DAbR: Dynamic Attribute-based Reputation scoring for Malicious IP Address Detection

Arya Renjan*, Karuna Pande Joshi†, Sandeep Nair Narayanan* and Anupam Joshi*
*Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County
†Department of Information Systems, University of Maryland, Baltimore County
Email: {arenjan1, karuna.joshi, sand7, joshi}@umbc.edu

***Abstract*—To effectively identify and filter out attacks from known sources like botnets, spammers, virus infected systems etc., organizations increasingly procure services that determine the reputation of IP addresses. Adoption of encryption techniques like TLS 1.2 and 1.3 aggravate this cause, owing to the higher cost of decryption needed for examining traffic contents. Currently, most IP reputation services provide blacklists by analyzing malware and spam records. However, newer but similar IP addresses used by the same attackers need not be present in such lists and attacks from them will get bypassed. In this paper, we present Dynamic Attribute based Reputation (DAbR), a euclidean distance based technique, to generate reputation scores for IP addresses by assimilating meta-data from known bad IP addresses. This approach is based on our observation that many bad IP's share similar attributes and the requirement for a lightweight technique for reputation scoring. DAbR generates reputation scores for IP addresses on a 0-10 scale which represents its trustworthiness based on known bad IP address attributes. The reputation scores when used in conjunction with a policy enforcement module, can provide high performance and non privacy-invasive malicious traffic filtering. To evaluate DAbR, we calculated reputation scores on a dataset of 87k IP addresses and used them to classify IP addresses as good/bad based on a threshold. An F-1 score of 78% in this classification task demonstrates our technique's performance.**

I. Introduction

Web traffic encryption is a widely adopted technique to protect users' privacy. However, it is reported that the number of attacks using encryption to conceal its malicious attack vectors [1] is also increasing at a fast rate. Zscaler [2] , a cloud-based information security company, reports that there is a 30% increase in the encrypted malicious traffic and many new attack payloads are being delivered over encrypted communication channels. Decryption of such web traffic for inspection is not only expensive in cost and time, they also invade users' privacy. This situation forces many of the existing technologies like deep packet inspection (DPI) impractical. Hence there is a critical requirement to use unencrypted features of web traffic for identifying and filtering attacks. IP address is one such feature which could be utilized for filtering out malicious traffic.

IP address based web traffic filtering is widely used as the first line of defense in many Intrusion Detection Systems (IDS) like Snort. Lists of bad or blacklisted IP addresses are available from a plethora of sources like FireHOL[1], Palo Alto[2], Cisco Talos[3], Spamhaus[4], VirusTotal IP reputation[5], etc. Some vendors even provide options to contest against the inclusion of certain IP addresses in their lists, if its owner feels that their IP is wrongly listed. A major shortcoming of using such lists for filtering attacks is that many of them are based on known and reported malicious incidents. For example, if an IP address is found in a malware specimen, they are added to such blacklists. A known bad IP address from such lists arriving at a filtering gateway is currently filtered out by simple address matching. However, attackers often take control of newer targets and use them for spreading malware or initiating attacks. Such cases are overlooked by these filtering systems.

We envision an IP reputation system which sits at a firewall or a gateway and works in conjunction with a policy enforcement module. The policy enforcement module can either filter out malicious traffic or add them as candidates for further inspections. We identified several practical design considerations for such systems. Firstly, it should be capable of handling a large number of requests because even mid-sized enterprises and websites process several thousands of connection requests. Secondly, such a system should be able to adjust fast to the dynamic threat landscape. The frequently updated bad IP address lists are attestations to these dynamic environments. Thirdly, many attacks are targeted and are sometimes socially or politically motivated. A side effect is that a model generated for one organization or enterprise may not be applicable to another.

