Beneath the Facade of IP Leasing: Graph-Based Approach for Identifying Malicious IP Blocks

Zhenni Liu[1,2], Yong Sun[1,2(✉)], Zhao Li[1,2], Jiangyi Yin[1,2], and Qingyun Liu[1,2]

[1] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
{liuzhenni, sunyong, lizhao, yinjiangyi, liuqingyun}@iie.ac.cn

Abstract. With the depletion of IPv4 address resources, the prevalence of IPv4 address leasing services by hosting providers has surged. These services allow users to rent IP blocks, offering an affordable and flexible solution compared to traditional IP address allocation. Unfortunately, this convenience has led to an increase in abuse, with illegal users renting IP blocks to host malicious content such as phishing sites and spam services. To mitigate the issue of IP abuse, some research focuses on individual IP identification for point-wise blacklisting. However, this approach leads to a game of whack-a-mole, where blacklisted IPs become transient due to content migration within the IP block. Other studies take a block perspective, recognizing and classifying IP blocks. This enables the discovery of potentially malicious IPs within the block, effectively countering service migration issues. However, existing IP block identification methods face challenges as they rely on specific WHOIS fields, which are sometimes not updated in real-time, leading to inaccuracies. In terms of classification, methods rely on limited statistical features, overlooking vital relationships between IP blocks, making them susceptible to evasion. To address these challenges, we propose BlockFinder, a two-stage framework. The first stage leverages the temporal and spatial stability of services to identify blocks of varying sizes. In the second stage, we introduce an innovative IP block classification model that integrates global node and local subgraph representations to comprehensively learn the graph structure, thereby enhancing evasion difficulty. Experimental results show that our approach achieves state-of-the-art performance.

Keywords: IP blocks detection · Graph representation learning.

1 Introduction

In recent years, IPv4 address leasing services offered by hosting providers have become increasingly prevalent. These services often involve the rental of consecutive IP addresses, known as "IP Blocks," for hosting Internet services. However, recent research [1] has highlighted a challenge where illegal users exploit these IP blocks, frequently migrating malicious services within them, making it difficult for network authorities to efficiently identify malicious IPs.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

To tackle this issue, we propose BlockFinder, an automated framework designed for the identification of malicious IP blocks. BlockFinder focuses on identifying malicious IPs at the block level, effectively combating the issue of malicious service migration within IP blocks. The framework consists of two stages:

IP block identification. According to the previous work [6], the IP blocks leased from hosting providers may not be meticulously recorded by RIRs as with IP allocations. Relying on public data, such as IP WHOIS records, to obtain IP blocks is not practical. Therefore, we propose a novel method based on the service stability of IP blocks. This approach ensures that all IPs within each IP block belong to the same entity, rather than roughly treating the entire Autonomous System (AS) IP address space as a single IP block.

IP block classification. Current IP block classification methods [1] often rely on passive flow statistics, which are susceptible to evasion. We propose a model that utilizes a comprehensive analysis of statistical features and graph-based behaviors. For graph-based detection, we leverage the observation that services migrate between different blocks, and IP block subgraphs reveal communication patterns. Specifically, we integrate a novel combination of node and subgraph representation, enhancing identification effectiveness even in scenarios with isolated nodes.

2 Related Work

Existing methods for identifying malicious IPs can be categorized into those that recognize from the perspective of individual IPs and those that identify from the perspective of IP blocks.

Individual IP perspective methods, like those by Alvarez et al. [2] and Coskun et al. [3], cluster IPs based on communication destination to find similarities with blacklisted IPs. However, these methods may not promptly detect the migration of malicious content within blocks.

IP block perspective methods include AS or hosting provider reputation-based approaches and IP block reputation-based methods. The former calculates a maliciousness score using meta-information from AS or hosting provider IPs [5, 8], but may miss smaller abused IP blocks. IP block reputation methods evaluate services' maliciousness within smaller IP blocks. For instance, Alrwais et al. [1] identified sub-allocated IP blocks from reputable hosting providers exploited for hosting malicious content, using IP WHOIS and PDNS. However, challenges arise due to outdated WHOIS records [6].

