Doktorandský den ’06

Ústav informatiky
Akademie věd České republiky

Monínec, Sedlec–Prčice

20.– 22. září 2006

vydavatelství Matematicko-fyzikální fakulty
University Karlovy v Praze

Publikaci ”Doktorandský den ’06” sestavil a připravil k tisku František Hakl
Ústav Informatiky AV ČR, Pod Vodárenskou věží 2, 182 07 Praha 8

ISBN *– not yet –*

Obsah

A Security Model Based on Virtual Organizations for Various Distributed Environments

Post-Graduate Student:
ING. ROMAN ŠPÁNEK

Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague.

Department of Software Engineering, Faculty of Mechatronics,
Technical University of Liberec

spanek@cs.cas.cz
roman.spanek@vslib.cz

Supervisor:
ING. JÚLIUS ŠTULLER, CSC.

Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague

stuller@cs.cas.cz

Field of Study:
Elektrotechnika a informatika

The work was supported by the project 1ET100300419 of the Program Information Society (of the Thematic Program II of the National Research Program of the Czech Republic) “Intelligent Models, Algorithms, Methods and Tools for the Semantic Web Realization”, partly by the Institutional Research Plan AV0Z10300504 “Computer Science for the Information Society: Models, Algorithms, Applications” and by the Ministry of Education, Youth and Sports of the Czech Republic, Project No. 1M4674788502 “Advanced Remedial Technology and Processes”.

Abstract

The paper presents a new approach for treating security issues in various environments with special emphasis on Mobile databases, Semantic web and Grids. A brief overview on possible security models and a discussion on their advantages and disadvantages is given. Our model based on virtual organization and is build up on mathematical background based on hypergraphs. We show that hypergraphs are the way how to reduce space complexity of the model. The complexity is important with respect to target environments where number of users might be huge. To verify our model an experimental implementation was programed and some graphical outputs are mentioned.

`http://www.cs.cas.cz/hakl/doktorandsky-den/index.html`.

1. Introduction

Rapid evolution in many computing areas brings up many useful aspects, but also many problems and issues to be addressed. Nevertheless, in the rest of the paper we will concentrate only on the security issues of the following distributed environments:

- Mobile computing
- Semantic web
- Grid computing

Although these have some different features, they also have a lot of in common. Before we proceed to the common features, let us briefly overview the environments.

Mobile databases [1],[2],[3], offer the ability to access and exchange information anywhere, at any time. The possible network architectures can be summarized as:

- cellular networks

- multihop wireless networks - broadly known as ad hoc networks
- sensor networks

In the first case, some specialized nodes, called base stations, coordinate and control all transmissions within their coverage area - *a cell*. The base station grants access to the wireless channels in response to service requests received by the mobile nodes currently in its cell.

The primary characteristic of an *ad hoc network* architecture [4],[5],[6],[7] is the absence of any stationary structure. Ad hoc nodes can communicate directly with the nodes in their transmission range in a peer-to-peer fashion. Communication to distant nodes is achieved through other nodes in the network in multi-hop fashion. Therefore each ad hoc node acts also as a router, storing and forwarding packets on behalf of other nodes. The result is a generalized wireless network that can be rapidly deployed and dynamically reconfigured to provide on-demand networking solutions. Besides the fact that pervasive computing has a lot of advantages, it also has some challenges to cope with. Taking issues like power supply limits, limited bandwidth and unreliability of wireless lines, the security is one of the most important. Without having efficient and strong security solution it may be very hard to achieve all the possible advantages of ad hoc networks. *Sensor network* can be characterized as networks build up from tiny sensing units having some communication and computation capabilities. Such sensors cooperate in multi-hop communication to delivery data to a unit responsible for its further processing. As sensor networks are a bit specific we do not addressed them in the paper anymore.

