Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 2, pp. 928–937, DOI 10.12694/scpe.v25i2.2610

NETWORK TRAFFIC MONITORING AND REAL-TIME RISK WARNING BASED ON STATIC BASELINE ALGORITHM

ZHAOLI WU* AND JUNWEI LIU†

Abstract. With the rapid growth of network traffic, in order to monitor network traffic, the author proposes a baseline based traffic inspection method. The main objective is to develop a global system for identifying malicious traffic, rather than a precise method for detecting the types of worms produced by malicious traffic. Although traffic is caused by the causes, network administrators can use this international search technique to detect malicious traffic data. The system based approach mainly includes designing time based on the traditional traffic model, detecting various equipments and network traffic process, and configuring the traffic flow according to each time frame. This method uses Cisco's NetFlow Collector, a NetFlow Collector (NFC), to collect raw NetFlow data transmitted by the device through UDP every 5 minutes. the Then, three-dimensional data such as communication port, communication time, and traffic flow (bytes or packets) is used to filter, remove the different values, calculate the base values, and compare the real-time results with the base values to check the traffic defects in the current network. If there are differences between the monitoring data and the system configuration at the same time, the system will issue an abnormal warning, and as time accumulates, the alarm level will gradually escalate.

Key words: Static baseline algorithm, Network traffic monitoring, Real time performance, Risk warning

1. Introduction. With the rapid development of broadband internet in China, the network scale of major operators in the country is constantly expanding, the network structure is becoming increasingly complex, network services are becoming increasingly rich, network traffic is growing rapidly, and the network environment is becoming increasingly complex [16]. Operators need to use reliable and effective network traffic monitoring systems to conduct timely and accurate traffic and flow analysis of their networks and various services carried by them, in order to tap into the potential of network resources, control network interconnection costs, and provide a basic basis for network planning, optimization and adjustment, and business development. Mainly manifested in the following aspects: By analyzing the network outlet flow and direction, it is possible to gain a detailed understanding of the access of internal users to other external networks, thereby effectively selecting the interconnection method and location with other networks, and saving interconnection link costs. Master the user's access to other operators: By monitoring the traffic of Wulian with other networks, analyze the business characteristics and main traffic directions of internal users accessing other external networks, accurately grasp the interests of internal users on the external network, and find the most applied hot information content. Based on the analysis results, the corresponding network content is constructed, and the hot information content that users are interested in is placed in the internal network to reduce the pressure on the interconnection link. Evaluate the cost and value of branch networks: By monitoring the incoming and outgoing traffic of each branch network, analyzing the size, direction, and content composition of the traffic, we can understand the bandwidth usage of each branch network, reflect the network cost it occupies, and also understand its business development and make a value evaluation [10, 9]. IP based billing applications and Service Level Agreement (SLA) verification services: By monitoring and analyzing the traffic on major customer access circuits, parameters such as business type, service level, communication time and duration, and communication data volume can be calculated, providing data basis for IP based billing applications and SLA verification services. By monitoring specific traffic in the network for a long time, it helps network administrators understand

*1. Jiangsu Vocational Institute of Architectural Technology School of Information and Electronics Engineering, Yangzhou, Jiangsu, 225006, China; 2.Jiangsu Collaborative Innovation Center for Building Energy Saving and Construction Technology, 221000, China; 3. School of computer science and technology, China University of mining and Technology, Xuzhou, Jiangsu, 221000, China. Corresponding author, ZhaoliWu5@126.com

†School of Internet of Things Technology, Wuxi Vocational College of Science and Technology, Wuxi, Jiangsu, 214028, China. JunweiLiu6@163.com

