Survey on Privacy Oriented Web Service Composition

Sandhiya R, Joe Dhanith P R, Dr Gunasekaran S

Abstract— In the web service environment privacy entity is used to determine when to release the private information depending upon the service providers. The query rewriting approach is used for querying and automatically composing DP services by using RDF views. Service oriented data mashup application is used to integrate the data from multiple data providers. This reveals the sensitive information from other data providers, creating flexible dynamic business process application. The attack model is introduced to analyse the social information from query log in Daas for encrypting the user information from the trusted client. The data in original form contains sensitive information about individuals this violates the privacy in the Privacy-Preserving Data Publishing (PPDP) method.

Index Terms— Daas services, DP services, RDF views, Mashup application, PPDP method.

1 INTRODUCTION

WEB Services connect computers and devices with each other using the Internet to exchange data and combine data in new ways. Web Services can be defined as software objects that can be assembled over the Internet using standard protocols to perform functions or execute business processes.Privacy Data also called information privacy is the aspect of information technology that deals with the ability of an organization or individual has to determine what data in a computer system can be shared with third parties. Information privacy usually relates to personal data stored on computer systems. The need to maintain information privacy is applicable to collected personal information, such as medical records, financial data, criminal records, political records, business related information or website data.The information privacy relates to different data types they are Internet privacy (online privacy) all personal data shared over the Internet is subject to privacy issues. Most websites publish a privacy policy details that the website's intend the use of collected online and/or offline data.

2 LITERATURE SURVEY

The web services technology faces problem to effectively discover the services based on their capabilities.BoualemBenatallah etal.,2003, [1] proposed a approach DAML-S ontology services that is used tosolve the problem here. The best match is given as request to web services combination approach and this approach enables to select the combination of web services to the best match is given as request and the algorithm for hyper graph is derived.The experimental results of test bed prototype that allows to evaluate the performance by an algorithm and the tool that enables to generate random XML-based services associated with service request is used to evaluate. Here it focuses on three cases varying in size of the application, web service and of the query. There is great difference in performance of the different versions of the algorithm concerning the real time execution of services.Finally the match between a service and request is described in DAML-S and the service is determined by comparing all the inputs and outputs of the query. The difference between query and its rewriting is effectively computed to improve the web service.Here comes a problem in minimal rewriting the queries using the views.

ArifTumer etal.,2003,[2] proposed that in the development of privacy mechanism there is minimum reveal of information about users so the privacy frame work for web services allows automatically negotiating the personal information. The web services declare their input parameter as optional and allow users how much of their personal information made available to services.The user must specify their privacy in different permission levels based on domain specific service ontology DAML-S. The privacy framework defines the web services and this service is used to store and manage user personal user data and provide single identity to pass personal information more easily. The stored personal data is generally limited to user identification and the trusted user makes available to store personal information. Finally the privacy preferences of a user define the rules to control the read access of personal information. The privacy frame work for web services is proposed to declare alternate data request if the mandatory input is not given by the user so this way becomes possible to automate the negotiation process with web service. The problem is mandatory input must begiven by the user otherwise it becomes impossible to automate the negotiation process with

- *Sandhiya R is currently pursuing masters degree program in Software Engineering in Coimbatore Institute of Engineering and Technology, India, PH-9789781120. E-mail: sandhiyarajagopal@gmail.com.*
- *Joe Dhanith P R is currently working in Computer Science and engineering in CIET , PH-8903939076. E-mail: joe.dhanith@gmail.com*

IJSER

web services.

Wolf TiloBalkeeta etal.,2004,[3] proposed that the skyline queries plays in retrieving over central databases application in web information system due to distributed retrieval over web sources. The retrieval algorithm presents heuristic, in speed up of the retrieval in on-line information services and the sampling schema allows the skyline for subsequent query refinement.Theexperimental results in mobile client device are used for the user query. The application running on the web server query may be applied to set of internet sources, collecting the individual results and combining engine that runs an algorithm to compute the overall best matching object. These final results have been aggregated according to each individual's user specification and best answer will be returned to the user.The skyline queries problem in web information system is addressed. The algorithm allows retrieving the skyline over distributed data sources. The number of advanced heuristics is presented to improve the performance towards real-time applications. The problem here is skylines and its quality is compared to a correct random sample of the actual skyline.

