

## Chapter 2

# A Survey on Performance Evaluation Mechanisms for Privacy–Aware Web Search Schemes

**Rafi Ullah Khan**

*The University of Agriculture, Peshawar, Pakistan*

**Mohib Ullah**

*The University of Agriculture, Peshawar, Pakistan*

**Bushra Shafi**

*The University of Agriculture, Peshawar, Pakistan*

**Imran Ihsan**

*Air University, Islamabad, Pakistan*

### ABSTRACT

*Due to the exponential growth of information on the internet, web search engines (WSEs) have become indispensable for effectively retrieving information. Web search engines store the users' profiles to provide the most relevant results. However, the user profiles may contain sensitive information, including the user's age, gender, health condition, personal interests, religious or political affiliation, and others. However, this raises serious concerns for the user's privacy since a user's identity may get exposed and misused by third parties. Researchers have proposed several techniques to address the issue of privacy infringement while using WSE, such as anonymizing networks, profile obfuscation, and private information retrieval (PIR) protocols. In this chapter, the authors give a brief survey of the privacy attacks and evaluation models used to evaluate the performance of private web search techniques.*

DOI: 10.4018/978-1-6684-6914-9.ch002

## INTRODUCTION

Online user privacy is a delicate issue and yet has been considered one of the significant and basic needs of the user by web service provider companies (Khan, 2020). Keeping records of users' queries and other activities on the Web by the service providers is not considered a misdeed. Many web service providers use this information for financial gains, such as targeted advertisement, market research and others (Khan, Islam, Ullah, Aleem, & Iqbal, 2019; Kulizhskaya, 2017; Petit, 2017). These companies claim that they use user behaviour tracking to improve product quality. However, collecting users' behavioural information and activity tracking without their knowledge is an indecent act and a direct violation of their fundamental rights (Raza, Han, & Hwang, 2020). Web Search Engine (WSE) is one of the prevalent web services used all over the globe. In the case of web search services, the issue of user privacy is more severe as it can unveil users' personal information (Khan, Ullah, Khan, Uddin, & Al-Yahya, 2021). Web search queries may contain users' queries related to their health condition, financial status, political affiliation, gender orientation, and others. These queries may unveil hidden implicit information about the users (Khan, 2020; Khan & Islam, 2017; Khan et al., 2019; Khan et al., 2021). In 2006, a privacy infringement case happened when America Online (AOL) (Adar, 2007; Barbaro, Zeller, & Hansell, 2006) published three months of queries from over 650K users, and New York Times managed to identify some of the users and their very personal information. Similarly, in the same year, Google was asked to submit the queries of their users to the United States Department of Justice (Khan, 2020). The Court of Justice of the European Union took this matter of user privacy seriously by introducing the "right to be forgotten" option and bound web service providers to its enforcement (Gibbs, 2016).

Due to the exponential growth in data and information on the internet, WSE has become essential for finding relevant information (Khan, 2020; Ullah, 2020; Ullah et al., 2022). WSE collects different information about the users and their profiles and uses these user profiles to retrieve the most relevant information according to the history and interests of the user (Stone, 2022a, 2022b). However, this user history and profile may contain explicit or implicit sensitive information about the user, which may lead to a privacy breach (Earp, Antón, Aiman-Smith, & Stufflebeam, 2005; Rostow, 2017). Furthermore, web service providers use this information for their other services (location, shopping, media, and others), due to which the target user receives recommendations from other services, which may lead to even more disastrous consequences (Citron & Solove, 2022).

Many new web search service providers took advantage of this situation and launched their WSE with privacy-preserving features. These web search service providers claim that they do not maintain users' profiles or any other activity tracking. However, the terms and conditions of these web search service providers affirm that they collect user queries only for result ranking purposes (Khan, 2020), and yet it may breach users' privacy, as happened in the case of AOL (Adar, 2007; Barbaro et al., 2006). Examples of these WSEs are Start-Page<sup>1</sup>, Qwant<sup>2</sup>, DuckDuckGo<sup>3</sup> and others.

Instead of server-side privacy preservation solutions, the researchers have proposed a client-side privacy preservation mechanism. The client-side privacy preservation mechanism does not rely upon service providers' privacy features; instead, they rely on their mechanism. These privacy preservation mechanisms can be classified into four major classes: Unlinkability solutions, Indistinguishability solutions, Hybrid Solutions and Private Information Retrieval Protocols (Khan, 2020). The details and examples of solutions for each class are given in the next section.

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/a-survey-on-performance-evaluation-mechanisms-for-privacy-aware-web-search-schemes/322584](http://www.igi-global.com/chapter/a-survey-on-performance-evaluation-mechanisms-for-privacy-aware-web-search-schemes/322584)

## Related Content

---

### E-Cocreation of Knowledge through Informal Communications

Kazushi Nishimoto (2011). *E-Activity and Intelligent Web Construction: Effects of Social Design* (pp. 135-153).

[www.irma-international.org/chapter/cocreation-knowledge-through-informal-communications/53280](http://www.irma-international.org/chapter/cocreation-knowledge-through-informal-communications/53280)

### A Survey on Performance Evaluation Mechanisms for Privacy-Aware Web Search Schemes

Rafi Ullah Khan, Mohib Ullah, Bushra Shafiand Imran Ihsan (2023). *Protecting User Privacy in Web Search Utilization* (pp. 26-45).

[www.irma-international.org/chapter/a-survey-on-performance-evaluation-mechanisms-for-privacy-aware-web-search-schemes/322584](http://www.irma-international.org/chapter/a-survey-on-performance-evaluation-mechanisms-for-privacy-aware-web-search-schemes/322584)

### Time-Dependent QoS Aware Best Service Combination Selection

Ikbel Guidara, Nawal Guermouche, Tarak Chaari, Mohamed Jmaieland Said Tazi (2015). *International Journal of Web Services Research* (pp. 1-25).

[www.irma-international.org/article/time-dependent-qos-aware-best-service-combination-selection/126288](http://www.irma-international.org/article/time-dependent-qos-aware-best-service-combination-selection/126288)

### A Specification Framework for Big Data Initiatives

Anh D. Ta, Marcus Tanqueand Montressa Washington (2019). *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 639-656).

[www.irma-international.org/chapter/a-specification-framework-for-big-data-initiatives/217854](http://www.irma-international.org/chapter/a-specification-framework-for-big-data-initiatives/217854)

### Flexible Probabilistic QoS Management of Orchestrations

Sidney Rosario, Albert Benvenisteand Claude Jard (2010). *International Journal of Web Services Research* (pp. 21-42).

[www.irma-international.org/article/flexible-probabilistic-qos-management-orchestrations/42108](http://www.irma-international.org/article/flexible-probabilistic-qos-management-orchestrations/42108)