Context-aware Service Coordination for Mobile e-Health Applications

Federico Bergenti[1], Cesar Caceres[2], Alberto Fernandez[2], Nadine Fröhlich[3], Heikki Helin[4], Oliver Keller[5], Ari Kinnunen[6], Matthias Klusch[5], Heimo Laamanen[4], António Lopes[7], Sascha Ossowski[2], Heiko Schuldt[3], Michael Schumacher[8]

[1]FRAMeTech S.R.L.
Via San Leonardo 1, 43100 Parma, Italy
bergenti@frame-tech.it

[2]Artificial Intelligence Unit, University Rey Juan Carlos
Tulipan s/n., 28933 Mostoles, Spain
{cesar.caceres|alberto.fernandez|sascha.ossowski}@urjc.es

[3]Database and Information Systems Group. University of Basel
Bernoullistraße 16, 4056 Basel, Switzerland
{nadine.froehlich|heiko.schuldt}@unibas.ch

[4]TeliaSonera Finland Oyj
P.O.Box 970, 00051 Helsinki, Finland
{**heikki.j.helin|heimo.laamanen}@teliasonera.com**

[5]DFKI GmbH
Stuhlsatzenhausweg 3, 66123 Saarbruecken, Germany
{keller|klusch}@dfki.de

[6]EMA Group Ltd
Laivakatu 3, 00150 Helsinki, Finland
ari.kinnunen@ema.fi

[7]"We, the Body and the Mind" Research Lab of ADETTI
Av. das Forças Armadas, Ed. ISCTE, 1600-082 Lisboa, Portugal
antonio.lopes@we-b-mind.org

[8]Artificial Intelligence Laboratory, Swiss Federal Institute of Technology
1015 Lausanne, Switzerland
michael.schumacher@epfl.ch

Abstract: In this paper, we present a general architecture for service delivery and coordination in intelligent peer-to-peer (IP2P) environments that has been developed within the CASCOM research project. Our essential approach is an innovative combination of agent technology, Semantic Web Services, peer-to-peer, context-awareness, and mobile computing for intelligent peer-to-peer mobile service environments. Services are provided by software agents exploiting the coordination infrastructure to efficiently operate in highly dynamic environments. Our infrastructure includes efficient communication means, support for context-aware adaptation techniques, as well as dynamic service discovery and composition planning. For end users, the architecture provides seamless access to Semantic Web Services anytime, anywhere, and using any device. Our architecture is being evaluated using a sample ad-hoc emergency healthcare assistance application scenario. We deployed a prototype of an open IP2P service environment and expect results on methods for service provision, discovery, composition, and monitoring in mobile environments.

1 Introduction

The ever-growing number of services on the WWW provides enormous business opportunities. In particular, there is a huge potential for creating added value through service coordination. For this to happen, technology must be developed to be capable of pervasively providing and flexibly coordinating ubiquitous business application services to mobile users and workers in the dynamically changing contexts of open, large-scale and pervasive application domains.

One step toward the realization of this vision is the development of an intelligent agent-based peer-to-peer (IP2P) environment. IP2P environments are extensions to conventional P2P architectures with components for mobile and ad-hoc computing, wireless communications, and a broad range of mobile devices. Basic IP2P facilities come as Web Services, while their reliable, task-oriented, resource-bounded, and adaptive coordination-on-the-fly characteristics call for agent-based software technology. A major challenge in IP2P environments is to guarantee a secure spread of (personal) service requests across multiple transmission infrastructures and ensure the trustworthiness of services that may involve a broad variety of providers.

In this paper, we present a general architecture for service delivery and coordination in IP2P environments that has been developed within the CASCOM project. This architecture aims at providing support for business (semantic web) services for mobile workers and users across mobile and fixed networks. For end users, the architecture provides easy and seamless access to Semantic Web Services. This gives more freedom to mobile workers to do their job whenever and wherever needed. For network operators, it aims towards a vision of seamless service experience providing better customer satisfaction, which in turn helps to retain current customer relations as well as attract new customers. To service providers the architecture offers an innovative platform for various mobile business application services.

