Research of D-S Evidence Method in Network Attack Intention Recognition

Yajing zhang[1, a], Lina Yu[2,b] and Wei Li[3,c]

[1]Hebei College of Industry and Technology ,Department of Computer Technology Shijiazhuang,China

[2] Hebei College of Industry and Technology ,Department of Computer Technology Shijiazhuang,China

[3] Caida Securities CO.,LTD ,Department of Computer Shijiazhuang,China

[a]zyjhbivt@126.com, [b]hbyln@126.com, [c]cdzqlw@126.com

Keywords: D-S evidence method, network attack, intention recognition

Abstract. When the network security threats alarm as evidence appeared, from these alarm properties were then able to predict the future attack important information, such as attack source, is the object of attack and attack port, etc. But the information is not enough to reflect the invaders attack intention. This paper puts forward a kind of attack intention recognition called D-S evidence theory method.

Introduction

D-S (Dempster - Shafer) evidence theory is put forward in the 1960s by the Dempster, and then the Shafer system perfect a set of mathematical theory, it is to bayesian probability theory further expansion. D-S evidence theory does not need to know in advance the prior probability can be information fusion. And, through the d-s evidence synthesis rules can along with the arrival of the new evidence calculated each attack the credibility of the purpose of redistribution.

The uncertainty in the real world can be divided into two kinds: one kind is objective uncertainty, also known as accidental uncertainty, such uncertainty is caused by observation system according to random way operation; Another kind is subjective uncertainty, also called cognition uncertainty, produce this kind of uncertainty reason is lack of understanding about be observation system. Compared with the classical bayesian probability theory, the d-s evidence theory is more suitable for characterization due to lack of understanding from the subjective uncertainty. And the traditional bayesian theory, compared with d-s evidence theory dealing with the human according to the observed evidence, the problem of coke yuan produced by the "trust", not "probability".

Important concepts of D-S evidence method

Frame of Discernment.

Set U said proposition X all possible value of the theory field collection, and all in the U elements within a room is incompatible, then say U for X recognition framework:

$U=\{x_1, x_2, \cdots, x_n\}$.

Mass function.

D-S evidence theory is one of the most important function, known as the basic possibility assigned function m (mass function), which are defined as follows:

Given recognition frame U, a map m on the power set 2^U of U is: $2^U \rightarrow [0,1]$ in meet the following conditions:

$m(\phi)=0, \sum_{X \subset U} m(X)=1$,

called m(X) is X's basic possibility distribution function, it presented (in a certain evidence E) the conclusion X trust degree.

If $x \subseteq U$ and m(X)>0,we called X for focus elements .

Given recognition framework U, m:$2^U \rightarrow [0,1]$ is reliability distribution on U, defined the following function:

$$Bel : 2^U \rightarrow [0,1],$$
$$Bel(X) = \sum_{B \subseteq X} m(B) \ (\forall X \subset U) \tag{1}$$

We called bel is Belief function on U, it means all the possibility of X's subset sum of measure, namely to X's total trust.

D-S synthetic rules.

Set Bel1 and Bel2 are two trust functions based on the same recognition framework on U , a m and m are respectively corresponding basic reliable function, focus elements are $A_1,\ldots,A_k,\ldots A_K$ and $B_1,\ldots,B_n,\ldots B_N$ .Set :

$$K = \sum_{A_k \cap B_n = \phi} m_1(A_k) m_2(B_n) \tag{2}$$

After synthesis basic possibility distribution function is:

$$m(C) = \frac{\sum_{A_k \cap B_n = C} m_1(A_k) m_2(B_n)}{1-K}, \forall C \subset U,\ C \neq \phi \tag{3}$$

In equation (3), if K≠1, m determines a basic possibility distribution; If K = 1, then we think m_1 and m_2 existing contradiction, it's unable to combine the basic possibility distribution.

Through the type of evidence for two two comprehensive, get comprehensive possibility distribution value. For more evidence calculation can also according to the D-S evidence combination method recursion, and more evidence synthesis and order have nothing to do.

Attack intention recognition

In extracting the future attacks on the basis of the evidence, first define different types of evidence on the various attack intention the credibility of the distribution, and then, according to the appearance of evidence to attack intention credibility distribution using d-s evidence theory of synthesis, and ultimately get more evidence synthesis attack intention after the credibility of the distribution, the credibility of the highest attack intention as intention recognition output.

