International Journal of Engineering & Technology, 7 (2.32) (2018) 150-152

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Privacy Proliferation of Customized Web Search Engine

N. Arunachalam[1*], S.Radjou[2], P.Aravindan[2], T.Sivagurunathan[2]

[1]*Assistant Professor, Department of IT, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry - 605107, India*
[2]*Student, Department of IT, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry - 605107, India*
**Corresponding author E-mail: narunachalam85@gmail.com.*

Abstract

In last few years the illegal disclosure of user privacy in web search engine has become more serious. Protecting and Pre-venting user privacy from illegal disclosure is attracting the interest among researchers in recent times. Existing web search engines do not consider the privacy of the users. Search engines tend to collect all the information from the user. A system to ensure the privacy of the user is essential. Hence, the Personalized Web Search (PWS) method was put forward to take control over the amount of information that the user can provide to the search engines. This PWS provides privacy protec-tion in web search system and minimize the information dis-closure of the user related to privacy through a customizable web-search.

Keywords: *Web Service, Privacy Disclosure, Personalized Web Search (PWS).*

1. Introduction

Web Service is a domain where all the transaction occurs each day. But there is lot of room for improvement in the perspective of privacy, especially in search engines. Search Engines are most prevalent online activity act as gatekeepers and channeling con-sumers to the content they seek. In addition to the search results the search engines cover the IP address, Country, City, Zip, Lati-tude, Longitude, ISP, Autonomous System, Number, Browser, Operating System, and GeoIP Database. Typically, Search engines display advertisements based on search results and through prefer-ences, Geographic locations and websites visited. Finally, the integration of modern search engines has become a threat to user's privacy [1].

Most of the web services tend to collect all the information from the user without any means of customization or approval from the user. Majority of the users are concerned about their privacy on web and are reluctant in availing many services. So, it is inevitable to provide a mechanism for the users in which they can decide what are the information they are going to share in the web plat-form. This objective can be attained by a mechanism suggested in this work, where privacy disclosure on web is made possible through Personalized Web Search (PWS) [2].

2. Backgrounds and Related Work

Search Engines like Bing, Yahoo and Google which displays ad-vertisements carrying search results where the right to privacy has been violated by these actions. Semantic Web which arose as a novel structure called Ontology and Further developed using Web Ontology Language (OWL) [3]. Ontology focuses on Security aspects for Semantic Web Services, Denker [4]. According to the authors of [5], Privacy is "the right to be let alone under all cir-cumstances". Ontology has been created for storing privacy poli-cies of the service providers incorporating many privacy related parameters [6].

3. Personalized Searching Methods

Personal Privacy is an important aspect of human life. Privacy is the right of a person to specify what information about him is disclosed to others. While accepting the privacy policy of the service provider, most of the users think that they are protecting their privacy by accepting the policy which informs them about the privacy rights they are surrendering to service providers. The following section provides a brief scenario of current trends adapted in providing privacy of user.

3.1. Based on Search Histories

Based on the user interaction with a Search Engine the user pro-file is constructed. Google uses a wrapper for monitoring the user activities called Google Wrapper. This Wrapper logs the queries, search results and clicks which are performed per user. User profiles are constructed based on Wrapper Information and it is weighted based on topic. When a user query was submitted then the results were ranked based on user profiles [7].

3.2 Search Engines Privacy Policy

Search Engine is used to search and share information to com-municate with other people. Information that search engine collect from us like name, email address, telephone number, credit card number, device information, log information, location information, Unique Application Number, Local storage etc. These are the information which almost all search engines operate and share user's personal data's [8].

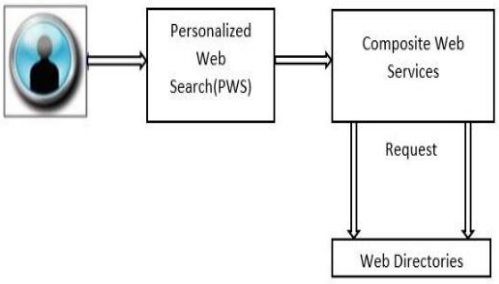

Fig1. Existing Search Engine Privacy Policy without customization

4. Research Directions

The impact of privacy on web environment is possible with the above said personalized web sphere methods where each one is unique of their own. But there are inconsistencies as well in each method. So, the best solution would be opting out the combinations to overcome all the misfortunes that may arise in the future. The combining of methods would result in better privacy protection that several authors have come up with. But one solution that is capable of out running all the inconsistencies is the privacy protection at the level where the application is being created, which is privacy features being embedded with the design process. One such method that suits is Ontology based approach. And combining that approach with an algorithm that is best of its kind in privacy enhancing and preserving is K-Anonymity algorithm. Where it works by generalizing the user data elements with the various level of preservation features that is about to be made. This results in an area where, even if the system gets into the hands of intruder, it is impossible to locate the user who is been responsible for the whole data. So, this is the one among the best solution that can produce elevated level of granularity with the amount of data specified to the system. The granularity level here is the level of privacy protected.

