Jurnal Kejuruteraan 37(3) 2025: 1349-1362
https://doi.org/10.17576/jkukm-2025-37(3)-20

BGP Security Analysis Using Network Simulation: An Impact Study of Cyber Attacks

Nassir S. Kadhim[a,b*], Kalaivani Chellappan[a*]& Nor Fadzilah Abdullah[a]

[a]*Faculty of Engineering and Built Environment; The National University of Malaysia (UKM), Malaysia*

[b]*Ministry of Communication (MOC) / ITPC Company, Iraq.*

[*]*Corresponding Author: kckalai@ukm.edu.my*

Received 4 April 2024, Received in revised form 4 October 2024
Accepted 4 September 2024, Available online 30 May 2025

ABSTRACT

BGP (Border Gateway Protocol) is the standardized routing protocol for the internet, enabling the exchange of routing information between autonomous systems (ASs). Despite its critical role in ensuring global routing stability and rapid convergence, BGP remains vulnerable to diverse and increasingly sophisticated anomalies, including Hijacking, Denial-Of-Service (DOS), and outages. Although the recent advancements in machine learning (ML) hold promise for accurate BGP anomaly detection, existing publicly available datasets often contain outdated information regarding past cyberattacks, hindering models of novel threats. Furthermore, the network topology criteria are also often neglected for anomaly identification. These shortcomings render inadequate training and robust ML models in BGP security applications. In this work a realistic BGP network topology is modelled to examine the routing behaviour of BGP traffic in the presence of different attack scenarios. We proposed a network simulation that included collecting BGP refreshed data, extracting, testing, and verifying (20) BGP features, and visualizing the features that are most impacted by each cyber-attack scenario based on contrasting the graphic patterns. Additionally, we created and formatted our own datasets to be used as input into a ML detection model. According to statistical analysis, the findings showed that six BGP features were the most significant regarding the effect of BGP cyber-attacks; nine features had moderate significance; seven features were less significant; and two features were found to be unaffected by cyber-attacks. the leveraging of obtained results will assist in building an accurate and efficient ML model for detection BGP anomalies.

Keywords: BGP traffic; network simulation; datasets; cyber-attacks; statistical analysis

INTRODUCTION

The Border Gateway Protocol (BGP) plays a critical role in the Internet's routing infrastructure, enabling efficient and reliable data exchange between autonomous systems (Zhao, Xi et al. 2021). This protocol is not just responsible for disseminating reachability but also plays a significant role in inter-domain routing decisions (H. Guo et al. 2009). Despite that, the inter-domain routing system is facing a combination challenge of supporting authenticity and security measures. Some of the BGP network cyber-attacks events, especially when they are unexpected or anomalous, might lead to an outage of Internet connectivity or sudden traffic shifts and degrade network performance (Aceto et al. 2018). BGP cyber threats or communication interruptions and failures (like marine cable cutting, link failure) can cause such as these unexpected events. Therefore, early detection of abnormal events is vital for mitigating their impact on BGP infrastructure. Over the past few years, ML methods have improved BGP anomaly detection by utilizing volume and path features of BGP update messages. From the perspective of machine learning, the BGP anomaly detection problem can be reduced to a two - class classification problem, regular class, and anomaly class (M. Cheng et al. 2018).

The BGP protocol was not designed with security in mind. There is plenty of evidence that it is inherently fragile (Wang Cun et al. 2016), (Goldberg Sharon, 2014) because the BGP mechanism is insufficient for ensuring the security of BGP messages. While BGP messages are exchanged among network peers, BGP anomalies can occur due to a variety of events, such as router misconfigurations, Hijacking, Denial-Of-Service (DOS) and Outages. Detecting and mitigating these anomalies in a timely and accurate manner is crucial for maintaining network security and stability. Many techniques have been employed to detect BGP anomalies (Bahaa Musawi et al. 2016). Many anomaly detection approaches in machine learning rely on selecting BGP features and datasets, which have limitations and are not appropriate for inputting into most ML models. Our work introduces proposed an advanced approach for extracting refreshed BGP features for anomaly detection model through network simulation. Our proposed method uses querying techniques and visualization to select and test effective features in various anomaly conditions by virtual environment that emulates real-world network topology. The primary goal of this research is to identify the most impacted features in cyber-attack conditions and create unique datasets based on the significance of features in a binary classification context and predict a binary outcome (Normal, Anomalous) that will be used to achieve the highest detection accuracy by utilizing various ML algorithms.

Several previous anomaly detection techniques chose BGP features associated with a small number of past security events that are mostly out of date. Furthermore, these methods do not consider the entire BGP network topology. These limitations make it difficult to determine whether their models are effective in real-world scenarios where the classifier must classify new events (Paiva Thales et al. 2021). By collecting actual BGP network information (assisted by MOC-ITPC) and leveraging network simulation, we provide a more realistic and dynamic testing environment allowing us to capture a wide range of BGP anomalies and evaluate their impact on network performance. Furthermore, we extracted (20) features from AS-PATH and BGP update message attributes. We then visualized the most affected features by cyber-attacks and generated our own (24) datasets, which will aid in the development of a quick and accurate anomaly detection system. Our method's effectiveness is based on extensive simulations and comparisons of various cyber-attack conditions for actual BGP network topology, as well as assessing the sensitivity to changes on topological characteristics.

The main contribution of this paper is summarized as follows:

1. Examined the most common BGP anomaly types, like direct (hijacking) and indirect (outages) and is not limited to any specific cyber-attack event.
2. Created and formatted our own private datasets of different BGP network conditions, which are composed of a variety of simulated events.
3. Take in consideration all the entire physical and logical topologies of the BGP realistic network (Iraqi internet gateways network).

The simulation tool utilized has successfully generated a realistic topology that accurately represents the actual network of (IRAQI Internet Gateways (IIGW)) as Iraqi national backbone network (fiber optic cables) main portion. This is the first research study into the security of this BGP network encompassing a comprehensive security analysis that has identified vulnerabilities susceptible to cyber-attacks. The outcomes of this research have significant implications for network operators, service providers, and security professionals, offering an improved understanding of BGP anomalies and enabling proactive measures to mitigate their impact. By enhancing the accuracy and efficiency of BGP anomaly detection based on the actual topology information and own datasets, we can contribute to the comprehensive security and resilience of the BGP infrastructure. Overall, this research presents an advanced strategy to extract the features BGP anomaly detection using programs platform highlighting its ability to advance the field and address evolving challenges of the cyber threats on BGP networks.

The rest of the paper is structured as follows: the next part is the background and related work, followed by our approach methodology. Then after, the results and simulation outcomes are outlined, followed by the discussion. Finally, the paper ends with the conclusion.

