 CMC, vol.59, no.1, pp.167-180, 2019

System Architecture and Key Technologies of Network Security Situation Awareness System YHSAS

Weihong Han[1], Zhihong Tian[1,*], Zizhong Huang[2], Lin Zhong[3] and Yan Jia[2]

Abstract: Network Security Situation Awareness System YHSAS acquires, understands and displays the security factors which cause changes of network situation, and predicts the future development trend of these security factors. YHSAS is developed for national backbone network, large network operators, large enterprises and other large-scale network. This paper describes its architecture and key technologies: Network Security Oriented Total Factor Information Collection and High-Dimensional Vector Space Analysis, Knowledge Representation and Management of Super Large-Scale Network Security, Multi-Level, Multi-Granularity and Multi-Dimensional Network Security Index Construction Method, Multi-Mode and Multi-Granularity Network Security Situation Prediction Technology, and so on. The performance tests show that YHSAS has high real-time performance and accuracy in security situation analysis and trend prediction. The system meets the demands of analysis and prediction for large-scale network security situation.

Keywords: Network security situation awareness, network security situation analysis and prediction, network security index, association analysis, multi-dimensional analysis.

1 Introduction

Network security incidents occur frequently in cyberspace and have a great impact, such as the “Stuxnet” incident in Iran and the “blackout incident in Ukraine” and so on. From the perspective of their respective goals and needs, various departments in cyberspace, such as government affairs, finance, e-commerce, banks and transportation, have deployed security products such as firewalls, intrusion detection and anti-virus. At present, all security products will give an alarm against network attacks. However, users still lack a macroscopic view of the overall situation of the network. Network attacks are related to asset vulnerabilities, resource consumption attacks are also related to the state of the system, and complex network attacks often run across systems and across administrative domains. Therefore, multi-channel data comprehensive analysis is needed to accurately discover network events. Therefore, it is difficult to discover large-scale and complex network attacks for locally deployed network security products.

The network security situation awareness system YHSAS introduced in this paper is oriented to the needs of space network security. Based on large data acquisition and

[1] Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou, 510006, China.

[2] Computer School, National University of Defense Technology, Changsha, 410073, China.

[3] Electrical and Computer Engineering, Rice University, 77051, USA.

[*] Corresponding Author: Zhihong Tian. Email: tianzhihong@gzhu.edu.cn.

storage management technology, it uses data analysis, mining and intelligent deduction methods to discover security incidents, assess their hazards, and predict their development. It also makes a multi-level and multi-granularity grasp of the overall network security situation from micro to macro, and gives a global view to provide decision support for cyberspace security. The main challenges faced by large-scale network security situation analysis include: 1) There are many kinds of security attacks against cyberspace. At present, there are at least 50000 species, which are evolving and emerging. How can they be judged in real time and accurately? 2) Network system security involves many factors, such as attacks, vulnerabilities, assets, networks, and so on, and the correlation is complex. How to give its threat and security situation in real time, quantitatively and understandably? 3) Network attack incidents break out instantaneously and are extremely harmful? How to predict them in advance so as to take relevant preventive measures? In view of the above challenges, this paper analyzes the system architecture and the key technologies involved in the implementation of the network security situation awareness system YHSAS. The test and practical use verify the effectiveness of the YHSAS system and its key technologies, which can provide reference value for the implementation of large-scale network security situation awareness system.

2 Related research

Around the research on the large-scale network security situation analysis and prediction system, the United States, Japan, the European Union and China have established a national network security incident monitoring system. The United States has developed the Global Early Warning Information System [GEWIS (2018)]. The National Cybersecurity Protection System, commonly known as the “Einstein Plan” [NCPS (2018)], is an important component of the Comprehensive National Cyber Security Initiative (CNCI) of the United States, which is designed and operated by the United States Department of Homeland Security to provide global, local and operational levels of network event monitoring, analysis, early warning and situation awareness. Japan has developed an Internet Scan Data Acquisition System [ISDAS (2018)]. The European Lobster [LOBSTER (2018)] belongs to the European infrastructure pioneer experimental program, which uses sensors deployed in some schools, research organizations and some telecom operators to obtain relevant information for accurate Internet communication flows monitoring, and uses deep packet inspection and deep flow inspection to identify 0day worm propagation, recognize dynamic port applications, and measure Internet services. WOMBAT (Worldwide Malicious Behavior and Attack Threat Observatory) [Dacier, Leitaand and Thonnard (2010)] is a project funded by the European Union to collect and analyze current and emerging threats (especially malicious code) in the network by means of honey pots, crawlers and external data sources. China has developed the 863-917 network security monitoring platform to analyze the network security situation of the national backbone network. The above systems are self-contained and realize monitoring and early warning for specific problem areas.

