Jestr

Journal of Engineering Science and Technology Review 10 (3) (2017) 123-127

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Alert Fusion of Intrusion Detection systems using Fuzzy Dempster shafer Theory

Vrushank Shah[1,*], Akshai Aggarwal[2] and Nirbhay Chaubey[3]

[1]Indus University, Ahmedabad, India.
[2]Gujarat Technological University, Chandkheda
[3]S.S. Agarwal Institute of Computer Science, Navsari

Received 17 October 2016; Accepted 13 June 2017

Abstract

The Distributed intrusion detection systems are often used to enhance the performance and reliability of inference over single intrusion detection system. The Distributed IDS system uses a evidence theory to combine the evidences from multiple sources of information to make inference about the presence of an attack. The traditional evidence theory accounts for handling the uncertainty due to randomness. However, in the distributed IDS the inference provided by individual IDS are usually fuzzy in nature. The present work shows design of a framework for the fusion of alerts from multiple IDS involving both types of uncertainities. The modified framework is designed by incorporating fuzzy theory into the existing evidence theory and has been demonstrated against DARPA99 dataset.

Keywords: Fuzzy Logic, DARPA99, IDS, Fusion, Evidence Theory

1. Introduction

Intrusion Detection system is a system that detects abnormality in the network traffic and raises an alert [8]. The Distributed Intrusion Detection Systems are often used to enhance the performance and reliability over single Intrusion detection system [1]. The Distributed IDS system uses alert fusion method to fuse the alert raised by multiple IDS systems. However, improving the ability of intrusion detection using alert fusion is still an open issue [6]. Evidence theory is efficient method for reasoning under uncertainty and has been proved as an robust method in many realistic applications [3]. It is known that the situation arising in distributed IDS are very complex and can be characterized by not only randomness but also fuzziness. However, the evidence theory accounts only for uncertainty due to randomness. While, uncertainity due to fuzziness is still an open issue. The idea behind present work is to design a framework for the fusion of alert from multiple IDS involving both types of uncertainities. The modified framework is designed by incorporating fuzzy theory into the existing evidence theory.

The work in this paper first shows how to design the fuzzy membership function of an intrusion detection system and then shows the method to calculate the mass of IDS for a particular intrusion. The mass obtained from multiple IDS systems are then fused using the evidence theory. The proposed framework are tested against realistic network traffic data.

2. Background

Evidence theory is a mathematical theory used to combine the evidence from multiple sources of information to calculate the probability of an event. The Dempster-Shafer theory proposed by Arthur dempster in 1968 [2] and modified by Glenn Shafer in 1976 [10] is the first mathematical theory propose to combine uncertain information of sources to make an inference. The fusion rule proposed under dempster-shafer framework is called as Dempster-Shafers rule. Dempster-shafers rule has been a topic of debate for researchers working in the field of information fusion [11].

Table 1. Two Features of two Attacks DOS and R2L

Feature	**DOS Attack**	**R2L Attack**
Low NOA	$NOA_{DOSl} = 100$, $\sigma_{DOSl} = 50$	$NOA_{R2Ll} = 200$, $\sigma_{R2Ll} = 50$
High NOA	$NOA_{DOSh} = 500$, $\sigma_{DOSh} = 50$	$NOA_{R2Lh} = 500$, $\sigma_{R2Lh} = 50$

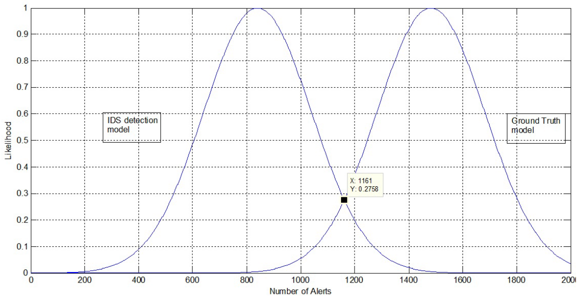

Fig. 1. Fuzzy Likelihood Model

The fusion theory is used to combine masses from n evidence sources and outputs a fused decision. For number of evidence sources n ≥ 2 let $\Theta = \{\theta_1, \theta_2, \theta_3, \ldots \theta_n\}$ be the frame of discernment for the fusion problem under consideration having n exclusive and exhaustive hypothesis. The sets of all subsets of Θ is called power-set of Θ and is denoted by 2^Θ. In shafer's framework [10] the basic belief

