BGP Behavior Monitoring and Analysis

Yihua Liao, Ke Zhang
Department of Computer Science
University of California, Davis
One Shields Avenue, Davis, CA 95616
{yhliao, kezhang}@ucdavis.edu

Abstract

Border Gateway Protocol, an important inter-domain routing protocol, has a number of vulnerabilities. Little is known about how BGP actually performs in today's Internet. We designed a framework, BGP Assistant, to monitor and analyze BGP traffic. Number of BGP Updates and Route convergence time are used to characterize BGP behavior. Preliminary results with the Oregon Route Views BGP show that BGP Assistant can help the network operator to diagnose the network, identify the anomalous ASes or IP prefixes, and respond to them in real time. Further work is in progress to extend its functionalities and better understand BGP behavior.

1 Introduction

The Internet is comprised of thousands of Autonomous Systems (AS), where each AS is a set of routers under a single technical administration and has its own routing policies. Most autonomous systems exchange routing information through the Border Gateway Protocol (BGP). Therefore as a widely deployed inter-domain routing protocol, BGP plays a critical role in the operation of the Internet.

The current implementation of BGP, the BGP-4 protocol [1], has a variety of vulnerabilities and intrinsic weakness. First of all, BGP packets are transmitted over TCP/IP without any encryption and authentication mechanisms. Communications between BGP peers are subject to active and passive wiretapping attacks. Transmission of fictitious BGP messages, modification or replay of valid messages could occur during the routing information exchange process. A false BGP route may cause deny of service to a destination or redirect that destination's traffic to another insecure location. Secure-BGP (S-BGP) [2] was proposed to enhance the security of BGP by verifying the authenticity and authorization of the BGP control traffic. However, the processing and bandwidth overhead coupled with the storage

requirement poses nontrivial challenges to the adoption of SBGP to the Internet. Second, as a path vector protocol, BGP limits the distribution of a router's reachability information to its neighbor routers. A large number of temporary routing table fluctuations may be generated in response to a single link failure, change in AS topology or change in routing policy. Slow routing convergence can cause the delay or drop of data packets, and thus degrade the efficiency and reliability of the Internet infrastructure [3].

Despite the critical importance of BGP, relatively little work has been done on its real-world performance. It is unclear how well (or poorly) BGP actually performs in today's Internet. Analyzing BGP behavior systematically is a difficult task. For one thing, there is a lack of solid analytic models that can precisely describe the normal behavior of BGP. Secondly, voluminous BGP data make it impossible for network operators to manually scrutinize BGP messages and detect anomalous or malicious activities in real time. In this paper, we propose a framework, **BGP Assistant**, to monitor and analyze BGP behavior. This framework aims to automate the process of interpreting BGP messages at a particular AS, building statistical profiles for this AS and reporting anomalous events. It can assist the network operator to diagnose the network, find the trouble-making BGP peers or IP addresses, and respond to them in real time.

The rest of the paper is organized as follows. In Section 2 we provide additional background on BGP and some related work. Section 3 describes the structure of BGP Assistant and some features we choose to characterize BGP behavior. Section 4 explains our experiments with the Route View Project data and preliminary results. We discuss our future work in Section 5. Finally, Section 6 concludes the paper.

2 Background

BGP is the de facto inter-autonomous system routing protocol. An AS uses an interior gateway protocol and common metrics to route packets within itself. At the boundary of each autonomous system, peer border routers exchange network reachability information with other autonomous systems through BGP. BGP uses TCP as its transport protocol. Two BGP speakers form a transport protocol connection between one another. They exchange messages to open and confirm the connection parameters. When connection is first established, a BGP speaker sends its entire routing table to the peer. During the following BGP session, the incremental updates are sent as the routing table changes. Routes are stored in the Routing Information Base (RIB) in

BGP speakers. It consists of three distinct parts: Adj-RIBs-In, Loc-RIBs, Adj-RIBs-Out. The Adj-RIB-In stores routing information that has been learned from other peers. Using Local decision algorithms, A BGP router applies its local policies to select the routes stored in Adj-RIBs-In and put them into Loc-RIB. Routes that will be advertised to other BGP speakers are present in Adj-RIBs-Out.