The network security situation analysis system involves many key technologies, including network security data collection[Wang, Tian and Zhang (2018); Tian, Cui and

An (2018)] network security big data storage and processing[Li, Sun, Jiang et al. (2018); Han, Tian, Huang et al. (2018); Qiu, Chai, Liu et al. (2018)], network security data sharing[Sun, Li , Su et al.(2018); Tian, Su, Shi et al. (2018); Chen, Tian, Cuiet al. (2018)], network attack detection[Cui, Zhang, Cai et al. (2018); Wang, Liu, Qiu et al. (2018); Yu, Tian, Qiu et al. (2018); Tan, Gao, Shi et al. (2018)], and so on. The research on these key technologies has also achieved a lot of results, laying a technical foundation for the construction of the network security situation analysis system.

The development of network security situation analysis technology has gone through three stages. The first phase focuses on feature-based security incident detection [Cai, Zhang and Li (2018); Liu, Meng and Wu (2013)], with the launch of the Einstein Plan of the United States in 2003 as the representative. The second stage focuses on the study of correlation analysis and threat quantitative assessment for complex security incidents [Niu and Wang (2011); Sang (2012); Zhang, Shiand and Chen (2013)], represented by Einstein Plan 2 of the United States in 2009. The third stage focuses on intelligent analysis for complex attacks, quantitative evaluation based on index system and development trend prediction research [Wang (2017); Luo (2017)], with the Einstein Plan 3 of the United States in 2013 as the representative.

3 YHSAS system architecture

The architecture of network security situation awareness system YHSAS is shown in Fig. 1. The main functions of the system include 1) Security information collection: it can collect all kinds of data, including files, packets, streams, sessions, memory information, registry information, address information, protocol information, service information, load transmission information, etc. It supports 10PB data storage scale and can integrate 187 types of network security devices. 2) Security attack detection: it can detect network scan attack, password attack, Trojan horse attack, buffer overflow attack, tamper information attack, forgery information attack, denial of service attack, e-mail attack and other conventional attacks and APT attacks, and the coverage rate is 92.3%. 3) Situation quantification calculation: It is a quantifiable security index system, which can describe the current macro overall security situation of the national Internet. 4) Security Situation Analysis: It can analyze and discover network security incidents in depth, calculate current network security situation and output multi-mode and multi-dimensional visualization. 5) Security situation prediction: It can accurately predict the security trend in a certain period of time in the future. The prediction module of calculation can predict Trojan Horse attack propagation, DDoS attack, virus situation, botnet, and APT attack, and the prediction is in good agreement.

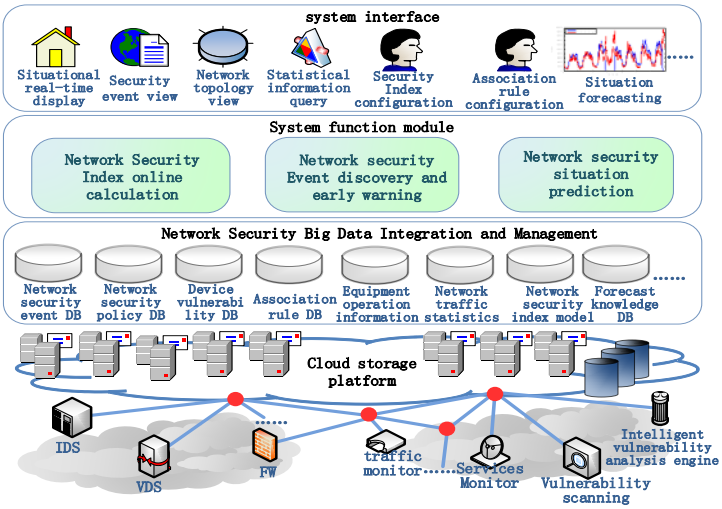


Figure 1: YHSAS system architecture

4 Key technologies of the system

The key technologies of large-scale network security situation awareness system mainly include: network security-oriented total factor information collection and high-dimensional vector space analysis technology, knowledge graph technology supporting super-large-scale network security knowledge representation and management, multi-level, multi-granularity and multi-dimensional network security index system construction method and multi-mode and multi-granularity network security event prediction technology based on adaptive prediction model.

