Privacy for Semantic Web Mining using Advanced DSA – Spatial LBS Case Study

Dr.D.Sravan Kumar[1] and M.Upendra Kumar[2]

[1]Principal and Professor CSE and IT KITE Women's College of Professional Engg Sciences Hyd.India

[2] Research Scholar JNTUH and Associate Professor CSE MGIT Hyd.India

Abstract: **The Web Services paradigm promises to enable rich flexible and dynamic interoperation of highly distributed, heterogeneous network enabled services. The idea of Web Services Mining that it makes use of the findings in the field of data mining and applies them to the world of Web Services. The emerging concept of Semantic Web Services aims at more sophisticated Web Services technologies: on basis of Semantic Description Frameworks, Intelligent mechanisms are envisioned for Discovery, Composition, and contracting of Web Services. The aim of semantic web is not only to support to access information on the web but also to support its usage. Geospatial Semantic Web is an augmentation to the Semantic Web that adds geospatial abstractions, as well as related reasoning, representation and query mechanisms. Web Service Security represents a key requirement for today's distributed interconnected digital world and for the new generations, Web 2.0 and Semantic Web. To date, the problem of security has been investigated very much in the context of standardization efforts; Personal judgments are made usually based on the sensitivity of the information and the reputation of the party to which the information is to be disclosed. On the privacy front, this means that privacy invasion would net more quality and sensitive personal information. In this paper, we had implemented a case study on integrated privacy issues of Spatial Semantic Web Services Mining. Initially we improved privacy of Geospatial Semantic Layer. Finally, we implemented a Location Based System and improved its digital signature capability, using advanced Digital Signature standards.**

Keywords: Semantic Web, Spatial Web mining, Scientific Data Mining, Privacy, Advanced Digital Signature Standards.

1. Introduction

Web Services are becoming more and more complex, involving numerous interacting business objects within considerable processes. Web Services are self-contained, self-describing, modular applications that can be published, located, and dynamically invoked across the Web [1]. In order to fully explore Web Service business opportunities while ensuring a correct and reliable modeling and execution, analyzing and tracking Web Services interactions will enable them to be well understood and controlled. The Web Services paradigm promises to enable rich, flexible and dynamic interoperation of highly distributed, heterogeneous network enabled services. The idea of Web Services Mining that makes use of the findings in the field of data mining and apply them to the world of Web Services and Service Oriented Architectures.

With the advent of evolution of WWW from Web searching to Web Mining, many business applications that utilize data mining techniques are extracting useful business information on the web. The web is transforming from a web of data to a web of both semantic data and services. This trend is providing us with increasing opportunities to compose potentially interesting and useful services from existing services. Using open web API's like Google (Spatial Google Map API version 2), Amazon and eBay we can implement data mining algorithms like Apriori for various business applications like CRM with ease. The emerging concept of Semantic Web Services aims at more sophisticated Web Services technologies: on basis of Semantic Description Frameworks, Intelligent mechanisms are envisioned for Discovery, Composition, and contracting of Web Services. The Semantic Web Services on basis of the Web Services Modeling Ontology (WSMO) and related initiatives. Ontology is a representation of a set of concepts within a domain and the relationships between those concepts. In other words, ontology defines the vocabulary of a domain.

A review of Semantic Web use in the EU shows that its related technologies are finding their way into real-world applications. Rather than being a fashionable research issue, the semantic web is slowly but surely becoming a reality. Semantic Web is an extension of the current Web in which data and information on the Web are defined and linked in a way that it can be used by computers not only for display purposes, but for automation, integration, and reuse of data across various applications. The Semantic Web allows computers to make more sense of the information on the Web with the result of facilitating better cooperation between computers and people. The semantic web is a recent initiative to take the World Wide Web much further and develop it into a knowledge representation and computing. The aim of semantic web is not only to support to access information on the web but also to support its usage. The semantic web encourages and facilitates the formulation of shared vocabularies and shared knowledge in the form of ontology's. The construction and design of ontology's for functions of web services is an area that is currently topic of active research: Requests to a service and the reaction of the server can be collected and learning methods, Identification of user behavioral patterns, Teaching aids etc.

The goal of Semantic Web Services research and development is to introduce semantics for service-descriptions and to enable automation for various tasks of a service integration process. Recent developments in Semantic Web services aim to introduce bottom up approach to service modeling allowing to build incremental layers on top of existing service descriptions while at the same time enhance existing SOA (Service Oriented Architecture) technologies.

