Dynamic Forensics Model based on Ontology and Context Information

Baoxian JIA[1], Weiqiang Yang[2]
[1] Liaocheng University, Liaocheng ,China 252059

[2] Liaocheng Wenxue Wen Xuan High School, Liaocheng ,China 25200

Abstract

The existing Forensics model already could not satisfy the need of the computer forensics. Key technology which can implement Forensics Model was studied in this paper. Honeypot forensics, IDS, Ontology technologies were combined effectively in order to obtain forensics data beforehand. A complete dynamic forensics system which can replay computer crime was designed. Non-repeated varying probability packet marking scheme was proposed. The paper proposed dynamic forensics model based on ontology and context information. Dynamic forensics will inevitably produce large amounts of disorganized data having many drawbacks such as non-uniform format, so the paper proposed the high-precision data mining model based on ontology. The efficiency and accuracy of spam filtering are improved.

Keywords: Computer Crime, Computer Forensics , Dynamic Forensics Model, Honey-pot, Ontology, Intrusion Detection; Context Information.

1. Introduction

With the rapid development of Internet and e-commerce, computer crime becomes more serious. The features of computer crime make it very difficult to obtain evidence. It resulted in that a large number of computer crime can not be found and not be arrested which greatly Contributed to the arrogance of computer crime. The current methods and techniques of computer forensics are relatively simple. It tends to forensic after in the crime has occurred to compensate for the loss. With the social development, computer technology has gradually increased. Forensic after the crime occurred can not meet the crime forensics requirements. We need dynamic forensic against computer crimes[1].

The paper studied the key technologies of computer forensics included honeypot, intrusion detection, context information, Ontology and so on. Integrated some of these key technologies, the paper proposed a real-time dynamic forensic model based on Ontology and context information. The model can obtain the evidence in advance. The model can take the initiative action to record the evidence of all aspects of crime in the criminal process and extract evidence

2. key technologies of dynamic forensics

2.1 Context Information

Context information is a concept in pervasive computing [2]. Abowd and Dey in 1999 proposed a general definition of context information [3]: "It is used to indicate the object context information of environmental characteristics and the object can be a person, place or the user interacts with the application associated with objects including user and the application itself."

With the participation of the context information, the traditional computing environment is converted into a perception environment. When the context information, the system access results changes. In computer forensics, the role and permissions of users will follow the corresponding changes of context information [4]. According to the current dynamic environment obtained by context information, the application of the model can forensic in real-time ways.

2.2 Ontology

Ontology is a conceptual specification. ontology was first developed in Artificial Intelligence (AI) to facilitate knowledge sharing and reuse, and had gained wide popularity in the early 1990s in several field applications, such as knowledge engineering, natural language processing, and knowledge representation. Nowadays, ontology is also a popular research topic in knowledge management, cooperative information systems, electronic commerce, information retrieval, intelligent information integration and medicine, among others[5].

The integration between data mining and ontology enables the data consistency to solve some problems such as data heterogeneous of forensic data. At the same time, the user identity under the premise of mining in the semantic rules improve the effectiveness of the excavation while the bulk evolution rules can also be easily increased and changed. In the mining process, the ontology is used to help users constitute a valid DM process (executable programs) collection[6].

2.3 Honeypot

Honeypot is a security resource whose value lies in being probed, attacked and damaged. Honeypot is a trap technology. On one hand the honeypot can effectively protect critical servers from hackers. Under normal circumstances, financial institutions or community banks will commit the honeypot trap outside the firewall. Usually, hackers focus at and attack the mirror honeypot, while the real server can be safe. At the same time, the use of honeypot can fully understands the hacker culture including hacking records, hacking techniques and attacking tools. Studying these results collected through honeypot can improve the server's defenses and reduce the losses caused by attacks to avoid the next attack. Honeypot is the platform to forensic. Usually we install the forensics system on the distributed honeynet system not on the real server. Once hacker attacks successfully, the consequences could be disastrous. Because the honeypot is a trap technology which can effectively protect critical servers from hackers.

3. Dynamic Forensics Model based on Ontology and Context Information

3.1 Tables and Figures

According to the corresponding visitor method and its licensing strategy, the model will determine the access legality of the user according through retrieving the current application environment. if it is the illegal invasion, the model will retrieve and update context information repository and record the invasion information and reason to implement operations.

In the model, with the change of context information, user roles and permissions will follow change accordingly. Compared with traditional forensic methods, the application of the model based on the current dynamic environment changes in real-time way. In addition, the entities involved, strategy, authority and rules in the model were described by the ontology language, to achieve the unity of the concept and the norm.Dynamic forensics will inevitably produce large amounts of disorganized data having many drawbacks such as non-uniform format, so the paper proposed the high-precision data analysis model based on ontology. According to the excavation results, the model can provide early-warning services. For example, based upon the results of data mining, the data analysis model can execute relevance semantic recommendation. Due to a subsidiary of the same ontological existing correlation between various