4.1 Network security oriented total factor information collection and high-dimensional vector space analysis

Traditional security devices and products usually collect data according to their local objectives, lacking support for global, unknown and complex security events analysis, therefore, YHSAS adopts a network security-oriented total factor information acquisition model, and greatly improves the support ability for accurate and real-time detection of complex security events by extracting and analyzing the security features of the multi-dimensional and multi-level high-dimensional vector full information.

(1) Multi-level and multi-dimensional network security information total factor acquisition model

Aiming at the problem that traditional security devices lack support for global, unknown and complex security event analysis, YHSAS adopts a multi-level and multi-dimensional total factor acquisition model, as shown in Formula 1. By combining active acquisition with passive reception, fine-grained full information acquisition and extraction of 13-dimensional objects, such as files, packets, streams, sessions, protocols and network objects, is carried out from six levels: protocol layer, behavior layer, sensitive behavior layer, attack layer, broad spectrum content layer and accurate content layer, and a high-

dimensional space vector reflecting the full information of the security object is obtained.

$$R_m^n = \{C_1, C_2 \cdots C_m \Lambda V_1, V_2 \cdots V_n, m = 13, n = 6\} \to R_o = \left\{ S_1, S_2 \cdots S_o, o = (\sum_{i=1}^{m} C_i \cup c_k) * \sum_{j=1}^{n} V_j \cup v_l) \right\}$$

$$\to \lambda_{R_o} \left\{ f(\lambda_{R_o}) = \left| R_o - \lambda_{R_o} E \right| \right\} = \left\{ \lambda_1 * \lambda_2 * \cdots * \lambda_p = \left| R_o \right| \right\} \to \mathrm{Re}\, sult_q = \left\{ T_1, T_2 \cdots T_q \right\}$$

Formula. 1 Total factor collection model

(2) Analysis method of security event feature information in high-dimensional vector space

In order to solve the problem of huge computational complexity in high-dimensional vector space caused by total factor information acquisition, YHSAS proposes a feature information extraction method for network security events based on high-dimensional vector space, as shown in Formula 2. Firstly, in this method, massive data samples are clustered on streams, and then filtered according to the characteristics of the generated categories. On the one hand, the scope of the subsequent analysis can be focused to reduce the computational complexity. On the other hand, suspicious new event categories can be found by clustering. Secondly, facing the category information generated, the trained feature recognition neural network is used to identify the feature information, so as to construct the feature vector space of network security event information and lay a foundation for the next event research and judgment.

$$dis(R, R') = \frac{1}{2} \sum_{x \in X} \left(\sqrt{p_R(x)} - \sqrt{p_{R'}(x)} \right)^2, v_R = \left(\sqrt{p_R(x_1)}, \sqrt{p_R(x_2)}, \cdots \sqrt{p_R(x_{|X|})} \right)^T, \sum_{x \in X} v_k(x) = 1$$

$$\min J(\Pi, R) = \sum_{k=1}^{K} \sum_{S \in \pi_k} \sum_{x \in X} \left(\sqrt{p_S(x)} - \sqrt{c_k(x)} \right)^2 + \sum_{k=1}^{K} \lambda_k \left(1 - \sum_{x \in X} c_k(x) \right), k = \arg\min_{k'=1,2,\cdots K} \left\| v_S - c_{k'} \right\|^2$$

Formula. 2 Information analysis method of clustering and feature training in high-dimensional vector space

(3) Automatic deployment technology of network security information collection agent

In view of the super heterogeneous complexity and on-line evolution of network system, as well as the huge scale characteristics of data acquisition agents, YHSAS proposes an online insertion and extraction technique of data acquisition agent based on lightweight component technology. Firstly, this technology carries out component-based packaging for all kinds of data acquisition agents, and through the integration of component application server technology, the online insertion and extraction of agents is realized. Secondly, the target data is extracted and integrated through the regular expression-based configuration file, and the automatic generation of configuration files and the automatic conversion of data patterns are supported. This technology can efficiently integrate network security devices and data. YHSAS system supports 187 kinds of network security devices, and its performance is second-level real-time.

4.2 Knowledge representation and management of super large-scale network security

In view of the large-scale, on-line evolution and space-time correlation characteristics of network security knowledge, YHSAS adopts the super knowledge graph model of network security knowledge representation and management, breaks through the automatic/semi-automatic construction method of multi-modal knowledge graph, as well as the key technologies of online evolution and fast matching, and builds a large-scale network security knowledge graph, breaking through the accurate and real-time detection technology of network security incidents. On the standard test set, the system has 99.8% de-duplication rate, 0.01% false alarm rate and 0.2% missing alarm rate.

