonable accuracy is to use a statistical test. This process is called model diagnosis. This section explains the model diagnosis made on the SMUT. In this context, the most serious problem we can find is the lack of real data [32]. Fortunately, as explained above, data is available from three real subjects.

From these preprocessed data, some histograms for different behaviours and people are shown in Figure 8. The sample pdf is shown with a solid line. The dashed line shows the pdf of the theoretical distribution (see section 4) which best fits that behaviour.

Graphs (a) and (b) in Figure 8 show two monotonous behaviours: sleeping and having dinner (having lunch is roughly similar). This kind of behaviour is modelled with a gamma distribution in this approach because it is a suitable distribution as a probability model for waiting times. Figure 8 (c) shows a non monotonous behaviour; going to the toilet. This behaviour is fitted with an exponential distribution in this work, as a special case of the gamma distribution. Finally, Figure 8 (d) shows an anytime behaviour; spare time. This behaviour is often interrupted by other behaviours (e.g. going to the toilet). These interruptions make the probability density function of the sample heavy-tailed and fitted by the Lomax distribution. At this point, it is already possible to visually check that the sample density is similar to a theoretical density for all the four graphs. However, more statistical evidence supporting the use of gamma, exponential and Lomax distributions for monotonous, non monotonous and anytime behaviours respectively is given now.

Table 1
P-values for each subject used in the study and the corresponding behaviours.

	Behaviour				
Person	Sleep	Dinner	Eat	Toilet	Spare time
A	0.404	0.311	0.361	0.111	0.086
B	0.488	0.467	0.542	0.079	0.108
C	0.337	0.489	0.575	0.137	0.103

It is possible to estimate the distance between a time series generated by a sample (i.e. real data) and over generated by simulation of theoretical distributions (i.e. theoretical distributions which are used in the SMUTs). One of the most well-known statistical tests to be applied in this case is the Kolmogorov-Smirnov (K-S) test [42]. The K-S test is a nonparametric and distribution-free goodness-of-fit test. It does not rely on parameter estimation or precise distributional assumptions [54]. Considering that the SMUT, as designed, does not assume any concrete probability distribution and does not require parameter estimation, then the K-S test is suitable here. Note that, although both K-S and the χ^2 tests are the most commonly used and for large size sample both tests have the same power, the χ^2 test requires a sufficient sample size to obtain a valid chi-square approximation [39,17]. The K-S is a goodness-of-fit test to indicate whether it is reasonable or not to assume that a random sample comes from a specific distribution. It is a form of hypothesis testing where the null hypothesis says that sample data follow the stated distribution. The hypothesis regarding the distributional form is rejected if the test statistic, D_n, is greater than the critical value (i.e. if the p-value is lower than the significance level). The significance level is fixed in this work at 0.05, which is the value usually given in statistical literature.

Table 1 shows the p-values obtained from the K-S test for each validated behaviour with the adequate distribution. The null hypothesis is that the behaviour sample data come from the stated distribution and it is rejected if p-value is lower than the significance level. From these results, none of the stated null hypothesis can be rejected. Therefore behaviour of the three considered subjects (i.e. A, B and C) can be fitted by the specified pdfs, as described above.

From the table, all simulated behaviours statistically occur in a similar manner to real behaviours from the three subjects. Thus, the SMUT

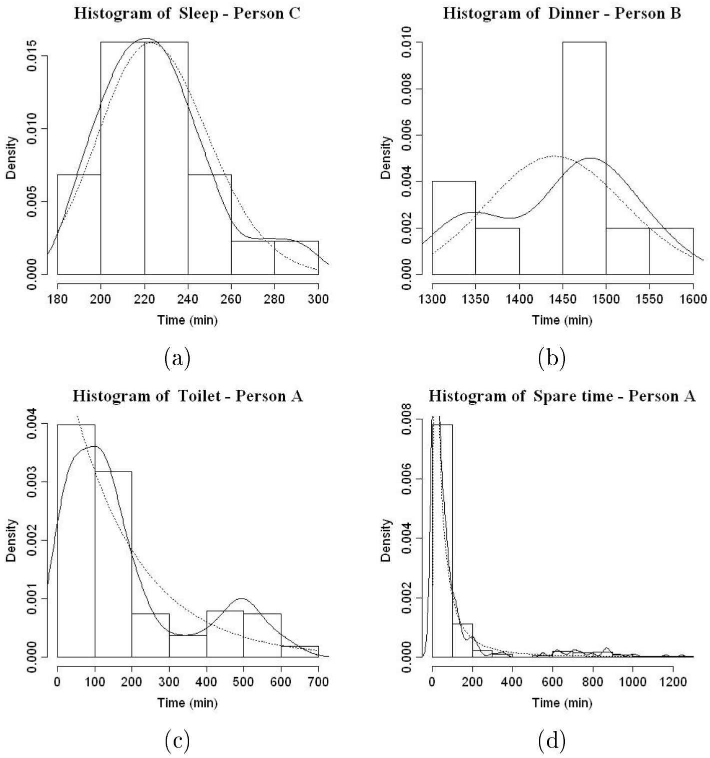


Fig. 8. (a) Minutes between 21:00 hours and the instant subject C goes to bed, (b) Time between dinners for subject B, (c) Time between uses of the toilet for subject A, (d) Time between spare time for subject A

is validated. By validating this SMUT, the UbikSim tool is ready to work on the validation of SUTs which need such subject behaviour to comply with the functionality required. Specific examples of SUT development may be found in other papers [24,25,52].

6. Related Works

A number of approaches dealing with testing, validation and verification of AmI systems can be found in the literature. All of them are based on using support tools. A well-known tool is Ubiwise [8]. Ubiwise focuses on the use and analysis of environment models for ubiquitous computing systems. For this purpose, sensors and its communications can be defined. People are considered individually on a simulation engine based on Quake II, where real people, acting as players in a game, generate information about their own context, which is captured through simulated sensors. Multiple users can link up to the same server to create interactive ubiquitous computing scenarios. However, the virtual subjects are not autonomous, since they must be controlled by the users. The subjects interact with the environment in order to validate its deployed services.

TATUS [47] is another tool which allows experimentation of adaptive ubiquitous computing systems. It is based on a graphic engine called Half-life. In this case, the main novelty is that multiple SUT may be connected to the engine. The SUT is adaptive so it makes decisions about changing its behaviour in reaction to user's movements and behaviour. Other environmental factors such as network conditions, ambient noise or social setting can also be considered. In this case, some virtual subjects could behave autonomously due to simple AI scripts, although the tool is mainly focused on user-controlled characters and their interactions with the SUT.

Another interesting tool is UbiREAL [44]. It lets the users intuitively grasp how devices are controlled, depending on temporal variation of contexts in a virtual space. The main contribution of UbiREAL is that it simulates physical quantities (e.g. temperature or humidity) and includes a network simulator. This network simulator allows communication between virtual devices and real devices. This is a very important aspect since it is possible to inject reality into the simulations. The behaviour of the virtual subjects must be preconfigured (it is possible to define a route and some actions to perform). Those existing simulators only simulate behaviours manually or based on simple automaton. UbikSim contributes to the area with the introduction of autonomous virtual humans to the simulations.

Regarding engineering realistic human behaviours, various approaches have been proposed to create autonomous characters. The approach to behavioural autonomy presented in section 4 is based on the proposal presented by Garcia-Valverde et al. [25], where behaviours are defined as states in a hierarchical automata and transitions between behaviours are probabilistic. The proposed environment was an office building where the workers did usual routines in a day of work. The simulations obtained allow them to build AmI advanced services which are able to adapt to the movements of the agents within a building.

In approaches [57] and [9], the behavioural models described also use a hierarchical structure of finite state automata similar to the model described in [25]. In these cases, each behaviour of a behaviour sequence is called a behaviour cell. There is a behaviour entity with a finite state automaton composed of at least one behaviour cell at the top of the structure. An elementary behaviour is situated at the bottom of the hierarchical decomposition and encapsulates a specialized behaviour which directly controls one or more actions. The lack of both approaches is the absence of probability. In the work of Chittaro et al. [12], the behaviour sequences are modelled through probabilistic automata (Probabilistic Finite-State Machine, PFSMs). Probabilistic personality influence implies that one cannot fully predict how a character will react to a stimulus. Because of that, a probabilistic approach gives more realism to simulated behaviours. The work of Garcia-Valverde et al. [25] includes both ideas. They define a probabilistic hierarchical automaton to achieve realistic and complex behaviours.

An important feature added to our model is the idea of a list of prioritized events. This idea is based on Temime et al. [56], where human agents have a pending task list. At each time tick, agents of the users check their task list in order to outstanding tasks. Following this approach, the more important tasks must be done before the lower ones, as the list is ordered by priorities. In the paper presented, the same idea is applied to transitions in the automata. They are weighted with a constant priority value. Then, if a transition to a new state is generated but its priority is lower than the current state, this transition is stored in a pending task list. These processes which are based on priorities give more realism to human behaviours.

Other related approaches have been reviewed. For example, Arthur [6] develops agents that act and choose in the way actual humans do. The agents are represented using parameterised decision algorithms, and choose and calibrate these algorithms so that the agents' behaviour matches real human behaviour observed in the same decision context. For this purpose, they use a parameterised learning automaton with a vector of associated actions that can be weighted to choose actions over time the way humans would. The structure of the automata is similar to the one presented in [25] but this work is more focused on machine learning. The virtual humans take decisions and learn from them. For this purpose, a detailed study about reactions to different stimulus of real subjects must be done and this is not within the scope of this paper. Anastassakis et al. [5] present an approach where every character is provided with a small KBS (Knowledge-Based System) for intelligent reasoning. A reasoning system is also used by Noser et al. [45]. Such methods are very flexible, but defining the knowledge base is a complex and time-consuming task.

SS models have already been used for engineering AmI systems in the literature. Reynolds et al. [49] simulate sensors, actuators, and the environment in an initial work on the design of a generic simulation tool. Liu et al. [36] propose a scalable framework for testing mobile context-aware applications based on the use of a multi-agent system which models complex and dynamic user's behaviour. Martin et al. [37] propose a simulation which separates agents, environment and a context defined by variables and maps. UbikSim is more complete than the aforementioned approaches because (1) it covers environment, context-aware, users and adaptation models and not a subset of these parts of an AmI service; (2) the users model can easily be extended by adding levels or states to the hierarchical automaton; (3) it provides a methodological proposal which guides the process of validating an AmI service.

7. Conclusion and Future Works

This work proposes a general methodology for the validation of AmI services and applications. It is a contribution to the engineering process of ubiquitous systems in general and AmI systems in particular. It is based on the simulation paradigm. When a new AmI system is going to be built, the typical physical environment, the devices (i.e. sensors and actuators) to be used, kind of users, services and applications must be identified. Then, the environment, devices and users are modelled in the form of SMUTs and validated. This validation must be combined, by applying the methodology, with the validation of the corresponding services and applications (i.e. the SUTs). Validation is based on simulation. Simulation runs generate data, which must be analyzed by means of forensic analysis. This forensic analysis gives clues about the maturity of the software component being validated. Also it is interesting to note that, if the SMUT is too complex as a model, a divide-and-conquer strategy is used to divide it into simple SMUTs in which the methodology is applied again.

We have shown the support tools which come with the methodology. UbikSim is a realistic environments simulator which can be used to test SMUTs and SUTs, to verify and validate them under the application of the methodology. The application of the methodology has been shown by means of a running example. This example was focused on the Necesity product and how to validate a particular SMUT. This validation can be used afterwards to validate SUTs (e.g. an adaptive algorithm to detect abnormal situations at the elderly people's home). The approach used to generate realistic user behaviour is shown in the paper along with how the corresponding simulation model of subjects is validated using a statistical test based on goodness of fit (i.e. the Kolmogorov-Smirnoff test).

The methodology and UbikSim tool are currently being applied in a number of projects related to ambient intelligence. Obviously, Necesity is a clear example. But other projects also use our proposal. For example, in CARONTE[6], the methodology is being used to provide assistance in emergency situations which may occur in geriatrics. CARDINEA[7] develops AmI services for caregivers in hospitals. The approach of this paper is being used for verification and validation of services. Note that in these projects the scenarios are geriatrics or hospitals, i.e. multi-user environments. In fact, we are working on the model of users in these kinds of environments where the methodology and users' model are still suitable, taking into account the relations and interactions between the agents of the simulation. This is possible thanks to the use of social simulations. Although more experiments are still needed, preliminary tests have shown that our approach is able to model the behaviour of the users in multi-inhabitant environments.

Future works also include a deep study of source data. Incorporating data from all the available subjects into a deeper study would help to create a taxonomy of elderly people's behaviours (i.e. kind of mobility, habits and so on). Such a taxonomy would be useful for an automatic parameter tuning of the models of elderly people. The user of the simulator, instead of configuring parameters by hand, would simply choose between a catalogue of elderly people's patterns of behaviour.

Acknowledgements

This research work is supported by the Spanish Ministry of Science and Innovation under grants AP2007-04269 and AP2007-04080 of the FPU program and in the scope of the Research Projects TSI-020302-2009-43 and TSI-020302-2010-129, through the Fundación Séneca within the Program "Generación del Conocimiento Cien-

[6]CARONTE website:
`http://caronte.germinus.com/inicio`
[7]CARDINEA website:
`http://cardinea.grupogesfor.com/home`

tífico de Excelencia" (04552/GERM/06) and the Research Project TIN2011-28335-C02-02 "Fundamentos para el Desarrollo de Servicios y Aplicaciones AAL".

References

[1] E. Aarts and B. de Ruyter. New research perspectives on ambient intelligence. *Journal of Ambient Intelligence and Smart Environments, IOS Press*, 1(1):5–14, 2009.

[2] E. Aarts and J. L. Encarnacao. True visions: Tales on the realization of ambient intelligence. In *Into Ambient Intelligence, Chapter 1*. Springer Verlag, Berlin, Heidelberg, New York, 2005.

[3] W. R. Adrion, M. A. Branstad, and J. C. Cherniavsky. Validation, verification, and testing of computer software. *ACM Comput. Surv.*, 14(2):159–192, 1982.

[4] OSGI Alliance. *Osgi service platform, release 3*. IOS Press, Inc., 2003.

[5] G. Anastassakis, T. Panayiotopoulos, and T. Ritchings. Virtual agent societies with the mVITAL intelligent agent system. In *Intelligent Virtual Agents*, pages 112–125. Springer, 2001.

[6] W. B. Arthur. On designing economic agents that behave like human agents. *Journal of Evolutionary Economics*, 3(1):1–22, 1993.

[7] R. Axelrod. Advancing the art of simulation in the social sciences. *Complex.*, 3(2):16–22, 1997.

[8] J. Barton and V. Vijayaraghavan. Ubiwise: A ubiquitous wireless infrastructure simulation environment. *HP Labs*, 2002.

[9] P. Bécheiraz and D. Thalmann. A behavioral animation system for autonomous actors personified by emotions. In *Proceedings of the 1998 Workshop on Embodied Conversational Characters*, 1998.

[10] J. A. Botía, A. Villa, J. T. Palma, D. Pérez, and E. Iborra. Detecting domestic problems of elderly people: simple and unobstrusive sensors to generate the context of the attended. In *First International Workshop on Ambient Assisted Living, IWAAL*, Salamanca, Spain, 2009.

[11] F. Cardinaux, D. Bhowmik, C. Abhayaratne, and M. S. Hawley. Video based technology for ambient assisted living: A review of the literature. *Journal of Ambient Intelligence and Smart Environments, IOS Press*, 3(3):253–269, 2011.

[12] L. Chittaro and M. Serra. Behavioral programming of autonomous characters based on probabilistic automata and personality. *Computer Animation and Virtual Worlds*, 15(34):319–326, 2004.

[13] D. J. Cook, J. C. Augusto, and V. R. Jakkula. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277–298, 2009.

[14] D. J. Cook, M. Youngblood, E. O. Heierman III, K. Gopalratnam, S. Rao, A. Litvin, and F. Khawaja. Mavhome: An agent-based smart home. In *Pervasive Computing and Communications, 2003.(PerCom 2003). Proceedings of the First IEEE International Conference on*, pages 521–524. IEEE, 2003.

[15] G. Darmois. Sur les lois de probabilit a estimation exhaustive. *CR Acad. Sci. Paris*, 260:1265–1266, 1935.

[16] S. Dashtinezhad, T. Nadeem, B. Dorohonceanu, C. Borcea, P. Kang, and L. Iftode. Trafficview: a driver assistant device for traffic monitoring based on car-to-car communication. In *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, volume 5, pages 2946–2950. IEEE, 2004.

[17] F. N. David and N. L. Johnson. The probability integral transformation when parameters are estimated from the sample. *Biometrika*, 35(1-2):182, 1948.

[18] P. Davidsson. Agent based social simulation: A computer science view. *J. Artificial Societies and Social Simulation*, 5(1), 2002.

[19] F. Doctor, H. Hagras, and V. Callaghan. A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 35(1):55–65, 2005.

[20] A. Drogoul, D. Vanbergue, and T. Meurisse. Multiagent based simulation: Where are the agents? In Jaime Simão Sichman, François Bousquet, and Paul Davidsson, editors, *MABS*, volume 2581 of *Lecture Notes in Computer Science*, pages 1–15. Springer, 2002.

[21] S. Ferguson. How computers make our kids stupid. *Maclean's*, 118(23):24–30, June 2005.

[22] P. A. Fishwick. *Simulation model design and execution : building digital worlds*. Prentice Hall international series in industrial and systems engineering. Prentice Hall, Englewood Cliffs, NJ, 1995.