Two types of BGP messages are important to BGP operation. First is KeepAlive message, which is sent periodically to ensure the liveness of the connection. If the peers can't receive KeepAlive messages in a preset period of time, the BGP connection has to be closed. Physical connectivity failure (link failure, router crash), transient connectivity problems due to congestion, or even manual reboots, may result in the delay of KeepAlive message to the peers. Sequentially, when BGP session restart, the peers have to send the full routing table again. Second is the Update message. Update messages are used to exchange routing information change between two peers. Usually it has two forms: *announcement* and *withdrawal*. *Announcement* messages carry on the prefix (destination) and ASPATH (a sequence AS number of intermediate autonomous systems between source and destination routers that form the directed path for the route). Upon receipt of a new *Announcement*, each router evaluates the path vector and use local decision algorithms to select the best route among all of the backup routes to that prefix. If the router is a transit router, it will append its unique AS number to the ASPATH and propagate to the downstream BGP speakers. Route *withdrawal*s are sent when a router makes a new local decision that a network is no longer reachable. We distinguish between *explicit* and *implicit* withdrawals. Explicit withdrawals are those associated with a withdrawal message; whereas an implicit withdrawal occurs when an existing route is replaced by the announcement of a new route to that destination prefix.

In an optimal, stable wide-area network, routers should only generate routing updates for relatively infrequent policy changes and the addition of new physical networks. Frequent BGP updates indicate the network routing instability. Routing instability has a number of possible origins, including problems with leased line, router failures, network congestion, software implementation and configuration errors. After one or more these problems affects the availability of a path to a set of prefix, the routers topology close to the failure will detect the fault, withdraw the route and make a new local decision to find a new route, if any, to the set of prefix. These routers will propagate the topology update information to the routers within the same autonomous systems. The boundary routers in the network may also propagate the updated information to the other AS routers. During the updated information propagating in

the Internet, all affected routers will suffer convergence problem. We expect the convergence will not last very long. However, Labovitz et al [4] pointed out in some case BGP routers suffer *slow convergence*. In their experiments, they found when a withdrawal to a destination received, a BGP router would explore a large number of backup routes, which might already be invalid. They observed this delayed convergence was three minutes on average and some took up to 15 minutes. Varadhan et al [5] and Griffin et al [6] explored another kind of routing instability, so called *divergence.* As mentioned before, BGP is policy based routing protocol, which allows administrator of an autonomous system to specify arbitrarily complex policies. In [6], it was showed that it is possible for autonomous systems to implement "unsafe" or mutually unsatisfiable policies, which will result in persistent route oscillations. Besides the slow convergence and divergence problem, Labovitz et al [4] also showed that most of BGP Update messages don't correspond to legitimate network topology changes.

To address these problems of BGP routing, some solutions have been proposed. Griffin et al [6] described modifications to BGP that can guarantee that the protocol will not diverge. Dan et al [7] added a new community attribute [8] in BGP withdrawal messages that can explicit tell if the withdrawal is failure withdrawal or policy withdraw, therefore, can help router to remove some invalid backup routes. These solutions all need to modify the existing well deployed BGP protocol. However in some cases they will encounter some restrictions.

In this paper, we focus on the network routing instability. We designed BGP Assistant, a framework that can help network operators to monitor and analyze BGP behavior, evaluate the network running condition, identify the root causes of some suspicious or malicious BGP behaviors, and take appropriate actions.