In this paper, we propose a simple and computationally fast reputation scoring technique, DAbR (Dynamic Attribute-based Reputation), for scoring unknown IP addresses which are not present in existing bad IP address lists. DAbR ingests the lists of bad IP addresses from different sources, extracts various meta-characteristics about them from intelligence sources like MaxMind, and generates a model for bad IP addresses. When an IP address arrives, our proposed system tags a reputation score with it, by calculating the similarity of its meta-characteristics to the generated model. We purposefully chose simpler, but effective model generation and reputation scoring

[1] http://iplists.firehol.org/

[2] https://blog.webernetz.net/palo-alto-external-dynamic-ip-lists/

[3] https://www.talosintelligence.com/reputation_center

[4] https://www.spamhaus.org/

[5] https://www.virustotal.com/

978-1-5386-7848-0/18/$31.00 ©2018 IEEE

steps in DAbR to accommodate fast and high volume traffic. We collected a dataset of around 87k IP addresses and their attributes to evaluate DAbR and clocked an F-1 score of 78%.

The rest of this paper is organized as follows: Section II describes a brief literature review and Section III explains the DAbR system architecture and its two phases of operation. In Section IV, the data collection statistics and results of DAbR evaluation are discussed before concluding the paper with future directions in Section V.

II. Literature Review

Reputation scoring is studied and used in a variety of fields like search engine query [3], e-commerce [4], social network rankings for enterprise profitability calculation [5], network trust-worthiness [6] etc. Page et al. [3] devised "PageRank" algorithm that uses reputation scoring to rank the search results based on the backlinks from each webpage. Katkar et al. [4] developed a Trust Reputation System (TRS) for e-commerce applications and used data mining to perform a semantic analysis of feedback, for calculating the reputation score. In the field of online and social media reputation, Marrakhi et al. [5] devised a technique IRMS (Intelligent Reputation Measuring System) to rank a brand's presence in the social media. Their scoring is based on the number of citations of the brand, its reach, impact, and influence in social media.

Reputation scoring is also discussed in the field of networking [6] to ensure trustworthiness of the participating nodes. Zhou et al. [7] devised a "Gossip" based reputation aggregation in peer-to-peer networks called 'GossipTrust' in which, each node receives reputation vectors from other nodes in the network and selectively integrates vectors to determine the trustworthiness of participating nodes. Mui et al. [8] used a bayesian probabilistic approach for formulating ratings in distributed networks. In the field of wireless sensor networks, Kim et al. [9] used a fuzzy logic based approach to score trust levels of each node, based on multiple degrees of trustworthiness in each node pair.

Reputation scoring also finds application in detecting malicious activities on Internet. Esquivel et al. [10] used reputation scoring of IP addresses for e-mail spam filtering by checking the SPF (Sender Policy Framework) resource records of SMTP senders. They were able to identify legitimate servers, spam servers and end hosts among the SMTP senders using their technique. In another research, Anderson et al. [11] used the meta-characteristics of TLS (Transport Layer Security) connections for TLS based malware detection. Studies have been done to detect malicious websites/domains using reputation scoring. Hegli et al. [12] used Maximum Entropy Discrimination (MED) classifier for reputation scoring of websites, based on data regarding domain registration, service hosting, IP address, domain creation date, popularity rank, number of hosts, etc. Chiba et al. [13] developed and evaluated a method to detect malicious websites using SVM based analysis of the octet-based, extended octet-based and bit string-based features of IP addresses. Another related work is from Antonakakis et al. [14] where they developed "Notos" reputation system that uses the unique DNS characteristics to filter out malicious domains based on their previous involvement with malicious or legitimate internet services. They used clustering analysis of network-based features, zone-based features and evidence-based features for each domain for reputation scoring. In many of these existing techniques, the scoring is done offline and hence can employ processing-intensive techniques. However, in this paper, we focus on developing a lightweight technique which uses the similarity of an IP address to the existing bad IP's to determine reputation scores.