BACKGROUND AND RELATED WORK

The Internet is a global network composed of tens of thousands of autonomous systems (ASes) that are interconnected administrative domains (Alotaibi et al. 2022). An AS consists of a collection of routers responsible for disseminating a set of IP routing prefixes. Routers on the Internet use an interdomain routing protocol called the Border Gateway Protocol (BGP) to make the Internet work by enabling data routing (Fonseca, Paulo et al. 2019). The purpose of BGP is to help an organization's business goals by providing network reachability information to other organizations. Any BGP activity that does not contribute to or undermines those business goals is considered anomalous. Unfortunately, determining whether a particular activity is or is not furthering those goals system can be difficult.

Detecting BGP anomalies accurately and quickly is crucial for maintaining network reliability and preventing potential threats. BGP anomalies are classified into four types: direct intended, direct unintended, indirect, and network failures (Bahaa Musawi et al. 2016). Traditional approaches to BGP anomaly detection often rely on manual analysis or simplistic rule-based methods, which are time-consuming and prone to false positives. As networks grow larger and more complex, there is a need for more advanced and efficient anomaly detection models (Sanchez et al. 2019). In recent years, several works have been developed to improve BGP anomaly detection systems based on ML algorithms. In machine learning approaches, BGP features are used for anomaly detection. (J. Mai et al. 2008) used a single BGP feature for anomaly detection. (Shivani Deshpande et al. 2008 &2009), employed several features regarding BGP volume and AS-PATH length to detect BGP anomalies. (Marijana et al. 2018) utilized 15 BGP features for the detection of BGP anomaly, and the authors took care of imbalanced data sets by resampling to improve classification accuracy. Table 1 summarizes the number of ML approaches that have been used to detect Border Gateway Protocol (BGP) anomalies based on extracting BGP features to achieve the highest model accuracy (Takhar & Trajković 2023). We proposed a network simulation-based method for detecting BGP anomalies accurately and quickly. Using advanced querying strategies and python scripts techniques, our method aims to provide real-time anomaly detection by evaluating and validating 20 BGP used by different BGP anomaly detection techniques related to the number of BGP updates and AS-PATH attributes. All the extracted features tested under three of the cyber-attack scenarios, hijacking, denial of service (DOS), and the outage.

TABLE 1. A summary of BGP anomalies detection based on machine learning approaches by BGP features extraction

Author-Year	No. of extracted features	Type of anomaly	Best Accuracy
H. Kalra et al. 2021	37	Direct unintended anomaly, system	91.31%
Paiva Thales et al. 2021	17	Indirect, direct or link failure of BGP anomalies	91%
N. Al-Rousan & Trajkovi´c, 2012	37	Indirect anomaly	91.4%
Lulu et al. 2014	9	Direct unintended anomaly	95%
Allahdadi et al. 2017	18	DDoS attacks, power outage	98%
M. Cheng et al. 2018	33	DDoS attacks	99.5%
Sanchez et al. 2019	14	RTL	94%
X. Dai et al. 2019	37	DDoS attacks	96.03%
M. Cosovic et al. 2018	15	RTL, DDoS attacks, power outage	99.8%
(Z. Li et al. 2022)	37	DDoS attacks, WannaCrypt, power outage	81.77%
K. Hoarau et al. 2021	32	Google leak	90%
Kevin Hoarau et al. 2022	13 Features graph	Forged AS Path	67%
Rahul D. Verma et al. 2023	10	Indirect anomaly	85.39%

METHODOLOGY

BGP INFRASTRUCTURE IN IRAQ

Iraq's geographical location is significant in the field of communication and the internet because it is one of the Arab world's easternmost countries, where shares borders with Turkey to the north, Iran to the east, Syria and Jordan to the west, and Saudi Arabia and Kuwait to the south. Iraq has 36 miles (58 km) of coastline along the northern end of the Arab Gulf, giving it a sliver of territorial sea used for connecting with submarine cables (John E. Woods et al. 2023). In addition to the submarine cable landing station, Iraq's Ministry of Communication (MOC) operates and controls a distributed Internet gateway network (BGP infrastructure) at border locations to manage international internet traffic transiting through all of Iraq's surrounding nations Figure 1.

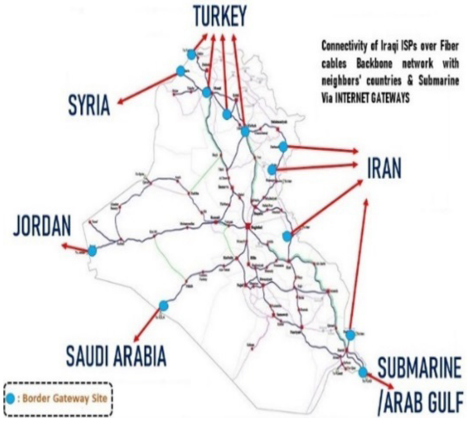

TURKEY
Connectivity of Iraqi ISPs over Fiber cables Backbone network with neighbors' countries & Submarine Via INTERNET GATEWAYS
SYRIA
IRAN
JORDAN
SAUDI ARABIA
SUBMARINE /ARAB GULF
: Border Gateway Site

FIGURE 1. Iraqi Fiber cable backbone and Gateways (Kassim AL-Hassani 2019)

DESIGNING NETWORK SIMULATION

This research uses a network design reference taken from the current Internet design, namely BGP Internet Full Route (Huston 2023), consisting of routers that use the EGP routing protocol. Each router has an identity number to represent an organization/company that has a connection in the internet network, called the AS Number. In this research, we allocated a sequence range of AS numbers to represent local ISPs (Internet services provider)/ ISPs inside Iraq, then selected a different AS number range to represent internet networks in other neighboring countries and global internet networks. Furthermore, we assigned a specific number to represent the AS number for the Ministry of Communication (Iraq), which is configured in all gateway routers.

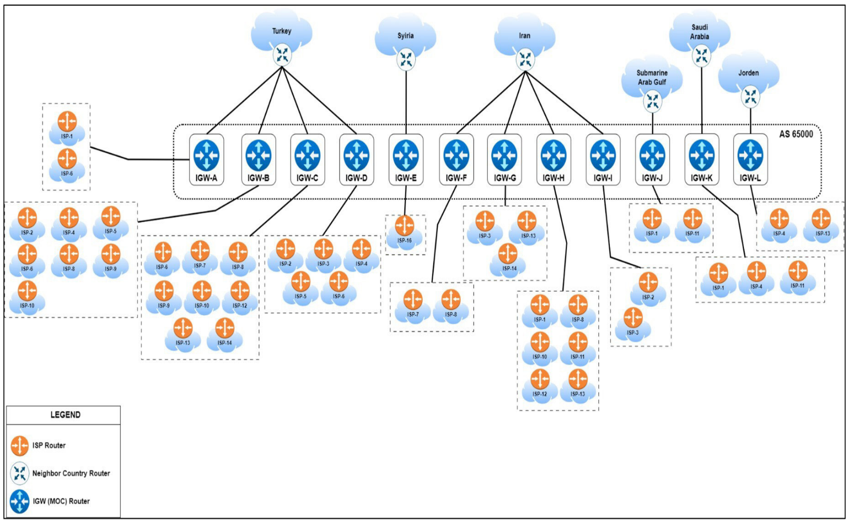


FIGURE 2. Iraqi gateway network high level design (HLD)

