

The Changing Role of Pervasive Middleware: from Discovery and Orchestration to Recommendation and Planning

Gabriella Castelli, Marco Mamei, Franco Zambonelli

DISMI – Università di Modena e Reggio Emilia – Via Amendola 2 – Reggio Emilia – ITALY
{gabriella.castelli, marco.mamei, franco.zambonelli}@unimore.it

Abstract

Future pervasive computing scenarios will be characterized by an increasing diversity and dynamics of services and of contextual data sources, and by an increasing exploitation of crowdsourcing for social sensing and human computation. Accordingly, the role of middleware should no longer be limited to facilitating interactions and compositions via discovery and orchestration, but should approach that of a recommendation engine capable of dynamically and adaptively planning patterns of service interaction and composition on a best-effort basis. Along these lines, this position paper firstly elaborates on the limitations of traditional middleware infrastructures in meeting the new requirements of the emerging pervasive computing scenarios. Then, it introduces two case study scenarios to motivate and clarify the concepts expressed. Finally, it identifies some key research challenges for future pervasive middleware infrastructures.

Keywords: *Middleware, Human Computation, Crowdsourcing, Recommendation, Planning, Privacy*

1. Introduction

The pervasive computing scenario envisioned by Mark Weiser is here to come. For it to become practical, though, it is fundamental to have general middleware infrastructures capable of effectively supporting interactions among many heterogeneous components and devices, while hiding the many complexities inherent in pervasive applications and services.

For instance, a large portion of pervasive middleware research so far has focuses on issues such as: (i) providing effective, possibly semantically-enriched, discovery services for resource lookup, specifically tuned to the physically-situated characteristics of pervasive services [Ton08]; (ii) defining models and tools to promote context-awareness and adaptation to context by services [Cas08]; (iii) service orchestration for enabling

resources sharing and composition [Lag10]; (iv) addressing non-functional issues such as security, privacy, and energy consumption [Maj10]. Without minimizing the relevance and impact of such large body of work, our opinion is that a number of additional very relevant issues arise when considering how the pervasive computing scenario is evolving.

We are assisting to the increasing availability of pervasive devices that are capable of providing services and of generating contextual information. This trend is exacerbated by two specific tendencies.

First, as pervasive computing is getting more and more integrated in our everyday physical and social worlds, the very concept of “service” is evolving, to include physical resources and services (e.g., a parking lot, or a seat in a restaurant can be regarded as physical resources/services), which may have characteristics and availabilities very different from traditional hardware and software services.

Second, while the power of involving humans via participatory sensing models (i.e., users making available their own sensing devices on need) is already recognized [Das10, Alt10, YueCK09], some recent work on crowdsourcing let us envision that humans and their peculiar sensing, actuating, and computing capabilities become primary components/devices of pervasive computing scenarios [Sha10]. Again, in these cases, the recruitment of humans in the context of pervasive applications and the orchestration of their activities in mixed ICT-human services exhibits peculiar characteristics that can be hardly supported by traditional middleware models and architectures [Red10].

In this position paper we try to elaborate on the idea that, to tackle the challenges risen by this scenario, the traditional role of pervasive middleware should change. Such role should shift from simply supporting service discovery, context-awareness, and orchestration, to acting as a recommendation engine (supporting components and services in finding what they need to operate) and being also capable of dynamically planning the resulting interactions depending on current needs, characteristics, and availability of service components.

Accordingly, the key contributions of this paper are:

- To discuss the limitations of the current approaches for pervasive middleware and identify the key changing requirements (Section 2)
- To introduce two representative case studies to better ground the ideas and to analyze some promising approaches (Section 3);
- To identify some key research challenges for future pervasive middleware research (Section 4).

Eventually, Section 5 concludes.

2. Limitations and Changing Requirements

Based on the sketched general scenario, we now focus on some key limitations of current approaches, and on the related changing requirements that must be tackled by future pervasive middleware.