III. DAbR (Dynamic Attribute based Reputation Scoring) for IP Addresses

Reputation scoring is an effective technique to identify attacks from known attack sources like botnets, spammers, virus infected systems, etc. [15]. However, most traditional IP reputation services rely on manually updated lists of bad/malicious IP addresses and cannot detect attackers using IP's which are not yet reported as bad or malicious. In this paper, we associate any IP address with a reputation score which is indicative of its malicious behavior. In the literature, Elkhannoubi et al. [16] point out several interesting observations regarding malicious IP addresses and factors which potentially helps their proliferation. Their research points towards different social, economic, and political factors which affect malware spread. Our preliminary analysis of bad IP address lists and their attributes (eg. IP address origin country, ASN, ISP, etc.) also showed that many bad IP's share similar attributes. Hence in this paper, we develop DAbR, a lightweight technique (computationally fast) for IP reputation scoring, by extracting and assimilating the similarity of IP address attributes from existing bad lists. Each DAbR score (the calculated reputation score) is a measure of trustworthiness of that IP address calculated from its attributes and their presence in existing bad IP address lists. DAbR creates a vector space for bad IP addresses and when a new IP address is available, it is projected to this vector space for calculating its DAbR score.

A. System Architecture

The general architecture of DAbR is described in Figure 1. It operates in broadly 2 phases: a model generation phase that generates a model for bad IP addresses and a reputation scoring phase which uses the generated model for scoring new and unobserved IP addresses. Sections III-A1 and III-A2 delve into their specifics.

1) Model Generation Phase: This is an offline phase which generates a model for bad IP addresses based on existing bad IP address attributes. It uses a lightweight technique for model generation to accommodate the ever-changing threat landscape. This phase should be repeated and the model should be updated whenever the known lists of bad IP addresses are altered significantly.

The main inputs to this phase are the lists of known bad IP addresses and their corresponding attributes from intelligence sources. As shown in Figure 1, this phase has 3 major tasks, each performed by a separate module in our implementation.

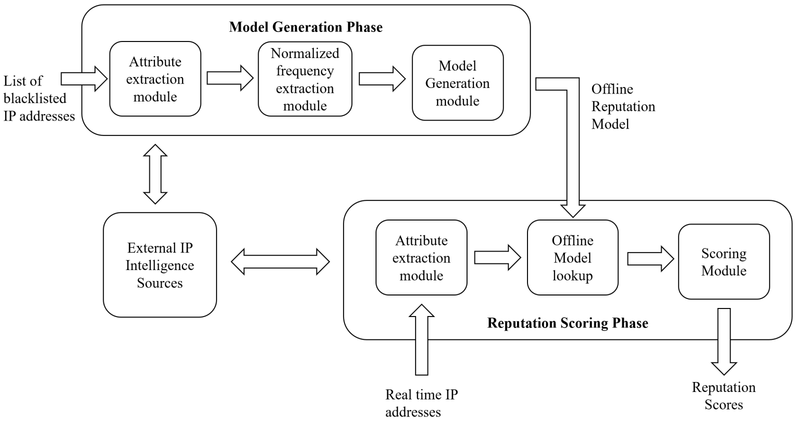

Fig. 1: System Architecture

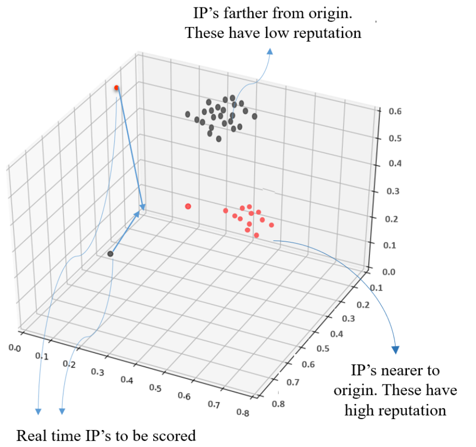

Fig. 2: A Sample IP Vector Space

The first task is attribute extraction which gathers intelligence about each IP address in the ingested list of bad IP addresses. Let $AttributeList = \{attr_1, attr2, ..., attr_i, ..., attr_n\}$ be the list of n attributes we extract for each IP address. Current DAbR implementation extracts the following attributes about each IP address.

