PERSONAL DISTRIBUTED ENVIRONMENT - SECURING THE DYNAMIC SERVICE PLATFORMS BEYOND 3G

S K Goo, J M Irvine, R C Atkinson

University of Strathclyde, UK

ABSTRACT

Future mobile systems are expected to offer users flexible access to information and services using a combination of different end-user devices in a Personal Distributed Environment (PDE). With PDEs able to operate over multiple air interfaces and heterogeneous networks, requiring seamless and rapid service provision, a flexible and fair trading of communication services is required. For this reason, a Digital Marketplace (DMP) is proposed. The DMP is based on an agent framework capable of enabling real-time service negotiation over disparate networks according to users' price and QoS requirements. This paper discusses the security threats and challenges to the PDE, and also to a DMP implementation.

INTRODUCTION

As systems evolve beyond 3G, it is anticipated that the user will receive delivery of a multitude of advanced services via a combination of different terminals (devices) in a dynamically changing mobile environment. These terminals, available services and user data will form the user's "Personal Distributed Environment" (PDE) [1]. The delivery of the user's service can be as simple as through a fixed telephone network or as complicated as through a combination of mobile radio systems and a digital broadcast system, depending on the services' availability and coverage at that PDE location. A single, possibly multi-mode, gateway terminal may be used, but it is more likely that the PDE will use a number of different devices interconnected by one or more Personal Area Networks (PANs). The DMP concept permits terminals (on behalf of users) to negotiate for service provision at call set-up. To permit the networked devices to utilise heterogeneous access technologies, a new business model is required from the service provider perspective. This approach allows the requested service to be offered over a network infrastructure by various independent competing operators. The negotiation scheme employed in the proposed model also needs to meet demands such as immediate service provision and service handover of the user devices entering and leaving the PDE as the location of service execution and operating environment changes.

As an example, imagine a user who is away from home whose set top box has recorded their favourite television programme in their absence. As part of their PDE, the set top box could send a message to the user informing them of the receipt of the programme. If that user was currently travelling by train, they may wish to view the episode on one of the train's video display units. To do so, the user has to invite the video display unit to join their PDE and instruct their set-top box to send the video stream to the train display. This requires secure methods of interacting with public devices and including them in the PDE, and it also requires methods of accessing resources for the transfer. The latter may involve a UMTS network link to the user's terminal, followed by Bluetooth to the display, or use of a broadcast network, or if the train provided an internal network, a link with that network.

The PDE concept requires that the service provider/ network operator to be able to identify which devices are common to a single user, since content can be delivered to any one of a user's devices. Privacy mechanisms are required to preserve anonymity when interacting with public devices. Furthermore, as the trend of separating a service provider from the network operator becomes more apparent, the underlying security issues between the two business roles become more significant, especially in protecting the regulation of the market and ensuring free competition among providers and operators.

As the example has shown, the true potential of the PDE concept requires the provision of service over multiple terminals, air interfaces and even fixed networks. While operation from a single network operators networks is possible, such provision ideally requires a flexible and fair trading system allowing connections to be established over an available infrastructure with real-time negotiation. The Digital Marketplace (DMP) [2] is a promising approach using software agents to administer service negotiation in a multi-vendor and multi-technology environment. It permits other marketplaces to be distributed and interconnected throughout the infrastructure, perhaps at various geographical locations. For example, the user's PDE could be split into several geographical locations with different marketplaces, where each marketplace has jurisdiction over service negotiation within different regions of a city. With this proposal, the user can be assured of a choice of highly competitive price and service qualities while the existing network resources are efficiently exploited and reused.

For end to end links, there are two possible ways for service providers to negotiate Quality of Service (QoS) contracts. The first is to negotiate the contract with a network provider who will undertake to carry the connection to the final destination, either on its own network or by negotiating in other marketplaces on his behalf. Alternatively, the service provider can take the full responsibility by negotiating several QoS contracts at different marketplaces that form part of the route to the destination. The security issues will vary between these two solutions, due to the different co-ordination of the end-to-end security levels, procedures and mechanisms within the DMP. Therefore, it is important for the security risk analysis to incorporate these scenarios.