One of the foci of the project, in addition to the technical work, is to run detailed validation and trial activities in a real-world setting. For this, the area of medical emergency assistance has been chosen. Validation and trial is strongly supported by

CASCOM partner EMA (Emergency Medical Assistance) and TILAK, the umbrella organization of the state hospitals of the Austrian state Tyrol.

The paper is organized as follows. In Section 2 we propose some e-health real-world use case scenarios to illustrate the kind of domains that the proposed architecture aims at, and the types of services that it is to provide. Section 3 describes the structure of the CASCOM architecture and elements in further detail. Section 4 presents details on the trial and validation activities at TILAK and EMA which are currently being planned. Finally, Section 5 concludes.

2 e-Health Application Scenario

We have developed several use case scenarios in order to gain a deeper understanding of requirements, behaviour, and needs of mobile users and workers. Further, these use case scenarios are important to investigate detailed requirements for the conceptual architecture design, and the implementation and execution of agents and multi-agent systems in next-generation IP2P environments.

The healthcare domain is to be one of the most viable and growing application fields for intelligent mobile service coordination, not just for the technical requirements that it imposes, but also for its huge economic and social relevance. The specified e-health use case scenarios are (1) an emergency healthcare scenario, and (2) a telemonitoring and e-inclusion scenario. These two scenarios are outlined below, but detailed information about them can be found in [3].

Emergency Healthcare Scenario

The emergency healthcare scenario is based on the fact that people on the move, for example, travelling in foreign countries for business or holidays, may get into situations where they need medical assistance because of a sudden disease or emergency. Currently, these sorts of episodes are neither tackled nor realized in this form in practice and no software system is presently widespread in use to address them. Our emergency healthcare scenario consists of two stories. In the first story, a man on a business trip suddenly observes serious ailments with his stomach unknown to him. Fortunately he has installed the CASCOM system to his smart phone and he is directed to a sophisticated healthcare institution where our infrastructure is also available. In addition, our infrastructure allows for accessing information from the record of this person stored in his home country in order to avoid redundant an unnecessary examinations and even to negotiate on the coverage of costs with his health insurance. In the second story, tourists from Finland are having summer vacation abroad. During the trip, one of them seriously suffers from pain in the upper part of her body. Although she has installed our system to her PDA, our infrastructure is not fully supported at the present location, which changes the story significantly from the first one. However, even in this case, our system proves to be helpful. Essentially, although repeated examinations cannot be avoided, the availability of our system allows the patient for instance to contact another physician for a second opinion. In addition, a service provider in the home country can be contacted for organizing the transfer back home.

