Classifying Anomalous Events in BGP Datasets

Marijana Ćosović and Slobodan Obradović
Faculty of Electrical Engineering, East Sarajevo
Bosnia and Herzegovina
marijana@etf.unssa.rs.ba, slobo.obradovic@gmail.com

Ljiljana Trajković
Simon Fraser University
Vancouver, British Columbia, Canada
ljilja@cs.sfu.ca

Abstract—**Border Gateway Protocol (BGP), which enables Internet interconnectivity, is susceptible to various anomalous events that may affect the Internet performance. Understanding the nature of anomalous events (unintentional or malicious) and their effects helps classify future events and improve the Internet robustness. Determining the rate and causes of these anomalous events is important for assessing loss of data and connectivity. BGP update messages contain network reachability information stored in a Routing Information Base (RIB). In this paper, we use datasets of known malicious attacks and a power outage event and employ machine learning algorithms to identify traffic anomalies.**

Keywords—BGP; machine learning; classification; malicious attacks; misconfigurations; power outage

I. Introduction

Propagation of the BGP routing information is susceptible to various anomalous events such as misconfigurations, power outages, malicious attacks, and worms. These events may spread false routing information throughout the Internet by either dropping data or directing traffic through unauthorized autonomous systems (ASs) and hence risking eavesdropping. An administrative domain consists of one or more networks that have uniform routing policies and operate independently. BGP is a centralized protocol that connects distinct Internet sites by announcing reachability information to other networks and, as such, could be exposed to malicious attacks.

Large-scale power outages and configuration errors in BGP routers may induce anomalous routing behavior leading to network instabilities. Configuration errors are considered accidental unless unconventional, malicious practices are intended. Table leak and prefix hijack events are examples of BGP configuration errors that may lead to large-scale connectivity issues in the Internet. A table leak occurs when an AS such as an Internet service provider (ISP) announces a prefix from its Route Information Base (RIB) that violates previously agreed routing policy. A prefix hijack arises as a consequence of an AS originating a prefix that it does not own. Large-scale power outages may affect ISPs due to a lack of reliable power backup. This could be further manifested in network equipment failures leaving affected networks isolated and the service disrupted.

Critical services depending on the Internet have been steadily increasing and, hence, anomalous events and their effects are having increasing economic consequences. Determining the anomalous events and their causes is an important step needed for assessing loss of data and connectivity. Internet connectivity and stability are affected by anomalous routing. Hence, it is important to classify anomalous events and avoid their effects on BGP. BGP data have been analyzed to identify anomalous events and discover rules that may be used in anomaly predictions [1]-[3]. Recent trends in designing BGP anomaly detection systems rely more frequently on machine learning techniques [1]-[4]. In this paper, we use known classifiers [3] and test their ability to reliably detect network anomalies in datasets of known BGP network anomalies.

The paper is organized as follows. In Section II, we describe BGP. Extraction of BGP features and the BGP datasets used in this study are described in Section III. Classification models and their performance measures are discussed in Section IV. We conclude with Section V.

II. Border Gateway Protocol

BGP (an upgrade of the EGP protocol) is an interdomain routing protocol used for routing packets in networks consisting of a large number of ASs. BGP version 4 allows classless interdomain routing (CIDR), aggregation of routes, incremental additions, better filtering options, and the ability to set routing policies based on business concerns. BGP protocol employs the Path Vector protocol, a modified version of the Distance Vector protocol. BGP is a standard for the exchange of information between the ISPs and between ISPs and major users such as university networks.

BGP relies on the Transport Control Protocol (TCP) protocol to establish a TCP connection (port 179) between the routers. BGP router establishes a TCP connection with directly connected routers residing in different ASs. Because of their size, BGP routing tables are exchanged only once between the peering routers when they first connect. They are afterwards exchanged only to communicate updates regarding new prefixes or withdrawals of the existing prefixes.

BGP allows ASs to exchange reachability information with peering ASs to transmit information about the availability of all routers within the AS. Based on the exchanged information and routing policies, it determines the most appropriate path to destination. Hence, BGP allows each subnet to announce its existence to the Internet and to publish its reachability information. Thanks to BGP, all sub-networks are connected and are known to the Internet.

