Department of Information and Communication Engineering
Graduate School of Information Science and Technology
THE UNIVERSITY OF TOKYO

Master Thesis

BLT: A Taxonomy and Classification Tool for Mining BGP Update Messages

(BLT: BGP アップデートメッセージ解析のためのメッセージの分類手法の提案と実装)

Tomoyuki Kitabatake
北畠 知行

Supervisor: Professor Hiroshi Esaki

February 2018

Abstract

The Border Gateway Protocol (BGP) is a key component in Internet routing. Consequently, monitoring BGP messages is essential to identify changes that are detrimental to networks reachability. This is however a complicated task, mainly due to the stateful and noisy nature of BGP. One need to keep track of the entire routing table to really understand the meaning of a single BGP message. And significant bursts of messages may be completely redundant. In this work, we propose a complete taxonomy of BGP update messages and its corresponding classification tool called BLT. We also introduce a simple anomaly detector based on BLT that pinpoints surge of selected classes of messages. We illustrate the benefits of this detector with five case studies that validate its ability to identify meaningful events.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

The Border Gateway Protocol (BGP) is Internet's key protocol for achieving inter-domain routing. Using BGP, Autonomous Systems (ASes) can globally advertise their IP space and the routes they learnt from other ASes. To keep track of routing changes, border routers maintain a local Routing Information Base (RIB) that consists of a set of BGP attributes (e.g. AS path) for each globally routed IP prefix. If the network undergoes changes, routers exchange BGP update messages to inform the new attributes. Depending on a router decision process these new attributes can be reflected in the router's RIB or not.

Monitoring BGP updates is crucial for network operators and researchers trying to track Internet dynamics and identify important changes that can compromise users connectivity. This is however a complicated task because BGP conceals routing process details (e.g. routing policies or complete network topology) and, at the same time, BGP is very noisy for certain network changes and instabilities, sometimes referred as BGP churn [1, 2].

1.2 Objective

In this work, our goal is to provide a general framework to assist operators and researchers in monitoring the Internet routing dynamics. Namely, we aim to classify and annotate BGP messages based on their effect on the routing process.

1.3 Approach

To achieve this goal we identified 17 different changes that update messages cause to routers' RIB. These 17 types of update are organized in a hierarchical taxonomy that provides and increasing level of details. In addition, we provide a classification tool, called BLT, that fetches BGP data and labels each message based on the proposed taxonomy. Since the labels convey detailed functions

of the messages, it greatly helps one to filter out superfluous messages and focus only on relevant messages.

1.4 Contributions

We demonstrate the benefits of BLT with a simple application, an anomaly detector that reports surge of messages of a certain class. Using this anomaly detector we present five case studies of BGP route leaks and Internet outages that are easily identified as a surge of one specific type of message.

1.5 Constitution of this thesis

The main contributions of this work consist of a complete hierarchical taxonomy of BGP update messages (Chapter 4), an open source classification tool for BGP data (Chapter 5) and an anomaly detector identifying surges of certain types of messages (Chapter6).

Chapter 2

BGP

2.1 The Internet

The Internet is a set of connected network. To route packets to distant networks beyond their network's border, the router in the network have to know pathes to other networks. The routing information is exchanged between routers in each network using routing protocols. Obtaining information of path to other network dynamically with routing protocols makes the Internet flexible.

To send or recieve traffic each other, all hosts connected to the Internet need to be identified uniquely. In Internet Protocol(IP) network, IP address is an identifier of an end host. End hosts in the same network are addressed with a common most significant bit group (prefix) in their IP address. This enables different networks identified uniquely.

Autonomous System(AS) is a collection of the networks under the same administrative control. Routers in the same AS run the same routing algorithm, called IGP (Internal Gateway Protocol). IGP include RIP (Routing Information Protocol), IS-IS (Intermediate System to Intermediate system) and OSPF (Open Shortest Path First). On the other hand, a routing algorithm used in inter-AS routing is called EGP (External Gateway Protocol). Nowadays BGP specified in RFC 4271 is the de fact standard routing protocol for EGP [3].

