15

Signal Processing Methodology for Network Anomaly Detection

Rafał Renk[2,3] , Michał Choraś[1,3],
Łukasz Saganowski[1,3], Witold Hołubowicz[2,3]
[1]*University of Technology and Life Sciences, Bydgoszcz*
[2]*Adam Mickiewicz University, Poznań*
[3]*ITTI Ltd. Poznań*
Poland

1. Introduction

Intrusion Detection Systems (IDS) can be classified as belonging to two main groups depending on the detection technique employed:

– signature-based detection,

– anomaly detection.

Currently, most IDS systems have problems in recognizing new attacks (0-day exploits) since they are based on the signature-based approach. In such mode, when system does not have an attack signature in database, such attack is not detected. Another drawback of current IDS systems is that the used parameters and features do not contain all the necessary information about traffic and events in the network (Coppolino at al., 2009).
On the other hand, anomaly detection techniques rely on the existence of a reliable characterization of what is normal and what is not, in a particular networking scenario. More precisely, anomaly detection techniques base their evaluations on a model of what is normal, and classify as anomalous all the events that fall outside such a model.
In this paper, a new solution for Anomaly Detection System (ADS) system based on signal processing algorithm is presented. ADS analyzes traffic from internet connection in certain point of a computer network. The proposed ADS system uses redundant signal decomposition method based on Matching Pursuit algorithm.
Our original methodology for network security anomaly detection based on Matching Pursuit is presented and evaluated using network data traces. We also compared Matching Pursuit approach to Discrete Wavelet Transform used by other researchers.
The paper is structures as follows: firstly, in Section 2 the motivation to use signal processing techniques in intrusion and anomaly detection systems is provided. In Section 3 the anomaly detection system based on the Matching Pursuit is presented in detail. The effectiveness of the proposed system is evaluated in Sections 4 and 5 where the comparison to the state-of-the-art method based on Discrete Wavelet Transform is shown.

Feature ID	Feature Description
1	ICMP flows/time period
2	ICMP in bytes/time period
3	ICMP out bytes/time period
4	ICMP in frames/time period
5	ICMP out frames/time period
6	TCP flows/time period
7	TCP input bytes/time period
8	TCP out bytes/time period
9	TCP in frames/time period
10	TCP out frames/time period
11	UDP flows/time period
12	UDP in bytes/time period
13	UDP out bytes/time period
14	UDP in frames/time period
15	UDP out frames/time period

Table 1. Network traffic parameters

2. Signal processing techniques applied to anomaly detection

Signal processing techniques have found application in Network Intrusion Detection Systems because of their ability to detect novel intrusions and attacks, which cannot be achieved by signature-based approaches (Esposito et al., 2005). It has been shown that network traffic presents several relevant statistical properties when analyzed at different levels (e.g. self-similarity, long range dependence, entropy variations, etc.) (Esposito et al., 2005)(Cheng et al., 2002).

Approaches based on signal processing and on statistical analysis can be powerful in decomposing the signals related to network traffic, giving the ability to distinguish between trends, noise, and actual anomalous events. Wavelet-based approaches, maximum entropy estimation, principal component analysis techniques, and spectral analysis, are examples in this regard which have been investigated in the recent years by the research community (Cheng et al., 2002)(Barford et al., 2002)(Huang et al., 2001)(Li & Lee, 2003)(Dainotti et al., 2006). However, Discrete Wavelet Transform provides a large amount of coefficients which not necessarily reflect required features of the network signals.

Therefore, in this paper we propose another signal processing and decomposition method for anomaly/intrusion detection in networked systems. We developed original Anomaly Detection Type IDS algorithm based on Matching Pursuit. As to our best knowledge, we have not met any other IDS system based on matching pursuit.

ADS based on Matching Pursuit uses Dictionary of Base Functions - BFD to decompose input 1D traffic signal (1D signal may represent for example packets per second) into set of based functions called also atoms. The proposed BFD has ability to approximate traffic signal.

In the proposed system we use 15 network traffic parameters shown in Table 1.