III. BGP DATASETS

Internet routing data analyzed in this paper are acquired from two projects that provide valuable information to networking research: the Routing Information Service (RIS) project initiated in 2001 by the Réseaux IP Européens (RIPE) Network Coordination Centre (NCC) and the Route Views project at the University of Oregon, USA. Both projects are collecting and storing chronological routing data that provide a unique view of the Internet topology. These routing data are beneficial when attempting to detect Internet anomalies.

Anomalous events considered in this paper are power outage and BGP router configuration error events shown in the Table I. They create interdomain routing instabilities manifested by sharp and sustained increases in the number of announcement or withdrawal messages exchanged by BGP routers. We only consider BGP update messages collected by the Remote Route Collectors (RRCs) and stored in the multi-threaded routing toolkit (MRT) format. After raw data from RIPE and RouteViews are downloaded, messages are preprocessed and valuable information extracted.

TABLE I. ANOMALOUS EVENT DATASETS

Event	Date	RRC	Peers
Moscow Power Blackout	May 2005	RIS 05	AS1853, AS12793, AS13237
AS9121 Routing Table Leak	Dec. 2004	RIS 05	AS1853, AS12793, AS13237
AS3561 Improper Filtering	Apr. 2001	RIS 03	AS3257, AS3333, AS286
Panix Domain Hijack	Jan. 2006	Route Views	AS12956, AS6762, AS6939, AS3549
AS Path Error	Oct. 2001	RIS 03	AS3257, AS3333, AS6762, AS9057
AS3356/AS714 De-peering	Oct. 2005	RIS 01	AS13237, AS8342, AS5511, AS16034

A. Data Preprocessing

In this paper, we use BGP update messages that originated from RIS route collectors: rrc01 (LINX, London), rcc03 (AMS-IX, Amsterdam), rrc04 (CIXP, Geneva), and rrc05 (VIX, Vienna). We also used data from the Route Views project data collector. Only data collected during the periods of Internet anomalies are considered. We use BGPdump library on Linux platform to transform BGP update messages from MRT into ASCII format. In order to process data files in batches, we use a bash script. BGPdump is a C library maintained by the RIPE Network Coordination Centre and used to analyze dump files produced by MRT.

BGP update messages are loaded into tables of an Oracle database using the SQL loader utility. All BGP update messages are imported sequentially into the database. Feature statistics, volume features, and AS-PATH features are computed every minute during five-day periods for six well known anomalous Internet events: AS 9121 Routing Table Leak [5], Moscow Power Blackout [6], AS 3561 Improper Filtering [6], Panix Domain Hijack [6], AS Path Error [6], and AS 3356/AS 714 De-peering [7]. Details regarding time of the event, remote route collectors used for acquiring data, and peers observed are listed in Table I.

We consider five-day periods, the days of the event (anomalous data points) and two days prior and two days after the event (regular data points). Choosing fifteen features was shown suitable for detecting anomalous events [3]. Hence, for each anomalous event, the dimension of the feature matrix is 7,200x15.

B. Anomalous Events

The Moscow Power Blackout occurred on May 25, 2005 and lasted several hours. Moscow Internet exchange was shut down during the power outage. Routing instabilities were observed due to loss in connectivity of some ISPs peering at this exchange. This effect was apparent at the RIS remote route collector in Vienna (rrc05) through a surge in announcement messages arriving from peer AS 12793, as shown in Fig. 1. Hence, we use volume of announcements as one of the features to detect instabilities.

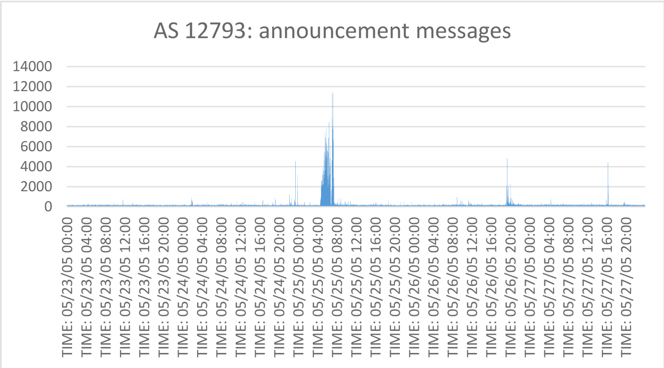

Fig. 1. Surge of announcement messages at the AS 12793 peer during the Moscow Power Blackout.