In the example, there is the issue of how, if the video transmission from the set-top box to the train's display unit is unsuccessful, can the service platform allow the user to recover the transmission while still maintaining the rights to a single viewing only? Restrictions incurred from the revocation of rights and licensing issues must also be taken into account during the design of the PDE service platform. When trust is revoked from a device that was formerly attached to the PDE, how will the marketplace manage this occurrence? It is essential that these issues and their associated security requirements are properly addressed so that providers and operators will have sufficient confidence to support the DMP concept in addition to the PDE architecture.

In the following section, the DMP architecture is briefly described. This allows the investigation of attacks to be conducted and hence, possible security threats are deduced at the different stages of the service transaction. The security requirements and services are then highlighted. This is followed by some discussions on the proposed trust model that is necessary in the security architecture of the DMP before the concluding section.

SECURITY ISSUES

DMP Entities

The DMP exists within the service layer of a four layer system consisting the user application layer, service negotiation layer, network resource layer and medium (communication) layer [2]. The three main trading players/actors in the DMP are the user, the service provider, which represents the user in the market and provides the link to the application, and the network operator, which provides the communication link. The DMP also consists of a market provider, which oversees transactions using market agents. The market provider's main role is to facilitate negotiations between service providers and network providers so their customers receive resources (specified according to QoS) at a competitive price. If any of the network providers do not fulfil their contractual commitments, the market provider re-evaluates their reputation values. Each of these identified actors negotiates through the use of autonomous agents, and each actor is entitled to assign and clone them. The agents may either have a fixed location (static agents) or move (mobile agents) around the specified area where the usage pattern in the DMP is homogeneous. For example, the user agents consist of a User Service Agent (USA) at the negotiation platform of one particular marketplace, a User Terminal Agent (UTA) at one of his device/ terminal and a User Home Agent (UHA) at the service provider server.

Networks providing coverage require two different types of agent: a Network Home Agent (NHA) at the site of the network operator server and a Network Operator Agent (NOA) for making bids to the respective service agents. The Service Provider Agent (SPA) is created by a service provider and travels to the marketplace where the user is located. The remaining two agent types are Market Controller Agent (MCA) and Market Interface Agent (MIA) and they belong to the market provider. All these agents are configured and implemented in such a manner to allow easy global interconnection (Figure 1) to any distributed communication infrastructure.

DMP Security Threats

This section will identify security threats and challenges that may pose a threat to a DMP system. The aim is to eliminate fraud and misrepresentation (and misinterpretation) from service negotiations when a DMP system is used with a PDE environment. The analysis can be examined from the perspectives of different key actors in a DMP such as the user (or consumer), the network operator, the service provider and the market provider as well as the terminal provider (or/and manufacturer). The anticipated security threats can be initiated either from the misbehaving agents, or from a malicious platform, and directed against any operating agents.

One important concern among the different players (e.g. bidders, suppliers and the market itself) is the trust relationship that must exist between the respective agents and how some of these parties may transgress prescribed rules of behaviour, for example by colluding to raise or depress prices during service negotiation. The colluding could occur a) between the network agents, b) between the network and service agents, c) between the market agent and its preferred network agent, d) between the service agent and the market agent or e) between the service agents themselves.

<u>Establishing a connection to the respective marketplace</u>
Prior to call set-up, users signal their QoS requirements to all parties in the geographical area. The signalling is conducted via a Logical Market Channel (LMC), provided by one of the network operators; the various operators may utilise various heterogeneous access