3. Anomaly detection system based on matching pursuit

Matching Pursuit is a known signal processing technique used for instance in audio compression, image and video compression (Mallat et al., 1993)(Neff et al., 2002)(Figureas

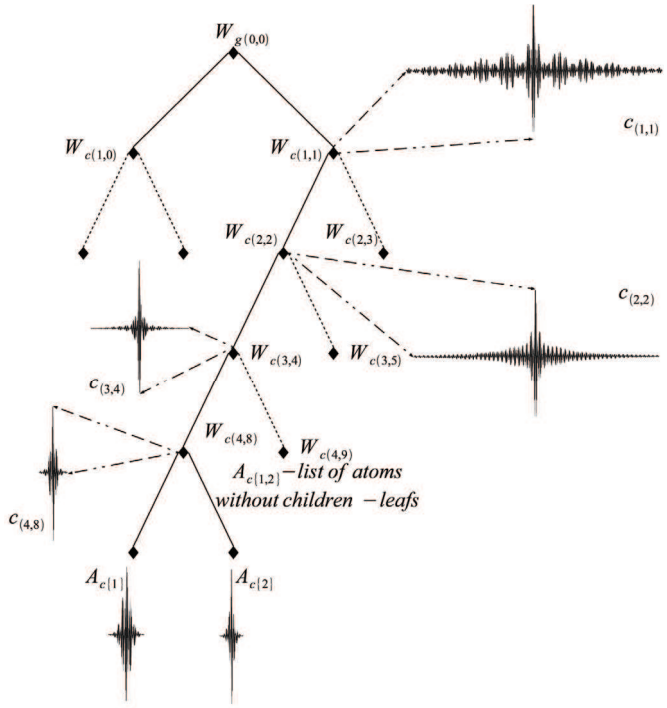


Fig. 1. Example dictionary structure

et al., 2006)(Shaopeng et al., 2002)(Daudet, 2010). However, as to our best knowledge, we are the first to use Matching Pursuit for intrusion and anomaly detection in computer networks. Matching Pursuit signal decomposition was proposed by Mallat and Zhang (Mallat et al., 1993). Matching Pursuit is a greedy algorithm that decomposes any signal into a linear expansion of waveforms which are taken from an overcomplete dictionary D. The dictionary D is an overcomplete set of base functions called also atoms.

$$D = \{\alpha_\gamma : \gamma \in \Gamma\} \tag{1}$$

where every atom α_γ from dictionary has norm equal to 1:

$$\|\alpha_\gamma\| = 1 \tag{2}$$

Γ represents set of indexes for atom transformation parameters such as translation and scaling. Signal s has various representations for dictionary D. Signal can be approximated by set of atoms α_k from dictionary and projection coefficients c_k:

$$s = \sum_{n=0}^{|D|-1} c_k \alpha_k \tag{3}$$

In the basic Matching Pursuit algorithm atoms are selected in every step from entire dictionary which has flat structure. In this case algorithm causes significant processor burden. Therefore, a dictionary with internal structure was used in our coder. Such dictionary is built from: atoms and centered atoms. Centered atoms group such atoms from D that are as correlated (to each other) as possible.
To calculate the measure of correlation between atoms, the function $o(a,b)$ can be used (Jost et al., 2005):

$$o(a,b) = \sqrt{1 - \left(\frac{|\langle a,b \rangle|}{\|a\|_2 \|b\|_2} \right)^2} \tag{4}$$

The quality of centered atom can be estimated according to (5):

$$O_{k,l} = \frac{1}{|LP_{k,l}|} \sum_{i \in LP_{k,l}} o\left(A_{c(i)}, W_{c(k,l)}\right) \tag{5}$$

$LP_{k,l}$ is a set of atoms grouped by centered atom. $O_{k,l}$ is mean of local distances from centered atom $W_{c(k,l)}$ to the atoms $A_{c(i)}$ which are strongly correlated with $A_{c(i)}$.
Example dictionary structure was presented in Figure 1. Atom tree consist of root node W_g and every centroid consist of two children. Parameter A_c represents leaf nodes (without children).
Centroid $W_{c(k,l)}$ represents atoms $A_{c(i)}$ which belong to the set $i \in LP_{k,l}$. List of atoms $LP_{k,l}$ should be selected according to the Equation 6:

$$\max_{i \in LP_{k,l}} o\left(A_{c(i)}, W_{c(k,l)}\right) \leq \min_{t \in D \setminus LP_{k,l}} o\left(A_{c(t)}, W_{c(k,l)}\right) \tag{6}$$

In the proposed ADS solution $1D$ real Gabor base function (Equation 7) was used to build dictionary (Troop, 2004)(Gribonval, 2001).