Intention recognition of d-s evidence theory framework as shown in figure 1 shows, the main points of the following phases:

(1) Structure recognition framework, which determine the problem may be involved in the attack intention set $U = \{x_1, x_2, \cdots, x_n\}$;

(2)Generation evidence E_i, (i=1,2,...,k);

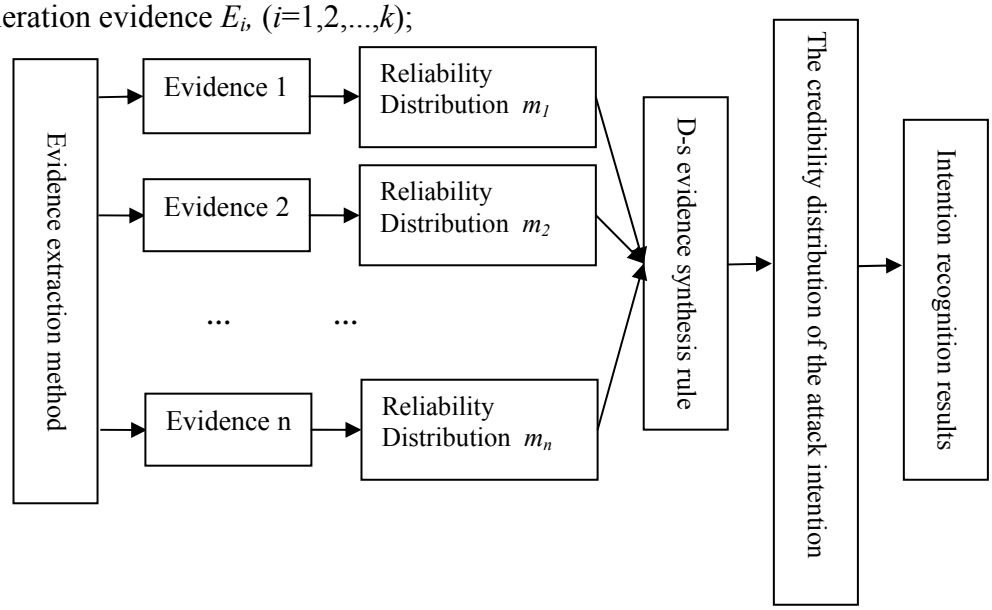

Figure 1. Attack intention recognition of D-S evidence theory inference method process

(3)Determine each evidence's basic probability distribution m_i,(i=1,2,...,k), $m_i = \{m_i(x_1), m_i(x_2), \cdots, m_i(x_n)\}$, $m_i(x_n)$ stands for the trust degree E_i to x_n.

(4)Using D-S evidence synthesis rules to compute two evidence body under the joint action of recognition framework each attack intention credibility;

(5)According to the arrival of the new evidence, the d-s evidence synthesis rules of credibility redistribution;

(6)Through the final attack intention to determine the credibility of the distribution of the output of the intention recognition.

Experiment

We take the experiment according to the DARPA2000 LLDoS1.0 attack scene on 172.16.115.20 attack.

Set against the intention may have: information steal (A_1), information manipulation (A_2), super user permissions (A_3), DDoS attack (A_4), the structure of the attack intention recognition framework is $U = \{A_1, A_2, A_3, A_4, \theta\}$,$\theta$ stands for the unsure attack intention.

From 172.16.115.20 warning we can extract in the future evidence of attack as follows:

Evidence 1 -- 22:51, IP scanning (alarm 384 - ICMP PING)

Evidence 2 -- 23:15, Sadmind scanning (alarm 12626585195 7)

Evidence 3 -- 23:34, Sadmind loophole attack (alarm 1911225 28)

Evidence 4 -- 23:50, using RSH (remote Shell command) installation program

According to expert analysis, the four kinds of evidence to attack the intentions of the basic probability assignment set respectively as follows:

$m_1 = m_1(A_1, A_2, A_3, A_4, \theta) = (0.25, 0.25, 0.2, 0.2, 0.1)$

$m_2 = m_2(A_1, A_2, A_3, A_4, \theta) = (0.2, 0.15, 0.3, 0.3, 0.05)$

$m_3 = m_3(A_1, A_2, A_3, A_4, \theta) = (0.1, 0.1, 0.4, 0.3, 0.1)$

$m_4 = m_4(A_1, A_2, A_3, A_4, \theta) = (0.1, 0.1, 0.2, 0.5, 0.1)$

First of all, calculating the credibility of the attack intention distribution in evidence 1 and 2 evidence, which showed in Table 1.