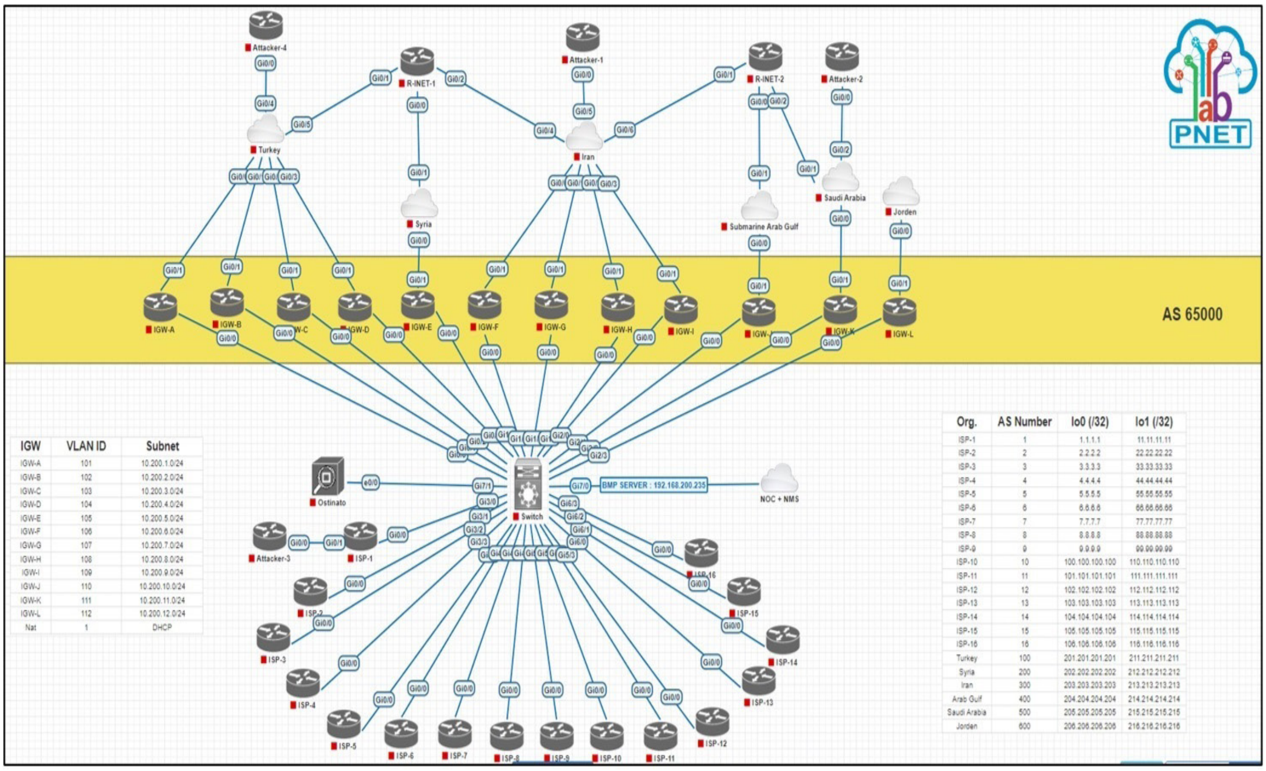
The high-level network design (HLD) for Iraqi internet gateways network used in this study is shown in Figure 2.

SIMULATION SOFTWARE

Network simulation software such as GNS3 and Cisco Packet Tracer can be used to simulate networks by running routers, switches, firewalls and even servers to test a network technology/protocol. It is commonly used as a simulator for learning and perfecting best-practice concepts in network design, including the Internet. However, these simulators software have limitations for our research work, such as not supporting gigabit interfaces, also not supporting running dozens of devices, and requiring very large CPU and memory resources in addition to supporting old device versions only.

As a result, Packet Network Emulator Tool Lab (PNETLAB) was used as a simulator for the purpose of bypassing GNS3 limitations and Packet Tracer. PNETLab is a platform that allows you to download and share labs with the community (Pasquale Chiacchio et al. 2007). PNETLab is a virtual machine. It is installed on the local machine and the Lab will run on it, so the lab's speed is unconstrained. PNETLab can be accessed from a web platform with hundreds of free Labs in the fields of networking, database, system at the PNET Lab store. In PNETLab, there is a QEMU feature to run network devices based on original OS / Firmware. Also, can be customized network bandwidth, and gigabit support. PNETLab can also run dozens, if not hundreds, of simulation devices concurrently. This study requires at least 50 devices to be operational simultaneously.

We have simulated the BGP network topology of Iraq by means of building the HLD (High Level Design) of MOC IGW peers' topology Figure 3 and create Lab environment for simulating the attacks scenario using PNETLAB. We have used Cisco IOS Firmware. (XR, Vios, IOL), and make similar tuning to the actual network devices which are Cisco Router 9k and 10k.Then we configured the routers to establish eBGP Routing protocol between the network layers.



IGW	VLAN ID	Subnet
IGW-A	101	10.200.1.0/24
IGW-B	102	10.200.2.0/24
IGW-C	103	10.200.3.0/24
IGW-D	104	10.200.4.0/24
IGW-E	105	10.200.5.0/24
IGW-F	106	10.200.6.0/24
IGW-G	107	10.200.7.0/24
IGW-H	108	10.200.8.0/24
IGW-I	109	10.200.9.0/24
IGW-J	110	10.200.10.0/24
IGW-K	111	10.200.11.0/24
IGW-L	112	10.200.12.0/24
Nat	1	DHCP

Org.	AS Number	lo0 (/32)	lo1 (/32)
ISP-1	1	1.1.1.1	11.11.11.11
ISP-2	2	2.2.2.2	22.22.22.22
ISP-3	3	3.3.3.3	33.33.33.33
ISP-4	4	4.4.4.4	44.44.44.44
ISP-5	5	5.5.5.5	55.55.55.55
ISP-6	6	6.6.6.6	66.66.66.66
ISP-7	7	7.7.7.7	77.77.77.77
ISP-8	8	8.8.8.8	88.88.88.88
ISP-9	9	9.9.9.9	99.99.99.99
ISP-10	10	100.100.100.100	110.110.110.110
ISP-11	11	101.101.101.101	111.111.111.111
ISP-12	12	102.102.102.102	112.112.112.112
ISP-13	13	103.103.103.103	113.113.113.113
ISP-14	14	104.104.104.104	114.114.114.114
ISP-15	15	105.105.105.105	115.115.115.115
ISP-16	16	106.106.106.106	116.116.116.116
Turkey	100	201.201.201.201	211.211.211.211
Syria	200	202.202.202.202	212.212.212.212
Iran	300	203.203.203.203	213.213.213.213
Arab Gulf	400	204.204.204.204	214.214.214.214
Saudi Arabia	500	205.205.205.205	215.215.215.215
Jorden	600	206.206.206.206	216.216.216.216

FIGURE 3. Network design via PNETLab Simulation

DEVICE ROLE AND FUNCTIONALITY

There are 5 types of routers based on their roles used in our network topology as shown in Table 2.