The *Semantic Web* is often believed to be the successor of the current web. Its main idea is to describe resources in the form of machine processable meta-data allowing automation of the requested tasks connected with the retrieval and usage of these resources. Although the main focus of previous work was aimed at the creation of knowledge representation languages (RDF-S, DAML+OIL [8], OWL [9]), reasoning systems, and also at the tools helping to embed web pages with semantic markup, the emerging commercial applications such as e-commerce, banking or travel services face a lot of security issues. Without a secure solution, it would be very hard to exploit all promising features of semantic web vision. The first possible approach is to extend the current security mechanisms used in distributed systems (Kerberos [10], PGP [11], SPKI [12] etc.). These technologies, however, cannot be seamlessly transferred due to the fully decentralized nature of the web, extremely large number of resources, services, agents and users, and their heterogeneity. Moreover, the number of entities accessing sources and interacting with themselves can be very large and can rapidly change.

The *Grid computing* paradigm can be characterized by a large number of interconnected users and sites cooperating on the common task. Users in a Grid are usually organized in *Virtual Organizations* (VOs). A Virtual Organization is a temporary or permanent coalition of geographically dispersed individuals, groups, organizational units or entire organizations that pool resources, services and information to achieve common objectives. The Dynamic Virtual Organizations Membership and structure of such a VO may evolve over time to accommodate changes in requirements or to adapt to new opportunities in the business environment. Considering this, it is straightforward that grid computing strategies can be used in the web environment for security improvements.

Even thought the mentioned areas do have some specifications typical for them, such as huge amount of pages in the case of semantic web, mobility of users in mobile computing paradigm, or heterogeneity of connected sources in grids, they also have some common specifications, e.g. usage of computer agent technologies. Further, while all of them offer ability to share resource, support communication and cooperation between users (but not only the humans users), the security is the crucial issue being common for all mentioned areas. Therefore it is natural to expect some solutions that might solve the problem with security in all of them.

The rest of the paper is organized as follows: section 2 briefly introduces the related security models and our security model is then described more in details in the next section. Section 4.2 contains some experimental result. The paper is then concluded.

2. Security Models and Approaches

The security, no doubt, is one of the key concerns in many areas. On the other hand, whenever humans became users, the security gains on importance. While all three mentioned areas (mobile computing, semantic web, grids) are primarily for human users, the security should be sufficiently solved. The security can be treated on two separate levels:

- cryptography level
- trust level

2.1. Cryptography

On the first level strong cryptography algorithms are the basis, taking responsibility for shielding transmitted data against man-in-the-middle attack, threat of tap, etc. Cryptography plays also important rule in certificates (PKI [13]) allowing users to communicate and find some useful information without having any knowledge of themselves. Cryptography, however, is enough only when consider tasks like sending messages, sharing files, etc. with only accent on secure transmission of sent data. This approach, nevertheless, suffers by lack of additional abilities required by human users, like when do share data, when to trust the sender, etc. Therefore the next level mentioned in the previous list is the level being responsible for the trust management.

2.2. Trust Managing Security Models

Security solutions based on a strong cryptography should be the basis. But, there is still a space for improvements on the second level of the trust management. The trust management approaches build an enhanced security level on underlying cryptography level. The main task is to build, preserve and manage relationships between users. The relationship are usually build up on the trust.

Definition: Trust of a party A to a party B for a service X is a measurable belief of A in that B behaves dependably for a specified period within a specified context (in relation to service X)

This approach is similar to the well known term of creating Virtual Organization in grid environment. The necessary condition for practical evolution of VOs is to have a strong mechanism which preserves their overall security. Here we propose a system based on the bottom-up (local) preservation of security. More specifically, the involved entities build up the security from mutual relations among them. The distributed mechanisms of VOs check and globalize these relations.

3. Related Work

This section gives a brief overview of the trust management approaches proposed for VO. Two main approaches are currently available for the *trust management*:

Policy-based approach has been proposed in the context of open and distributed services architectures [14],[15],[16],[17],[18] as well as in the context of Grids [19] as a solution to the problem of authorization and access control in open systems. Its focus is on trust management mechanisms employing different policy languages and engines for specifying and reasoning on rules for trust establishment. In addition, it is possible to formalize trust and risk within rule-based policy languages in terms of logical formulae that may occur in rule bodies. Currently, policy-based trust is typically involved in access control decisions. Declarative policies are very well suited for specifying access control conditions that are eventually meant to yield a boolean decision (the requested resource is either granted or denied). Systems enforcing policy based trust typically use languages with well-defined semantics and make decisions based on ”nonsubjective“ attributes (e.g., requester’s age or address) which might be certified by certification authorities (e.g., via digital credentials). In general, policy-based trust is intended for systems with strong protection requirements, for systems whose behavior is guided by complex rules and/or must be easily changeable, as well as for systems where the nature of the information used in the authorization process is exact.

