Asian Journal of Information Technology 6 (5): 614-617, 2007

An Intrusion Detection Expert System with Fact-Base

Yuan Yuan and Dai Guanzhong
College of Automation, Northwestern Polytechnical University,
Xi'an Shaanxi 710072, Pepole Republic, China

Abstract: This study designs an intrusion detection expert system with fact-base(FIDES) which includes some important files and directories that are vulnerable to certain types of attack scenarios. FIDES matches and categorizes audit data with fact-base component. Inference component of FIDES adopts misuse detection techniques or anomaly detection technique for different audit data according to the result of categorization. The experiments show that FIDES could estimate the unknown user activity accurately and the False Negative Rate and the False Positive Rate have been reduced effectively.

Key words: Linux, FIDES, fact-base, misuse detection, anomaly detection, expert, system, intrustion

INTRODUCTION

The security manager must establish and maintain a security environment that ensures three requirements: The confidentiality, the integrity and the availability of information resources. But most computer systems are vulnerable to two different groups of attacks: Insider attacks and outsider attacks (Ilgun, 1993). A system could be secure to an outsider attack by access control, such as passwords, Access Control List (ACL). However, even the most secure systems are vulnerable to abuse by insiders who misuse their privileges, so it can't rely on access control mechanisms in every case to safeguard against a penetration or insider attack. Audit trails may be the only mean of detecting authorized but abusive user activity. These attacks are usually detected by tools referred to as Intrusion Detection System (IDS).

At present, intrusion detection techniques can be categorized into misuse detection and anomaly detection. Misuse detection techniques, for example (Ilgun *et al*., 1995) use patterns of well-known attacks or weak points of the system to identify intrusions. The main shortcomings of such systems are that known intrusion patterns have to be hand-coded into the system and they are unable to detect any future (unknown) intrusions that have no matched patterns stored in the system. So, for misuse detection techniques the False Negative Rate (FNR) is high.

Anomaly detection techniques, such as (Karlton and Mohammed, 2002) firstly establish normal user behavior patterns (called profiles) and then try to determine whether deviation from the established normal profiles should be flagged as intrusions. The main advantage of anomaly detection systems is that they can detect new types of unknown intrusions. In recent years, the continual emergence of new attacking methods has caused great loss to the whole society. So, the advantage of detecting future attacks has specially led to an increasing interest in anomaly detection techniques. But it will send an alarm when it detects an unknown normal activity. Consequently, for anomaly detection techniques the False Positive Rate (FPR) is high.

The Intrusion Detection Expert System (IDES), being developed at SRI's Computer Science Laboratory (CSL), is a comprehensive system that uses innovative statistical algorithms for anomaly detection, as well as an expert system that encodes known intrusion scenarios (Lunt *et al*., 1988, 1990). IDES is the result of research that started in CSL in the early 1980s and that led to the IDES prototype that is running at SRI.

In this study, we design an intrusion detection expert system with fact-base (FIDES) working on Linux. It is a host-based IDS and combines both of the 2 above detection techniques. FIDES categorizes the audit data by the fact-base component firstly and then uses relevant detection technique for different audit data. The experiments show that it is very effective to estimate the unknown user activity.

FIDES STRUCTURE

FIDES is designed to support only a single Linux host. Data engine collects the user behaviors and translates them into the defined format of audit data. Inference engine ascertains whether the observed activity is abnormal by using misuse detection techniques or anomaly detection techniques. Decision engine determines which of the alerts should be filtered and send

Corresponding Author: Yuan Yuan, College of Automation, Northwestern Polytechnical University, Xi'an Shaanxi 710072, Pepole Republic China

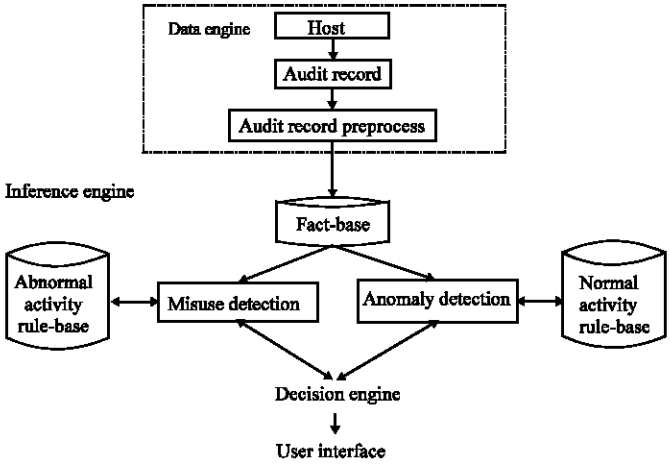

Fig. 1: FIDES structure

Table 1: Format of audit data

Item	Description
Id	Sequence number of event
Time	Time of event generated
User	User who generated the event
Pid	Pid of event
Process	Name of event
File_list [MAX]	File list used by event
Description	Description about event

the final result to user interface. User interface shows the security manager the detailed report about alert and the method to process it. Figure 1 shows the structure of FIDES.