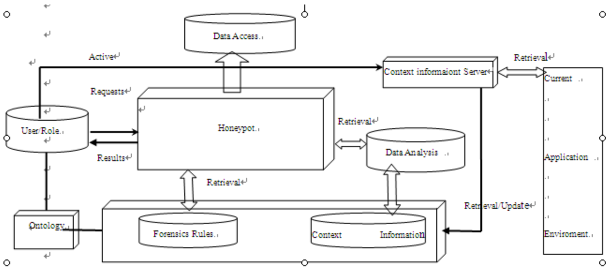

Fig. 1 Model of dynamic forensics

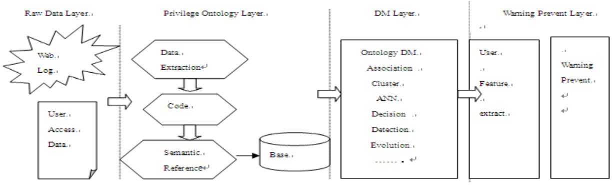

Fig. 2 overall architecture

4. Data Analysis based on Ontology

Through the above analysis of ontology construction and data mining, we initially proposed Data mining based on Ontology architecture divided into four levels as shown in fig 2: the raw data layer, the privilege ontology layer, data mining layer, warning prevent layer.

（1） The raw data layer stored the original data of computer crime .For example, it includes the daily user access log information gathered. These logs information collection and excavation is essential for warning prevent level, because through the Web site usage mining, user -access to content, residence time and frequency we can obtain the behavior and the way users access to the general knowledge to improve the Web site services design. More importantly, through the characteristics of these users, they can conduct targeted personalized service.

（2） Privilege ontology layer. This level is the proposal and optimization of the original data layer. That is, through semantic extraction, semantic tagging methods the original chaotic, non-structured or semi-structured data can be structured to the reusable, non-ambiguous understanding of ontology library, providing the foundation excavation for data mining layer [7].

（3） data mining layer. The level makes deep-level ontology excavation using various mining technology including association rules, cluster analysis, decision tree, discrete point detection, neural networks, evolution analysis. For example, mining association rules can be applied to dynamic forensic ontology mining. In this process we deal with a URL address as a project, a browsing process as a record. Then from such a database, we can find out the URL address of the association rules which is helpful for the Internet search engine and thus these rules associated with the network address of a series can be presented to the user. When a user is reading a web page, personalized service in advance the user might visit later in the page can be provided in accordance with association rules.

（4）Based upon the results of data mining, the warning prevent layers can execute relevance semantic recommendation. Due to a subsidiary of the same ontological existing correlation between various areas, when a user is interested in a particular sub-areas, such correlation makes it possible to recommend between the different areas. Due to the different manifestations to preferences property, we can identify the relevant attributes of various types of forensic resources to achieve a match between the different forensic resources and make such a recommendation By constructing a preferences feature model, we can match the different sub-areas. Therefore, through determining what ontology the area belongs to, we may infer what could be real intrusion. The current view of the sub-ontology does not belong to the same root, we think it has nothing to do with the intrusion, so there is no need to research it.

5.Conclusion

Honeypot, forensics, IDS, Ontology technologies were combined effectively in order to obtain forensics data beforehand. A complete dynamic forensics system which can replay computer crime was designed. Non-repeated varying probability packet marking scheme was proposed. The paper proposed dynamic forensics model based on ontology and context information. The efficiency and accuracy of spam filtering are improved. Computer forensics and intrusion detection having many drawbacks such as non-uniform format, so the paper proposed the high-precision data mining model based on ontology. According to the excavation results, the model can provide early-warning services. The paper specified the process of data mining based on ontology using association rules and used Bayesian network model to calculate the correlation degrees between ontologies and realize the association mining.

Acknowledgments

Shandong science research and development plan(J12LN39); Shandong Soft Sicence development plan(J12RKBO1159)

References

[1] Y.Wang,Cannady,J.Rosenbluth,J.Foundations of computer forensics:A technology for the fight against computer crime,Computer Law and Security Report,2005,21(2):119-127.

[2]Sandhu,R.S.,etal.,Role-based access control models.Computer,1996.29(2):p.38-47.

[3]Sandhu,R.,D.Ferraiolo,and R.Kuhn,NIST model for role-based access control:
Towards a unified standard. Proceedings of the ACM Workshop on Role-Based Access
Control,2000:p.47-63.

[4] Jiang Hua,Zhang Shasha.The Network Identity Authentication System Based on Iris Feature Identification[J],Modern Applied Science,2009(5):127-130.

[5] Corcho O,Lopez M F,Perez A G,etal.Methodologies,tools and languages for building ontologies,where is their meeting point[J].Data and Knowledge Engineering,2003,46：41-64.

[6] XIE Sheng-xian,JIA Bao-xian.Data Mining Based on Ontology in the CRM Decision Analysis [J].Journal of Liaocheng University: Natural Science Edition,2010,23(4):96-99

[7] Davies J，Fensel D Toward the Semantic Web: Ontology driven knowledge management[J]．John Wiley & Sons Ltd，2003：310-316

First Author. Baoxian Jia (1982-), master candidate, research direction: Ontology , semantic Web and E-hospital.

Second Author. Weiqiang Yang, research network and data mining.