Reputation-based approach has emerged in the context of electronic commerce systems, e.g. eBay. In distributed settings, reputation-based approaches have been proposed for managing trust in public key certificates, in P2P systems, mobile ad-hoc networks, and recently, in the Semantic Web, such as [20],[21],[22],[23],[24],[25]. Typically, reputation-based trust is used in distributed networks where any involved entity has only a limited knowledge about the whole network. In this approach, the reputation is based on recommendations and experiences of other users/sites.

In the following we will put a strong emphasis on creating the underlying VO by "evolution". In order to describe it we need a model which can efficiently capture the grid changes. This model is described in the subsequent section.

4. Security Model Based on Virtual Organization

As was mentioned in the previous sections, VO can be useful model for treating the trust between users. Further, such a model is useful in all related environments (mobile databases, semantic web, grids). On the other hand, VO model can be limited by some specific features of the environments:

- *mobile database* environment in addition to mobility of users also poses severe limitations to storage and computation capabilities of devices
- *semantic web* environment with almost unlimited number of users poses requirements on storage complexity
- *grids* were the target environment for VO, therefore the main issue of heterogeneity of sources had been addressed

From the list of additional limitations, it can be shown that a model having the following specification is required:

1. The model should be able to store large amount of users in low storage complexity. VOs are very often modeled and depicted as (oriented) weighted graphs. But the complexity of storing information about all members in VO might be very high. The given complexity is $O(n^2)$, where n is the amount of vertexes. This is, however, unacceptable in case of mobile database environment and also semantic web might very quickly exceed storage capacity of particular node.
2. The model should have some level of autonomy in building relationships and the trust among users. The autonomous feature of the model is crucial when considering environments where users' relationships became complicated or agent technologies are used. Such feature is highly useful when users would like to create strongly connected groups *on-the-fly*. Nowadays approaches usually assume that such groups are created by somebody and usually manually. We consider such creation as a bottleneck of these models.
3. The model should be implementable in distributed (heterogeneous) environment. A distributed implementation is the key factor influencing model capabilities and usefulness.

Our approach is therefore build up on the previous list of requirements. The next section describes the very base of our proposal.

4.1. The Security Model

Let us shortly describe a mathematical model that we found to be useful. Hypergraphs are commonly quaternion (V, E, W_v, W_e), where V is a set of vertexes, E is set of edges ($E \in 2^V$), W_v is a set of vertexes' weights and finally W_e is a set of edges' weights. The main difference between graphs and hypergrahs is

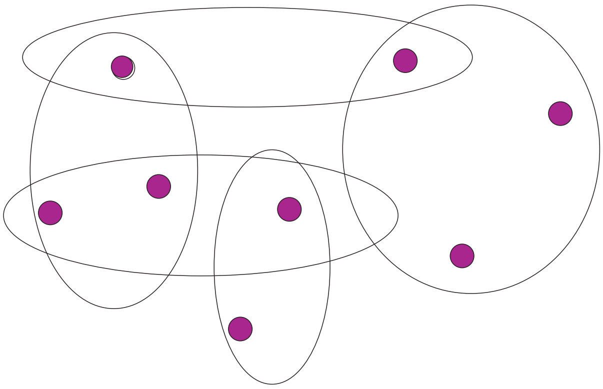

Figure 1: An example of hypergraph.

that an hyperedge can be incident to more then two vertexes. An example of a hypergraph is in Figure 1. It is example of hyperpraph containing 5 hyperedges and 8 vertexes.

In Figure 2 the same situation is sketched, but now using graphs instead of hypergraphs. The edges in Figure 2 are shown in different colors and styles according to hyperedges from Figure 1.