the network's traffic model. The benchmark data formed can be used by network administrators to correctly analyze the network usage status, and timely issue abnormal alarms. Preventive measures can be implemented before fault events occur or expand, thereby improving the overall quality and efficiency of the network [3]. The implementation of communication network vulnerability analysis focuses on the prevention of denial of service (DDoS) attacks and the spread of major diseases. Real-time analysis of network traffic quality helps to identify network traffic quality in time and identify the specific characteristics of network traffic quality. By comparing with the traditional principle of network communication, administrators can quickly determine whether the abnormal communication is the protection of network security, determine the type of security precautions, measure the potential danger and the potential to affect various kinds of attacks, develop and implement anti-accident measures in an emergency system. and implement anti-accident measures.By analyzing the traffic flow, the data base can be provided for network performance evaluation such as multi-outlet load balance, critical link bandwidth location, route selection, QoS location, etc. the network performance evaluation method can be used to evaluate the network performance of the network system. With the development of computer and communication technology, the scale and complexity of electric power enterprises are increasing, and more and more industrial applications are also becoming more and more complex. As a result, the possibility of various types of traffic jams is also increasing, and various traffic jams are also following. The generation of traffic jam not only affects the performance of network and reduces efficiency, but also can increase information confidentiality of enterprises, affect their development [17]. Therefore, differential detection has become an important challenge that power companies are facing. The main flow recording technologies mainly include monitoring protocol technology based on mirror (valve online) flow, distributed monitoring technology (based on hardware analysis technology), NetFlow based on monitoring technology, and SNMP based on monitoring technology.Compared to NetFlow monitoring technology, other technologies also have their own disadvantages. Automobile telescope can only be used for single link and is not suitable for wide area network inspection. The hardware probe technology is limited by upper limit of intersection speed. SNMP technology mainly collects some status information of equipment and related information. Because of the monotonicity of the data and some errors, data analysis is only possible on the data of network layer 2 and 3 and the status of the future equipment. NetFlow monitoring technology is involved in a unified process that does not rely on certain links.At the same time, it has high efficiency in collecting data, a wide range of network applications, lower costs compared to others, and high cost-effectiveness. In response to this research issue, Gong, Q. et al. a passive, nonautonomous intrusion detection system for 100G research networks is proposed. Lightning safety technology uses multi- core and GPU to achieve high quality packet and pipeline detection. It extracts the physical and physical conditions of the real time data in the network and gives them to the network anomaly detection system [4]. Bollmann, C. A. et al. this paper describes a new application of robust estimation in fault diagnosis of computer network traffic volume.The proposed algorithm is based on the position and distribution of the model, and is derived from the unknown zero order numbers. This test is not parametric and suitable for large-scale data analysis of external traffic network. Using two different global denial of service attacks including the actual ability of backbone network traffic to verify the effectiveness of these measures.Because of the heavy tails in the network traffic system (a special kind of real network traffic system), the simulation results are better than the traditional test means such as interpolation and interpolation. Monte Carlo analysis was used to evaluate the efficacy and apparent improvement rate from 7% to 11% in false negative. The test request is also shown to be associated with an average (a similar test case) [2].

On the basis of current research, the monitoring method proposed by the author, although lacking in accuracy and unable to fully monitor abnormal traffic, has been greatly simplified due to its highly singular mechanism. It is very useful for a complex and large network of a large enterprise, with improved efficiency and better real-time performance.

2. Methods.

2.1. Static baseline algorithm.

(1) Definition and Principle. The 'static baseline algorithm' refers to setting the same upper limit (upper baseline) or lower limit (lower baseline) that does not change over time within a 24-hour cycle range for a certain indicator, and dividing the normal range and abnormal area of the indicator value [20]. The algorithm principle is as show in Figure 2.1.

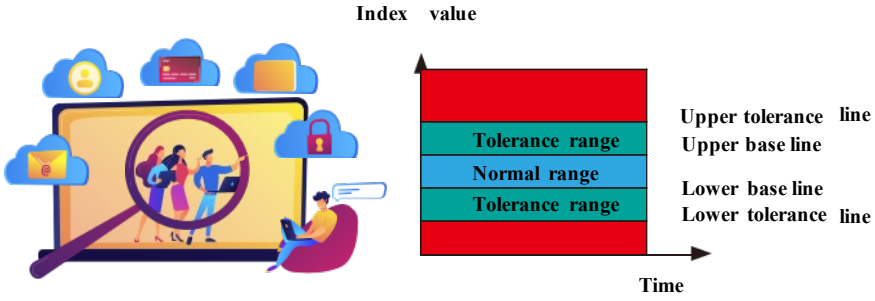

Fig. 2.1: Static baseline algorithm

For certain indicators, there may only be upper tolerance line or lower tolerance line, such as CPU load; Other indicators may require simultaneous attention, with both the upper and lower tolerance lines reflecting abnormal situations. The upper and lower tolerance line settings for distinguishing between normal and abnormal severity levels come from network operation and maintenance experience, management requirements, or device capacity limitations, and the accuracy of the settings determines whether the algorithm can work. Excessive tightness (such as a small upper limit value or a large lower limit value) may lead to false positives [6, 15]; Excessive looseness (such as high upper limit values and low lower limit values) may lead to false positives, increase the workload of monitoring personnel, affect the enthusiasm of maintenance personnel, and ultimately reduce the effectiveness of this technology. At the same time, as the number of indicators, network elements, and business systems included in the active monitoring scope increases, the number of upper and lower tolerance lines that need to be set will also increase sharply, indicating a practical need to improve work efficiency. Therefore, on the premise of pursuing accuracy, finding the automatic threshold setting mode becomes an important factor for the successful application of this algorithm, which needs to be constantly adjusted and optimized in the operation and maintenance.

