An Alert Fusion Method Based on Grey Relation and Attribute Similarity Correlation

http://dx.doi.org/10.3991/ijoe.v12i08.5958

Wei Liang[1], Zuo Chen[2], Ya Wen[2] and Weidong Xiao[1]
[1] Xiamen University of Technology, Xiamen, China
[2] Hunan University, Changsha, China

***Abstract*—Various security devices which produce a large volume of logs and alerts have been used widely. It is such a troublesome and time-consuming task for network managers to analyze and deal with the information. This paper presented an improved alerts aggregation method based on grey correlation and attribute similarity method. We used grey correlation to ascertain the importance of alert attributes in network security, and considered it as the weight of attributes. Then we combined with the attribute similarity method and calculated the overall feature similarity in order to complete alert aggregation. Experiments results showed that this method had a strict mathematical theory basis and a higher practical value, which can effectively reduce raw alerts and reduce redundancy for alert data fusion.**

***Keywords*—Grey correlation analysis; Attribute similarity; Aggregation; Hyper alerts.**

I. INTRODUCTION

With the development of computer technology, humans have a closer relationship with the network, especially in our entertainment such as study and work. At the same time, the characteristics of the network such as diversity, openness and connectivity make the network vulnerable to various attacks. Although intrusion detection system, firewall and other security instruments have been widely used, the complex information not only unable to clearly outline the network situation, but also can't make network administrators accurately understand the threat degree or grasp the network security situation for making the right decision.

Data fusion technology is applied in a large-scale network environment to collect and integrate security status data of multi-sensor heterogeneous networks, which can achieve comprehensive monitoring of large-scale networks for grasping the network situation and real-time monitoring of network security status. Methods in multi-sensor data fusion model can be divided into three levels: data layer fusion, feature layer fusion and decision layer fusion [1]. Feature layer fusion, which aims at correlation analysis and fusion on the feature information after preprocessing, is in the middle layer of data process. This paper is aimed at data aggregation and correlation in the feature layer to reduce data redundancy based on the data fusion layer idea.

In recent years, analysis of aggregation and correlation techniques of network security events, which mainly on correlation algorithm researches, have become a hot topic in the field of network security and made a lot of meaningful achievements. Scholars have done a large amount of works on correlation analysis method of event, and proposed many methods such as alert correlation based on attribute similarity [2], alert correlation based on known scenarios [3-5], alert correlation based on prerequisite and consequence relationship [6-8] etc. According to the theory research of network correlation and characteristics of network traffic, this paper proposed an improved attribute similarity method of security event correlation analysis. The basis of attribute similarity method theory is cluster, which aggregates and classifies those events that are satisfied with certain similarity degree to remove redundancy or duplication and improve network administrators' efficiency of alerts analysis.

There are some limitations while using attributes similarity correlation analysis method. On the one hand, to calculate the overall similarity of all attributes, the selection of attributes and improving calculation function are parts that need to be optimized. Commonly, selected attributes include attack type, time stamp, source IP address, destination IP address, source port and destination port. And there is a need to define different attributes similarity function for different attributes. As is known to us all, different calculation functions will get different results. On the other hand, traditional method of attribute similarity correlation does not take into account the weight of attribute. It just gives each attribute equal weight or according to expertise, which neglects the objective weights of attributes in system. However, weights according to experts' definition can't find out correlation information from alerts that have less attributes. All these have a negative effect on evaluation result.

In this paper, based on grey correlation analysis method we analyze the importance value of main factors that affect the network, and normalize the value as the weight. Using this method can obtain the objective weight. On the basis of the number of attack alerts that produced in the whole time and change of each attribute to get dynamic attribute importance value. Then we use attribute similarity to achieve the alert aggregation. The results evince that it can effectively reduce redundancy and aggregate repetitive and similar alerts which are produced by the same attack. Traditional attribute similarity methods are quite different in selecting attributes and defining calculation methods of each attribute. Combined with the existing methods, this paper optimizes the similarity function of each attribute. It turns out that our method can more effectively compress the number of alert information, and the aggregation rate is higher than traditional methods.

In Section II, some of the related works and theoretical basis in alert correlation are reviewed. The detail of the proposed correlation frame work is presented in Section

III, while its performance in alert correlation is discussed in Section IV. Finally, the conclusions and some suggestions for future work are given in Section V.