5. Proposed Framework

The proposed system contains the privacy customization feature added with the actual design process of the system itself, which is made possible with an approach called Ontology Based Approach (OBA). OBA on representing privacy-related context information in a formal and unambiguous way, the privacy ontology servers' act as a shared model for exchanging privacy policies between users and context-aware environments, and among users in the environments. More importantly, we expect to take advantage of logic-based inference capability, which is inherent in ontology-based context models, to facilitate automated processes in privacy interaction between users and the context-aware system. Before describing the privacy ontology, we present a summary of literature survey on individual concerns over privacy and privacy protection in context-aware ubiquitous computing environments. It helps justify the need of the privacy ontology solution.

Through this approach, the user gains the privilege of selecting the type of information that they intend to provide to the system. Further this approach is enhanced by integrating Privacy enhancement feature. Thus, this Ontology-based approach aims at providing all the privacy features at the design process itself, so that a complete customization on privacy is integrated with the system. In the existing web services, whenever a user submits a request, it is directed through the web directory and the user receives the response. But in the proposed work, a middleware is introduced between the user and the webservice. The middleware Personalized web search aid the user in customizing the information, they intend to share with the web service. The system architecture of the proposed system is shown in the figure 2.

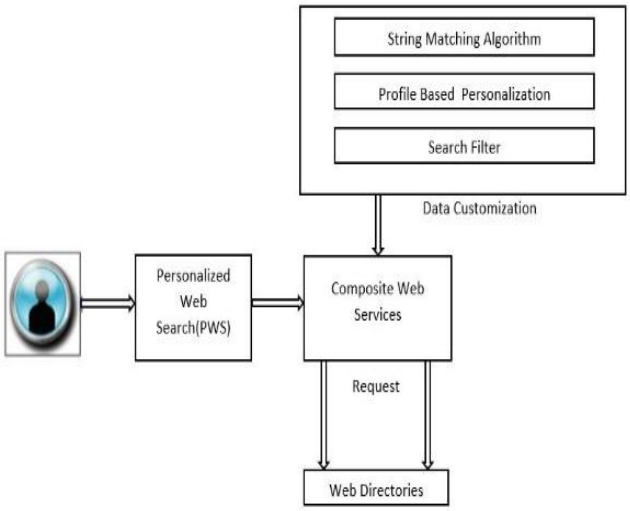

Fig 2. Proposed Search Engine with Privacy Customization

The PWS is the major element which assists the user in customizing the information they wish to share with the webservice .The PWS system is accompanied by data customization unit (DC). The three major components of DC are String Matching algorithm (SMA), Profile based personalization (PBP) and a search Filter (SF). The DC unit is the major entity in our proposed system and it is briefed in the succeeding section.

5.1. String Matching Algorithm

The SMA calculates the similarity between two strings as described below. We have adapted the Algorithm suggested by Oliver []. This implementation does not use a stack as in Oliver's pseudo code, but recursive calls which may or may not speed up the whole process. The complexity of this algorithm is O (N**3) where N is the length of the longest string.

```
int similar text (string $first , string $second[
float &$percent ] )
where,
first – the first string
second – the second string
percent – by passing a reference as third
argument,
```

Procedure1. Matching Function

Whenever this function similar text () finds similar data, it will calculate the similarity in percentage [9].

5.2. Profile Based Personalization (PBP)

PBP Framework facilitates the user to select the amount of information that they provide. Configured selections on framework collects the data where other data omitted from collecting [10]. The current search engines that collects all the information are entitled in the design process here as in format of checkboxes to select, so whatever the checkboxes is selected, all that information is collected from the user whereas the other information will not be collected [11]. There is also room for when the user changes mind, and again deselects the checkboxes that has been already selected, if that is done so, then the information that is collected previously also will be deleted. Examples of information that are opted to collected are IP Address, Location, ISP, Geo IP Database Information Of the country and etc [12].

Fig 3. User Customizable Web Search.

5.3. Search Filter (SF):

SF Framework provides data refinement with two options. The first option does not collect any of the data from the user and keyword specific data entered by user won't be collected. There will a checkbox option which when checked collects none of the data from the user [13]. This option leaves no room for search recommendations for user in the future. The second option paves the way for not collecting the data regarding option. This won't collect data regarding the items that the user has entered. The user can add/delete items of their wish at any time. And the last option is when the user wishes to delete all the information that has been entered into the system's database previously can be deleted if that delete previously collected data option is selected [14].

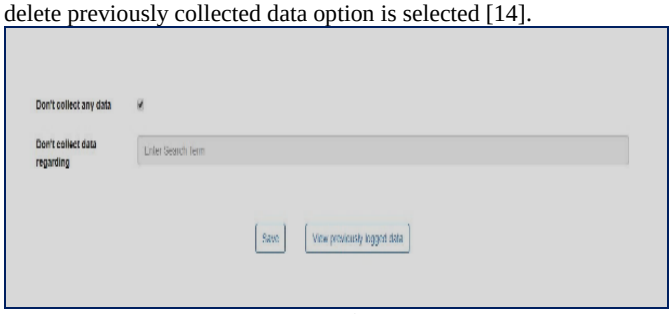

Fig 4. Search Refined Page.