(2) Method for determining the baseline. Traditionally, maintenance personnel determine thresholds for various indicators based on experience and set them manually in network management systems and other means, abbreviated as the "manual setting" method [19]. This method can adapt to work needs when there are few indicators and small fluctuations. However, with the inclusion of active monitoring indicators and a large number of equipment, this method is difficult to adapt to operational needs and has low work efficiency; For indicators with large fluctuations, it is easy to have arbitrary settings and strong subjectivity, which is not conducive to system maintenance. Of course, manual setting can absorb the experience of operation and maintenance personnel, and has a certain degree of flexibility. The indicators related to management requirements and equipment capabilities applicable to static baseline algorithms can be manually set to monitor whether they are below the assessment value or exceed the corresponding proportion of processing capacity; Other indicators applicable to this algorithm, due to their certain fluctuation range and compliance with certain statistical laws, can be automatically learned by the system to generate corresponding thresholds based on the principle of dynamic baseline algorithm, and automatically set. This method is referred to as the "automatic setting method". Therefore, both methods are applicable to different types of indicators, with the latter having a wider scope of application.

Manual setting method: Maintenance personnel refer to the historical performance, management requirements, or equipment capabilities of a certain indicator, and combine their own work experience to set the baseline: manually write it into the system configuration file or fill it in the configuration window to set the baseline [5].

Automatic setting method: In summary, the "automatic setting method" adopts a principle similar to the dynamic baseline algorithm, based on historical statistical data, regardless of time period differences, calculates a threshold that does not change with 24 hours of time, and automatically sets it as the baseline of normal data. The algorithm description is as follows: historical data values; A total of N pieces of data within all

Table 2.1: NetFlow field description

byte	content	describe
0∼3	srcaddr	IP address of the source
4∼7	dstaddr	IP address of the destination
8∼11	ne: xthop	IP address of the next network segment router
12∼15	input and output	SNMP index for input and output interfaces
16∼19	dPkts	Packets in this information flow
20∼23	dOctets	The total number of Layer3 bytes in a packet of information flow
24∼27	First	SysUptime at the beginning of information flow
28∼31	Last	SysUp time when the last packet of the information flow is received
32∼35	srcport and dstport	TCP/UDP source and destination port numbers
36∼39	pad1 ,prot and tos	Unused (i.e. content 0) bytes, IP protocol (e.g. 6=TCP, 17=UDP), and IP service type
40∼43	Flags, pad2, pad3	The cumulative OR of TCP flags, pad2 and pad3 are unused (i.e. bytes with content 0)
44∼48	reserved	Unused (i.e. bytes with content 0)

collection granularity for M consecutive days: If the collection granularity is 15 minutes, then the number of data N=MX24X (60/15); If the collection granularity is 30 minutes, then the number of data N=MX24X (60/30). N data are denoted as x1∼xN. Exclude abnormal data; Based on the maintenance records during this statistical period, eliminate data during holidays, malfunctions, and errors. It is also possible to arrange a set of data in a fixed proportion.When the system is implemented, the probability of normal data can be manually adjusted and then calculated. Determine the range of baseline values; Priority should be given to the probability distribution algorithm in the dynamic baseline algorithm section, followed by the sorting method.

2.2. NetFlow introduction. Cisco's NetFlow is a switching technology introduced under the IOS system [14]. It utilizes seven attributes with segment identification, including source IP address, target IP address, TOS byte, and layer 3 protocol type. At the same time, while being supported by Cisco routers, it quickly and accurately distinguishes different types of Flow, and then tracks, measures, and analyzes them to obtain information such as time, type, and size. NetFlow services can provide some other efficient and high-quality services, such as data statistics under the optimal exchange path cooperating with routers, efficient data statistics under the maximum limit of information interaction between routers and switches, and multi type data statistics such as users, protocols, ports, and service types. In addition, NetFlow can also be deployed anywhere in the network as a pathable device, and to some extent, security services can be implemented through packet filtering.