Geospatial Semantic Web is an augmentation to the Semantic Web that adds geospatial abstractions, as well as related reasoning, representation and query mechanisms. Geospatial data and information that identify the geographic location (using precise scientific coordinates), and the characteristics of natural or constructed features and boundaries on the Earth. Geoprocessing is the manipulation of geospatial data and the analysis of these data to derive information. Geoprocessing is used to generate high-quality data, perform quality control checks on data, and undertake modeling and analysis.

Web Service Security represents a key requirement for today's distributed interconnected digital world and for the new generations, Web 2.0 and Semantic Web. To date, the problem of security has been investigated very much in the context of standardization efforts; these efforts however, have dealt mainly with adopting existing security techniques such as encryption for use in Web Services. The standards also focused on addressing the problem of security interoperability thought the development of standard formats for security assertions, tokens and credentials. Interoperability is certainly an important issue for Web Services in that easy and flexible service composition requires that security relevant information be seamlessly transmitted across different services. Often sensitive information is exchanged among users, Web Services, and software agents. This exchange of information has highlighted the problem of privacy. A large number of strategies employed to preserve people's privacy require users to define their respective privacy requirements and make decisions about the disclosure of their information. Personal judgments are made usually based on the sensitivity of the information and the reputation of the party to which the information is to be disclosed. In the absence of a comprehensive privacy preserving mechanism, no guarantees about the information disclosure can be made. The emerging Semantic Web is expected to make the challenge more acute in the sense that it would provide a whole infrastructure for the automation of information processing on the Web. On the privacy front, this means that privacy invasion would net more quality and sensitive personal information.[2]

2. Semantic Web and Semantic Web Services

The Semantic Web would be an extension of the current one, in which information is given well defined meaning, better enabling computers and people to work in cooperation. Thus the current web is basically comprised of documents, presented by computers and read by man would also include data and information that would automatically be handled by agents and utilities. [3-5] Researchers from universities and company's about semantics divided into three groups: The first deals with Semantic Web Services; the second presents Semantic Web processes; and the third deals with the applications for the Semantic Web. Exposing functionality in the form of Web Services is generally more attractive for market participants than publishing all relevant facts directly on the web. Refer to Table 1 which provides an overview of emerging semantic web services technologies.

TABLE 1. AN OVERVIEW OF EMERGING SEMANTIC WEB TECHNOLOGIES

Technology	Definition
OWL-S	OWL-S (formerly DAML-S) is a services ontology that enables software agents to discover, invoke, compose, and monitor Web resources.
OWL 2	OWL 2 extends the Web Ontology Language (OWL) with a small but useful set of features (EL, QL, and RL) that enable effective reasoning.
WSMO	The Web Service Modeling Ontology (WSMO) provides a conceptual framework and a formal language for semantically describing all relevant aspects of Web Services to facilitate the automation of discovering, combining, and invoking electronic services over the Web.
WSML	The Web Services Markup Language (WSML) provides a formal syntax and semantics for the WSMO and consists of several variants, such as WSML-Core, WSML-DL, WSML-Flight, WSML-Rule, and WSML-Full.
SWRL	The Semantic Web Rule Language aims to be the Semantic Webs standard rule language and is based on a combination of the OWL DL, OWL Lite, RuleML and so on.
RuleML	RuleML constitutes a modular family of Web sublanguages including derivation rules, queries and integrity constraints as well as production and reaction rules.
RIF	The Rule Interchange Format (RIF) aims to be the standard rule language of the Semantic Web for Rule Interchange.

3. Geo Spatial Semantic Web

Figure 1 represents the Semantic Web Layer (SWLayer) Cake. This is a very simplified representation of the principles and technologies involved in implementing the vision of the semantic web. [6-7] at the base of the cake, Unicode allows computers to consistently represent and manipulate text in most of the worlds writing systems, while URI (Uniform Resource Identifier) to identify resources on the web. XML (extensible Markup Language) Schemas are used to facilitate the sharing of structure data across different systems on the web. RDF Schema is an extensible knowledge representation language that provides basic elements for the descriptions of ontologies. SPARQL defines a standard way in which to communicate with RDF based services on the Web, and in this way creating interoperability between different RDF tools and services. "Web of Trust" will come into existence by propagating trust transitively. Finally, at the top of SWLayer cake is the issue of interfaces to interact with the semantic web. Geo Spatial semantic web can be viewed as an augmentation to the semantic web that adds Geospatial abstractions as well as related reasoning, representation and query mechanisms. [8].