$$\alpha_{u,s,\xi,\phi}(t) = c_{u,s,\xi,\phi} \alpha\left(\frac{t-u}{s}\right) \cos(2\pi\xi(t-u) + \phi) \tag{7}$$

where:

$$\alpha(t) = \frac{1}{\sqrt{s}} e^{-\pi t^2} \tag{8}$$

$c_{u,s,\xi,\phi}$ - is a normalizing constant used to achieve atom unit energy.
In order to create overcomplete set of $1D$ base functions dictionary D was built by varying subsequent atom parameters: Frequency ξ and Phase ϕ, Position u, Scale s. Base functions dictionary D was created with using 10 different scales (dyadic scales) and 50 different frequencies.
Traffic Parameters are used to create one dimensional signal. This signal is decomposed with the use of Matching Pursuit transformation. After MP decomposition we achieved projection coefficients c_k which are used for creating normal traffic profiles.

Network Traffic Feature	Total number of of attack	Detected number of attack	Detection Rate [%]
ICMP flows/minute	73	61	84.93
ICMP in bytes/minute	73	31	43.83
ICMP out bytes/minute	73	39	54.79
ICMP in frames/minute	73	59	82.19
ICMP out frames/minute	73	65	90.41
TCP flows/minute	73	68	94.52
TCP in bytes/minute	73	32	46.57
TCP out bytes/minute	73	31	45.20
TCP in frames/minute	73	57	79.45
TCP out frames/minute	73	54	76.71
UDP flows/minute	73	41	58.90
UDP in bytes/minute	73	52	73.97
UDP out bytes/minute	73	73	100.00
UDP in frames/minute	73	52	73.97
UDP out frames/minute	73	70	98.63

Table 2. Detection Rate for W5D1 (Fifth Week, Day 1) (DARPA, 2000) trace

Matching Pursuit (Mallat et al., 1993),(Jost et al., 2005) algorithm (stop condition of the MP algorithm) was modified in order to encode only sufficient number of atoms. This number of atoms may be different for a given traffic (signal) analysis window (for e.g. for $10min$. analysis window number of encoded atoms may be between 3 and max 10). This operation causes a significant reduction of the algorithm execution time.
Normal traffic profiles are calculated using input traffic without attack and anomalies. Normal profiles are calculated separately for every traffic feature. Our ADS system compares current traffic traces (to be analyzed) with the reference profiles calculated during normal work

Network Traffic Feature	Total number of of attack	Detected number of attack	Detection Rate [%]
ICMP flows/minute	68	49	72.06
ICMP in bytes/minute	68	56	82.35
ICMP out bytes/minute	68	54	79.41
ICMP in frames/minute	68	59	86.76
ICMP out frames/minute	68	56	82.35
TCP flows/minute	68	37	54.41
TCP in bytes/minute	68	41	60.29
TCP out bytes/minute	68	23	33.82
TCP in frames/minute	68	31	45.58
TCP out frames/minute	68	32	47.05
UDP flows/minute	68	66	97.05
UDP in bytes/minute	68	62	91.17
UDP out bytes/minute	68	60	88.23
UDP in frames/minute	68	62	91.18
UDP out frames/minute	68	60	88.24

Table 3. Detection Rate for W5D5 (Fifth Week, Day 5) (DARPA, 2000) trace

TCP trace (packet/second) (MAWI, 2005)	Window1 MPMP	Window2 MPMP	Window3 MPMP	MPMP for trace	MPMP for normal trace
Mawi 2004.03.06 tcp	210.34	172.58	239.41	245.01	240.00
Mawi 2004.03.13 tcp	280.01	214.01	215.46	236.33	240.00
Mawi 20.03.2004 tcp (attacked: Witty)	**322.56**	**365.24**	**351.66**	346.48	240.00
Mawi 25.03.2004 tcp (attacked: Slammer)	**329.17**	**485.34**	**385.50**	400.00	240.00

Table 4. Matching Pursuit Mean Projection - MP-MP for TCP trace (20 min. analysis window)

(stored in a database). ADS system makes an alarm when difference between profiles exceed certain threshold (we usually use the threshold value equal to 30%).