NetFlow has a relatively small impact on routers due to its use of special switching technology [18]. The query process only targets the first group. After the Flow is identified and distinguished, the subsequent groups default to being a part of it. This is oriented towards connection based processing, avoiding the operation of access lists, and thus achieving data collection under small impact. By further analyzing the field descriptions of NetFlow, its superiority can be better reflected. As shown in Table 2.1, it can be found that the flow record contains information with data flow identification such as the source/destination IP address, transport layer source, destination port number, etc., which includes tcpflags for data security related data. Address related data such as source address, destination address, source autonomy number, and destination autonomy number, etc. NetFlow provides many details about the data, it can analyze various information of data in detail.

In summary, NetFlow technology, due to its unique formal architecture, does not repeatedly search for each data packet, but analyzes the data flow, improving the previous inefficient traffic monitoring and high impact on devices, making it more suitable for large networks.

2.3. Baseline analysis. Basis analysis is a multi- period averaging algorithm which deals with the sequence of time, calculates the average value of the variables in each phase, and links their results to make a practical basis. It represents the normal value of the system, and when used for vehicle maintenance, it intro-

Table 2.2: Description of abnormal types of network abuse

Abnormal	Definition	Illustrate
ALPHA	Unusual high byte rate transmission from point to point	Bandwidth measurement experiment
DOS,DDOS	(Distributed) Denial of Service Attack	A large number of data packets are sent to a specific port of a separate destination IP (such as port 0)
Flashing crowding	Unusual high volume of requests for resources or services	Large number of web requests for a single IP (port 80)
scanning	Scan a host for vulnerable ports (port scanning) or scan the network for an attack target (network scanning)	Network Scan for Port 139 (NetBIOS)
worm	Autobroadcast code can be spread across the network through security vulnerabilities	Attack via Port1433 (MSSQL Snake worm)
Point to multipoint	Content distribution from one server to many users	A single server broadcasts to a large number of destination address sets on port 119
Transmission loss	Traffic reduction event caused by traffic exchange between OD flow pairs	Shutdown of nodes on the scheduled network backbone and measurement failure of nodes
Entrance transfer	Consumers transfer traffic from one entrance to another	Online consumers transfer their business traffic from one network to another

duces a traditional vehicle maintenance system [8]. The author adopts a baseline analysis method to monitor: certain specific ports (blacklisted) of well-known ports (port numbers below 1024), namely some published ports related to abnormal traffic, such as some worm viruses (135 ports used by shockwave viruses); All other non well-known ports (port numbers above 1024), excluding ports for certain specific applications (such as some internal communication software of enterprises) (whitelist method). By conducting baseline analysis on the historical and current records of ports that meet the above conditions, perform the following steps:

Read NetFlow data and process log data; Computing principles from historical data sources; Determine the importance of state value; Identify the host epidemic. The specific process form is described below.

(1) Select a baseline to analyze NetFlow records and perform statistics. From the data flow collected by NetFlow, locate the destination destination port (dst port) and identify it to determine whether it is the destination destination destination for the specified data [11]. If it is a known port, first check whether the target port is a blacklist port. If so, calculate the traffic data of the port and collect them into historical data for future calculation;If it is an unknown port, first check whether the destination port belongs to the whitelist port. If it is not, then count the traffic records of all relevant ports as a whole (with port number -1 indicating these overall ports).Network avoidance can include three main types: victim avoidance, lightweight avoidance, and performance avoidance. Most types of abuse in enterprises are abnormal, as shown in Table 2.2.

Table 2.2 shows the relationship between different traffic and the number of packets or bytes. Taking a worm as an example, there will be more data packets or bytes in terms of network traffic. However, the number of bytes transmitted by the worm is much smaller than that of traditional data, indicating that during transmission of the worm, the number of data packets in the network will increase, while byte statistics will not change much [12]. Therefore, at this point, counting the number of packets according to the response rate is more effective, which can detect the presence of abnormal worms and other network traffic. On the contrary, for flash memory congestion, such as some downloading attributes, the number of packets or bytes will increase, but the impact of downloading traffic costs on network performance is more significant. the effect of network performance is more significant.Therefore, in this case, the value of the byte count can have a better effect on whether the network is uniquely or unequally distributed.