The AS 9121 Routing Leak occurred on December 24, 2004 when AS 9121 announced to peers that it could be used to reach almost 70% of all prefixes (106k+). As a consequence, tens of thousands of networks had either misdirected or lost traffic. The AS 9121 started announcing prefixes to peers around 9:20 GMT and the event lasted shortly past 10:00 GMT. The AS 9121 continued to announce bad prefixes throughout the day. The announcement rate reached the second peak at 19:47 GMT.

The AS 3561 Improper Filtering is a BGP misconfiguration error that occurred on April 6, 2001. AS 3561 allowed improper route announcements from its downstream customers, which created connectivity disruptions. Surge of announcement messages coming from peer AS 3257 was observed at the RIS rrc03.

Panix, the oldest commercial ISP in New York, was hijacked on January 22, 2006 and its services were unreachable from the greater part of the Internet, leading to BGP instability. Con Edison (AS 27506) advertised routes that it did not own at the time. Previously, Panix was a customer of Con Edison, which was once authorized to offer advertised routes. Even though AS 27506 was the autonomous system that originated improper routes, major downstream ISPs did not configure filters properly and propagated those improper routes, leading to excess number of update messages.

The AS PATH Error occurred on October 7, 2001. It was caused by an abnormal AS-PATH (AS 3300, AS 64603, AS 2008) that contained private AS 64603 that should not have been included in the path. At the time, AS 3300 and AS 2008 belonged to INFONET Europe and INFONET USA, respectively. The path was distributed to the network via misconfigured routers and caused the leakage of the private AS numbers. Shown in Fig. 2 is the increase of withdrawal messages around 20:00 GMT, peaking around 21:00 GMT and slowly decreasing during the following four hours.

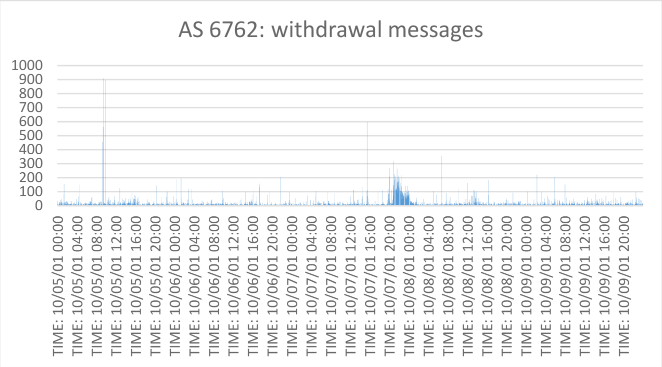

Fig. 2. Surge of withdrawal messages at the AS 6762 peer during the AS PATH Error Event.

The AS 3356/AS 714 De-peering event occurred on October 5, 2005. Even though the Level3 Communications Inc. (AS 3356) notified the Cogent Communications Inc. (AS 714) two months in advance of de-peering, the event created reachability problems for many Internet locations. Mostly affected were single-homed customers of Cogent (~2,300 prefixes) or Level3 (~5,000 prefixes). De-peering resulted in partitioning of around 4 % of prefixes in the global routing table.

IV. Classification Models and Performance

We have developed [3] six models based on Naïve Bayes (NB) and Decision Tree (J48) classifiers implemented in Weka (v. 3.7.11). We used meta learning, a subfield of machine learning, to automatically detect data models. Filter and wrapper methods for feature selection were employed within meta classifiers to evaluate sets of relevant features. Filter methods use subset selection based on general characteristics of the data. This process is independent of the learning algorithm. Wrapper methods employ a subset evaluator to create all possible subsets of features. The subset selection is based on learning algorithm used to train the model itself. Hence, learning algorithm is wrapped into selection process.

Our aim is to identify an anomaly class, usually represented with fewer samples than the regular class. We use performance parameters such as precision, recall, F-measure, and Matthews correlation coefficient (MCC) to evaluate performance of the classifiers [3] and confirm that the receiver operating characteristics (ROCs) may be misleading in cases of class imbalance by providing overly optimistic results [8]. Hence, we also use appropriate precision-recall (PR) curves.

Models NB-1 and NB-2 are classifiers trained on discretized datasets and on datasets with the F-measure optimized, respectively. Models NB-3 and NB-4 are filter methods employing a feature evaluator based on correlation with greedy stepwise search method and a feature evaluator based on gain ratio with ranking of individual features, respectively. Models NB-5 and NB-6 are wrapper models using a classifier as a parameter for evaluation of sets of features on training data and using a 5-fold cross-validation internally to estimate the accuracy of the learning scheme for a set of features, respectively. We consider datasets of known network anomalies and test classifiers' ability to reliably detect those network anomalies for given datasets.