4. Evaluation of the proposed anomaly detection system based on matching pursuit

In our previous work we showed the first results of the Matching Pursuit methodology for anomaly detection using only one network traffic metric (feature), namely packets per second (Saganowski et al. , 2009). Hereby, we in our anomaly detection system, we correlate much more network traffic features (Table 1).
Performance of our approach was evaluated with the use of the following trace bases:

– (DARPA, 2000),
– (MAWI, 2005),
– (CAIDA, 2004),
– (UNINA, 2009) (University of Napoli network traces).

The test data contains attacks that fall into four main categories (Wei et al., 2009) such as:

1. DOS/DDOS: denial-of-service, e.g. syn flood,
2. R2L: unauthorized access from a remote machine, e.g. guessing password,
3. U2R: unauthorized access to local superuser (root) privileges, e.g., various ”buffer overflow” attacks,
4. PROBING: surveillance and other probing, e.g., port scanning.

UDP trace (packet/second) (MAWI, 2005)	Window1 MPMP	Window2 MPMP	Window3 MPMP	MPMP for trace	MPMP for normal trace
Mawi 2004.03.06 tcp	16.06	13.80	17.11	15.65	16.94
Mawi 2004.03.13 tcp	20.28	17.04	17.40	18.24	16.94
Mawi 20.03.2004 tcp (attacked: Witty)	**38.12**	**75.43**	**61.78**	58.44	16.94
Mawi 25.03.2004 tcp (attacked: Slammer)	**56.13**	**51.75**	**38.93**	48.93	16.94

Table 5. Matching Pursuit Mean Projection - MP-MP for UDP trace (20 min. analysis window)

TCP trace (packet/second) (CAIDA, 2004)	Window1 MPMP	Window2 MPMP	Window3 MPMP	MPMP for trace	MPMP for normal trace
Backscatter 2008.11.15	147.64	**411.78**	**356.65**	305.35	153.66
Backscatter 2008.08.20	**208.40**	161.28	153.47	147.38	153.66

Table 6. Matching Pursuit Mean Projection - MP-MP for TCP trace with DDoS attacks (20 min. analysis window)

For experiments we chose 20 minutes analysis window because most of attacks (about 85%) ends within this time period. We extracted 16 traffic features in order to create 1D signals for Matching Pursuit - Mean Projection analysis. Traffic features were calculated with the use of 1 minute time period.
In Table 2 and Table 3 detection rates achieved for DARPA benchmark trace base are presented. These are results achieved for two test days. Detection results were compared to the list of attacks which should exist in this two testing days.
In Table 4 and Table 5 there are results for MAWI test base. Bold numbers in tables point to existence of anomalies/attacks in certain window.
In Table 6 there are results achieved for CAIDA test base. Traces consist of DDoS attacks and every trace represents 1 hour of the network traffic.

5. Comparison of the matching pursuit with standard DWT using 15 traffic parameters

The basic idea of wavelet transform is to decompose the input signal into family of some specific functions that are called wavelets. Wavelets are functions that are generated through a process of dilations and translations of one single function, which is usually called "mother wavelet". The concept of wavelet transform was defined in (Grossman & Morlet , 1985). In case of IDS system signal represents parameters of network traffic (such as number of packets per second). Anomaly Detection System based on DWT and using 15 network features was presented in (Wei et al., 2009).
In Table 7 comparison of our anomaly detection methodology to state-of-the-art DWT based (Wei et al., 2009) signal processing ADS was presented. Both solutions were tested with the use of the same DARPA (DARPA, 2000) test traces. DARPA benchmark traces consist of attacks which belong to every layer of TCP/IP protocols stack.
In Table 7 the results for W5D1 (Week 5 Day 1) testday are reported. We used all 15 traffic parameters presented in Table 1 during systems testing.
Detection rate and false positive rate achieved by our methodology based on Matching Pursuit is better than in DWT based system presented in (Wei et al., 2009).
Detection rate is changing depending on the particular traffic feature. To recognize 100% of anomalies for DARPA testbed we have to use 1 to max 4 traffic features.
We also significantly reduced false positive parameter in comparison to DWT-based ADS (Wei et al., 2009). It is very important parameter in ADS systems, since the number of false positives can not be overwhelming.