(2) Calculate baseline values from historical table data. If the members of a company or an organization have a similar behavior, their network traffic still has a close relationship, that is, due to the work and rest of the members of the organization, the network traffic has a consistent pattern and the peak release status. According

to this characteristic, we divide the algorithm into two types: working day and non- working day. Week: Within 5 minutes, take the sample points of the corresponding port for the current minute of the previous 20 working days, calculate their average and standard deviations, and use the Grabbs method to eliminate negative results. Non working days: In units of 5 minutes, take a port sampling point corresponding to the current minute of the first 8 non working days, remove the maximum and minimum 5% data, and then take the average to obtain the current hourly baseline. Taking workdays as an example, the specific steps are described as follows:

Sort the flow values of the device in the first 20 working days for each port that meets the criteria, and obtain the following sequence:

$$\{X_{1_i}, X_{2_i}, \ldots, X_{20_i}, i_j \subset \{1, 2, \ldots, 20\} X_{1_i} \leq X_{2_i} \ldots \leq X_{20_i}\} \tag{2.1}$$

Remove the maximum and minimum values by 5%, leaving the following sequence:

$$\{X_{1_i}, X_{2_i}, \ldots, X_{19_i}, i_j \subset \{1, 2, \ldots, 20\} X_{1_i} \leq X_{2_i} \ldots \leq X_{19_i}\} \tag{2.2}$$

Find the average value, which is the baseline value corresponding to a port that meets the conditions during that period.

(3) Determine dynamic critical state values. Set T_h as the baseline threshold for a time period subdivision, and b_1 as the baseline value. The larger the value of b_1, the greater T_h should be, that is, b_1 is proportional to T_h. There are:

$$T_h = K \times b_1 \tag{2.3}$$

Assuming K is a constant greater than 1, make the necessary modifications to K value based on historical data. Simple analysis shows that the higher the K value, the easier it is to determine the transmission behavior, but it can also lead to an increase in detection rate, making some less parasitic bacteria can be detected [13]; On the contrary, the smaller the K value, the more convenient the transmission behavior, which can also increase the false value.Some normal information flows will also be misjudged as abnormal traffic. The system established by the author initially sets the K value to 2, meaning that abnormal alarms are only detected when the current value is twice the baseline value.

(4) Identify the source of abnormal traffic. When the traffic is monitored more than the critical value described in step 3 in a certain period of time, this cycle is called an abnormal cycle. Track the data during this period for further analysis and identification of host generated traffic defects based on NetFlow data. The key step is to identify raw data on NetFlow, sort the count byte or count packet at that time, port, or port from the top, and identify the master IP address and port for further processing.

2.4. Real time performance monitoring related technologies.

(1) Basic Principles. Real time performance monitoring technology includes the following categories: Focusing on performance indicator monitoring, comparing with indicator thresholds, historical statistical baselines, etc., to discover the degradation of business and network performance; Associate alarms with engineering cutover and configuration information, eliminate abnormal alarms, reduce invalid alarm rates, and improve work efficiency; Assisted by business dial testing and signaling detection to locate or discover faults; Real time performance monitoring technology focuses on discovering network anomalies and alarm generation mechanisms, but does not elaborate on the processing measures to be taken after the alarm is generated.

(2) Indicator selection. Real time performance monitoring captures the overall operational status of the business, promptly detects business or network anomalies, and monitors network operational status from the perspective of user business. Correspondingly, the selection of real-time performance monitoring performance indicators should reflect the operational status of the business or network. A feasible method for selecting performance indicators is to analyze business processes, decompose and refine them layer by layer from the perspective of business processes, and extract performance indicators in business process analysis. The selection of real-time performance monitoring indicators should follow the following principles and basic methods: Following the logical path of "Purpose>Method>Indicator Set", starting from analyzing the correspondence between business processes and network element devices, select complete indicators that match the positioning. Starting from customer perception, propose performance indicators that reflect the "end-to-end" business quality;

Propose supplementary performance alert indicators to address the limitations of network element capabilities and insufficient network management functions; Starting from the purpose of reducing the occurrence rate of faults, select performance indicators that can timely reveal the hidden dangers of faults, such as statistical channel utilization indicators within the base station range, in order to explain channel issues, if the scope is expanded to BSC, it is possible that a channel issue cannot affect the entire indicator, leading to valuable information being hidden; Meeting the principle of minimizing the set of indicators, that is, when multiple indicators can reflect the same fault or business degradation, only the indicators that can directly reflect the problem are selected, and other indicators are only queried and called during in-depth analysis, based on the high incidence of faults and complaints about network elements, select the key points. According to the fact that 75% of the complaints on the GSM network side belong to the wireless part, the traffic network will focus on the wireless part; The sources of various indicators need to be clearly defined (such as existing network management, Yu Ling method, automatic reporting of network elements, business call testing, etc.), collection granularity, latency requirements, feasibility, and the impact of monitoring work on network elements needs to be evaluated.