Discovery. Traditional solutions, based on matching property listings, will be increasingly ineffective. First, the number of available resources and services matching a lookup request can become overwhelming (e.g., a request for sensing noise in a street can be associated to hundreds of smart phones and dozens of embedded microphones). It is unthinkable to leave to applications the burden (and the costs) of selecting among many possible choices. Second, it may be difficult for applications to select among nearly equivalent services and resources that differ for e.g., their spatial location, the dynamics of the service, response times, the involvement of human-service-providers. These may require adopting very long property lists subject to frequent and costly updates, in order to describe them. Third, service components in an unknown environment may be unaware of how to specify exactly what they want, and of all the possibilities and constraints there existing. Accordingly, service and resource discovery will need more flexibility and intelligence than that provided by current semantic property-based approaches. It should include fuzzy matching and matching against incomplete or incorrect specification, and should integrate ranking algorithms to reduce a possible long list of options. In this perspective, service discovery will have to evolve into a sort distributed recommendation engine, relieving services from the complexities involved in specifying very complex constraints and then in choosing among many possibly options, and able to learn from history what can better suit in a given environment.

Situation awareness. Effective applications and services require high level of awareness about the context in which they operate. Also in this case, though, the amount of information available about the context can become so large, and account for so many useless

(from the service/application viewpoint) facets of the context to be not informative *per se*. This may force applications and services to engage in complex information selection and reasoning to get the needed clue out of it [Cas08]. Accordingly, the middleware should take charge of digesting such large amount of contextual information, elaborating and aggregating them in order to generate comprehensive and sound view of specific situations. And eventually, in a way similar to the one we envision for discovery services, to recommend and make available to services those specific – compact yet expressive and ready to use – views about the context that may fit specific application scenarios, hiding irrelevant data.

Orchestration. Also composition and orchestration of services will have to account for a very large number of diverse services and resources to be potentially involved in complex composite applications. For instance, multiple instances of an application A may require services of type $T1...TN$ and resources of type $R1...RM$ to be composed and orchestrated together. The issue then is not finding the needed resources and services, since many can be available. Rather, the issue is planning which combination of services and resources, among the many of different nature (e.g., human or ICT provided) and of different efficiency and availability (closer or farther, ready to use or to be recruited), to recommend to A , also based on the specific context of each instance of A and on the possibly incomplete specifications by A . The middleware should become capable of dynamically planning the most proper composition based on what it knows about the multiple and possible contrasting goals of A and on the current situation. In doing so, the middleware should also orchestrate the overall execution accounting for the intrinsic dynamics of services and resources and the unpredictability of humans if involved, possibly relying on what it has learnt about the past and it envision about the future.

Security and Privacy. While security has and will always played a relevant role in pervasive middleware research, the involvement of physical resources (typically owned by some human) and of humans themselves (either to make their ICT resources available or to directly play an active sensing, actuating, or computing role) will make privacy a critical success factor. People will contribute to pervasive services (by making available their personal devices, some privately owned physical resource, or their own work) only if this will not unveil some private sensible information about them (e.g., a person making available its microphone to sense noise in a street may not wish that its current location is made public). The middleware should be able to perform resource management, recommending

services and planning their orchestration, also accounting for the specific privacy needs of the components involved. For instance, the middleware should be ready to dynamically play the role of mediator in service interactions whenever such interactions involves human activities that has to remain anonymous (as, e.g., enforced by the Amazon Mechanical Turk crowdsourcing tools).

3. Case Study Scenarios

We now introduce two clarifying scenarios to exemplify the concept expressed, and to show that the identified requirements have been faced by some other proposals as application-specific issues (and not as general middleware features, as we think they should).

3.1. E-mobility and Parking Management

Technologies for vehicular sensing, traffic monitoring, adaptive traffic lights and traffic signs are paving the way for a new paradigm of traffic management in urban environment [Lee09, Bie10]. This “e-mobility” scenario will support adaptive traffic routing strategies, facilitate the adoption of multimodal mobility model, and effectively support transit sharing.

In this scenario, a specific application problem, to which a lot of research effort has been already devoted, is the research and reservation of available parking spaces. Finding an available parking lot can be stressful, time consuming and can lead to traffic congestion and delays. Thus, one can think at enriching a car navigation service with support for finding suitable parking spaces.

This is indeed a challenging problem that exhibits many of the characteristics of the envisioned future pervasive scenarios. First, the application involves the use of physical resources, inherently dynamic in terms of availability. It may also consider the possibility of dynamically recruiting parking spaces via crowdsourcing (a privately own parking made available by the owner for some time slots). Second, the discovery and selection of appropriate parking spaces among a multitude have to account for several parameters and constraints.