[*] E-mail address: vrushank26@yahoo.in

ISSN: 1791-2377
doi:10.25103/jestr.103.17

assignment (bba) is a function m from 2^{Θ}, the power set of Θ to [0,1]. The mass assignment will satisfy the property

$$m(\varphi)=0 \text{ and } \sum_{A\varepsilon 2^{\theta}} m(a)=1 \tag{1}$$

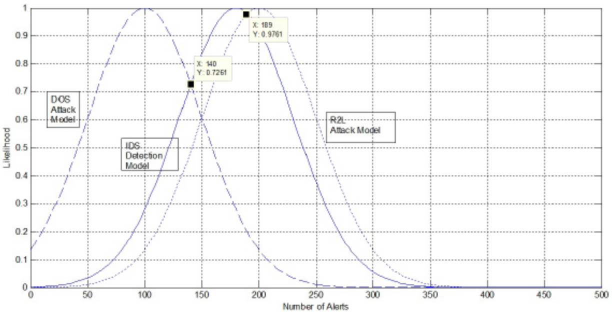

Fig. 2. Fuzzy likelihood model with low NOA

Let, $m_1(B)$ and $m_2(C)$ are two independent masses from two sources of evidence then the combined mass m(A) obtained by combining $m_1(B)$ and $m_2(C)$ through the rule,

$$m(A)=\frac{\sum\limits_{B\cap C=A} B_1 C\varepsilon 2e^{m_1(B)m_2(C)}}{1-\sum\limits_{B\cap C=\varnothing} B_1 C\varepsilon 2e^{m_1(B)m_2(C)}} \tag{2}$$

$$m(\varphi)=0 \tag{3}$$

3. Fuzzy Membership of the IDS

The Fuzzy membership of Intrusion detection system is modeled when the prior distribution of all kind of attacks in the frame of discernment are known. When IDS sniffs the incoming network traffic, alerts are generated and these alerts is used to update prior distribution to get posterior distribution. As the alerts created by different IDS may be fuzzy. Thus, the problem is how to get posterior distribution with fuzzy data which can give us the membership of IDS.

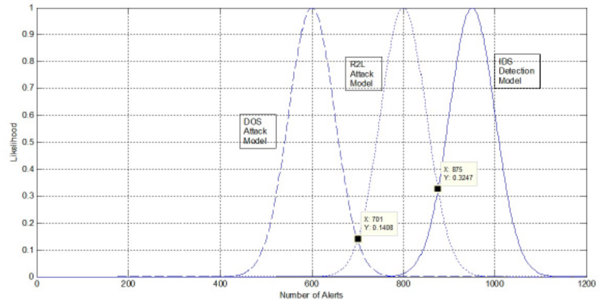

Fig 3. Fuzzy likelihood model with High NOA

Let Θ = {θ_1, θ_2, θ_3, ... θ_n }be the frame of discernment for the fusion problem under consideration having n exclusive and exhaustive list of known attack category. θ is an individual attack category from Θ. Let S_θ be the set of feature for attack θ in Θ. The likelihood function $p_i(d/\ \theta)$, where θ means target and d is detected attack by the IDS. Let the number of features of attack is M, then M likelihood functions denoted by p_1, p_2,..., p_M are available for each IDS. To simplify the problem, we assume that each attack has a single feature i.e, M=1 namely, number of alerts (NOA) and thus only one likelihood is available from IDS for each attack. To understand the method of deriving the fuzzy likelihood function, let θ be the target whose feature is NOA. This feature can be modelled using gaussian fuzzy membership function defined as,

$$A_r = e^{-\left(\frac{NOA-NOA_r}{2\sigma^2\rho}\right)} \tag{4}$$

$$A_r = e^{-\left(\frac{NOA-NOA_r}{2\sigma^2 t}\right)} \tag{5}$$

Here, NOA is the Number of alerts, NOAr is number of real alerts which are known and NOAt is the number of true alerts generated by an IDS. Ar is the gaussian model of the attack which shows the ground truth of the attack's feature that is the Number of alerts. At is the gaussian model of the IDS detecting the attack's feature. While, σt represents the accuracy of IDS's alert from estimated mean value. An attack is detected when there is highest degree of matching between the gaussian model of ground truth of the feature favouring that attack and gaussian model of IDS detecting the feature of that attack.