TABLE 2. Types and quantity of utilized routers

Role	Device Type	Quantity
IGW (Internet Gateway)	Router Cisco	12
Local ISP (Internet Service Provider)	Router Cisco	16
Neighbor Country	Router Cisco	6
Attacker	Router Cisco	4
ISP Tier-1 (International)	Router Cisco	2

The internet gateway Router (IGW) functions as a gateway for internet traffic from within Iraq to abroad, and vice versa. The IGW Router is connected to a local ISP, the ISP is located inside Iraq and has the need for a global internet connection. The IGW Router is also connected to the Neighbor Country. That is a country that borders directly with Iraq. In this topology, Iraq is geographically linked to five other countries uses it to connect to international ISPs, namely Turkey, Syria, Iran, Saudi Arabia, and Jordan, in addition to marine tape across the Arab Gulf using submarine cables.

To adapt to the original internet environment, All the neighboring countries above are connected to international ISPs, there are 2 international ISP Routers as simulations to represent the Internet globally - tier level (R-INET 1 / AS123 and R-INET-2 / AS456). These internet ISP routers connect several countries and provide several services on the internet such as hosting the main Web Servers, and there are also devices that become traffic generators in our topology scenario.

TRAFFIC GENERATION AND DISTRIBUTION

Traffic Generator uses a customized Python Script that runs on Ubuntu Linux OS. There are 2 Traffic Generators in our topology (Traffic Generator & Traffic Generator 2), each of which is connected to two of International Internet Routers.

TABLE 3. Actual & simulated distributed BGP traffic

ROUTER MOC	Traffic Distribution			
	Actual Traffic Distribution		Simulated Traffic Distribution	
	GB	Percentage	KB	Percentage
IGW-A	396	24%	440	22%
IGW-B	76	5%	94	5%
IGW-C	94	6%	116	6%
IGW-D	120	7%	148	7%
IGW-E	0	0%	80	4%
IGW-F	31	2%	38	2%
IGW-G	118	7%	146	7%
IGW-H	505	31%	500	25%
IGW-I	36	2%	44	2%
IGW-J	0	0%	100	5%
IGW-K	228	14%	281	14%
IGW-L	16	1%	20	1%
Total	1620	100%	2008	100%

The Traffic Generator function here is to simulate actual internet traffic. When the python script is executed, the Traffic Generator will send traffic in the form of UDP Packets (connectionless) to the intended IP address. The router will automatically send the packets to the destination address using the path from the IGW route to achieve normal traffic results identical to the actual network, then the traffic will be distributed to all IGWs.

This traffic's distribution has been predetermined based on data from the Gateway network Routers belonging to the Ministry of Communication Iraq (MOC). The data from the original IGW Router against the simulated traffic distribution is shown in Table 3. Actual Traffic is the real traffic from each IGW measured in GB units. Because Simulation Software cannot generate that much traffic, we must reduce it to the MB or KB level. In this case the (Simulated Traffic) in KB units with a balanced percentage ratio will be used.

PEERING BGP

IGW router needs to be mapped for BGP peering to a local ISP. The BGP configuration used in this network simulation is eBGP (external BGP) Because each router will peer with

a router with a different AS Number. Topologically followed the actual network pattern shown below:
Local ISP-MOC(IGW)-Neighbor-International peer

Each Local ISP conducts peering at least one IGW and conducts advertisements / prefix announcements. The prefix made by the announcement is the prefix of the IP Loopback. IP Loopback is used as a representation of the Internal Network at the Local ISP. IGW routers only perform BGP peering to Local ISP and neighboring Country. There is no own prefix announcement or advertisement going from this router because its function is only as a gateway. The Neighbor Country routers work as transit level to the tier level providers.

After all the routers have been configured, the connection must be tested using the Traffic Generator to determine whether the traffic has passed through the IGW. To view traffic that has passed through IGWs, Network monitoring tool is required, and it is necessary to use a specific network monitoring software. In this research, we use Grafana as a Network Monitoring System (NMS). Grafana NMS is used to monitor IGW routers traffic through the device's interfaces. Grafana can monitor routers using the SNMP and BMP protocols (W. Shao 2021).

BGP FEATURES EXTRACTION

The BGP routing protocol uses attributes to determine the best path selection, and common attributes include Weight, Local Preference, AS- PATH, MED, ORIGIN. Data for anomaly detection was collected in this study based on the AS -PATH attribute and the BGP updates message (Hammood & Musawi 2021).

Table 4, show all the possible features used by various BGP anomaly detection techniques previously which are extracted (Bahaa Musawi et al. 2016), monitored, tested, and verified in our work. Volume (Update message) and AS -PATH features are the two types of features extracted from BGP messages. Volume characteristics provide information about the absolute quantity of each type of message in relation to the type of information it communicates (e.g., reannouncements, withdrawals). AS -PATH features (e.g., longer paths, rare Ases in paths) provide insight into the behavior of announced AS -PATH (Fonseca, Paulo et al. 2019). We used PostgreSQL software, and a SQL query script to extract the 20 features from OpenBMP (our work database), which is used to gather BGP data.

THE RESULTS

For monitor BGP traffic behavior across all gateway devices, the graph tool is adopted in both normal condition and cyber-attack conditions. Cyber Attack needs to be simulated in order to get the result of behavioral analysis. We used three different Attack Scenarios: Hijacking, Denial of Services (DOS), and Outages.

NORMAL CONDITIONS

We collected all the data from normal conditions circumstances using the Grafana monitoring tool (Dewo et al. 2023), including traffic usage (SNMP) and extracted BGP features (BMP). The traffic statistics are relatively stable, indicating the normal traffic from local ISPs. The BGP features were also monitored using different dashboards and graphs.

PREFIX-HIJACKING CONDITION

In this type of hijack, an attacker configures its BGP router to announce a prefix belonging to another AS. BGP allows any BGP speaker to announce any route regardless of whether the route exists or not (Haeberlen et al. 2009); therefore, the attacker's neighbors will adopt it as a new route. There is another type of prefix hijacking that is related. Which is Sub-Prefix Hijack: In this scenario an attacker announces a sub-prefix that belongs to a victim AS. BGP selects the most specific address or longest address match (Bahaa Musawi et al. 2016).