Let us now analyze the case study with respect to the identified requirements.

Discovery. Traditional discovery cannot easily apply to the finding of parking space, since there can be a wide number of possibilities that can match the drivers’ request (or, vice versa, there could be no one fully satisfying but a number of minimally unsatisfactory), so the support of the middleware in selecting and recommending the best solutions plays an important role. Also, one should consider that a driver might not know

how to express his preferences, either because detailing too much is not practical or simply because he has no idea of what’s available in a city. Therefore, it is not surprise that some existing proposals in this application area already suggests adopting recommendation engines. For instance, in the e-mobility scenario proposed in [Cho08], a multi agent system for parking negotiations is proposes that – by taking into account few relevant driver preferences – selects, i.e. recommends, the best matching parking to the driver and the route to reach it.

Situation awareness. This involves continuous gathering and processing of data about car locations, road conditions, and traffic flow, to properly route with accurate schedule towards the available parking space. However, many other data related to situations happening (or likely to occur) in the urban environment could be fruitfully taken into account. For instance, this is the case of a parade taking place close to where the user desires parking. Many proposals attribute the importance of context-awareness in suggesting parkings, but most of them accounts for reaching awareness directly related to the current [Mat10] or likely future [Kla10] situation of parking slots, without accounting for higher and more comprehensive levels of situation-awareness.

Orchestration. When recommending a parking lot to a car, there could be a number of factors and constraints to be taken into account, and possibly a number of activities to be triggered. For instance, to reach on time a reserved parking lot there may be need of properly orchestrate the actuation traffic lights and signs along with the car navigation systems. In doing so, the system should try to accommodate the competing needs of a multiplicity of drivers, and possibly account for more global situation such as traffic congestion, pollution, and the maximization of the parking space occupation. That is, it requires planning a complex orchestration of components. A proposal going in that direction can be found in [Del09], where an orchestrated reservation of parking spaces (computed via ad hoc vehicular networks) is planned by accounting several parameters in the attempt of maximizing both the single driver experience and the overall network congestion.

Privacy. Issues of privacy arise in this scenario because its realization requires the availability of the position and destinations of vehicles. In addition, they arise when a user offers his privately owned parking when not used. For both cases, the middleware should guarantee that the information about users’ habits is properly protected, possibly preventing those direct access to sensible data. Also, it should possibly adopt further strategies to prevent more subtle ways of unveiling users’ habits (e.g., recommending multiple time to the same car and in the same periods of the day

the same privately owned parking lot). The issue of enforcing privacy in e-mobility system has been extensively faced, and some solutions are available. [Hoh08], for instance, introduces the concept of virtual trip lines (geographical markers that triggers the cars' location update) to mediate interactions between vehicles. Yet, more general solutions are to be investigated to support all the intricacies of the scenario.

3.2. Participatory Sensing Campaigns

Another important area of development for pervasive applications is the exploitation of the mobile sensing capabilities provided by smartphones via crowdsourced (or participatory) sensing [Das10, Alt10], an approach that will extend in the future to include the exploitation of human specific sensing, actuating, and computing capabilities (as, e.g., in the DARPA network challenge [Pic10], for which crowdsourced teams of people were asked to find out the locations of ten weather balloons placed somewhere in the USA).

In general, the proliferation of smartphones with heterogeneous sensing capabilities, makes them very useful components of pervasive infrastructures. In fact, they can be exploited to acquire information about an environment, or perform targeted research campaigns on need, without having to deploy costly sensing infrastructures and rather relying on the likely availability of persons with smartphones to support the required sensing needs. Independently of whether such exploitation is participatory (i.e., requiring the active intervention of smart phone owners) or opportunistic (i.e. exploiting smart phone capabilities transparently to their owners), the people-centric nature of this form of sensing raises, once again, peculiar challenges to pervasive computing applications. These challenges relates the need of involving and orchestrating a multitude of human actors that are not under the control of any authority, that express different degree of willingness in participating, and that have devices with different sensing capabilities.

Let us now analyze the case study with respect to the identified requirements.