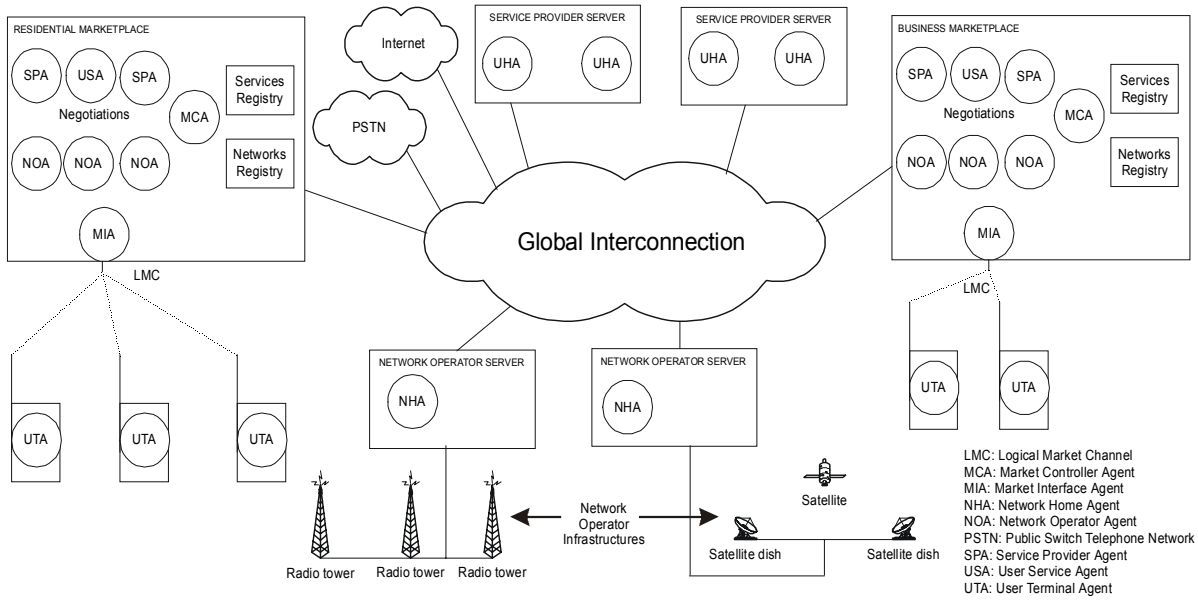

Figure 1: Digital Marketplaces [2]

technologies. The market provider could contract this LMC to a network provider that is different from the one that a given user has actually contracted their service with. When this happens, covert attacks and threats may hence be instigated through this second operator by reducing its security mechanisms. Hence, an illegitimate user could impersonate the contracted user, and obtain a service via the LMC without payment. It also invites passive attack such as eavesdropping on the user's service requirements and/or service/user profile information transmitted to the market agent via the LMC.

Auction, Bidding and Negotiation

The auction involves one buyer – the user or their service provider – and one or more sellers – network operators. As the number of entities involved at this transaction stage is much larger than in the earlier stage, the anticipated collusion problems and the associated security issues are also more severe. These include masquerade/impersonation of other entities (e.g. a registered network and service agents), a dishonest auctioneering process, denial of execution (DoE), illegal execution of auction rules, manipulation of user requirements and failure to follow through on commitments.

Other concerns include the prevention of active attacks from illegitimate providers gaining trust to alter or delete bids, placing spurious bids in a Denial of Service attack (DoS), or a dishonest market agent from itself colluding with either service provider agents or network operator agents to pass private information (such as reputation and the bidding strategies).

The offered price in the DMP may be related to the ratio between the demand and the supply of radio resources in each marketplace. The demand depends on the number of consumers or bidders while the overall market supply (or network resources) is heavily dependent on the number of registered network operators within the marketplace. If any of the parties are guilty of deception and provide any wrong information to the demand and supply, this would create an imbalance in the ratio and thus the offered price. Aside from collusion, or the alteration of bids, agents could fabricate spurious resource requests to increase demand. Supply would be restricted if bids were not passed to all eligible network operators. In the selection process, a network's reputation, QoS and price influence the awarding of a contract to a network operator. If the network's reputation is inaccurately penalised by an unfair market agent, the network operator may experience a lower reputation from the service agents and may subsequently react by lowering its offered price.

Real time solution to the problems of bid manipulation and collusion is difficult, but offline solutions using logging is possible if the market agent maintains a read-only log of bids and offers, which becomes public after a time delay. Agents can then verify that bids and offers they placed in the market were properly recorded and can check that they were communicated to them honestly at the time. Collusion outside the market is much more difficult and must be addressed by regulatory strategies similar to those used in commercial markets today.

