International Journal of Cyber-Security and Digital Forensics (IJCSDF) 5(4): 197-207
The Society of Digital Information and Wireless Communications (SDIWC), 2016 ISSN: 2305-0012

Intrusion Detection System with Spectrum Quantification Analysis

Yusuke Tsuge and Hidema Tanaka
National Defense Academy of Japan
Hashirimizu 1-10-20, Yokosuka, Kanagawa, Japan 239-8686
Email: {em54027, hidema}@nda.ac.jp

ABSTRACT

Intrusion Detection System (IDS) is a countermeasure against network attacks. There are mainly two types of detections; signature-based and anomaly based. Signature-based IDS detects intrusion packets by comparing contents of captured packets with the signature which is characteristic of intrusion packets. On the other hand, anomaly-based IDS detects them from normal behavior that is defined to distinguish normal communications from abnormal ones. Since attackers change their technique rapidly, anomaly-based detection draws research interest nowadays. However, since it is difficult to define normal behavior effectively, some anomaly-based IDS depends on visual identification of operator. To solve these problems, we propose a method using Detection-table which can be determined either normal or abnormal sessions. This method uses Discrete Fourier Transform and Shannon-Hartley theorem to analyze spectrum of each session. They assume fluctuation of spectrum in normal sessions as random and abnormal sessions as biased. To quantify difference between each spectrum and the standard one, we can obtain entropy using Shannon-Hartley theorem. Therefore, from the assumption, when entropy is small, we judge the session is normal, and when it is large, we judge the session is abnormal. By spectrum analysis based on such assumption, it is possible to derive the Detection-table. And we also find out that our quantification method will discover unknown abnormal sessions.

KEYWORDS

Intrusion Detection System (IDS), Discrete Fourier Transform (DFT), Shannon-Hartley theorem, window function, Kyoto2006+ dataset

1 INTRODUCTION

Intrusion Detection System (IDS) is a countermeasure against network attacks [1] [2]. Research of IDS has been conducted in many cases; such as [3] [4] [5] [6] [7] [8] [9] [10] [11]. The detection methods of IDS are divided into two types; signature-based IDS and anomaly-based IDS.

In signature-based IDS, characteristic of intrusion packets are stored as signatures in database. By comparing contents of captured packets with the signatures, intrusion packets can be detected. Snort [12] [13], Bro [14] [15], Swatch [16] and LogSurfer [17] are known as freeware-signature-based IDS. Snort is the most typical freeware of signature-based IDS and have a high detection rate. Bro enables to make signature to suit the purpose by using simple script. Since Swatch and LogSurfer get log-data by using syslog, they can detect intrusion packet by monitoring log-data. This type of IDS can judge recent sessions which are almost known attacks and are already analyzed.

However, this type does not detect unknown attacks. So in general, signature-based IDS has large false negative. And this type needs huge size of database of signature.

In anomaly-based IDS, normal behavior is defined to distinguish normal communications from abnormal ones. Therefore, it may detect unknown attacks. There are some existing methods; Wang et al. method [18], Imai et al. method [19], Sato et al. method [20] and Enkhbold method [21]. Wang et al. method are unsupervised using Mahalanobis distance. Sato et al. and Imai et al. method are also unsupervised using cluster analysis. Enkhbold method are spectrum analysis using Discrete Fourier Transform described in Section 2. This type of IDS is difficult to define normal. So in general, anomaly-based IDS has non-negligible false positive. And this type is difficult to operate. In fact, since almost methods depend on visual identification of operator, it is difficult to compare effec-

tiveness fairly and to quickly determine packets which are whether normal packets or abnormal ones.