Table 2. Two IDS detection Scenario

Scenario	Number of Alerts
S_{low}	NOA_l = 200 , σ_l = 50
S_{high}	NOA_h = 1100, σ_h = 50

Table 3. Likelihood of various IDS Scenario

Feature	p(d/DOS)	p(d/R2L)
Low NOA	0.7261	0.9761
High NOA	0.1408	0.3247

Table 4. Snort Alert against DARPA99 Dataset

	ICMP			UDP			TCP		
Days	NOA	NOA_t	NOA_r	NOA	NOA_t	NOA_r	NOA	NOA_t	NOA_r
4th week Monday	138	44	25	0	0	0	84	27	15
4th week Tuesday	50	19	12	0	0	0	228	85	55
4th week Wednesday	218	59	37	0	0	0	314	84	54
4th week Thursday	292	28	11	0	0	0	455	43	16
4th week Friday	363	181	100	0	0	0	294	146	81
5th week Monday	577	106	64	0	0	0	182	5	3
5th week Tuesday	228	121	58	0	0	0	310	164	79

5th week Wednesday	202	32	23	0	0	0	358	58	40
5th week Thursday	1236	57	29	0	0	0	427	102	53
5th week Friday	297	57	40	0	0	0	338	64	45

Table 5. PHAD Alerts against DARPA99 Dataset

	ICMP			UDP			TCP		
Days	**NOA**	**NOA_t**	**NOA_r**	**NOA**	**NOA_t**	**NOA_r**	**NOA**	**NOA_t**	**NOA_r**
4th week Monday	398	56	27	4	1	0	105	15	7
4th week Tuesday	0	0	0	0	0	0	340	104	37
4th week Wednesday	253	37	24	12	2	1	706	104	66
4th week Thursday	0	0	0	2	0	0	828	71	26
4th week Friday	253	115	32	12	5	2	454	206	58
5th week Monday	272	104	63	14	1	0	128	6	4
5th week Tuesday	0	0	0	2	1	1	717	284	189
5th week Wednesday	0	0	0	8	1	0	796	89	49
5th week Thursday	0	0	0	15	2	1	382	157	60
5th week Friday	0	0	0	11	1	0	291	120	42

Suppose the values are as follows: NOA_r = 1482; NOA_t = 840; σ_r = 200 and σ_t = 200. Then, A_r and A_t can be modelled as shown in figure-1. The value of likehood probability $p_i(d/\theta)$ is determined from the intersection point of IDS detection model with ground truth model. The value at intersection between A_r and A_t is (1161,0.2761). Thus, the value of likelihood of $p_i(d/\theta)$ is 0.2761 which shows that there are 27.61 % chance that the attack d is detected by IDS given as attack θ in ground truth. The level of intersection between A_r and A_t determines level of consistency between IDS detection and ground truth.

Table 6. Likelihood Values of Snort and PHAD for 4th week Monday

IDS name	ICMP	UDP	TCP
Snort	0.78	0	0.75
PHAD	0.37	0.51	0.69

To understand the applicability of the proposed method to intrusion detection framework, Let us assume that we have the frame of discernment □ = {DOS,R2L} where, DOS stands for Denial of service attacks and R2L stands for remote to local attacks and also assume that there are two IDS namely signature based producing less number of alerts and anomaly based producing large number of alerts.

Table 7. Mass Values for SNORT and PHAD for 4th week Monday

IDS name	m_{ICMP}	m_{UDP}	m_{TCP}	$m_{Uncertainty}$
Snort	0.4457	0.0000	0.4285	0.1257
PHAD	0.1968	0.2712	0.3670	0.1648
Fused Mass using Dempster Shafer	0.3676	0.0000	0.4067	0.2257

The ground truth NOA feature is as shown in table-1 and IDS detecting feature is as shown in table-2. It can be observed from table-3 the likelihood values of low NOA intrusion detection systems are higher compared to high NOA intrusion detection systems. Also, it can be concluded from the results in table-3 that p(d/R2L) has the higher likelihood values compared to p(d/DOS) which shows that chances of occurance of R2L attack is higher compared to DOS attack.

4. Mass Calculation Method

In Computer Networks, the identification of real intrusion is determined based on the features observed by the IDS. The features observed depend upon whether the IDS is anomaly based or signature based. The signature based IDS is designed to compare the signatures of known attacks loaded in the database and raises an alert for abnormal packet. While, anomaly based IDS is trained with normal profile of IDS, features are extracted from incoming packet and extracted features are compared with features of normal profile. IDS is used to generate the likelihood function which describes the probability of occurence of an attack given collected fuzzy data. The likelihood function are calculated for each IDS system and are converted in to basic probability assignment (bpa) or mass value which is then fused using evidence theory. These section explains the method to calculate the mass of an attack based on its fuzzy membership values.