Knowledge Management	Semantic Web layer cake	Main Purpose
Wise Use	SW Interface	Interaction with user
	Trust (Signature, Encryption)	Authentication. Trust worthiness of statements.
Knowledge	Proof	Explain reasoning and conclusions
	Rules, Logic	Infer unstated facts
	SPARQL, OWL, RDF Schema	Definition of vocabularies and shared meanings. SPARQL: Standardized queries for RDF databases.
Information	RDF	Create self-describing documents and metadata. Data interchange.
	XML Schema	
	XML	
Data	URI, Unicode	Identify Processes

FIGURE 1. SEMATIC WEB LAYER CAKE

4. Semantic Web Interoperation Security and Geospatial Semantic Web Privacy

Refer Figure 2 which represents the Architecture of the Semantic Web. Trust is an integral part of the semantic web architecture. [9-11] Trust management and Negotiations play an important role in semantic security. Often, sensitive information is exchanged among users, Web Services and software agents. This exchange of information has highlighted the problem of privacy. A large number of strategies employed to preserve people's privacy requires users to define their respective privacy requirements and make decisions about the disclosure of their information. Personal judgments are usually made based on the sensitivity of the information and the reputation of the party to which the information is to be disclosed. On the privacy front, this means that privacy invasion would net more quality and sensitivity personal information.

Trust		
Proof		Digital Signatures
Logic		Digital Signatures
Ontology Vocabulary		Digital Signatures
RDF + RDF schema		Digital Signatures
XML + NS + XML schema		
Unicode	URI	

FIGURE 2. ARCHITECTURE OF THE SEMANTIC WEB

5. Implementations and Validations

a. *Implementation of Spatial Location Based Service*

We had implemented a Location Based Service (LBS) which is an information service, accessible with mobile devices through mobile networks. [12] It includes services to identify an object in a particular location such as discovering a list of colleges or hospitals in a particular location. This application consists the logic for SMS receiving and sending through SMS push and pull mechanism. Any mobile user who wants to use these services can simply send a message stating whereabouts of his requirements to a cell which in turn is connected to the LBS server. Reply will be sent according to the requirements of the user by considering the respected database through the mobile connected to the server. Figure 3 provides the architecture of our location based service. Figure 4 provides output of the spatial application.

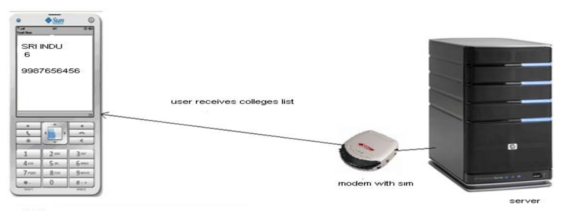

FIGURE 3. THE ARCHITECTURE OF OUR LOCATION BASED SERVICE

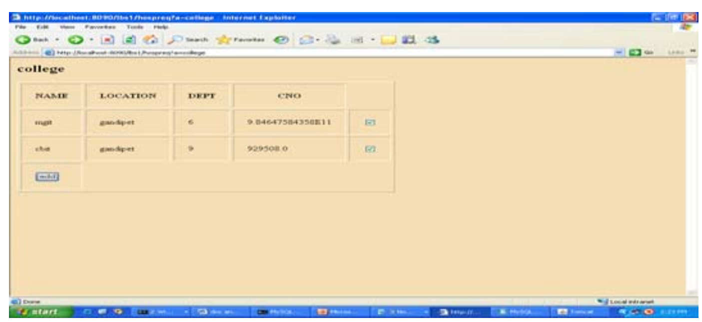

FIGURE 4. OUTPUT OF THE SPATIAL APPLICATION

b. *Implementation of Advanced Scheme of Digital Signature Standards*

We had implemented and integrated advanced scheme of Digital Signature Standards into the Semantic Web Security architecture. [13] We used ECC (Elliptic Curve Cryptography) digital signature scheme and a new forward-secure digital signature scheme is proposed in order to reform the limitations of DSA. In this new scheme, although the digital signature's private key is under the control of a one-way function and continually changed in different durations with time goes by; its public key remains the same. The attacker could not fake the older signature even if the private key is leaked out in some period of time. In this way this scheme makes sure of the security of former phases. The validity of the new scheme is proved and the security is analyzed in this implementation. Figure 10 represents the overall class diagram of the advanced DSA. Figure 5 shows the output screen shot of the application.

