# Web Security and Vulnerability: A Literature Review

**H Yulianton[1,2*], H L H S Warnars[1], B Soewito[1], F L Gaol[1] and E Abdurachman[1]**

[1]Computer Science Department, BINUS Graduate Program - Doctor of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480
[2]Faculty of Information Technology, Universitas Stikubank, Semarang, Indonesia 50243

[*] heribertus.yulianton@binus.ac.id

**Abstract**. The web continues to grow and attacks against the web continue to increase. This paper focuses on the literature review on scanning web vulnerabilities and solutions to mitigate web attacks. Vulnerability scanning methods will be reviewed as well as frameworks for improving web security. This research is the basis for future work that will end with the elaboration of web scanning and security with the aim of proposing better innovations.

## 1. Introduction

Until now the web continues to grow both in terms of the number of users and in terms of the technology used. Along with that there has also been an increase in attacks on the web [1]. The attacks on the web cannot be separated from the vulnerabilities that exist on the web. Therefor we need to look for any vulnerabilities that exist on the web.

As a first step, we can use two open source tools to scan for web vulnerabilities, OWASP WAP and RIPS [2]. In addition to these two tools, several other methods have been found to improve the accuracy of the scanning results, such as using static and/or dynamic code analysis [3-4], machine learning [3,5-6], data mining [4], and some penetration testing tools [7-8].

In this paper we review the article about scanning for vulnerabilities found on the web and solutions to eliminate them. We use the SLR method according to Kitchenham and Charters [9] to review articles about web security and vulnerabilities.

We collect articles related to web security and vulnerability to answer the following research questions:

RQ1: What methods are used to scan for web vulnerabilities?

RQ2: What areas of research have been carried out in web security and vulnerabilities?

Furthermore, this paper will be divided into five sections, namely: methods, results, discussion, conclusions and further research.

## 2. Methods

The first stage of the literature review is finding literature that is relevant to the topic discussed. Literature search sources come from:

- IEEEXplore Digital Library (ieeexplore.ieee.org)
- ACM Digital Library (dl.acm.org)

In accordance with the research questions that have been defined in the introduction, keywords are chosen to be used to search for articles. The chosen keywords are: "web AND security AND vulnerability". The search process is carried out in October 2019. Articles obtained in this stage are referred to as "studies found"

Furthermore, screening is done with the criteria for publication date only after 2015 and only journal articles. Articles obtained in this stage are referred to as "selected studies". The results of the two stages can be seen in Table 1.

**Table 1**. Steps for selecting articles.

| Sources | Studies Found | Selected Studies |
|---|---|---|
| IEEEXplore Digital Library | 1338 | 31 |
| ACM Digital Library | 647 | 10 |
| Total | 1985 | 41 |

## 3. Results

### *3.1. Publishing outlets*

Articles included in selected studies were published in 24 different journals. The five journals with the most articles are: IEEE Access, ACM Transactions on Privacy and Security, IEEE Transactions on Information Forensics and Security, Computer Journal and IEEE Transactions on Reliability. The names of other journals can be seen in Table 2.

**Table 2**. Source of publications.

| Journal name | # | % |
|---|---|---|
| IEEE Access | 5 | 12.20% |
| ACM Transactions on Privacy and Security | 4 | 9.76% |
| IEEE Transactions on Information Forensics and Security | 4 | 9.76% |
| Computer Journal | 3 | 7.32% |
| IEEE Transactions on Reliability | 3 | 7.32% |
| IEEE Internet of Things Journal | 2 | 4.88% |
| IEEE Transactions on Dependable and Secure Computing | 2 | 4.88% |
| IEEE Transactions on Parallel and Distributed Systems | 2 | 4.88% |
| ACM Transactions on Embedded Computing Systems | 1 | 2.44% |
| ACM Transactions on Intelligent Systems and Technology | 1 | 2.44% |
| ACM Transactions on Internet Technology | 1 | 2.44% |
| ACM Transactions on Programming Languages and Systems | 1 | 2.44% |
| ACM Transactions on the Web | 1 | 2.44% |
| IEEE Communications Surveys and Tutorials | 1 | 2.44% |
| IEEE Latin America Transactions | 1 | 2.44% |
| IEEE Transactions on Learning Technologies | 1 | 2.44% |
| IEEE Transactions on Network and Service Management | 1 | 2.44% |
| IEEE Transactions on Services Computing | 1 | 2.44% |
| IEEE Transactions on Software Engineering | 1 | 2.44% |
| IEEE Transactions on Vehicular Technology | 1 | 2.44% |
| IEEE/ACM Transactions on Networking | 1 | 2.44% |
| IET Communications | 1 | 2.44% |
| Proceedings of the ACM on Programming Languages | 1 | 2.44% |
| Tsinghua Science and Technology | 1 | 2.44% |

### *3.2. Publication trends*

The frequency of articles published from year to year since 2015 can be seen in Table 3. 2016 was the year with the highest number of publications of 10 articles and 2017 was the year with the smallest number of publications of 6 articles.