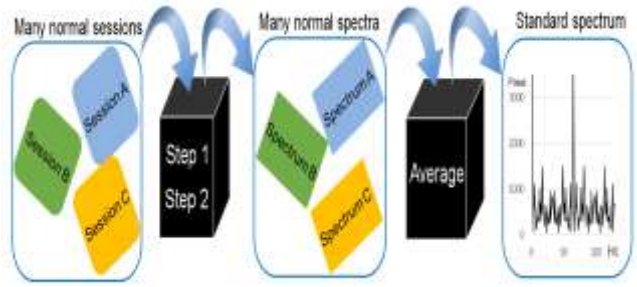

Figure 1. Outline of preparation

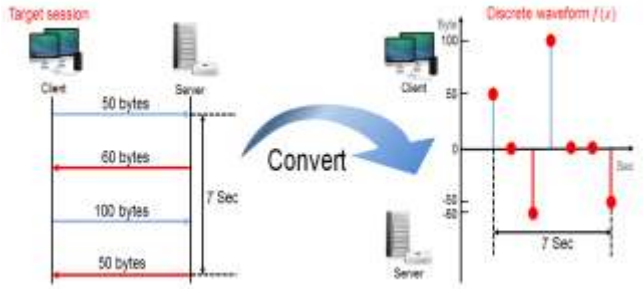

Figure 2. Outline of step1

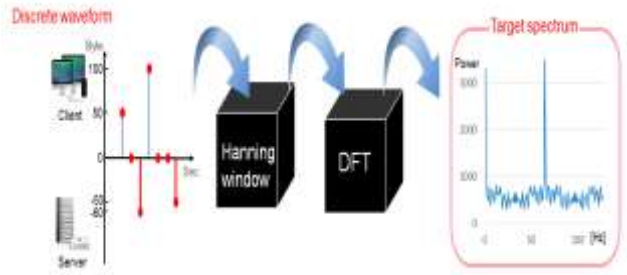

Figure 3. Outline of step2

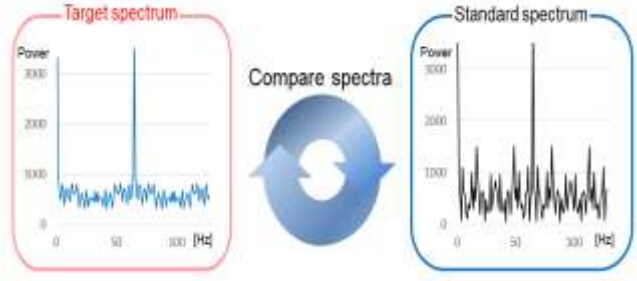

Figure 4. Outline of step3

Nowadays, the speed of complication and evolution of attack technique is fast, so necessity of anomaly-based IDS is increasing, in especially for critical communication system. Constructing IDS is not to choose signature-based IDS or anomaly-based IDS, we need to combine both of them efficiency. However, as mentioned above, it is possible to operate signature-based IDS automatically but is difficult for anomaly-based IDS. To realize this purpose, we need to solve the problem of anomaly-based IDS depended on decision of operator. In this paper, we solve this problem by proposing quantification method.

As mentioned above, there are many methods on anomaly-based IDS. In this paper, for solving the above problem, we focus on the technique of Enkhbold [21]. This method uses spectrum analysis of sessions by Discrete Fourier Transform (DFT). There are some methods using DFT (e.g., Zhou et al. [22]). They are different in focusing on the features in sessions ("time-interval" or "time-interval and payload"). In Enkhbold method [21], discrete waveforms are made from fluctuation of payloads, then each spectra of session is derived using DFT. By comparing spectra of sessions with standard spectrum which is derived from average of spectra of normal sessions, we can distinguish normal ones from abnormal ones.

And our previous improvement embedded window function into Enkhbold method to improve the efficiency in visual identification [23] (in the followings we call it previous method). However, since visual identification has no objectivity, we cannot compare it correctly. Also, previous method takes a long time to derive a spectrum (see Section 3.2).

To solve the problems, we propose quantification method using Shannon-Hartley theorem in this paper. In Section 2, we show the outline of previous method. Section 3 shows the basic idea of our proposal method using Shannon-Hartley theorem. In Section 4 and 5, we show our proposal method and example operation. In Section 6, we show our discussion. In Section 7, we discuss the advantageous of our method.

2 PREVIOUS METHOD AND PROBLEM

We define "session" the total communication set between one client and the server. Figure 1～4 shows outline of previous method [23]. It consists of followings.

Preparation (Figure 1): Make a discrete waveform from a payload and time elapsed of normal session. The standard spectrum derived from an average of the spectra.