1) Autonomous System Number (ASN)
2) Internet Service Provider (ISP)
3) Country where the IP is registered
4) Usertype (residential or commercial user)
5) Country where the IP is located
6) Subdivision in the country where the IP is located
7) City in the country where the IP is located

We represent each IP address $X = \langle x_1, x_2, x_3, .., x_i, .., x_n \rangle$ as a vector of these n attributes, where x_i corresponds to the value of i^{th} attribute in the $AttributeList$. Most of the IP address attributes, x_i's, are nominal in nature and does not have a linear relationship between them. For example, the attribute '$CountryLocation$' takes values like '$country_a$', '$country_b$', etc. If we represent these values in a one-dimensional space, their relative positioning does not provide any valuable information due to their nominal nature. Hence, our second task in the model generation phase is to project such nominal values to a new n dimensional space such that their relative positioning is directly related to their reputation. We use normalized frequency to generate such a reputation scale for each nominal IP address attribute. Let $UX_i = ux_{i,1}, ux_{i,2}, .., ux_{i,j}, .., ux_{i,m}$ be the m unique nominal values possible for an attribute $attr_i$. The $NormalizedFrequency(NF)$ for each unique attribute $ux_{i,j}$ is defined as described in Eq: 1.

$$\begin{aligned} &NF(ux_{i,j}) = N_{ux_{i,j}}/N_{total} \\ &where, \\ &\quad N_{ux_{i,j}}\textit{: number of IP addresses having } ux_{i,j} \\ &\qquad\quad \textit{as an attribute value} \\ &\quad N_{total}\textit{: total number of IP addresses} \end{aligned} \tag{1}$$

The normalized frequency, $NF(ux_{i,j})$, always lies in the range $[0, 1]$ and their relative positioning has the following semantics: the attributes which occur more frequently in the input list have a higher value of normalized frequency (closer to 1) compared to the ones that occurred less frequently (closer to 0). This implies that higher the value of $ux_{i,j}$, the more frequently it is present among the attributes of bad IP addresses. Conversely, if the value is closer to 0, we have not seen this attribute frequently among that list of known bad IP attributes. We hypothesize that larger the value of $NF(ux_{i,j})$, the larger is its tendency to indicate a malicious activity.

Algorithm 1 Model Generation Phase

```
Input:
    $BlackListedIPList$
Output:
    $ReputationModel$

badIPVectorArray = getAttributes(BlackListedIPList)
for i in badIPVectorArray do
    allUniqueAttributes = getUniqueAttributes[i]
    for j in allUniqueAttributes do
        NF = GetNormalizedFrequency[j] using Eq: 1
        append NF to ReputationModel
return ReputationModel
```

The final task in this phase is model generation. A DAbR $ReputationModel$ is an aggregation of the normalized frequencies for all nominal attributes present among the bad IP attributes. This module now generates $NF(ux_{i,j})$ for each nominal attribute $attr_i \in AttributeList$ and packs them into the model. The generated $ReputationModel$ is used by the reputation scoring phase for real-time scoring of unknown IP addresses. An algorithmic representation of this phase is presented in Algorithm 1.

2) Reputation Scoring Phase: The objective of the reputation scoring phase is to associate a numeric reputation score with an IP address based on the generated model. Whenever an IP address is available, DAbR associates a reputation to it in 3 steps as shown in Figure 1. First, DAbR looks up external

IP intelligence sources and collects attributes about the IP address. It is represented as a vector $X' = \langle x'_1, x'_2, x'_3, ..., x'_n \rangle$ where x'_i corresponds to the value of $attr_i \in AttributeList$. The second step is offline model lookup in which X' is projected to the new n dimensional space and generate a new vector $NX' = \langle nx'_1, nx'_2, ..., nx'_i, ..., nx'_n \rangle$ using the $ReputationModel$ generated in the Model Generation Phase. Each nx_i is the normalized frequency of x'_i corresponding to $attr_i$ from $ReputationModel$. If the value x'_i is not present in the $ReputationModel$, the value is taken as 0 because we have never seen that attribute among the reported bad IP address attributes.