3 Methodology

3.1 System Structure

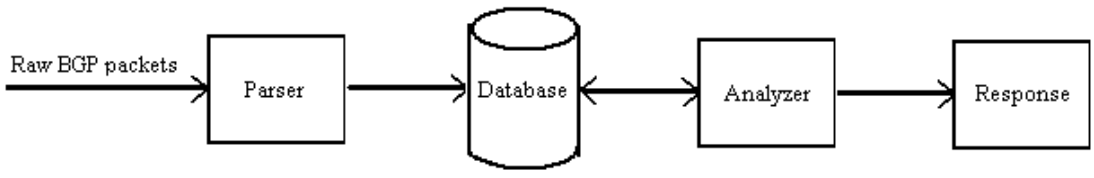


Figure 1: Structure of BGP Assistant

Figure 1 presents the overall structure of our BGP Assistant. The binary BGP packets received from its peers at an AS are read into the system. So far we have only used the BGP Update messages, which provide the most important routing information. It's the parser's responsibility to transform the raw packets into a human-readable text format. Depending on whether the packet was an announcement or withdrawal, various important attributes are extracted from each packet and loaded into a BGP data warehouse. Table 1 lists all the attributes BGP Assistant uses. Information is appended into a table, where each row simply represents a BGP Update on a particular prefix. Therefore, if a BGP Update contains multiple route announcement or withdrawals, multiple rows will be added to the table after parsing this Update. The database was implemented with MySQL. Retrieving an arbitrary subset of the data can be achieved with simple queries.

Attribute	Description
Date	Update receiving date
Time	Update receiving time
AdFlag	'A'/'W', announce or withdraw
SrcAS	AS that sent the Update
DesAS	Receiving AS
Prefix	IP address
PrefixLen	Prefix length
AsPath	ASPATH to the prefix
OriginAS	Origin AS of the prefix
NextHop	IP address of the first AS on the path
UdpMED	Multiple Exit Descriptor
LocalPref	Local preference

Table 1. List of attributes extracted from BGP Updates. Some of them, for example, asPath and nextHop, are not applicable for withdrawals.

Querying the database is performed through the analyzer, which is a crucial component of BGP Assistant. Similar to intrusion detection [9], there are two general approaches to analyzing the BGP events. First, knowledge of known vulnerabilities, such as slow convergence, can be built into the analyzer. Then pattern-matching methods can be employed to identify instances of known vulnerabilities. Second, the analyzer can learn the normal BGP behavior and build statistical profiles. Here BGP behavior is characterized in terms of several statistical metrics and models, which are described in Section 3.2. The analyzer will compare a new observation with the corresponding profile and flag if significant deviation occurs. These two approaches

complement each other and can be integrated together. Meanwhile, the analyzer periodically updates the statistical profiles of the AS and deletes the old entries of the database to save storage space. The analyzer reports the current running condition of the network, and any AS or prefixes that are causing the instability of the network. Based on the output of the analyzer, the response module can notify the administrator, and recommend or automatically take appropriate actions. For example, it can send a traceroute message to the suspicious destination that appears in a BGP Update message and track in real time the impact of the received BGP messages.

3.2 Statistical Metrics

So far we have used the following metrics to characterize BGP behavior, each of which represents a quantitative measure accumulated over a period.

- Number of updates within a time interval. This can be measured at different levels. It can be the number of updates issued by all BGP peers, or by a particular AS; number of updates regarding all prefixes, or a specific prefix.
- Convergence time, the time it takes for the route to a prefix to converge to a stable state. A variation of this is the time difference between the first update and the last update regarding the same prefix within a time window.

It is straightforward to employ standard statistical models, such as Mean and Standard Deviation or Chi-Square, to determine whether a new observation is abnormal with respect to the previous observations. Giving more weight to more recent observations and Constantly updating the profiles may generate adaptive profiles of these metrics.

4 Experiments

4.1 Dataset

We used the BGP data of the Route Views Project [10] of University of Oregon in this study. The Route Views Project provides real-time information about the global routing system from the perspectives of several different backbones and locations around the Internet. Currently, the Route Views router uses multi-hop BGP peering sessions with 54 BGP routers in 43 different ASes, including backbones in America, Asia, Europe and so on. Route Views uses AS6447 in its peering sessions. Routes received from neighbors are never passed on nor used to forward traffic, and Route Views itself does not announce any prefixes.

Although BGP Assistant was originally conceived as a tool to monitor and analyze BGP behavior in real time, we performed our data analysis offline. For our preliminary experiments, we obtained about 90-minute BGP Update data from the Route Views Project, which was originally collected on October 26, 2001. The binary MRT format [11] Update messages were fed into the BGP Assistant parser. Update attributes were extracted and inserted into the database. There are totally 59735 entries in the database, representing 59735 prefix updates received by AS6447, the Route Views router. There are 12 distinct AS peers and 19678 distinct prefixes within this dataset. Then the analyzer was used to query the database and analyze the BGP behavior.