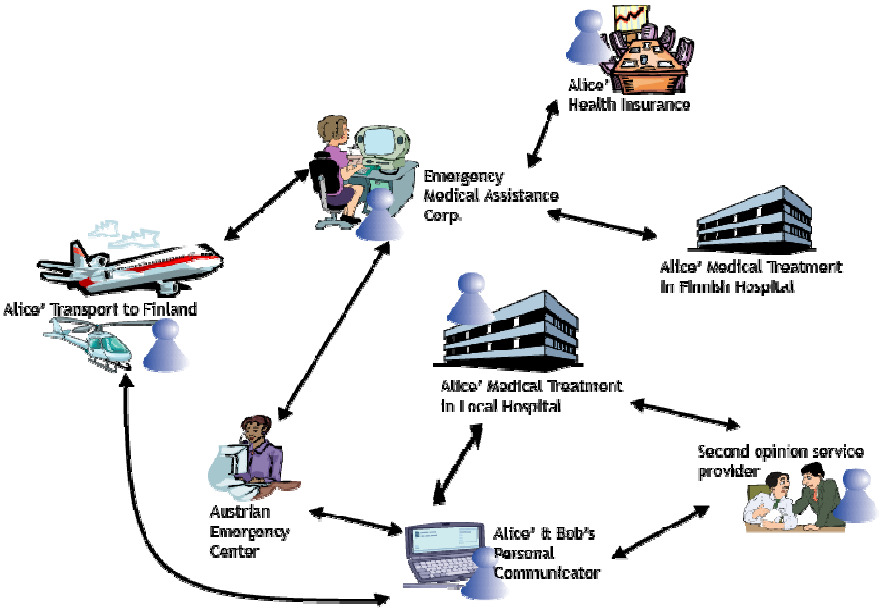

Figure 1: Emergency assistance scenario description

As results of a medical emergency, persons (patients) not only need medical treatment, they also need informational as well as sometimes transportation assistance either directly or indirectly. In the latter case, assistance in the form of information is usually required by the physicians, hospitals, or healthcare professionals involved. One implication of these complex requirements is the need for on-demand initiation, co-ordination, and supervision of various activities represented either through persons, or non-human actors (i.e., agents and services).

Telemonitoring and e-Inclusion Scenario

The telemonitoring and e-inclusion scenario relates to improving people's living conditions and to reduce the costs of long hospitalizations. In this scenario, the underlying idea is to allow elderly people, chronic patients or high risk patients to stay at home and to benefit from a remote and automated medical supervision. The scenario assumes the existence of universal communication infrastructures mainly used by mobile devices. The infrastructure in use has to provide a flexible platform for different monitoring applications and must guarantee a high level of reliability. A core task of monitoring systems in healthcare is to co-ordinate and process various data streams. These data streams are either automatically generated by different sensors or manually by user (patient) input. Data has to be collected, propagated, processed, and stored in a distributed manner. Intelligent personal and infrastructure agents are candidates to fulfil data processing. In addition, these agents not only have to detect critical situations from the sensor data streams but also to determine how to handle these situations (e.g., proposal of nearest healthcare institution based on the current location of a patient, etc.).

However, an important fact is that the system has to support intermitted connectivity, which means that disconnected parts of the system are still able to operate locally. This applies especially to all devices or agents in the patient's immediate vicinity.

Apart from the infrastructure for transferring different medical measurements to the place where decisions about medical treatments are made, the system should also support health care service providers to attend large numbers of patients efficiently. Throughout this scenario, connections between end users are normally characterized by a one-to-one communication. Thus, with every new patient, the efforts for the care-giver to provide thoroughly attendance increase and eventually become very extensive. Extending the communication pattern to a one-to-many communication allows the caregiver to build virtual groups of patients with similar characteristics, thus opening up new possibilities for medical attendance. It is important that this extension is compatible with the highly individual and personal physician-patient relationship, that is, this relationship should not be impaired. Instead, it should be an orthogonal add-on.

In both scenarios security is essential for technically preventing unauthorised persons from getting access to information. Privacy is a concept closely related to security. In a context-aware environment privacy means that a user has control and information over the information that is collected, stored and transmitted about her/him. Security and privacy are important in the CASCOM architecture.

3 CASCOM Architecture

The proposed architecture aims at the innovative combination of intelligent agent technology, Semantic Web Services, peer-to-peer, and mobile computing for intelligent peer-to-peer mobile service environments. For this purpose, conceptually, the architecture relies on a layered approach (see Figure 2). The CASCOM generic architecture comprises four components. The *Networking Layer* is an Intelligent P2P network that provides an efficient, secure and reliable communication independent of the access technology. The *Service Coordination Layer*, situated above the networking layer, provides flexible semantic service discovery, matchmaking, composition and execution functionalities. Orthogonal to both layers, the *Context Subsystem* is in charge of acquiring, storing and providing context information to both layers. *Security and Privacy Subsystem*, also orthogonal to the Networking and Service Coordination Layer, is responsible for ensuring security and privacy of information throughout the different components of the CASCOM infrastructure. The four main components of this architecture link the applications layer with the underlying networks and are described in more detail below.