[23] P. A. Fishwick. Computer simulation: growth through extension. *Trans. Soc. Comput. Simul. Int.*, 14:13–23, March 1997.

[24] T. Garcia-Valverde, A. Garcia-Sola, and J. A. Botia. Improving RFID's location based services by means of hidden markov models. *European Conference of Artificial Intelligence (ECAI' 10). Lisbon, Porto.*, pages 4–7, 2010.

[25] T. Garcia-Valverde, A. Garcia-Sola, F. Lopez-Marmol, and J. A. Botia. Engineering Ambient Intelligence Services by Means of MABS. *Trends in Practical Applications of Agents and Multiagent Systems*, pages 37–44, 2010.

[26] T. Garcia-Valverde, E. Serrano, J. A. Botia, A. F. Gomez-Skarmeta, and J.M. Cadenas. Social simulation to simulate societies of users inmersed in an ambient intelligence environment. *1st IJCAI Workshop on Social Simulation (SS@IJCAI2009)*, 2009.

[27] N. Gilbert and K. G. Troitzsch. *Simulation for the Social Scientist*. Open University Press, February 2005.

[28] IBM. Flowcharting techniques. *IBM GC20-8152-1 edition*, 1969.

[29] S. S. Intille. Designing a home of the future. *Pervasive Computing, IEEE*, 1(2):76–82, 2002.

[30] T. Ishida, Y. Nakajima, Y. Murakami, and H. Nakanishi. Augmented experiment: participatory design with multiagent simulation. In *IJCAI'07: Proceedings of the 20th international joint conference on Artifical in-*

telligence, pages 1341–1346, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.

[31] C. Kidd, R. Orr, G. Abowd, C. Atkeson, I. Essa, B. MacIntyre, E. Mynatt, T. Starner, and W. Newstetter. The aware home: A living laboratory for ubiquitous computing research. *Cooperative buildings. Integrating information, organizations, and architecture*, pages 191–198, 1999.

[32] F. Klügl. A validation methodology for agent-based simulations. In *Proceedings of the 2008 ACM symposium on Applied computing*, pages 39–43. ACM, 2008.

[33] B. O. Koopman. On distributions admitting a sufficient statistic. *Transactions of the American Mathematical Society*, pages 399–409, 1936.

[34] A.M. Law, W.D. Kelton, and W.D. Kelton. *Simulation modeling and analysis*. McGraw-Hill New York, 1991.

[35] X. Li, W. Mao, D. Zeng, and F. Wang. Agent-based social simulation and modeling in social computing. In *Proceedings of the IEEE ISI 2008 PAISI, PACCF, and SOCO international workshops on Intelligence and Security Informatics*, PAISI, PACCF and SOCO '08, pages 401–412, Berlin, Heidelberg, 2008. Springer-Verlag.

[36] Y. Liu, M. O'Grady, and G. O'Hare. Scalable Context Simulation for Mobile Applications. In *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, pages 1391–1400. Springer, 2006.

[37] M. Martin and P. Nurmi. A generic large scale simulator for ubiquitous computing. In *Mobile and Ubiquitous Systems: Networking & Services, 2006 Third Annual International Conference on*, pages 1–3. IEEE, 2007.

[38] S. Marzano. People as a source of breakthrough innovation. *Design Management Review*, 16(2):23–29, 2005.

[39] F. J. Massey Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[40] D. Midgley, R. Marks, and D. Kunchamwar. Building and assurance of agent-based models: An example and challenge to the field. *Journal of Business Research*, 60(8):884–893, August 2007.

[41] G. J. Myers, C. Sandler, T. Badgett, and T. M. Thomas. *The Art of Software Testing, Second Edition*. Wiley, June 2004.

[42] H.R. Neave and P.L. Worthington. *Distribution-free tests*. Routledge London, 1989.

[43] I. Nieto, J. A. Botía, and A. F. Gómez-Skarmeta. Information and hybrid architecture model of the OCP contextual information management system. *Journal of Universal Computer Science*, 12(3):357–366, 2006.

[44] H. Nishikawa, S. Yamamoto, M. Tamai, K. Nishigaki, T. Kitani, N. Shibata, K. Yasumoto, and M. Ito. UbiREAL: Realistic smartspace simulator for systematic testing. *UbiComp 2006: Ubiquitous Computing*, pages 459–476, 2006.

[45] H. Noser and D. Thalmann. Towards autonomous synthetic actors. *Synthetic Worlds. TL Kunii and A. Luciani, John Wiley and Sons, Ltd*, 1995.

[46] The Institute of Electrical and Electronics Engineers. *IEEE 90: IEEE Standard Glossary of Software Engineering Terminology*. IEEE, 1990.

[47] E. O'Neill, M. Klepal, D. Lewis, T. O'Donnell, D. O'Sullivan, and D. Pesch. A testbed for evaluating human interaction with ubiquitous computing environments. In *Testbeds and Research Infrastructures for the Development of Networks and Communities, 2005. Tridentcom 2005. First International Conference on*, pages 60–69. IEEE, 2005.

[48] A. Pentland. Perceptual environments. In D. J. Cook and S. K. Das, editors, *Smart Environments: Technology, Protocols and Applications.*, pages 345–359. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2005.

[49] V. Reynolds, V. Cahill, and A. Senart. Requirements for an ubiquitous computing simulation and emulation environment. In *Proceedings of the first international conference on Integrated internet ad hoc and sensor networks*, page 1. ACM, 2006.

[50] M. Richiardi, R. Leombruni, N. Saam, and M. Sonnessa. A common protocol for agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, 9(1):15, 2006.

[51] L. Rudolph. Project oxygen: pervasive, human-centric computing–an initial experience. In *Advanced Information Systems Engineering*, pages 1–12. Springer, 2001.

[52] E. Serrano and J. A. Botia. Validating ambient intelligence based ubiquitous computing systems by means of artificial societies. *Information Sciences*, November 2010. In press.

[53] E. Serrano, J. A. Botia, and J. M. Cadenas. Ubik: a multi-agent based simulator for ubiquitous computing applications. *Journal of Physical Agents*, 3(2):39, 2009.

[54] D. Sheskin. *Handbook of parametric and nonparametric statistical procedures*. CRC Pr I Llc, 2004.

[55] Y. Shi, W. Xie, G. Xu, R. Shi, E. Chen, Y. Mao, and F. Liu. The smart classroom: merging technologies for seamless tele-education. *Pervasive Computing, IEEE*, 2(2):47–55, 2003.

[56] L. Temime, Y. Pannet, L. Kardas, L. Opatowski, D. Guillemot, and P.Y. Boëlle. NOSOSIM: an agent-based model of pathogen circulation in a hospital ward. In *Proceedings of the 2009 Spring Simulation Multiconference*, pages 1–8. Society for Computer Simulation International, 2009.

[57] D. Thalmann, S.R. Musse, and M. Kallmann. Virtual Humans' Behaviour: Individuals, Groups, and Crowds. *Proceedings of Digital Media Futures*, pages 13–15, 1999.

[58] K.G. Troitzsch. Validating simulation models. *Networked Simulations and Simulated Networks*, pages 265–270, 2004.

5.3 Improving RFID's Location Based Services by means of Hidden Markov Models

Title	Improving RFID's Location Based Services by means of Hidden Markov Models
Authors	Garcia-Valverde, T. and Garcia-Sola, A. and Botia, J.A.
Type	Conference
Conference	Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence
Relevance (2010)	Core Rank A+
Address	Lisbon, Porto
Publisher	IOS Press
Pages	1045–1046
Year	2010
Month	August
ISBN	978-1-60750-605-8
State	Published

Conference details: European Conference on Artificial Intelligence
Biennial European Conference on Artificial Intelligence (ECAI)
Leading conference in the field of Artificial Intelligence in Europe
Publisher: IOS Press
Core Rank A+ in the ERA Conference Ranking
Website 2010: `http://ecai2010.appia.pt/`

Authors – Personal details	
Name	**María Teresa Garcia Valverde**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Alberto García Sola**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Juan Antonio Botía Blaya**
Position	Lecturer of the Department of Information and Communications Engineering
University/ Organization	University of Murcia

Contribution of the PhD student
The PhD student, María Teresa García Valverde, declares to be the main author and the major contributor of the paper *Improving RFID's Location Based Services by means of Hidden Markov Models*

Improving RFID's Location Based Services by means of Hidden Markov Models

Teresa Garcia-Valverde and **Alberto Garcia-Sola** and **Juan A. Botia**[1]

Abstract.
Services in Ambient Intelligence systems must adapt to context information and users. So, the development of these services become in a complex task. In this context, we present AmISim: a methodology for the engineering of AmI services. Using AmISim, a LBS based on RFID technology is built. Engineering the service is based on Hidden Markov Models (HMMs) for locations within an intelligent building. The results show that the resolution of a RFID based LBS can be significantly improved and allows by means of the simulator optimally configure the number and location of RFID antennas.

1 Introduction

Ambient Intelligence (AmI) services acquire context knowledge from the environment allowing us personalize and adapt to different users and situations. Besides, to predict contexts allows us to provide proactive services and improve the accuracy of the system. So, the development of AmI systems is a difficult task. To resolve that problem, the approach proposed here is based on AmISim [4], a MABS (Multi-Agent Based Social Simulation) [1] simulator.

Specifically, in this work, the engineering of an RFID (Radio Frequency Identification) Location-Based Service (LBS) is presented for an intelligent building. The work rest on the assumption that users usually have some degree of regularity in their movements through the environment. Under such basis, the system may predict their motion and offer such prediction by a LBS which adapt to particular users. With such approach, a probabilistic framework based on HMMs to location prediction and appropriately deploying all the RFID apparatus are proposed. So, we may improve the system in two ways: 1) obtaining the best location and to trying to optimize the number of antennas and 2) it could be possible, probabilistically, to augment the resolution of the location system.

The rest of the paper is structured as follows. Section 2 describes the methodology for model learning and service building based on AmISim. Section 3 presents the predictive location model with HMMs. Section 4 gives empirical results. Finally, in section 5 some conclusions and future works are given.

2 Model learning and service building

In this approach, we use AmISim, a simulator of entire AmI environment which uses Ubik. Ubik is the simulated part of AmISim. Ubik is a MABS which has capabilities to simulate and to obtain a set of simulated data according to the real environment. The Ubik simulator is integrated with the real part of AmISim, an entire framework which is able to use semantic technologies for modelling the context model and uses it to provide adaptive and intelligent services.

AmISim is used here in a scenario in which a service can predict workers' locations in an office building. So, we can built the prediction framework based on HMMs and use the set of obtained data from Ubik for training and validating our model. When the HMMs are checked, we can deploy the services over a real environment.

An abstract model of the space in the building is used here. In the abstract model, an undirected graph is compound from the building, where the edges of the graph are doors and crosses, and the nodes are rooms, corridors and stairs. This way, the RFID antennas are situated in the nodes of the graph and each user wears an identification card with a RFID label which identifies his location in a node. In the simulation environment the user's location is known at each moment of time, however this assumption is not true in a real environment.

3 Predictive location model: HMMs

The above section leads us to deduce that an approximation for each not sensorized area is needed. So, a probabilistic framework based on HMMs to built predictive location models is proposed. A HMM for each possible not sensorized destination is modelled and implemented using JAHMM [3], a Java implementation of HMMs.

In these location prediction models, the hidden states model the different possible locations according to the followed path by a worker since the origin to the destination. It is not necessary a direct correspondence between hidden states of the HMMs and nodes of the graph. The observable states are defined by the set the sensorized nodes and they compound sequences of active RFID antennas in sensorized nodes through the worker's trajectory.

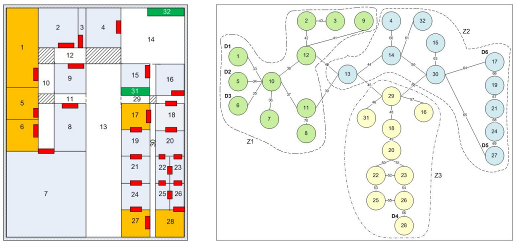
Ubik is used to generate the movements of the workers that allow us training and validating the models. Training sets for the HMMs are built by using sequences of active RFID antennas as generated through worker following some path, called traces τ. τ ends when the worker reaches his destination. A node is a destination when the worker spends there more than a threshold for such location. Note that each τ is built from the worker's path starting from his desk (worker's basic location). The traces are classified by their destination and each set of them is a training set of a HMM.

HMM parameters are estimated by using the Baum-Welch algorithm[2]. Thus, learning the HMMs consists on finding the maximum-likelihood estimate of the parameters of the HMM given a group of training samples for a single destination. When the models have been trained, the probability of an observation sequence can be estimated by the forward-backward algorithm[5]. Given a trace, the HMM with the higher probability for such trace allows us to predict the worker's location.

[1] University of Murcia, Spain, email: {mtgarcia, agarciasola, juanbot}@um.es

4 Experiments

The scenario of the experiments is showed in figure 1 a). Figure 1 a) represents a floor of a building and figure 1 b) is the graph built from the building, where the nodes are numbered and the edges, doors and crossings, are represented in red and lined, respectively.



a) b)

Figure 1. a) Floor of the building b) Graph of the floor

Destinations are represented by label D_i in Fig. 1 b). In the experiments the building is divided into zones: Z1, Z2 and Z3. The division into zones shows the fact that in a real environment the workers of the zone Zi usually will go to destinations in this zone, the nearest destinations. D1, D4 and D5 are relaxation rooms of the zones 1, 2 and 3 respectively. D2 is a meeting room and, D3 and D6 are warehouses. Each destination is modelled with a HMM.

In order to determine how the decrease of antennas affects the prediction capabilities of the system, in each test, the number of antennas is decreased. Firstly, the antennas of the rooms from each zone are removed. Then, the antennas of the corridors are taken because, as it will be seen in the results, the corridors give more information to the prediction.

The generated HMMs are able to describe all observation sequences of the training set (despite their lengths) and are robust with respect to an unbalanced training data distribution. This feature allows the system a better adaptation to changes in the workers' pattern behaviours, which often occurs in dynamic scenarios like this.

How the number of antennas decreasing influences the precision of the technique allows to decide the minimum number of antennas in order to have an acceptable quality. The results are showed in figure 2: a) represents error rate evolution segregated by destination (y axis) when the number of antennas decreases (x axis) and b) represents averaged error rate (left scale) and averaged precision (right scale).

It is clear from the results that there are dependent destinations. For example, D_1, D_2, D_3 are dependent because they share an important part of the graph structure. So, when the service does not predict D_1 when it should, is that it predicts either D_2 or D_3 (see the corresponding error rates in Fig.2 a)).

The decreasing the number of antennas does not affect until approximately a reduction of 23% because the removed antennas have little presence in traces. The same evolution, but inverse, is showed for the precision. As a conclusion, it can be said that removing antennas representing starting points of traces is affordable. However, corridors (i.e. connecting paths of the graph) should not be removed.

5 Conclusions

This work shows how a LBS based on RFID technology is engineered without using any real set up. An AmI simulator is used to help in the initial engineering of such AmI system. It is illustrated how the simulator is used to generate training data. Such data is used to estimate HMM for any location uncovered by RFID antennas. Experimentation shows that this approach may help on deciding the number and the position of RFID antennas while maintaining an acceptable location power for workers in a building. From this work, one future goal is modelling working activities in order to construct a more detailed context of workers.

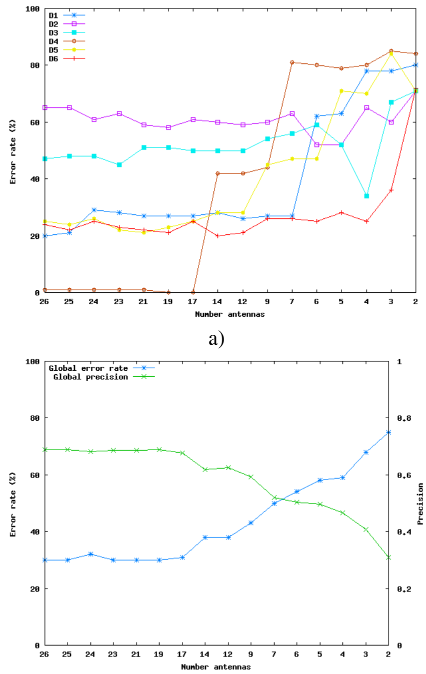


Figure 2. a) Error rates by destination. b) Averaged error rate/precision.

ACKNOWLEDGEMENTS

This research work is supported by the FPU grant AP2007-04269, the 09493/FPI/08 grant and in the scope of the Research Projects TSI-020302-2009-43, TIN2008-06441-C02-02, through the Fundación Séneca within the program 04552/GERM/06

REFERENCES

[1] Robert Axelrod, 'Advancing the art of simulation in the social sciences', *Complex.*, **3**(2), 16–22, (1997).

[2] L. Baum, 'An inequality and associated maximization technique in statistical estimation of probabilistic functions of Markov processes', *Inequalities*, **3**, 1–8, (1972).

[3] J.M. Francois, 'Jahmm-hidden markov model (hmm): An implementation in java', *URL: http://www. run. montefiore. ulg. ac. be/francois/software/jahmm*, (2006).

[4] T. Garcia-Valverde, E. Serrano, J.A. Botia, Gomez-Skarmeta A., and J.M. Cadenas, 'Social simulation to simulate societies of users inmersed in an ambient intelligence environment', *1st SS@IJCAI 2009*, (2009).