The rise in hosting malicious content on IP blocks from reputable providers underscores the need for effective identification methods. Yet, existing WHOIS-based methods face limitations due to delayed updates [6]. Meanwhile, relying solely on PDNS for IP block classification lacks data diversity and is vulnerable to evasion. Thus, more effective identification methods are necessary.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

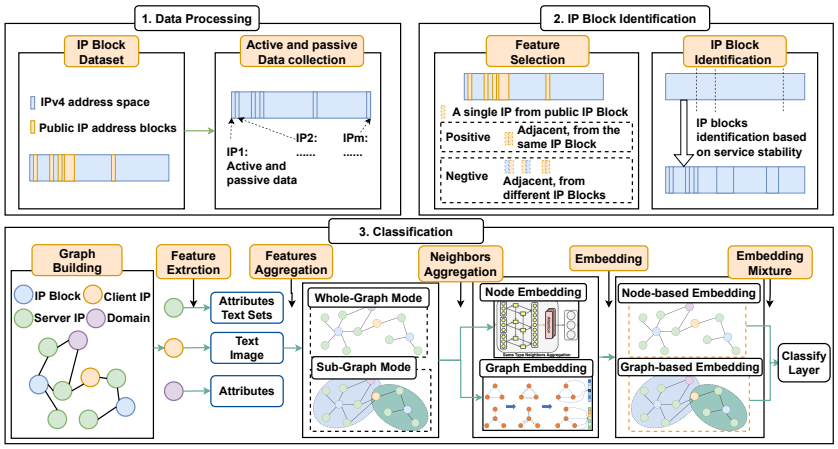

Fig. 1: The architecture of BlockFinder.

3 Approach

3.1 IP Block Identification

We propose a technical approach to achieve fine-grained IP block identification based on the service stability of IP blocks. Through extensive observation of both active probing and passive traffic data, we have identified that the stability of services within the IP block is inherently maintained, demonstrating in two aspects:

- **Temporal stability.** Significant changes in services within the block are uncommon over short periods. As shown in the Fig.2, on IP Block1 consisting of two consecutive IP addresses 108.*.*.249[3] and 108.*.*.249, the service set {zimm.*.com, harri.*.com} is observed at T1, and {harri.*.com, zimm.*.com} is observed at T2. Despite dynamic services changes on individual IPs, the overall service set of the entire IP block remains the same, indicating high stability.
- **Spatial stability.** The service set distribution within an IP block remains stable. Distinct service sets are observed on different IP blocks. As depicted in Fig.2, the service set {zimm.*.com, harri.*.com} consistently resides on IP Block1, while IP Block2 hosts the stable service set {luc.*.com, weiss.*.com}.

Based on the description, we define the stability contribution resulting from dividing two IPs into the same IP block as $W(IP_p, IP_q)$, shown in Eq.1. Here, $Sim(IP_p, IP_q)$ represents the similarity of their service sets, and $I(IP)$ represents the information amount of the service set. $S(i, j)$ denotes the sum of stability contributions brought by each IP in the IP block. The objective of IP block identification is to determine division locations in the continuous IP address space to maximize the overall stability value of identified IP blocks.

[3] In order to protect user privacy, we use "*" to represent key locations.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

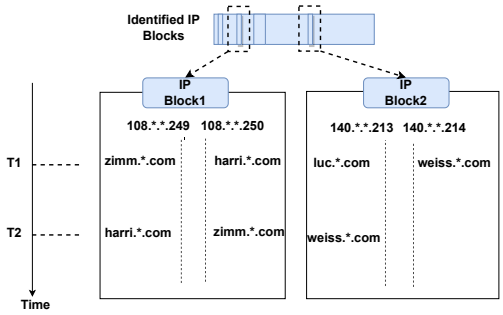

Fig. 2: Service stability.