2.2 BGP Overview

BGP allows each AS to advertize existence of its own prefix to remaining ASes in the Internet and makes sure that they can know how to get there. A router which recieved advertisement about a prefix from other AS propagates it to make the rest of ASes in the Internet reach packets to the prefix.

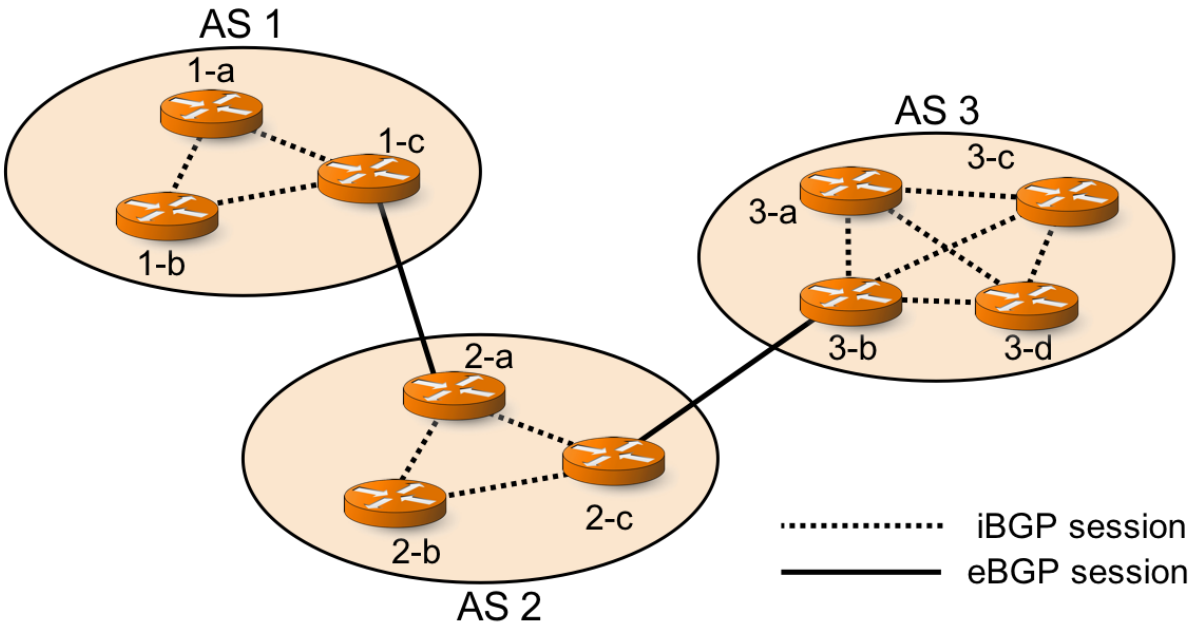

Fig. 2.1: eBGP and iBGP sessions

2.3 BGP sessions

A pair of BGP routers exchanges routing information over a TCP connection established between them. This TCP connection is called BGP session. After a router receives message from a neighbor router and saves routing information to routing table, the router modifies attributes of the message and sends it to other routers having BGP session. Repeating this process causes networks connectable each other.

Figure 2.1 shows an example of BGP sessions. There are TCP connections between two gateway routers 1-c and 2-a and between two gateway routers 2c and 3-b. These connections are established beyond the border of two different ASes. In contrast to that, there are some TCP connections between two routers in the same ASes. A BGP connection between routers in different ASes is called external BGP session(eBGP session) and a BGP connection between routers in the same ASes is called internal BGP session(iBGP session). In figure 2.1, the eBGP sessions are denoted with black straight lines whereas the iBGP sessions are shown with black dotted lines.