(1) The super knowledge graph model supporting large-scale network security

To solve the problem of huge scale, high evolution and real-time utilization of cyberspace security knowledge, YHSAS adopts the super knowledge graph knowledge representation model. On the basis of the triad of the traditional knowledge graph, attributes and rules are added, and five tuples $< Concept, Instance, Relation, Proertites, Rule >$ are proposed to solve the problem of large scale, high evolution and real-time utilization of the cyberspace security knowledge. Instances are concrete instances of concepts, each of which has its own attributes and can evolve, which is the concretization of the relationship of network security events. Based on the proposed super knowledge graph, a large-scale network security knowledge graph is implemented, which covers 93578 vulnerability information, 51300 attack methods, 151 mainstream operating systems, more than 200 mainstream applications, and 3507 malware information, breaking through the difficult problem of real-time and accurate research and judgement.

(2) Automatic construction method of super knowledge graph

Aiming at the bottleneck of knowledge acquisition of network security knowledge graph, YHSAS proposes a method of building large-scale network security knowledge graph based on multi-modal network security data by entity word recognition, relation extraction and entity link. The extended entity set E' is obtained by recognizing other candidate entity words in free text based on syntactic dependency rules and other candidate entity words in table data based on pattern reasoning. As for relation extraction, firstly, the original data fragments where knowledge in G has appeared are found out from the network security big data, and the classifier Ct of t is obtained by training with LSTM depth learning model. Then, the data fragments where E' has appeared in the original data are computed by Ct, and the possible relationship types between the entity words in E' are obtained. In terms of entity link, entity genes are constructed according to the attributes and relations of knowledge. According to the similarity between entity genes and context features, the link between entity words mentioned in the data context and known entities in G is realized. For entity words that cannot be linked, they are regarded as new entities added to G to achieve disambiguation fusion and knowledge expansion.

(3) Knowledge automatic evolution based on tensor decomposition and path sorting

Aiming at the problems of large scale, high evolution, space-time attributes and real-time utilization of cyberspace security knowledge, YHSAS adopts the evolutionary method of network security super knowledge graph. Based on tensor decomposition, an edge-and-

attribute oriented knowledge automatic evolution algorithm is proposed, that is, according to the attribute values of the nodes and adjacent nodes, the values of unknown attributes are predicted, and the possible new feasible edges between the two nodes are predicted based on the feasible path ordering method and all feasible paths between the two nodes. During the evolution of network security knowledge graph, the candidate entity recognition, entity relationship classification and entity determination techniques are used to automatically evolve and reason from network security vulnerability database and utilization methods. Based on automatic evolutionary reasoning algorithm, the data flow detected by the network security can be quickly studied and judged, which breaks through the difficult problem of real-time and accurate judgment of security events.

(4) The network security event detection based on the incremental subgraph matching of tolerance K

Based on the method of combining subgraph matching with activity pattern, YHSAS proposes an incremental fast attack subgraph matching research and judgement algorithm based on tolerance K, which realizes the detection of botnet and slow DDoS. On the experimental data set, a similarity measure method based on dynamic time warping distance is proposed, and the accuracy rate of Botnet migration detection is 92%. The results of the botnet collaborative detection method based on the botnet malicious behavior target and the time correlation analysis are as follows: when the time span is 2 months and the number of Botnet IP is more than 40, the missing alarm rate of the detection method is 0. A slow DDoS attack detection method based on traffic and service collaborative detection is proposed in order to solve the problem of difficulty and inefficiency of detecting slow DDoS attacks with strong concealment. In order to evaluate the effectiveness of this method, the network simulation experiment is carried out to test the detection system. The experimental results show that the correct detection rate is 99.7%, the missing alarm rate is 0.4%, the false alarm rate is 0.3%, and the performance of the detection system is good.

4.3 Multi-level, multi-granularity and multi-dimensional network security index construction method

There are many factors influencing large-scale network security situation analysis, and their importance is different. Therefore, the construction method of multi-level, multi-granularity and multi-dimensional network security index system, and the configurable, real-time calculation and online evolution method of its index is given to accurately describe and quantify large-scale network from macro to micro network security situation.