$$W(IP_p, IP_q) = \begin{cases} Sim(IP_p, IP_q) & I(IP_p) > \alpha_i, I(IP_q) > \alpha_i, \\ & Sim > \alpha_s \\ 0 & I(IP_p) < \alpha_i or I(IP_q) < \alpha_i \\ punish(punish < 0) & I(IP_p) > \alpha_i, I(IP_q) > \alpha_i, \\ & Sim < \alpha_s \end{cases} \tag{1}$$

$$S(i,j) = \frac{\frac{1}{2}\sum_{p\in[i,j]}\sum_{q\in[i,j]} W(IP_p, IP_q)}{j-i+1}, (p \neq q) \tag{2}$$

3.2 IP Block Classification

Based on the obtained IP block in the previous section, the constructed heterogeneous graph is depicted in Fig. 3. Building upon this graph, we introduce the HGNT-Net algorithm (Heterogeneous Graph Node and Topology representation learning Network). This algorithm advocates for incorporating both the node representation of IP blocks and the heterogeneous subgraph topology centered around IP blocks during the learning phase.

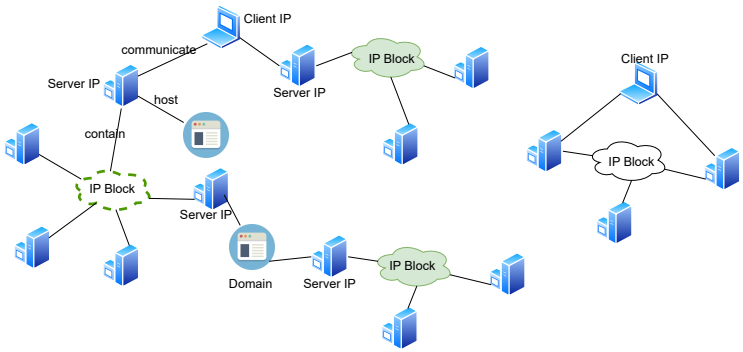

Fig. 3: Example of IP Block Graph.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

Node representation. To effectively extract node representations, an attention mechanism is introduced. For instance, nodes with more communication sessions and longer durations should have higher weights, as shown in Eq.4. Here, || denotes feature concatenation, and the final node representation is determined as shown in Eq.5.

$$Session(i,j) = [sessions, duration\_ms, c2s\_pkt, s2c\_pkt, c2s\_byte, s2c\_byte] \quad (3)$$

$$e_{ij}^{lk} = \begin{cases} a([Wh_i \| Wh_j]), j \in \mathcal{N}_i \ k \in [1,2] \\ Session(i,j), j \in \mathcal{N}_i \quad k = 3 \end{cases} \quad (4)$$

$$h_i^{l+1} = \sigma(\|_{k=1}^{K} \sum_{j \in \mathcal{N}_i} \frac{exp(LeakyReLU(e_{ij}^{lk}))}{\sum_{q \in \mathcal{N}_i} exp(LeakyReLU(e_{iq}^{lk}))} W_k^l h_j^l + W_0^l h_i^l) \quad (5)$$

Subgraph representation. The heterogeneous subgraph representation learning begins by extracting the heterogeneous subgraph centered on each IP block. This subgraph exhibits unique topology and content characteristics. Content representation h_{gc} focuses on the content of all relevant nodes in the IP block subgraph. The topology representation h_{gt} primarily depicts the network structure of the IP block subgraph, including its degree, network density, diameter, and clustering coefficient. Finally, the node representation and subgraph representation are merged and fed into the classification layer, as shown in the Eq.6.

$$h = \beta h_n \| \gamma(h_{gc} \| h_{gt}) \quad (6)$$

4 Experimental Evaluation

4.1 IP Block Identification

We collected a total of 999 IP blocks from FireHOL[4] IP List within the 122 /24 IP prefixes. Since IP block identification results are represented as ranges and cannot be evaluated using traditional accuracy metrics, we utilized modified evaluation indicators commonly used in clustering scenarios: the IP block outline coefficient (S-Block) and Davies-Bouldin coefficients (DB coefficients), as shown in Eq.7 and 8. Additionally, we employed the coverage indicator Coverage-Block to assess the degree of deviation in IP address identification results by comparing test IP blocks with verified block intervals. We compared our method with a classical IP block identification approach based on signal pulses [9], which relies on the observation that IPs within the same block exhibit closely connected feature distances among adjacent IPs.