Data engine: The most important data for FIDES is the audit data. Linux audit data can be collected from several sources, such as operating system, network traffic flowing, application logs and so on. Each one has its advantages and disadvantages with respect to the types of intrusion, the complexity and volume of the data and the capability of appealing to an intuitive understanding of what is h appening when an anomaly is detected (Lunt *et al*., 1990).

The preferred audit data for host-based IDS is operating system audit record. The reasons are:

The architecture of operating system audit record and the protection for them have been considered when audit system was designed, so the security of audit record could be ensured.

The audit record provides the events happened in kernel level, which reflect the status of system and provide some detailed information. This makes IDS identify each user's activity.

So, in FIDES we use operating system audit record as the audit source. But the audit record has to be preprocessed to the appointed format that is required by this system. FIDES audit data format is defined in Table 1.

Inference engine: Inference engine is the logical component through which an expert system evaluates an audit data (called fact in IDES) against the production rules. It retrieves audit data from the data engine and compares them with user's rules. If this behavior matches some rule, a report will be generated and sent to decision engine.

Rule-base: In this context, a fact is a statement that is asserted into the system and whose validity is accepted. Facts are often implemented as attributes and values that represent the state of the environment. A rule is an inference formula of the form P1,...Pn infer Q. Inference formulae can be alternatively expressed as production rules, such as IF...THEN.... Production rules are the basic elements through which an expert system could interpret and discover meaning from environmental signals that it receives, as in:

If user TOM changes to user root then it is a privilege elevation.

A production rule consists of two parts, the antecedent (left-hand side) and the consequent (right-hand side). When the conditions in the antecedent are satisfied, the rule is activated.

In FIDES there are two rule-bases which describe the normal and abnormal user activity, respectively. New rules may be compiled, installed and then added to the real-time analysis processing. Rules already available to FIDES may be turned ON or OFF dynamically during runtime.

Misuse detection component: The FIDES misuse detection component uses abnormal activity rule-base that characterizes known intrusion types to raise an alarm if observed activity matches any of its encoded rules. This type of analysis is intended to detect attempts to exploit known security vulnerabilities of the monitored

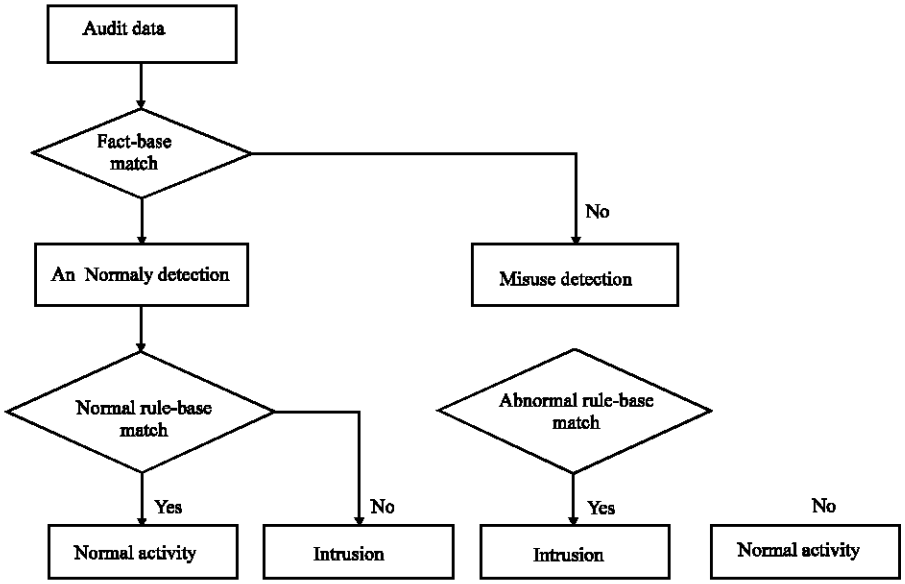

Fig. 2: Work flow of inference engine

systems and intruders who exhibit specific patterns of behavior that are known to be suspicious or in violation of site security policy. Observed activity that matches any of these predefined behaviors is flagged.

Anomaly detection component: The FIDES anomaly detection component maintains historical statistical profiles called normal activity rule-base for each user and raises an alarm when observed activity departs from established patterns of use for an individual. The historical profiles are updated regularly, so that FIDES adaptively learns each user's behavior. This component is intended to detect intruders masquerading as legitimate users. Anomaly detection component may also detect intruders who exploit previously unknown vulnerabilities and who can not be detected by any other means. Anomaly detection component can turn up interesting and unusual events that could lead to security-relevant discoveries upon investigation by a security manager.

Fact-base: It is obviously that it could be detected correctly when we match an unknown normal activity using abnormal activity rule-base and vice versa. Now, the problem is which rule-base should be using to match the unknown activity.