Discovery. In this scenario, discovery involves identifying those users that can contribute to a specific sensing task, typically based on geographic co-location and temporal availability, and possibly based on additional requirements (e.g., the resolution of the smart phone camera) depending on the specific sensing task. However, it is not always possible or feasible to rely on real-time data about users location and sensing possibilities as the number of potential users to involve can be very large in a dense urban environment. If all the

users at a given location are recruited to provide sensed data, the risk is in receiving back too much data most of which unusable [Der10]. On the opposite, both to support the lack of data and to discriminate among which users to involve, it makes sense to study the typical users behavior in order to rely on users that are likely to be fulfill the application requirements. For instance, [Alt10, Red10] propose recruitment frameworks relying on the known past behavior of participants, there included their past sensing contributions, to discriminate among many potential choices. Again, this configures the need for a middleware to support crowdsourced sensing by properly elaborating recommendation on the users to hire for specific tasks.

Situation awareness. It is crucial in this application scenario to understand the situations in which the users are involved and their ability/willingness to participate. Reaching high levels of situation awareness is fundamental to support applications, and also to support the middleware itself in its recommendations and planning activities. With this regard, the large body of work in activity recognition via monitoring of smart phone usage [Mil08, Luk10] can be relevant.

Orchestration. Let us consider the case of sensing campaigns, where snapshot of a large area has to be built with regard to specific aspects and with time constraints (e.g., mapping the sound levels of a town at midnight), and requiring the involvement of a large number of users. There, orchestrated involvement of users must take place, which may require careful global plans. Such plans should evaluate how many users to involve, with which density, and with which accuracy to conduct the study. It should also account to the different situations in different regions of the area, and the possible existence of budget constraints, in the case economic incentives are provided to promote participation and recruitment [Pic10]. A very extensive discussion on the issues involved in panning and executing sensing campaign is reported in [Red10]. Again, our claim is that these – so far application-specific – kinds of planning activities should become general mechanisms of the middleware infrastructure.

Privacy. According to many [Kap09, Shi10], privacy issues are a key limitation in the full exploitation of crowdsourcing sensing. Users – in many cases and for many diverse personal reasons – may be willing to participate only being guaranteed anonymity. In this case the middleware should play a major role in acting as a mediator between users and services to ensure privacy. Most of current proposals with this regard address the issue via various forms data aggregation [Hua10, Shi10] or data obfuscation [Gan08]. However, also in this application scenario, we think there is need for strategies

to tackle privacy issue in a more general way.

4. Research Challenges

For middleware to embrace the identified changing requirements, the following challenges (among many we may have not identified) have to be faced.

A common model for representing generalized services and data. There is the need for a model that allows to effectively represent in a uniform way both hardware, software and physical resources, and human services. This model should avoid complex and highly-structured formats that would be difficult to be encoded and maintained, and should rather prefer expressive and flexible formats. For instance, the use of pragmatic (i.e., tag-based) ontologies to encode such diverse information could support both an effective creation of such descriptions in the first place and an effective use by the middleware [Ram07].

Algorithms for distributed recommendation. In future pervasive computing scenarios, it will be necessary to identify general-purpose algorithmic approaches for ranking available services and resources in order to both recommend only the best matching solution(s) and to integrate them to provide effective combined services. On the one hand, since having a single mean for properly recommending in a variety of scenarios is illusionary, such algorithms should define sort of general mechanisms flexibly and easily tunable to application-specific needs. On the other hand, given the inherent decentralization of pervasive scenarios, they should be capable of handling distributed data processing and be in themselves distributed [Orm10].

Algorithms for planning and orchestration. In the above scenarios, it is important to develop novel approaches to organize, plan and coordinate the interactions among services and resources. In particular, autonomic algorithms to dynamically assign distributed resources in cloud computing architectures provide interesting research guidelines in this direction [Pat09]. However, pervasive computing scenarios will be highly dynamic, with very diverse services and applications running in parallel and competing to accessing shared resources. To this end multi-criteria and multi-constraints planning mechanisms should be taken into consideration. Moreover, planning and orchestration will have to deal with the specific issues arising with services carried on by human users, such as long latency and inaccuracy. Again, the idea of having general mechanisms tunable to specific needs stands.