We ran Sub-Prefix hijacking scenario through the network topology with the traffic generator turned on and running. Then, from the existing prefix, we created a new longest prefix that we can announce from any local router in the network, and the existing prefix was rerouted to the selected router. For example, an existing prefix IS (x.x.x.x/n), then we created a new prefix (x.x.x.x/n+1) and announced it in the BGP Networks from the selected gateway router.

After the attack is launched, the impact of the traffic can be monitored by Grafana NMS, where we can clearly see the effects of the attack on the gateways network BGP traffic, which illustrates the difference in network BGP traffic usage under normal and hijacking attack conditions. When we looking at network behavior from the gateway layer, we noticed that the gateway device IGW-G loses traffic related to the hijacking prefix, and the traffic is rerouted to another IGW, (in this case, IGW-H) BGP traffic will increase in this device due to the addition of hijacking traffic to the original device traffic, proving that the attack was successfully implemented and has an impact on network traffic (Azab et al. 2024). Additionally, we utilize the OpenBMP database and Grafana monitoring tool, all twenty features related to (AS-PATH) and (BGP update message) in the condition of (Hijacking) were monitored, visualized, and compared to the features in the network's normal conditions.

TABLE.4 BGP features from AS-PATH attribute and number of BGP updates

No	Feature Name	Definition
		BGP features from AS-PATH
1	Announcements to Longer/ Shorter paths	Counting AS -PATH length, then determine what is the max number (longer) and min number (shorter). provides insights into dynamics of AS-PATH changes in BGP network.
2	Average/Maximum AS Path Length	Determine the average AS -PATH Length or the Maximum number of AS-PATH length. This feature is useful for detecting if any hijacking method successfully
3	Average/Maximum unique AS path length	Almost same with the feature number 2 except count the unique value of AS Number in AS -PATH.
4	Average/Maximum edit distance	The number of edits required (add, delete, change) to make two strings equal.
5	Number of rare ASes.	The average value of Total AS Number in AS-PATH, this average will be the threshold of what kind we called "rare".
6	Average/maximum number of rare ASes	Calculated based on previous feature; the differences are only what data aggregation is calculated. helpful in identifying potential BGP route hijacks.
7	Prefix origin change	Detect the changes of origin AS, one of the more impacted features by cyberattack to determine the anomaly of Hijacking Attack.
8	AS-PATH changes according to geographic location	Compares the change of AS -PATH with the geographic location based on AS Number information. it provides valuable insights into the routing behavior at different geographical regions.
9	Number of new paths announced after withdrawing an old path.	Calculating by collecting data prefix, current A, and previous AS-PATH, then we count the prefix that announce a path after withdrawing an old path.
10	Number of new-path announcements	Collects prefixes and their unique AS -PATH in the time range (last modified), then counts only the prefixes that announce the new AS-PATH.
		BGP features from the number of BGP update
11	Number of announcements/ Withdrawals	Counting the number of announcements and withdrawals from (ip_rib_log) table. The Fluctuations indicate the changes in network reachability
12	Number of duplicate announcements / withdrawals	Calculating based on the previous feature, with the additional logic to detect whether any of the announcement or withdrawal are duplicate.
13	Number of new announcements after withdrawal	Calculates based on if there is any withdrawal of a prefix, and then that prefix is announced back, how many times it announced.

continue ...

... cont.

14	Number of plan new announcements	Calculates the number of announcements with 0 times withdrawals, meaning that prefix is plain new announcements in the time of range.
15	Number of implicit withdrawals with same/different path	Creates the new field to state whether an AS-PATH is the same path or different as before in the range of time (last modified).
16	Number of IGP/EGP/ INCOMPLETE messages	Counts the number of each origin type (IGP/EGP/INCOMPLETE). The origin types represent how the prefix is announced.
17	Number of ORIGIN changes	Counts when announcement method is changed, there are two origin types used as a normal origin for BGP Prefix, IGP and Redistribute.
18	Number of announced prefixes	Counts the number of prefixes that announced by IGW Routers, this will consist of Local ISP Prefixes and Internet / Global ISP Prefixes.
19	Maximum/average announcements per prefix	Indicates the maximum and average values of the general number of updates across prefixes.
20	Maximum/average announcements per AS	Similar to the previous features, but with respect to ASes.

According to the scenario of hijacking, after verifying and comparing all twenty BGP features, it was discovered that there are a total of 6 features (feature numbers 1, 6, 8, 11, 17, and 18) that have a significant difference based on contrasting the graphic patterns between normal and hijacking conditions, two of the impacted features are shown in Figure 4 below as an example from Table 4, feature number One from the AS-PATH attribute and feature number Eleven from the BGP update message attribute.