Billing

The service provider will usually determine billing. It will consolidate the billing information from the respective marketplaces and collect the charges from consumers on behalf of the network operators. Accountability and non-repudiation are two significant problems to address in the billing. Issues such as the ability to verify the accuracy of data, the capability to rectify the fraud and how much data are sufficient for billing and legal purposes are necessary to be resolved before the DMP concept can be considered feasible. An

event logger located in each marketplace will help to keep track of all activities performed by the agents. This will work alongside another suggested component: a system security agent in the DMP. Another aspect to be considered is if a user who has no current network connectivity or any established relationship with a service provider and uses micro-payment for up front payment, what are the additional required security measures?

Device Identification

Though the terminal provider (or/and manufacturer) is not directly involved in any of the service transactions, the devices provided to be used in the service negotiation and transaction within the PDE have to be as compatible as possible with other devices while maintaining ease of usability and configurability. Secure hardware, such as SIM or smart cards, is often used to provide security in existing systems but network operators and service providers are unlikely to share these, therefore designing devices to accept multiple smart cards, for example, will make them more expensive and complex. The manufacturers will also need to consider ways of configuring devices so that they can work with one or more PDEs, and not allow information leakage between them or over the marketplaces. Furthermore, as the PDE is distributed over a number of devices, traditional solutions such as smart cards, which are used to perform the cryptographic operations and identity verification of a user within a device, may no longer be appropriate, so a software solution for authentication and identification services is required.

To summarise, the predicted threats can be classified and based on four main areas such as agent's execution (e.g. data exchange and computation), the agent's intercommunication with other agents (e.g. collusion), the agents themselves and the storage of data (i.e. the location (confidentiality) and the storing methodology (confidentiality and integrity)). Due to the space restrictions, the discussions in this paper are limited to the security services of concern to the user since the PDE will be concentrating extensively on the viewpoint of user-centric applications.

SECURITY SERVICES

DMP users require some form of identification. On some occasions, the users may wish to conceal their real identities from untrusted devices, so some degree of anonymity tuning may have to take place. However, there are limitations to the extent to which this can be achieved when accountability and non-repudiation are involved. Alternatively, users can use a virtual identity (VID) to protect their privacy needs during correspondence with other parties. Identification in conjunction with authentication can assist the user to gain access to the service. In this case, mutual authentication of the entity (instead of the data origin) is most likely to be involved and helps to avoid masquerading, DoS, theft of rights, replay attacks and resource misuse. The security agent which acts like a Trusted Third Party (TTP) can use signatures to certify and distribute the public and session keys to the trading agents (which have no previous formal relationship) to get authenticated. This is vital as the PDE topology will be dynamic.

A confidentiality/privacy service is required to safeguard the user location and the user profiles such as service usage profile and monetary information. This is especially important when there is a need to propagate user information through several network links that are owned by different network operators. The purpose of this service is to prevent eavesdropping and theft of rights.

Appropriate rights and permissions for services must be specified. As for access rights to user information (i.e. access control), some control of rights and restriction must be enforced. Both components provide user a degree of protection from DoS and misuse of resources, and prevent illegitimate access of entities.

To avoid unauthorised manipulation of user tariffs relayed by network operators to the service providers, and the user's charging information provided by the service providers, data integrity is required. This service is also used to prevent replay attack, theft of rights and replication should the situation be identified.

Trust Model

The DMP-PDE security architecture is dependent on many components (e.g. cryptographic functions) and security services, but one important factor is the trust relationships between these agents (and how to prevent unpredicted foul play). In a traditional mobile communication system, the user can simply put their trust on their mobile provider. However, in the PDE case, the user (and the content providers) may need an extra trust model, particularly when there is an increase in the number of interconnection of different devices and public equipment. The service provider and network operator could also add this model to their existing security techniques and architectures, where suitable. The purpose of this proposed trust model is to help each entity to evaluate the trustworthiness of its corresponding entity; it can (indirectly) force the other entities to behave in the service negotiation.