6. Results & Discussions

The user profile was created using information collected implicitly. User interests on different web pages were modelled to correlate them with the search results user clicked [15]. The concept of personalized anonymity prevented privacy intrusion. The effect of personalization under different circumstances was studied. And a novel approach based on ontology was adopted for profile construction and support of the user interest was used for measuring privacy. The user privacy was provided using this novel method [16].An analysis was done based on the trust value provided by the user for the web service. Trust value of the web-service was also calculated for other privacy providing techniques and comparison is made. The graphical illustration given below infers that our proposed algorithm performs better than others in the perspective of trust value provided by the user [17].

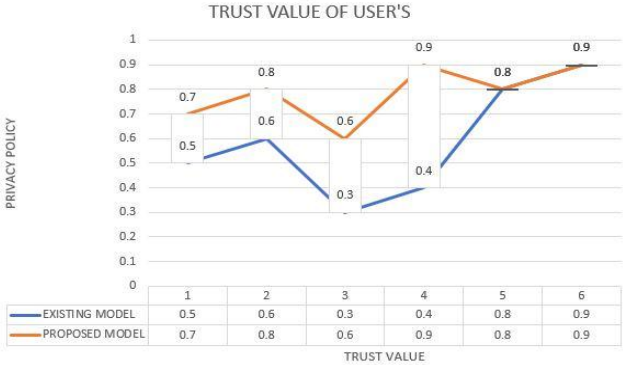

Fig 5.Comparison of User's Trust Value.

7. Conclusion

Personalized Web Search (PWS) improves the quality of search services on the Internet [18]. The need for privacy and privacy risks related to the different approaches of personalization were studied. Privacy preservation methods are used in PWS to prevent leakage of personal information on the Internet. The paper presents a novel technique used for personalization [19]. The PWS approaches gives option for user to customize data related to privacy, where in the user opts out the type of data that they intend to provide for the system that they use. This paves way for privacy preservation in the web environment.

References

[1] International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014) Conference Organized "Privacy Dis closure in Web Services Paradigm" Rekha Bhatia, Manpreet Singh The Value of Web Search PrivacySeptember/October 2015 Co-published by the IEEE Computer and Reliability Societies 1540-7993/15/$31.00 © 2015 IEEE.

[2] The World Wide Web Consortium (W3C): OWL Web Ontolo-gy Language Overview (February 2004).

[3] Denker G, Kagal L, Finin T, Paolucci M, Sycara K, "Security for daml web services: Annotation and matchmaking", Second International Semantic Web Conference, pp. 335-350, 2003.

[4] S.D. Warren and L.D. Brandeis, "The Right to Privacy," Har-vard Law Review, vol. 4, no. 5, 1890, pp. 193-220.

[5] Rekha Bhatia and Manpreet Singh Privacy Issues in Web Ser-vices: An Ontology based Solution .

[6] International Conference on Research Advances in Integrated Navigation Systems (RAINS - 2016), April 06-07, 2016,978-1-4673-8819-8/16/$31.00 ©2016 IEEE PERSONALIZATION AND PRIVACY IN PROFILE-BASED WEB SEARCH .

[7] The 5th International Conference on Electrical Engineering and Informatics 2015 August A New String Matching Algorithm Based on Logical Indexing

[8] R.Bhatia, and M. Singh. Trust Based Privacy Preserving Ac-cess Control in Web Services Paradigm. In: the Second IEEE International Conference on Advanced Computing, Networking and Security, ADCONS (2013), 243-246

[9] C. Farkas, S. Jajodia, "The Inference Problem : A Survey". SIGKDD Explorations, Vol. 4, Issue 2, pp 6-11.

[10] M. Li, H. Wang and D. Ross. Trust-based Access Control for Privacy Protection in Collaborative Environment, the 2009 IEEE International Conference on eBusiness Engineering, Macau, China, (ICEBE 2009), 425-430, 2009.

[11] S.D. Warren and L.D. Brandeis, "The Right to Privacy," Har-vard Law Review, vol. 4, no. 5, 1890, pp. 193-220.

[12] A.F. Westin, The Right to Privacy, Atheneum, 1967.

[13] A.Novotny and S. Spiekermann. Personal information mar-kets and privacy: a new model to solve the controversy, pp 102-120.IOSPress, 2013.

[14] R.Bhatia and M.Singh, "A Novel Trust-Based Privacy Pre-serving Access Control Framework in Web Services Paradigm," Springer Book Series Advances in Intelligent Systems and Computing, Volume 308, pp 441-453, 2014.

[15] Agrawal, R., Kiernan, J., Srikant, R. and Xu, Y. 2002, "Hippocratic databases". In Proceedings of Very LargeData Base, Hong Kong, China, pp 143-154.

[16] Rotenberg, M., The Privacy Law Sourcebook 2000, United States Law, International Law, and Recent Developments, Electron-ic Privacy Information Center (2000)

[17] R.Bhatia and M.Singh, "Preserving Privacy In Health Care Web Services Paradigm Through Hippocratic Databases," Springer Book Series Advances in Intelligent Systems and Computing, Volume 308, pp 177-188, 2014.