Nanzhang etal.,2005, [4] proposed that the issues related to sharing information in a distributed system which holds a private database. In semi-honest, behaviour has widely adopted for adversarial threats it also includes the malicious adversaries. The protocols effectively and efficiently protect privacy against the different kinds of malicious adversaries.The protocols are used to protect privacy against different kinds of malicious adversaries and architecture is proposed to share information using the trusted third party services. The experimental results of numerical measurement are used to demonstrate system security for all possible attacks and methods cannot be exhausted. There is no privacy disclosure occurs when μ>=2. Finally the issues related to privacy protection are information sharing whichbecomes an important and common application in distributed system. The problem in information sharing intersection in comparison to other operations in dealing with multiple parties in the system which contains correlated attacks from multiple adversaries.

Paolo Traverso etal.,2008, [5]proposed that SOC utilizes services that supports the low-cost composition for distributed applications in heterogeneous environments. The application components are assembled into network services that can be loosely coupled for flexible dynamic business process. The requirement of complex applications that requires the use of the service oriented computing paradigm.The SOC technique provides service foundation, service composition, service management and monitoring service-oriented engineering.The results related to services research roadmap introduce a logical service-based architecture to create an adaptive IT environment. It is used to provide the facilities for ensuring consistency across the organization and high availability of services and security and this provides a good quality of service QOS. This finally is used to create flexible dynamic business process application and the complex application requires the use of SOC. The problem is that it provides multiple-channel access to service and also manages and monitors it.

Mahmoud Barhamgi etal.,2008,[6] proposed that the data providing DPS, allows accessing the data through web service in the query like manner. The user queries require several service compositions. The model is introduced for the description of DPS and specification of service-oriented queries. The DPS is modelled over the mediated ontology in RDF views. Then the query rewriting algorithm is processed for queries over DPS.This approach is for querying and automatically composing of DPS is done. Then the query rewriting algorithms is used for processing queries over DPS.The architecture is used in the system for querying and compressing the DPS of data sources with different interfaces are expressed as DPS in WSDL with RDF views. The WSDL files each operation element is associated with RDF view and to each operation element "rdf query" is contained. The query mediator implements the service query model with the five components for the interactive query formulated for web based query interface and helps users to specify their RDF queries. Finally,this approach for querying, composing DPS is described as RDF views and the RDF query rewriting algorithms used to compose DPS. The problem is in composite service that cannot necessarily be reused to answer the same query for different values.

Junpei Kawamoto etal.,2009,[7] proposed that Data-as-a-Service (DaaS) is essential component of the cloud computing framework. The DaaS brings new risk in which data are currently stored and managed by the service provider and the user information is encrypted at a trusted client. The attack model is introduced to analyse the social information from query log in Daas and a solution is provided to a problem. The hashing method is used for optimizing and conversion.The experimental result is used in open dataset for evaluating query processing on a p2p network. It has a hierarchical categories and each category has more segments. The people might belong to some interest group in the dataset. The queries are created in the users groups to calculate the similarity.The new privacy problem arises for the security of social information to prove the Daas service which already provide. The attack method is used for expressing social information by analysing and rewriting queries. The problem is when social information differs from personal information because personal information is identified by individual users.

Benjamin C.M Fung etal.,2010, [8] proposed that the data in original form contains sensitive information about the individuals and such data will violate the privacy. The privacy preserving data publishing (PPDP) provides methods and tools for publishing useful information while preserving the data privacy and many approaches have been proposed for different data publishing and to evaluate the PPDP.The experiment is done on the data that is so sensitive in data mining

IJSER

patterns cannot be generated. The sensitive values are chosen by individual record the owners are sufficient because an attacker can use association rules from data to estimate values. The heuristic algorithm is used to suppress minimum set of values to such attacks.Finallythe privacy-preserving data publishing approach is used for information sharing while preserving individual privacy and protecting sensitive information. There is a problem in degradation of data/service quality, loss of valuable information is increases the cost and complexity.