[5] L.R. Rabiner, 'A tutorial on hidden Markov models and selected applications in speech recognition', *Readings in speech recognition*, **53**(3), 267–296, (1990).

5.4 Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach

Title	Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach
Authors	Garcia-Valverde, T. and Garcia-Sola, A. and Botia, J.A. and Gomez-Skarmeta A.
Type	Journal
Journal	Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on
Impact factor (2010)	2.105
Pages	1-6
Year	2012
Month	April
Volume	PP
Issue	99
ISSN	1094-6977
DOI	`10.1109/TSMCC.2012.2187443`
URL	`http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6185688`
State	Published

Journal details: IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews
Editor-in-Chief: Vladimir Marik
ISSN: 1094-6977 (print version)
ISSN: 1558-2442 (electronic version)
Publisher: IEEE
Impact factor (2010): 2.105
Website: `http://www.ieeesmc.org/`

Authors – Personal details	
Name	**María Teresa Garcia Valverde**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Alberto García Sola**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Juan Antonio Botía Blaya**
Position	Lecturer of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Antonio Fernando Gómez Skarmeta**
Position	Professor of the Department of Information and Communications Engineering
University/ Organization	University of Murcia

Contribution of the PhD student
The PhD student, María Teresa García Valverde, declares to be the main author and the major contributor of the paper *Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach*

Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach

T. Garcia-Valverde, A. Garcia-Sola, J.A. Botia and A. Gomez-Skarmeta

***Abstract*—Services in Ambient Intelligence environments should adapt to contextual information (Context-aware), of users and environment in a non-intrusive and natural way. Location-aware, i.e. the user location, is one of the most important pieces in context-aware. According to this premise, a Location Based Service (LBS) using Radio Frequency Identification (RFID) technology is presented. The service is based on Hidden Markov Models for location within an intelligent building. This problem leads to a multiobjective optimization problem, in which, the best configuration of antennas that minimizes the set of antennas but maximizes the precision of the prediction should be found. Specifically, this work presents a Memetic approach for multiobjective improvement of LBS in AmI environments. The Memetic Algorithm provides in this problem the exploitation of domain knowledge and the combination of metaheuristics. Experimental results show that the approach obtains a configuration of antennas which optimally configures the number and position of the antennas while keeping a high quality of the precision of the location prediction.**

***Index Terms*—Ambient Intelligence (AmI), Hidden Markov Models (HMM), Location Based Services (LBS), Memetic algorithm, Multiobjective Optimization.**

I. Introduction

Context-aware is one of the key factors in Ambient Intelligence (AmI) environments [1]. An AmI environment needs to know the context in order to provide the ability to personalize and adapt to the users and their behaviours in different situations. According to this requirement, location information is one of the most essential contextual information in context-aware systems [2]. Location-awareness can provide nearest features of interest to users and adapt to them.

This work arises as needed from our previous work [3], in which a probabilistic framework based on Hidden Markov Models (HMMs) to predict location and appropriately deploy all the RFID (Radio Frequency Identification) apparatus is proposed. Then, the system can offer users' motion prediction by a RFID LBS which adapts to particular users. The work uses an AmI simulator, AmISim [4], in order to reproduce the whole AmI environment. The simulated data is used to estimate the HMM for each location uncovered in order to predict the users' destination.

Experimentation showed that this approach may help to optimally configure the number and the position of the RFID antennas while maintaining an acceptable location prediction for workers in a building. In fact, a reduction of the 23% of the antennas was made without a significant reduction in the quality of the prediction. However, the removal of antennas in corridors caused a significant reduction in the precision. From these results, it is clear, that the number of antennas used and their respective locations are deciding factors for the location prediction. Notice that the order of removal among rooms and among corridors was made intuitively. Because of that, emerges the necessity of finding the best configuration of antennas in the building, in which the set of antennas is minimized, but the precision of the prediction is maximized. Then, this problem implies two subproblems: the optimization of the set of antennas and the prediction using HMMs.

In this respect, evolutive algorithms have been used extensively to effectively address a wide range of this kind of problems [5]. There are several approaches regarding the use of genetic algorithms (GA) to solve optimization problems in ubiquitous computing environments, specially, in ubiquitous networks [6], [7], [8]. These works are focused on ubiquitous computing and they do not take in to account the user as main actor in the system, which is the main issue in AmI. On the other hand, an AmI environment is provided of knowledge from context. It is generally accepted that the incorporation or embedding of knowledge of the problem domain improves the performance of an optimization algorithm [9]. In such way, in this work, a memetic multiobjective algorithm is proposed to solve the presented problem.

Memetic algorithms (MAs) define an optimization paradigm based on the exploitation of problem knowledge and the combination of different metaheuristics [10], [11]. They use the notion of a meme defined as a unit of cultural evolution that is capable of local refinements [12]. MAs have shown remarkable success for solving optimization problems in NP-hard combinatorial problems and multi-objective memetic search [9]. Examples of MAs applied to real world problems can be found in the recent literature: network design problems [13], flow shop problems [14], robot control [15] or VLSI floorplanning [16].

In such sense, some hybrids or memetic algorithms have been proposed to resolve problems in ubiquitous networks [17], [18]. However, as it is described above, they do not take into account the requirements of the AmI environments.

The multi-objective memetic algorithm presented here is able to optimize the configuration of antennas in a building while maximizing the precision of the users' location prediction using HMMs. The algorithm takes domain information to successfully solve both objectives in an AmI environment.

The rest of the paper is structured as follows. Section II defines the problem in the appropriate context. The next section sets the proposed memetic multiobjective algorithm. Section IV gives empirical results of the approach. Finally, some conclusions and future work are given in Section V.

II. Multiobjective Optimization for the Antennas Location

As described above, finding the best configuration of antennas implies obtaining the minimum number of antennas maintaining a high precision of the prediction. So, this problem is stated as a multiobjective optimization problem (MOP) with two subobjectives: *maximizing the precision of the prediction by the HMMs* and *minimizing the number of allocated antennas*. A general MOP can be defined as follows:

$$Maximize\, f(x) = (f_1(x), f_2(x), ..., f_k(x)),\; x \in X \tag{1}$$

where $f(x)$ is the objective vector, $f_i(x)$ is the i-th objective to be maximized, x is the decision vector, and X is the feasible region in the decision space. In most of the MOPs there is no solution which causes an optimum in all objectives. A solution can not be improved on one objective without degrading other. Then, the concept of optimum is changed and a solution in which the objectives are satisfied in an acceptable factor is searched. In order to solve the problem of the conflicts among the objectives in MOPS, several techniques have been proposed, like the weighted-sum approach, ϵ-constraint method or evolutionary approaches [19], [20].

Department of Information and Communications Engineering, University of Murcia, Spain. E-mail:{mtgarcia,agarciasola,juanbot,skarmeta}@um.es

This research work is supported by the Spanish Ministry of Science and Innovation under the grants AP2007-04269 of the FPU program, the 09493/FPI/08 Séneca grant and in the scope of the Research Projects DIA++ (TRA2009_0141) and CARONTE (TSI-020302-2010-129).

Manuscript received December 31, 2010; revised January 27, 2012.

One of the classical and simplest approaches to solve a MOP is to translate the problem into a single objective problem by using the weighted-sum approach. According to this approach, in this work, a weight w_i is assigned to each normalized objective function. Given that, one objective in this problem must be minimized and the other must be maximized. For simplicity, the objective of minimizing the number of antennas is turned into one of maximizing. So, the objective function is defined by two maximizing objective functions, $f_1(x)$, *maximizing precision of HMMs*, and $f_2(x)$, *minimizing number of antennas*, as:

$$f(x) = w_1 f_1'(x) + w_2(1 - f_2'(x)) \tag{2}$$

where $f_1'(x)$ and $f_2'(x)$ are the normalized objective functions and w_1, w_2 are weights subject to the following conditions: $0 \leq w_1, w_2 \leq 1$ and $w_1 + w_2 = 1$.

In this work, the main goal is developing a location service using RFID technology. But, in this particular problem, there are areas uncovered by RFID antennas (less antennas means a considerably cheaper deployment). Thus, when the user is in one of such areas, the service must predict which one is most probable. HMMs are used for this prediction.

A HMM is a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden) [21]. In this problem, a HMM for each possible destination is built, where a destination is a non sensorized place which is interesting to predict. The hidden states represent the different possible locations according to the possible locations where any user might be located at (e.g. its office). The observable states are the active RFID antennas in the workers' trajectory (across the set of sensorized places). In each hidden state, observable information from the worker's trajectory is shown according to probability distributions.

In previous works, the problem of how a worker moves through an office building was addressed [3]. The main result of such works was a computational model of such displacements. This computation model was integrated in the AmISim simulator to reproduce the movements of the workers across their paths. The simulated paths allow the simulator to obtain sequences of active antennas (traces). So, it is possible to obtain a data set for training and validating the HMMs.

HMM parameters are estimated by using the classical Baum-Welch algorithm [22] and the probability of an observation sequence is estimated by the forward-backward algorithm [21]. Then, given a trace, the HMM with the highest probability for such trace predicts the worker's location. Finally, for the validation of the HMM, the validation set (1/3 randomly selected from the complete set) is used in order to calculate the precision of each HMM.

Precision depends on the number of antennas and their locations. Our previous work shown that reducing the number of antennas without getting worse the prediction is possible. But this reduction is only possible until a certain point and then, the precision starts to get worse. Furthermore, different configurations of the position of the antennas over the scenario obtain different precision values. In such way, the **first objective** for this problem is finding the configuration (number and position) of antennas which *maximizes the precision of the HMMs*. It is formulated as follows:

$$Maximize\, f_1(x) = \frac{\sum_{i=1}^{n}(p_{d_i})}{n} \tag{3}$$

where $D = \{d_1, d_2, \ldots, d_n\}$ is the set of uncovered destinations which should be predicted and p_{d_i} is the obtained precision by the HMM for the destination d_i.

Precision of the HMMs is determined by the number and location of these RFID antennas. This result implies two problems. On the one hand, getting the best configuration implies having a priori knowledge of the environment and its users in order to find the fittest configuration. On the other hand, it is not possible, in a real situation, to test different configurations over a real building with people who are working. Notice, besides, that the number of possible configurations is given by the factorial number of places where an antenna can be placed. As a matter of fact, this optimization problem is NP-complete.

In location based on RFID each worker carries a RFID tag and the building is equipped with RFID antennas in strategic places. When the worker passes near of an antenna, the antenna detects and identifies the worker. So, the worker's location is updated with the room in which the antenna is placed. An abstract model based on an undirected graph is used for the building model. Thus, each room, corridor or stair is a node of the graph and each door or crossing is an edge. Using this model, the movements of workers are simulated by the use of the best path between an origin and a destination using Dijkstra algorithm [23]. The RFID antennas can be situated in the nodes of the graph. An antenna in an edge is not useful because is a space of transition.

In such situation the workers' location is only known when they are inside a node with an antenna. Provided that it is not affordable yet unnecessary to place RFID arcs in every place, the location system lacks of the necessary information in non sensorized places.

This problem is defined as: given the total number of nodes where an antenna can be placed, select the minimal subset the antennas which allows us to obtain a quality of prediction as high as possible. In such way, in this problem $R = \{r_1, r_2, \ldots, r_t\}$ is the set of nodes in the building, i.e. all rooms, stairs and corridors and $U \subset R$ is the subset of these nodes where it is not allowed to placed an antenna, for example, a destination or a restricted area due to privacy reasons. In this work, it is considered that the only places which are not allowed to place an antenna are the destinations, i.e. $U = D = \{d_1, d_2, \ldots, d_n\}$.

In order to simplify the model, it is assumed that only one antenna can be placed in each node. So, the maximum number of antennas in a configuration is the size, $(t - n)$, of the set of nodes where an antenna can be placed. Whereas, the minimum number of antennas is 2. This restriction is set from the requirements of the HMMs, in which a configuration with less than two antennas does not allow the HMMs to be trained.

Then, let r_i to be a node $\in R$ and s_{r_i} the number of antennas in such node. In such way, the **second objective**, *minimizing the number of antennas*, can be formulated as follows:

$$Min.\, f_2(x) = \sum_{i=0}^{t-n} s_{r_i},\ subject\, to \begin{cases} 2 \leq \sum_{i=0}^{t-n} s_{r_i} \leq (t-n) \\ r_i \notin D \\ s_{r_i} = 0|1. \end{cases} \tag{4}$$

III. MEMETIC MULTIOBJECTIVE ALGORITHM

In an AmI system the interaction between users and environment is made in a non-intrusive and natural way, being aware of context. That is, the specific knowledge of the context is essential to provide intelligence and proactivity. So, it is clear that, an optimization algorithm could benefit from this knowledge. On the other hand, according to the *no free lunch* theorem [24], optimization and search algorithms are dependents of the quantity and quality of the problem knowledge that they incorporate. Thus, this is another reason to consider domain knowledge in the optimization process.

According to these premises, in this work, a multiobjective memetic algorithm is used to resolve the presented problem. MAs have been used in a variety of classical NP-hard optimization problems [25]. These works have shown that the use of MAs is often better than that obtained by the GA alone, especially when prior knowledge on suitable problem-specific memes is available [9]. The following sections describe the proposed algorithm to solve the presented problem presented.

A. Encoding of individuals and Population initialization

A solution or a chromosome of the population must be defined in terms of how their genes are encoded (using GA terminology). Here, a chromosome is represented as $C = \{c_1, c_2, \ldots, c_n\}$, where C defines a solution in the population, i.e., a configuration of antennas of the building. Each gene c_i represents the presence ($c_i = 1$) or lack ($c_i = 0$) of an antenna in the node i of the graph.

From equation 4, a number of antennas within the range of the maximum and the minimum must be preserved. Then, if the operators cause unfeasible solutions, a repairing process is applied: randomly remove or add antennas until the number is in the allowed range. Finally, the initial population is generated randomly preserving the allowed number of antennas.

B. Fitness Function

The quality of a configuration of antennas is measured by the fitness function. This quality is set according to the objectives of the problem by the equation 2. Notice that calculating the fitness for an individual (or agent) of the population implies to recalculate the HMMs for this configuration of antennas.

This approach uses a weighted sum of the two objectives as a fitness function. In this approach the main difficulty is selecting the weight vector. The weights set the relative importance for each objective. As described above, previous experiments have shown that reduction on the number of antennas is possible until a critical amount, where the precision starts to get worse. Progressively reducing the number of antennas means, in this context, to reach the prediction error that the user can afford. This is the reason why the weighted-sum approach is the most suitable in this problem. Nevertheless, a deeply analysis of the weight is shown in Section IV-B.

C. Mutation, Respectful crossover and Reproduction

Two mutation operators which change the individual are defined. By using the first mutation operator, the number of antennas is modified. A number among the allowed range of antennas is randomly selected. Following the same procedure as in the repairing process, antennas are removed and added until reaching the randomly selected number of antennas.

The second mutation modifies the position of antennas. This mutation is made using a classical mutation process. Two different indexes of the chromosome between 1 and the maximum number possible of antennas are randomly selected. The values (presence or lack of antenna) are exchanged and the number of antennas is updated. If the mutated individual is unfeasible, the repairing process will be applied. This operator can also cause a reduction or increment of one antenna. However, the first mutation is still needed in order to achieve a better covering of the search space.

The crossover operator is a classical crossover technique. Two chromosomes C_1 and C_2 and two different indexes inf and sup are randomly selected. Then, all genes between inf and sup of the chromosome C_1 are changed by the genes of the chromosome C_2. This operator is based on the feature of respectful. A recombination operator is respectful when the descendents inherit the common features of their parents. The property of respect refers in this case to the transmission of all antennas shared by parents to descendents. The repairing will be applied if an unfeasible solution is generated. Finally, in the reproduction phase, the classical roulette wheel approach is used to generate the offspring.

D. Heuristic selection

Traditional selection is based on the fitness function which provides a measure of the quality of the solution. However, some domain knowledge is available and the use of such heuristic knowledge can be addressed. This statement is based on the fact that when the meme is considered valuable enough, it should be passed on the offspring.

According to this fact, the concept of nodes more visited (NMV) is defined. Then, a node is a place where an antenna is placed. A NMV is calculated according to the number of times that a worker crosses this node. The basic idea is to consider that those nodes which are most visited by the workers provide more information to the HMMs. In this selection operator two individuals are randomly pre-selected. Then, the selection operator evaluates the quality of their configurations according to the NMV. So, the individuals with highest values of NMV are selected to the offspring. This operator is compared with the classical binary tournament selection to check how the domain knowledge can improve the results.

E. Local search

Using Local Search (LS) in a MA is usual, but not a must within the MA paradigm. However, LS uses to be one of the components that most contribute to good results [26]. Here, a LS technique is proposed in order to guide the evolution of the population to solutions which are known to improve.

The proposed LS exploits the following fact learnt from previous experiments: antennas in corridors are more critical than in rooms. Corridors are space of crossing, because of that they are usually visited for going to other places, i.e., they give more information to the prediction. Thus, the corridors are treated as special nodes. In each individual, iteratively each corridor from all corridors is added or exchanged by other antenna (randomly selected index of antenna). The decision of adding or exchanging corridors is taken by a random boolean process. After adding or exchanging the new antenna of the corridor, the fitness function is evaluated. If the fitness is higher the change is maintained, otherwise, the change will be reverted. The technique is simple, but it can improve considerably the results (see Section IV-A).