This new vector NX' is a point on the new n dimensional vector space. Since we use the normalized frequency value, the new vector space has the property that *closer the point is to the origin, the lesser we have seen it in the bad IP address; or conversely, farther the point is from the origin, the vector is more likely similar to some of the known bad IP address*. We use this property in the scoring module to generate a reputation score for each IP address. The '$DAbRScore$' or the reputation score generated by DAbR for an IP address is the inverse of the euclidean distance of NX' from the origin in the new n dimensional space. It is calculated using Eq: 2. The operation of this module is presented in Algorithm 2.

$$\begin{gathered} ED = \sqrt{ux_{1,j}^2 + ux_{2,j}^2 + ... + ux_{n,j}^2} \\ DAbRScore = (1.0 - (ED/ED_{max})) * 10 \end{gathered} \tag{2}$$

$where,$

ED : Euclidean Distance from origin

ED_{max}: Maximum value of ED computed

Algorithm 2 Reputation Scoring Phase

```
Input:
    $ReputationModel$, $liveIPList$
Output:
    $DAbRScoreList$

liveIPVectorArray = getAttributes(liveIPList)
for i in liveIPVectorArray do
    DAbRVector = generateEuclideanVector[i]
    DAbRScore = getScore(DAbRVector) using Eq: 2
    append DAbRScore to DAbRScoreList
return DAbRScoreList
```

B. Reputation Score Interpretation

The DAbR score of an IP address, defined in Eq: 2, will always be a value in the range 0 - 10. DAbR scores closer to 0 implies very low reputation and those closer to 10 implies very good reputation. Mathematically, the score is inversely proportional to the euclidean distance of IP vector from origin in the new vector space. The vector components are the normalized frequencies of occurrences of their respective attributes in the bad IP attribute list. Hence higher values indicate higher similarity toward IP addresses in bad lists. If we map all bad IP addresses (from the bad list used for model generation) into the vector space, they will not be closer to the origin. Similarly, if the euclidean distance of the IP vector is very small (closer to origin), we will get a higher reputation score.

1) Example Scenario: Consider an example scenario with 1000 bad IP addresses and 3 attributes - Country, ASN, and, ISP. Let's assume that 500 IP's have Country as $CountryX$ and 250 of these 500 IP's have ISP as $ISPX$. Now when we generate the model, $CountryX$ will get a very high NormalizedFrequency value of 0.5 (500/1000 = 0.5) and $ISPX$ will get a NormalizedFrequency value of 0.25 (250/1000 = 0.25). Figure 2 depicts a representative figure of the new 3-dimensional state space generated by the Model Generation phase. If we map all the IP's with country $CountryX$ and ISP $ISPX$ from the training set, they will be positioned farther from the origin as shown. All those points will get very low reputation scores by default using the Eq: 2. When a new IP address is available for finding reputation, its attributes are fetched from the sources. Let's assume the new IP address is from country $CountryX$ and ISP $ISPX$. Now when this IP address is mapped to the euclidean space, it will be positioned farther from the origin and it will get a lower reputation. On the other hand, if the IP has attributes which are not at all present in any of the bad IP attributes, it will be placed at the origin, because our threat intelligence has not found any malicious activities corresponding to them. DAbR will give a very high reputation score for such IP addresses.

IV. Evaluation

DAbR associates a reputation score for unknown IP addresses using a model generated from known bad IP address lists. In this section, we evaluate the performance of DAbR using real datasets. We created a dataset of about 87k IP addresses to evaluate DAbR. Sections IV-A and IV-B discuss the new dataset, its characteristics, and DAbR's evaluation using this dataset.

A. Data Collection

As discussed in Section III, the proposed technique is performed in two phases; a model generation phase and a reputation scoring phase. Bad IP addresses and their attributes are required for model generation. We also require another combined set of bad and good IP addresses for evaluating DAbR.