Anne H.H etal.,2010, [9] proposed that the client server application initially was not designed for the need of integration. The composite application is reconfigured by changing the set of components. The service-oriented architecture (SOA) is used to integrate the component that has been provided by web services and for building composite applications. The web-service based integration requires extra programming when integrating with non web service components. The standard semantic web matching algorithm is used as components for non-web-service-based components.The standard semantic web matching algorithm is used as equivalent components for non-web-service-based components.The result of graphical user interface component was created which is used to send the currently executing composite application drives and test cases. The "view filter" and "display filter" allows the user to set graphical user interface search criteria. The WSDL is included in the search request to match the composite application individually by the various techniques.Finally the semantic web and web services matching composition framework problem has been solved by using various technologies. There is a problem when component is not being annotated as priori at runtime it is impossible to mash up the capability within composite application.

Roman Vaculin etal.,2010, [10] proposed that the web service provide the access to the structure data in the data sources. The data providing services (DPS) is the service available for interactions with services requesters such as for composition and mediator. The RDF Model view is to used represent the content provided by the DPS and the match making algorithm is developed for match between the DPS as RDF view based on the data providing services.The DPS uses the SOAP version of the SPAROL protocol for dynamic configuration and allow easy interactions of service requesters with DPS.The service requester with three inputs of types A,B,C and this requires two outputs D & E to be returned. First the DPS must be registered with match maker by sending a (message1) and also describes the data is provided by DPS. The DPS should contain data to allow the output type as D & E to be retrieved based on inputs. The search request R is sent to the match maker(message2) and DPS must identify the possible match for the request after discovering(message 3) the service requester must be able to communicate with DPS to formulate the SQL or SPARQL query for retrieving the data and it execute the process to sending appropriate SOAP messages to 4&6. Then finally the given search request R, the DPS must be able to identify the query Q to retrieve the requested information with the provided inputs. Thus the data providing service using RDF view provide flexible modelling of structure and data services and their easy integration within the web service infrastructure and the service requesters.The problem arises when the information can be combined with degrees of match is not clear between the input and output.

Salah EddineTbahriti etal.,2011, [11] proposed that the web service privacy model deals with the input data, output data and operation. The matching protocol is used for partial and total privacy compatibility. The negotiation model is used to reconcile client's requirements which provide in case of incompatibility. Theexperimental results in service-based system are that client/provider specifies how to handle private information. The privacy policy (PP) and privacy requirement (PR) is used for set privacy practices. The PCM algorithm is used for cost-model. Finally privacy deals with data and operation that client/providers specify the privacy practices. The negotiation model is used to reconcile the incompatibility between the client and providers privacy. The problem is that the technique used in this does not give proper privacy-preserving service composition.

Benjamin C.M. Fung etal.,2012, [12] proposed that the mashup is a web technology which allows different service provides to flexibility integrate and deliver high services to customers. The integrating data from multiple sources brings three challenges, the first one is joining multiple data sets together could contain sensitive information to other data provides, the second one is integrated mashup data reveal their specific sensitive information that was not available before mashup and the third one ismashup data from multiple sources contain many data attributes. To resolve this privacy problem a service-oriented architecture with privacy-preserving data mashup algorithm is used to address the challenges used for preserving both privacy information on the mashup data. The mashup is implemented in a distributed web service environment and each data provider is connected to LAN. The adult census data set is used it has six numerical attributes and eight categorical attributes and binary class attributes to represent the two income levels. The two private tables Ta and Tb. Ta contain first q attributes Tb remaining five attributes. Here we consider the divorced and marital-status as sensitive and remaining 13 attributes as QID. The common UID is added to both tables for joining. Finally the data mashup application is used to generalize their privacy information and the problem of privacy-preserving data mashup and is achieved by the mashup data process. There is an insufficient problem for anonymous data to satisfy the privacy requirement.