II. RELATED WORKS AND THEORIES

The basic idea of attribute similarity method is checking attributes of the alert information and calculating their degree of similarity. Then combined with the weight of each attribute, we calculate the overall similarity. Alerts similar enough will be aggregated into a super alert to reduce the number of repeat alerts and similar alerts. There have been some researches and theoretical basis on this aspect.

A. Researches of Attribute Similarity and Weight Determination Method

Researchers had made some related researches about attribute similarity and weight decision methods. Valdes [9] for the first time calculated the alert attribute similarity value on probabilistic framework. Through calculating the similarity value of some common attributes such as IP address, IP port, and time attribute, and giving different weight for each attribute to compute the overall similarity. Finally, the paper made a conclusion whether the two alerts can be aggregated or correlated. Although the article presented a framework of alert aggregation algorithm, it did not specifically compare attribute similarity and discuss weight assignment. The algorithm proposed in paper [10] contained attribute similarity calculation. However, the algorithm considered only the attributes of exact match, and did not consider time attribute which made results inaccurate. A method based on fuzzy comprehensive evaluation was proposed in [11], which was based on Valde's framework and just used fuzzy matrix in the final judgment. And the feature extracted from the event was also as one of the attributes to be compared. The experimental results may be better, but the attributes calculated in the experimental model included only the time and source IP.

Thus, when using the attribute similarity method, the definition of attribute weight is the key point. At present there are many ways to determine index weight at home and abroad, mainly divided into three categories: one is subjective weighting method, such as AHP (Analytic Hierarchy Process), Delphi, and Fuzzy Analysis Method. But the main drawback of the subjective weighting method is too subjective and arbitrary. Different weights will be obtained if we depend on different experts, and the results are also vulnerable to the impact of decision makers who lack of enough knowledge. One is objective weighting method, such as the Maximum Entropy Method, Principal Component Analysis (PCA), etc. The use of objective weighting method for solving the index weight is usually based on more complete mathematical theories and methods. Another is subjective and objective comprehensive method, such as Compromise Coefficient Comprehensive Weighting Method, Frank-Wolfe. Mathematics theory foundation of subjective and objective comprehensive method is relatively perfect and it also has got some preliminary research results. But the complexity of the algorithm is generally higher, which affects its application to a certain extent.

Among these objective weighting methods, grey correlation analysis method has smaller error, high reliability characteristics and is easy to compute. A lot of experiments show that the result of grey correlation method is very close to practical experience. Qu [12] used grey correlation to determine the major factors which reflect network security events. Li [13], combining with AHP, applied an improved grey correlation method to the network security situation assessment to determine index weight.

B. Theory of Grey Relation Analysis

Grey correlation analysis method is on the basis of the similar degree of sequence curve geometry shape to judge whether the link between the two sequence curves are close. The closer the curve is, the greater the correlation between corresponding sequence is, vice versa. We determine the evaluation index system according to the purpose of evaluation. Firstly, we collect evaluation data and determine the reference sequence $X_0 = \{x_0(1), x_0(2), \mathrm{L}\ x_0(n)\}$. The reference sequence should be an ideal comparison standard, which is chosen that can best reflect the system characteristics. It can be the optimal value (or the worst value) of the index to form the reference sequence and we can also choose other reference value according to the evaluation purpose. Then the comparative sequence is defined as:

$X_i' = \{x_i(1)', x_i(2)', \mathrm{L}\ x_i(n)'\}$, where *i=1, 2, m*, *m* is the index number.

The absolute difference of the corresponding element of the reference sequence and the comparative sequence can be described as $\Delta(j) = |x_0(j) - x_i(j)|$,where $i = 1, 2\mathrm{L}\ m$; $j = 1, 2\mathrm{L}\ n$.

The correlation coefficient of $X_0\ (j)$ and $X_i\ (j)$ is as the following equation: [14]

$$r_{0i} = \frac{\min\limits_i \min\limits_j \Delta_i(j) + \rho * \max\limits_i \max\limits_j \Delta_i(j)}{\Delta_i(j) + \max\limits_i \max\limits_j \Delta_i(j)} \quad (1)$$

In equation (1), ρ is the distinguish coefficient, which is among (0, 1), and the smaller the ρ is, the stronger the ability of correlation coefficient to distinguish. Usually ρ is 0.5 or given based on the practical situation. $\min\limits_i \min\limits_j \Delta_i(j)$ is two levels of minimum differential, and $\max\limits_i \max\limits_j \Delta_i(j)$ is two levels of maximum differential.