IV. Computational Experiments

In this section the most important results of the approach are shown, but first the scenario and parameters for the experiments are presented. To properly compare and check the improvement of this approach the same scenario, see Fig.1, from our previous work is used [3]. In Fig. 1 a), a plan of the floor is presented. Rooms are in blue, corridors in white and stairs in green. Doors are in red and crossings are lined. Fig. 1 b), shows a graph based on the representation of the scenario.

Destinations (orange) are labelled in the graph as D_i. Destinations $D1$, $D4$ and $D5$ are relaxation rooms, $D2$ is a meeting room and $D3$ and $D6$ are warehouses. They are spread among three zones (green, blue and yellow) in order to obtain a more realistic simulation, i.e., workers have a higher rate of visits in those destinations within their zones. In the graph of Fig. 1 b) the nodes are numbered. The number

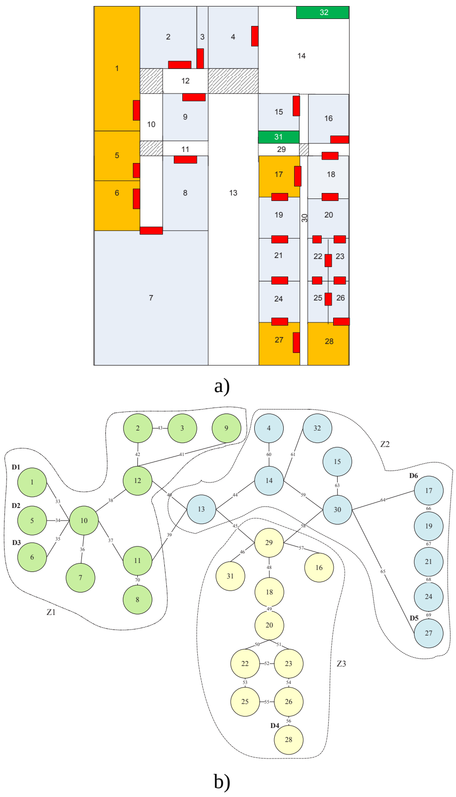


Fig. 1. a) Floor of the building b) Graph of the floor

of these elements will define the specific gen in the chromosome (0 if it does not have an antenna and 1 if otherwise).

A prediction model based on HMM is created for each destination. Since the HMMs are probabilistic, the creation of the HMMs is performed 3 times and the precision is averaged. The algorithm runs on a quad-core 2.66GHz with 4GB of RAM. The MA is also probabilistic. Thus, the execution of the memetic, for each different combination of configuration parameters, is performed 10 times, and the obtained results are averaged. The size of the initial population is 50 and it is maintained in each generation. Finally, the stopping criterion is the following: it stops when no significant improvement is obtained over 15 generations.

The following sections show the obtained results of the experiments in order to analyze the benefit of the different operators and the use of LS (Section IV-A), other multiobjective approaches and how the weights affect to the solution (Section IV-B), and, finally, the achieved improvements with respect to the previous results (Section IV-C).

A. Different configurations

Different experiments are performed in order to check the suitability of each configuration. Each normalized objective $f_1'(x)$ and $f_2'(x)$ from equation 2 is weighted, where the weights w_1 and w_2 are set to 0.9 and 0.1. From our previous work, it is known that setting the relative importance of the precision to 0.9 reduces the number of antennas keeping the precision nearly unaffected.

Table I shows the four combinations tested: Binary Tournament (BT), Nodes More Visited (NMV), algorithm without Local Search (noLS) and with it (LS). In the table is shown the average fitness, $f(x)$. The results for the two objectives functions are also shown, $f_1(x)$ is the obtained average precision and $f_2(x)$ is the average number of antennas. Besides, the best, $f_B(x)$ and the worst, $f_W(x)$, individuals fitness are shown as well. Finally, the table includes the average number of generations in each experiment, *"Ng"*, the HMM calculations, *"HMMs"*, and the average time spent per generation, *"Tg"*.

From the results, it is clear that the BT obtains worse results without LS than with it. In the case without LS, the use of the NMV operator improves a 7% with respect to the use of BT. But in the case with LS, the improvement is higher, up to 10% of improvement is obtained by the use of NMV operator. In fact, in this case the number of antennas is much higher for BT (18) than for NMV (13). The worst and the best cases also obtain better fitness when they include the use of NMV.

As observed, the use of BT causes a higher convergence of the algorithm and, therefore, it finishes with fewer generations than NMV. BT uses as guide function the fitness of the individuals, while NMV uses knowledge of the domain. So, NMV manages to escape from the local optimum and it improves the solutions. Besides, according to the time spent per generation, it is observed that the increment of time caused by the NMV is insignificant respect to the BT.

Regarding LS the results are clear. Usually, when a MA includes LS, it improves its results. This search usually increments the convergence of the algorithm while guiding the solutions to the optimal. In these experiments, the results are similar. On the one hand, LS improves the results for both selection operators. On the other hand, LS also produces a convergence of the algorithm which causes a much lower number of generations. Besides, the LS approach needs to calculate more HMMs in order to evaluate if the introduction of a corridor in an individual improves the precision and this implies an increment of the spent time per generation. However, this increment of time is offset by the lower number of generations, and then, the global execution time is affordable. The experiments show how the use of the MA obtains the best solutions, whereas the use of a basic GA (BT-noLS) obtains an average result 1.1% worse and 1.6% worse in the best case.

B. Multiobjective approaches and Weight evolution

In this section the weighted-sum approach is compared with other well-known state-of-the-art multiobjective evolutionary algorithms (MOEAs): NSGA-II [19] and MOEA/D [20]. Finally, how the variability of the weights affects the weighted-sum approach in this problem is shown.

NSGA-II is a fast non dominated sorting and crowded distance estimation procedure for ranking qualities of solutions. The majority of existing MOEAs (this includes NSGA-II) are based on Pareto dominance. Therefore, in order to evaluate a different approach, MOEA/D is also used. It decomposes the problem into a number of subproblems and optimizes them simultaneously. Two of the most used decomposition approaches have been used in this paper: Weighted Sum (W-MOEA/D) and Tchebycheff approach (T-MOEA/D) [27] (in principle, any decomposition approach can serve [28]).

Table II shows most relevant results from MA (NMV_LS), NSGA-II and MOEA/D algorithms. Same configuration and parameters as in Section IV-A are used. The table shows the best and the worst individuals' fitness, $f(x)$, and their results on each objective, $f_1(x)$ and $f_2(x)$. In NSGA-II the fitness is obtained by weighting the first and second objective of the solutions of the Pareto-optimal front obtained. Again, the weights are 0.9 and 0.1 in order to obtain a fair comparison.

Algorithm	$f(x)$	$f_1(x)$	$f_2(x)$	$f_B(x)$	$f_W(x)$	Ng	HMMs	Tg
BT_noLS	0,6828	0,7283	18	0,6917	0,6627	39	831	02’18”
NMV_noLS	0,6895	0,7195	15	0,6977	0,6771	44	1111	02’28”
BT_LS	0,6843	0,7266	18	0,6993	0,6733	32	949	03’00”
NMV_LS	**0,6945**	0,7178	13	**0,7079**	**0,6790**	35	1230	03’05”

TABLE I
RESULTS FROM DIFFERENT CONFIGURATIONS

Algorithm	$f_B(x)$	$f_{1_B}(x)$	$f_{2_B}(x)$	$f_W(x)$	$f_{1_W}(x)$	$f_{2_W}(x)$
NMV_LS	0,70794	0,73105	13	0,67904	0,72030	18
NGA-II	0,696156	0,71795	13	0.671824	0,70801	17
W-MOEA/D	0,92312	0,32309	2	0,15563	0,68234	24
T-MOEA/D	0,87463	0,38517	3	0,20082	0.68494	23

TABLE II
MEMETIC ALGORITHM, NSGA-II AND MOEA/D

From the results, it is clear that the MA reaches better results than NSGA-II. This is because the MA uses domain knowledge which improves the solutions. In the case of MOEA/D, the best individual for W-MOEA/D was reached when the weights were 0 for the first objective and 1 for the second one and the worst when the weights were $(0.13, 0.87)$. In the case of T-MOEA/D, the weights for the best and the worst individuals were $(0.02, 0.98)$ and $(0.15, 0.85)$ respectively.

Notice that when the antennas’ weight is the highest, the value of fitness reaches the best value. This is caused by the disparately scaled objectives. Although the objectives are normalized, the minimum number of antennas is 2 and this is the theoretical minimum too. However, the maximum value for the prediction is 1 and this value can not be reached because a perfect prediction is not realistic. It is well-known that the uniformness of the objectives influences on the performance of MOEA/D[20]. Despite good fitness values are reached, none of the solutions are suitable because a precision lower than 0.5 is not acceptable.

In fact, the setting of weights is analyzed in our proposal in order to evaluate how each weight affects to the problem. Fig. 2 shows the performed experiments obtained, using the MA, by varying the value of the weights of the objectives between 0 and 1, at intervals of 0.1. The X-axis defines the weight for the precision of the HMM, w_i. The value of the weight for the antennas is $(1 - w_i)$. The Y-axis shows the value of the fitness, the precision of the HMMs or the number of antennas.

According to Fig. 2, incrementing the antennas’ weight implies a decrement in the precision of HMMs. Therefore, the minimum precision which can be accepted is a parameter of this problem. The relative importance of the weights will depend on the budget and how important is a mistake in the location service.

C. Improvement by Memetic algorithm

In this section, the proposed MA is compared with the obtained results from our previous work [3]. In the previous work was shown the possibility to reduce the number of antennas from 26 to 17 (23%) without affecting the prediction. Thus, in this section is compared the obtained precision by the best solution using MA (precision of HMM 0,7311 and 13 antennas) and the precision of two interesting solutions from our previous work. The first one is a configuration with the same number of antennas: 13, it obtains a 0,6226 of precision. The second one is the best solution from our previous work: with 17 antennas and 0,6735 of precision.

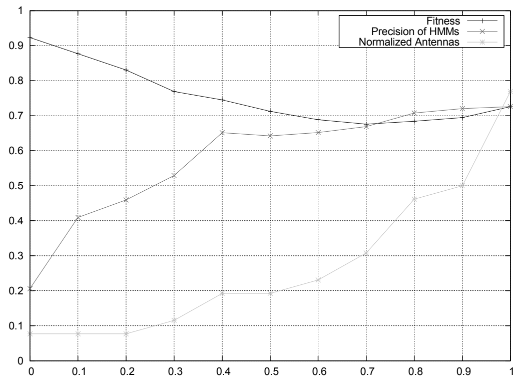


Fig. 2. Weight evolution

The MA obtains a solution with higher precision than the previous work with the same number of antennas. In fact, the improvement of the MA is over 10.8%. If the best solution from the MA is compared with the best solution from our previous work (with 17 antennas), the precision will be better again. In that case the reached improvement is only a 5.7% better. However it has been reduced the number of antennas to 50% with respect to the maximum number of antennas and to 27% with respect to the best individual from previous work.

These results show that the use of a MA in this problem can obtain better solutions. These solutions improve the precision of the prediction while the number of antennas can be reduced more than the best solutions of our previous work.

V. CONCLUSIONS AND FUTURE WORK

This paper has presented how the process of deploying a LBS based on RFID technology can be assisted by a multiobjective memetic algorithm. This algorithm introduces problem domain knowledge combined with local search. Experimentation shows that this approach helps on finding the best configuration of RFID antennas more effectively and efficiently than our previous work or other multiobjectives approaches. In the memetic algorithm, the combined use of the selection operator NMV and the LS using problem domain knowledge, obtains the best solutions. In fact, the best configuration achieves a minimum number of antennas while keeping an acceptable indoor location prediction using HMMs

This work may advance in a number of directions. One future goal is the introduction of additional constraints regarding the places where an antenna can be placed in an AmI environment. For instance, antennas should be invisible to the user in order to obtain a non-intrusive interaction. Future work also includes experimentation over the real world. Although they are still under development, some very preliminary proofs of concept have been performed in a specific real world scenario using WiFi antennas instead RFID and the results support the ideas showed in this paper.

REFERENCES

[1] K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J. C. Burgelman, "Scenarios for Ambient Intelligence in 2010," IST Advisory Group, Tech. Rep., Feb. 2001. [Online]. Available: ftp://ftp.cordis.lu/pub/ist/docs/istagscenarios2010.pdf

[2] M. Weiser, "Hot topics-ubiquitous computing," *Computer*, vol. 26, no. 10, pp. 71–72, 1993.

[3] T. Garcia-Valverde, A. Garcia-Sola, and J. Botia, "Improving RFID's location based services by means of Hidden Markov Models," in *European Conference of Artificial Intelligence (ECAI' 10)*, 2010.

[4] T. Garcia-Valverde, E. Serrano, J. Botia, G.-S. A., and J. Cadenas, "Social simulation to simulate societies of users inmersed in an ambient intelligence environment," in *Proc. IJCAI Workshop on Social Simulation*, 2009, pp. 14–25.

[5] D. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-wesley, 1989.

[6] J. Kim, J. Lee, and J. Jo, "Hierarchical spanning tree network design with Nash genetic algorithm," *Computers & Industrial Engineering*, vol. 56, no. 3, pp. 1040–1052, Apr 2009.

[7] S. Das, N. Banerjee, and A. Roy, "Solving Optimization Problems in Wireless Networks using Genetic Algorithms," *Handbook of bioinspired algorithms and applications*, p. 219, 2006.

[8] S. Pal, S. Das, and A. Basak, "Design of time-modulated linear arrays with a multi-objective optimization approach," *Progress In Electromagnetics Research B*, vol. 23, pp. 83–107, 2010. [Online]. Available: http://www.jpier.org/PIERB/pier.php?paper=10052401

[9] Y. Ong, M. Lim, and X. Chen, "Memetic computation-past, present & future [research frontier," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, pp. 24–31, May 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5447938

[10] N. Krasnogor and J. Smith, "A tutorial for competent memetic algorithms: Model, taxonomy and design issues," *IEEE Trans. Evol. Comput.*, vol. 9, no. 5, pp. 474–488, Oct 2005.

[11] Y. Ong and A. Keane, "Meta-Lamarckian learning in memetic algorithms," *IEEE Trans. Evol. Comput.*, vol. 8, no. 2, pp. 99–110, Apr 2004.

[12] Y. Ong, M. Lim, N. Zhu, and K. Wong, "Classification of adaptive memetic algorithms: A comparative study," *IEEE Trans. Syst., Man Cybern., B, Cybern.*, vol. 36, no. 1, pp. 141–152, Feb 2006.

[13] C. Lo and W. Chang, "A multiobjective hybrid genetic algorithm for the capacitated multipoint network design problem," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 30, no. 3, pp. 461–470, 2002.

[14] B. Liu, L. Wang, and Y. Jin, "An Effective PSO-Based Memetic Algorithm for Flow Shop Scheduling," *IEEE Transactions on Systems Man and Cybernetics Part B (Cybernetics)*, vol. 37, no. 1, pp. 18–27, Feb 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4067067

[15] F. Neri and E. Mininno, "Memetic compact differential evolution for cartesian robot control," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, pp. 54–65, May 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5447943

[16] M. Tang and X. Yao, "A memetic algorithm for VLSI Floorplanning," *IEEE Transactions on Systems Man and Cybernetics Part B (Cybernetics)*, vol. 37, no. 1, pp. 62–69, Feb 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4067090

[17] D. Shi, W. Jiang, and D. Xu, "Research on a Novel Multicast Routing Hybrid Optimization Algorithm and Its Application," in *Embedded and Ubiquitous Computing*, vol. 1. IEEE, 2009, pp. 339–345.

[18] A. Konstantinidis, K. Yang, H. Chen, and Q. Zhang, "Energy-aware topology control for wireless sensor networks using memetic algorithms," *Computer communications*, vol. 30, no. 14-15, pp. 2753–2764, Oct 2007. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0140366407002058

[19] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimisation: NSGA-II," in *Proceedings of the 6th International Conference on Parallel Problem Solving from Nature*, ser. PPSN VI. London, UK, UK: Springer-Verlag, 2000, pp. 849–858. [Online]. Available: http://dl.acm.org/citation.cfm?id=645825.668937

[20] Q. Zhang and H. Li, "MOEA/D: A multiobjective Evolutionary Algorithm Based on Decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, Dec 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4358754

[21] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.

[22] L. Baum, "An equality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," *Inequalities*, vol. 3, pp. 1–8, 1972.

[23] E. Dijkstra, "A note on two problems in connection with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, Dec 1959.

[24] D. Wolpert and W. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, 1997.

[25] K. Tang, K. Tan, and H. Ishibuchi, "Guest editorial: Memetic algorithms for evolutionary multi-objective optimization," *Memetic Computing*, vol. 2, no. 1, pp. 1–1, Mar 2010. [Online]. Available: http://www.springerlink.com/index/10.1007/s12293-010-0034-7

[26] P. Moscato and C. Cotta, "A gentle introduction to memetic algorithms," *Handbook of metaheuristics*, pp. 105–144, 2003.

[27] K. Miettinen, *Nonlinear Multiobjective Optimization*. Kluwer, 1999, vol. 12.

