International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 7 4815 - 4817

A Survey Paper on Implementing Service Oriented Architecture for Data Mining

Vasant Patil
Department of Computer Engineering,
SKN Sinhgad Institute of Technology and Science,
Savitribai Phule Pune University.
vasantpatil11@gmail.com

Kalpana Kadam
Department of Computer Engineering,
SKN Sinhgad Institute of Technology and Science,
Savitribai Phule Pune University.
more.kalpana@gmail.com

***Abstract*—** Web service is working with the web with an object or component to achieve the communication between the distributed applications and between the different platforms through a series of protocols. Web Service provides a set of standard types systems, rules, techniques and internet service-oriented applications for communication between the different platforms, different programming languages and different types of systems to achieve interoperability. This survey paper gives the application of web service for data mining also we build a data mining model based on Web services and going forward it is possible to build a new data mining solution for security according to the prototype of a dynamic web service based data mining process system.

Keywords- *SOAP; SOA; Web Mining; Database;*

*****

I. INTRODUCTION

With Web technology, data on networks has become increasingly large and complex. No matter managers or network users are required from these complex Web data to get useful information and knowledge, so Web data mining is necessary to adapt this demand. Web data mining can extract undiscovered unexpected possibly useful information and knowledge from much incomplete noise ambiguous random practical application data on WWW network[6]. It is a new commercial information mining technology. Its main characteristic is to extract key data to support business decision-making from business database through extraction, conversion, analysis and other models transaction.

II. LITERATURE SURVEY

Jonathan Lee, Shin-Jie Lee, and Ping-Feng Wang, A Framework for Composing SOAP, Non-SOAP and Non-Web Services[2]. Recently, there is a trend on developing mobile applications based on service-oriented architecture in numerous application domains, such as telemetric and smart home. Although efforts have been made on developing composite SOAP services, little emphasis has been put on invoking and composing a combination of SOAP, non-SOAP, and non web services into a composite process to execute complex tasks on various mobile devices. Main challenges are twofold: one is how to invoke and compose heterogeneous web services with various protocols and content types, including SOAP, RESTful, and OSGi services; and the other is how to integrate non-web services, like Web contents and mobile applications, into a composite service process. In this work, we propose an approach to invoking and composing SOAP, non-SOAP, and non-web services with two key features: an extended BPEL engine bundled with adapters to enable direct invocation and composition of SOAP, RESTful and OSGi services based on Adapter pattern; and two transformation mechanisms devised to enable conversion of Web contents and Android activities into OSGi services. In the experimental evaluations, we demonstrate network traffic and turnaround time of our approach are better than those of the traditional ones.

San-Yih Hwang, Chien-Ching Hsu, Chien-Hs iang Lee, Service Selection for Web Services with Probabilistic QoS [3]. Most of the former works in Web service selection and recommendation treat the QoS values as constants. However, QoS values of a service as perceived by a given user are intrinsically random variables because QoS value prediction can never be precise and there are always some unobserved random effects. In this work, author have address the service selection problem by representing services. QoS values as discrete random variables with probability of satisfying constraints imposed on the composite service is high and the execution time is reasonable. This paper gives a method starts with an initial Web service assignment and incrementally adjusts it using simulated annealing. This paper conduct several experiments and the results show the authors approach generally performs better than previous work, such as the integer programming method and the cost-driven method.

Fumiko Satoh, Michiaki Tatsubori, Yuichi Nakamura, Nirmal K. Mukhi, Kouichi Ono, Methodology and Tools for End-to-End SOA Security Configuratin,[4].The configuration of non-functional requirements, such as security, has become important for SOA applications, but the configuration process has not been discussed comprehensively. In current development processes, the security requirements are not considered in upstream phases and a developer at a downstream phase is responsible for writing the security configuration. However, configuring security requirements properly is quite difficult for developers because the SOA security is cross-domain and all required information is not available in the downstream phase. To resolve this problem, we clarify how to configure security in the SOA application development process, and define the developers roles in each phase. Additionally, supporting technologies to generate security configurations are proposed: Model-Driven Security and Pattern-based Policy Configuration. This paper gives a methodology for end-to-end security configuration for SOA applications and tools for generating detailed security configurations from the requirements specified in upstream

phases model transformations, making it possible to configure security properly without increasing developers workloads.