2.4 Path attributes and BGP routes

BGP is a path vector protocol which keeps the path information of each destination prefixes that can update dynamically. Maintaining the paths to any prefixes in routing table in routers in each AS enables packets to be routed correctly. To propagate these path information, BGP message has

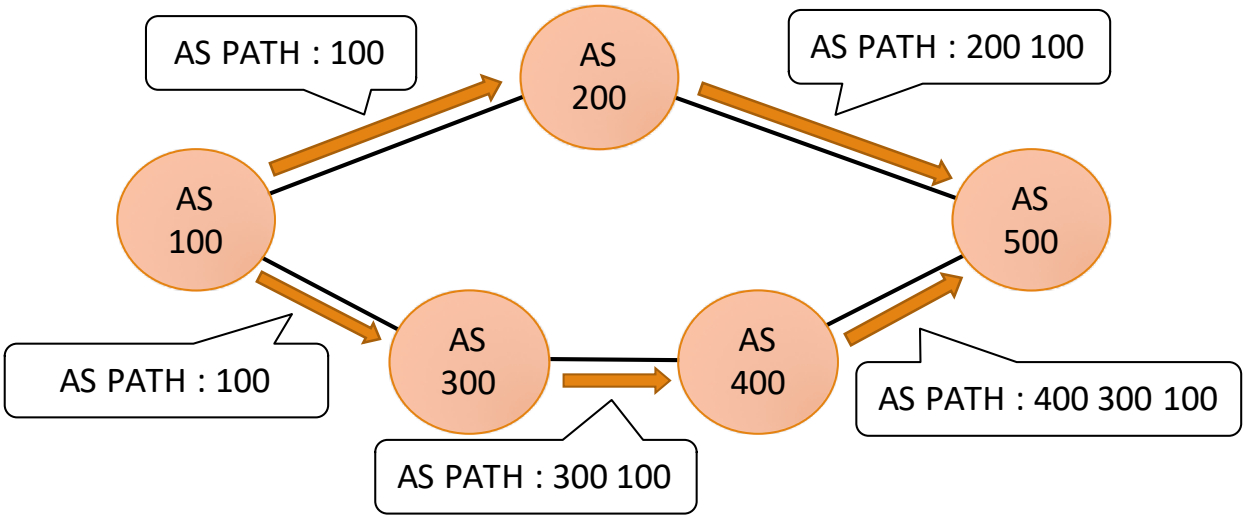

Fig. 2.2: An example of propagation of BGP messages.

AS-PATH attribute which provides the path to destination prefixes. In addition to the AS-PATH attribute, BGP message has a lot of other attributes in order for the organizations to apply their policies. The wide variety of attributes makes it difficult to decide the best path in each router. In other words, this charactaristics mean that BGP has high flexibility. This is why BGP is regarded as a flexible routing protocol.

Figure 2.2 depicts how to propagate BGP messages. Assume the following:

- There is an undirected graph in which five ASes are connected: AS100, AS200, AS300, AS400, AS500.
- AS100 has eBGP session with AS200 and AS300.
- AS200 has eBGP session with AS100 and AS500.
- AS300 has eBGP session with AS100 and AS400.
- AS400 has eBGP session with AS300 and AS500.
- AS500 has eBGP session with AS200 and AS400.

When AS100 obtains new IP prefix {1.2.3.0/24}, it tells connected ASes that new prefix is in AS100 which mean the path to the {1.2.3.0/24} is AS100. AS200 and AS300 which recieved the messages from AS100 keep this information in their routing table, add own AS number and propagate it to other ASes connected. In this phase, the path to the new prefix sended from AS200 is {200, 100}. AS400 behaves in the same way. Finaly, AS500 recieves two BGP messages having different route to the new prefix. The path to the new prefix in the message from AS200 is {200, 100}, whereas that from AS400 is {400, 300, 100}. AS500 chooses the second path because of the short length of the path, and maintains it in its routing table. If AS500 has more connections, AS500 will propagate the path information to them again.

