UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES AND ACADEMIC ADVANCEMENT

A Comparative Study between Web Log Expert , alter Wind analyzer and Nihuo Weblog Analyzer to Analyze Abnormal Usage Behavior

By

Mohamed Ahmed Fadl Allah Ebraheem

A Thesis

Submitted to the College of Graduate Studies and Academic Advancements

In Partial Fulfillment of the Requirement for the Degree of Master of Science in Information Technology

Supervisor

Dr. Hassan Fadlallah

April 2017

Abstract

Web usage mining tools can be applied to analyze abnormal usage on web sites. most of web usage mining tools are very costly. Selection of wrong tool is expensive both in terms of money and loss of time. Because there is a lack of a review and comparison among these tools in analyzing abnormal usage behavior so we compared Web Log Expert, Alter Wind Log analyzer Professional and Nihuo Web Log Analyzer to measure their ability of analyzing abnormal usage.

We are using analytic experimental method by conducted two experiments first one scans vulnerabilities and brute force attack to the targeted website. Second experiment DDOS attack to the targeted website in order to measure abnormal usage. Our methodology based on built a model that contains comparative processes that has three phases. The first phase used the log before abnormal usage with the selected tools and keeps the results of this phase. The second phase used the log after abnormal usage with the selected tools and keeps the results of this phase. The third phases do the process of comparison based on the results of the previous phases.

In our result we found these tools have the applicability for analyzing abnormal usage especially that based in scan vulnerability and brute force attack. Our comparison results was done on the two sides .First side is the changes that happened in the number values of the results to these characteristic side is that the characteristics that do not contain numbers. We compared these characteristics depending on the degree of details provided by these characteristics. Second. Side is that the characteristics that do not contain numbers we compared these characteristics depending on the degree of details provided by these characteristics. We have found that the Nihou log analyzer is best than others selected tools to analyze abnormal behavior accuracy in our experiments.

المستخلص

أدوات تعدين استخدام الويب يمكن تطبيقها لتحليل الاستخدام غير العادي علي مواقع الويب .معظم أدوات تعدين استخدام الويب مكلفة جدا. اختيار أداة خاطئة مكلفة سواء من حيث المال وفقدان الوقت ولأن هناك عدم وجود استعراض والمقارنة بين هذه الأدوات في تحليل سلوك الاستخدام غير العادي لذلك قمنا بمقارنة Web Log Expert ، Alter Wind Log analyzer Professional و Nihuo Web Log Analyzer لقياس قدرتها على تحليل الاستخدام غير العادي. لقد إستخدمنا طريقة التحليل التجريبي من خلال إجراء تجربتين الأولي قمنا فيها بفحص نقاط الضعف مع هجوم ال brute force على الموقع المستهدف. التجربة الثانية هجوم منع الخدمه المتعدد علي نفس الموقع من أجل قياس الاستخدام غير العادي. منهجيتنا قامت على أساس بناء نموذج يحتوي على عمليات مقارنة تتكون من ثلاث مراحل. في المرحلة الأولى هناك السجل قبل الاستخدام الغير العادي مع الأدوات المحددة والإحتفاظ بنتائج هذه المرحلة. المرحلة الثانية السجل بعد الاستخدام غير العادي مع الأدوات المحددة والإحتفاظ بنتائج هذه المرحلة. أما المراحل الثالثة فتتم فيها عملية المقارنة استنادا إلى نتائج المراحل السابقة.

كنتيجة لهذه التجارب وجدنا أن هذه الأدوات لديها القدره لتحليل الإستخدام غير العادي خصوصا في التجربه الأولي .نتائج المقارنه تمت في جانبين .الجانب الاول يعتمد علي التغيرات التي حدثت علي قيم هذه الخصائص .الجانب الثاني اعتمد علي التفاصيل ودرجه التنوع المقدمه في هذه الخصائص .ولقد وجدنا أن الأداة Nihuo Web Log Analyzer أفضل من غيرها من الأدوات المختارة لتحليل دقه السلوك الغير عادي في هذه التجارب التي قمنا بها .

1.1 Introduction

Web mining is the application of data mining techniques to extract knowledge from Web data including Web documents, text on web, images on web, usage logs of web sites etc. Web Mining can be categorized into three broad areas of mining. Web Content Mining (WCM) is mining of information from the contents of web like text, images etc. Web Structure Mining (WSM) deals with the structure of the web pages and effect of this structure on traversal through web pages. Web usage mining is a research field that focuses on the development of techniques and tools to study users web navigation Behavior. Web Usage Mining is also called web log mining. [1]

Web Usage Mining is that part of Web Mining which deals with the extraction of useful knowledge from the secondary data derived from the interactions of the users while interacting with the Web.

The Web usage data includes the data from Web server access logs, proxy server logs, browser logs, user profiles, registration data, user sessions or transactions, cookies, user queries, bookmark data, mouse clicks and scrolls, and any other data as the results of interactions. The scope of Web usage mining is local, which means that the scope of Web usage mining spans an individual Web site. [2]

In today's business environment almost all companies have their computers connected to the public Internet. As the number of companies with computers and services accessible to the Internet increases, a corresponding increase in the number of attacks against these businesses is also observed. Network based attacks on business computers have been increasing in frequency and severity over the past several years. [3]

1.2 Problem statement

Most of Web usage mining tools are very costly. Selection of wrong tool is expensive both in terms of money and loss of time and because there is a lack of a review and comparison among these tools in analyzing abnormal usage behavior. We Compare Web Log Expert, Alter Wind Log Analyzer Professional and Nihuo Web Log Analyzer to measure their ability of analyzing abnormal usage.

1.3 Research Objectives

- To reduce the risk of lost times and cost for the users to select appropriate tools that help in analyzing abnormal usage.
- To investigate the accuracy of the selected tools for analyzing abnormal usage
- To build a model for the comparison process
- To validate the proposed model

1.4 Methodology

This is an analytic experimental method. In this research we conducted two experiments first one scans vulnerabilities and brute force attack to the targeted website. Second experiment makes DDOS attack to the same targeted website in order to measure abnormal usage. The method used the following steps in two experiments, first starts by collected log before and after attack from www.htrshat.com , second we used selected tools (Log Expert, Alter Wind Log Analyzer Professional and Nihuo Web Log Analyzer) to analyze log before and after attacks. Finally we compare the results to extract common and uncommon characteristics and using this characteristics to compare selected tools.

1.5 Data set

The logs in this research have been collected from www.htrshat.com website in date from 31/1/2017 to 21/2/2017 to first Experiment and from 28/2/2017 to 2/3/2017 to the second Experiment.

1.6 Research Tools

In this research, we used Web Log Expert, Alter Wind Log Analyzer Professional and Nihuo Web Log Analyzer.

1.7 Thesis Organization

In this thesis there are five chapters. First chapter include introduction, problem statement, objective of research, data set, methodology, research tool. Second chapter includes background and related work that deals with web mining; specially web usage mining .Third chapter includes the methodology of the research .Fourth chapter includes the implantation of the proposed methodology and experiments. Finally the fifth chapter contains conclusion and recommendation for future work.