Of six datasets that were considered in this study, four datasets (Moscow Power Blackout, AS 9121 Routing Table Leak, Panix Hijack, and AS Path Error) collected at rrc peers AS 2793, AS 13237, AS 12956, and AS 6762, respectively, have produced significant results. Performance measures for the Moscow Power Blackout event are shown in Table II.

TABLE II. Performance of NB and J48 Classifiers using the Moscow Power Blackout Dataset

Data set	Model	F-measure	MCC	ROC	PR
Moscow Power Blackout	NB-1	0.848	0.846	0.971	0.912
	NB-2	0.877	0.874	0.971	0.903
	NB-3	0.877	0.874	0.969	0.842
	NB-4	0.848	0.846	0.971	0.912
	NB-5	0.900	0.898	0.971	0.911
	NB-6	0.892	0.890	0.982	0.907
	J48-1	0.894	0.893	0.896	0.795
	J48-2	0.796	0.804	0.903	0.800
	J48-3	0.876	0.874	0.926	0.808
	J48-4	0.905	0.903	0.931	0.849
	J48-5	0.902	0.901	0.934	0.829
	J48-6	0.896	0.894	0.941	0.835

Wrapper methods based on Naïve Bayes models (NB-5 and NB-6) show better performance than other NB models, although with a small margin. Filter and wrapper methods based on Decision Tree model (J48-4 and J48-5) outperform other decision tree models for the Moscow Power Blackout dataset.

Performance measures for the AS 9121 Routing Table Leak event are shown in Table III. Wrapper methods based on Naïve Bayes models (NB-5 and NB-6) show better performance than other NB models, although with a larger performance margin than in the case of the Moscow Power Blackout dataset. Filter and wrapper methods based on the Decision Tree models (J48-4 and J48-6) outperform other decision tree models for this case.

Performance measures for the Panix Hijack event are shown in Table IV. The wrapper method based on Naïve Bayes model NB-5 shows the best performance followed by the NB-2 model. Filter and wrapper methods based on Decision Tree models (J48-4 and J48-6) outperform other decision tree models for the Panix Hijack event.

TABLE III. PERFORMANCE OF NB AND J48 CLASSIFIERS USING THE AS 9121 ROUTING TABLE LEAK DATASET

Data set	Model	F-measure	MCC	ROC	PR
AS 9121 Routing Table Leak	NB-1	0.901	0.902	0.999	0.961
	NB-2	0.899	0.898	0.999	0.949
	NB-3	0.888	0.888	0.998	0.888
	NB-4	0.901	0.902	0.999	0.961
	NB-5	0.950	0.950	0.993	0.950
	NB-6	0.956	0.956	0.992	0.962
	J48-1	0.906	0.905	0.958	0.847
	J48-2	0.672	0.694	0.958	0.846
	J48-3	0.930	0.929	0.934	0.844
	J48-4	0.955	0.955	0.955	0.898
	J48-5	0.938	0.938	0.967	0.873
	J48-6	0.944	0.944	0.967	0.893

TABLE IV. PERFORMANCE OF NB AND J48 CLASSIFIERS USING THE PANIX HIJACK DATASET

Data set	Model	F-measure	MCC	ROC	PR
Panix Hijack Event	NB-1	0.706	0.721	0.999	0.918
	NB-2	0.820	0.821	0.999	0.911
	NB-3	0.800	0.804	0.998	0.874
	NB-4	0.706	0.721	0.999	0.918
	NB-5	0.848	0.848	0.998	0.905
	NB-6	0.794	0.793	0.994	0.865
	J48-1	0.946	0.946	0.992	0.945
	J48-2	0.864	0.870	0.992	0.944
	J48-3	0.877	0.876	0.970	0.874
	J48-4	0.962	0.962	0.977	0.888
	J48-5	0.855	0.854	0.938	0.739
	J48-6	0.947	0.946	0.988	0.919

Performance measures for AS-PATH error event are shown in Table V. The NB-2 model with a classifier that is trained on datasets with F-measure optimized shows the best performance followed by wrapper methods based on Naïve Bayes classifiers namely, NB-6 and NB-5. Filter and wrapper methods based on Decision Tree models (J48-4 and J48-5) outperform other decision tree models for the AS-PATH error dataset.