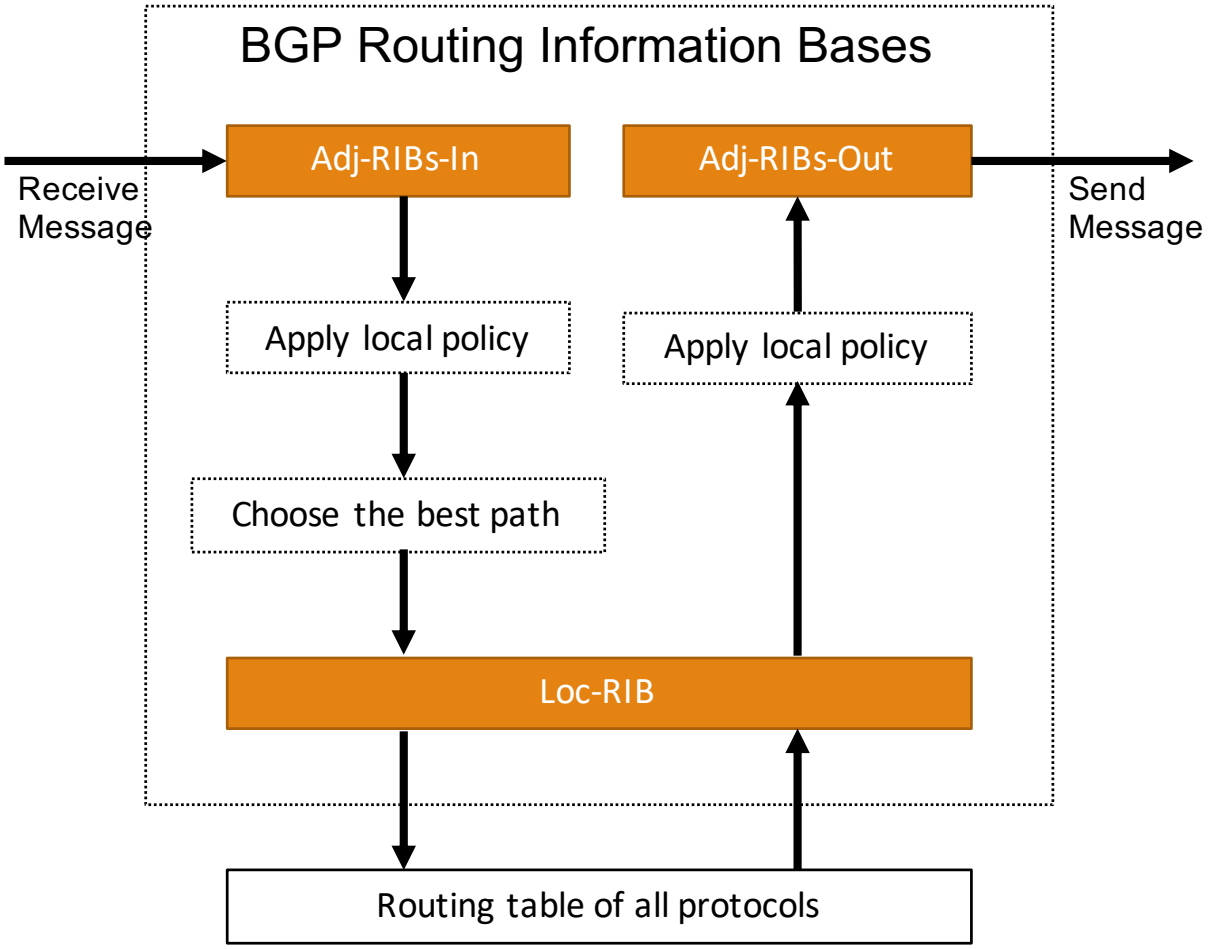

Fig. 2.3: Conseptual model of RIBs

2.5 BGP Routing Table

BGP routing information is stored in the BGP Routing Information Bases(RIBs). RIBs not only maintain routes but also controll routing information. When a BGP router receives a BGP message, the router can apply local policy to route or filter it in RIBs before the route is stored. Before sending a message to neighbors, the BGP speaker can do the same thing.