(1) Network security index extraction based on R clustering and factor analysis

YHSAS adopts the method of R clustering and factor analysis combined with principal-subordinate analysis to extract the network security index. Firstly, the principal component analysis is used to determine the main factors and reasonable levels that affect the network situation. Secondly, Delphi method is used to determine the number of layers of network security situation index system, and then the indicators in the same layer are classified by R clustering, so that different categories represent different aspects of network situation assessment. Finally, indexes with large factor loading in each category are selected by factor analysis method, so that a few indexes can reflect the security

situation of the whole network. The established network security situation index system only uses 16% indexes, reflecting 99% of the original information, which can effectively and objectively measure the network security situation.

(2) Multi-mode network security index system calculation model

By analyzing the characteristics of different network security factors and according to the characteristics of different network security indicators, different network security index quantification methods including extremum method, statistical standardization method, anti-cotangent function method, intermediate variable method and logarithm method are given. For the quantified network security index, the aggregation algorithm is used to aggregate the sub-indexes and calculate them into the upper-level index to form a hierarchical network security index system. The main calculation models include: weighted average method, which is intuitive and easy to understand; maximum method, which takes the maximum of one of them as the index result after aggregation by depicting the most serious degree locally; harmonic triangular norm method, which can reflect both global and local characteristics.

$$f(w_1 * x_1, \cdots w_n * x_n) = \sum_{i=1}^{n} w_i * x_i \quad \text{In which} \quad \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} = 1$$

$$f(w_1 * x_1, \cdots w_n * x_n) = Max(w_i * x_i) \quad \text{In which} \quad \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} = 1$$

$$f(w_1 * x_1, \cdots w_n * x_n) = \prod_{i=1}^{n} (w_i * x_i) / \prod_{i=1}^{n} (1 - w_i * x_i) \quad \text{In which} \quad \sum_{\mathrm{i}=1}^{n} w_{\mathrm{i}} = 1$$

Formula. 3 Multi-mode network security index system calculation model

(3) Self-evolving technology of network security index system based on deep learning

Aiming at the problem of continuous innovation and evolution of network security attack and the adaptability change of existing index system, the self-learning and self-evolution technology of network security index system based on convolutional neural network is adopted. Firstly, this technology constructs the evaluation method of index system to measure the coincidence between the index system and the reality from three aspects of correctness, stability and redundancy, so as to guide the quantitative method of deep learning algorithm on the network security index system and the feedback adjustment of the weights of aggregation operators and various parameters. The test shows that the coincidence rate between the index system based on this method and the reality is over 90%.

4.4 Multi-mode and multi-granularity network security situation prediction technology

To solve the problem that it is difficult to predict the development trend of current technology to network security, YHSAS proposes a multi-mode and multi-granularity network security situation prediction technology based on adaptive prediction model, including: network security situation prediction technology based on organic combination of multiple forecasting methods, prediction technology based on frequent episodes of

characteristic event sequences, prediction technology based on wavelet decomposition and ARMA model, and multi-dimensional entropy anomaly detection method based on improved support vector regression prediction to realize the accurate prediction of network security.

(1) Network security situation prediction framework supporting multiple prediction modes

There are many factors that affect the evolution of network security situation, and it is difficult to predict only by single prediction technology, so YHSAS adopts a network security situation prediction system architecture which combines multiple prediction methods. The related technologies of time series data prediction are applied to the field of network security. According to the characteristics and application requirements of different network security data, a reasonable prediction model is selected, and modeling is carried out by using historical security event data. Then, according to different prediction models, multiple-granularity prediction is carried out for different security data sources. For short-term prediction, the development rule of recent historical data is mainly considered for modeling prediction. For medium-term and long-term prediction, the seasonal factors and the overall long-term trend of historical security events over a long period of time are mainly considered. Tests show that the system supports short-term, medium-term, long-term and other time granularity prediction, supports Trojan horse, worm, botnet and other major network security events prediction, and the prediction effect is ideal.

(2) Prediction technology of time series data based on frequent episodes of characteristic events

Botnets, worms and other network security events with long-term propagation characteristics often have the characteristics of self similarity, so YHSAS proposes a new solution for time series data prediction: Firstly, the time series data is transformed into event sequences by segmenting the time series data and discretizing the time series sub-segment features. Then, the related concepts and methods of frequent episodes in the field of event sequence processing are introduced to extract the knowledge needed for prediction, and then the future development of time series data is predicted by using these knowledge. The specific prediction process of the proposed method can be divided into two stages: knowledge extraction and prediction: in the prediction stage, the extracted frequent episode prefix events are used to match the characteristic event sequences formed by the recent time series data, and then the selected frequent episode suffix events are used to predict the characteristic events on the future time series sub-segments. Practical application shows that the prediction based on characteristic time frequent episodes can improve the prediction accuracy of botnets and worms by about 15% in long-term multi-step prediction episodes.