As discussed in section-3, let $\Theta = \{\theta_1, \theta_2, \theta_3, ... \theta_n\}$ be the frame of discernment for the fusion problem under consideration having n exclusive and exhaustive list of known attack category. θ is an individual attack category from □. Let S_θ be the set of feature for attack θ in □. The likelihood function $p_i(d/\theta)$, where θ means target and d is detected attack by the IDS. Let the number of features of attack is M, then M likelihood functions denoted by p_1, p_2,…,p_M are available for each IDS. To simplify the problem, we assume that each attack has a single feature i.e, M=1 and thus only one likelihood is available from IDS for each attack. For example, if □ = {p,q,r,s,….}, then the likelihood functions are denoted as $p_i(p/\theta)$, $p_i(q/\theta)$, $p_i(r/\theta)$, $p_i(s/\theta,)$…. For each ith IDS.

The basic probability assignment is done as follows:

- Select the maximum value in p_i, then

  $x = \max(p_i(p/\theta), p_i(q/\theta), p_i(r/\theta), p_i(s/\theta),...)$

- The likelihood function of uncertainty is defined as $pi(u/\theta)=1-x$

- The mass for attack p is calculated as follows:

$$m(p)=\frac{p_1(p/\theta)}{p_1(p/\theta)+p_1(q/\theta)+p_1(r/\theta)+p_1(s/\theta)+...+p_1(u/\theta)} \quad (6)$$

Similarly for other attacks.

$$m(p)=\frac{p_1(p/\theta)}{p_1(p/\theta)+p_1(q/\theta)+p_1(r/\theta)+p_1(s/\theta)+...+p_1(u/\theta)} \quad (7)$$

$$m(p)=\frac{p_1(r/\theta)}{p_1(p/\theta)+p_1(q/\theta)+p_1(r/\theta)+p_1(s/\theta)+...+p_1(u/\theta)} \quad (8)$$

$$m(p)=\frac{p_1(s/\theta)}{p_1(p/\theta)+p_1(q/\theta)+p_1(r/\theta)+p_1(s/\theta)+...+p_1(u/\theta)} \quad (9)$$

For uncertainty, the mass value will be

$$m(u)=\frac{p_1(u/\theta)}{p_1(p/\theta)+p_1(q/\theta)+p_1(r/\theta)+p_1(s/\theta)+...+p_1(u/\theta)} \quad (10)$$

When the mass value for all the attacks mentioned in Θ is calculated from all the IDS systems, the data can be fused based on evidence fusion rule, as described in section-3.

5. Results

The robustness of our proposed fuzzy-DS rule can be proved with the help of testing and evaluation of system in real or online enviornment. However, It is difficult and very costly to perform online evaluation for new IDS or new proposed methodology. The wide spread research in the field of IDS along with very high cost for development of these systems has led to perform online evaluation [6]. DARPA dataset is the first and a benchmarking dataset used to the research community in field of IDS [5]. DARPA has provided a number of datasets including 1998, 1999 and 2000 datasets. To evaluate our proposed system we have used DARPA99 dataset against two different IDS systems. The complete DARPA99 dataset was 5 weeks long. First week and third week data is used for training purpose. Second week data has labelled attacks. While for testing the IDS, 4th and 5th week data is used. The dataset consists of Denial of service (DOS), Remote to local (R2L), User to root (U2R), Probe and finally data attacks. We have analyzed the network traffic belonging to ICMP, TCP and UDP category against signature based IDS namely SNORT and anomaly based IDS called PHAD. The detailed explanation of various intrusions/attacks present in DARPA99 along with normal traffic is explained in detail in [5] by Kendall.

The simulation environment consists three 3rd Generation Intel Core i5 processor (1.6GHz), Operating system installed is Linux Ubuntu with 4GB RAM. One machine deployed with Signature based IDS such as SNORT [7].

Another machine deployed with anomaly detector such as PHAD [9]. Third machine acts as an attacker machine having dataset loaded and is being replayed using TCPreplay [13]. The fuzzy membership functions for SNORT and PHAD are calculated from the number of alerts generated in each of the ICMP, UDP and TCP category for each day of 4th and 5th week data. The calculated fuzzy membership values are used to find the mass value for each category and for both types of Intrusion detection systems. The mass values are then fused using Dempster-shafer theory to make an inference. Table-4 and Table-5 shows the alerts generated by SNORT and PHAD against DARPA dataset for ICMP, TCP and UDP services. Table-6 shows the likelihood values of Snort and PHAD calculated using fuzzy membership for 4th week Monday. Table-7 shows mass values calculated for ICMP, UDP and TCP traffic using equations [6-10] for SNORT AND PHAD. The mass values calculated for SNORT IDS is explained here.