(3) Relevant regulations on indicator data. Indicator data source: From an analysis perspective, most performance indicators have been collected and stored in the network management system, and meet the time and spatial granularity requirements for indicator collection. These indicators can be directly extracted based on the existing network management system. For indicators that have been collected and stored in the network management system, but do not meet the time or spatial granularity requirements for indicator collection, or that cannot be provided by the existing network management system yet, this can mainly be achieved by modifying the network management system, if the network element's capabilities permit, increase the collection time density and refine the spatial granularity of the collection. In the case where the network management system renovation cannot meet the real-time performance monitoring requirements, it can be considered to achieve this by directly connecting to the network element, using instruction interaction, collecting device operation logs, and logging in to the device to obtain statistical reports. Data collection principle: For situations where instructions are required to retrieve network element data, the impact of the instructions on the network element itself should be considered. It is recommended to set the switch (protection threshold) for obtaining data in the real-time performance monitoring function module or network management system. When the system load is too high, the real-time performance monitoring module will automatically shut down the collection of this indicator. After the system load is normal, collection can be resumed to avoid further aggravation of the system load.

The principle for setting the granularity of collection time: Time granularity is the minimum period for extracting indicator data. The set value should be suitable for the monitoring needs. If it is too small, it will cause excessive system load pressure but not generate actual value; Excessive size may lead to loss of monitoring significance [1]. At present, the business quality data provision capabilities of mainstream devices include different granularity such as 1 minute, 5 minutes, 15 minutes, and 1 hour. The granularity for selecting business quality indicators for real-time performance monitoring is generally 15 minutes. In special cases, finer or more relevant granularity can be considered. The system connection rate indicator of MSC is a comprehensive indicator that can comprehensively reflect network issues, covering the overall performance of wireless and switching systems. The user perception is obvious, and considering the impact of collection on the system, it is recommended to use a time granularity of 15 minutes. The location update success rate of HLR reflects the successful login of users to the network, which can reflect the use of illegal cards, SIM card errors, office data errors, etc. Considering the impact of its collection on business systems, it is recommended to use a time granularity of 1 hour. The principle for determining the collection delay requirement: The delay of business quality alarms is directly related to the size of real-time performance monitoring capabilities. However, business quality alerts require a complete process of data generation, extraction, analysis, and presentation. From the perspective of balancing real-time requirements and feasibility, the delay cannot exceed one data cycle. For data with a granularity of 15 minutes, the delay should not exceed 15 minutes. The system presentation must be completed before 10:30 from 10:00 to 10:15.

3. Result analysis. This experiment mainly collects Net Flow data from the backbone router of a certain power enterprise information network, and uses running Oracle stored procedures to calculate baseline values

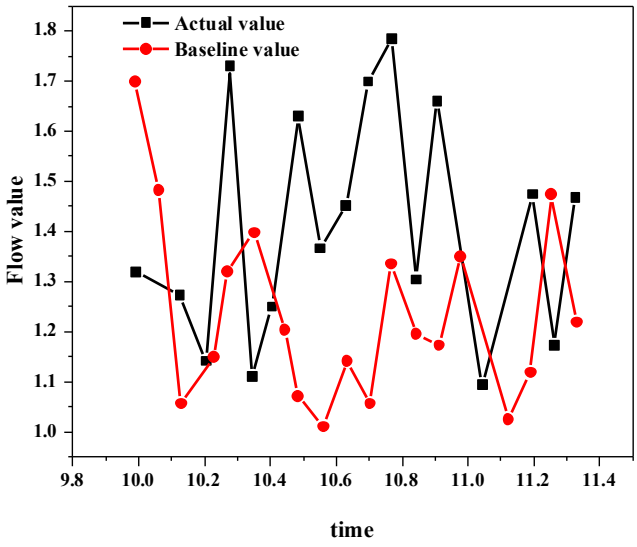

Fig. 3.1: Unknown port traffic value 1

for result analysis. Taking the data results from working days as an explanatory analysis, the situation is similar for non working days. Figure 3.1 shows the unnamed port traffic values for a period of real traffic and calculated baseline values from 10:00 to 11:35 on a certain working day, which is a normal situation and there is no alarm situation.

Figure 3.2 shows the unnamed port traffic values for a period of real traffic and calculated baseline values from 14:00 to 15:35 on a certain working day. The actual value exceeds the baseline value by more than twice, which is an abnormal situation and requires an alarm. The IP and port that generated large traffic need to be analyzed through the original NetFlow record, and the reason should be analyzed.

Figure 3.3 shows the actual traffic and calculated baseline values for TCP port 135 on a certain working day from 16:00 to 17:35. This port is used for shock waves, with a byte count of 48 [7]. From the graph, we can see that at 16:45, the traffic value of Port 135 suddenly increased, and then returned to normal. It may be that a certain machine was infected with the shockwave virus. When the machine tried to infect other machines, it was blocked by a firewall, causing the traffic value to return to normal.