It is discovered that 90% of intrusion activity would access some important files or directories. So, we put forward the conception of fact-base. Fact-base consists of groups of files or directories that share some characteristics that are vulnerable to certain types of attack scenarios, such as /etc/rc.local or the ELF files with s sign. We performed some initial experiments to get this list of such files and directories.

For each audit data we match its file_list with fact-base firstly. If it matches some item in fact-base, it is possible an intrusion. Then we detect it with the anomaly detection component and use normal activity rule-base matching it. If the audit data doesn't match any item in fact-base, we match it with the misuse detection component and use abnormal activity rule-base matching it. The work flow is shown in Fig. 2. By this mechanism, FIDES could verdict the unknown normal or abnormal user behavior correctly. With the expanding of fact-base, normal activity rule-base and abnormal activity rule-base, the result of detection will be accurate more and more and the FNR and FPR will be reduced effectively.

Decision engine: The FIDES decision engine component filters the alarms generated by the inference engine component before reporting them to the security manager. Typically, 100 of audit records can be generated by a single user action. An unusual action could result in hundreds of alarms in rapid sequence. To avoid flooding the security manager with redundant alarms, the decision engine should filter some alarms to remove such redundancies. The result filter configuration option has been provided in FIDES. For example, the security manager can turn off alert reporting for specific users, if it is known that they will be doing something unusual and would generate a lot of false alarms. Although filtered alerts are not reported, they are still logged.

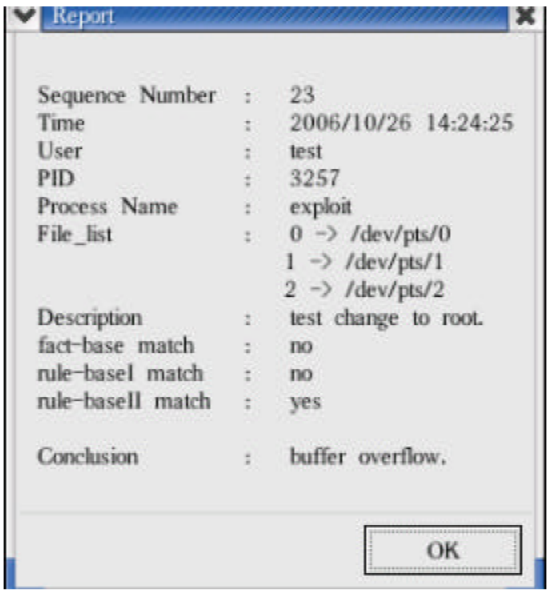

Fig. 3: An anomaly report window

User interface: One of important part of FIDES is user interface which provides the security manager with a powerful and comprehensive view of the target host being monitored. FIDES user interface is written using the GTK Library and operated under the X-Window system.

The user interface provides time-varying graphical displays of target system activity, as well as the ability to zoom on abnormal behavior by selecting from among several built-in database queries. When the system shows that abnormal activity is occurring, the system manager can quickly and easily determine which user is generating this anomaly.

Each instance is presented in its own window. Figure 3 shows an example of a window that the security manager could observe on the screen.

CONCLUSION

FIDES is a host-based intrusion detection expert system for Linux. It is capable of detecting anomalous behavior, especially, the access sequence with safety risk to file system or the alteration to system configure file. FIDES could generate detailed analysis report and send it to the user interface. By the design and development of fact-base, FIDES can estimate the unknown activity accurately and reduce the FNR and FPR evidently.

Future work: FIDES supports only Linux at present. We plan to improve the format of audit data to support other operating systems, such as Solaris, HP-UX etc. In addition, the veracity of detection result has great relation to fact-base and rule-base. The future research is expanding them through the learn mechanism of expert system to reduce the FNR and FPR more.

The following issue is the security of FIDES. The audit data must be protected from modifying. The rule-base must be protected from undesired reading and unauthorized modification. FIDES must be protected from malicious denials of service. These would make FIDES more complicatedly but more effectively.

REFERENCES

Ilgun K., 1993. Ustat. A Real-time Intrusion Detection System for UNIX. Proc. IEEE Symp. On Research in Security and Privacy Oakland. CA., pp: 16-28

Illgun., K., R. Kemmerer and A. Philips, 1995. State Transition Analysis: A Rule-based Intrusion Detection Approach. IEEE. Trans. Software Eng., 2: 181-199.

Karlton, S. and Z. Mohammed, 2002. ADMIT: Anomaly-based Data Mining for Intrusions. In Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM. Press, Edmonton Alberta Canada, pp: 386-395.

Lunt, T. F., R. Jagannthan and M. Park, 1988. A Prototype Real-time Intrusion-detection Expert System. IEEE Symposium on Security and Privacy[C]. Oakland: IEEE. Comput. Soc., pp: 59-65.

Lunt, T.F., A. Tamaru and F. Gilham *et al*., 1990. IDES: A Progress report.Proceedings of Annual Computer Security Applications Conference[C].IEEE. Comput. Soc. Press, pp: 273-285.