6. Conclusions

In the article our developments in feature extraction for Anomaly Detection Systems are presented. The major contributions of our work is a novel algorithm for detecting anomalies

Traffic Feature	MP-MP DR[%]	MP-MP FP[%]	DWT DR[%]	DWT FP[%]
ICMP flows/minute	68.49	20.54	14.00	79.33
ICMP in bytes/minute	79.45	27.39	83.33	416.00
ICMP out bytes/minute	73.97	32.87	83.33	416.00
ICMP in frames/minute	78.08	27.39	32.00	112.00
ICMP out frames/minute	72.60	30.13	32.00	112.00
TCP flows/minute	89.04	34.24	26.67	74.67
TCP in bytes/minute	47.94	32.87	8.67	23.33
TCP out bytes/minute	80.82	27.39	8.67	23.33
TCP in frames/minute	36.98	26.02	2.00	36.00
TCP out frames/minute	38.35	27.39	2.00	36.00
UDP flows/minute	89.04	41.09	10.00	74.67
UDP in bytes/minute	98.63	41.09	11.33	66.67
UDP out bytes/minute	100.00	46.57	11.33	66.67
UDP in frames/minute	98.63	39.72	12.67	66.67
UDP out frames/minute	100.00	46.57	12.67	66.67

Table 7. Proposed MP-MP ADS in comparision to DWT based ADS (Wei et al., 2009). Both solutions were tested with the use of DARPA (DARPA, 2000) testbed (results in table are for Week5 Day1 testday; DR-Detection Rate [%], FP-False Positive [%])

based on signal decomposition. In the classification/decision module we proposed to use original matching pursuit features such as mean projection. As to our best knowledge, we are the first to propose anomaly detection system based on Matching Pursuit.
We tested and evaluated the proposed approach and showed that experimental results proved the effectiveness of the proposed method. We also provided the comparison of the Matching Pursuit methods to DWT-based anomaly detection and reported better results in terms of detection rate and false positives.
Our developments can be used in many deployments and applications, for instance for military network security enhancement or critical infrastructures information systems security.

Test Days	W5D1	W5D2	W5D3	W5D4	W5D5
DR [%] for all attack instances - DWT ((Wei et al., 2009))	94.67	66.1	49.52	74.33	26.7
DR [%] for all attack instances - MPMP (Matching Pursuit Mean Projection)	100	100	100	100	100
DR [%] attack types (DoS, U2R,R2L,PROBE) - DWT ((Wei et al., 2009))	100	75	71.43	88.89	74.1
DR [%] attack types (DoS, U2R,R2L,PROBE) - MPMP (Matching Pursuit Mean Projection)	100	100	100	100	100

Table 8. Cumulative DR - detection rate takes into consideration attacks recognized by all traffic features presented in Table 1. Results were compared to results achieved for fifth test week (Week5 Day1-5) of (DARPA, 2000) traces.

7. Acknowledgment

This work partially supported by Polish Ministry of Science and Higher Education funds allocated for the years 2010-2012 (Research Project number OR00012511).

8. References

Coppolino, L.; D'Antonio, L.; Esposito, M.; Romano L. (2009). Exploiting diversity and correlation to improve the performance of intrusion detection systems, *Proc. of IFIP/IEEE International Conference on Network and Service*, 2009.

Esposito, M.; Mazzariello, C.; Oliviero, F.; Romano, S.P.; Sansone, C. (2005). Evaluating Pattern Recognition Techniques in Intrusion Detection Systems, *PRIS*, pp. 144-153, 2005.

Esposito, M.; Mazzariello, C.; Oliviero, F.; Romano, S.P.; Sansone, C. (2005). Real Time Detection of Novel Attacks by Means of Data Mining Techniques, *ICEIS*, pp. 120-127, 2005.

Barford, P.; Kline, J.; Plonka, D.; Ron, A. (2002). A signal analysis of network traffic anomalies, *ACM SIGCOMM Internet Measurement Workshop*,2002.

Huang, P.; Feldmann, A.; Willinger, W. (2001). A non-intrusive, wavelet-based approach to detecting network performance problems, *ACM SIGCOMM Internet Measurement Workshop*, Nov. 2001.

Li, L. & Lee, G. (2003). DDos attack detection and wavelets, *IEEE ICCCN03*, pp. 421-427, Oct. 2003.

Dainotti, A.; Pescape, A.; Ventre, G. (2006). Wavelet-based Detection of DoS Attacks. *IEEE GLOBECOM*, San Francisco USA, Nov 2006.

Neff, R.; Zakhor, A.; (2002). Matching Pursuit Video Coding - Part I: Dictionary Approximation. *IEEE Transactions on Circuits and Systems for Video Technology*, vol.12, no. 1, pp. 13-26, 2002r.