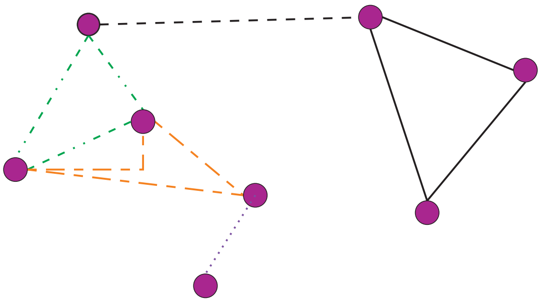

Figure 2: An example of graph showing groups of vertexes.

It is clear that graphs are subset of hypergraphs. It is also clear that hypergraphs are richer structure than graphs. On the other hand, the richness of hypergraph brings some implementations issues (note that the issues are out of scope of the paper).

Let now return to the list of three additional limitations from the previous section. The first item on the list was the low space complexity. In the case of graphs the space complexity is $O(n^2)$, which is unacceptable. On the other hand, in Figure 1 you can see how VO can be stored as hypergraph. In that case VO is not stored as sets of vertexes, edges and their incidence, but simply by a membership of hyperedges. Therefore hypergraphs can be very useful for modeling VO reducing the space complexity.

The second requirement is a kind of autonomy. One of possible solutions is to have set of rules that take care of all edges and also vertexes in the VO. We, hence, propose such a set of rules. Due to space limitation the rules are not mentioned here (see [26] for details on the rules).

Third item on the list requires implementation in distributed environment. When try to build up a list of all possible distributed implementation, we should start with implementations based on Remote Procedure Call (RPC), like CORBA[27] or JavaRMI [28]. Another technology worth mentioning are services. As an example let us mention web services based on WSDL[29] and SOAP[30]. One of the last possibility is to use message passing. The main advantage is of message passing is its simple and environmentally independent implementation.

With respect to our needs and also to target environments the best choice is message passing with its simple, straightforward and efficient implementation.

4.2. Experimental Application SecGRID

An experimental implementation SecGRID was programed in ANSI/C and its aim was to verify that proposed algorithms for edge reevaluation preserve consistency of the VO. By the consistency of the VO we mean that the structure will:

- not degenerate to one huge VO containing all nodes
- not degenerate to huge amount of very small VOs
- preserve relationships (expressed by an edge weight) between users

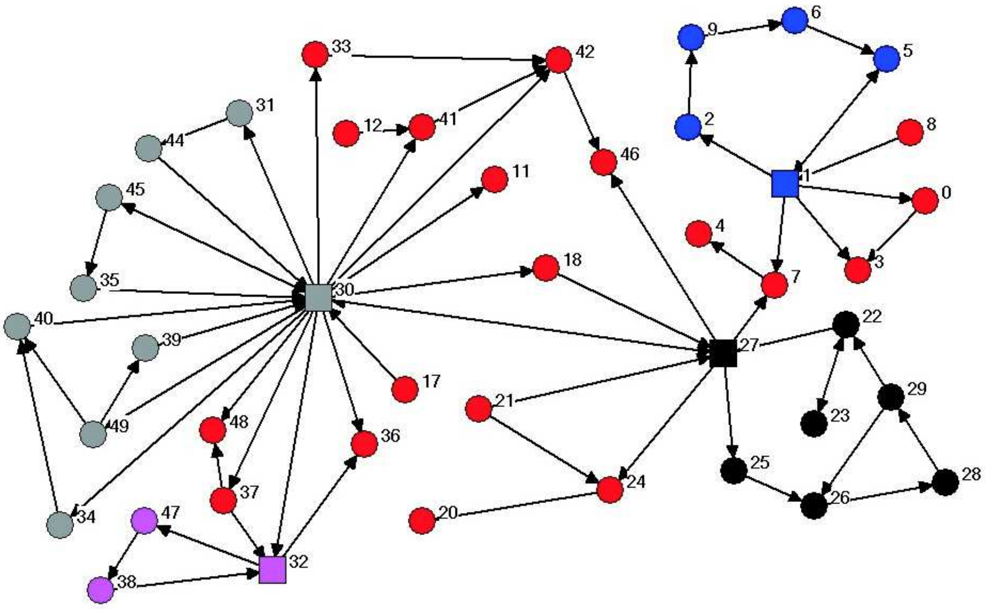

Figure 3: Experimental implementation output.