4.2 Instability

Frequent BGP updates indicate network routing instability. An obvious measure of routing instability is the number of Updates within a time period. We set a time window of 30 minutes, and observed the update messages received by AS6447 within this time window. Figure 2 presents the numbers of Update messages issued by several AS neighbors within three consecutive time windows. Clearly, AS3130 and AS2914 sent more Updates than any other ASes. In particular, AS3130 was the dominant one within the second time period. This might result from router configuration errors or physical and data link problems within that AS domain. Similarly, Figure 3 shows the numbers of Update messages regarding different prefixes within the time windows. This type of query to the database can clearly show to the network operators which AS or prefix was causing large unusual amount of BGP Update messages.

Once we know which prefix is suspicious, we can find out which AS contributed more Update messages regarding this particular prefix. Figure 4 presents the number of Update messages with regard to prefix 207.45.205.0/24 issued by different ASes. Again, AS2914 and AS3130 sent more BGP Updates.

4.3 Convergence Time

Convergence time is another metric we used to characterize BGP behavior. Here convergence time is defined as the time it takes for a route to a prefix to converge. We consider a route is stable if there is no route change within the next 15 minutes. The final Update message could be a withdrawal or Announcement. Figure 5 presents the

convergence time distribution of our dataset for these two different categories. While most routes can converge within 5 minutes (300 seconds), some can take as long as more than 15 minutes, which is much longer than expected.

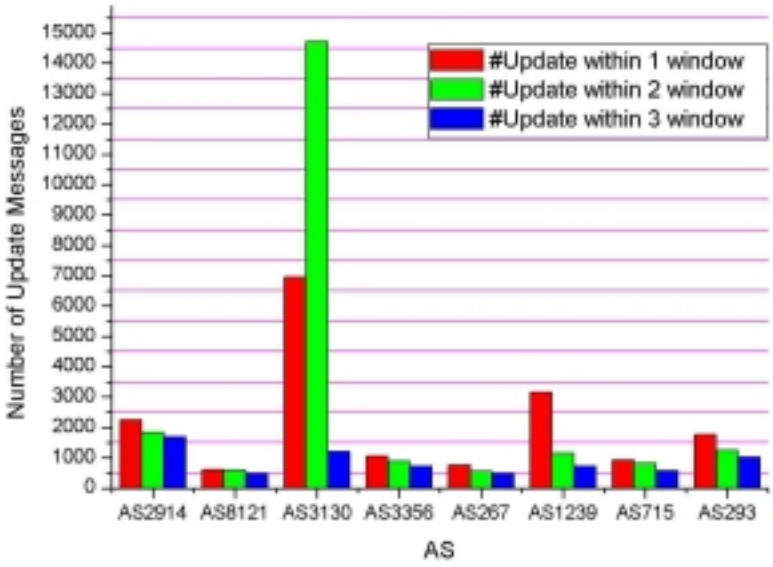
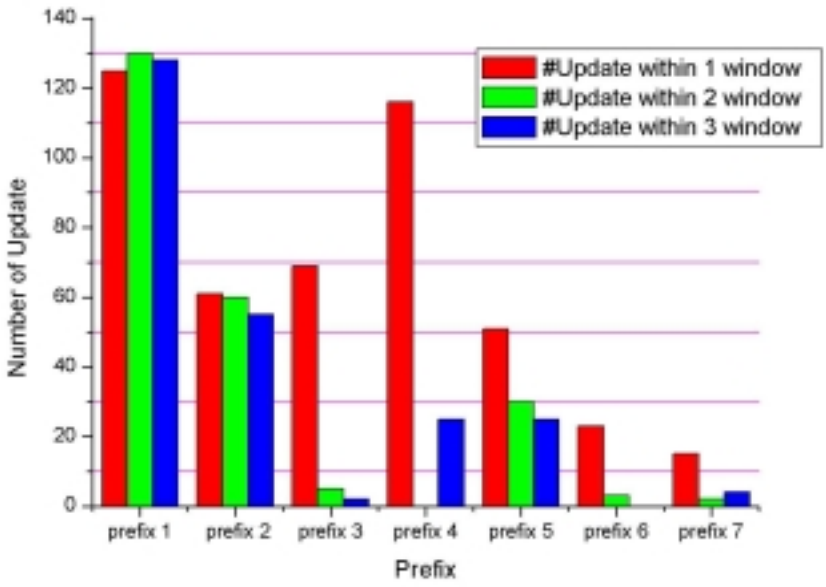