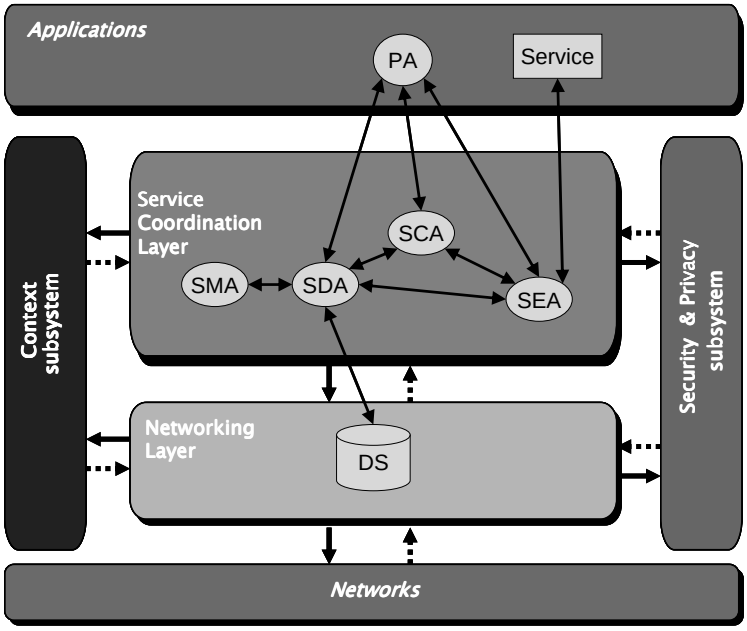

Figure 2: Layered model of the architecture
(PA: Personal Agent, SMA: Service Matchmaking Agent, SDA: Service Discovery Agent, SCA: Service Composition Agent, SEA: Service Execution Agent, DS: Directory Service)

Networking Layer

The Networking Layer is to provide a generic, secure, and open Intelligent P2P network infrastructure. It takes into account varying Quality of Service (QoS) of wireless communication paths, limitations of resource-poor mobile devices, and contextual variability of nomadic environments. It provides the following functionality:

- Efficient, secure, and reliable agent message transport communication over wireless (and wireline) communication paths independently of the access technology,
- Provision the context subsystem with network-related context information (e.g., QoS of a network, network availability, etc.),
- Low-level service discovery in IP2P environment, and
- Agent execution environment for resource-constrained mobile devices

Efficient agent message transport is an essential feature for our network architecture. Bit-efficient message encoding has been chosen due to the requirement for efficient communication over slow communication networks and resource limited devices. The simplicity and ease of implementation on small devices has been a key issue to choose FIPA-HTTP as the message transport protocol. We are aware that the performance of this protocol may be a concern, especially in wide area wireless networks (WWAN) which we address by using HTTP 1.1 persistent connections.

Context information about network/communication environment, including data about (available) networks and QoS of data communication can be used within the networking layer to adapt the CASCOM network to changing environmental situations.

Although the semantic service discovery in this infrastructure happens in higher layers of the architecture, some support for it is also needed in the networking layer. The service discovery in the networking layer considers mainly the "low-level" service lookup IP2P environments. Heterogeneous services from different providers are made public. A requestor searches dynamically for such service descriptions that match its needs. For that, the requestor interacts with a service directory in order to retrieve interesting published service descriptions. Directory services (DS) play a central role in order to link meta-services and functional services. Directories can be centralized or distributed, where a federation can be built. Services are represented as structured objects within the directory. They may be represented as OWL-S [8] (WSMO [10]) could also be used but we are not concentrating on it now). Entries are described in FIPA-SL0 [4] language because it is independent from the descriptions of the Web Services. Also SL0 is general enough and has a strong expressiveness.

Agents in our architecture need an agent platform that is usable in resource constrained devices. Among the different agent platforms considered, JADE/LEAP [1] is chosen as the most appropriate for CASCOM because it follows the FIPA standard and allows agents to be efficiently executed on small devices. Furthermore, as it is an active open source project, we were able to adapt it and overcome some limitations of JADE/LEAP. In particular, this agent platform expects infrastructure components (known as main containers) whose existence cannot be assumed, for example, in pure IP2P network architectures.