To verify the model the simulation comprised two phases:

1. random generation of edges
2. intentional edge generation with special accent of problematic situation in the VO evolution

One of the resulted graphical representation, which was obtained using NetDRAW [31], is shown in Figure 3. Figure shows a situation after few steps of adding edges into the structure have been done. Users are depicted as circles in the color according to group membership, apart from the reds that are members of no group. Every group has also a leading member that has additional responsibilities, e.g. outer group communication support. The leaders are shown as squares in the color corresponding to group membership. For the sack of lucidity, the edges weights are not shown.

From the figure it can be seen that the consistency is preserved and the vertexes are uniformly distributed into groups. Note that group corresponds to a hyperedge in our hypergraph model and that the implementation uses hypergraphs with hyperedge incidence 2.

5. Conclusions

The aim of the paper was to propose a new security model for mobile databases, semantic web and grids. The paper begins with a brief overview on two separate levels of the security (crysptography and the trust) followed by a list of features specific to the target environments. Having sumarized all requirements we describe our proposal based on Virtual organization model for the trust security level. Our model uses hypergraph theory as its mathematical basis, while the hypergraphs have abilities that enable us to reduce the space complexity of the model. The experimental results obtained through an experimental implementation are given to verify the "evolution" phase of the proposed model showing that it does not degenerate to any of the limiting cases. Althought the model is based on hypergraphs with full cardinality of hyperedges the experimental application is based on hypergraph with hyperedges' cardinalities reduced to 2.

References

[1] S. DasBit and S. Mitra, "Challenges of computing in mobile cellular environment a survey", *Elsevier B.V.*, 2003.

[2] Y. Lu, B. Bhargava, W. Wang, Y. Zhong and X. Wu, "Secure Wireless Network with Movable Base Stations", *IEICE Trans. Community*, vol. E86-B, 2003.

[3] Y. Zong, B. Bhargava and M. Mahoui, "Trustworthiness Based Authorization on WWW", *IEEE Workshop on Security in Distributed Data Warehousing*, 2001.

[4] P. K. Behera, P. K. Meher, "Prospects of Group-Based Communication in Mobile Ad hoc Networks", *Springer-Verlag Berlin Heidelberg*, 2002.

[5] A. Flaxman, A. Frieze, E. Upfal, "Efficient communication in an ad-hoc network", *Elsevier*, 2004.

[6] S. Basagni, "Remarks on Ad Hoc Networking", *Springer-Verlag*, Berlin Heidelberg, 2002.

[7] R. Molva, P. Michiardi, "Security in Ad Hoc Network", *IFIP International Federation for Information Processing*, 2003.

[8] "http://www.daml.org/"

[9] "http://www.w3.org/TR/webont-req/"

[10] J. Steiner, C . Neuman, and J.I . Schiller, "Kerberos : An Authentication Service for Open Network Systems, " *in Proc. Winter USENIX Conference*, Dallas (1988).

[11] Philip R. Zimmermann. "The Official PGP User's Guide". *MIT Press*, Boston, 1995.

[12] Carl M. Ellison, Bill Frantz, Butler Lampson, Ron Rivest, Brian M. Thomas, and Tatu Ylonen. "SPKI certificate theory," *Internet RFC 2693*. October 1999.

[13] Arsenault, A. and Turner, S. "Internet X.509 public key infrastructure: roadmap internet draft", *draft-ietf-pkix-roadmap-09.txt*, July 2002. http://www.ietf.org/internet-drafts/draft-ietf-pkix-roadmap-09.txt.

[14] P. Bonatti and P. Samarati, "Regulating service access and information release on the web.", *In CCS ´00: Proceedings of the 7th ACM conference on computer and communications security*, pages 134–143. ACM Press, 2000.