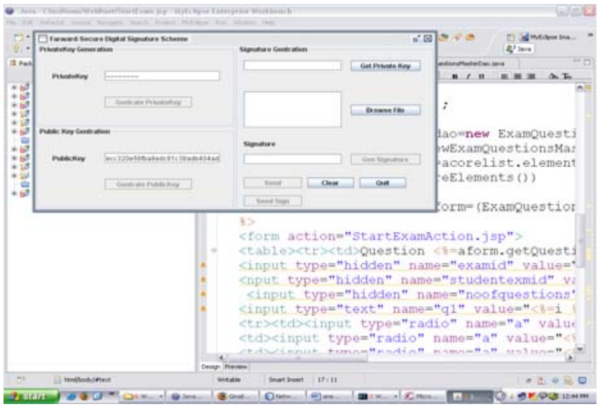

FIGURE 5. OUTPUT SCREEN SHOT OF THE ADVANCED DSA APPLICATION.

For details of implementation, source code and detailed UML diagrams, Please refer to the web site http://sites.google.com/site/upendramgircse

6. Conclusions

In this paper, we discussed various privacy issues integrating web services mining, semantic web, and geospatial semantic web by implementing a case study. Further work pertaining to Privacy Preserving data mining includes data perturbation and SMC (Secure Multiparty Computation). Privacy enforcement in web services and semantic web using decentralized models like distributed, registry-based and peer to peer reputation models. Finally geospatial privacy includes geoRBAC. The widespread deployment of location based services and mobile applications, as well as the increased concern for the management and sharing of geographical information in strategic applications like environmental protection and homeland security, have resulted in strong demand for spatially aware access control systems. These application domains impose interesting requirements on access control systems. In particular, the permissions assigned to users depend on their position in a reference space; users often belong to well-defined categories; objects to which permissions must be granted are located in that space; and access control policies must grant permissions based on object locations and user positions.

7. Acknowledgements

The authors wish to thank the following members for implementing these concepts: D.Shravani, P.Nagamani, P.Praneeth, Ch.Venkatabhilash, V.Deepthi, R.Naveen Kumar, Sesha Sai Ega.

8. References

[1] Hsinchun Chen, Xin Li, Michael Chau, Yi-Jen Ho and Chunju Tseng, “Using Open Web APIs in Teaching Web Mining,” *IEEE Transactions on Education*, vol. 52 no. 4, pp. 482–490, November 2009.

[2] Ravi Kothuri, Albert Godfrind and Euro Beinat, *Pro Oracle Spatial for Oracle Database 11g* Apress publishers, 2007.

[3] Valentina Janev and Sanja Vranes, “Semantic Web Technologies: Ready for Adoption?” IEEE IT Pro, pp. 8-16, September/October 2009.

[4] Tomas Vitvar, Jacek Kopecky, Jana Viskova, Adrian Mocin, Mick Kerrigan and Dieter Fensel, ”Semantic Web Services Architecture with Lightweight Description of Services”, Advances in Computers, Vol. 76, pp. 177 – 223, Elsevier Inc. 2009

[5] Martin Hepp,”Semantic Web and Semantic Web Services: Father and Son or Indivisible Twins?” IEEE Internet Computing, pp. 85 – 88, March April 2006.

[6] Rafael Moreno-Sanchez, “The Geospatial Semantic Web: What are its Implications for Geospatial Information Users?” pp. 588- 609, IGI Global, 2009.

[7] Sarah Spiekermann and Lorrie Faith Cranor “Engineering Privacy,” *IEEE Transactions on Software Engineering*, vol. 35 no. 1, pp. 67—82, January-February 2009.

[8] Lixin Fu, Hamid Nemati and Fereidoon Sadri *Privacy-Preserving Data Mining and the Need for Confluence of Research and Practice* IGI Global, 2009.

[9] Bhavani Thuraisingham, Natasha Tsybulnik and Ashraful Alam *Adminstering the Semantic Web: Confidentiality, Privacy and Management.* IGI Global, 2009.

[10] Yi Zhao, Xia Wang and Wolfgang A.Halang, “Security in Semantic Interoperation”, IGI Global, 2009.

[11] Shinsaku Kiyomoto, Keith M.Martin, and Kazuhide Fukushima,” Privacy-Aware Location Database Service for Granular Queries”, WISTP LNCS 5746, pp. 28—37, IFIP 2009.

[12] Rahul katarya and Dr.Satbir Jain, “Web Search with mobile incorporation Geographic information system” ; CSI Communications, Vol No 33, Issue No.8, November 2009.

[13] Neil daswari, Christoph kers and Anita kesavan, “Foundations of Security, What every Programmer needs to know” A press publishers, 2007.