The correlation value can be calculated after getting the correlation coefficient. Correlation coefficient is a correlation degree value between the reference sequence and the comparative sequence at each moment, so it is more than one number. But the decentralized information is not convenient to compare the overall correlation. So it is necessary to integrate the correlation coefficient of every moment into a value that is to calculate its average as the quantitative representation of correlation degree between the reference sequence and the comparative sequence.

$$\mathrm{r}(X_0, X_1) = r_1 = \frac{1}{n} | \sum_j^n r_{0i}(j) \quad (2)$$

Value of grey correlation degree obtained by grey correlation model is the sort of importance of the factors affecting in the system. The normalized correlation value can be weight of affecting factors. Weight of traditional attribute similarity method is mainly depending on expert experience. The weight obtained by grey relation degree reflects the importance of attributes, and is more objective.

III. ALERT AGGREGATION METHOD

The process of alert aggregation method presented in this paper contains several parts. Firstly, we preprocess the alerts for deleting some invalid information and extracting information of the key attributes. Secondly, we obtain the weight according to the received alert information. Finally, the similarity function is used to complete aggregation. The process diagram is as follows:

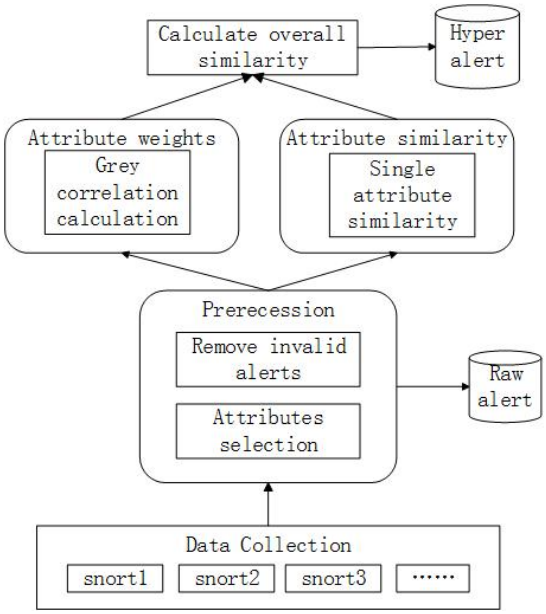

Figure 1. The proposed alert aggregation framework

A. Grey Correlation to Determine the Attribute Weight

In order to eliminate the effect of different orders of magnitude and facilitate the calculation and comparative analysis, we first need dimensionless processing. Dimensionless method includes average processing, preliminary processing, centralized processing, etc. This paper refers to [13] and the process is as follows:

The original data matrix: $\begin{Bmatrix} x_0(1)' & x_0(2)' & \mathrm{L} & x_0(n)' \\ x_1(1)' & x_1(2)' & \mathrm{L} & x_1(n)' \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ x_m(1)' & x_m(2)' & \mathrm{L} & x_m(n)' \end{Bmatrix}$

1) Range processing.

$$\mathrm{x}_i(j) = \frac{\mathrm{x}_i(j)' - p}{Q - P} \quad i = 1,2\mathrm{L}\ m\ ;\ j = 1,2\mathrm{L}\ n \tag{3}$$

$$p = min\{x_i(1) \quad x_i(1) \quad \mathrm{L} \quad x_i(n)\}$$

$$Q = max\{x_i(1) \quad x_i(1) \quad \mathrm{L} \quad x_i(n)\}$$

The result of dimensionless matrix is as follows:

$$\begin{pmatrix} X_0 \\ X_2 \\ \mathrm{M} \\ X_m \end{pmatrix} = \begin{Bmatrix} x_0(1) & x_0(2) & \mathrm{L} & x_0(n) \\ x_1(1) & x_1(2) & \mathrm{L} & x_1(n) \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ x_m(1) & x_m(2) & \mathrm{L} & x_m(n) \end{Bmatrix}$$

2) Correlation value calculation.

Use grey correlation equation (1) and equation (2) to calculate the correlation value.

$$r = \{r_1, \quad r_2, \quad \mathrm{L}\ , \quad r_m\} \tag{4}$$

3) Index weight vector.

