Tamper Detection Based on Webpage Change Features

Hanji Shen [ab], Jun Li [a], Wei Wan [a], Chun Long [a], Jingdeng Zhou [a], Yuhao Fu [a] , Xiaofan Song [a]

a Computer Network Information Center, Chinese Academy of Sciences, China

b University of Chinese Academy of Sciences, China

shenghanji@cnic.cn, jlee@cstnet.cn, wanwei@cnic.cn, longchun@cnic.cn, zhoujingdeng@cnic.cn, fuyuhao@cnic.cn, songxiaofan@cnic.cn

***Abstract*—With the rapid development of the Network, the webpage tampering has become a problem that can not be ignored. In October 2019, for example, the number of websites tampered with in China reached more than 20,000, with the largest share coming from Beijing, Shandong and Guangdong respectively [1]. This shows that webpage tampering is no longer a mere simple problem, but requires greater attention. In order to find a better way to detect web page tampering, the website tamper-proof technology receives much attention in the area of web security. This paper proposes a method of webpage tampering detection based on the webpage change features through analyzing the features of webpage changes and the illegal tampering purpose. Webpage changes will be determined before detecting. The detection model is decided by webpage change time, webpage change code amplitude, webpage change frequency, webpage change content location and webpage change content relationship. To be more specific, the detection process includes two training and detection stages, training phase and detection phase. In the training phase, the effective detection evidence only suitable for the webpage can be obtained through using the data set (multiple changes data in the webpage) to train model. In the detection phase, the detection model will detect the particular webpage according to the detection evidence, and then gives the detection result. If the result is misdeclaration, the detection evidence will be retrained. Furthermore, the simulation tampering experiment is designed to verify the feasibility of the detection method. And the optimal number of the experimental webpage changes firstly into the training phase is determined according to the accuracy and recall rate of the test results. After Verifying, the accuracy and recall of the test results were 98.32% and 99.12% respectively. The best number of changes was 55 and the risk value was 1.**

***Keyword*—Webpage detection model, webpage tamper detection, web features**

Hanji Shen is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. He received M.S. degree in Engineering in the field of Computer Technology from University of Chinese Academy of Sciences, and he is currently studying for a Ph.D. in University of Chinese Academy of Sciences, majoring in Computer Software and Theory. He has been an IEEE member since 2014. His research interest is Cyber Security.

Jun Li is currently a researcher, deputy chief engineer and doctoral supervisor in Computer Network Information Center, Chinese Academy of Sciences, China. He received Ph.D. degree in Computer System Structure from University of Chinese Academy of Sciences. He is one of the earliest experts engaged in computer network technology research in China. He has been engaged in scientific research and engineering practice in the field of computer network for a long time. His research interests include Network Architecture and Cyber Security.

Wei Wan is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. He received Ph.D. degree in Computer Software and Theory from University of Chinese Academy of Sciences. He has presided over and participated in many national projects. Moreover, he developed Investigation of state division in botnet detection model, Botnet detecting method based on activity similarity and so on. His research interests include Information Security, Cyber Security, Cyber Risks and Web Vulnerabilities.

Chun Long is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. He received Ph.D. degree in Computer Software and Theory from University of Chinese Academy of Sciences. He developed Compound Attack Prediction Method based on the Attack Graph and Multi-source Security Event Fusion method based on EA-DS Evidence Theory. His research interests include Information Security, Cyber Security, Cyber Risks and Web vulnerabilities.

Jingdeng Zhou is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. He received B.S. degree in Network Engineering from Jinggangshan University. His research interests include Network Security Guarantee Technology and Cyber Space Security.

Yuhao Fu is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. He received B.S. degree in Computer Science and Technology from Chongqing University of Post and Telecommunications. His research interests include Network Security Guarantee Technology and Cyber Security.

Xiaofan Song is currently a full engineer in Computer Network Information Center, Chinese Academy of Sciences, China. She received M.S. degree in Cyber Security from University of Southampton. Her research interests include Information Security, Cyber Security, and Cyber Risks.