Service Coordination Layer

Setting out from the services of the networking layer, and based on the functionalities offered by both the Context and the Security & Privacy subsystems, the Service Coordination Layer takes an agent-based approach towards flexible Semantic Web Service discovery and coordination. Its main functionality is twofold:

- Semantic service discovery (service discovery + semantic matchmaking)
- Service coordination (service composition + execution monitoring and replanning)

In the proposed architecture, semantic service discovery functionality is realized by two different types of agents: Service Discovery Agents (SDA) and Service Matchmaking Agents (SMA). This was done for reasons of efficiency and flexibility, as in some application domains the matchmaking functionality may not be necessary. In much the same way, the service coordination functionality is realized by Service Composition Agents (SCA) and Service Execution Agents (SEA).

SDAs provide the means for the discovery of required services in the whole system network, considering contextual information. They are to handle abstract service descriptions, concrete service groundings, or both. Typically, SDAs receive service

specifications from the Personal Agent (PA) installed in the user's mobile device. SDAs then acquire relevant contextual information from the Context Subsystem and, using both the service specification and the acquired context information, they make use of the service discovery functionality of the networking layer and the semantic matching functionality of SMAs, to determine services (descriptions or providers) that fulfill the received service discovery request. As a result, SDAs are able to return a set of descriptions/providers and their correspondent service process model and/or grounding.

SMAs provide the means to compare service specifications in a context dependent fashion. Several semantic matchmaking approaches have been proposed [9] [13]. We use an OWL-S service matchmaker called OWLS-MX [6]. The OWLS-MX matchmaker takes any OWL-S service as a query, and returns an ordered set of relevant services that match the query each of which annotated with its individual degree of matching, and syntactic similarity value. The user may extend the query by specifying the desired degree, and syntactic similarity threshold.

SCAs are capable of creating value-added composite services that match service specifications. Once SCAs receive service specifications, they contact SDAs to discover existing services in a given domain, constrained to the current context, and plan a composite value-added service matching the received specification. The SCAs make use of the OWLS-Xplan [7]. Xplan is a heuristic hybrid search planner based on the FF-planner [5]. The generated composite service will orchestrate one or more simpler services. In a typical interaction, when no single service is found matching a given service specification, the service composition functionality is used to create a composite service matching the service specification, the output of which is a service description. These service descriptions may be stored in some directory for later use.

SEAs manage the execution of composite value-added service descriptions generated by SCAs. Since the received compound service description relies on simpler services, the execution will also coordinate the execution of these simpler services. Whenever necessary, SEAs will use SDAs to discover appropriate available service providers for each of the simpler services evoked from the compound service description. The execution model is based on principles of the OSIRIS process management system [11].

Context Subsystem

The context subsystem, orthogonal to the previously described layers, is in charge of acquiring, storing, and providing context information to both layers. Generally speaking, there will be contextual information for each of the system components. Each of them will be able to acquire it using the following set of functionalities:

- Discovery and acquisition of context information
- Subscription of context listeners and acquisition of context events/changes in the environment
- Access to context information repository (representing historical context information)

However, the context subsystem relies on the other layers to obtain the context information, thus working as a gateway of context information between layers.

Security and Privacy Subsystem

The Security and Privacy Subsystem, also orthogonal to the Networking and Service Coordination Layer, is responsible for ensuring security and privacy of information throughout the different components of the infrastructure. One of the main things we need to protect is the information (data) that every node of the network maintains. In detail, data confidentiality, integrity, and availability (CIA) [12] are topics of concern that any approach to security must address.