For bad IP addresses, we used known sources like *Talos IP blacklist feeds*[6]. To improve the total number of IP addresses, we also downloaded around 100k blacklisted domains from services like hpHosts[7] (hpHosts is a hosts file for Windows that allows protection against access to spammer, scammer, pornographic, spoofed and malicious websites.) and converted

[6] http://talosintel.com/feeds/ip-filter.blf

[7] https://www.hosts-file.net/

them to IP addresses. Unlike bad IP addresses available directly from intelligent sources, the good IP addresses are harder to collect. We downloaded a list of top domains from services like OpenDNS[8] and top one million ranked websites as ranked by Alexa[9]. We also performed a preliminary analysis to verify if these domains are mentioned among the blacklisted domains collected from hpHosts. These cleaned domains are then converted to IP addresses to get good IP addresses. In total, we gathered around 100k blacklisted IP addresses and 20k good IP addresses. We collected less number of non-black listed IP addresses because they are not being used for model generation and are only used for evaluation.

The next step is the collection of attributes corresponding to each IP address. We used services like Maxmind GeoIP[10] (Maxmind is an IP Intelligence tool that provides various network and geographical features of user requested IP addresses) databases for collecting various attributes. The collected attributes include Autonomous System Number (ASN), Internet Service Provider (ISP), IP registration country, usertype (residential/commercial user), IP location country, subdivision, and cityname. We faced many challenges during data collection and cleaning. First, attributes were not directly available for many bad IP addresses. Some of these IP addresses were not in service also. Second, the intelligence sources provided limited information about many IP addresses owing to the fact that some providers do not use standard formats for publishing their information. After the attribute collection and data cleaning, the final dataset consisted of 70635 bad IP addresses, 17101 good IP addresses, and attributes associated with each of these collected IP addresses.

B. Results

We evaluate DAbR using the newly generated dataset of bad and good IP addresses. For this, we convert the reputation scoring task to a classification task using a threshold. The classification task is to identify a given IP address as a bad IP address or a good IP address. First, a DAbR score is calculated using Eq: 2 for each input IP address after generating a model described in section III-A1. If the score falls below a specific threshold, $ReputationScore_{thr}$, it is detected as bad. Otherwise, it is detected as good.

$ReputationScore_{thr}$	Precision	Recall	Accuracy	F-1 Score
0.0	0.491	1	0.491	0.659
2.0	0.516	0.999	0.538	0.680
4.0	0.740	0.816	0.768	0.776
6.0	0.802	0.554	0.713	0.655
8.0	0.733	0.113	0.543	0.196

TABLE I: Evaluation metrics for thresholds 0.0 - 8.0

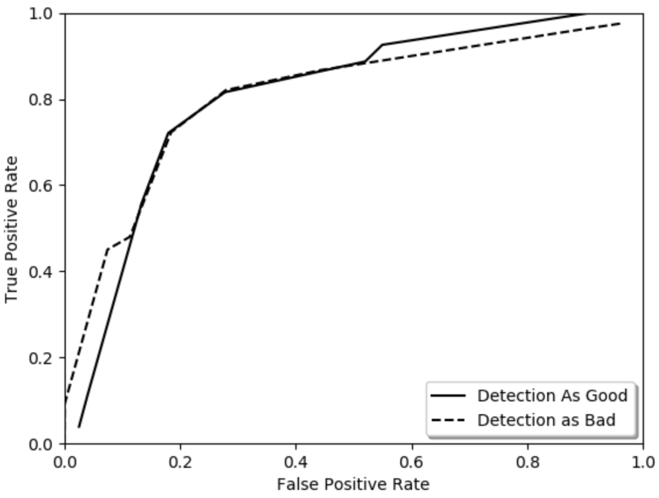

Fig. 3: ROC Curve

$ReputationScore_{thr}$	Precision	Recall	Accuracy	F-1 Score
4.4	0.740	0.815	0.768	0.776
4.5	0.741	0.815	0.768	0.776
4.6	0.749	0.810	0.773	0.779
4.7	0.752	0.809	0.775	0.779
4.8	0.754	0.807	0.776	0.780
4.9	0.756	0.793	0.773	0.774
5.0	0.795	0.721	0.771	0.756