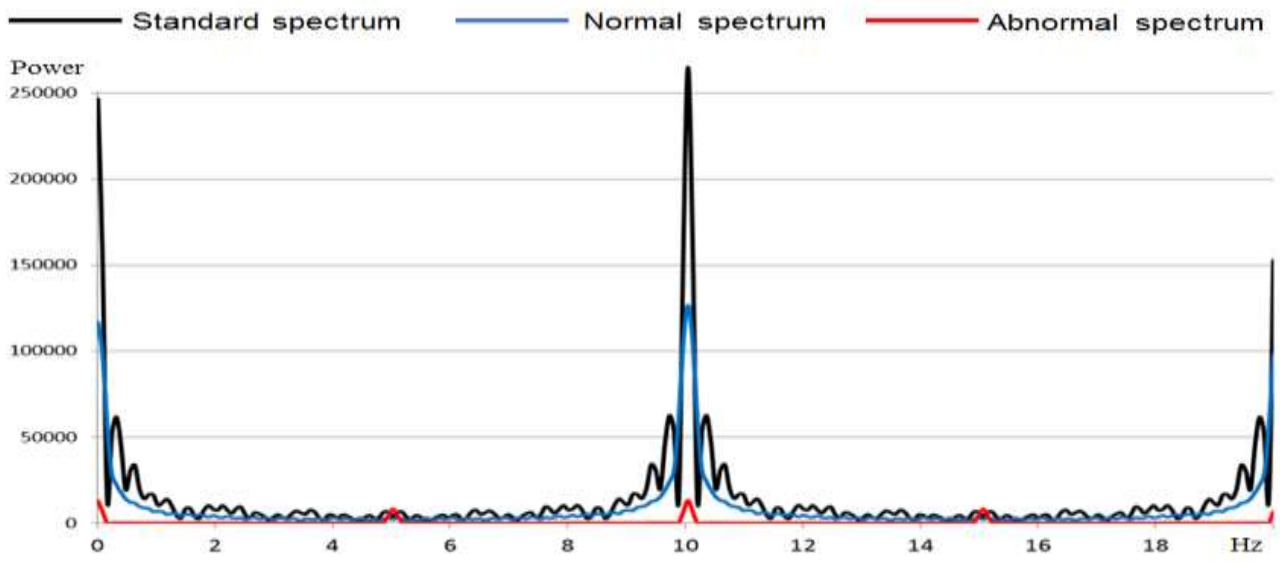

Figure 5. Example of abnormal detection [23]

Step-1(Figure 2): Make discrete waveform from target session.

Step-2(Figure 3): Apply window function to the discrete waveform. Perform DFT to the resultant.

Step-3(Figure 4): Compare the spectrum with the standard spectrum.

In Preparation, we make the standard spectrum. Its process is the same as the procedure of Step-1 and Step-2. We derive a lot of spectra from normal sessions, and the standard spectrum is derived from an average of the spectra. Note that normal sessions mean the sessions which are already checked as normal by other methods.

In Step-1, we make discrete waveform by regarding positive values as payload from client and negative value as payload from server. We make discrete waveform $f(x)$ based on time elapsed in transmission as shown in Figure 2. Let T be the session time from start to end ($0 \leq x \leq T$). Since the value of T changes for each session, when we perform DFT to any $f(x)$s of Step-2, each resultant spectrum has various frequency range. As the result, we cannot compare among spectra in Step-3. To solve this problem, in previous method, we normalize each session with $1/T$ and derive discrete waveform. In this process, when we take 10^{-m} of minimum scale (m decimal places of $1/T$), the discrete waveform has $N = 10^m$ points.

In Step-2, we perform DFT to discrete waveform $f(x)$ and make spectrum as follows.

$$|F(k)| = \sum_{n=0}^{N-1} \left(f(n) \times W_{han}(n)\right) e^{\frac{-i2\pi kn}{N}} \quad (k = 0,1,\cdots N-1) \tag{1}$$

$$W_{han}(n) = 0.5 - 0.5\cos\frac{2\pi n}{N-1} \tag{2}$$

where $|F(k)|$ is power of the spectrum. And $W_{han}(n)$ is Hanning window. Previous method shows the detailed analysis for the reason why Hanning window is effective for IDS using DFT [23].

In Step-3, we compare the spectrum derived in Step-2 with the standard spectrum. Figure 5 shows an example detection. We use visual identification in Figure 5, and focus on status of spectra between 0 [Hz] and 10 [Hz]. The behavior of standard spectrum becomes random in the frequency range.

However, abnormal spectrum which are derived from abnormal session has almost constant comparing with the standard spectrum, and two large peaks

Figure 6. Behaviors of normal sessions

are found around 5[Hz] and 10[Hz]. As the result, we can distinguish normal spectra from abnormal ones.

However, the previous method has two problems. Firstly, the method uses visual identification. Since this scheme has no objectivity, detection results become ambiguous. Therefore, we cannot compare a spectrum correctly. Also, since it requires significant human effort, it is not efficient as a detection method. Secondly, the method has to take a certain time to derive a spectrum (see Section 3.2). This is a critical issue for IDS which is required quick detection. To solve these problems, we show two basic ideas in the following section.

3 BASIC IDEA

We show two ideas to solve the problems. The first idea is a solution for the visual identification problem. The second idea is a solution for the problem which takes a lot time to derive a spectrum.

3.1 First idea

To solve the visual identification problem, we propose quantification method using Shannon-Hartley theorem. The principle of previous method is based on the following assumptions.

1. behaviors of normal sessions are various seems and to be random.
2. behaviors of abnormal sessions have some characteristics and biases.

From the viewpoint of spectrum analysis, the spectra of normal sessions become noise spectrum (Figure 6) and ones of abnormal sessions becomes biased spectrum (Figure 7).

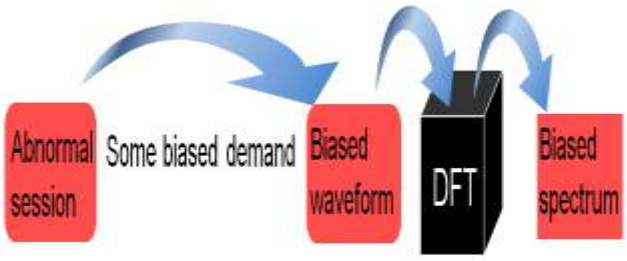

Figure 7. Behaviors of abnormal sessions

Therefore when we define the standard spectrum as noise, we can calculate entropy using Shannon-Hartley theorem [24].
Then, if the target spectrum is abnormal, the value of entropy will be large, on the other hand, if it is normal, the value will be small. This assumption is the quantification for IDS using DFT. Shannon-Hartley theorem is shown as follows.