**Table 3**. Frequency of publications

| Year | # | % |
|---|---|---|
| 2015 | 7 | 17.07% |
| 2016 | 10 | 24.39% |
| 2017 | 6 | 14.63% |
| 2018 | 9 | 21.95% |
| 2019 | 9 | 21.95% |

## 4. Discussion

Not all articles in the "selected studies" category talk about scanning web vulnerabilities. In fact, only six articles [3-8] discuss the scanning of web vulnerabilities. Some articles discuss how to defend against various attacks, such as phishing attacks [10], XSS attacks [5,11-13], node scanning [14] and multistep attacks [15].

From the object studied it turns out that not all articles use the web as an object. Exactly there are nine articles [3-5,8,11-12,16-18] that use web applications as objects and two articles [7,19] use web services as objects. Other research objects include: mobile platform [10,20-22], operating system [23-26], web user [27-28], tor network [29-30], IoT [31-32], security system [6,25,33-34], cloud data centers [35], location-based services [36], compilers [37], java application [38], DNS traffic [39], DBMS [40], 3D printers [41] and M-learning environment [42]. Table 4 summarizes research objects of this study.

**Table 4**. Research objects

| Research object | # | % |
|---|---|---|
| Web application | 9 | 21.95% |
| Web service | 2 | 4.88% |
| Mobile platform | 4 | 9.76% |
| Operating system | 4 | 9.75% |
| Web user | 2 | 4.88% |
| Tor network | 2 | 4.88% |
| IoT | 2 | 4.88% |
| Security system | 4 | 9.76% |
| Cloud data center | 1 | 2.44% |
| Location-based service | 1 | 2.44% |
| compiler | 1 | 2.44% |
| Java application | 1 | 2.44% |
| DNS traffic | 1 | 2.44% |
| DBMS | 1 | 2.44% |
| 3D printer | 1 | 2.44% |
| M-learning environment | 1 | 2.44% |
| Others | 7 | 17.07% |

Some articles also discuss the benchmarks of existing tools, such as vulnerability detection tools [19], web crawler [43] and static analysis tools [44].

## 5. Conclusion

Although there is an increase in the number and types of web attacks, research on web security is still low. The methods used to scan for web vulnerabilities include: static or dynamic analysis, data mining, machine learning and penetration testing. Although many studies use web applications as objects, there are also many other interesting objects.

## 6. Further Research

For our next research we will combine static and dynamic analysis methods, machine learning and penetration testing to scan for network vulnerabilities. We hope that by combining this method can get better results compared to the state-of-the-art method.