TABLE II: Evaluation metrics for thresholds 4.4 - 5.0

We used a modified 4 fold cross validation for evaluation. In each fold of a traditional 4 cross validation, 3 out of 4 parts of the input data is used for model generation and the remaining part is used for testing. However, DAbR uses only bad IP address's attributes for model generation. Hence in our evaluations, for each of the 4 folds in the 4 cross validation, we use 3 out 4 parts of the input bad IP dataset for model generation. For reputation scoring, we used a union of the remaining one part of the bad IP dataset and the complete set of good IP's. The Receiver Operating Characteristic (ROC) curve, Figure 3, is plotted using the average of true positive rates and false positive rates across the 4 folds in cross validation. In this figure, we plotted the performance of DAbR by varying the $ReputationScore_{thr}$ from 0 to 10. Two ROC curves are depicted in Figure 3. The solid line represents the ROC curve, when we consider True Positive as "Good IP detected as Good" and False Positive as "Bad IP detected as Good". The dotted lines represent the ROC curve when True Positive is "Bad IP detected as Bad" and False Positive is "Good IP detected as Bad". The ROC curve shows that the classification task performs well with a True Positive rate of around 80% along with a False Positive rate of below 20%. Other classic performance metrics for a classification task like F-1 score, accuracy, precision, and recall are also reported. Table I reports these metrics for $ReputationScore_{thr}$ from 0.0 to 8.0. Classification metrics for more granular values of $ReputationScore_{thr}$ from 4.4 to 5.0 is reported in Table II, where we achieve maximum F-1 scores. As seen from the tables I and II, the maximum value of F-1 score = 78.034% is seen at a threshold of 4.8. At this threshold for reputation

[8] https://github.com/opendns/public-domain-lists
[9] http://stuffgate.com/stuff/website
[10] https://www.maxmind.com/en/home

		Detection outcome		
		Good IP	Bad IP	Total
Actual Value	Good IP	55268	13136	68404
	Bad IP	17979	52656	70635
	Total	73247	65792	

TABLE III: Confusion Matrix

scoring, the average true positive rate is about 76.5% and the average false positive rate is about 22%. Also, Precision = 75.45%, Recall= 80.79% and accuracy of detection is 77.6% at threshold of 4.8. The confusion matrix at this threshold is presented in Table III. In this confusion matrix, the total number of IP's is the cumulative sum across all the 4 folds in our cross validation.

In a large enterprise with a very large number of connections, a 20% false positive rate is high. However, it should be noted that DAbR is not a classification technique. The main task of DAbR is to quickly associate a numeric value corresponding to every incoming IP address, otherwise considered normal. The numeric value indicates how similar or dissimilar the IP address is to the existing blacklisted IP addresses. As described in Section I, DAbR is designed to work in conjunction with a policy enforcement module. Instead of detecting it as either Good or Bad, it delegates the response action to the enforcement module. In existing systems, 100% of incoming packets need to be sent for further investigation. With the introduction of TLS 1.3 and wider adoption of encryption, techniques like deep packet inspection will be highly costly and privacy-invasive because it could involve actual decryption. The DAbR score enables to reduce this cost by avoiding such inspections for a large share of packets according to a predefined policy. This would be a very substantial performance improvement and allows penalizing only suspicious connections rather than everyone connected.