[28] H. Li and Q. Zhang, "Multiobjective Optimization Problems With Complicated Pareto Sets, MOEA/D and NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 284–302, Apr 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4633340

5.5 An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments

Title	An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments
Authors	Garcia-Valverde, T. and Garcia-Sola, A. and Hagras H. and Dooley J. and Botia, J.A. and Callaghan V. and Gomez-Skarmeta, A.
Type	Conference
Conference	IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012)
Relevance (2010)	Core Rank A+
Address	Brisbane, Australia
Publisher	IEEE
Year	2012
Month	June
State	In press

Conference details: IEEE International Conference on Fuzzy Systems
Annual IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
Leading event in the field of Fuzzy Systems.
Core Rank A+ in the ERA Conference Ranking
Website 2012: `http://www.ieee-wcci2012.org`

Authors – Personal details	
Name	**María Teresa Garcia Valverde**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Alberto García Sola**
Position	PhD student of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Hani Hagras**
Position	Professor of the School of Computer Science and Electronic Engineering
University/ Organization	University of Essex
Name	**Dr. James Dooley**
Position	Member of the Intelligent Environments Group at the School of Computer Science and Electronic Engineering
University/ Organization	University of Essex
Name	**Dr. Juan Antonio Botía Blaya**
Position	Lecturer of the Department of Information and Communications Engineering
University/ Organization	University of Murcia
Name	**Dr. Victor Callaghan**
Position	Professor of the School of Computer Science and Electronic Engineering
University/ Organization	University of Essex
Name	**Dr. Antonio Fernando Gómez Skarmeta**
Position	Professor of the Department of Information and Communications Engineering
University/ Organization	University of Murcia

Contribution of the PhD student

The PhD student, María Teresa García Valverde, declares to be the main author and the major contributor of the paper *An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments*

An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments

Teresa Garcia-Valverde, Alberto Garcia-Sola, Antonio Gomez-Skarmeta and Juan A. Botia
Dept. of Information and Communications Engineering
University of Murcia
Murcia, Spain
mtgarcia@um.es, agarciasola@um.es, skarmeta@um.es, juanbot@um.es

Hani Hagras, James Dooley and Victor Callaghan
School of Computer Science and Electronic Engineering
University of Essex
Colchester, UK
hani@essex.ac.uk, jpdool@essex.ac.uk, vic@essex.ac.uk

***Abstract*—One of the important requirements for Ambient Intelligent Environments (AIEs) is the ability to localise the whereabouts of the user in the AIE to address her/his needs. The outdoor localisation means (like GPS systems) cannot be used in indoor environments. The majority of non intrusive and non camera based indoor localisation systems require the installation of extra hardware such as ultra sound emitters/antennas, RFID antennas, etc. In this paper, we will propose a novel fuzzy logic based indoor localisation system which is based on the WiFi signals which are free to receive and they are available in abundance in the majority of domestic spaces. The proposed system receives WiFi signals from a big number of existing WiFi Access Points (up to 170 Access Points) with no prior knowledge of the access points locations and the environment. The proposed system is able to adapt online incrementally in a lifelong learning mode to deal with the uncertainties and changing conditions facing unknown indoor structures with a few days of calibration at zero-cost deployment with high accuracy. The proposed system was tested in simulated and real environments where the system has given high accuracy (that outperformed the existing techniques) to detect the user in the given AIE and the system was able also to adapt its behaviour to changes in the AIE or the WiFi signals.**

I. INTRODUCTION

An Ambient Intelligent Environment (AIE) aims to provide us with seamless, invisible and proactive services which adapt to our preferences and needs [1]. These services are performed by acquiring the contextual information of users and the environment in a nonintrusive and natural way. Hence, there is a need to localise the user accurately to be able to infer the user context and hence provide the needed services to the user(s).

There are various localisation systems available for the outdoor environments like GSM/UMTS or GPS [12]. However, these outdoor localisation means cannot be used for indoor environments.

In the recent years, there has been a good progress in indoor localisation systems, but most of them present problems like requiring time consuming calibration process, poor robustness or high cost of installation of new hardware equipment [19]. In addition, to protect the user privacy, cameras are not desired in AIEs. These problems cause such indoor localisation systems to be unsuitable to be used in AIEs where it is needed to have non intrusive, ubiquitous and cheap systems which do not necessitate the installation of expensive hardware equipment. In that sense, nowadays some indoor localisation systems use WiFi Access Points (APs) to localise the user. This alternative is truly ubiquitous and inexpensive since almost every modern dwelling has several WiFi APs installed or accessible nearby. WiFi is suitable to infer the location by analyzing and measuring the signal properties, for example using RSSI (Received Signal Strength Indicator). There have been several efforts (both commercial and academic) that used the properties of wireless signals for localisation (RADAR [5], Nibble [8] or Active Campus [15]). It was also shown by previous works [5] that the RSSI is usually more indicative of location than other properties of the signal.

The RSSI is a measure of the power present in a received signal. However, the use of WiFi RSSI is subject to noise and uncertainty due to some problems like weather condition variations, artificial variations and interferences (like building materials or walls) or interferences with other signals like microwaves or Bluetooth devices [13]. Therefore, there is a need to have a system which is able to handle the encountered uncertainty, noise and imprecision.

When using RSSI, most of the existing techniques use stochastic approaches to infer location [24]. However, none of these stochastic techniques are easily understandable by human beings and, in some cases, they are not able to deal with inexact and uncertain information. Because of that, some works such as [9] or [3] used fuzzy logic in order to handle the encountered uncertainties and provide systems whose behaviour could be easily understood by the users. However, these works need to know the APs and their location which is not practical in most domestic environments where APs come and go and it is very difficult to know their exact locations. Without the need of prior knowledge of the APs, the work of Alonso et al. [4] uses a Fuzzy Decision Tree to obtain the position of a robot. Nevertheless, this system is not able to learn online and adapt to the highly dynamic intelligent environments.

As described above, the existing works are not able to work in the long term within real environments where the WiFi signals and their strengths are continuously changing and thus there is a need for continuous adaptation of the localisation systems. With this aim, we present a novel fuzzy logic system which is able to localise the users in an indoor environment using WiFi signals from a set of APs of unknown location and whose quantity, existence and signal strength changes continuously. Only a few days of training are needed to learn different scenarios, while no prior knowledge, new equipment or extra cost are required. Furthermore, the technique is able to automatically learn offline and online to adapt in order to deal with the environmental changes.

We have compared our proposed technique with current techniques in the literature and we will show that our system outperforms the other techniques in the offline comparison. The system is the first system to be able to operate online and adapt its behaviour in an incremental and lifelong approach with no prior knowledge. We will show the evaluation of our system in a real world space where we will show the results achieved by our system which has produced a very good performance with high accuracy.

In the following section, we will present our online lifelong learning fuzzy logic based system for indoor localisation using WiFi signals. Section 3 will presents the experiments and results in both simulated and real scenarios. Finally, Section 4 presents the conclusions and future work.

II. INCREMENTAL AND ADAPTATIVE ONLINE LIFE LONG LEARNING FUZZY SYSTEM

In this work we present a novel online lifelong learning fuzzy logic system for indoor localisation using WiFi signals which provides the ability of fuzzy rules to approximate independent local models for mapping a set of inputs to a set of outputs to provide transparent and flexible representations that are easily adaptable and interpretable [11].

With the aim of being able to handle the described changes and problems in the WiFi signals, APs and the environment, the proposed system includes an offline and online lifelong learning process which adapts to the changes of the highly dynamic intelligent environments. Both fuzzy learning processes are described in the following subsections.

A. The Fuzzy Offline Learning

1) Capturing Input/Output Data: The offline approach initially monitors the physical location of the user in the AIE (room, corridor, etc.) and collects the RSSI readings from all the accessible WiFi APs of unknown location. In such way, the system accumulates during a short period of time (less than a minute in each location is enough) 'snapshots' of the current inputs WiFi RSSI ($x^{(t)}$) and the associated physical location in the AIE $C_q^{(t)}$. Using this accumulated data, the offline approach learns a descriptive model for the user localisation. Therefore given a set of data pairs:

$$\left(x^{(t)}, C_q^{(t)}\right), t = 1, 2, \ldots, N \tag{1}$$

where N is the number of data instances, $x^{(t)} \in R^n$ and $C_q^{(t)} \in R^k$, the offline system extracts rules which describe how the k output locations are related to the n input RSSI variables $x = (x_1, \ldots, x_n)^T \in R^n$.

2) Fuzzy Membership Functions Generation: The accumulated data need to be categorized into a set of fuzzy membership functions in order to quantify the raw crisp values of the RSSI of the APs into linguistic labels: "*Close*", "*Medium*", "*Far*" and "*No Detected*". The fuzzy sets "*Close*" and "*Far*" are described by left and right shoulder membership functions, respectively. The fuzzy set "*Medium*" is described by a triangular membership.

We have used the Fuzzy c-means clustering algorithm (FCM) [6] to generate the fuzzy membership functions. The FCM defines a set of p clustered regions over the sampled data. Hence, there are p centres $\bar{c_1}, \bar{c_2}, \ldots, \bar{c_p}$ defined for these clustered regions. The number of clusters p is predefined and, in our case, p was set to 3 corresponding to "*Close*", "*Medium*" and "*Far*". Then, the distances of each instance from each cluster centre is calculated and used to assign each instance with a degree of membership to each cluster. These fuzzy sets provide an appropriate granularity of the fuzzy partition when the APs have been detected.

In the case of the fuzzy set "No Detected", it is described by a zero-one membership function, where $\mu_{A_{No\ Detected}}(x) = 0\ if\ x \in U$ and $\mu_{A_{No\ Detected}}(x) = 1\ if\ x \notin U$ and the interval U defines the range where the RSSI of the AP can be detected. Finally, another fuzzy set is needed to deal with the variability of the number of APs over time: "*Don´t Care*". This variability can cause incomplete IF rules. Note that "*No Detected*" is used when the AP was not detected in a specific place, but it had been detected previously, whereas "*Don´t Care*" is used when no information about the AP is obtained that day. The "*Don´t Care*" membership function, $\mu_{A_{Don't\ Care}}(x)$, is the same as the unit interval, where $\mu_{A_{Don't\ Care}}(x) = 1\ for\ x \in U$.

3) Fuzzy Rule Extraction: The defined set of membership functions are combined with the existing input/output data to extract the rules defining the user localisation behaviour based on the WiFi RSSI. The fuzzy sets for the antecedents and consequents of the rules divide the input and output space into fuzzy regions. In such way, the used fuzzy rules are defined for the n-dimensional pattern classification problem as follows:

$$R_q: IF\ x_1\ is\ A_1^{(q)}\ and \ldots and\ x_n\ is\ A_n^{(q)}\ then\ Class\ C_q\ with\ CF_q \tag{2}$$

where $q = (1, 2, \ldots, M)$, M is the number of rules and q is the index of the rules and $x = (x_1, \ldots, x_n)^T$ is a n-dimensional pattern vector (the n-dimension of the problem is defined by the number of APs in the scenario) representing the RSSI from the WiFi APs. $A_i^{(q)}, i = 1, \ldots, n$ is an antecedent fuzzy set for the i-th WiFi attribute (the i-th attribute of the rule R_q is the RSSI of the i-th AP in the scenario), C_q is a consequent location class (the possible locations in the scenario, like *kitchen* or *office1*). In such way, a room level granularity of location is used to infer the location of the user. A room level granularity has been shown to provide sufficient location information for many AIEs applications [16]. The consequent classes are updated automatically from the training set and if a new location is added after the training phase, the online learning algorithm is able to update this information in the system. Therefore, no prior information about the structure or map of the building is required in this approach. Finally, CF_q is rule confidence [18] which is explained in Step 2) below.

The different steps involved in the offline approach are as follows:

Step 1) For an input-output pair $\left(x^{(t)}, C_q^{(t)}\right), t = 1,2,\ldots,N$, compute the membership values $\mu_{A_i^k}(x_i^{(t)})$ for each membership function $k = 1,\ldots,V$ and for each input variable i $(i = 1,\ldots,n)$ find $k^* \in \{1,\ldots,V\}$, such that

$$\mu_{A_i^{k^*}}(x_i^{(t)}) \geq \mu_{A_i^k}(x_i^{(t)}) \quad (3)$$

for all $k = 1,\ldots,V$. Let the following rule be called the rule generated by $(x^{(t)}, C_q^{(t)})$:

$$R_q\colon IF\ x_1^t\ is\ A_1^{k^*}\ and \ldots and\ x_n^t\ is\ A_n^{k^*}\ then\ Class\ C_q^{(t)} \quad (4)$$

For each input variable x_i there are V fuzzy sets A_i^k, $k = 1,\ldots V$, to characterise it; so that the maximum number of possible rules that can be generated is V^n. However given the dataset only those rules among the V^n possibilities whose dominant region contains at least one data point will be generated. In Step 1) one rule is generated for each input–class pair, where for each input the fuzzy set that achieves the maximum membership value at the data point is selected as the one in the "IF" part of the rule, as explained in (3). The firing strength of the generated rule is calculated as follows:

$$w^{(t)} = \prod_{i=1}^n \mu_{A_i^q}(x_i\ (t)) \quad (5)$$

The firing strength of a rule $w^{(t)}$ is a measure of the strength of the points $x^{(t)}$ belonging to the fuzzy region covered by the rule.

Step 2) Step 1) is repeated for all the *t* data points from *1* to *N* to obtain the *N* data rules in the form of (2). Due to the fact that the size of the training set can be quite large consisting of many similar instances, the rule base can also be quite large and it can have many rules which are conflicting. That is, rules with the same antecedent fuzzy sets but different consequent values. In this work, a confidence grade of a rule, CF_q, is used to resolve these conflicts. The confidence grade of a rule can be thought as the degree of belief for that rule or a measure of the validity of the rule [20]. In such way, in this step each rule R_q is weighted with a confidence grade CF_q. Then, if a new input causes several rules to be fired with different consequent classes, the confidence of the rule will be used to determine which rule is the most reliable to classify this input.

The *N* rules are divided into groups, with rules in each group sharing the same "IF" part. If we assume that there are *H* such groups, let group *l* have N_l rules. Then, each group *l* is divided into groups, with rules in each group sharing the same consequent C_q. If we assume that there are *G* such groups, let group *b* have N_b rules. Thus, each rule in a group *b* shares same "IF" part and same consequent. In such way, the confidence of the fuzzy rule $A_q \Rightarrow C_q$ is defined as [18]:

$$CF_q = c\left(A_q\ \Rightarrow\ C_q\right) = \frac{\sum_{e=1}^{N_b} w^{(t_e^b)}}{\sum_{u=1}^{N_l} w^{(t_u^l)}} \quad (6)$$

where the consequent of the group b is C_q, $e = 1,\ldots,N_b$, and t_e^b is the index for the data points in the group *b*. And $u = 1,\ldots,N_l$, and t_u^l is the index for the data points in the group *l*.

Step 3) The rule base is created in the previous step and, therefore, the system is able to use it to classify a new input $(x_1,\ldots,x_n)^T$ to map it to a specific location. In this step, when a new input $x^{(t)}$ arrives, the system computes its compatibility grade with each rule R_q in the rule base by the firing of the rule, $w_q^{(t)}$. If $w_q^{(t)} > 0$ the rule q will be fired. Then, the single winner method is used for classifying the new input as it is described in [18]. Therefore, for each new input $x^{(t)}$, $t = 1,2,\ldots,N$, the system measures its compatibility grade $w_q^{(t)}$ with each rule $q = 1,\ldots,M$ and finds $q^* \in \{1,\ldots,M\}$, such that

$$w_{q^*}^{(t)} \cdot CF_{q^*} \geq w_q^{(t)} \cdot CF_q \quad for\ all\quad q = 1,\ldots,M \quad (7)$$

Finally, the consequent of the class of the winner rule R_{q^*} will be the classification of the input $x^{(t)}$. Note that if no rule is fired, the input will be rejected and no classification will be made.

Step 4) In the previous step, if none of the existing rules is fired for an input, it will be rejected. Considering that even a small scenario can have a high number of APs (around 50) and this problem is 4^n for an n-dimensional pattern classification scenario with *n* APs, it is unfeasible to have all the fuzzy rules corresponding to all the combinations of the antecedent fuzzy sets. Therefore, it is possible that in Step 3) there are a lot of misclassifications, i.e., rejected inputs. In order to solve this problem, a new approach based on similarity of fuzzy rules is proposed in this work. It is based on a function distance measured between the incoming inputs and the existing rules, where no prior information is needed.

Hence, if a new input $x^{(t)}$ is misclassified, the similarity process is initiated. In this process the compatibility grade of the input $x^{(t)}$ with each rule R_q is measured using similarity. The most similar rule is chosen to classify the input. For each misclassified input $x^{(t)}$ in Step 3), its membership values are computed according to (3), where the fuzzy set that achieves the maximum membership value at the data point is selected: $A_i^{k^*}(x_i\ (t))$, where $k = 1,\ldots,V$ and $i = 1,\ldots,n$. Hence, we obtain a rule for the input $x^{(t)}$ where its antecedent part is written as follows:

$$IF\ x_1^t\ is\ A_1^{k^*}\ and \ldots and\ x_n^t\ is\ A_n^{k^*} \quad (8)$$

The similarity between the antecedents of the generated rule by the input $x^{(t)}$ and each rule R_q is measured as a function distance $\mathfrak{D}\ (A_i^{k^*}, A_i^{(q)})$ which measures the distance between each $A_i^{k^*}(x_i\ (t))$ and the *i*-th antecedent of the "IF" part of that rule using the difference between the linguistic labels which are coded. For example, the difference between the linguistic labels "*Close*" and "*Medium*" is $\mathfrak{D}$ ("Close", "Medium") = 1, and in the same way for $\mathfrak{D}$ ("Medium", "Far") = 1, $\mathfrak{D}$ ("Far", "No Detected") = 1 and $\mathfrak{D}$ ("No Detected", "Don´t Care") = 1. In the case of "*Close*" and "*Far*" is $\mathfrak{D}$ ("Close", "Far") = 2 and so on.