3 Experiment Results

The comparison has been made between Service Dis-

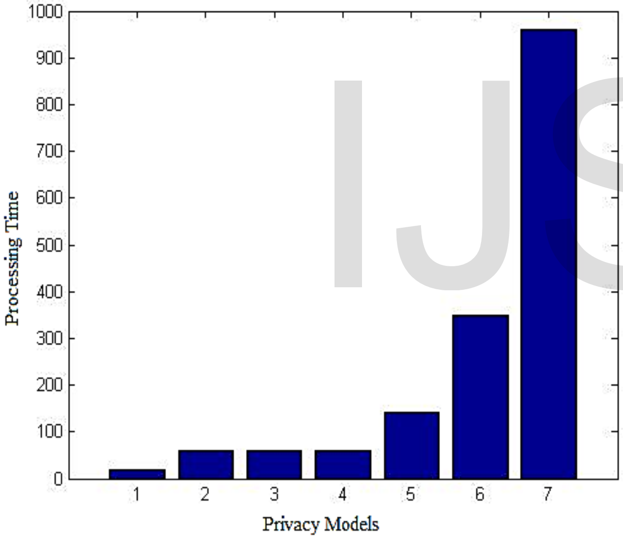
IJSER

covery, Semantic Matchin, Privacy compatibility ,Query Conversion, Private Data Mashup, RDF Query Rewriting and Privacy Model comparison in processing time and it is denoted in seconds.

Table.1The comparison result has shown as follows.

Methods	Processing time
Service Discovery	20s
Semantic Matching	60s
Privacy compatibility	60s
Query Conversion	60s
Private Data Mashup	142s
RDF Query Rewriting	350s
Privacy Model	960s

The above comparison results is shown in graph as follows:



4 CONCLUSIONS AND FUTURE WORK

In this paper, the various approaches like Service Discovery, Semantic Matching, Privacy compatibility, Query Conversion, Private Data Mashup, RDF Query Rewriting, and Privacy Model is used in web service privacy. Here Service Discoverycan be used to integrate and deliver high services its processing time is 20s,Semantic Matching can be used for privacy frame work of web services its processing time is 60s,Privacy compatibility can be used for privacy checking its processing time is 60s,Query Conversioncan be used for optimizing its processing time is 60s,Private Data Mashup can be used for preserving both privacy and information and its processing time is 142s,RDF Query Rewriting can be used for querying and composing the DP services its processing time is 350sand Privacy Modelcan be used to set privacy practices on the collected data its processing time is 960s. In future, measures should been taken to protect the composition results from privacy attacks before the final result returned by the mediator.

IJSER

REFERENCES

[1] B. Benatallah, M.-S.Hacid, C. Rey, and F. Toumani.Request rewriting-based web service discovery. In D. Fensel, K. P. Sycara, and J. Mylopoulos, editors, International Semantic Web Conference, volume 2870 of Lecture Notes in Computer Science, pages 242–257. Springer, 2003

[2] A. Tumer, A. Dogac, and I.H. Toroslu, ‘‘A Semantic-Based User Privacy Protection Framework for Web Services,’’ in Proc. ITWP, vol. 3169, Lecture Notes in Computer Science, B. Mobasher and S.S. Anand, Eds., 2003, pp. 289-305.

[3] Efficient Distributed Skylining for Web Information Systems-Wolf-TiloBalke, Ulrich Güntzer, Jason XinZheng.

[4] N. Zhang and W. Zhao, “Distributed Privacy Preserving Information Sharing,” Proc. 31s Int’l Conf. Very Large Databases (VLDB), pp. 889-900, 2005.

[5] M.P. Papazoglou and B. Kratz, “Web Services Technology in Support of Business Transactions,” Service Oriented Computing and Applications, vol. 1, no. 1, pp. 51-63, 2001.