$$C = Blog_2(1 + \frac{S(f)}{N(f)}) \quad (3)$$

where B denotes bandwidth [Hz] of the channel. $S(f)$ denotes the average received signal power and $N(f)$ denotes the one of the noise and interference over the bandwidth. However, since we have discrete values over frequency range, we rewrite above equation as follows.

$$I = \int_{f1}^{f2} \log_2 \left(1 + \frac{S(f)}{N(f)}\right) df \quad (4)$$

where $S(f)$ and $N(f)$ denote signal power and noise power of frequency f respectively. To satisfy the non-negativity of entropy and calculation for discrete waveform, we rewrite above equation as follows.

$$I_s = \sum_{f=0}^{N-1} \log_2(1 + \frac{\max\{S_s(f), S_t(f)\}}{\min\{S_s(f), S_t(f)\}}) \times \Delta f \quad (5)$$

Where $S_s(f)$ and $S_t(f)$ denote standard spectrum power and target spectrum power at point f. And where Δf denotes the unit frequency scale which is calculated as follows.

$$\Delta f = \frac{f_s}{N-1} \quad (6)$$

Figure 8. Outline of P1

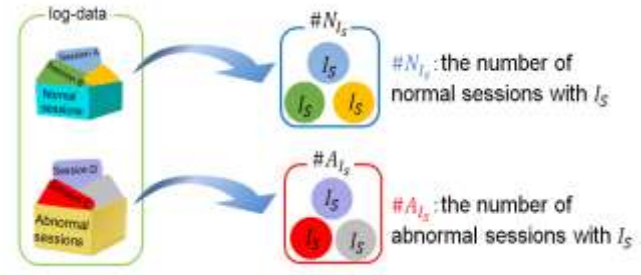

Figure 10. Outline of P3

where f_s denotes the sampling rate for a real network environment. In this paper, we determine it by the average of total number of sessions per unit time. In the followings, we call I_S evaluation-value. We can judge whether the target session is normal or abnormal using this evaluation-value, however, the value of normal and one of abnormal may never be different. Therefore we make the Detection-table which shows each range of evaluation-value is normal or abnormal with probability. In this procedure, we need trusted log-data which classify normal sessions and abnormal sessions.

3.2 Second idea

Since the previous method derives a spectrum by using the session time T. If the session time is long, we cannot judge immediately. In IDS which is required quick response, this is critical condition. Therefore we set initial session time in advance to short time for detection. As the result, we omit communications over initial session time. We need some changes in previous method. Firstly, we do not need normalization procedure in Step-1 because it is done in advance. Secondly, we should not apply window function in Step-2. By applying initial session time, the number of communication in a

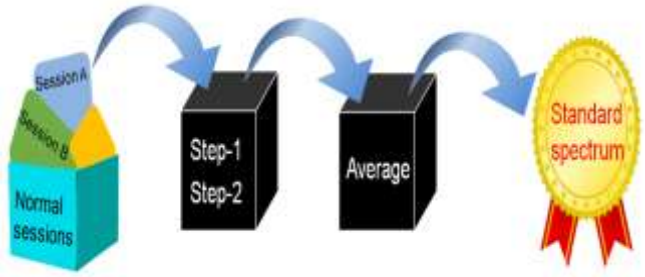

Figure 9. Outline of P2

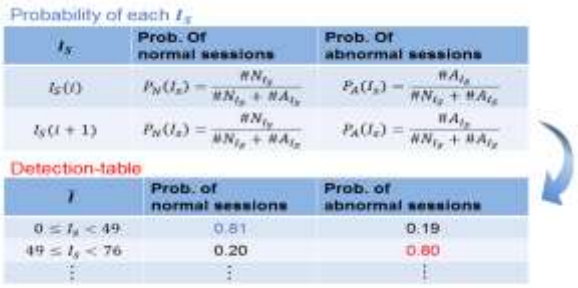

Figure 11. Outline of P4

session is decreased. So, we need to make characteristics of session stand out efficiently. On the other hand, since window function regards such characteristic in short time as noise, some significant feature will be lost. Therefore, we should not use window function. As the result, we conclude that our improvements for previous method is effective for visual identification [21]. However, they do not contribute to our quantification method for short time.

4 PROPOSAL METHOD

Our proposal method has two phase; Preparation phase and Detection phase. Figure 8～11 shows outline of Preparation phase and Figure 12 shows outline of Detection phase.