When the distances are measured, the similarity between the rule created by the input $x^{(t)}$ with each rule R_q is calculated as:

$$\mathcal{S}(x^{(t)}, R_q) = \frac{\sum_{i=1}^{n}\left(1 - \frac{\mathfrak{D}\left(A_i^{(t)}, A_i^{(q)}\right)}{V-1}\right)}{n} \quad (9)$$

where $\mathcal{S}(x^{(t)}, R_q) \in [0,1]$, V is the number of fuzzy sets and $i = 1, \dots, n$, where n is the number of RSSI values of the inputs which is the number of antecedents of the rule, i.e. n-dimensional problem.

Therefore, in this step, when a misclassification is produced by the input $x^{(t)}$, the system computes the similarity $\mathcal{S}(x^{(t)}, R_q)$ for each rule R_q, $q = 1, \dots, M$ and finds $R_{q^*}, q^* \in \{1, \dots, M\}$, such that

$$\mathcal{S}(x^{(t)}, R_{q^*}) \geq \mathcal{S}(x^{(t)}, R_q) \ \ for\ all\ q = 1, \dots, M \quad (10)$$

Finally, the consequent of the rule R_{q^*} will be the classification of the input $x^{(t)}$.

B. Incremental Online Learning

The offline algorithm is able to use the RSSI to predict locations with no prior information and to handle noise and uncertainties in the short term. However, over a long period of time, some changes can occur like discovering new APs different from the APs learnt or changes in the environment (change in the structure, APs, interferences, etc). The offline algorithm is not able to handle them. Because of that, the algorithm needs to be modified to provide adaptation of the system over time. In order to solve these problems, an online incremental lifelong learning is proposed in this work.

From the offline learning a fuzzy rule base $S = \{R_1, \dots, R_M\}$ is available to classify the location. Then, when a new day starts, the conditions may have changed and the system is adapted by triggering the online learning. In such way, when a new input is received its set of APs could have changed. Then, for an input data $x = (x_1, \dots, x_g) \in R^g$, g input WiFi RSSI variables are received based on the new input. Therefore, now a g-dimensional pattern vector is received and the rule base is n-dimensional. Hence, some changes are needed to classify using the process described in Step 3) of the offline learning.

Firstly, a new fuzzy rule base S' is created to represent the current information in the environment. This rule base is created as a copy of the original rule base S. Each rule $R_z \in S$, where $z = (1, 2, \dots, Y)$, Y is the number of rules in S' and z the index of the rules, is changed to reflect the new conditions. Let $A \in R^n$ be the set of AP identifiers in the rule base S' and $B \in R^g$ is the set of AP identifiers in the input. Then, for each AP identifier in B not included in A, a new antecedent is added to each rule with the value "*Don´t Care*" (i.e. the rule was incomplete regarding the new APs). When all the rules have been updated, the confidence of the rules in S' is calculated following (6).

Secondly, the new input $x = (x_1, \dots, x_g)$ also needs to be adapted. In such way, the membership values of the input are computed according to (3) and then, for each AP identifier in A not included in B, a new fuzzy set is added to the input with the value "*Don´t Care*".

After both updates, the rule base S' and the new input have the same dimension, let call this a f-dimension, where f is given by the new number of APs. Therefore, the system is able to classify using the same process than in the offline learning, but using the rule base S'.

However, note that new information in the scenario is available from the new APs and, hence, this information can be learnt by the system and used to improve the classification. In this work two kinds of learning are introduced to learn the new information: *non-automatic learning* (if there is intervention of the user) and *automatic learning* (if the user does not interact with the system).

Three new structures are used in these learning processes: two temporal fuzzy rule bases (fuzzy rule base of fails $S_{Fails} = \{R_{Fails_1}, \dots, R_{Fails_a}\}$, a = 1,..., Q, Q is the number of rules in S_{Fails} and fuzzy rule base of success ($S_{Sucess} = \{R_{Success_1}, \dots, R_{Success_d}\}$, $d = 1, .., E$, E is the number of rules is S_{Sucess} and d is the index of the rules) and a sliding window (W).

1) Non-automatic learning: This learning is raised when the user interacts with the system. That means that the user notifies a wrong classification to the system. A wrong classification in the system can be caused by an error in the fuzzy rule based system or by an error in the process of similarity caused by a misclassification. Both cases are different and, therefore, they are handled in different ways.

If an input $x^{(t)}$ is classified by the similarity process and the user notifies an error in this classification, that means that there was not enough information and because of that, no rule was fired and the similarity process was applied. In such case, the similarity process classified the input $x^{(t)}$ by the most similar rule to the input. Let such rule be called R_z, where $R_z \in S'$. Furthermore, the user is notified about the wrong classification and that the right class should be C_c. Then the system creates a new rule in the rule base S_{Fails} to reflect this new information. Following Step 1) of the offline learning for rule regeneration, a new rule is generated by $(x^{(t)}, C_c)$:

$$R_a: IF\ x_1\ is\ A_1^{k^*}\ and \dots and\ x_f\ \ is\ A_f^{k^*}\ then\ Class\ C_c\ with\ CF_a \quad (11)$$

where $a = (1, 2, \dots, Q)$, Q is the number of rules in S_{Fails}, a is the index of the rules. Furthermore, $R_a \in S_{Fails}$, $A_h^{k^*}(x_h\ (t))$ is the maximum membership value achieved at the data point, $k = 1, \dots, V$ and $h = (1, \dots, f)$, where f is the dimension of the pattern vector which represents the RSSI from the WiFi APs. The consequent of the rule is the class which was notified by the user as right for this input, C_c. Finally, the confidence of the rules in S_{Fails} is recalculated as in (6) for each rule in S_{Fails}. Notice that C_c can be one of the existing locations, where $C_c \in R^k$, or a new one defined by the user (e.g. a change in the structure of the building) and, then, the new location will be included in R^k.

Then, when another new input arrives, the system uses the rule base S_{Fails} to classify and if no rule is fired, the rule base S' will be used as usual. This process enables the system to correct immediately the error and use the new information for the following classifications.

On the other hand, if an input $x^{(t)}$ is classified by a fired rule R_z in the rule base S' and the user notifies an error in this classification, a non-automatic learning is also produced. However, in this case, the fired rule R_z is introduced in the sliding window *W*. We take a sliding window of σ fired rules. Each rule in the window is recorded as a triple $\{R_z, C_c, \mathscr{v}\}$, where $R_z \in S'$, C_c is the right consequent notified by the user and $\mathscr{v}$ is the number of times that the rule has produced this error. In such way, if the user notifies an error with the right consequent C_c and the error was produced by a fired rule $R_z \in S'$, the system will check if the pair (R_z, C_c) was already in *W*. If this is the case, $\mathscr{v}$ will be incremented by one unit. Otherwise, the triple $(R_z, C_c, 1)$ will be introduced in *W*.

Each time that a new triple is introduced in *W*, a process of window updating is performed. The oldest triple $\{R_z, C_c, \mathscr{v}\}^*$ in *W* is checked. If its $\mathscr{v}$ value is higher than $\mathscr{p}$, the consequent of R_z is updated for C_c in S'. Otherwise, if the $\mathscr{v}$ value is lower, no updating will be performed and the error will be forgotten. This process, called "learning inertia" [11], prevents the system from adapting its rules in response to "one off" user actions that do not reflect a marked change. Here, we consider that $\mathscr{p} = 3$ means a real change.

2) Automatic learning: In this kind of learning, it is considered that the classification was right because the user does not notify an error. However, if the classification was made by the similarity process, new information can be obtained (the similarity process is used if there is not enough information). The system learns new information as follows.

Let the rule which was used to classify the input $x = (x_1, \ldots, x_g)$ by the similarity process be called R_z (the most similar rule to the input), where $R_z \in S'$ and C_z is the consequent class of the rule R_z which was used to classify. If the user does not notify an error, a right classification is considered and the system will create a new rule in the rule base $S_{Success}$. Then, following Step 1) of the offline learning for rule regeneration, the new rule is generated by $(x^{(t)}, C_z)$:

$$R_d: IF\ x_1\ is\ A_1^{k^*}\ and \ldots and\ x_f\ \ is\ A_g^{k^*}, then\ Class\ C_z\ with\ CF_d \quad (12)$$

where $d = (1, 2, \ldots, E)$, E is the number of rules in $S_{Success}$ and d is the index of the rules. Furthermore, $R_d \in S_{Success}$, $A_j^{k^*}(x_j(t))$ is the maximum membership value achieved at the data point, $k = 1, \ldots, V$, $j = (1, \ldots, g)$ and the consequent C_z of the rule is the consequent of R_z. Finally, the confidence of the rules in $S_{Success}$ is recalculated as in (6) for each rule in $S_{Success}$. This process enables the system to learn situations not captured by the existing rules and create new rules that reflect these conditions to be used in the following classifications.

Both non-automatic and automatic learning add rules to the rule base of fails, S_{Fails}, or to the rule base of successes, $S_{Success}$. Besides, it could have modified the rule base S' by the sliding window *W* and the discovery of new APs. Hence, at the end of the day, this new information is updated in the original rule base *S*. All the rules in S', S_{Fails} and $S_{Success}$ are combined into *S* and the temporal structures S_{Fails}, $S_{Success}$ and S' are cleaned.

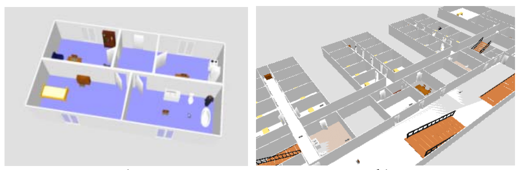
With the incremental online learning described, the system is able to change both existing rules as adding new rules. This allows rules to continue operating even if there are changes in the APs or in the environment conditions. Furthermore, even if there is a situation in the environment which is not captured by the existing rules, the system will automatically create new rules that satisfy the new conditions. Hence, the system adopts life-long learning, where it adapts to the environmental changes while a zero cost and easy deployment is kept.

III. Experiments and Results

This section presents the experiments and results performed to validate our proposal. The experiments are made, firstly, using a simulator of intelligent and ubiquitous environment, UbikSim [7], and then, we report our experiments in real world settings. In the following subsections, we will describe these experiments and show the feasibility of the proposal.

A. Simulation Experiments

The use of the UbikSim simulator (shown in Fig. 1) as a first phase provides an easy and low cost solution to test and validate the system after its deployment on the real world. UbikSim includes tools to simulate the physical environment (building, sensors, WiFi APs, etc.) and humans able to move through the environment [14].



a) b)

Figure 1. a) Home b) Office block

Two different scenarios have been simulated using UbikSim: one small to test the system and its ability of learning and a second one more complex. The first scenario represents a typical house (see Fig. 1a)) where an AP is located in each room and the corridor does not have any AP. The second scenario (see Fig. 1b)) is an office block which contains 73 places (rooms, offices and facility rooms) where 41 APs are placed in 41 of these places.

Finally, a simple model of the WiFi signal is set to simulate the signal propagation. The strength of the signal of the AP decreases by a factor randomly selected from an interval which depends on the distance between the AP and the receiver. Therefore, the RSSI is smaller when the distance between the AP and receiver is bigger. In order to achieve a more realistic model, a Gaussian noise is added to the signal to simulate the variability of the RSSI and some values randomly are removed to simulate signal-loss.

In this stage of the experiments, only the offline learning is used since the simulated environment does not change, nevertheless a 6% of noise was introduced. Our proposal is compared with two decision tree algorithms (J48 [21] and REPTree [23]) and with two decision rule algorithms (JRIP [10] and PART [22]), which use crisp values in the rules. These algorithms are chosen because they are classical techniques which have been proven to show good results in this kind of problems. Furthermore, these algorithms share similar features with our approach, like the use of rules and a higher understandability than other kind of techniques. Other techniques were tested (such IBk [2], Naïve Bayes, Bayesian Network or OneR [17]). However, the techniques we will compare against obtained better results and hence we will compare only with these techniques in the paper due to the space constraints.

Table 1 and Table 2 summarize the results for the simulated home and the building block respectively. Both tables describe the percentages of success, errors, not classified instances, size of the model (number of rules in the rule algorithms and number of nodes in the decision tree algorithms) and finally, the time taken to build the model. Note that in the experiments an error in the classification means a fail in the prediction location. Other measures which take into account the distances between rooms could be considered. Nevertheless, it will depend on the level of granularity required for the AIE. As described above, a room level granularity is considered sufficient in most AIEs systems [16].

TABLE I. RESULTS OF SIMULATED HOME

Algorithm	Success	Errors	No classifies	Size	Time (ms)
Fuzzy Based System	**81,21%**	18,79%	0%	665	1,05
Fuzzy Based System without Similarity	79,9%	15,04%	5 %	665	1,03
J48	80,86%	19,14%	0 %	53	0.13
REPTree	79,71%	20,29%	0 %	43	0.15
JRIP	80,30%	19,70%	0 %	12	2.24
PART	80,60%	19,44%	0 %	27	2.04

TABLE II. RESULTS OF SIMULATED BUILDING

Algorithm	Success	Errors	No classifies	Size	Time (ms)
Fuzzy Based System	**82,06%**	17,94%	0 %	1065	3.55
Fuzzy Based System without Similarity	77,30%	15,92%	6,78 %	1065	3.51
J48	81,71%	18,29%	0 %	661	1.33
REPTree	78,51%	21,49%	0 %	215	70.4
JRIP	77,07%	22,93%	0 %	419	37.99
PART	81,38%	18,62%	0 %	9185	23.98

From the results, the best performances are obtained by our system (using the idea of similarity described), J48 and PART, while the other techniques show worse results and longer times to build their models. A longer time to build the model is also taken by PART algorithm which makes impossible to use it in real environments. Among the best three algorithms, our system obtains the best results. However, although J48 and PART algorithms have as good performance as our approach, they are less understandable. Furthermore, unlike our proposal both J48 and PART are not able to learn and adapt online to changes of the environment, which is needed in real environments. Therefore, the results show that our technique is able to classify the location and obtain as good results as the other techniques while being build in a relatively fast time and maintaining the understandability of the system.

B. Real World Experiments

The simulation assumes behaviours that are not always true in the real world. Therefore, real world experiments were needed to validate the technique thoroughly.

In the real experiments, the collected data is defined as follows. Every time the system is launched, a new set A to contain the different APs and their RSSIs is created and initialised. Every time new information is retrieved from the system, a 'snapshot' of all the APs with their RSSI seen so far is recorded. Note that due to the features of RSSI, it is not possible to differentiate between a lost AP and an AP which is running but the signal is too low to be discovered. Because of that we record it but with RSSI equivalent to zero. Therefore, to the system, APs can disappear only when the application is re-launched, which is done once a day.

The real world experiments were performed in the iSpace in the University of Essex, UK. The iSpace is an intelligent apartment acting as a test bed for AIEs [11] (see Fig. 2). The rooms are small and very close to each other, which leave little distance between each room. Furthermore, there are a lot of interferences due to the big number of APs (from 70 to 90 APs), other localisation systems and some metallic walls which cause signal reflection.

Figure 2. iSpace

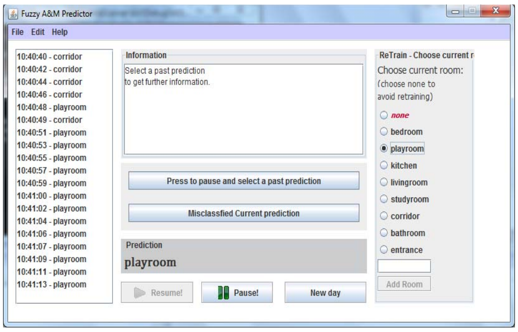


Figure 3. Desktop predictor software

Fig. 3 shows the software program used to capture input data, which allows the system to perform the offline and online training. It also shows the output data which permits the user to interact with the system (in the right part the user can notify the right location if there was an error, or even introduce new locations). Finally, the output is used to validate the system.

1) Offline Experiments: Firstly, the offline technique was tested over the real scenario in order to determine if it is able to handle the features of the real world without the online learning. In such way, the collected data for the experiments were collected in the same day where 2/3 of the data were used for training and the rest 1/3 for testing. Table 3 describes the obtained results. Like the simulated experiments, our proposal was compared with other traditional techniques, however J48 is the only technique which obtained good results in a reasonable time.

TABLE III. OFFLINE RESULTS

Technique	%Success
J48	75,57 %
Fuzzy Based System	**76,40 %**

The results show that our technique outperforms the J48.These experiments are performed using data from one day. However, experiments using a different data set from different days to train and test the system show that the results get worse down to 10%. This worsening happens because the offline technique is not able to handle new APs and environmental changes. Hence, online learning is needed in order to provide a good performance over a long period.

2) Online Experiments: This subsection describes the experiments over the real world to validate if the whole technique was able to learn over time and adapt to the changes. The experiments were carried out using different data sets from different days obtained over a period of one month. In the iSpace, after 8 days the system gets stable and no more training data is needed.

The experiments are performed as follows. A fixed offline learning data set, DS_0, is used for all the tests (the data set obtained the first day). After this first day, a period of 19 days is given without taking data to check the performance after a period without learning. Then, different data sets in the following days are collected to test the online learning: $\{DS_1, DS_2, \dots DS_{\varpi-1}\}$, where each DS_α is the data set collected the α-th day and is taken a day after the data set $DS_{\alpha+1}$, and $\alpha = 1, \dots \varpi$, where ϖ is the day in which the system gets stable. Hence, data sets are collected until the system gets stable. Finally, the data set of the day in which the system gets stable, DS_ϖ, is used to validate the system each day. Using these data sets different tests are performed. In all tests the data set DS_0 is used for the offline learning and the last collected data set DS_ϖ is used to validate the system.