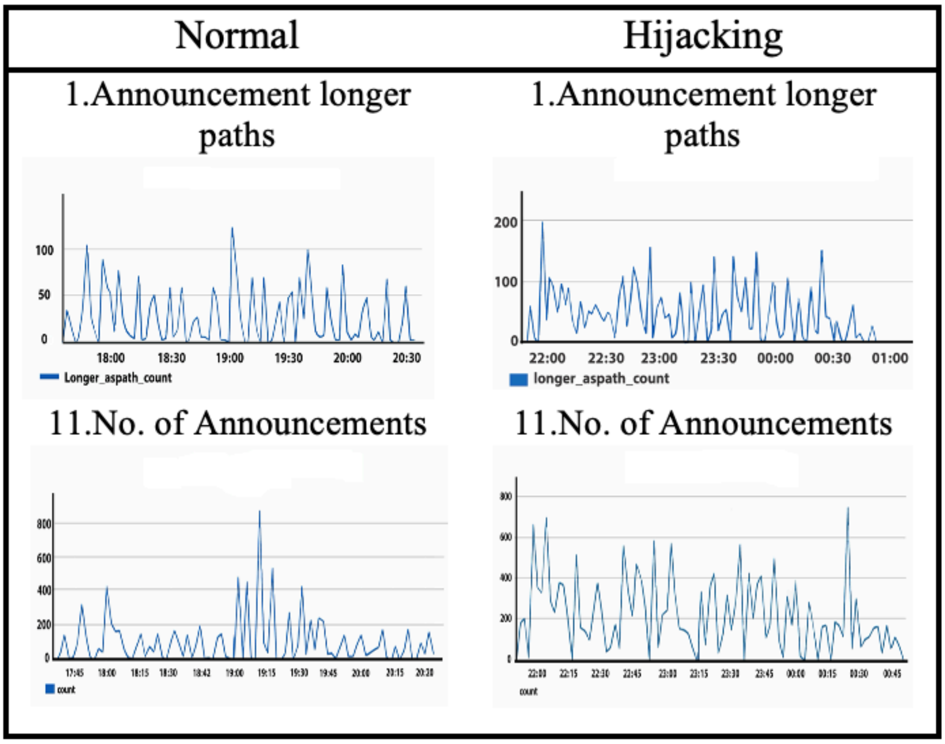


FIGURE 4. Normal & Hijacking BGP feature comparison

DOS (BLACKHOLING) CONDITION

Denial of Service is the most common Cyber Attack in BGP Network or Internet Infrastructure. Denial of Service has a significant impact on network performance and availability (Farasat & Khan 2020). One of the techniques that emerged to mitigate this kind of attack is called BGP Blackholing, i.e., dropping traffic to a destination (prefix) (Copstein et al. 2020). This scenario is used to simulate a Denial-of-Service attack against our resources in the local ISPs layer. The response to this attack is to enable Route Blackhole for the targeted resources. The impact of this response will be outages for the traffic destined to the resources, but it will effectively stop DOS attacks from flooding the local ISP environment.

All twenty features related to (AS-PATH) and (BGP update message) in the condition of (DOS) were monitored, visualized, and compared to the features in the network's normal conditions, After verifying and comparing all twenty BGP features, we found that nine of them (features 1, 6, 7, 10, 11, 14, 16, 17, and 18) have a significant difference based on contrasting the graphic patterns between normal and (DOS) conditions and are more impacted by this kind of cyber-attack. Figure 5 below shows feature number Six from the AS-PATH attribute and feature number 16 from the BGP update message attribute Table 4, as an example.

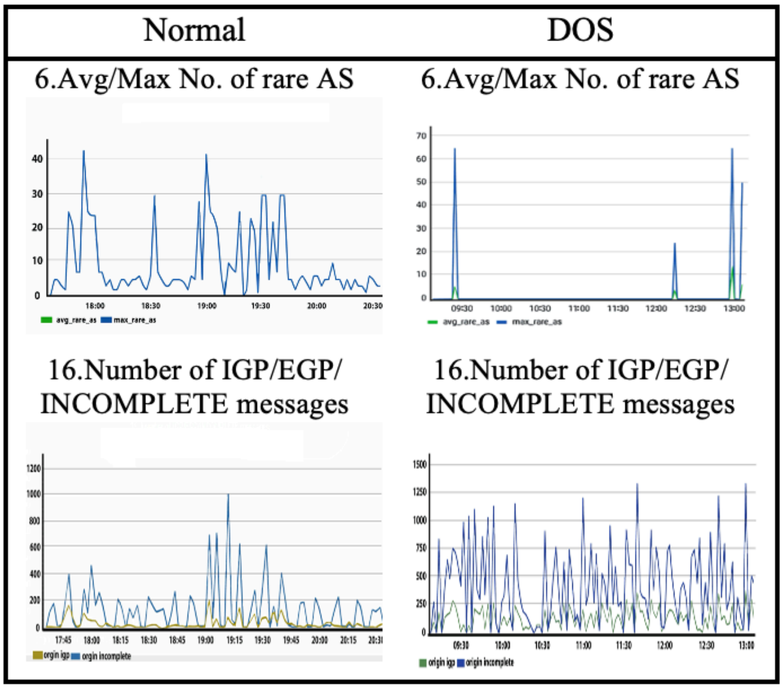


FIGURE 5. Normal & DOS BGP feature com parison

OUTAGES CONDITION

Previous network outages and security analyses involving BGP have shown that the current Internet routing infrastructure is extremely vulnerable. BGP's limited guarantees can contribute to significant instability and outages.

The goal for implementing the attack is to show its effect on the Gateway Border Network / IGW Router behavior. We can recognize which gateway more impacts regarding failover, time to effect, traffic flapping, etc. Based on the geographical map of Iraq and the BGP infrastructure, this scenario is divided into four parts. The first outage scenario is to shut down the EAST internet border gateways ((All border gateways in the east of Iraq with east neighboring countries are shut down)), the second outage scenario is to shut down all the WEST internet border gateways. The same applies for the third and fourth scenarios, which are to shut down the NORTH border gateways and the SOUTH border gateways ((All border gateways of Iraq with the north and south neighboring countries are shut down, respectively.).

1. Outage all East border gateways: In our simulated topology, the Iraqi east side borders include IGW routers connected to (Iran) neighbor routers; these devices are IGW - (F, G, H, I). These routers were turned off, and we tested the network conditions and monitored IGW traffic on the remaining operating gateways. Af-

ter this scenario applied, the traffic got lost in all the EAST side borders, and by using Grafana monitoring tools, we examined the traffic of all other IGWs and found that BGP traffic from the east borders is rerouted to IGW-C & D (North side borders) because these two routers have more peers than the other devices in the network topology.

2. Outage all West border gateways: The gateways network on the west side includes all IGW routers that are linked to neighboring countries routers in Syria, Jordan, and Saudi Arabia. In this scenario, traffic is distributed to IGW-E, L, and K, respectively. When the attack lunched, we did notice that traffic via IGW-E (with the Syria border) was not rerouted because the ISPs connected to this border do not have backup peering with other gateways, as is the actual network topology. while the other gateways BGP traffic is routed to different IGWs that are connected to different neighboring countries.
3. Outage all North border gateways: All IGW routers connected to Turkey's neighboring country (IGW-A, B, C, D) are located on the network's northern side. These routers were shut down according to this scenario, and we then tested and verified network behavior and BGP traffic route changes. In this condition, all traffic capacity is rerouted to IGW-F, G, and H on the eastern network side, which borders Iran.
4. Outage the south border gateway: The south network side includes the cable landing station and one IGW router connected to submarine Arab Gulf cables (IGW-J). To apply this part of the outage scenario, this router was shut down using the PNETLAB platform. The BGP traffic was rerouted to the (IGW-H) on the eastern border with Iran's side after the attack was launched.

All twenty features related to (AS-PATH) and (BGP update message) under the condition of (Outage 4 parts) were monitored, visualized, and then compared to features of the network under normal conditions. There are a total of eight features with the numbers (1,4,6,10,14,15,16,17) that have a significant change based on contrasting the graphic patterns between normal and outage conditions; these features are more impacted by this type of cyberattack. Two graphs comparing the impact of the features are displayed in Figure 6 below. Feature number Four from the AS-PATH attribute, and feature number Fourteen from the BGP update message attribute in Table 4.

DATASETS COLLECTION AND ANALYSIS

Each anomaly detection method requires its own feature extraction and dataset generation because there are no publicly available datasets in a format suitable for feeding ML detection models. (Fonseca, Paulo et al. 2019). for this reason, we added visualization to all twenty features during feature extraction by using the Grafana tool. This tool displays the peak detection in the time series of each feature graph, assisting in identifying the most affected features in each case of (cyber-attack condition) versus normal condition and help to create proper data sets.