Figueras, R.M.; Vandergheynst, P.; Frossard, P. (2006). Low-Rate and Flexible Image Coding With Redundant Representations. *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 726-739, march 2006r.

Shaopeng, S.; Junxun, Y.; Yongcong, Y.; Raed, A.M. (2002). A low bit-rate audio coder based on modified sinusoidal model. *IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions Proceedings* , vol.1, pp. 648- 652, July 2002.

Daudet, L. (2010). Audio Sparse Decompositions in Parallel. *IEEE Signal Processing Magazine*, vol.27, no.2, pp.90-96, March 2010.

Troop J.A. (2004). Greed is Good: Algorithmic Results for Sparse Approximation. *IEEE Transactions on Information Theory*, vol. 50, no. 10, October 2004.

Gribonval R. (2001). Fast Matching Pursuit with a Multiscale Dictionary of Gaussian Chirps. *IEEE Transactions on Signal Processing*, vol. 49, no. 5, May 2001.

Cheng, C.M.; Kung, H.T.; Tan, K.S. (2002). Use of spectral analysis in defense against DoS attacks, *IEEE GLOBECOM 2002*, pp. 2143-2148.

DAPRA, (2000). DARPA traces, *http://www.ll.mit.edu/mission/communications/ist/corpora/* .

MAWI, (2005). MAWI traces, *WIDE Project: MAWI Working Group Traffic Archive at tracer.csl.sony.co.jp/mawi/* .

CAIDA, (2004). CAIDA traces, *The CAIDA Dataset on the Witty Worm - March 19-24, Colleen Shanon and David Moore, www.caida.org/passive/witty.*

UNINA, (2009). UNINA traces, *http://www.grid.unina.it/Traffic/Traces/ttraces.php.*

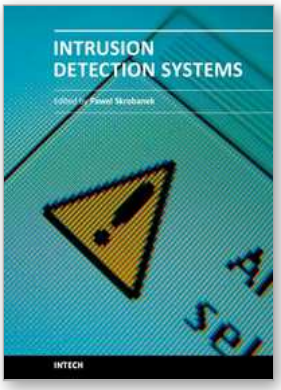
Wei, L.; Ghorbani A. (2009). Network Anomaly Detection Based on Wavelet Analysis, *EURASIP Journal on Advances in Signal Processing*, Volume 2009, Article ID 837601, 16 pages.

Mallat, S.; Zhang, (1993).Matching Pursuit with time-frequency dictionaries., *IEEE Transactions on Signal Processing*, vol. 41, no 12, pp. 3397-3415, Dec 1993.

Jost, P.; Vandergheynst P.; Frossard, P. (2005). Tree-Based Pursuit: Algorithm and Properties., *Swiss Federal Institute of Technology Lausanne (EPFL), Signal Processing Institute Technical Report*, ,TR-ITS-2005.013, May 17th, 2005.

Grossman, A.; Morlet, J. (1985). Decompositions of Functions into Wavelets of Constant Shape, and Related Transforms. *Mathematics and Physics: Lectures an Recent Results*, L. Streit, 1985.

Saganowski, Ł.; Choraś, M.; Renk, R.; Hołubowicz, W. (2009). A Novel Signal-Based Approach to Anomaly Detection in IDS Systems, and Related Transforms. *M. Kolehmainen et al. (Eds.): ICANNGA 2009* ,LNCS 5495, pp. 527−536, Springer 2009.



Intrusion Detection Systems

Edited by Dr. Pawel Skrobanek

ISBN 978-953-307-167-1
Hard cover, 324 pages
Publisher InTech
Published online 22, March, 2011
Published in print edition March, 2011

The current structure of the chapters reflects the key aspects discussed in the papers but the papers themselves contain more additional interesting information: examples of a practical application and results obtained for existing networks as well as results of experiments confirming efficacy of a synergistic analysis of anomaly detection and signature detection, and application of interesting solutions, such as an analysis of the anomalies of user behaviors and many others.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Rafał Renk, Michał Choraś, Łukasz Saganowski and Witold Hołubowicz (2011). Signal Processing Methodology for Network Anomaly Detection, Intrusion Detection Systems, Dr. Pawel Skrobanek (Ed.), ISBN: 978-953-307-167-1, InTech, Available from: http://www.intechopen.com/books/intrusion-detection-systems/signal-processing-methodology-for-network-anomaly-detection

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821