Security concerns not only data, but also the software that deals with the data. This holds especially for network-centric systems. Within the realm of computer security, misuse, theft, and unauthorized usage of computing resources are well-studied phenomena. The security and privacy requirements identified are: identification, authentication, authorization, single sign-on, as well as local and network security. We also need to guarantee the integrity of transmitted data, non-repudiability, traceability, privacy, delegation, and nationalization. In order to guarantee the correct treatment of data, we introduce two novel architectural abstractions [2]: Validation-Oriented Ontologies (VOO) and Guarantors. A VOO as a signed set containing: (i) An ontology that models a domain; and (ii) A set of runtime tools capable of asserting properties of individuals of this ontology. Guarantors are agents that, in some sense, play the role of middleman in interactions. Guarantors are trusted by all interacting parties and they are in charge of supporting interactions by providing (under their responsibility) all necessary VOOs.

4 Trial and Validation of the CASCOM Infrastructure

The aim of the trial is to evaluate the CASCOM architecture and its implementation by using real network services, resources, devices, and terminals.

To show and prove the CASCOM architecture and the benefits of service coordination on the top of an intelligent agent-based peer-to-peer network a demonstrator will be built and evaluated. The full fledged demonstrator will be based on the CASCOM IP2P platform, encompass handheld devices and operate in a wired as well as wireless environments and support context-awareness in the coordination of services. The objective is not only to prove the CASCOM architecture and its implementation but evaluate the acceptance of the system and the suitability of the selected technology in the chosen application domain. Moreover it will be verified whether the CASCOM service coordination framework meets the business needs of multiple service providers and network operators in a concrete setting and application.

As application to be considered, we have chosen the emergency assistance scenario briefly introduced in Section 2. In particular, it is planned to concentrate in the trials to patients with heart failures.

An important prerequisite of the trial is the specification of the requirements (functional and non functional) the application has to meet. Security and privacy functionality has to be taken into account. Essentially, it leads to a refinement of the general application scenario which will be tailored to the partners that support the trials, the environment they have available and the data and services relevant for them.

Functional Requirements

The use cases developed for the emergency assistance scenario serve as a starting point for the detailed trial plan. From the functional point of view, the scenario has to be linked to the concrete environment available. The following list of services that need to be made available includes in the different steps that can be identified in the application scenario:

1. Tracking the location of the patient; directing the patient to the nearest healthcare center. These services are used by the patient when being out of hospital (e.g., a tourist traveling abroad)
2. Giving the treating emergency physician access to the relevant part of the medical history of a patient. In our concrete setting (emergency patients with heart failures), this means giving access to:
   - Previous medications
   - Previous diagnoses
   - Previous laboratory results
   - ECG data
   - Medical images (e.g. X-Ray data)

   These services are made available to the treating emergency physician, either in the emergency department of a hospital which is equipped with WLAN technology or to an emergency physician who is equipped with a UMTS device in an ambulance car
3. Checking whether the treatment is covered by the local health insurance; providing information on the health insurance to the treating healthcare organization; organization of transport back to the home country. These services are provided by organizations that support patients being sick abroad (like the EMA Group in Finland)
4. Services for post-treatment care in the home country. This again includes access to the relevant data both in the home country (medical record) and in the country where a patient has been treated.

Non-functional Requirements

In the trial activities, the following non-functional requirements are being considered:

- **Easy to use and ergonomic UI**: Complex and important decisions need to be made in a timely manner on the basis of potentially large volumes of data or data in different formats (e.g., ECG data, previous medications, etc.). Therefore, it is

crucial, especially in step 2, to have interfaces people can easily work with. This point is much more important than in other application areas.

- **Performance**: Service execution (and therefore data access by means of services) has to be at least faster than the manually handling of a certain situation. It is to consider that in emergency situations, the time is an important factor and it is not possible to have longs delays e.g., due to bootstrapping the system, load the software, and initialize connections.
- **Robustness, Availability, Reliability**: Recurring failures endanger people's health or life directly. Therefore, the system has to contain methods that guaranty execution of actions even if parts of the system are inoperable and it has to detect and recover from failures. That is very challenging especially in mobile environments.
- **Security and Trustworthiness**: Because patient's data are very sensible, no unauthorised recipient should have access to this data.
- **Adaptability and Scalability:** From the network service provider's point of view the architecture must be open to any network service provider because it is almost impossible to make prognoses about who wants to join the network as shown in the scenario. Besides, the architecture must be able to scale to serve any number of users and services.