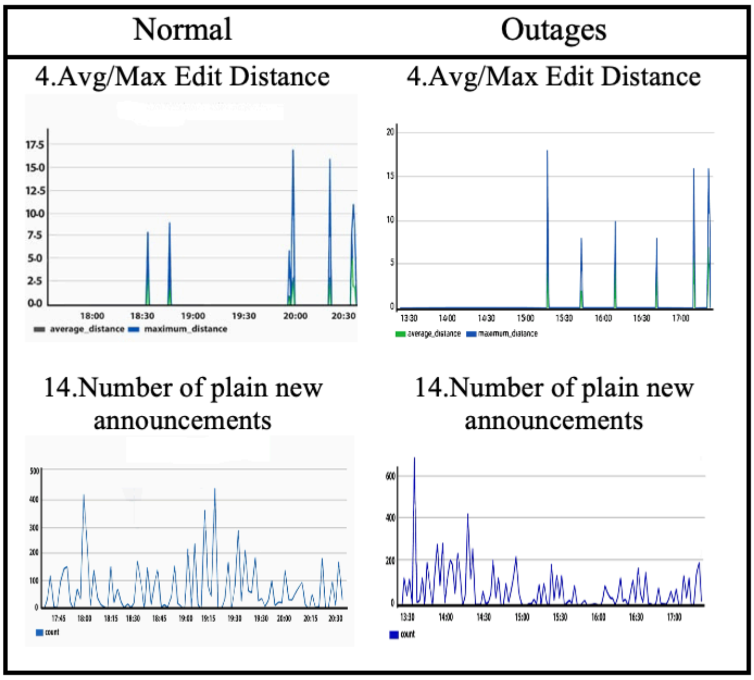


FIGURE 6 Normal & Outage BGP features comparison

DATASETS COLLECTION AND ANALYSIS

TABLE 5. Simulator's operating hours and samples

Normal scenario	Anomaly scenario
16hours x 60 samples/hour = 960 samples	Hijacking: 6 hours x 60 samples/hour = 360 samples DOS: 6 hours x 60 samples/ hour = 360 samples Outages: 6 hours x 60 samples/hour = 360 samples Total: 1,080 samples
Total: 2,040 samples	

We extracted the desired twenty features from the raw BGP AS-PATH and BGP update messages Table 4, by running the simulation program (PNETLAB) and producing a sampling of normal and anomaly (cyber-attacks) conditions. Three anomaly events in addition to the normal condition were utilized to generate four data sets based on one-minute time interval per sample considering the number of working hours of the simulator, as presented in Table 5.

Based on the feature samples collected at per minute intervals, we organized data sets tables as inputs to the ML binary classification algorithms for regular (normal condition) for BGP AS-PATH and BGP update message features as well as (Hijacking, DOS, Outages) datasets.

DISCUSSION

Understanding Border Gateway Protocol (BGP) weaknesses and identifying potential cyber-attack scenarios are crucial for better protecting this vital infrastructure (Fonseca, Paulo et al. 2019). The BGP protocol security enhancement has been the subject of numerous attempts in the past, but BGP is still susceptible to a wide variety of anomalous events (Scott et al. 2024). BGP anomalies are rare, but they can cause great damage when they occur. Identifying BGP anomalies is a challenging task where BGP traffic has been characterized as voluminous and complex. Furthermore, unstable BGP traffic has the effect of masking anomalous traffic (W. Shao,2021).

In our work, we focused on analyzing BGP traffic behavior based on real-world network data and extensive network testing. We examined BGP traffic under three types of BGP cyberattack scenarios. The BGP features were extracted from the BGP database of the simulator. PostgreSQL scripts were used to get (10) features from the BGP AS-PATH attribute and (10) features from the BGP UPDATE message attribute. As per the data produced by simulation programs, the AS-PATH attribute contains three features that are categorized into two sections of the dataset labeled as "average" and "maximum.". Furthermore, there is one feature divided into two areas, specifically labeled "longer" and "shorter.". Therefore, the AS-PATH comprises a total of 14 feature datasets.

To assess the significance of features in a binary classification context and predict a binary outcome (Normal, Anomalous) based on extracted features, Statistical analysis was employed to identify the most impactful features by the cyber-attacks, we used ANOVA-Test and generated the F-Statistic and the p-values. Based on the findings of statistical analysis (p-value less than (0.005), it has been observed that Six of BGP features are most significant by BGP cyber-attacks' effect within the range of F-statistic values (F > 90), the second group of Nine features showed moderate significance by range of F-statistic values (30 < F < 90), and the third group of Seven showed a low significance than the rest of the features (F < 30). While Two features were shown to be unaffected by cyber-attacks; the p-value is greater than (0.005). As a result, the outcomes of the extracted BGP features and the degree to which cyberattacks influence them would support the development of an accurate and efficient ML model for detecting BGP anomalies.

CONCLUSION

This study introduces a new way to use network simulations based on realistic network considering all the actual topological parameter of the (Iraqi Internet Gateways Network) to create datasets for normal and different BGP cyberattack scenarios. Twenty BGP features related to the AS-PATH and the number of BGP update attributes are extracted to format four groups of (24) datasets. By combining network simulation techniques with advanced software, we were able to visualize the BGP features that were more impacted by each cyberattack scenario (Hijacking, DOS, Outage) determined by comparing the patterns in the graphics.

The generated and formatted four data sets were analyzed statistically to show the most significant extracted features in terms of binary classification and determining whether the BGP traffic is normal or anomalous. Our findings will contribute to the creation of an accurate and efficient ML model for identifying BGP abnormalities in our next work. Furthermore, this approach may be regarded as a proactive procedure for internet network operators, allowing them to detect and address anomalies rapidly and effectively.

ACKNOWLEDGEMENT

Thanks go to the Iraqi Ministry of Communication (MOC) - ITPC company-Iraq for their assistance to gather information and data of this manuscript.

DECLARATION OF COMPETING INTEREST

None

REFERENCES

Allahdadi, A. & Morla, R. & Prior, R. 2017. A Framework for BGP Abnormal Events Detection.

Alotaibi, H. S., Gregory, M. A., Li, S. & Ali, I. 2022. Multidomain SDN-based gateways and border gateway protocol. *Journal of Computer Networks and Communications* 2022: 1-23.

Andra Lutu & Bagnulo, M. & Cid-S., Jesús & Maennel, O. 2014. Separating wheat from chaff: Winnowing unintended prefixes using machine learning. *Proceedings - IEEE INFOCOM*. 943-951.