Yang Li, Shi ZhongZhi, An Efficient Data Mining Framework on Hadoop using Java Persistence API [5]. Data indexing is common in data mining when working with high dimensional, large-scale data sets. Hadoop, a cloud computing project using the MapReduce framework in Jave, has become of significant interest in distributed data mining. A feasible distributed data indexing algorithm is proposed for Hadoop data mining, based on ZSCORE binning and inverted indexing and on the Hadoop Sequence File format. A data mining framework on Hadoop using the Java Persistence API(JPA) and MySQL Cluster is proposed. The framework is elaborated in the implementation of a decision tree algorithm on Hadoop. In this paper the data indexing algorithm with Hadoop MapFile indexing , which performs a binary search , in a modest cloud environment. The result show the algorithm is more efficient than native MapFile indexing. This paper compare the JDBC and JPA implementations of the data mining framework. The performance show the framework is efficient for data mining on Hadoop.

III. PROPOSED APPROACH: FRAMEWORK AND DESIGN

A. *Existing System*

Existing system consist of SOAP messages exchange includes similar messages. Also different evaluation encoding technique have been emerged for SOAP messages performance and security. By checking and investigating the impact of security policy evaluation on WS performance. This security policy evaluation consists of checking and verifying the access and usage security constraints defined in Fig. Proposed System Architecture of SOAP message.

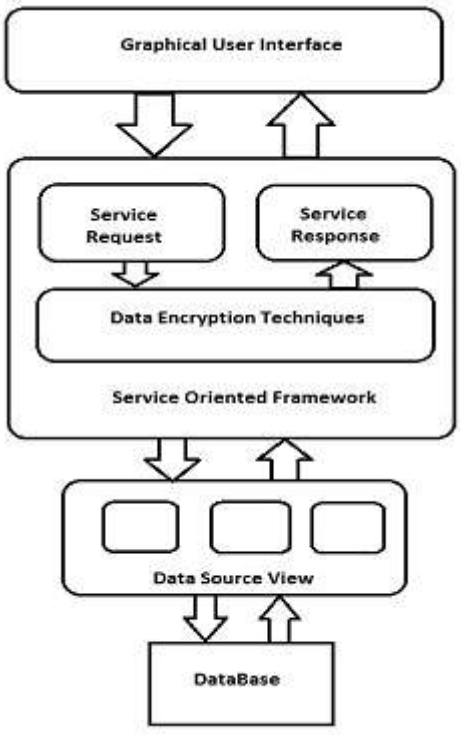

Fig 1. Proposed System Architecture

The security evaluation (i.e. checking and verifying) the access and usage security constraints which is performed at both end points of client and server end points. This is achieved by using set of rules (i.e. actions) and security constraints (e.g. authorization, signatures, encryption)[1]. So the security policy rules are characterized by 3-tuple entity. (a) Subject: This specifies the users to whom the rule applies. (b) Object: This identifies which message or portion of the corresponding policy should be applies. (c)Rule: The rule specifies the action(it can be access, signature or encryption) authorized for the policy subject on the policy object. Current system/existing system checks various techniques for optimization of WS-Security performance including (a)Digest based caching, per-hashing, and (b)on-demand canonization[1].

B. *Proposed Work*

The proposed system uses an encryption/decryption techniques. The core technologies of Web services framework include SOAP (simple object access protocol), WSDL (web services description language) and UDDI (universal description discovery and integration) and also their expressions standard XML documents. Service-Oriented Architecture (SOA) of Web services architecture is shown in Figure 1. Web Service provider describes its services through the WSDL(i.e. interface) and ask Web services to register server about the description. Registration server updated its directory and release on web according to the WSDL interface/description and UDDI agreement. Users of this system request for registration on server before using Web services to access to Web services and then communicate through connections set used by the SOAP protocol and Web services providers. The core framework of the given proposed web data mining system is based on web service which is similar to the Web data miming process system, including a Single Integrated User Interface, Web Data Mining Process Designer, and Web Data Mining Process Executor. Each activity in Web data mining process is viewed as a web service provided by Web data mining service providers on web. The Web data mining service providers uses common KDD functions for data per-processing, algorithms of Web data mining and visualization analysis. The aim of using web services in this approach is to achieve universal/independence, interoperability between applications uses Web standards. Web Services are loosely coupled integration system which allows flexible integration of heterogeneous systems of different domains consist of business-to-consumer, business-to business and enterprise application integration.