Table 1. Attack intention synthesis under Evidence 1 and Evidence2

Attack intention	A_1(0.25)	A_2(0.25)	A_3(0.2)	A_4(0.2)	θ(0.1)
A_1(0.2)	A_1(0.05)	ϕ(0.05)	ϕ(0.04)	ϕ(0.04)	A_1(0.02)
A_2(0.15)	ϕ(0.0375)	A_2(0.0375)	ϕ(0.03)	ϕ(0.03)	A_2 (0.015)
A_3(0.3)	ϕ(0.075)	ϕ(0.075)	A_3(0.06)	ϕ(0.06)	A_3(0.03)
A_4(0.3)	ϕ(0.075)	ϕ(0.075)	ϕ(0.06)	A_4(0.06)	A_4(0.03)
A_5(0.05)	A_1(0.0125)	A_2(0.0125)	A_3(0.01)	A_4((0.01)	θ(0.005)

According to D-S evidence synthesis rules, after synthesis of evidence 1 and evidence 2 the credibility of the distribution is:

$m_{(1,2)} = m_{(1,2)}(A_1, A_2, A_3, A_4, \theta) = (0.23, 0.18, 0.28, 0.28, 0.03)$

Then on the basis of evidence 3 and 4 evidence fusion, get the credibility of the distribution :

$m_{(1,2,3)} = m_{(1,2,3)}(A_1, A_2, A_3, A_4, \theta) = (0.13, 0.11, 0.42, 0.33, 0.01)$

$m_{(1,2,3,4)} = m_{(1,2,3,4)}(A_1, A_2, A_3, A_4, \theta) = (0.08, 0.06, 0.27, 0.58, 0.01)$

Conclusion

Based on D-S theory, in the moment evidence 3 happened, speculated that the most likely attack intention for super user permissions is A_3, credibility is 0.42, at this time the attacker makes Sadmind buffer loophole attack, attack intention is user permissions; in the moment evidence 4 happened, the most possible attack intention is DDoS attack A_4, credibility is 0.58, the evidence proved that the attacker installed the malicious programs, likely to attack other host. The result of reasoning with facts coincide.

References

[1] Steinberg, A., Bowman, C., White, F. Revisions to the JDL Data Fusion Model. SPIE 3719. 1999, 430-441.
[2] Bass, T. Intrusion Detection System and Multisensor Data Fusion: Creating Cyberspace Situational Awareness. Communications of The ACM. 2000, 43(4):90-105.
[3] Julisch, K. Mining Alarm Clusters to Improve Alarm Handling Efficiency. in: Proceedings 17th Annual Computer Security Applications Conference. New Orleans, LA, USA: IEEE Comput. Soc, 2001. 12.
[4] Goodall, J. R., Lutters, W. G., Komlodi, A. The Work of Intrusion Detection:Rethinking the Role of Security Analyst. in: Proceedings of the Tenth Americas Conference on Information Systems. New York: 2004.
[5] Ammann, P., Wijesekera, D., Kaushik, S. Scalable, Graph-based Network Vulnerability Analysis. in: Proceedings of 9th ACM Conference on Computer and Communications Security (CCS 2002): 2002.
[6] ARDA. Advanced Research and Development Activity. Exploratory Program Call for Proposals 2006. USA. 2005.
[7] Porras, P. A., Fong, M. W., Valdes, A. A Mission-Impact-based Approach to INFOSEC Alarm Correlation. in: Poceedings of the 5th International Symposium on Recent Advances in Intrusion Detection(RAID2002). Zurich, Switzerland: Springer-Verlag, 2002. 95.
[8] Valdes, A., Skinner, K. Probabilistic Alert Correlation. in: Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection (RAID2001): Springer-Verlag, 2001. 54~68.
[9] Ning, P., Cui, Y., Reeves, S., Xu, D. Techniques and Tools for Analyzing Intrusion Alerts. ACM Transactions on Information and System Security. 2004, 7(2): 274.
[10]Yanmei Li, Jingmin Wang. Risk Assessment Model of Smart Grid Project Based on Variable P recision RS and LSSVM. Advances in Information Sciences and Service Sciences (AISS). Vol.3, N o.10, 2011,11 :375～383.

Published by Atlantis Press, Paris, France.