4. Conclusion. A baseline algorithm based on the method for detecting traffic anomalies in enterprises is proposed, and a baseline algorithm with multiple time-frequency segments and multiple communication ports is proposed. The baseline algorithm is proposed. Based on this dynamic principle, the analysis is made on the traffic data of operation and failure date in order to achieve the purpose of detecting the traffic malfunction. Based on the experimental results, the advantages of this method are as follows: Adopting baseline analysis method has the advantage of singularity. The baseline based abnormal traffic monitoring method proposed by the author has a high degree of singularity, which is analyzed in multiple time periods of working and non working days. Only the baseline is used as the sole criterion for analysis in monitoring, avoiding the tedious and complex monitoring methods of other methods. Different detection methods will not be used according to the reasons for abnormal traffic generation, this overall mechanism presents a high degree of singularity in the analysis method. The monitoring method proposed by the author, although lacking in accuracy and unable to fully monitor abnormal traffic, has been greatly simplified due to its highly singular mechanism. It is very useful for a complex and large network of a large enterprise, with improved efficiency and better real-time performance. However, the first base algorithm proposed by the author is simple, and the importance of multi-bases has a significant impact on the accuracy of the algorithm. In the future, it is necessary to improve the basis generating algorithm while control complexity to realize traffic congestion.

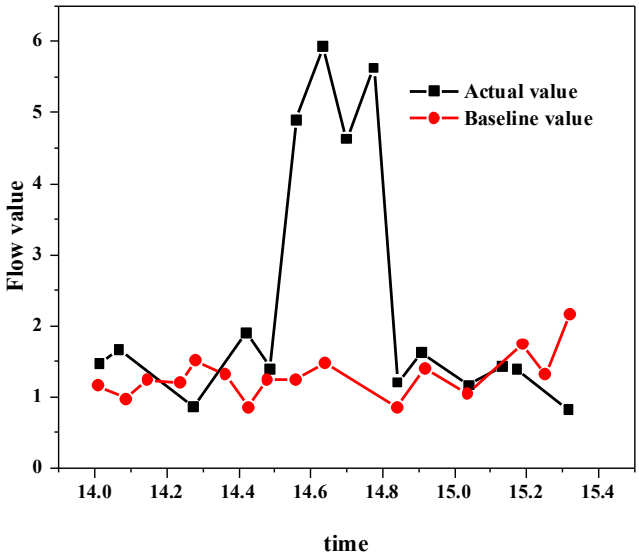

Fig. 3.2: Unknown port traffic value 2

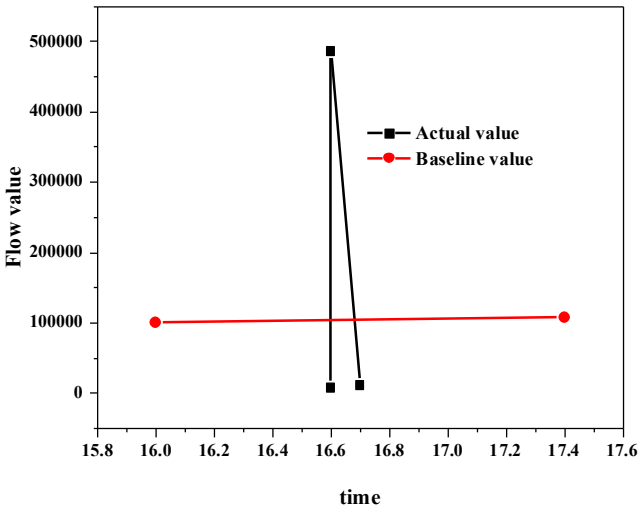

Fig. 3.3: Flow Value of Famous Port (135 Port)

Acknowledgement. Jiangsu Collaborative Innovation Center for Building Energy Saving and Construction Technology:XT-Research and application of public building energy consumption monitoring system based on big data, Project No: SJXTY1603, Project Leader: Zhaoli Wu.

REFERENCES

[1] J. Azimjonov, A. Özmen, and M. Varan, *A vision-based real-time traffic flow monitoring system for road intersections*, Multimedia Tools and Applications, (2023), pp. 1–20.

[2] C. A. Bollmann, M. Tummala, and J. C. McEachen, *Resilient real-time network anomaly detection using novel non-parametric statistical tests*, Computers & Security, 102 (2021), p. 102146.