The RIBs consists of three tables: namely, the Adj-RIBs-In, Loc-RIBs and the Adj-RIBs-Out (see Figure 2.3).

1. Adj-RIBs-In
   Routing information in inbound update messages that were recieved from neighbor BGP router are stored here. Entries of this table are classified by from which peer a message came or groups in which have a common local policy. The inputs of the Decision Process are these contents stored in Adj-RIB-In.

2. Loc-RIB
   The best paths which are selected after applying a local policy to routing information stored in Adj-RIBs-In are kept in the Loc-RIB. These BGP routing information are installed to the routing table in which routing information of all protocols (RIP, OSPF, static, etc.). They are maintained and used in that table to route packets beyond the border of AS.
3. Adj-RIBs-Out
   The paths in the Loc-RIB are compared with that in the routing table of all routing protocols. This process produces a best path used in outbound update message to neighbor routers. Adj-RIB-Out keeps routing information that are adopted local policy to the routing information in Loc-RIB. They are available for outbound update message.

If there are several paths to the same prefix in Adj-RIB-In, the best path is choosed and stored in Loc-RIB and used for routing packets. In the Decision Process, following priorities are adopted.

1. Network Layer Reachability Information (NLRI)
   If a router doesn't know the route of the next hop router of a new path, or there is no reachability to the next hop router of a new path, this path is ignored and never installed to the Loc-RIB.
2. LOCAL_PREF attribute
   LOCAL_PREF attribute of a message is compared with that of others. The path to which the highest LOCAL_PREF attribute is added takes precedence.
3. AS_PATH attribute
   If LOCAL_PREF attributes in the all path to the same prefix are the same, AS_PATH attributes are compared. The route to which the shortest length of AS_PATH is added takes precedence.
4. ORIGIN attribute
   If superiority can not be obtained in the process until comparing routes in AS_PATH attribute, Origin attribute is used for Decision Process. IGP has higher priority than EGP. Incomplete is lowest priority of them.
5. MED attribute
   If the best path cannot been determined by comparison of ORIGIN attribute, MED attribute is compared. The lower MED value has higher priority.
6. type of peer
   IEGP peer has higher priority than that of IBGP.
7. IGP cost up to the next hop
   The path which has lowest IGP cost up to the next hop is used.
8. router ID
   If superiority can not be obatained in the all processes above, the ID of the router in the peer has highest priority.

2.6 Phenomenons

2.6.1 Duplicate announce

A BGP speaker announces routes only when the path or other attributes were changed. However, a lot of redundant messages have been observed. These redundant messages are called *duplicate announce*. Too many duplicate announces cause low readability of the log for operators and large unnecessary operation load to a CPU. Several works focused on BGP duplicates after first observation of duplciate announces in 1998 [4]. Duplicate announces occupied about 15% of the messages obtained from RIPE monitors in 2007 [5]. More recently, in 2012, duplicates are responsible for about 40% of update messages found in monitors located in ASes of different size from previous study [1]. One of the causes of duplicate announces was proved to unintended interaction between eBGP and iBGP [6, 7].

2.6.2 Path hunting

Sometimes BGP withdrawal messages trigger a sharp increase of the number of messages [8]. This event is happend in the following cases. A BGP speaker recieves a withdrawal message propagated by neighbor router that discovered failure reachability to a prefix. The BGP speaker has more routes about the prefix. Another path, therefore, is selected in the BGP speaker, and it propagates *announce messages* that mean the path was changed to the prefix to its peers. After that, the BGP speaker receives a withdrawal message which mean the reachability failure of the prefix from another peer whose associated AS is the nexthop AS to the prefix in the Loc-RIB in the BGP speaker. Since the BGP speaker temporarily recognizes it that the current path to the prefix became not available, it propagates another path that was in the Adj-RIB-In. This event may continue for several times, in the worst case, for as many paths as the BGP speaker has in the Adj-RIB-In. As a result, the event generates the churn of the BGP messages.