In such way, in the first test (*Test 1*) only offline learning is made using DS_0 and then, the system is validated using the data set DS_ϖ. In the second test (*Test 2*) the offline learning is made using DS_0 and then, an online learning is made using DS_1, finally, the validation is made using DS_ϖ. In the third test (*Test 3*), the offline learning and the validation are made in the same way than the previous tests, but two online learning are made using $\{DS_1, DS_2\}$. Therefore, each test incrementally adds one more online learning data set as follows:

$$Test_\beta = Offline\ learning(DS_0) + Online\ learning\ \{DS_1, \dots, DS_{\beta-1}\} + Validation\ (DS_\varpi) \quad (13)$$

where $\beta < \varpi$. Finally, a last test is made to check the performance of the system after a long period without any update, i.e. since the stabilization of the system. Then, a new data set is taken two weeks after the day in which the last data set, DS_ϖ, was taken. The new data set is used to validate the system. Results of these tests are shown in Fig. 4.

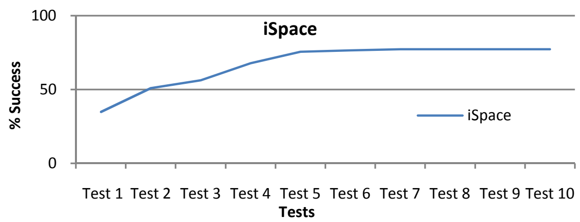


Figure 4. Online learning

From the results, it is clear that in *Test 1* a low success rate is obtained. As we described above, *Test 1* is performed using only the offline learning and then, it is validated with the data set DS_ϖ. Therefore, the system has not been updated for more than a month. In the following tests, the system continues learning using the online learning. In fact, the system improves until it gets 76.47% in the iSpace.

Finally, two weeks after these experiments, the last test was made in order to check the performance after a long period of time without updating the system. The system reached 77.22% in the iSpace. In fact, the system is able to stabilize with a very good accuracy.

It is clear that the system has learnt and adapted to medium-term changes over time and is able to provide a good performance in a long-term when abrupt changes occur using online-learning. Despite the real world presents some difficulties such as noise or changes in APs, the obtained results are very promising, obtaining a good success rate.

IV. CONCLUSIONS AND FUTURE WORK

This paper has presented a novel fuzzy online learning and adaptation localisation system for indoor environments using WiFi RSSI. The system provides a zero cost solution which uses the deployed WiFi APs with a short training period and with no knowledge about the APs and its location. The system is able to adapt and handle short and long term uncertainties and environmental changes in a lifelong learning mode.

A novel similarity process is presented here to deal with the problem of having a high number of APs. This process allows the system to continue classifying when the environmental conditions change, even if there are no rules corresponding to this new situation. Furthermore, in order to provide a lifelong localisation solution a novel online learning is also presented. Online learning is able to extend the rule base to adapt unobtrusively the system to new environment conditions. The overall solution provides high accuracy at zero cost in real world living spaces.

With the aim to show the feasibility of our proposal, we have performed different experiments in simulated environments as in real AIEs like the iSpace in the University of Essex. The performance of our system is compared with other approaches and we have shown that our proposal outperforms the other techniques while being able to operate online in a long term mode to learn and adapt to changes of the environmental conditions.

The system provides a truly ubiquitous solution with high accuracy, good resilience to changes and a low cost and easy deployment which can be used in very different scenarios of ambient intelligence. Among our future works more experiments in different scenarios are included. In fact, we are currently testing the system over a domestic apartment in the town centre and in an office block. Both scenarios cover two very different environments regarding location and number of APs. Preliminary results show a very good performance with a high accuracy. Other future works include extending the current system, which is based on a type-1 fuzzy system to a type-2 fuzzy system. This extension provides a more appropriate framework for modelling and handling the short and long uncertainties arising in WiFi environments.

ACKNOWLEDGMENT

This research work is supported by the Spanish Ministry of Science and Innovation under the grant AP2007-04269 of the FPU program, the 09493/FPI/08 grant and in the scope of the Research Projects TRA2009 0141 and TSI-020302-2010-129.

REFERENCES

[1] G. Acampora and V. Loia. "Fuzzy control interoperability and scalability for adaptive domotic framework." IEEE Transactions on Industrial Informatics, vol. 1, no. 2, pp. 97–111, May 2005.

[2] D. Aha. "Tolerating noisy, irrelevant, and novel attributes in instance-based learning algorithms." Int J Man-Machine Studies, Vol. 36, 267–287, 1992.

[3] U. Ahmad, A. Gavrilov, Y. Lee, and S. Lee, "Context-aware, self-scaling Fuzzy ArtMap for received signal strength based location systems," Soft Computing-A Fusion of Foundations, Methodologies and Applications, vol. 12, no. 7, pp. 699–713, 2008.

[4] J. Alonso, M. Ocaña, M. Sotelo, L. Bergasa, and L. Magdalena, "WiFi localization system using fuzzy rule-based classification," Computer Aided Systems Theory-EUROCAST 2009, Las Palmas de Gran Canaria, Spain, pp. 383–390, 2009.

[5] P. Bahl, and V. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv , Israel, 2000, pp. 775–784.

[6] J.Bezdek, Pattern recognition with fuzzy objective function algorithms. New York: Plenum, 1981.

[7] F. Campuzano, T. Garcia-Valverde, A. Garcia-Sola, and J. Botia. "Flexible simulation of ubiquitous computing environments." International Symposium of Ambient Intelligence, Salamanca, Spain, 2011, pp. 189–196.

[8] P. Castro, P. Chiu, T. Kremenek, and R. Muntz, "A probabilistic room location service for wireless networked environments," in Ubicomp 2001: Ubiquitous Computing. Springer, Atlanta, Georgia, 2001, pp. 18–34.

[9] C. Chen, J. Yang, G. Tseng, Y. Wu, and R. Hwang, "An Indoor Positioning Technique Based on Fuzzy Logic," Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, 2010, vol. 2, pp. 854–857.

[10] W. Cohen. "Fast effective rule induction" in Proceedings of the 12th International Conference on Machine Learning, Lake Tahoe City, California, pp. 115-123, 1995.

[11] F. Doctor, H. Hagras, and V. Callaghan. "A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments." IEEE Transactions on Systems, Man, and Cybernetics, Part A, vol. 35, no. 1, pp. 55–65, 2005.

[12] P. Enge, and P. Misra, "Special issue on global positioning system," Proceedings of the IEEE, vol. 87, no. 1, pp. 3–15, 1999.

[13] Farpoint Group. "The Invisible Threat: Interference and Wireless LANs". White Paper. Document FPG 2006-321.1. Oct, 2006.

[14] T. Garcia-Valverde, A. Garcia-Sola, and J. Botia. Improving RFID's location based services by means of Hidden Markov Models. In European Conference of Artificial Intelligence (ECAI' 10), Lisbon, Portugal, 2010.

[15] W. Griswold, R. Boyer, S. Brown, T. Truong, E. Bhasker, G. Jay, and R. Shapiro, "Activecampus-sustaining educational communities through mobile technology" University of California, San Diego, Dept. of Computer Science and Engineering, Technical Report, 2002.

[16] A. Haeberlen, E. Flannery, A. Ladd, A. Rudys, D. Wallach, and L. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in Proceedings of the 10th annual international conference on Mobile computing and networking. ACM, Philadelphia, PA, 2004, pp. 70–84.

[17] R. Holte. "Very simple classification rules perform well on most commonly used datasets." Machine Learning, vol. 11, pp. 63–91, 1993.

[18] H. Ishibuchi, and T. Yamamoto. "Rule weight specification in fuzzy rule-based classification systems", IEEE Transactions on Fuzzy Systems, vol 13, no. 4, pp. 428-435, 2005.

[19] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 37, no. 6, pp. 1067–1080, 2007.

[20] J. Mendel, "Fuzzy logic systems for engineering: A tutorial," Proc. IEEE, vol. 83, pp. 345–377, Mar. 1995.

[21] R. Quinlan. "C4.5: Programs for Machine Learning". Morgan Kaufman, San Mateo, 1993.

[22] I. Witten, and E. Frank. "Generating accurate rule sets without global optimization" in Proceedings of the 15th International Conference on Machine Learning, Madison, Wisconsin, 1998, pp. 144-151.

[23] I. Witten, and E. Frank. "Data Mining – Practical Machine Learning Tools and Techniques". Elsevier Inc., 2005.

[24] G. Zaruba, M. Huber, F. Kamangar, and I. Chlamtac, "Indoor location tracking using RSSI readings from a single Wi-Fi access point," Wireless networks, vol. 13, no. 2, pp. 221–235, 2007.

Chapter 6

Acceptance letters

Combining the real world with simulations for a robust testing of Ambient Intelligence services

Decision on your manuscript #AIRE-519R1

Asunto: Decision on your manuscript #AIRE-519R1
De: "Artificial Intelligence Review (AIRE)" <geetha.bhaskar@springer.com>
Fecha: 08/04/2012 11:53
Para: "Teresa Garcia-Valverde" <mtgarcia@um.es>

Dear Ms. Teresa Garcia-Valverde:

We are pleased to inform you that your manuscript, "Combining the real world with simulations for a robust testing of Ambient Intelligence services" has been accepted for publication in Artificial Intelligence Review.

You will receive an email from Springer in due course with regards to the following items:

1. Offprints

2. Colour figures

3. Transfer of Copyright

Best regards,
Springer Journals Editorial Office
Artificial Intelligence Review

Figure 6.1: Acceptance letter from Journal of Artificial Intelligence Review

Full reference:

T. Garcia-Valverde, E. Serrano, and J.A. Botia. Combining the real world with simulations for a robust testing of ambient intelligence services. *Artificial Intelligence Review*, pages 1–24, 2012. 10.1007/s10462-012-9340-4. Impact factor (2010): 0.429. Available since 29 April 2012.

Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach

05/02/12 mstracker.com/reviews.php?id=19675&aid=23852

Journal of Ambient Intelligence and Smart Environments

MANUSCRIPT 11-163-RR

"Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach"

DECISION LETTER

We would like to thank you for the revised submission of your paper entitled

"Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach"

to the Thematic Issue: 'A Software Engineering Perspective on Smart Applications for AmI' of the IOS Journal on Ambient Intelligence and Smart Environments. We are happy to confirm the final approval of your paper. Congratulations!

With kind regards from the guest editors,
Davy Preuveneers, Paulo Novais and Juan M. Corchado

Figure 6.2: Acceptance letter from Journal of Ambient Intelligence and Smart Environments

Full reference:

T. Garcia-Valverde, F. Campuzano, E. Serrano, A. Villa, and J. A. Botia. Simulation of human behaviours for the validation of ambient intelligence services: a methodological approach. *Journal of Ambient Intelligence and Smart Environments*, 2012. In-press. Impact factor (2010): 1.5.

Improving RFID's Location Based Services by means of Hidden Markov Models

Full reference:

T. Garcia-Valverde, A. Garcia-Sola, and J.A. Botia. Improving RFID's location based services by means of hidden markov models. In *Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pages 1045–1046, Lisbon, Porto, August 2010. IOS Press.

Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach

Asunto: SMCCSI-10-12-0520.R1: acceptance, correspondence
De: smcc-eic@labe.felk.cvut.cz
Fecha: 27/01/2012 8:12
Para: mtgarcia@um.es
CC: smcc-assist@labe.felk.cvut.cz

Prof. Vladimir Marik,Editor in Chief
Department of Cybernetics, FEE
Czech Technical University in Prague
Technicka 2
166 27 Prague 6, Czech Republic

27-Jan-2012

Dear Miss Teresa Garcia-Valverde,

ID: SMCCSI-10-12-0520.R1
Title: Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach
Authors: Garcia-Valverde, Teresa; Garcia-Sola, Alberto; Botia, Juan A.; Gomez-Skarmeta, Antonio

Type: TECHNICAL CORRESPONDENCE
Standard length: 6 I-EEE formatted pages

It is a pleasure to inform you that your manuscript has been accepted for Publication as a TECHNICAL CORRESPONDENCE in the Transactions on Systems, Man, and Cybernetics--Part C: Applications and Reviews.

Figure 6.3: Acceptance letter from IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews

Full reference:

T. Garcia-Valverde, A. Garcia-Sola, J.A. Botia, and Gomez-Skarmeta A. Automatic design of an indoor user location infrastructure using a memetic multiobjective approach. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, PP:1–6, 2012. Impact factor (2010): 2.105. Available since 17 April 2012.

An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments

Asunto: FUZZ-IEEE 2012 Paper #6 Decision Notification
De: James M.Keller <fuzzieee2012@ieee-cis.org>
Fecha: 21/02/2012 19:08
Para: mtgarcia@um.es

Dear Author(s),

Congratulations! On behalf of the FUZZ-IEEE 2012 Technical Program Committee and the program chairs, we are pleased to inform you that your paper:

Paper ID: 6
Author(s): Teresa Garcia-Valverde, Alberto Garcia-Sola, Hani Hagras, James Dooley, Juan A. Botia, Victor Callaghan and Antonio Gomez-Skarmeta
Title: An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments

has been accepted for presentation at the 2012 IEEE International Conference on Fuzzy Systems and for publication in the conference proceedings published annually by IEEE. This email provides you with all the information you require to complete your paper and submit it for inclusion in the proceedings. A notification of the presentation format (oral or poster) and timing of that presentation will be sent by the end of May.

Figure 6.4: Acceptance letter from IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012)

Full reference:

T. Garcia-Valverde, A. Garcia-Sola, Hagras H., Dooley J., J.A. Botia, Callaghan V., and A. Gomez-Skarmeta. An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012)*, Brisbane, Australia, 2012. IEEE. In press.

Chapter 7

Publications relevance

Combining the real world with simulations for a robust testing of Ambient Intelligence services

The article entitled 'Combining the real world with simulations for a robust testing of Ambient Intelligence services', has been accepted in the journal of *Artificial Intelligence Review*, whose relevance and impact factor can be observed next:

ISI Web of Knowledge℠

Journal Citation Reports®

WELCOME | HELP | RETURN TO JOURNAL

2010 JCR Science Edition

Rank in Category: ARTIFICIAL INTELLIGENCE REVIEW

Journal Ranking

For **2010**, the journal **ARTIFICIAL INTELLIGENCE REVIEW** has an Impact Factor of **0.429**.

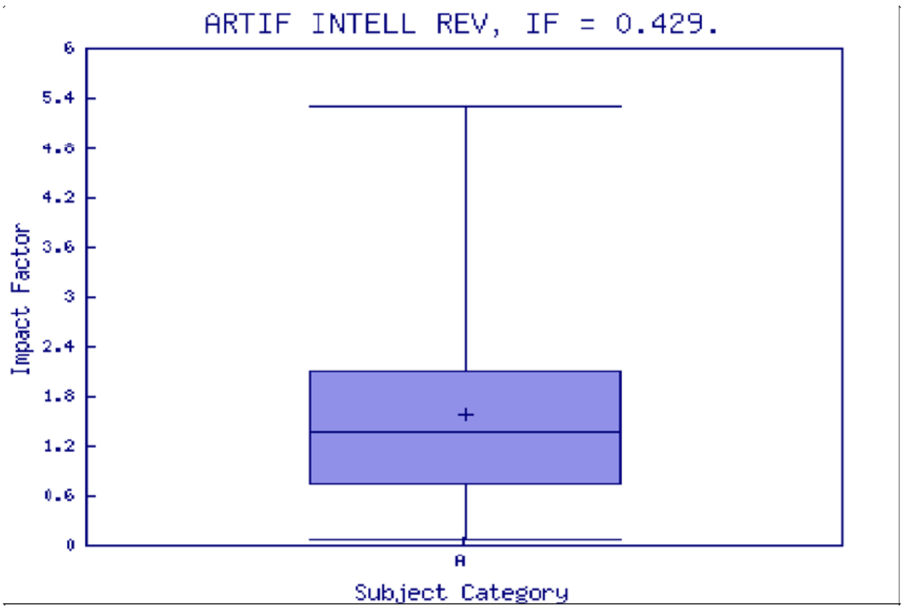
This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	108	97	Q4

Category Box Plot

For **2010**, the journal **ARTIFICIAL INTELLIGENCE REVIEW** has an Impact Factor of **0.429**.

This is a box plot of the subject category or categories to which the journal has been assigned. It provides information about the distribution of journals based on Impact Factor values. It shows median, 25th and 75th percentiles, and the extreme values of the distribution.



Key

A - COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE

Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach

The article entitled 'Simulation of Human Behaviours for the Validation of Ambient Intelligence Services: a Methodological Approach', has been accepted in the *Journal of Ambient Intelligence and Smart Environments*, whose relevance and impact factor can be observed next:

ISI Web of Knowledge℠

Journal Citation Reports®

WELCOME | HELP | RETURN TO JOURNAL

2010 JCR Science Edition

Rank in Category: Journal of Ambient Intelligence and Smart Environm...