(3) Multi-dimensional entropy prediction method based on support vector regression model

Aiming at the characteristics of noise and disturbance factors in large-scale network traffic data, YHSAS proposes a method of applying least squares support vector machines (LSSVM) to the prediction of entropy values in various dimensions of network traffic data, which can effectively shield the noise and disturbance factors in network

traffic data and detect the abnormal value of traffic entropy in time. The main technological breakthroughs include: Fast multi-dimensional entropy calculation. The detection accuracy is improved by correlating through entropy mutation on multiple dimensions, and for large-scale network anomaly detection, massive traffic data needs to be processed in real time. Genetic algorithm improves support vector regression. Adaptive crossover and mutation operators are used to cross and mutate all individuals in the population, which improves the searching ability of the algorithm, and only the individuals with large fitness are retained, ensuring the direction of evolution, accelerating the convergence speed and avoiding the degeneration of excellent individuals produced by crossover caused by variation. Tests show that the early detection and early warning of DDoS attacks and worm attacks that may cause abnormal traffic flow are effective.

5 Performance analysis

5.1 Accuracy of network security index

We use Blade IDS Data Set to test the accuracy of network security index construction and calculation. Blade IDS send 10,241attack packets last a day, and the attack packets have 101 different types of alerts. The alerts are divided into 15 categories, which cover the main categories of network security events. Network security experts have analyzed the network security situation of one day by the way of the manual classification, and give the index of network security, as shown by red lines in Fig. 2.The blue line shows the network security situation calculated by YHSAS. It is shown by the test that YHSAS basically reflects the trends of network security situation.

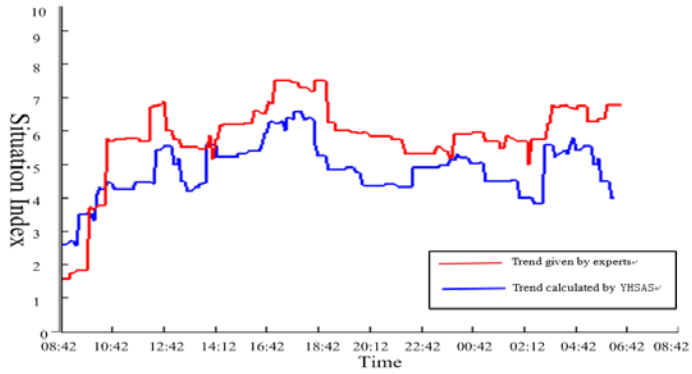


Figure 2: Accuracy of network security index

5.2 Sensitivity of network security index

By simulating a Dos attack, we verify whether the network security index can effectively reflect the changes in network security status when the network is facing a network attack. We use TFN2K [Barlow and Thrower (2000)] to initiate a DoS attack. TFN2K can perform a denial-of-service attack on the target server through ICMP flooding, SYN flooding, and smurf. We first use ICMP flooding to consume the resources of the target service, and continuously increase the intensity of the flood attack per minute to observe

the change of the network security index. The attack strength is sent from 10 attack packets per minute to Send 300 attack packets per minute, and then we gradually reduce the intensity of the ICMP flood attack until the attack stops completely. Fig. 3 shows the change of the network security index in the simulated DoS attack in the experimental environment. The point on the curve is obtained by the YHSAS system every 6s. Through the trend of the curve, it can be seen that the basic running index suddenly increases when the simulated attack strength reaches 250, because the attacked host reaches the performance bottleneck then. When the attack strength is weakened to about 150 again, The basic running index value returns to its normal level. The trend of risk index and composite index mainly increases with the increase of attack intensity, and decreases with the decrease of attack intensity. The vulnerability in the network did not change during the simulated attack, so the vulnerability index showed a stable trend throughout the test. Therefore, the network security index of YHSAS reflects the change of network security situation in real time and has good sensitivity.