A possible trust model for DMP could be based on the perspective of social mechanism (of the DMP actors), as used in [3,4]. The two involved "social" components for the trust model are risk assessment (RA) and recovering rate from attack (RR) and the proposed trust level can be based on:

$$T = 1.0 - RA - RR \quad \text{--- (1)}$$

where RA is defined as how much risk that the agent/host/communication path will incur to the communicating entities and RR is defined as the reputation of not recovering quickly from the malicious attack or also known as the slow recovery rate. The maximum value for both RA and RR is 0.5.
The risk assessment (RA) can be calculated via the votes whereby

$$RA = \frac{\sum_{n=1}^{n=\infty} V_n}{n} \quad \text{--- (2)}$$

where n = 1, 2, 3, … and V_n is the vote cast for a particular object from its corresponding objects. The value of this "trust" vote, V_n, is ranged between 0 (min) & 0.5 (max). Consider a communication example:

Agent A ⇔ host ⇔ Agent B,
Agent B ⇔ Agent C ⇔ Agent A,
Agent D ⇔ Agent A.

From the above, the risk assessment (RA) for Agent A will be based on the votes from the host, Agent C and Agent D. If the host and the two agents gave an average risk assessment of 0.1, for example, it means Agent A has incurred a low threat or risk to its corresponding entities.

It is a requirement that at least one registered agent is involved in each risk assessment. This is because all the risk values that are given by the communicating agents and hosts have to be assessed through the stored votes in the event logger.

As for the recovery reputation (RR), it is different from the DMP's reputation (which is the ability to deliver the promised service) because the value of RR is determined by the security agent in the system. If RR is at its maximum value of 0.5, it means the entity has the poorest recovery rate possible. The recovery rate defined in this paper can be quite subjective as it depends on whether the number of hours or the number of days is used in determining how fast that these "attacked" agents/hosts/communication paths are cleared from the malicious attack/virus. If the value of the (slow) recovery rate is high, it may also indicate that the agents/hosts/communication paths are not updated with the latest security mechanisms, to deal with new attack. Having explained how the two trust components work in the trust calculation, there are also other possible social mechanisms, which can be added into the trust model. One example is reliability, which acts a monitoring protocol to the trust evaluation. The reliability component can be used to evaluate how reliable the votes (given in the risk computation) are.

If the trust level is 1 (which is the maximum), it means that the entity has the lowest risk level & lowest reputation of not recovering (which means high immunity towards the malicious virus/attack).

Further to these security services, using good key distribution with regular key refresh, securing the task delegation and isolating mobile agents from malicious hosts could also add (double) counter threats which are yet to be predicted. Finally, to realise "real" end-to-end security in DMP-PDE, all encrypted programs/code must be executed without decrypting them (even if the intermediate nodes/ hosts of a route are securely protected). This will automatically append the privacy and integrity services for the exchanging of data and the executable code.

CONCLUSIONS

This paper has proposed an agent-based framework to facilitate the provision of mobile services in a personal distributed environment. While this proposal ensures both the consumers and the suppliers can negotiate in a fair, competitive and dynamically changing environment, it has also identified important security issues which closely relate to the DMP actors during the service transaction. As a result, essential security services and trust models may then be addressed. These discussions will reveal how the underlying transport network, the types of service and the future application development in the PDE can be achieved. The detailed security specifications for the mechanisms and protocols for DMP-PDE architecture will then be developed based on these concepts.

ACKNOWLEDGEMENT

The work reported in this paper has formed part of the PDE area of the Core 3 Research Programme of the Virtual Centre of Excellence in Mobile & Personal Communications, Mobile VCE, www.mobilevce.com, whose funding support, including that of EPSRC, is gratefully acknowledged. Full detailed technical reports on this research are available to Industrial Members of Mobile VCE. Also, we would like to thank our colleagues within Mobile VCE for their helpful advice and suggestions.

REFERENCES

[1] Dunlop, J., Atkinson, R. C., Irvine, J., Pearce D., "A Personal Distributed Environment for Future Mobile Systems", IST Mobile & Wireless Communications Summit 2003, to appear in June 2003.
[2] Irvine, J., "Adam Smith Goes Mobile: Managing Serviced Beyond 3G with the Digital Marketplace", European Wireless 2002, Florence, Italy, February 2002.
[3] Padovan, B., Sackmann, S., Eymann T., Pippow, I., "A Prototype for an Agent-based Secure Electronic Marketplace including Reputation Tracking Mechanisms", in Proc. IEEE 34th Annual International Conference on System Sciences, 2001, Hawaii.
[4] Castelfranschi, C., Falcone R., "Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification", in Proceedings of International Conference on Multi Agent Systems, 1998, Pages: 72–9