[15] N. Li and J. Mitchell, "A Role-based Trust-management Framework.", *In DARPA Information Survivability Conference and Exposition (DISCEX)*, Washington, D.C., Apr. 2003.

[16] R. Gavriloaie, W. Nejdl, D. Olmedilla, K. E. Seamons, and M. Winslett, "No registration needed: How to use declarative policies and negotiation to access sensitive resources on the semantic web.", *In 1st European Semantic Web Symposium (ESWS 2004)*, volume 3053 of Lecture Notes in Computer Science, pages 342-356, Heraklion, Crete, Greece,Springer, May 2004.

[17] M. Y. Becker and P. Sewell, "Cassandra: distributed access control policies with tunable expressiveness.", *In 5th IEEE International Workshop on Policies for Distributed Systems and Networks*, Yorktown Heights, June 2004.

[18] P. A. Bonatti and D. Olmedilla, “Driving and monitoring provisional trust negotiation with meta-policies.”, *In 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2005)*, pages 14-23, Stockholm, Sweden, Jun 2005. IEEE Computer Society.

[19] J. Basney, W. Nejdl, D. Olmedilla, V. Welch, and M. Winslett, “Negotiating trust on the grid.”, *In 2nd WWW Workshop on Semantics in P2P and Grid Computing*, New York, USA, May 2004.

[20] K. Aberer and Z. Despotovic, “Managing trust in a peer-2-peer information system.”, *In Proceedings of 10th International Conference on Information and Knowledge Management*, pages 310-317, 2001.

[21] E. Damiani, S. D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante, “A reputation-based approach for choosing reliable resources in peer-to-peer networks.”, *In Proceedings of ACM Conference on Computer and Communications Security*, pages 202-216, 2002.

[22] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, “ Eigenrep: Reputation management in p2p networks.”, *In Proceedings of 12th International WWW Conference*, pages 640–651, 2003.

[23] C. Duma, N. Shahmehri, and G. Caronni, “Dynamic trust metrics for peer-topeer systems.”, *In Proceedings of 2nd IEEE Workshop on P2P Data Management*, Security and Trust (in connection with DEXA ´05), August 2005.

[24] L. Kagal, T. Finin, and A. Joshi, “A policy based approach to security for the semantic web.”, *In Proceedings of the 2nd International Semantic Web Conference*, Sanibel Island, Florida, USA, Oct. 2003.

[25] G. Tonti, J. M. Bradshaw, R. Jeffers, R. Montanari, N. Suri, and A. Uszok, “Semantic web languages for policy representation and reasoning: A comparison of KAoS, Rei and Ponder.”, *In Proceedings of the 2nd International Semantic Web Conference*, Sanibel Island, Florida, USA, Oct. 2003.

[26] R. Spanek, M. Tuma, “Secure Grid-based Computing with Social-Network Based Trust Management in the Semantic Web”, *submitted to NNW*, 2006.

[27] “OMG Specifications”, *http://www.omg.org/technology/documents/spec_catalog.htm*.

[28] “Java RMI Specification”, *http://java.sun.com/j2se/1.4.2/docs/guide/rmi/spec/rmiTOC.html*.

[29] “WSDL specification”, *http://www.w3.org/2002/ws/desc/*.

[30] “SOAP specification”, *http://www.w3.org/TR/soap12-part0/*.

[31] S.P. Borgatti, “NetDraw: Graph Visualization Software.”, *Harvard: Analytic Technologies*, 2002, http://www.analytictech.com/netdraw.htm.

Ústav Informatiky AV ČR
DOKTORANDSKÝ DEN ’06

Vydal
MATFYZPRESS
vydavatelství
Matematicko-fyzikální fakulty
University Karlovy
Sokolovská 83, 186 75 Praha 8
jako svou *– not yet –* . publikaci

Obálku navrhl František Hakl

Z předloh připravených v systému LaTeX
vytisklo Reprostředisko MFF UK
Sokolovská 83, 186 75 Praha 8

Vydání první

Praha 2006

ISBN *– not yet –*