Journal Ranking

For **2010**, the journal **Journal of Ambient Intelligence and Smart Environm...** has an Impact Factor of **1.500.**

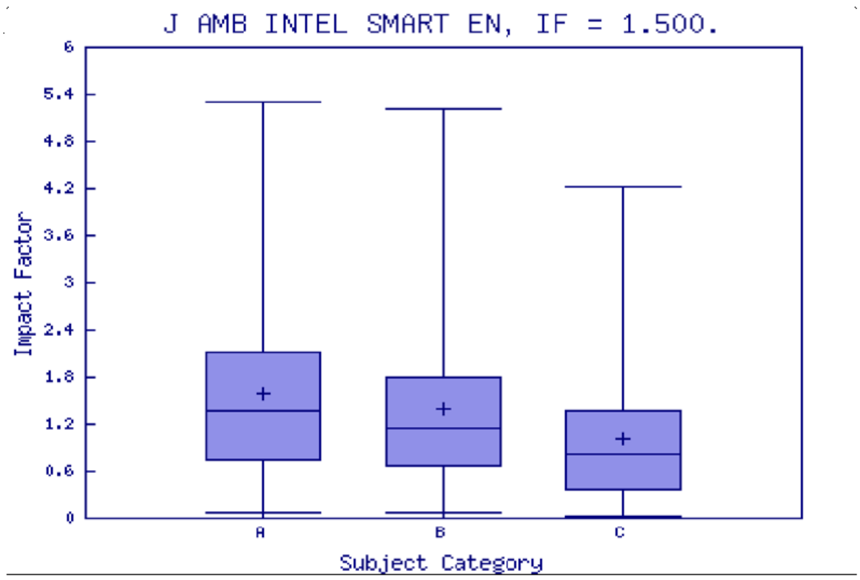
This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	108	48	Q2
COMPUTER SCIENCE, INFORMATION SYSTEMS	128	47	Q2
TELECOMMUNICATIONS	80	17	Q1

Category Box Plot

For **2010**, the journal **Journal of Ambient Intelligence and Smart Environm...** has an Impact Factor of **1.500.**

This is a box plot of the subject category or categories to which the journal has been assigned. It provides information about the distribution of journals based on Impact Factor values. It shows median, 25th and 75th percentiles, and the extreme values of the distribution.



Key

A - COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE

B - COMPUTER SCIENCE, INFORMATION SYSTEMS

C - TELECOMMUNICATIONS

Improving RFID's Location Based Services by means of Hidden Markov Models

The article entitled 'Improving RFID's Location Based Services by means of Hidden Markov Models', has been published in the Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence.

ECAI is the leading Conference on Artificial Intelligence in Europe, and is a biennial organization of ECCAI (European Coordinating Committee for Artificial Intelligence).

The ECAI 2010 review process was extremely selective: of the 607 papers submitted, only 22% were accepted as full papers, with a further 18% accepted as short papers/posters. ECAI is considered as Core Rank A+ in the ERA Conference Ranking.

Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach

The article entitled 'Automatic Design of an Indoor User Location Infrastructure using a Memetic Multiobjective Approach', has been accepted in the journal of *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, whose relevance and impact factor can be observed next:

ISI Web of Knowledge℠

Journal Citation Reports®

WELCOME | HELP | RETURN TO JOURNAL

2010 JCR Science Edition

Rank in Category: IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS P...

Journal Ranking

For **2010**, the journal **IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS P...** has an Impact Factor of **2.105**.

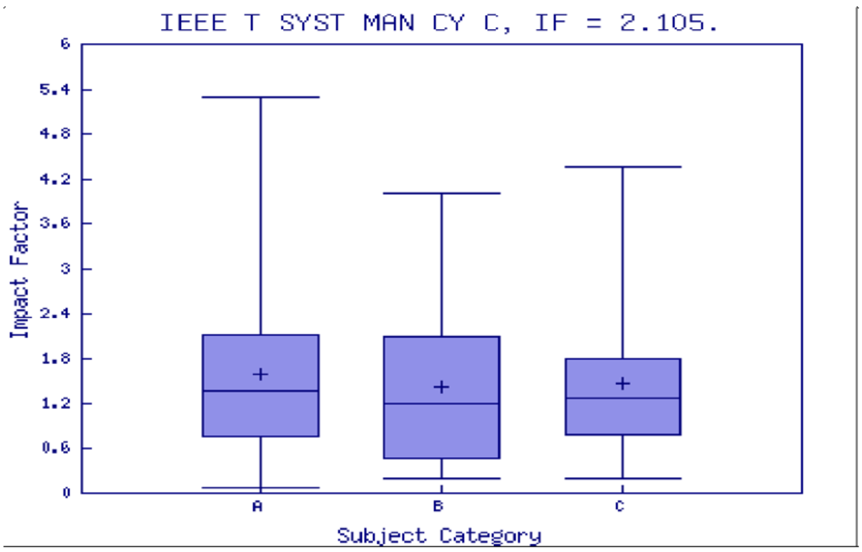
This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	108	27	Q2
COMPUTER SCIENCE, CYBERNETICS	19	4	Q1
COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS	97	18	Q1

Category Box Plot

For **2010**, the journal **IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS P...** has an Impact Factor of **2.105**.

This is a box plot of the subject category or categories to which the journal has been assigned. It provides information about the distribution of journals based on Impact Factor values. It shows median, 25th and 75th percentiles, and the extreme values of the distribution.



Key

A - COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE

B - COMPUTER SCIENCE, CYBERNETICS

C - COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS

An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments

The article entitled ‘An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments’, has been published in the Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012). The annual FUZZ-IEEE is one of the leading events in the field of fuzzy systems.

FUZZ-IEEE is considered as Core Rank A+ in the ERA Conference Ranking. This papers has received a grant from the IEEE Computational Intelligence Society. The grant awarded process, which was based on the ranking by the reviewers, was extremely selective: a total of 141 applications and 45 grants have been awarded.

Chapter 8

Bibliography

8.1 References

[1] E. Aarts and R. Wichert. Ambient intelligence. *Technology Guide*, pages 244–249, 2009.

[2] Gregory D. Abowd, Anind K. Dey, Peter J. Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 304–307, London, UK, 1999. Springer-Verlag.

[3] Juan Carlos Augusto, Hideyuki Nakashima, and Hamid Aghajan. Ambient intelligence and smart environments: A state of the art. In Hideyuki Nakashima, Hamid Aghajan, and Juan Carlos Augusto, editors, *Handbook of Ambient Intelligence and Smart Environments*, pages 3–31. Springer, New York, 2010.

[4] Matthias Baldauf, Matthias Baldauf, Schahram Dustdar, and Schahram Dustdar. A survey on context-aware systems. *International Journal of Ad Hoc and Ubiquitous Computing*, 2004.

[5] J. Benedicto, S. E. Dinwiddy, G. Gatti, R. Lucas, and M. Lugert. Galileo: satellite system design and technology developments, european space agency, 2000.

[6] Gaetano Borriello, Matthew Chalmers, Anthony LaMarca, and Paddy Nixon. Delivering real-world ubiquitous location systems. *Commun. ACM*, 48(3):36–41, March 2005.

[7] P. J. Brown. The stick-e document: a framework for creating context-aware applications. In *Proceedings of EP'96, Palo Alto*, pages 259–272, 1996.

[8] C. Darwin. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. J. Murray, sixth edition, June 1872.

[9] Richard Dawkins. *The Selfish Gene*. Oxford University Press, September 1990.

[10] K. Ducatel, M. Bogdanowicz, F. Scapolo, J. Leijten, and J. C. Burgelman. Scenarios for ambient intelligence in 2010. Technical report, IST Advisory Group, February 2001.

[11] P. Enge and P. Misra. Special issue on global positioning system. *Proceedings of the IEEE*, 87(1):3–15, 1999.

[12] Carlos M. Fonseca and Peter J. Fleming. Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In *Genetic Algorithms: Proceedings of the Fifth International Conference*, pages 416–423. Morgan Kaufmann, 1993.

[13] O. Gabel, L. Litz, and M. Reif. NCS testbed for ambient intelligence. In *2005 IEEE International Conference on Systems, Man and Cybernetics*, volume 1, 2005.

[14] Abdulbaset Gaddah and Thomas Kunz. A survey of middleware paradigms for mobile computing. *Techical Report, July*, 2003.

[15] Tao Gu, Hung K. Pung, and Da Q. Zhang. Toward an osgi-based infrastructure for context-aware applications. *IEEE Pervasive Computing*, 3(4):66–74, 2004.

[16] H. Hagras, V. Callaghan, M. Colley, G. Clarke, A. Pounds-Cornish, and H. Duman. Creating an ambient-intelligence environment using embedded agents. *Intelligent Systems, IEEE*, 19(6):12–20, November 2004.

[17] K. Henricksen and J. Indulska. Modelling and using imperfect context information. In *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, pages 33–37. Ieee, 2004.

[18] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press, April 1992.

[19] J. Hong, E. Suh, and S.J. Kim. Context-aware systems: A literature review and classification. *Expert Systems with Applications*, 36(4):8509–8522, 2009.

[20] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, and Mike Dean. Swrl: A semantic web rule language combining owl and ruleml. Technical report, National Research Council of Canada, Network Inference, and Stanford University, 2004.

[21] Daniel Katz and Robert L Kahn. The social psychology of organizations. *The American Journal of Psychology*, 80(3):480, 1967.

[22] Kristian Ellebæk Kjær. A survey of context-aware middleware. In *Proceedings of the 25th conference on IASTED International Multi-Conference: Software Engineering*, SE'07, pages 148–155, Anaheim, CA, USA, 2007. ACTA Press.

[23] N. Krasnogor and J. Smith. A tutorial for competent memetic algorithms: model, taxonomy, and design issues. *Evolutionary Computation, IEEE Transactions on*, 9(5):474–488, 2005.

[24] M. Luck, P. McBurney, O. Shehory, and S. Willmott. *Agent technology: computing as interaction (a roadmap for agent based computing)*. University of Southampton, 2005.

[25] M.T. Maybury and W. Wahlster. *Readings in intelligent user interfaces*. Morgan Kaufmann, 1998.

[26] P. Moscato. A gentle introduction to memetic algorithms. In *Handbook of Metaheuristics*, pages 105–144. Kluwer Academic Publishers, 2003.

[27] J. Nehmer, M. Becker, A. Karshmer, and R. Lamm. Living assistance systems: an ambient intelligence approach. In *Proceedings of the 28th international conference on Software engineering*, pages 43–50. ACM, 2006.

[28] I. Nieto, J.A. Botía, and A.F. Gómez-Skarmeta. Information and hybrid architecture model of the OCP contextual information management system. *Journal of Universal Computer Science*, 12(3):357–366, 2006.

[29] J. Pérez, M. Arenas, and C. Gutierrez. Semantics and complexity of SPARQL. *The Semantic Web - ISWC 2006*, pages 30–43, 2006.

[30] L.R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[31] P. Remagnino and G. L. Foresti. Ambient intelligence: A new multidisciplinary paradigm. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(1):1–6, January 2005.

[32] G. Riva. Ambient intelligence in health care. *CyberPsychology & Behavior*, 6(3):295–300, 2003.

[33] G. Riva, F. Vatalaro, F. Davide, and M. Alcañiz. Ambient intelligence, 2005.

[34] E. Serrano, J.A. Botia, and J.M. Cadenas. Ubik: a multi-agent based simulator for ubiquitous computing applications. *Journal of Physical Agents*, 3(2), 2009.

[35] T. Strang and C. Linnhoff-Popien. A context modeling survey. In *Workshop Proceedings*, 2004.

[36] N. Streitz and P. Nixon. The disappearing computer. *Communications of the ACM*, 48(3):32–35, 2005.

[37] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung. Ontology based context modeling and reasoning using owl. *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, pages 18–22, 2004.

[38] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):66–75, September 1991.

[39] M. Weiser. Hot topics-ubiquitous computing. *Computer*, 26(10):71–72, 1993.

[40] Michael Weiss and Babak Esfandiari. Patterns for negotiating actors. In *European Conference on Pattern Languages of Programs (EuroPLoP*, 2005.

8.2 Publications

[41] J. A. Botia, E. Serrano, **T. Garcia-Valverde**, and A. Gomez-Skarmeta. Incremental deployment of large-scale AmI system by means of social models. In *International CAEPIA-TTIA Workshop on Simulation of Complex Social Systems (SiCoSSys 2009)*, Sevilla, Spain, November 2009.

[42] A. Caballero, **T. Garcia-Valverde**, J.A. Botia, and A. Gomez-Skarmeta. A trust and reputation model as adaptive mechanism for multi-agent systems. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, 13(42):3–11, 2009.

[43] F. Campuzano, **T. Garcia-Valverde**, A. Garcia-Sola, and J. A. Botia. Towards the standardization of ubiquitous computing simulations. In *Workshop on the Reliability of Intelligent Environments 2011 (WoRIE'11)*, Salamanca, Spain, April 2011.

[44] A. Garcia-Sola, **T. Garcia-Valverde**, and J. A. Botia. Reasoning on a semantic web based context-awareness middleware. *Trends in Practical Applications of Agents and Multiagent Systems*, pages 147–155, 2010.

[45] A. Garcia-Sola, **T. Garcia-Valverde**, and J. A. Botia. Distributed rule reasoning in a context-aware system by means of semantic web. In *Workshop on the Artificial Intelligence Techniques for Ambient Intelligence (AITAmI'11) on The 7th International Conference on Intelligent Environments (IE' 11)*, Nothingham, United Kingdom, June 2011.

[46] A. Garcia-Sola, **T. Garcia-Valverde**, F. Lopez-Marmol, and J.A. Botia. On the efficiency of semantic web based context computing infrastructures. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[47] J. J. Martinez, **T. Garcia-Valverde**, F. Campuzano, P. Campillo-Sanchez, A. Garcia-Sola, and J. A. Botia. *Ambient Intelligence and Smart Environments*, volume 12, chapter Multi-Agent based Social Simulation applied to Validation of Location Services. IOS Press, 2012.

[48] **T. Garcia-Valverde**, A. Caballero, J.A. Botia, and A. Gomez-Skarmeta. Semantic description of multimedia contents for the optimization of the advertising impact on TV program grids. In *IEEE International Joint Conference on Neural Networks (IJCNN 2010)*, pages 1–7, Barcelona, Spain, July 2010. IEEE. Core Rank A+.

[49] **T. Garcia-Valverde**, F. Campuzano, E. Serrano, and J. A. Botia. Human behaviours simulation in ubiquitous computing environments. In *Workshop on Multi-Agent Systems and Simulation at MALLOW*, Lyon, France, September 2010.

[50] **T. Garcia-Valverde**, F. Campuzano, E. Serrano, A. Villa, and J. A. Botia. Simulation of human behaviours for the validation of ambient intelligence services: a methodological approach. *Journal of Ambient Intelligence and Smart Environments*, 2012. In-press. Índice de Impacto JCR en 2010: 1.5.

[51] **T. Garcia-Valverde**, A. Garcia-Sola, and J.A. Botia. Improving RFID's location based services by means of hidden markov models. In *Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pages 1045–1046, Lisbon, Porto, August 2010. IOS Press. Core Rank A+.

[52] **T. Garcia-Valverde**, A. Garcia-Sola, and J.A. Botia. Predicting user behaviour for location services engineering in AmI environments. In *4th Symposium of Ubiquitous Computing and Ambient Intelligence (UCAmI 2010)*, Valencia, Spain, September 2010.

[53] **T. Garcia-Valverde**, A. Garcia-Sola, J.A. Botia, and Gomez-Skarmeta A. Automatic design of an indoor user location infrastructure using a memetic multi-objective approach. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, PP:1–6, April 2012. Índice de Impacto JCR en 2010: 2.105.

[54] **T. Garcia-Valverde**, A. Garcia-Sola, J.A. Botia, and A. Gomez-Skarmeta. Ubiquitous Deployment Configuration of Indoor Location Services. In *IEEE International Joint Conference on Neural Networks (IJCNN 2012)*, Brisbane, Australia, 2012. IEEE. In press. Core Rank A+.

[55] **T. Garcia-Valverde**, A. Garcia-Sola, Hagras H., Dooley J., J.A. Botia, Callaghan V., and A. Gomez-Skarmeta. An Adaptive Learning Fuzzy Logic System for Indoor Localisation using Wi-Fi in Ambient Intelligent Environments. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012)*, Brisbane, Australia, 2012. IEEE. In press. Core Rank A+.

[56] **T. Garcia-Valverde**, A. Garcia-Sola, H. Hagras, J. Dooley, V. Callaghan, and J. A. Botia. A fuzzy logic based system for indoor localisation using wifi in ambient intelligent environments. *IEEE Transactions on Fuzzy Systems*, 2012. Under review. Índice de Impacto JCR en 2010: 2.695.

[57] **T. Garcia-Valverde**, A. Garcia-Sola, F. Lopez-Marmol, and J. A. Botia. Engineering Ambient Intelligence Services by Means of MABS. *Trends in Practical Applications of Agents and Multiagent Systems*, pages 37–44, 2010.

[58] **T. Garcia-Valverde**, E. Serrano, and J. A. Botia. Social simulation for AmI systems engineering. *International Conference on Hybrid Artificial Intelligence Systems (HAIS 10)*, pages 80–87, 2010.

[59] **T. Garcia-Valverde**, E. Serrano, J. A. Botia, A. Gomez-Skarmeta, and Cadenas J. M. Social simulation to simulate societies of users inmersed in an ambient intelligence environment. In *1st IJCAI Workshop on Social Simulation (SS@IJCAI2009)*, Pasadena CA, USA, July 2009.

[60] **T. Garcia-Valverde**, E. Serrano, and J.A. Botia. Combining the real world with simulations for a robust testing of ambient intelligence services. *Artificial Intelligence Review*, pages 1–24, April 2012. 10.1007/s10462-012-9340-4. Índice de Impacto JCR en 2010: 0.429.