Based on the result of grey correlation analysis, we can obtain by normalizing $r = \{r_1, \quad r_2, \quad \mathrm{L}\ , \quad r_m\}$.

$$w_i = \frac{r_i}{\sum_{i=1}^{m} r_i}$$

$$W = (w_1, w_2, \mathrm{L}\ , w_m) \tag{5}$$

W is the index weight vector. The smaller the value is, the smaller the weight is, and the less it works.

B. Attribute similarity definition and calculation function

Long et al. [15] used IDMEF format to represent alerts produced by snort system, and defined the distance calculation method between IDMEF alert documents, then use the distance as a basis for clustering. Refer to IDMEF format and the attributes we got from the alert, five alerts attributes were selected, that is alert type, time stamp, source IP address, destination IP address, source port, destination port [16]. When using attributes similarity to implement clustering, there is a need to define each attribute similarity function. Different similarity function will be applied depending on the difference of alert information attributes. Ultimately we take the sum of values obtained by each similarity function with weight, and the greater the value returned, the more similar the two alerts are.

1) Alert Type Similarity Calculation

The purpose of alert aggregation is correlating multiple original alerts produced by one attack as much as possible. So these original alerts aggregated must have similar attack types. We can determine whether two types of attacks are the same directly. The return value is 1 if the types are the same, and if not the return value is 0.

$$ClassSim = \begin{cases} 1 & if \quad Alert1.class = Alert2.class \\ 0 & if \quad Alert1.class \neq Alert2.class \end{cases} \tag{6}$$

ClassSim represents type similarity. *Alerti.class* represents the attack type of *i*-th alert.

2) Time Similarity Calculation

Alert time is one of the important factors in alert correlation. Time similarity has an important influence in calculating the global similarity. In [17], a time similarity computing framework was used. Firstly, we compute the time interval $T_{interval}$ between two alerts (*Alert1* and *Alert2*).

$$T_{\mathrm{int}erval} = | Alert1.timestamp - Alert2.timestamp | \tag{7}$$

Then we compare the time interval with presupposed minimum threshold t_{min} and maximum threshold T_{max}. If time interval is less than t_{min}, the similarity value is 1. If time interval is more than $t_{max,}$ the similarity value is 0, and if time interval is between t_{min} and T_{max}, the value is calculated by the formula in Equation 8. Values of t_{min} and T_{max} are different in a variety of papers. Literature [18] thought it should be set to different values according to alert types. Literature [19] gave T_{max} -t_{min}=300s based on common experience, in [20] the values were given t_{min}=10Min and T_{max}=60Min.In this paper the threshold value is set on the

basis of [20] and practical situation that is the t_{min}=*10Min*, T_{max}=*60Min*. Timestamp similarity can be expressed as:

$$TimeSim = \begin{cases} 1 & ,if \quad T_{interval} \le t \\ \dfrac{T_{\max} - T_{interval}}{T_{\max} - t_{\min}} & ,if \quad t < T_{interval} < T \\ 0 & ,if \quad T_{interval} \ge T \end{cases} \tag{8}$$

3) Similarity Computation of IP Address

IP address is analyzed based on Classless Inter-Domain Routing (CIDR) format. It is a 32-bit binary number, which is usually divided into four 8-bit binary number (or 4 bytes), and includes the network part and host part that are distinguished from the subnet. In this paper, IP address similarity was calculated without taking into account the subnet parameter. In order to obtain the probability of two IP addresses of the same subnet, *L* is the number of equal dimension of two IP addresses from high bit to low bit continuously. The two IP addresses are completely different if *L* is 0, else they are completely the same if *L* is 1. So the formula of IP address similarity is as follows:

$$IPSim = (IP_A, IP_B) = \frac{L}{32} \tag{9}$$

4) Similarity Computation of IP Port

The attacker must know whether the port is open before detecting vulnerability of a certain port's service. So this attribute is very important to attack correlation with the same port, which can be represented with 0 or 1 whether the port is the same or not. Therefore, the similarity between source port and destination port can be obtained by the following formula:

$$PortSim = \begin{cases} 1 & if \quad X = Y \\ 0 & if \quad X \ne Y \end{cases} \tag{10}$$

X and *Y* are ports. Whether alerts can be aggregated or correlated depends on the overall similarity.