IV. ALGORITHM OF PROPOSED SYSTEM

A. *Process of SOAP Messages signing*

```
// Instance creation of the proxy
WeblogProxy prxy = new WeblogProxy();
// Creation of a KerberosSecurityToken
string targetPrincipal = ”host/” + new Uri(prxy.Url).Host;
KerberosToken kebrtkn =
new     KerberosToken(targetPrincipal);
// Adding SecurityToken in the Request
proxy.RequestSoapContext.Security.Tokens.Add(kebrtkn);
//Process of Signing the message with a signature object
MessageSignature mssig = new MessageSignature(token);
```

```
proxy.RequestSoapContext.Security.Elements.Add(mssig);
```

B. *Process of SOAP Messages encryption*

```
// Get the X509SecurityToken
SecurityToken Sectkn = GetX509Token();
// Create an instance of the proxy
WeblogProxy prxy = new WeblogProxy();
// Create and add the encrypted username token
EncryptedData encpt = new EncryptedData(Sectkn);
prxy.RequestSoapContext.Security.Elements.Add(encpt);
// Send the request
confirmed = prxy.AddEntry(newEntry);
```

V. CONCLUSIONS

Web data mining is a upcoming technique emerging with fast development of internet. Some web data structures are complex, hence a challenge for data mining is how to communicate between the distributed and heterogeneous databases. Handling requirements of SOA applications already implemented but the security configuration processes have not been well defined. So the proposed system implements the security configuration and helps to improve the system performance better and extend it in to the area Web structure mining, Web content mining and Web usage mining integrated system.

REFERENCES

[1] Joe M. Tekli, Member, Ernesto Damiani, Richard Chbeir, Gabriele Gianini:SOAP Processing Performance and Enhancement,IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 5, NO. 3, JULYSEPTEMBER 2012.

[2] Jonathan Lee, Shin-Jie Lee, and Ping-Feng Wang, A Framework for Composing SOAP, Non-SOAP and Non-Web Services. DOI 10.1109/TSC.2014.2310213, IEEE Transactions on Services Computing.

[3] San-Yih Hwang, Chien-Ching Hsu, Chien-Hs iang Lee, Service Selection for Web Services with Probabilistic QoS, DOI 10.1109/TSC.2014.2338851, IEEE Transactions on Services Computing.

[4] Fumiko Satoh, Michiaki Tatsubori, Yuichi Nakamura, Nirmal K. Mukhi, Kouichi Ono, Methodology and Tools for End-to-End SOA Security Configuratin, 2008 IEEE Congress on Service 2008-Part I.

[5] Yang Li, Shi ZhongZhi, An Efficient Data Mining Framework on Hadoop using Java Persistence API, 2010 10th IEEE International Conference on Computer and Information Technology (CIT 2010).

[6] Chunying Chen, Xiongwei Zhou and Jianzhong Zhang Web Data Mining System Based on Web Services, 2009 Ninth International Conference on Hybrid Intelligent System.

[7] N. Abu-Ghazaleh and M.J. Lewis, Differential Deserialization for Optimized SOAP Performance, Proc. ACM/IEEE Conf. Super- computing, pp. 21-31, 2005.

[8] N.Abu-Ghazaleh, M.J.Lewis, and M.Govindaraju, Performance of Dynamic Resizing of Message Fields for Differentisl Serialization of SOAP Message, Proc. Intl Symp. Web Service and Applications. Pp. 783-789.2004.

[9] N. Abu-Ghazaleh, M. Govindaraju, and M.J. Lewis, Optimizing Performance of Web Services with Chunk-Overlaying and Pipelined-Send, Proc. Intl Conf. Internet Computing (ICIC), pp. 482-485,2004.

[10] R. Chinnici, J.J. Moreau, A Ryman, and S. Weerawarana, Web Service Description Language (WSDL) Version 2.0 Part 1:Core Language, W3C recommendation, http://www.w3.org/TR/wsdl20, Aug.2009.

[11] M.G. Kostoulas, M. Matsa, N. Mendelsohn, E. Perkins, A. Heifets, and M. Mercaldi, XML Screamer: An Integrated Approach to High Performance XML Parsing, Validation and Deserialization, Proc. 15th Intl Conf. World Wide Web (WWW 06), pp. 93-102, 2006.

[12] Web Services Security:SOAP Message Security 1.1, http://www.oasisopen. org/committees/download.php/16790/wss-v1.1-spec-os-SOAPMessageSecurity.pdf.

[13] IBM and Microsoft, Security in a Web Services World: A Proposal Architecture and Roadmap, http://www.ibm.com/developerworks/webservices/library/specification/wssecmap/,2002.

[14] T. Bray, J. Paoli, C. Sperberg-McQueen, Y Mailer, and F.Yergeau, Extensible Markup Language(XML) 1.0, fifth ed. W3C, http://www.w3.org/TR/REC-xml, Nov.2008.