Trial Details

In addition to CASCOM partner EMA, the Tiroler Landeskrankenanstalten GmbH (TILAK) will support the trial from a healthcare organization's perspective. Emergency physicians (one from the emergency department of TILAK and an emergency physician in the ambulance car) will test the application from the physician's point of view. From the healthcare systems (clinical information systems) point of view, the trial is supported by the IT department of TILAK which allows accessing anonymized patient data via the health@net system (a regional network of healthcare providers in Tyrol) from within CASCOM services. Additional activities, according to steps 1-4 identified above, will take place in Finland, hosted by the EMA Group.

From a network perspective, the WLAN which is available in the emergency department of TILAK will be used. In addition, for providing access to physicians when being in an ambulance car, UMTS will be considered as well. From a hardware perspective, the different users in the trial will be equipped with PDAs.

The actual trial phase is planned to last about two months and will start in spring 2007. Feedback on the trial activities will be given in the form of interviews. Therefore, convenient questionnaires are to be developed to get comparable results.

5 Conclusions

In this paper, an innovative generic architecture has been defined supporting intelligent peer-to-peer mobile service environments. This architecture has been described in detail and a scenario of emergency assistance in the healthcare domain was proposed. Over this it is described how the architecture and the belonging prototype are evaluated.

Until now, we have implemented a prototype based on the presented architecture, focusing on the emergency assistance scenario described. At the end of the project, we will come up with a fully fledge demonstrator for the emergency assistance scenario.

References

[1] Bergenti, F.; Poggi, A.: LEAP: A FIPA Platform for Handheld and Mobile Devices, Intelligent Agents VIII, Springer, 2002, pages 436–446.

[2] Bianchi, R.; Fontana, A.; Bergenti, F.: A Real-World Approach to Secure and Trusted Negotiation in MASs. Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (AAMAS). 2005

[3] CASCOM Consortium. Deliverable D3.1: Use Case Scenarios. http://www.ist-cascom.org. 2005

[4] Foundation for Intelligent Physical Agents. FIPA SL Content Language Specification. Geneva, Switzerland, Dec. 2002. Specification number SC00008I.

[5] Hoffmann, J.; Nebel, B.: The FF Planning System: Fast Plan Generation through Heuristic Search. Journal of Artificial Intelligence Research (JAIR) (14): 2001, pages 253–302.

[6] Klusch, M.; Fries, B.; Khalid, M.; Sycara, K: OWLS-MX: Hybrid Semantic Web Service Retrieval. Proceedings 1st International AAAI Fall Symposium on Agents and the Semantic Web, Arlington VA, USA, November 2005

[7] Klusch, M.; Gerber, A.; Schmidt, M.: Semantic Web Service Composition Planning with OWLS-Xplan. Proceedings 1st International AAAI Fall Symposium on Agents and the Semantic Web, Arlington VA, USA, November 2005.

[8] OWL-S Home Page. http://www.daml.org/services/owl-s/

[9] Paolucci, M.; Kawamura, T.; Payne, T.; Sycara, K.: Semantic matching of web services capabilities. In Proceedings of the First International Semantic Web Conference on The Semantic Web. Springer-Verlag, 2002, pages 333-347.

[10] SDK WSMO working group http://www.wsmo.org/.

[11] Schuler, C.; Weber, R.; Schuldt, H.; Schek, H.-J.: Scalable Peer-to-Peer Process Management – The OSIRIS Approach. In Proceedings of the 2nd International Conference on Web Services (ICWS'2004), San Diego, CA, USA, July 2004, pages 26–34. IEEE Computer Society

[12] SecPedia. http://www.infosecpedia.org/pedia/index.php/Main_Page

[13] Sycara, K.; Widoff, S.; Klusch, M.; Lu, J.: LARKS: Dynamic Matchmaking among Heterogeneous Software Agents in Cyberspace. Kluwer Academic Press, 2002.