At the time of calculating the overall similarity, we set a weight *W* and a minimum similarity expectation *H* for each attribute. Weight is used to measure the importance of the property when calculating overall similarity and minimum similarity expectation is used to control the overall similarity between alerts. In this paper, the weight of each attribute is gotten by grey relation analysis. The overall similarity *TotalSim* between *Alert1and Alert2* is:

$$TotalSim = \sum_{i}^{n} w_i * Sim(Alert1_i, Alert2_i) \tag{11}$$

Where *i* is alert attribute index, *n* is the index number, and w_i is the weight of the *i*-th attribute of the alert. *Sim* $(Alert1_i, Alert2_i)$ is the similarity value between *Alert1*and *Alert2*.

IV. EXPERIMENTAL ANALYSIS

A. Attribute Weight Calculation

We replay five days' data of first week, and divide the alerts in appropriate time slicing. In this paper it is divided into five periods (T_1 to T_5 in table 1) based on days as the unit. According to the results of ACID and the statistics of alert information from MySQL database, we form behavioral sequence of related factors (or Reference Sequence) X_i and behavior sequence of system features (or compare sequence) X_0. We will get the data in table 1 after dimensionless and normalization processing on the original data.

Calculation of grey correlation between X_0 and each X_i to get the correlation coefficient matrix is showed in table 2.

Correlation degree is obtained by grey relation analysis finally, *R*= {0.8189, 0.8786, 0.5870, 0.7343, 0.8045, 0.6690}, Normalize *R* to get the weight of each index, *W*= {0.1823,0.1956,0.1307,0.1635,0.1791,0.1489}.

B. Analysis of Alert Aggregation Effect

In order to measure the effect of the alert aggregation, the alert aggregation rate is defined as evaluation standard in experimental analysis. Assuming that the number of original alert is *N*, and was changed into *n* after aggregation. So the alert aggregation rate can be expressed as:

$$\delta = (N - n) / N$$

The rate δ is used to reflect the efficiency of aggregation algorithm for reducing repetition and redundancy and providing higher quality data for the next data fusion layer.

In the experiment, we use the weight combining with the attribute similarity algorithm, and choose 3209 alerts produced in the first day of the second week as experimental data. To compare clustering effect, we train different expectation value *H*. Clustering results are showed in the table 3.

As shown in table 3, when *H* takes 0.8 to 1, the similarity of the aggregated alerts is very high which can effectively remove repetitive alerts and eliminate redundant alerts. When *H* is more than 0.6, it can combine the alerts of high similarity into the same class. The results of processing the original alerts collected from the Intranet (table 4) and Internet (table 5) on the second week are shown in table 4.

TABLE I.
DATA AFTER DIMENSIONLESS

	x_0	x_1	x_2	x_3	x_4	x_5	x_6
T1	0.8814	0.6667	1.0000	0.2000	1.0000	0.6702	0.4333
T2	0.9237	0.6667	0.8864	1.0000	0.2941	0.6702	1.0000
T3	1.0000	1.0000	0.9205	0.0000	0.8824	1.0000	0.0000
T4	0.0000	0.0000	0.0000	0.1000	0.0000	0.0000	0.0667
T5	0.8475	0.6667	0.7159	0.3000	0.5294	0.5851	0.6667

TABLE II.
INCIDENCE COEFFICIENT MATRIX

	r_1	r_2	r_3	r_4	r_5	r_6
T1	0.6996	0.8082	0.4232	0.8082	0.7031	0.5274
T2	0.6604	0.9305	0.8676	0.4426	0.6636	0.8676
T3	1.0000	0.8627	0.3333	0.8095	1.0000	0.3333
T4	1.0000	1.0000	0.8333	1.0000	1.0000	0.8824
T5	0.7344	0.7917	0.4773	0.6112	0.6559	0.7344
correlation degree	0.8189	0.8786	0.5870	0.7343	0.8045	0.6690

TABLE III.
CHANGE OF ALERTS AGGREGATION WITH DIFFERENT *H*

expectation value H	1	0.9	0.8	0.7	0.6	0.5
Aggregated	2388	2385	2020	1967	96	25
Aggregation rate (%)	25.58	25.68	37.05	38.7	97.01	99.22