4.1 Preparation phase

P1: Collecting and classifying log-data

(Figure 8)

The purpose of proposal method is to quantify the difference between the standard spectrum and target ones. Therefore we need both normal and abnormal session log-data. And also we need same methods to classify normal session and abnormal session correctly. For this procedure, it is desirable

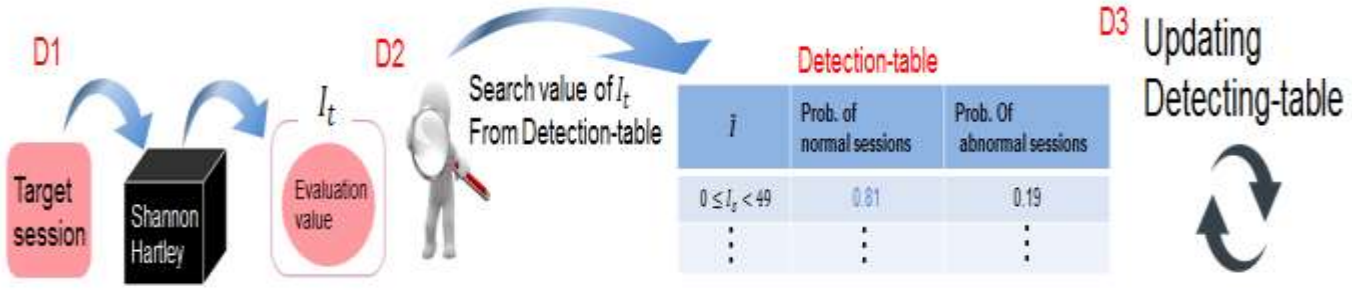

Figure 12. Outline of Detection phase

Table 1. Outline of honey pot

Type	Number of machines
Solaris 8 (Symantec based)	4
Windows XP(full patch)	1
Windows XP(no patch)	5
Windows XP SP2	2
Windows Vista	1
Windows 2000 Server	1
Mac OS X	2(one is mail server)
Printer	2
TV set	1
HDD recorder	1
SGNET honeypots	5
dedicated honeypots	4
Web Crawler	1
Black hole sensor /24	1
Black hole sensor /26	1

to hold log-data in a long term and use some methods such as signature type IDS which can detect definitely.

P2: Derivation of standard spectrum
(Figure 9)

Among the log-data which is prepared in P1, we use normal sessions to derive the standard spectrum using Step-1 and Step-2 of previous method.

P3: Calculation of evaluation-value
(Figure 10)

We calculate evaluation-values of all session in log-data using eq. (5) and eq. (6). We reclassify the result of log-data (normal or abnormal session) using each evaluation-value (I_s), and count the number of normal session ($\#N_{I_s}$) and the number of abnormal session ($\#A_{I_s}$) for each I_s.

P4: Construction of Detection-table
(Figure 11)

Let $P_N(I_S)$ be the probability of normal session with evaluation-value which is equals to I_S.And let $P_A(I_S)$ be the probability of abnormal session.

$$P_N(I_s) = \frac{\#N_{I_s}}{\#N_{I_s} + \#A_{I_s}} \tag{7}$$

$$P_A(I_s) = \frac{\#A_{I_s}}{\#N_{I_s} + \#A_{I_s}} \tag{8}$$

Let Q, $(0.5 < Q \le 1.0)$ be the threshold of successful detection probability. We search for the

range of evaluation-value $\tilde{I}_S$ which satisfies followings.

$$\sum_{I_S \in \tilde{I}_S} P_N(I_s) \geq Q \quad \text{or} \quad \sum_{I_S \in \tilde{I}_S} P_A(I_s) \geq Q \tag{9}$$

Table 2. Feature of Kyoto 2006+ dataset

Conventional features	Additional features
Duration	IDS detection
Service	Malware detection
Source bytes	Ashula detection
Destination bytes	Label
Count	Source IP address
Same srv rate	Source Port number
Serror rate	Destination IP address
Srv serror rate	Destination Port number
Dst host count	Start time
Dst host srv count	Duration
Dst host same src port rate	
Dst host serror rate	
Dst host srv serror rate	
Flag	

Table 3. Specification of computer experiment

Log-data	Kyoto2006+ dataset (15,746,592 sessions)
Sampling rate	f_s = 7.4 [Hz]
Threshold	Q = 0.8
Initial session time	T = 2 [sec]
OS	Windows 7 Professional
CPU	Intel Corei7-3770 3.4 GHz
RAM	16.0 GB
Programming language	Visual Basic for Applications

where

$$\tilde{I}_S = \{I_S | I_S(i) \leq I_S < I_S(j)\} \tag{10}$$

Note that $I_S(m)$ denotes m-th ($0< m$) value of I_S.

4.1 Detection phase (Figure 12)

D1: Calculation of evaluation-value of target session

We derive the spectrum of target session using Step-1 and Step-2 which use initial session time and calculate evaluation-value I_t using the standard spectrum.

D2: Look-up Detection-table

We search for the range $\tilde{I}_s$ which involves the value of I_t in Detection-table. Normal session or abnormal session is judged by the probability which exceeds the threshold Q. Then falsenegative (or false-positive) can be evaluated as $1 - Q$.

D3: Updating Detection-table

When the result of detection is confirmed true, we update Detection-table to improve successful detection probability.