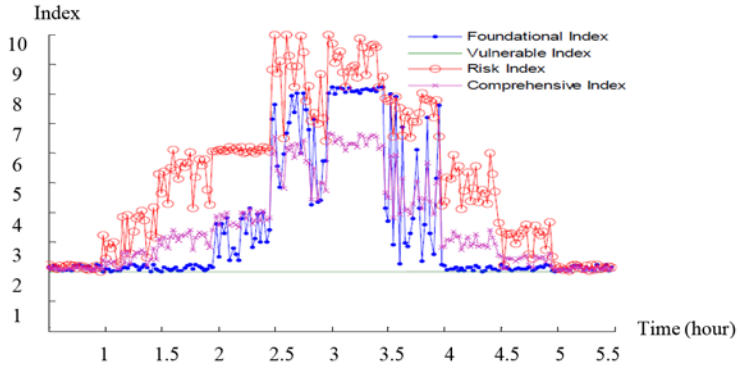
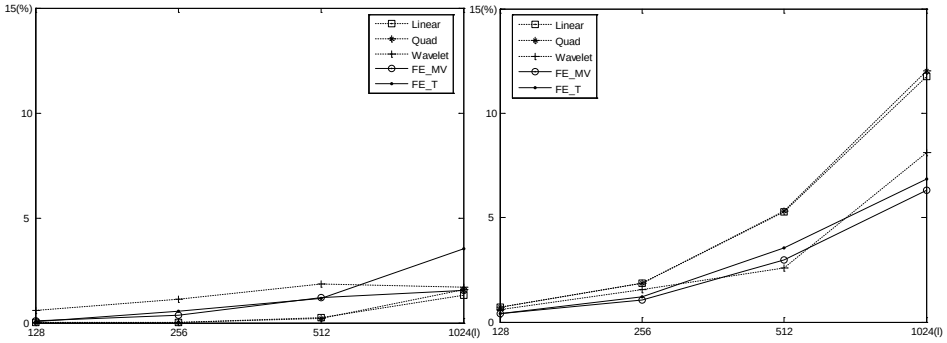


Figure 3: Sensitivity of network security index

5.3 Accuracy of network security situation prediction

We use UCI data set to compare the prediction accuracy between Short-term prediction and long-term prediction. The result is shown in Fig. 4(a) and Fig. 4(b). In the case of the single-step prediction, the regression-based method considers the continuity of data, which has a good effect in the short-term prediction. But its advantage reduces with the increases of prediction length. The errors of these prediction methods are all at a lower level. For the long-term projection, the data fitting curve cannot accurately reflect the data law after a period of time. At this time, regression-based prediction error increases rapidly and predictions based on wavelet neural network and frequent item sets are suitable for this prediction scene. Prediction accuracy has been significantly improved compared to a regression-based method. Because YH-SSAS integrates several different prediction methods, the short-term prediction accuracy is not less than 95% and the long-term forecast accuracy is not less than 90%.



(a) Short-term prediction error (b) Long-term prediction error

Figure 4: Accuracy of network security situation prediction

6 Conclusions

Large-scale Network Security Situation Awareness System (YHSAS) is developed for national backbone network, large network operators, large enterprises and other large-scale network. The system acquires, understands and displays the security factors which cause changes of network situation, and predicts the future development trend of these security factors. The key technologies of YHSAS system are deeply studied in this paper, including: network security oriented total factor information acquisition and high-dimensional vector space analysis technology, knowledge graph technology supporting super-large-scale network security knowledge representation and management, multi-level, multi-granularity and multi-dimensional network security index system construction method, multi-mode and multi-granularity network security event prediction technology based on adaptive prediction model, and so on. The performance tests show that YHSAS has high real-time performance and accuracy in security situation analysis and trend prediction, and meets the demands of analysis and prediction for large-scale network security situation.

Large-scale network security situation awareness system also faces many new challenges. In terms of accuracy of large-scale network security event prediction, there are too many factors affecting the occurrence of security incidents, new means of attack is unknown, and there are lots of operation situations of network hackers for various purposes, so it is difficult to accurately predict the occurrence and development trend of major network attack events. As one of the acknowledged worldwide problems in the field, it needs further study.

Acknowledgement: This work is funded by the National Natural Science Foundation of China under Grant U1636215 and the National key research and development plan under Grant Nos. 2018YFB0803504, 2016YFB0800303.

References

Cai, X.; Zhang, H.; Li, T. (2018): Network security threats situation assessment and analysis technology study. *International Conference on Measurement, Information and Control*, pp. 643-646.

Chen, J.; Tian, Z. H.; Cui, X.; Yin, L.; Wang, X. (2018): Trust architecture and reputation evaluation for internet of things. *Journal of Ambient Intelligence & Humanized Computing*, vol. 2, no. 9, pp. 1-9.