Table 8. Comparison of Fuzzy Evidence Rule versus Traditional Evidence Rule

Metric	Snort	PHAD	DS Rule	Proposed Rule
True Positives	127	118	131	143
True Negatives	2715	2681	2644	5324
False Negatives	140	149	136	267
False Positives	2784	2818	2855	32
True Positive Rate	0.4757	0.4419	0.4906	0.5356
False Positive Rate	.5063	0.5125	0.5192	0.060
Positive Prediction Value	0.0437	0.0402	0.0439	0.8171
Negative Prediction Value	0.9510	0.9473	0.9511	0.9522
Accuracy	0.4929	0.4854	0.4813	0.9481

The basic probability assignment is done as follows:

x = max(p_{snort}(ICMP); p_{snort}(UDP); p_{snort}(TCP)) = 0.78

The likelihood function of uncertainty is defined as

p_{snort}(Uncertainty) = 1-x = 0.22

$$m_{snort}(ICPM)=\frac{0.78}{0.78+0+0.75+0.22}=0.4457 \quad (11)$$

$$m_{snort}(UDP)=\frac{0}{0.78+0+0.75+0.22}=0.0000 \quad (12)$$

$$m_{snort}(UDP)=\frac{0.75}{0.78+0+0.75+0.22}=0.4285 \quad (13)$$

Table-7 shows the results of fused mass of SNORT and PHAD Intrusion detection systems using dempster-shafer theory as discussed in section-2. The fused inference of two IDS shows that there is maximum chance that a TCP protocol related attack exists in the network. Table-8 shows the performance of DS rule with proposed rule in terms of various metrics against DARPA99 dataset. In DARPA99 Experiment, we preprocessed the dataset and total 5766 packets where loaded on to the network. The Frame of discernment defined for this experiment is □ = *{TCPflood, -TCPflood, θ}*.

It is observed from the results that with alert fusion of SNORT and PHAD with fuzzy DS rule, we are able to achieve 94.8 % accuracy as compared to 49.2 % obtained with SNORT as an single IDS and 48.5 % with PHAD as a single IDS. It is also evident from the result that with our proposed method there is not much significant increase in true positive rate. However, there is significant reduction in false positive rate.

6. Conclusion

The work proposed in these paper shows that the inference in the alert fusion of distributed intrusion detection system can be achieved using fuzzy dempster shafer theory which not only incorporates the uncertainty of intrusion detection system but also handles the fuzziness in the system. The work shows method to calculate fuzzy membership function of an IDS and also explains the process of mapping of fuzzy membership to the mass values which can be used as input to the dempster-shafer fusion model. The present work can be extended and apply to the modified version of dempster-shafer rules proposed such as Yager's rule proposed in [14], DSmT rule proposed in [11] and Consensus operator in [4].

References

1. Bass, T.: Intrusion detection systems and multisensor data fusion. Communications of the ACM 43(4), 99{105 (2000)
2. Dempster, A.P.: A generalization of bayesian inference. In: Classic works of the dempster-shafer theory of belief functions, pp. 73{104. Springer (2008)
3. Hu, W., Li, J., Gao, Q.: Intrusion detection engine based on dempster-shafer's theory of evidence. In: 2006 International Conference on Communications, Circuits and Systems, vol. 3, pp. 1627{1631. IEEE (2006)
4. Josang, A.: The consensus operator for combining beliefs. Arti_cial Intelligence 141(1), 157{170 (2002)
5. Kendall, K.: A database of computer attacks for the evaluation of intrusion detection systems. Tech. rep., DTIC Document (1999)
6. Kenkre, P.S., Pai, A., Colaco, L.: Real time intrusion detection and prevention system. In: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014, pp. 405{411. Springer (2015)
7. Koziol, J.: Intrusion detection with Snort. Sams Publishing (2003)
8. Liao, H.J., Lin, C.H.R., Lin, Y.C., Tung, K.Y.: Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications 36(1), 16{24 (2013)
9. Mahoney, M.V., Chan, P.K.: Phad: Packet header anomaly detection for identifying hostile network traffic (2001)
10. Shafer, G., et al.: A mathematical theory of evidence, vol. 1. Princeton university press Princeton (1976)
11. Smarandache, F., Dezert, J.: Information fusion based on new proportional conflict redistribution rules. In: 2005 7th International Conference on Information Fusion, vol. 2, pp. 8{pp. IEEE (2005)
12. Thomas, C., Sharma, V., Balakrishnan, N.: Usefulness of darpa dataset for intrusion detection system evaluation. In: SPIE Defense and Security Symposium, pp. 69,730G{69,730G. International Society for Optics and Photonics (2008)
13. Turner, A., Bing, M.: tcpreplay tool (2012)
14. Yager, R.R.: Approximate reasoning as a basis for rule-based expert systems. IEEE Transactions on Systems, Man, and Cybernetics (4), 636{643 (1984)