Figure 2 The number of BGP Update messages issued by several ASes within 3 time windows.



Fgure 3. The number of BGP Update messages regarding different prefixes within 3 time windows.

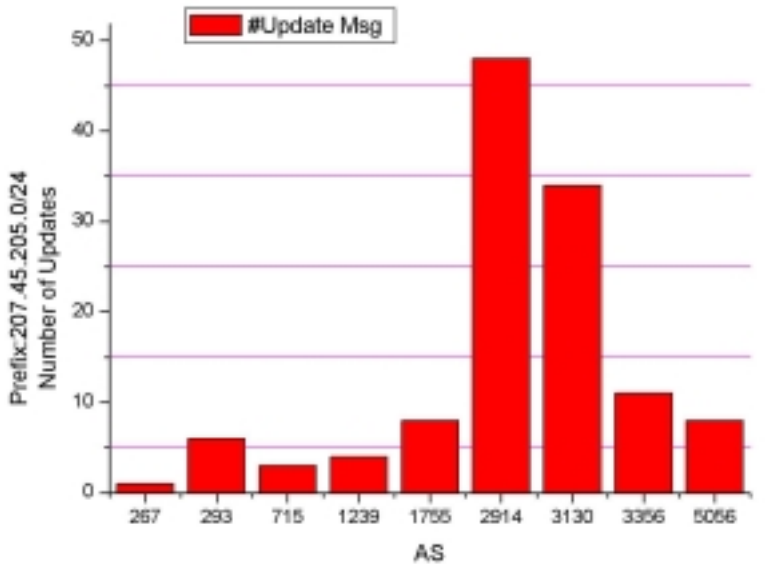


Figure 4. Numbers of BGP Update messages regarding prefix 207.45.205.0/24 sent by different ASes.

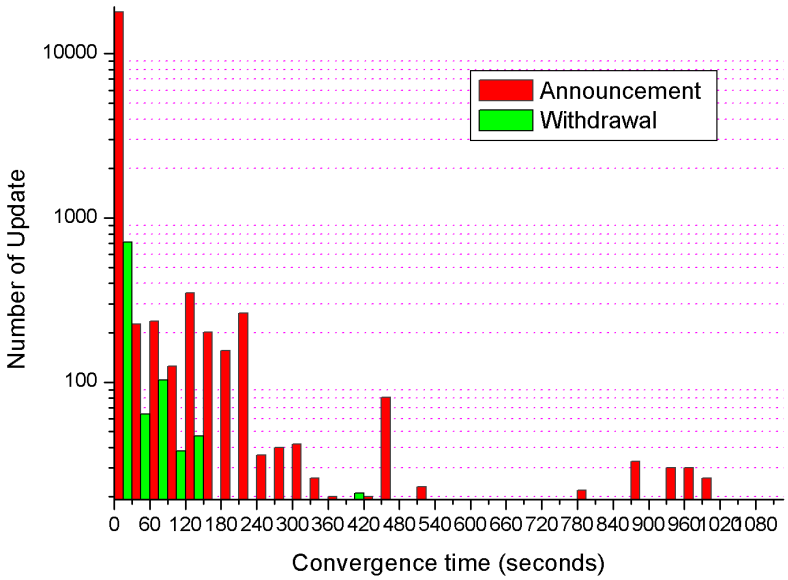


Figure 5. Convergence time distribution.