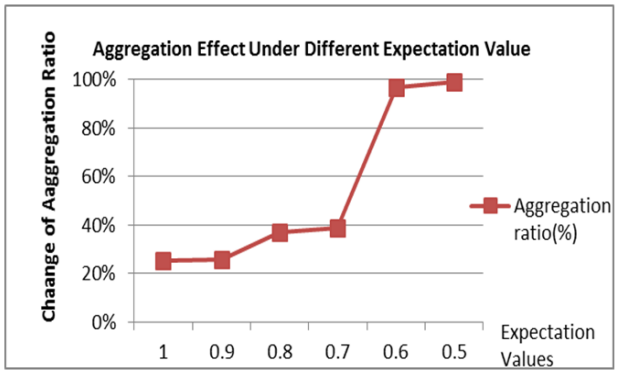

Figure 2. Aggregation Rate under Different Expectation Values

TABLE IV.
CHANGE OF ALERTS BEFORE AND AFTER AGGREGATION IN INTRANET WHEN *H*=0.8

	Day1	Day 2	Day 3	Day 4	Day 5
Raw Alerts	3209	4943	4468	8764	5718
Aggregated	2020	3095	3187	7066	5085
Aggregation rate (%)	37.05	21	28.67	19.37	11.07

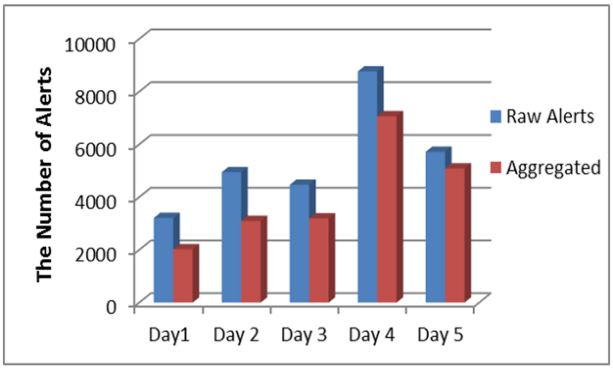

Figure 3. Changes of Alert Number in Intranet

TABLE V.
CHANGE OF ALERTS BEFORE AND AFTER AGGREGATION IN INTERNET WHEN *H*=0.8

	Day1	Day 2	Day 3	Day 4	Day 5
raw alerts	3422	7165	5015	7022	8976
aggregated	1921	5636	3841	5406	7944
Aggregation rate (%)	43.86	21.33	23.41	23.01	11.49

We can know from the data chart in Figure 4, there are a large number of repeat and redundant alerts collected by snort. By setting a high expectation value can effectively remove the redundancy and provide accurate analytical data for the next process of data fusion. And the changed expectation value can combine alerts of high similarity into a class for further analysis in the next step, and create a super alert database[21-22]. Thus we set expectation as 0.6 and observe changes of alert number after aggregated showed in the following table 6.

Table 6 shows that when $H = 0.6$, although the numbers of original alerts are quite different, the difference in the numbers after aggregation is very small. The reason is that that we use the snort system with the same filtering rule to detect alerts, so the results of the aggregation are with little difference.

The best aggregation effectiveness will be gotten when *H* is between 0.5 and 0.6 while the aggregated similarity degree is not very high. In our paper, *H* is 0.7. It can not only eliminate duplicate alerts but also aggregate similar alerts effectively. Compared our method with no-weight method and weight determined by the optimal sequence method in paper [20], we can get the result after analyzing five days of the alerts as follows in Table 7.

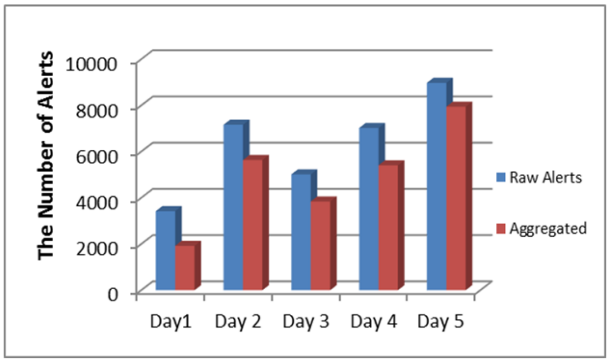

Figure 4. Changes of Alert Number in Intranet

TABLE VI.
CHANGES OF ALERTS BEFORE AND AFTER AGGREGATION WHEN H=0.6

	Day 1	Day 2	Day 3	Day 4	Day 5
Raw alerts	3209	4943	4468	8764	5718
Aggregated	96	79	90	95	32
Aggregation rate (%)	97.01	98.4	98.19	99.44	98.55