Figure 2.4 shows an example of path hunting. When the BGP session between AS1 and AS2 fails, AS2 sends withdrawal message W to neighbors (AS3, AS5). Since keeping backup paths to AS1 in Adj-RIB-In, AS5 that received withdrawal removes the path {5,2,1}, and installs the path {5,3,2,1} to its Loc-RIB-In, and advertizes new path {5,3,2,1} to peers. After that, AS3 forwards withdrawal to AS4 and AS5. AS5 receives withdrawal again, AS5 propagates new path {5,4,3,2,1} because maintaining more backup routes to AS1 in Adj-RIB-In. Finaly, AS4 sends withdrawal to AS5. Since AS5 doesn't have more backap routes, it recognizes unreachability to AS1, and sends withdrawal to its peers. During this time, AS5 has propagated three BGP messages to the Internet. This kind of process is occured in several ASes at once withdrawal is propagated.

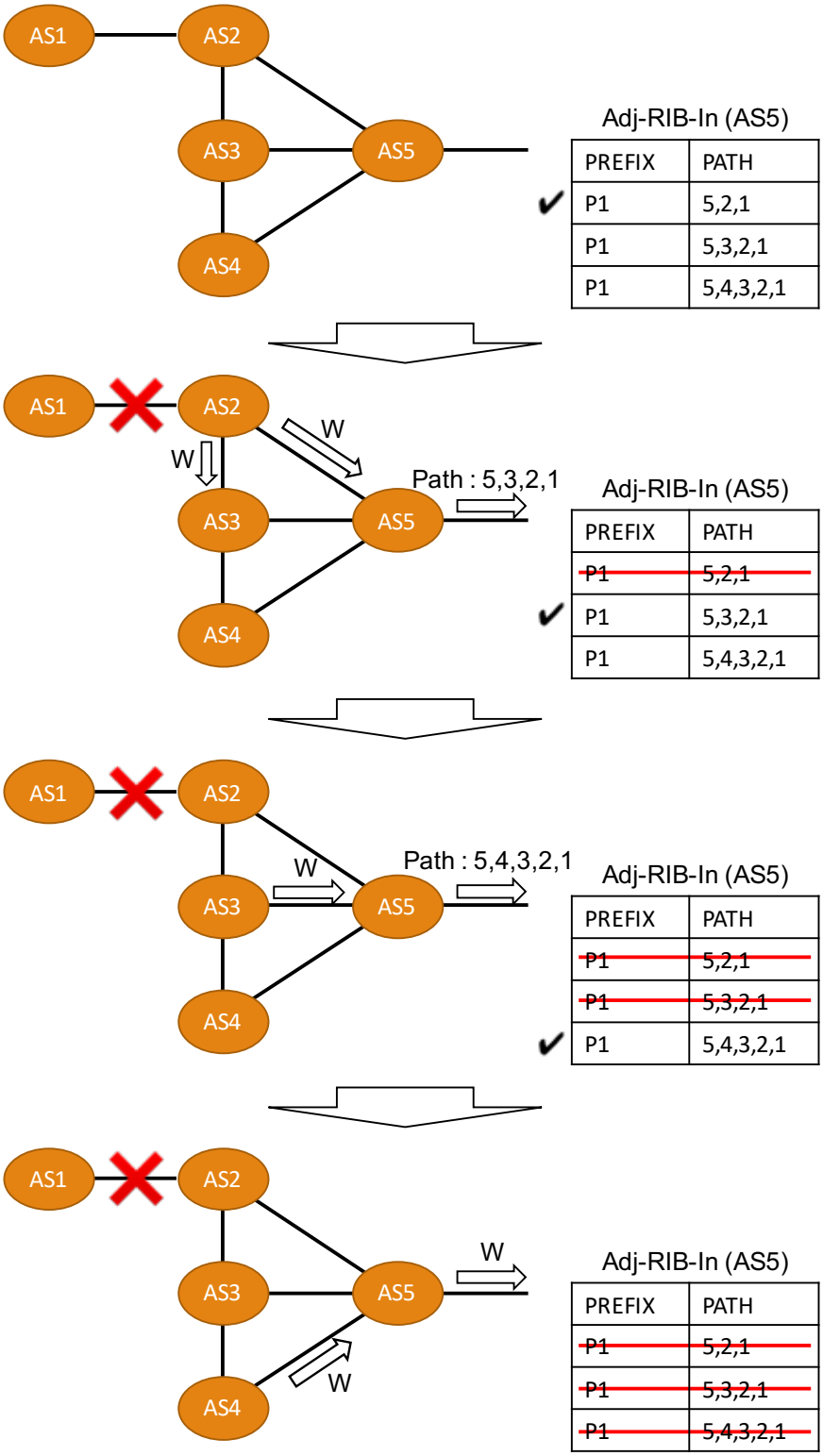

Fig. 2.4: An example of path hunting

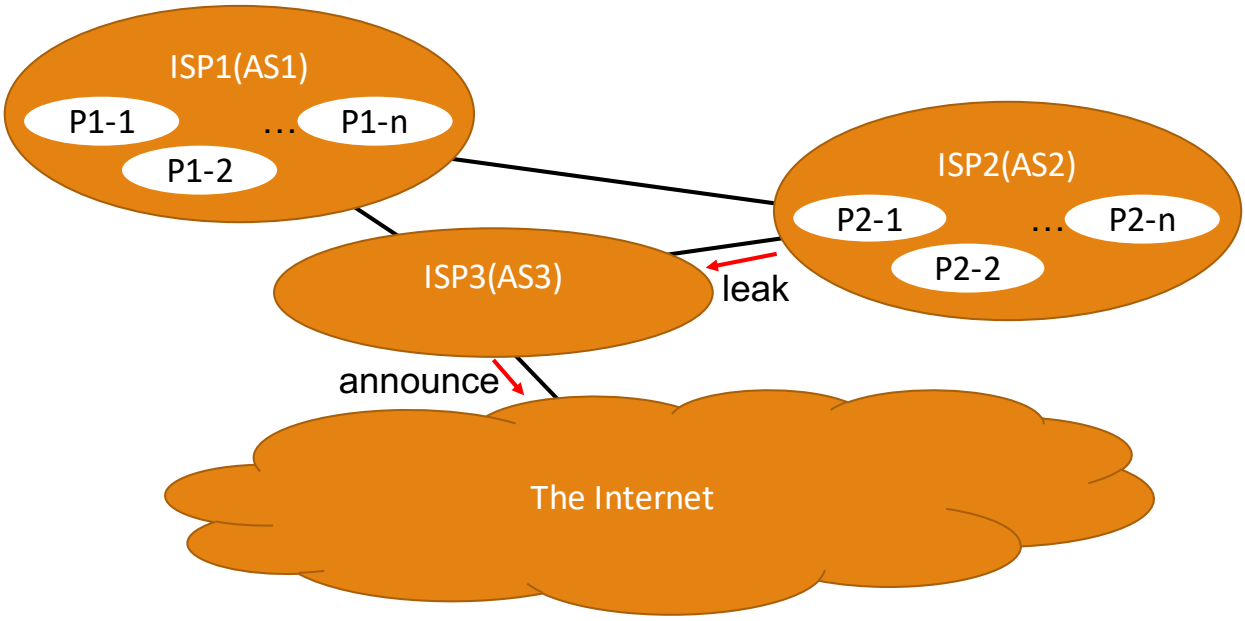

Fig. 2.5: An example of route leak

2.7 Type of Incidents

In chapter 7, we discuss results of the anomaly detection of BGP events based on the taxonomy as evaluation of it. Before discussing that, we address the type of BGP incidents to understand the following chapters.