Due to the time limit, we didn't implement any statistical models. However, it should be fairly straightforward to use methods such as Mean and Standard Deviation or Chi-Square, to determine whether a new observation is abnormal with respect to the previous observations.

5 Future Work

Our BGP Assistant is currently at its early stage. Further work is in progress to extend its functionalities and better understand BGP behavior. One focus of our future work will be BGP slow convergence. So far no specific patterns of slow convergence have been well defined. It is desirable to conduct a formal analysis of BGP slow convergence. There are two possible approaches. One is based on statistical analysis. Through analyzing a large amount of BGP data, one can find the normal behavior of BGP routing and define some statistical patterns of slow convergence. The other way is to infer AS topology from the Update messages and then find how slow convergence problems propagate among ASes. Potentially, if we can specify the patterns of BGP slow convergence, we will be able to use BGP Assistant to detect slow convergence instances. If slow convergence is detected earlier, invalid routes will not be propagated to downstream routers and thus we can save CPU time and network bandwidth.

BGP Assistant can also learn the AS paths to prefixes, especially those of great importance, for example, the prefixes that cover the IP addresses of root DNS servers. It could bring the whole Internet into jeopardy and cause catastrophic consequences if a malicious AS announces a false path to one of the root DNS servers. Learning the normal path(s) to a particular root DNS server will prevent such malicious announcements.

Finally, a visualization module can be built in our BGP Assistant. Using some emerging information visualization techniques [12], the visualization module will transform BGP behavior into a graphical representation. This will utilize human knowledge and facilitate the process of interpreting BGP behavior and identifying anomalous BGP activities.

6 Conclusions

A framework, BGP Assistant, has been designed to monitor and analyze BGP behavior in real time. It aims to help the network operators to detect abnormal BGP events, diagnose the network and identify the trouble-making ASes or prefixes. In our preliminary results with the Oregon Route Views Project data, the number of BGP Updates and route convergence time were used to characterize BGP data. Routing instability and route convergence were analyzed. Future work includes further study

on modeling BGP behavior and learning AS paths to prefixes and a visualization module.

7 Acknowledgement

The authors wish to thank Professor S. Felix Wu for his insightful comments and constructive suggestions.

Reference

[1] Y. Rekhter and T. Li, “a border Gateway Protocol 4 (BGP-4)”, RFC 1771, March 1995.

[2] S. Kent, C. Lynn, J. Mikkelson and K. Seo, “Secure Border Gateway Protocol (S-BGP) – Real World Performance and Deployment Issues”, in *Proceedings of the Network and Distributed System Security Symposium (NDSS 2000)*, San Diego, California, February 2000.

[3] C. Labovitz, G. R. Malan and F. Jahanian, “Internet Routing Instability”, in *Proceedings of the ACM SIGCOMM*, September 1997.

[4] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, “Delayed Internet Routing Convergence”, in *Proceeding of ACM SIGCOMM*, August 2000.

[5] K. Varadhan, R. Govindan, and D. Estrin, “Persistent Route Oscillations in Internet Domain Routing”, Technical Report USC CS TR.96-631, Department of Computer Science, University of Southern California, Feb. 1996.

[6] T. Griffin, F. Shepard, and G. Wilfong. “An Analysis of BGP Convergence Properties”, in *Proceedings of ACM SIGCOMM*, August 1999

[7] D. Pei, X. Zhao, L.Wang, D.Massey, A.Mankin, S.Wu, and L. Zhang, “Improving BGP Convergence through Consistency Assertions”, in *Proceeding of the IEEE INFOCOM*, June 2002.

[8] R.Chandra, P.Traina, “BGP Communities Attribute”, RFC 1997, August 1996.

[9] S. Axelsson, “Intrusion Detection System: A survey and Taxonomy”, http://citeseer.nj.nec.com/axelsson00intrusion.html, 2000.

[10] University of Oregon Route Views Project, http://antc.uoregon.edu/route-views/.

[11] MRT, http://www.mrtd.net/

[12] G. Geisler, “Making Information More Accessible: A Survey of Information Visualization Applications and Techniques”, http://www.ils.unc.edu/~geisg/info/infovis/paper.html, 1998.