TABLE VII.
H=0.7, COMPARE OUR METHOD WITH NO-WEIGHT METHOD AND OPTIMAL SEQUENCE METHOD

time period	**raw alarm**	**our method**		**no-weight**		**optimal sequence method**	
		aggregated	***Min-similarity***	***aggregated***	***Min-similarity***	***aggregated***	***Min-similarity***
Day1	3209	1956	0.7015	1969	0.7013	102	0.7004
Day2	4943	3021	0.7024	3035	0.7023	86	0.7015
Day3	4468	3145	0.7038	3153	0.7008	101	0.7020
Day4	8764	6968	0.7019	7000	0.7004	124	0.7001
Day5	5718	5051	0.7051	5056	0.7037	56	0.7048

From table 6, the number of alerts aggregated by our method is lower than no-weight method, and the aggregation rate is relatively higher but not very consistent with the actual situation. The result number is dozens of alerts when use the method in [19], which is quite different with the former ones. Although the aggregation rate is very high, it ignored the similarity in alerts after aggregation and will need further cluster splitting and cluster merger.

The minimum similarity refers to the minimum similarity of clusters after aggregating alerts. The greater the value is, the higher the similarity of aggregation will get. From the table it can be seen that the minimum similarity in our method is higher than no-weight method, and in method [20] is just 0.7 without changing. Comparison of three methods is shown in figure 5.

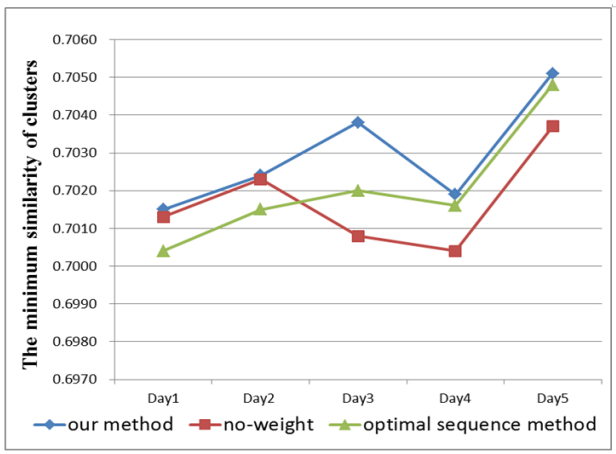

Figure 5. Comparison of Three Methods

V. CONCLUSION

The main idea in our paper is that the higher similarity of attack type between two alerts, the more likely it belongs to the same attack. The shorter the time interval between two alerts, the more likely it belongs to the same attack. The higher similarity of IP address and port, the more likely it belongs to the same attack. Traditional attribute similarity calculation does not consider the objective weight, and calculation function of each attribute is very different. This article put forward the Grey Correlation Analysis method to determine attribute index weight based on the current limitation of attribute similarity calculation methods, and combine attribute similarity method with the obtained weight to implement clustering and integration. Meanwhile, we optimize calculation function for each attribute. The experimental results show that our method is more reasonable than no-weight method and the method of literature [20], and aggregation effect is much better.

Furthermore, we need more analysis to the alerts after aggregation, and find the causal relationships between related alerts.

REFERENCES

[1] T. Bass, "Intrusion systems and multi-sensor data fusion: creating cyber space situational awareness," Communications of the ACM, vol. 43, no. 4, pp.99-105, 2000. http://dx.doi.org/10.1145/332051.332079

[2] K. Julisch, "Clustering intrusion detection alarms to support root cause analysis," ACM Transactions on Information and System Security, vol. 6, no. 4, pp.443-471, 2003. http://dx.doi.org/10.1145/950191.950192

[3] Ali. A. Ghorbani, W. Lu, and M. Tavallaee, "Network intrusion Detection and Prevention Concepts and Techniques," Springer, 2010.

[4] P. Ning, and Y. Cui, "An intrusion alert correlator based on prerequisites of intrusions," 2002.

[5] A. E. E. Taha, "Intrusion Detection Correlation in Computer Network Using Multi-Agent System,"Ain Shams Univrsity,2011.

[6] P. Ning, Y. Cui, and D. S. Reeves, "Constructing attack scenarios through correlation of intrusion alerts," In Proceedings of the 9th ACM Conference on Computer and Communications Security, pp.245-254, 2002. http://dx.doi.org/10.1145/586110.586144

[7] X. Feng, D. Wang, G. Ma, "Research on the key technology of reconstructing attack scenario based on state machine," The 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), pp. 42 -46, 2010.