V. Conclusion

In this paper, we proposed DAbR, a lightweight reputation scoring system for IP addresses. In this technique, we associate a DAbR score on the scale of 0-10 for each IP address in network traffic. The score is representative of the IP address's trustworthiness based on the existence of its attributes in known malicious attributes. To evaluate DAbR, we aggregated a dataset of 87k IP addresses and their attributes from various existing sources like Talos, hpHosts, OpenDNS, and Maxmind GeoIP. We used a threshold on DAbR score to classify good/bad IP addresses and observed a classification F-1 score of 78% with selected attributes. The results demonstrate DAbR score's usefulness in separating network traffic into different classes. A policy enforcement module can now be employed to filter malicious network traffic in a more efficient and limited privacy-invasive way. IP's with bad reputation scores should be sent for detailed evaluation while traffic corresponding to good reputation can pass through without penalization. Our technique does not use processing-intensive techniques and hence can meet requirements of high speed reputation scoring on high volume situations. In the future, we would like to extend our reputation scoring to accommodate specific traffic patterns expected on the network. That is, learning the general norms automatically from network traffic and associate higher reputation for them. Our technique can also be extended for scoring other network observables like domain, URL, etc. Another interesting domain to explore is to analyze and identify more attributes useful for reputation scoring.

VI. Acknowledgements

This project was partially funded by a research award from Cisco Systems. We gratefully acknowledge them for their support in the successful completion of DAbR.

References

[1] "Most cyber attacks now use encryption: Are you prepared for the good, the bad and the ugly?" https://www.venafi.com/blog/most-cyber-attacks-now-use-encryption-are-you-prepared-for-good-bad-and-ugly, accessed: 2018-06-22.

[2] "February 2018 zscaler ssl threat report," https://www.zscaler.com/blogs/research/february-2018-zscaler-ssl-threat-report, accessed: 2018-06-20.

[3] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web." Stanford InfoLab, Tech. Rep., 1999.

[4] S. Katkar, K. Pande, P. Patil, and Q. Ahmed, "Implementation of trust reputation system (trs) for e-commerce applications using data mining," *Imperial Journal of Interdisciplinary Research*, vol. 3, no. 4, 2017.

[5] M. El Marrakchi, H. Bensaid, and M. Bellafkih, "Scoring reputation in online social networks," in *10th International Conference on Intelligent Systems: Theories and Applications (SITA)*. IEEE, 2015.

[6] Y. Wang and J. Vassileva, "Trust and reputation model in peer-to-peer networks," in *Proceedings Third International Conference on Peer-to-Peer Computing (P2P2003)*, Sept 2003, pp. 150–157.

[7] R. Zhou and K. Hwang, "Gossip-based reputation aggregation for unstructured peer-to-peer networks," in *2007 IEEE International Parallel and Distributed Processing Symposium*. IEEE, 2007, p. 95.

[8] L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt, "Ratings in distributed systems: A bayesian approach," in *Proceedings of the Workshop on Information Technologies and Systems (WITS)*, 2001.

[9] T. K. Kim and H. S. Seo, "A trust model using fuzzy logic in wireless sensor network," *World academy of science, engineering and technology*, vol. 42, no. 6, pp. 63–66, 2008.

[10] H. Esquivel, A. Akella, and T. Mori, "On the effectiveness of ip reputation for spam filtering," in *Communication Systems and Networks (COMSNETS), 2010 Second International Conference on*. IEEE, 2010.

[11] B. Anderson, S. Paul, and D. McGrew, "Deciphering malwares use of tls (without decryption)," *Journal of Computer Virology and Hacking Techniques*, pp. 1–17, 2016.

[12] R. Hegli, H. Lonas, and C. K. Harris, "System and method for developing a risk profile for an internet service," uS Patent 8,438,386.

[13] D. Chiba, K. Tobe, T. Mori, and S. Goto, "Detecting malicious websites by learning ip address features," in *Applications and the Internet (SAINT), IEEE/IPSJ 12th International Symposium on*, 2012.

[14] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster, "Building a dynamic reputation system for dns." in *USENIX security symposium*, 2010, pp. 273–290.

[15] "Citrix: Ip reputation," https://docs.citrix.com/en-us/netscaler/11/security/reputation/ip-reputation.html, accessed: 2018-06-30.

[16] H. Elkhannoubi and M. Belaissaoui, "Assess developing countries' cybersecurity capabilities through a social influence strategy," in *2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, Dec 2016, pp. 19–23.