[3] L. Duan, J. Zhou, Y. Wu, and W. Xu, *A novel and highly efficient botnet detection algorithm based on network traffic*

analysis of smart systems, International Journal of Distributed Sensor Networks, 18 (2022), pp. 182459–182476.
[4] Q. Gong, P. DeMar, and M. Altunay, *Thundersecure: deploying real-time intrusion detection for 100g research networks by leveraging stream-based features and one-class classification network*, International Journal of Information Security, 21 (2022), pp. 799–812.
[5] J. Guo, X. Ding, and W. Wu, *Reliable traffic monitoring mechanisms based on blockchain in vehicular networks*, IEEE Transactions on Reliability, 71 (2021), pp. 1219–1229.
[6] S. He, L. Chen, S. Zhang, Z. Guo, P. Sun, H. Liu, and H. Liu, *Automatic recognition of traffic signs based on visual inspection*, IEEE Access, 9 (2021), pp. 43253–43261.
[7] A. M. Htut and C. Aswakul, *Development of near real-time wireless image sequence streaming cloud using apache kafka for road traffic monitoring application*, PLoS One, 17 (2022), p. e0264923.
[8] J. Jiao, X. Sun, Y. Zhang, L. Liu, J. Shao, J. Lyu, and L. Fang, *Modulation recognition of radio signals based on edge computing and convolutional neural network*, Journal of Communications and Information Networks, 6 (2021), pp. 280–300.
[9] A. Krishnakumar, S. Kadian, U. Heredia Rivera, S. Chittiboyina, S. A. Lelièvre, and R. Rahimi, *Organ-on-a-chip platform with an integrated screen-printed electrode array for real-time monitoring trans-epithelial barrier and bubble formation*, ACS Biomaterials Science & Engineering, 9 (2023), pp. 1620–1628.
[10] C. Li, Q. Peng, D. Wang, L. Luo, and H. Zuo, *Smart substation network quality monitoring and fault prediction*, 2108 (2021), p. 012061.
[11] D. Li, J. Bao, S. Yuan, H. Wang, L. Wang, and W. Liu, *Image enhancement algorithm based on depth difference and illumination adjustment*, Scientific Programming, 2021 (2021), pp. 1–10.
[12] X. Liu, Q. Li, W. Chen, P. Shen, Y. Sun, Q. Chen, J. Wu, J. Zhang, P. Lu, H. Lin, et al., *A dynamic risk-based early warning monitoring system for population-based management of cardiovascular disease*, Fundamental Research, 1 (2021), pp. 534–542.
[13] ———, *A dynamic risk-based early warning monitoring system for population-based management of cardiovascular disease*, Fundamental Research, 1 (2021), pp. 534–542.
[14] J. Pang and Z. Zhao, *Real-time monitoring of fluidized bed agglomerating based on improved adaboost algorithm*, 1924 (2021), p. 012026.
[15] X. Qi, Y. Ji, W. Li, and S. Zhang, *Vehicle trajectory reconstruction on urban traffic network using automatic license plate recognition data*, IEEE Access, 9 (2021), pp. 49110–49120.
[16] F. I. H. Sakiyama, F. Lehmann, and H. Garrecht, *A novel runtime algorithm for the real-time analysis and detection of unexpected changes in a real-size shm network with quasi-distributed fbg sensors*, Sensors, 21 (2021), p. 2871.
[17] Y. Wang, Q. Wang, D. Suo, and T. Wang, *Retraction note: Intelligent traffic monitoring and traffic diagnosis analysis based on neural network algorithm*, Neural Computing and Applications, 35 (2023), pp. 4183–4183.
[18] H. Wu, Y.-P. Zhao, T.-L. Yang, and H.-J. Tan, *An ensemble radius basis function network based on dynamic time warping for real-time monitoring of supersonic inlet flow patterns*, Aerospace Science and Technology, 111 (2021), p. 106551.
[19] X. Xiang, C. Chen, H. Wang, H. Lu, H. Zhang, and J. Chen, *A real-time processing method for gb-sar monitoring data by using the dynamic kalman filter based on the ps network*, Landslides, (2023), pp. 1–17.
[20] W. Yu, K. Ruan, H. Tang, and J. Huang, *Routing hypergraph convolutional recurrent network for network traffic prediction*, Applied Intelligence, 53 (2023), pp. 16126–16137.

Edited by: B. Nagaraj M.E.
Special issue on: Deep Learning-Based Advanced Research Trends in Scalable Computing
Received: Sep 14, 2023
Accepted: Nov 6, 2023