Algorithms for reasoning and learning from the context. Strictly related to the above challenge, there is the need for the middleware to understand what is

happening in the environment in order to provide an effective representation of situation to be used both for the use of applications that want to access contextual data from a high-level perspective and for the use of the middleware itself to perform recommendation and planning. In particular, novel data mining algorithms to deal with such heterogeneous and large-scale data coming from distributed devices should be created. Moreover, data mining algorithms and partial results should be integrated in a network of knowledge allowing them to take advantage of each other [Cas08].

Privacy-aware strategies for resource management and planning. The middleware should be in charge of dealing with privacy issues guaranteeing that resources manipulation and orchestration do not reveal the identities of services and human involved (unless explicitly approved), and possibly taking care of mediating interactions. At the same time, it should take care of hiding all the complexities of managing privacy from applications and users while enabling mutual discovery and interactions. For this reason, the mechanisms run by the middleware (such as recommendation and planning) should integrate privacy-aware strategies to deal with these issues effectively. Research on mechanisms enabling interactions while preserving such sensible data will be an important avenue for research [Shi10].

Incentive programs for boosting participation. Location-based crowdsourcing relies on the willing of all available components (devices, services, human actors) to put themselves at the service of the collectivity and to interact opportunistically and effectively with each other. Although in some cases it can be done on a voluntary basis [Der10], usually some sort of incentives should be designed to ensure users participation. Some interesting research directions in this area are based on auction mechanisms as proposed in [LeeH10], and on recursive mechanism to better support distributed recruitment [Pic10]. Not to exclude the necessity of defining mechanisms specifically tuned to the peculiar economics of participatory sensing and crowdsourcing.

A middleware for future pervasive scenario will have to integrate the above features yet allowing access via a simple API and also by resource-constrained devices.

5. Conclusions

Future pervasive computing scenarios will be increasingly complex, involving a myriad of devices, providing very elaborated services, and integrating human capabilities as an active part of the infrastructure. In these emerging scenarios, the role of pervasive middleware will have to evolve, to provide flexible

general-purpose support for recommendation-based discovery, advanced situation-awareness, planning of complex orchestrations, and privacy-aware strategies.

The issues sketched in this position paper are the current focus of the EU-funded project “SAPERE: Self-aware Pervasive Service Ecosystems” (www.sapere-project.eu).

Acknowledgments. Work supported by the FP7 project SAPERE (Grant No. 25873), FET Program of the EU.