Aceto, G. & Botta, A. & Marchetta, P., Persico, V. & Pescapè, A. 2018. A comprehensive survey on internet outages. *Journal of Network and Computer Applications* 113. 10.1016/j.jnca.2018.03.026

Azab, A., Khasawneh, M., Alrabaee, S., Choo, K.-K. R. & Sarsour, M. 2024. Network traffic classification: Techniques, datasets, and challenges. *Digital Communications and Networks* 10(3): 676-692.

Bahaa Musawi & Branch, P. & Armitage, G. 2016. BGP anomaly detection techniques: A survey. *IEEE Communications Surveys & Tutorials*. PP. 1-1. 10.1109/COMST.2016.2622240.

Copstein, R. & Zincir-Heywood, A. 2020. Temporal Representations for Detecting BGP Blackjack Attacks. 1-7. 10.23919/CNSM50824.2020.9269055.

Dewo, K., Yasin, V., Budiman, T., Sianipar, A. & Yulianto, A. 2023. IT infrastructure dashboard monitoring application development using Grafana and Promotheus. A case study at Astra Polytechnic School.

Farasat, T. & Khan, A. 2020. Detecting and analyzing border gateway protocol blackholing activity. *International Journal of Network Management* 31.

Fonseca, P. & Mota, E. & Bennesby, R. & Passito, A. 2019. BGP dataset generation and feature extraction for anomaly detection. 1-6. 10.1109/ISCC47284.2019.8969619.

Goldberg, S. 2014. Why is it taking so long to secure internet routing? *Communications of the ACM* 57: 56-63. DOI: 10.1145/2659899.

H. Geoff. 2022. BGP in 2022 – the routing table". https://blog.apnic.net/2023/01/06/bgp-in-2022-the-routing-table/.

H. Guo, S. Gao and H. Zhang. 2009. Inter-domain routing with AS number: A traffic engineering perspective. 2009 International Symposium on Computer Network and Multimedia Technology, pp. 1-4, doi: 10.1109/CNMT.2009.5374756.

H. Kalra, A. P. Singh and D. Sadhya. 2021. Anomaly detection in Border Gateway Protocol using supervised machine learning. 2021 IEEE Bombay Section Signature Conference (IBSSC), Gwalior, India, 2021, pp. 1-6, doi: 10.1109/IBSSC53889.2021.9673281.

Haeberlen, A. & Avramopoulos, I., Rexford, J. & Druschel, P. 2009. NetReview: Detecting when interdomain routing goes wrong. 437-452.

Hammood, N. & Musawi, B. 2021. Using BGP features towards identifying type of BGP anomaly.

J. Mai & Yuan, Lihua & Chuah, Chen-Nee. 2008. Detecting BGP anomalies with wavelet. 465 - 472. 10.1109/NOMS.2008.4575169.

John E. Woods, Gerald Henry Blake, Hugh Kennedy, Majid Khadduri, Richard L. Chambers, Utta Egghe, Jason Jason, 2023. https://www.britannica.com/place/Iraq

K. Hoarau, P. U. Tournoux and T. Razafindralambo. 2021. BML: An Efficient and Versatile Tool for BGP Dataset Collection. 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada.

Kevin Hoarau, P. U. Tournoux and T. Razafindralambo. 2022. Detecting forged AS paths from BGP graph features using Recurrent Neural Networks. IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA.

Kassim Mohammed AL-Hassani. 2019.https://itig-iraq.iq/wp-content/uploads/2019/05/Iraq-Submarine-cables-article-v2.pdf.

Marijana Cosovic & Slobodan, Obradovic. 2018. BGP anomaly detection with balanced datasets. *Tehnicki Vjesnik* 25: 766-775. 10.17559/TV-20170219114900.

M. Cheng, Q. Li, J. Lv, W. Liu and J. Wang. 2018. Multi-Scale LSTM Model for BGP Anomaly Classification, 2018. *IEEE Transactions on Services Computing* 14(3): 765-778.

Marijana Cosovic, & Slobodan, Obradovic & Junuz, Emina. 2017. Deep Learning for Detection of BGP Anomalies.

N. Al-Rousan, & Trajkovic, L. 2012. Machine learning models for classification of BGP anomalies. 2012 IEEE 13th International Conference on High Performance Switching and Routing, HPSR.

Paiva, T. & Siqueira, Y., Batista, D. & Hirata, R. & Terada, R. 2021. BGP Anomalies classification using features based on as relationship graphs. 1-6. 10.1109/LATINCOM53176.2021.9647824.

Pasquale Chiacchio, P., Basile, F. & Carbone, C. 2007. Simulation and analysis of discrete-event control systems based on Petri nets using PNetLab. *Control Engineering Practice - CONTROL ENG PRACTICE*.

Rahul Verma & Govil, Mahesh & Keserwani, Pankaj. 2023. ELM based Ensemble of Classifiers for BGP Security against Network Anomalies. 1-6. 10.1109/ESDC56251.2023.

Sanchez, Odnan Ref & Ferlin-Reiter, S & Pelsse, C. & Bush, R. 2019. Comparing machine learning algorithms for BGP anomaly detection using graph features. 35-41. 10.1145/3359992.3366640.

Scott, B., Johnstone, M., Szewczyk, P. & Richardson, S. 2024. Matrix Profile data mining for BGP anomaly detection. Computer Networks 242: 110257.

Shivani Deshpande & Thottan, M. & Sikdar, B. 2008. An online scheme for the isolation of BGP misconfiguration errors. *IEEE Transactions on Network and Service Management* 5: 78-90. 10.1109/TNSM.

Shivani Deshpande & Thottan, M., Ho, T. & Sikdar, B. 2009. An online mechanism for BGP instability detection and analysis. *IEEE Transactions on Computers* 58: 1470-1484. 10.1109/TC.

Takhar, H. K. & Trajković, L. 2023. BGP features and classification of internet worms and ransomware attacks. *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*.

W. Shao, 2021. BMP Server LAB. https://medium.com/@a13089/bmp-server-lab-3799d68e7941.

Wang, Cun & Li, Zhengmin & Huang, Xiaohong & Zhang, Pei. 2016. Inferring the average as path length of the Internet. 391-395. 10.1109/ICNIDC.2016.7974603.

X. Dai & Wang, Na & Wang, Wenjuan. 2019. Application of machine learning in BGP anomaly detection. *Journal of Physics: Conference Series* 1176: 032015. 10.1088/1742-6596/1176/3/032015.

Z. Li & Gonzalez Rios, Ana & Trajkovic, Ljiljana. 2020. Detecting internet worms, ransomware, and blackouts using recurrent neural networks. 2165-2172. 10.1109/SMC42975.2020.9283472.

Zhao, Xi & S. Band, Shahab & Elnaffar, Said & Sookhak, Mehdi & Mosavi, Amir & Salwana, Ely. 2021. The implementation of border gateway protocol using software-defined networks: A systematic literature review. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2021.3103241.