## References

[1] Symantech Corporation. 2019, *2019 Internet Security Threat Report*. (Mountain View: Symantech Corporation vol 24)
[2] Tyagi S and Kumar K. 2018, *2018 Fifth International Conference on Parallel, and Grid Computing (PDGC)*
[3] Shar, L. K., Briand, L. C. and Tan, H. B. K., 2015. *IEEE Transactions on Dependable and Secure Computing* **12** 688-707
[4] Medeiros, I., Neves, N. and Correia, M., 2016. *IEEE Transactions on Reliability* **65** 54-69
[5] Mokbal, F. M. M. et al., 2019. *IEEE Access* **7** 100567-80
[6] Dong, Z., Kane, K. and Camp, L. J., 2016. *ACM Transactions on Privacy and Security* **19** 1-31
[7] Salas, M. I. P. and Martins, E., 2015. *IEEE Latin America Transactions* **13** 707-12
[8] Liu, M and Wang, B., 2018. *IEEE Access* **6** 70983-8
[9] Kitchenham, B. et al. 2009. *Information and Software Technology* **51** 7-15
[10] Wu, L., Du, X. and Wu, J., 2016. *IEEE Transactions on Vehicular Technology* **65** 6678-91
[11] Mitropoulos, D., Louridas, P., Polychronakis, M. and Keromytis, A. D., 2019. *IEEE Transactions on Dependable and Secure Computing* **16** 188-203
[12] Das, D., Sharma, U. and Bhattacharyya, D. K., 2015. *Computer Journal* **58** 802-22
[13] Mitropoulos, D., Stroggylos, K., Spinellis, D. and Keromytis, A. D., 2016. *ACM Transactions on Privacy and Security* **19** 1-31
[14] Jajodia, S. et al., 2017. *IEEE Transactions on Information Forensics and Security* **12** 2532-44
[15] Zonouz, S. A. et al., 2015. *IEEE Transactions on Parallel and Distributed Systems* **26** 562-73
[16] Rezvani, M., Ignjatovic, A. and Bertino, E., 2018. *ACM Transactions on Internet Technology* **18** 55:-1-12
[17] Calzavara, S., Rabitti, A. and Bugliesi, M., 2018. *ACM Transactions on the Web* **12** 1-36
[18] Jan, S., Panichella, A., Arcuri, A. and Briand, L., 2019. *IEEE Transactions on Software Engineering* **45** 335-62
[19] Antunes, N. and Vieira, M., 2015. *IEEE Transactions on Services Computing* **8** 269-83
[20] Scheir, M. et al., 2015. *ACM Transactions on Embedded Computing Systems* **14** 85:1-25
[21] Conti, M., Mancini, L. V., Spolaor, R. and Verde, N. V., 2016. *IEEE Transactions on Information Forensics and Security* **11** 114-25
[22] Tao, D., Lin, Z. and Lu, C., 2015. *Tsinghua Science and Technology* **20** 537-44
[23] Pomonis, M. et al., 2018. *ACM Transactions on Privacy and Security* **22** 5:1-28
[24] Cao, Y. et al., 2018. *IEEE/ACM Transactions on Networking* **26** 765-78
[25] Min, B. and Varadharajan, V., 2016. *The Computer Journal* **59** 1735-48
[26] Kumar, P. et al., 2018. *IEEE Transactions on Network and Service Management* **15** 1545-59
[27] Heartfield, R., Loukas, G. and Gan, D., 2016. *IEEE Access* **4** 6910-28
[28] Neria, M. B., Yacovzada, N. S. and Ben-Gal, I., 2017. *ACM Transactions on Intelligent Systems and Technology* **8** 1-21
[29] Imani, M., Amirabadi, M. & Wright, M., 2019. *IET Communications* **13** 2723-34

[30] Tan, Q. et al., 2019. *IEEE Internet of Things Journal* **6** 1584-93
[31] Shwartz, O. et al., 2018. *IEEE Internet of Things Journal* **5** 4965-76
[32] Choi, C. & Choi, J., 2019. *IEEE Access* **7** 110510-7
[33] Yao, S. et al., 2019. *IEEE Access* **7** 6117-28
[34] Han, W., Li, Z., Yuan, L. and Xu, W., 2016. *IEEE Transactions on Information Forensics and Security* 11 258-72
[35] Cui, L. et al., 2017. *IEEE Transactions on Parallel and Distributed Systems* **28** 1163-75
[36] Argyros, G. et al., 2017. *ACM Transactions on Privacy and Security* **19** 12:1-31
[37] Donaldson, A. F., Evrard, H., Lascu, A. & Thomson, P., 2017. *Proceedings of the ACM on Programming Languages* **1** 1-29
[38] Spoto, F. et al., 2019. *ACM Transactions on Programming Languages and Systems* **41** 1-58
[39] Torabi, S., Boukhtouta, A., Assi, C. & Debbabi, M., 2018. *IEEE Communications Surveys & Tutorials* **20** 3389-415
[40] Medeiros, I., Beatriz, M., Neves, N. & Correia, M., 2019. *IEEE Transactions on Reliability* **68** 1168-88
[41] Do, Q., Martini, B. & Choo, K.-K. R., 2016. *IEEE Transactions on Information Forensics and Security* **11** 2147-86
[42] Kaiiali, M. et al., 2016. *IEEE Transactions on Learning Technologies* **9** 258-71
[43] Tatli, E. I. and Urgun, B., 2016. *The Computer Journal* **60** 555-72
[44] Nunes, P. et al., 2018. . *IEEE Transactions on Reliability* **67** 1159-75