[6] M. Barhamgi, D. Benslimane, and B. Medjahed, ‘‘A Query Rewriting Approach for Web Service Composition,’’ IEEE Trans.Serv. Comput., vol. 3, no. 3, pp. 206-222, July-Sept. 2010.

[7] J. Kawamoto and M. Yoshikawa, ‘‘Security of Social Information from Query Analysis in DaaS,’’ in Proc. EDBT/ICDT Workshops, 2009, pp. 148-152.

[8] B. C. M. Fung, K. Wang, R. Chen and P. S. Yu, “Privacy-Preserving Data Publishing: A Survey of Recent Developments”, In Journal of ACM Computing Surveys, 2010. vol. 42, no. 4, pp 1-53.

[9] A.H.H. Ngu, M.P. Carlson, Q.Z. Sheng, and H.-Y. Paik,‘‘Semantic-Based Mashup of Composite Applications,’’ IEEE Trans. Serv. Comput., vol. 3, no. 1, pp. 2-15, Jan.-Mar. 2010.

[10] R. Vaculı´n, H. Chen, R. Neruda, and K. Sycara, ‘‘Modeling and Discovery of Data Providing Services,’’ in Proc. IEEE Int’l Conf. Web Serv., Washington, DC, USA, 2008, pp. 54-61.

[11] S.-E. Tbahriti, B. Medjahed, Z.Malik, C. Ghedira, andM.Mrissa, ‘‘MeerkatVA Dynamic Privacy Framework forWeb Services,’’ in Proc. Web Intell., O. Boissier, B. Benatallah, M.P. Papazoglou, Z.W. Ras, and M.-S. Hacid, Eds., 2011, pp. 418-421.

[12] B.C.M. Fung, T. Trojer, P.C.K. Hung, L. Xiong, K. Al-Hussaeni, and R. Dssouli, ‘‘Service-oriented Architecture for High-Dimensional Private Data Mashup,’’ IEEE Trans. Serv. Comput.,vol. 5, no. 3, pp. 373-386, 2012.

[13] U. Bellur and R. Kulkarni. Improved matchmaking algorithm for semantic web services based on bipartite graph matching. In ICWS, pages 86–93. IEEE Computer Society,2007.

[14] M.J. Carey, “Data Delivery in a Service-Oriented World: The BEA aquaLogic Data Services Platform,” Proc. SIGMOD Conf., pp. 695-705, 2006.

[15] Privacy-Enhanced Web Service Composition Salah-EddineTbahriti, ChirineGhedira, BrahimMedjahed, and Michael Mrissa IEEE Transactions on Services Computing, vol. 7, no. 2, april-june 2014.

[16] S.-E. Tbahriti, M. Mrissa, B.Medjahed, C. Ghedira, M. Barhamgi, and J. Fayn, ‘‘Privacy-Aware DaaS Services Composi-

tion,'' in Proc. DEXA I, vol. 6860, Lecture Notes in Computer Science, A. Hameurlain, S.W. Liddle, K.-D. Schewe, and X. Zhou, Eds.,2011, pp. 202-216.

[17] Y. Lindell and B. Pinkas, "Secure Multiparty Computation for Privacy-Preserving Data Mining," J. Privacy and Confidentiality, vol. 1, no. 1, pp. 59-98, 2009.

[18] G. Wang, S. Yang, and Y. Han, "Mashroom: End-User Mashup Programming Using Nested Tables," Proc. 18th Int'l Conf. World Wide Web (WWW '09), pp. 861-870, 2009.

[19] L. Motiwalla and X. Li, "Value Added Privacy Services for Healthcare Data", In Proceeding of 6th World Congress on Services,SERVICES 2010.

[20] H. Kargupta, K. Das, and K. Liu, ''Multi-party, Privacy-Preserving Distributed Data Mining Using a Game Theoretic Framework,'' in Proc. 11th Eur. Conf. Principles PKDD, 2007,pp. 523-531.

IJSER