Cui, J. H.; Zhang, Y. Y.; Cai, Z. P. (2018): Securing display path for security-sensitive applications on mobile D\devices. *Computers, Materials & Continua*, vol. 55, no. 1, pp. 17-35.

Dacier, M.; Leita, C.; Thonnard, O. (2010): *Assessing Cybercrime Through the Eyes of the WOMBAT*. Cyber Situational Awareness.

GEWIS (2018): Global early warning information system.

http://www.acronymfinder.com/Global-Early-Warning-Information-System-%28GEWIS%29.html.

Han, W. H.; Tian, Z. H.; Huang, Z. Z.; Li, S. D.; Jia, Y. (2018): Bidirectional self-adaptive resampling in imbalanced big data learning. *Multimedia Tools and Applications*.

ISDAS (2018): JPCERT/CC, Internet Scan Data Acquisition System. http://www.jpcert.or.jp/isdas/.

Li, M. H.; Sun, Y. B.; Jiang, Y.; Tian, Z. H. (2018): Answering the min-cost quality-aware query on multi-sources in sensor-cloud systems. *Sensors*, vol. 12, no. 9, pp. 24-35.

Liu, P.; Meng, Y.; Wu, Y. (2013): Application of cluster analysis and outlier technology in network security situation. *Network Security Technology & Application*, vol. 5, no. 12, pp. 7-17.

LOBSTER (2018): Large-scale monitoring of broadband internet infrastructures. http://www.ist-lobster.org/downloads/index.html.

Luo, K. (2017): Research on network information security situation analysis and protection technology. *China Computer & Communication*, vol. 15, no. 5, pp. 53-61.

Barlow, J.; Thrower, W. (2000): TFN2K-an analysis. *Axent Security Team*, vol. 13, no. 2, pp. 21-29.

NCPS (2018)**:** National cybersecurity protection system. https://www.dhs.gov/national-cybersecurity-protection-system-ncps.

Niu, X. L.; Wang, R. (2011): Brief analysis of network security management platform situation and development in patent technology. *Computer Security*, vol. 15, no. 5, pp. 32-41.

Qiu, J.; Chai, Y. H.; Liu, Y.; Gu, Z. Q.; Li, S. D. et al. (2018): Automatic non-taxonomic relation extraction from big data in smart city, vol. 6, pp. 74854-74864.

Sang, Y. (2012): Current situation of campus network security and analysis of network security technology. *Intelligent Computer & Applications*, vol. 19, no. 8, pp. 107-118.

Sun, Y. B.; Li, M. H.; Su, S.; Tian, Z. H.; Shi, W. et al. (2018): Secure data sharing framework via hierarchical greedy embedding in darknets. *ACM/Springer Mobile*

Networks and Applications, vol. 12, no. 8, pp. 48-57.

Tan, Q. F.; Gao, Y.; Shi, J. Q.; Tian, Z. H.; Shi, W. et al. (2018): Towards a comprehensive insight into the eclipse attacks of tor hidden services. *IEEE Internet of Things Journal*.

Tian, Z. H.; Su, S.; Shi, W.; Yu, X.; Du, X. J. et al. (2018): A data-driven model for future internet route decision modeling. *Future Generation Computer Systems*, vol. 1, no. 6, pp. 4212-4221.

Tian, Z. H.; Cui, Y.; An, L. (2018): A real-time correlation of lost-level events in cyber range service for smart campus. *IEEE Access*, vol. 6, pp. 35355-35364.

Wang, X. (2017): Network security situation prediction method based on spatiotemporal analysis. *Jiangsu Science & Technology Information*, vol. 8, no. 19, pp. 43-52.

Wang, Y.; Tian, Z. H.; Zhang, H. L. (2018): A privacy preserving scheme for nearest neighbor query. *Sensor*, vol. 18, no. 8.

Wang, Z. H.; Liu, C. G.; Qiu, J. (2018): Automatically trace back RDP-based targeted ransomware attacks. *Wireless Communications and Mobile Computing*.

Yu, X.; Tian, Z. H.; Qiu, J. (2018): A data leakage prevention method based on the reduction of confidential and context terms for smart mobile devices. *Wireless Communications and Mobile Computing*.

Zhang, H.; Shi, J.; Chen, X. (2013): A multi-level analysis framework in network security situation awareness. *Procedia Computer Science*, vol. 17, no. 12, pp. 530-536.