5 EXPERIMENT

5.1 Kyoto2006+ dataset

In this paper, we use Kyoto2006+ dataset [25]. This dataset was derived from the actual data traffic during November 2006 to August 2009 by using the honey pot which is installed in Kyoto University. A structure of honey pot is shown in Table 1. This dataset consists of 14 conventional features and 10 additional features (see Table 2). These 14 features were extracted based on KDD Cup 99 data set [26] which is very popular and widely used performance evaluation data for IDS. We use Duration, Source IP address, Destination IP address, Source bytes, Destination bytes and Label

(shaded in Table 2). The label can classify as either normal session or abnormal session for each session. In fact, this dataset is old to use for evaluation of IDS performance. However, since this is an open public, it is possible for third persons to verify the

Table 4. Detection-table

$\tilde{I}_s$	Prob. of normal sessions	Prob. of abnormal sessions	Number of sessions
$0 \leq I_s < 49$	80.6%	19.4%	11,300,021
$49 \leq I_s < 76$	19.6%	80.4%	3,754,637
$76 \leq I_s < 80$	54.1%	45.9%	239,087
$80 \leq I_s$	20.0%	80.0%	452,847

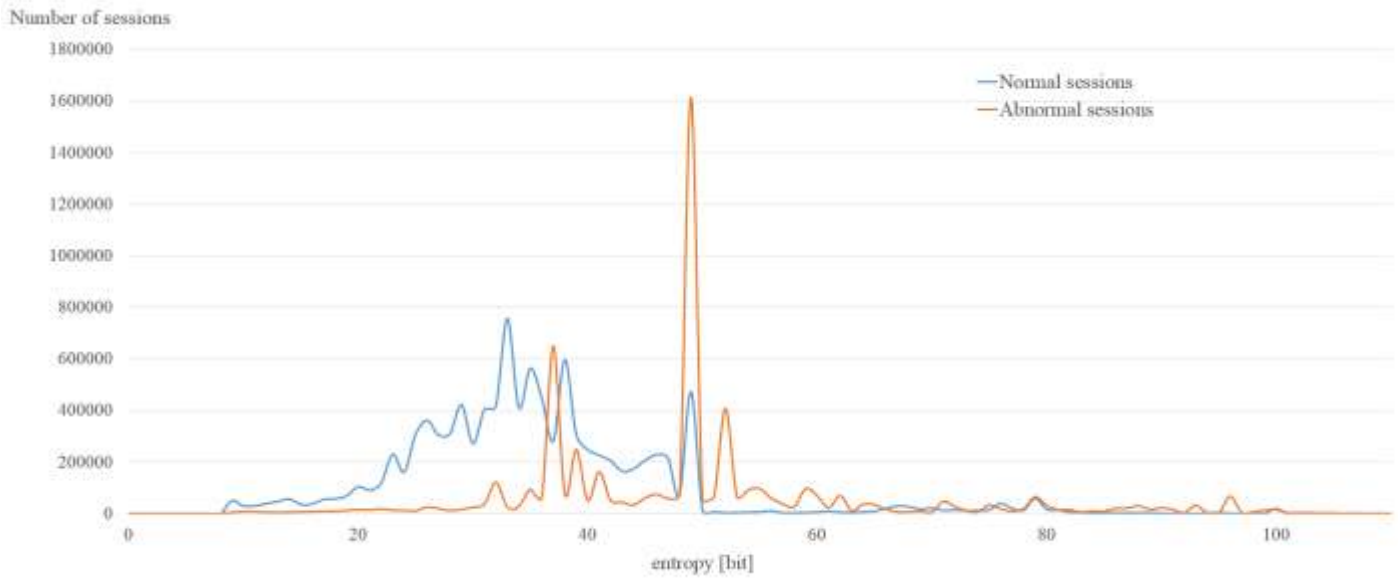

Figure 13. Distribution of number of sessions

effectiveness and compare the effectiveness among other methods. In addition, by using the label, we can confirm the successful detections.

5.2 Preparation phase

We omit 0 byte of payload and 0 second of session time in Kyoto2006+ dataset because these types of session cannot contribute to derivation of Detection-table. There are total 15,746,592 target sessions. As mentioned in Section 3, we define sampling rate f_s by the average of total number of sessions per unit time. Table 3 shows the specification of our computer environment.

5.3 Detection phase

When we use the Detection-table shown in Table 4, we can find following facts.

1) The assumption about evaluation-value is true.

In Section 3, we described our assumption that

when evaluation-value is small, the session will be normal, on the other hand, when it is large, the session will be abnormal. We can confirm that this assumption in Table 4; the range $\tilde{I}_s: 0 \leq I_s < 49$ is normal with probability of 80.6%, the range $\tilde{I}_s: 49 \leq I_s < 76$ is abnormal with probability of 80.4% and the range $\tilde{I}_s: 80 \leq I_s$ is abnormal with probability of 80.0%. In particular, our assumption can be confirmed clearly in detection of abnormal sessions.