TABLE V. PERFORMANCE OF NB AND J48 CLASSIFIERS USING THE AS-PATH ERROR DATASET

Data set	Model	F-measure	MCC	ROC	PR
AS PATH Error	NB-1	0.875	0.877	0.999	0.969
	NB-2	0.938	0.936	0.999	0.955
	NB-3	0.865	0.868	0.999	0.933
	NB-4	0.875	0.877	0.999	0.962
	NB-5	0.907	0.905	0.997	0.900
	NB-6	0.921	0.920	0.999	0.957
	J48-1	0.913	0.911	0.976	0.855
	J48-2	0.910	0.907	0.976	0.854
	J48-3	0.910	0.908	0.960	0.858
	J48-4	0.922	0.920	0.974	0.864
	J48-5	0.921	0.920	0.982	0.846
	J48-6	0.916	0.914	0.980	0.907

Wrapper models use classification algorithms to generate feature subsets and to evaluate them using threshold function applying classification results leading to the best performance measures. Wrapper classifiers are the best performing classifiers in our study. Thirteen out of sixteen wrapper classifiers have performance measure in the top half of all classifiers presented compared to filter classifiers that have performance measures in the top half of all classifiers only in six cases. Note that wrapper methods take longer training time than filter methods.

V. CONCLUSION

Machine learning techniques are one of the most promising approaches for detecting network traffic anomalies. They have been recently widely deployed in analyzing BGP behavior. In this paper, we analyze performance of BGP detection models based on Naïve Bayes and Decision Tree J48 classifiers. The Moscow Power Blackout, AS 9121 Routing Table Leak, Panix Hijack, and AS Path Error datasets are examples of known anomalies that have been tested for developing anomaly detection algorithms.

We have transformed datasets by employing feature discretization and feature selection using both filter and wrapper methods and have analyzed their effects on given datasets in order to improve classification performance. We found that performance of the classifiers is influenced by the employed datasets. Even though none of the classifiers that we analyzed perform the best across all specified datasets, filter and wrapper methods based on decision tree models have outperformed other models.

REFERENCES

[1] N. Al-Rousan and Lj. Trajković, “Machine learning models for classification of BGP anomalies,” in *Proc. 13th IEEE Int. Conf. High Performance Switching and Routing*, Belgrade, Serbia, June 2012, pp. 103–108.

[2] N. Al-Rousan, S. Haeri, and Lj. Trajković, “Feature selection for classification of BGP anomalies using Bayes models,” in *Proc. Int. Conf. Mach. Learning Cybern.*, Xi'an, China, July 2012, pp. 140–147.

[3] M. Ćosović, S. Obradović, and Lj. Trajkovic, “Performance evaluation of BGP anomaly classifiers,” in *Proc. Int. Conf. on Digital Inform., Networking and Wireless Commun.*, Moscow, Russia, Feb. 2015, pp. 115–120.

[4] Y. Li, H. J. Xing, Q. Hua, X.-Z. Wang, P. Batta, S. Haeri, and Lj. Trajković, “Classification of BGP anomalies using decision trees and fuzzy rough sets,” in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, San Diego, CA, USA, Oct. 2014, pp. 1331–1336.

[5] A. C. Popescu, B. J. Premore, and T. Underwood. (May 19, 2005). Anatomy of a Leak: AS9121. Renesys Corporation. Manchester, NH, USA. [Online]. Available: http://research.dyn.com/content/uploads/2013/05/renesys-nanog34.pdf.

[6] (Dec. 27, 2015) North American Network Operators Group Mailing List [Online]. Available:
https://www.nanog.org/mailinglist/mailarchives/old_archive/2005-05/msg00650.html;
https://www.nanog.org/mailinglist/mailarchives/old_archive/2001-04/msg00209.html;
https://www.nanog.org/mailinglist/mailarchives/old_archive/2006-01/msg00484.html;
https://www.nanog.org/mailinglist/mailarchives/old_archive/2001-10/msg00216.html.

[7] Y. Zhang, Z. M. Mao, and J. Wang, “A firewall for routers: protecting against routing misbehavior,” in *Proc. 37th Annu. IEEE/IFIP Int. Conf. on Dependable Syst. and Networks,* Edinburgh, UK, June 2007, pp. 20–29.

[8] J. Davis and M. Goadrich, “The relationship between Precision-Recall and ROC curves,” in *Proc. 23rd Int. Conf. Mach. Learning,* Pittsburgh, PA, USA, June 2006, pp. 233–240.