$$S - Block = \frac{b-a}{\max(a,b)} \quad (7)$$

$$DB - Block = \frac{1}{n} \sum_{i=1}^{n} \max(\frac{\sigma_i + \sigma_j}{dist(c_i, c_j)}) \quad (8)$$

Our method identified a total of 968 IP blocks within the same 122 /24 IP prefixes. The S-Block value for our result is 0.55, and the DB-Block is 8.7. In

[4] FireHOL. https://firehol.org/.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

contrast, the S-Block for the signal pulse method is 0.31, with a DB-Block of 11.9. This demonstrates that our method achieves better cohesion within identified IP blocks and better separation between blocks. The coverage results are provided in Table 1. Here, ValBlock represents the collected IP blocks, TestBlock represents the IP blocks identified by our method, and ValBlock rate represents the proportion of total ValBlocks. It's observed that nearly 90% of the ValBlocks are entirely encompassed within the TestBlock. This is primarily because ValBlock is manually collected and may not provide comprehensive coverage. Our method offers broader observation capabilities, hence including more IPs.

Table 1: Coverage-Block results.

Coverage Type	Coverage-Block	ValBlock Rate
$ValBlock \subseteq TestBlock$	100%	86.99%
$TestBlock = ValBlock$	100%	1.9%
$TestBlock \subseteq ValBlock$	59.09%	2.4%
$ValBlock \cap TestBlock \subseteq ValBlock$	61.5%	10.91%

4.2 IP Block Classification

The labeled IP block dataset is derived from the 968 IP blocks identified in Section 4.1. Among these, 300 IP blocks were labeled as malicious by FireHOL, while 100 blocks were labeled as benign. Table 2 shows the number of nodes and edges associated with the aforementioned 968 IP blocks. Evaluation indicators are presented in Eq.9, 10 and 11.

Table 2: Heterogeneous graph nodes.

Node Type	Node Count	Edge Type	Edge Count
ip_block	968	IP_block-server_ip	9399
server_ip	9399	server_ip-client_ip	14350
client_ip	14350	server_ip-domain	6485
domain	5251	-	-

$$Specificity = Recall@neg = \frac{TN}{TN + FP} \tag{9}$$

$$Precision@neg = \frac{TN}{TN + FN} \tag{10}$$

$$F1@neg = \frac{2 \cdot Precision@neg \cdot Recall@neg}{Precision@neg + Recall@neg} \tag{11}$$

To comprehensively verify the effectiveness of HGNT-Net, we conducted two sets of experiments. The first set compares our approach with classic heterogeneous graph node classification algorithms: RGCN [7] and HGT [4]. RGCN is a

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

foundational work on basic heterogeneous graph node classification, while HGT introduces an attention mechanism. Our algorithm, HGNT-Net, integrates both node and subgraph representation. The second set comprises ablation experiments, comparing: 1) using only node features without considering relationships between nodes; 2) considering relationships between nodes without incorporating subgraph topology; and 3) integrating both relationships between nodes and subgraph topology.

Experimental results are summarized in Table 3. Compared to RGCN, HGT introduces an attention mechanism allocating different attention scores to contributions of different paths. HGNT-Net not only integrates the attention mechanism for nodes but also incorporates the topological characteristics of subgraphs to further enhance model performance. Experiments demonstrate that our method outperforms others across various indicators.

Table 3: The experimental results of comparison algorithm.