2) Ranges of evaluation-value which are under the threshold Q exist.

The ranges $\tilde{I}_s: 76 \leq I_s < 80$ are under the threshold. In fact, we cannot solve this problem, and

compromised this result. For these ranges, where under the threshold, our detection results will become unstable, however, these cases are only 1.52% of the whole of log-data (Figure 13). From this fact, we conclude that the sessions whose evaluation-values are included in these range are too small in Kyoto

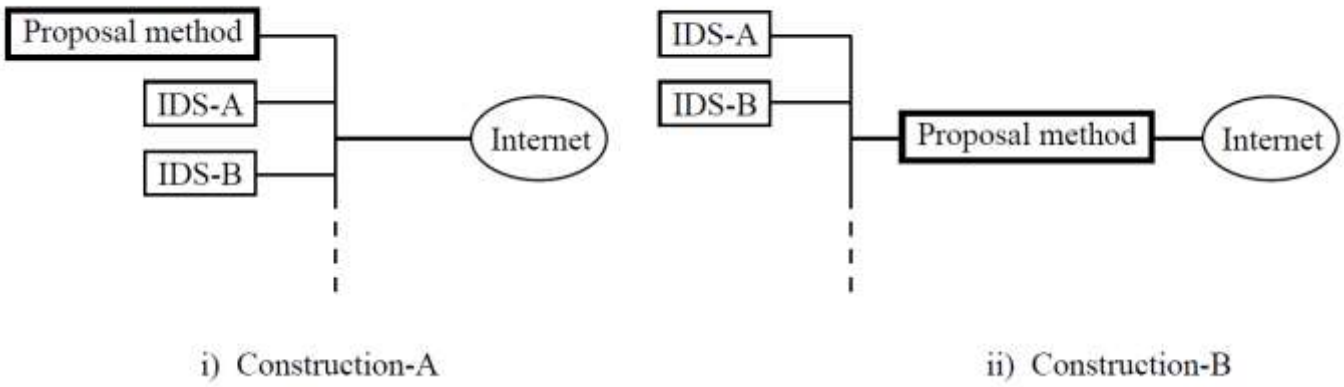

Figure 14. Position of proposal method in IDS construction

2006+ dataset. Therefore, we expect that this problem will be solved by updating Detection-table using more sessions.

3) Where is unknown abnormal sessions?

A possibility that we can find out the unknown abnormal sessions is low in the ranges which the threshold of successful detection is satisfied. Because these ranges have enough sample sessions to analyze in details. Therefore, we have to pay attention and analyze in the range where the threshold is not satisfied. As the result, we can expect that our proposal method will contribute the effective detection and analysis of unknown abnormal session. As mentioned in Section 4.2, Detection-table is updated using detection results, the probability of successful detection always keeps being improved. Therefore, the result shown in Table 4 is the initial state of our proposal method for the communication environment of Kyoto 2006+ dataset. At the same time, by resetting the threshold *Q* high, suspicious sessions will be found easily.

6 DISCUSSION

In Section 3, we assume that the spectra of normal sessions become random noise and ones of abnormal sessions have peaks. Based on this assumption, the previous method obtained result of Figure 5. In the previous method, operators conduct spectrum analysis by using visual identification. It is ambiguous to depend on visual identification of operator. Also, the previous method takes a long time to derive a spectrum. On the other hand, the proposal method defines the standard spectrum as noise, we can calculate entropy using Shannon-Hartley theorem. We can expect that normal spectrum will be small and abnormal one will be large. And we can make the Detection-table by using assumption. We can conduct spectrum analysis by using Detection-table without using visual identification of operator. Also, since the proposal method uses initial session time, deriving a spectrum finishes in short. From these improvements, proposal method solves problems of the previous method. Therefore, we conclude that our proposal method is more excellent than the previous one.

7 CONCLUSION

In this paper, we propose a quantification method for IDS using DFT and define the Detection-table. Using the Detection-table, we can operate IDS using DFT by deterministic algorithm. There are some concepts to construct IDS shown in Figure 14. Construction-A is the majority decision type. Obviously, our proposal method is not effective in this position. Construction-B is adequate position for our proposal method and it will work as proac-

tive detection as mentioned in Section 5.3. The possibility which the target session is unknown abnormal session is high when its evaluation-value is involved in the range where threshold Q is not satisfied. Unfortunately, Kyoto 2006+ dataset does not include any unknown abnormal sessions, we cannot confirm the effectiveness of our proposal method concerning to this feature. And as mentioned in Section 4.2 (D3), we can update the Detection-table. However, the restrictions in time and on the computer environments did not enable us to execute this procedure in this experiment. These are our future works.

In the operation of our proposal method, we need only the Detection-table as the dataset. Therefore we can conclude that our proposal method is very low-cost IDS and very fast computational method. These feature will enable to construct real-time detection. This is also our future work.