References

- [Alt10] F. Alt, A. S. Shirazi, A. Schmidt, U. Kramer, Z. Nawax, “Location-based Crowdsourcing: Extending Crowdsourcing to the Real World”, NordiCHI Conference, Reykjavik (IC), 2010.
- [Bie10] A. Biem, E. Bouillet, H. Feng, A. Ranganathan, A. Riabov, O. Verscheure, H.N. Koutsopoulos, M. Rahmani, B. Güç, “Real-Time Traffic Information Management using Stream Computing”, IEEE Data Engineering Bulletin, 33(3):64-68, 2010.
- [Cas08] G. Castelli, M. Mamei, F. Zambonelli, “Engineering Contextual Knowledge for Autonomic Pervasive Services”, Information and Software Technology, 50(1-2):36-50, 2008.
- [Cho08] S.-Y. Chou, S.-W. Lin, C.-C. Li, “Dynamic Parking Negotiation and Guidance Using an Agent-based Platform”, Expert Systems with Applications, 35(3):805-817, 2008.
- [Das10] T. Das, P. Mohan, V.N. Padmanabhan, R. Ramjee, A. Sharma, “PRISM: Platform for Remote Sensing Using Smartphones”, MobiSys Conference, San Francisco (CA), 2010.
- [Del09] T. Delot, N. Cenerario, S. Ilarri, S. Lecomte, “A Cooperative Reservation Protocol for Parking Spaces in Vehicular Ad-hoc Networks”, Mobility Conference, New York (NY), 2009.
- [Der10] M. Demirbas, M.A. Bayir, C.G. Akcora, Y.S. Yilmaz, H. Ferhatosmanoglu, “Crowd-sourced Sensing and Collaboration Using Twitter”, WOWMOM Symposium, Montreal (CDN), 2010.
- [Gan08] R.K. Ganti, N. Pham, Y.-E. Tsai, T.F. Abdelzaher, “PoolView: Stream Privacy for Grassroots Participatory Sensing”, SenSys Conference, New York (NY), 2008.
- [Hoh08] B. Hoh, M. Gruteser, R. Herring, J. Ban, D. Work, J. Herrera, A.M. Bayen, M. Annavaram, Q. Jacobson, “Virtual Trip Lines for Distributed Privacy-preserving Traffic Monitoring”, MobySys Conference, Breckenridge (CO), 2008.
- [Hua10] K.L. Huang, S.S. Kanhere, W. Hu, “Preserving Privacy in Participatory Sensing Systems”, Computer Communications, 33(11): 1266-1280, 2010.
- [Kap09] A. Kapadia, D. Kotz, N. Triandopoulos, “Opportunistic Sensing: Security Challenges for the New Paradigm”, COMSNET Conference, Piscataway (NJ), 2009.
- [Kla10] A. Klappenecker, H. Lee, J.L. Welch, “Finding Available Parking Spaces Made Easy”, 6th International Workshop on Foundations of Mobile Computing, Cambridge (MA), 2010.
- [Lag10] B. Lagesse, M. Kumar, M. Wright, “ReSCo: A Middleware Component for Reliable Service Composition in Pervasive Systems,” Percom Conference, Mannheim (D), 2010.
- [Lee09] U. Lee, M. Gerla, M. 2010, “A Survey of Urban Vehicular Sensing Platforms”, Computer Networks, 54(4):527-544, 2010.
- [LeeH10] J.-S. Lee, B. Hoh, “Dynamic Pricing Incentive for Participatory Sensing”, Pervasive and Mobile Computing, 2010, 6(6): 693-708, 2010.
- [Luk10] P. Lukowicz, O. Amft, D. Roggen, J. Cheng, “On-Body Sensing: From Gesture-based Input to Activity Driven Interaction”, IEEE Computer, 43(10):92-96, 2010.
- [Maj10] V. Majuntke, G. Schiele, K. Spohrer, M. Handte, C. Becker, “A Coordination Framework for Pervasive Applications in Multi-User Environments”, IE Conference, Kuala Lumpur (MAL), 2010.
- [Mat10] S. Mathur, T. Jin, N. Kasturirangan, J. Chandrasekaran, W. Xue, M. Gruteser, W. Trappe, “ParkNet: Drive-by Sensing of Road-side Parking Statistics”, MobiSys Conference, New York (NY), 2010.
- [Mil08] E. Miluzzo, N.D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S.B. Eisenman, X. Zheng, and A.T. Campbell, “Sensing Meets Mobile Social Networks: the Design, Implementation and Evaluation of the CenceMe Application”, SenSys Conference, New York (NY), 2008.
- [Orm10] R. Ormándi, I. Hegedüs, M. Jelasity: “Overlay Management for Fully Distributed User-Based Collaborative Filtering”, Euro-Par, Naples (I), 2010.
- [Pat09] N. Paton, M. Aragão, K. Lee, A. Fernandes, R. Sakellariou, “Optimizing Utility in Cloud Computing through Autonomic Workload Execution”, IEEE Data Engineering Bulletin, 32(1):51-58, 2009.
- [Pic10] G. Pickard, I. Rahwan, W. Pan, M. Cebrián R. Crane, A. Madan, A. Pentland, “Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy”, CoRR, 2010.
- [Ram07] R. Ramakrishnan, A. Tomkins, “Toward a PeopleWeb”, IEEE Computer, 40(8):63-72, 2007.
- [Red10] S. Reddy, D. Estrin, M. Srivastava, “Recruitment Framework for Participatory Sensing Data Collections”, Pervasive Conference, Helsinki (FI), 2010.
- [Sha10] D. Shahaf, E. Horvitz, “Generalized Task Markets for Human and Machine Computation”, AAAI Conference, Atlanta (GE), 2010.
- [Shi10] J. Shi, R. Zhang, Y. Liu, Y. Zhang, “Prisense: Privacy-preserving Data Aggregation in People-centric Urban Sensing Systems”, Conference on Information Communications, Piscataway, (NJ), 2010.
- [Ton08] A. Toninelli, A. Corradi, R. Montanari, “Semantic-based Discovery to Support Mobile Context-aware Service Access”, Computer Communications, 31(5), 2008.
- [YueCK09] M. C. Yuen, L. J. Chen, I. King, “A Survey of Human Computation Systems”, CSE Conference, Vancouver (CA), 2009.