Methods	Accuracy	Precision	Specificity	F1@neg	AUC
RGCN	0.88	0.76	0.76	0.82	0.85
HGT	0.83	0.81	0.8	0.74	0.9
Without using graph	0.77	0.75	0.75	0.77	0.82
Only node representation	0.85	0.83	0.83	0.73	0.86
HGNT-Net	**0.91**	**0.87**	**0.87**	**0.86**	**0.96**

4.3 Measurement

Using BlockFinder, we conducted measurements on the top 5 IP address spaces of hosting provider M247. Initially, we identified IP blocks within the target IP address spaces using our method based on service stability. Subsequently, we updated the heterogeneous graph based on the identified IP blocks. Finally, we utilized HGNT-Net to identify malicious IP blocks within these spaces. In total, we identified 4,672 IP blocks out of the 5 IP address spaces, comprising a total of 300,000 IPs, with 1,101 of them classified as malicious. We systematically measured the size and utilization rate of these malicious blocks. Our findings revealed that the average size of malicious IP blocks is approximately 32, ensuring swift evasion without arousing suspicion due to excessively large malicious scope. The utilization rate of nearly half of the IP blocks is only 40%, indicating underutilization of IPs within the blocks. In the event of an IP being blacklisted, services can be migrated to other IPs within the same block, thus prolonging the service's survival time. The measurements were conducted on a CPU AMD Ryzen 5 5600H with Radeon Graphics @ 3.30 GHz. Since IP block leases from cloud hosting providers typically have minimum durations, such as one month, IP blocks usually do not change frequently. Additionally, the number of IP block nodes is typically around 1/100 of individual IP nodes. Due to the relatively

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6

infrequent updates of IP blocks and the significant reduction in scale compared to individual IPs, efficient periodic recognition can be achieved.

5 Conclusion

In this paper, we introduce BlockFinder, a novel framework for identifying malicious IP blocks. Initially, we employ an IP block identification method based on service stability to identify blocks of varying sizes. Subsequently, our classification model, HGNT-Net, enhances performance in scenarios with isolated nodes and limited information dissemination by integrating node and subgraph representations centered on IP blocks. The algorithm's effectiveness is demonstrated through ablation experiments and comparisons with classic algorithms. Our framework effectively addresses the challenge of malicious services migrating within IP blocks by detecting malicious IPs from a block perspective.

Acknowledgments. This work is supported by National Key R&D Program of China (Grant No.2021YFB3101001) and the Scaling Program of Institute of Information Engineering, CAS (Grant No.E3Z0191101).

References

1. Alrwais, S., Liao, X., Mi, X., Wang, P., Wang, X., Qian, F., Beyah, R., McCoy, D.: Under the shadow of sunshine: Understanding and detecting bulletproof hosting on legitimate service provider networks. In: 2017 IEEE Symposium on Security and Privacy (SP). pp. 805–823. IEEE (2017)
2. Cid-Fuentes, J.Á., Szabo, C., Falkner, K.: An adaptive framework for the detection of novel botnets. Computers & Security **79**, 148–161 (2018)
3. Coskun, B.: (un) wisdom of crowds: Accurately spotting malicious ip clusters using not-so-accurate ip blacklists. IEEE Transactions on Information Forensics and Security **12**(6), 1406–1417 (2017)
4. Hu, Z., Dong, Y., Wang, K., Sun, Y.: Heterogeneous graph transformer. In: Proceedings of the web conference 2020. pp. 2704–2710 (2020)
5. Konte, M., Perdisci, R., Feamster, N.: Aswatch: An as reputation system to expose bulletproof hosting ases. In: Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication. pp. 625–638 (2015)
6. Noroozian, A., Koenders, J., Van Veldhuizen, E., Ganan, C.H., Alrwais, S., McCoy, D., Van Eeten, M.: Platforms in everything: Analyzing {Ground-Truth} data on the anatomy and economics of {Bullet-Proof} hosting. In: 28th USENIX Security Symposium (USENIX Security 19). pp. 1341–1356 (2019)
7. Schlichtkrull, M., Kipf, T.N., Bloem, P., Van Den Berg, R., Titov, I., Welling, M.: Modeling relational data with graph convolutional networks. In: The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15. pp. 593–607. Springer (2018)
8. Shue, C.A., Kalafut, A.J., Gupta, M.: Abnormally malicious autonomous systems and their internet connectivity. IEEE/ACM Transactions on Networking **20**(1), 220–230 (2011)
9. Xie, Y., Yu, F., Achan, K., Gillum, E., Goldszmidt, M., Wobber, T.: How dynamic are ip addresses? In: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications. pp. 301–312 (2007)

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_6