Obviously, it does not need to point out, the Detection-table is various in the communication environment. We need to more experiments in various communication systems.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 24560491.

REFERENCES

1. G. Bruneau, "The history and evaluation of intrusion de tection," https://www.sans.org/reading-room/whitepapers/detection/ history-evolution-intrusion-detection-344.

2. J. Anderson, "Computer security technology planning study volume ii," *Electronic Systems Division Technical Report*, pp. 73–51, 1972.

3. D. E. Denning, "An intrusion-detection model," *IEEE Transactions on software engineering*, no. 2, pp. 222–232, 1987.

4. V. V. K. Labib and V. R. Vemuri, "Anomaly detectionusing s language framework: Clustering and visualization of intrusive attacks on computer systems," in *Fourth Conference on Security and Network Architectures, SAR '05*. Citeseer, 2005.

5. J. Hochberg, K. Jackson, C. Stallings, J. McClary, D. DuBois, and J. Ford, "Nadir: An automated system for detecting network intrusion and misuse," *Computers & Security*, vol. 12, no. 3, pp. 235–248, 1993.

6. L. Portnoy, E. Eskin, and S. Stolfo, "Intrusion detection with unlabeled data using clustering," in *In Proceedings of ACM CSS Workshop on Data Mining Applied to Security (DMSA-2001*. Citeseer, 2001.

7. S. Kumar and E. H. Spafford, "A pattern matching model for misuse intrusion detection," 1994. [8] S. Staniford-Chen, S. Cheung, R. Crawford, M. Dilger, J. Frank, J. Hoagland, K. Levitt, C. Wee, R. Yip, and D. Zerkle, "Grids-a graph

8. S. Staniford-Chen, S. Cheung, R. Crawford, M. Dilger, J. Frank, J. Hoagland, K. Levitt, C. Wee, R. Yip, and D. Zerkle, "Grids-a graph based intrusion detection system for large networks," in *Proceedings of the 19th national information systems security conference*, vol. 1. Baltimore, 1996, pp. 361–370.

9. U. Lindqvist and P. A. Porras, "expert-bsm: A host-based intrusion detection solution for sun solaris," in *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*. IEEE, 2001, pp. 240–251.

10. D. Curry, H. Debar, and B. Feinstein, "Intrusion detection message exchange format data model and extensible markup language (xml) document type definition," *IDWG, February*, 2002.

11. E. Eskin, "Anomaly detection over noisy data using learned probability distributions," in *In Proceedings of the International Conference on Machine Learning*. Citeseer, 2000.

12. "Snort," https://www.snort.org/.

13. M. Roesch *et al.*, "Snort: Lightweight intrusion detection for networks." in *LISA*, vol. 99, no. 1, 1999, pp. 229–238.

14. "The bro network security monitor," https://www.bro.org/.

15. V. Paxson, "Bro: a system for detecting network intruders in real-time," *Computer networks*, vol. 31, no. 23, pp. 2435–2463, 1999.

16. "Swatch," http://swatch.sourceforge.net/.

17. "Logsurfer," https://www.cert.dfn.de/eng/logsurf/.

18. K. Wang and S. J. Stolfo, "Anomalous payloadbased network intrusion detection," in *International Workshop on Recent Advances in Intrusion Detection*. Springer, 2004, pp. 203–222.

19. K. Imai, S. Aoki, and T. Miyamoto, "Anomaly detection

based on clustering of network traffic characteristics considering results of signature based ids evaluation," *ICISS Technical Report*, vol. 489, no. 114, pp. 7–12, 2015.

20. M. Sato, H. Yamaki, and H. Takakura, "Unknown attacks detection using feature extraction from anomaly-based ids alerts," in *Applications and the Internet (SAINT), 2012 IEEE/IPSJ 12th International Symposium on*. IEEE, 2012, pp. 273–277.

21. E. Chimedtseren, K. Iwai, H. Tanaka, and T. Kurokawa, "Intrusion detection system using discrete fourier transform," *Proceedings on the 7th IEEE Symposium on Computational Intelligenece for Security and Defense Applications(CISDA)*, no. CS3-1, pp. 1–5, 2014.

22. M. Zhou and S.-D. Lang, "A frequency-based approach to intrusion detection," in *Proc. of the Workshop on Network Security Threats and Countermeasures*, 2003.

23. Y. Tsuge and H. Tanaka, "Intrusion detection system using discrete fourier transform with window function," *International Journal of Network Security & Its Applications (IJNSA)*, vol. 8, no. 2, pp. 23–34, 2016.

24. C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

25. J. Song, H. Takakura, Y. Okabe, M. Eto, D. Inoue, and K. Nakao, "Statistical analysis of honeypot data and building of kyoto 2006+ dataset for nids evaluation," in *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*. ACM, 2011, pp.29–36.

26. KDDCup1999Data, "The third international knowledge discovery and data mining tools competition dataset kdd99-cup," https://kdd.ics.uci.edu/dataases/kddcup99/kddcup99.html.