[8] S. H. Ahmadinejad, S. Jalili, and M. Abadi, "A hybrid model for correlation alerts of known and unknown attack scenarios and updating attack graphs", Compute Networks, Vol.5, pp. 2221-2240, 2011. http://dx.doi.org/10.1016/j.comnet.2011.03.005

[9] A. Valde, K. Skinner, "Probabilistic alert correlation," Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection, Spring-Verlag, pp. 54-68, 2001. http://dx.doi.org/10.1007/3-540-45474-8_4

[10] H. Debar, A. Wespi, et al., "Aggregation and correlation of intrusion-detection alerts," Proceedings of 4 the International Symposium on Recent Advance in Intrusion Detection (RAID), Lecture Note in Computer Science 2212, Berlin Springer-Verlag, pp. 85-103, 2001. http://dx.doi.org/10.1007/3-540-45474-8_6

[11] C. Mu, H. Huang, S. Tian, Y. Lin, and Y. Qin, "Intrusion-Detection Alerts Processing Based on Fuzzy Comprehensive Evaluation," Journal of Computer Research and Development, vol. 42, no. 10, pp. 1679- 1685, 2005 http://dx.doi.org/10.1360/crad20051006

[12] Z. Qu, X. Wang, "Application of Grey Relation in Analyzing Network Security Events," International Colloquium on Computing, Communication, Control, and Management, pp.238-241,2009.

[13] Y. Li, K. Wang, "Index Weight Technology in Threat Evaluation Based on Improved Grey Theory," International Symposium on Intelligent Information Technology Application Workshops, pp. 307-310, 2008. http://dx.doi.org/10.1109/IITA.Workshops.2008.104

[14] J. Si, K. Wang, W. Wang, D. Man, W. Yang, "Study of Index Weight in Network Threat Evaluation Based on Improved Grey Theory," Pacific-Asia Workshop on Computational Intelligence and Industrial Application, pp.9-13, 2008. http://dx.doi.org/10.1109/paciia.2008.134

[15] J. Long, D. Schwartz, S. Stoecklin, "Distinguishing False from True Alerts in Snort by Data Mining Patterns of Alerts," Department of Computer Science, Florida State University, Tallahassee, FL, USA, 2006. http://dx.doi.org/10.1117/12.665211

[16] S. Lee, B. Chung, H. Kim, Y. Lee, C. Park, H. Yoon, "Real-time analysis of intrusion detection alerts via correlation," Computers & Security, vol. 25, pp. 169-183,2006. http://dx.doi.org/10.1016/j.cose.2005.09.004

[17] X. Peng, Y. Zhang, S. Xiao, Z. Wu, J. Cui, L. Chen and D. Xiao, "An Alert Correlation Method Based on Improved Cluster Algorithm," Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol. 12, pp. 342 - 347, 2008. http://dx.doi.org/10.1109/paciia.2008.285

[18] F. Guo, J. Ye, M. Yu, "Design and implementation of distributed IDS alert aggregation model," Application Research of Computers, vol.26, no.1, pp. 975-980, 2009.

[19] C. Long, H. Shen, J. Li, J. Ge, "An SR-ISODATA Algorithm for IDS Alerts Aggregation," IEEE International Conference Information and Automation (ICIA), pp. 92-97, 2014. http://dx.doi.org/10.1109/ICInfA.2014.6932632

[20] J. Tian, "Research on Automatic Intrusion Response Decision-making System Based on Clustering," Harbin: Harbin Industrial University,2006.

[21] http://en.wikipedia.org/wiki/Classless_Inter-Domain_Routing.

[22] http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/data/1999data.html.

AUTHORS

W. Liang is with the Department of Software Engineering, Xiamen University of Technology, Xiamen, Fujian, 361024, China (e-mail: Wliang@xmut.edu.cn).

Z. Chen is with College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, 410082, China (e-mail: chenzuo@hnu.edu.cn).

Y. Wen is with College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, 410082, China (e-mail: weny@hnu.edu.cn).

W. Xiao is with the Department of Software Engineering, Xiamen University of Technology, Xiamen, Fujian, 361024, China (e-mail: wdong@xmut.eeedu.cn).

This work was supported by National Natural Science Foundation of China (No. 61202439), and the Research Project supported by Xiamen University of Technology (YKJ15019R). Submitted 20 June 2016. Oublished as rsubmitted by the authors 27 July 2016.