2.7.1 Route leak

A BGP has vulnerabilities that have not been fixed yet. Misconfiguration of the BGP router or intentional attack to it causes unexpected behaiors in the BGP routing system. One of those is known as *route leaks*. Although route leaks has some patterns, we describe the basic common knowledge of it using one example. Classification of the route leaks is defined in RFC7908 [9].

Figure 2.5 depicts an example of route leaks. Route leaks occure as follows.

1. An AS (such as AS2 in Figure 2.5) leaks routing information that it should not notannounce in violation of intended policies.
2. The ASes receiving messages (such as AS3 in Figure 2.5) doesn't detect the leak and propagate them to its customers or peers.

Occurrence of route leaks may causes reachability problems and create Internet congestion [10] Suppose that AS2 leaks routing information of prefixes in AS1 (such as P1-n in Figure 2.5), the routes to P1-n are changed as path through AS2 and then all trafic destinated to P1-n becomes to

go through AS2. If AS2 does not have the environment that can deal with it, most packets to P1-n are lost in AS2. As a result, P1-n appears to disapper from the Internet.

For one more example, AS2 leaks routing information about prefixes its private ASes have (such as P2-n in Figure 2.5). AS3 accepts these update messages and propagates them. As a consequence, excessive BGP churn is produced. Generally, prefixes in the private ASes are aggregated and are not known to external ASes. Misconfiguration of the router may cause this event.

The massive events of route leak include Google route leak in August 2017[*1], route leak from Level(3) in November 2017[*2] and Telekom Malaysia route leak in June 2015[*3].

2.7.2 BGP hijacking

BGP hijacking is classified as one pattern of the route leak (Type 5 in RFC7908). It is occured when an AS propagates the prefixes to an upstream AS as if origin of the prefixes is itself. Although data packets to these prefixes intend to go to the AS, the AS does'nt have these prefixes. Reachability of these prefixes for other ASes is lost. This kind of events are produced by misconfiguration or intentional attack.

The famous events of BGP hijacking include BGP hijack of Innofield AG in April 2016[*4] and BGP hajack out of India in November 2015[*5].

[*1] https://dyn.com/blog/large-bgp-leak-by-google-disrupts-internet-in-japan/
[*2] https://blog.thousandeyes.com/comcast-outage-level-3-route-leak/
[*3] https://blog.thousandeyes.com/route-leak-causes-global-outage-level-3-network/
[*4] https://bgpmon.net/large-hijack-affects-reachability-of-high-traffic-destinations/
[*5] https://bgpmon.net/large-scale-bgp-hijack-out-of-india/

Chapter 3

Related Work

BGP has been widely studied by the research community. The scalability of BGP received a lot of attention, and in particular, the growth of routing tables [11] and BGP churn [1, 2].

BGP data has also been used in various monitoring systems. For example, Argus [12] is a prefix hijack detection system that identifies anomalous changes in BGP data and triggers pings from several vantage points to characterize the detected anomalies. A recent study also uses BGP data to detect Infrastructure outages [13], that approach relies on BGP communities to map AS paths to facilities and BGP update messages to track vanishing facilities. Detected changes are also characterized with extra data plane measurements.

Closer to our work, BGPMon is a service provided by OpenDNS that helps network operators to monitor their IP prefixes. This service relies mainly on BGP data and consists in a set of involved heuristics*[1], for example modeling the business relationships between difference ASes. This system mainly focuses on the origin ASes thus it may fails to detect important events where the origin ASes are not changing (e.g. the BGP route leak from Google presented in Chapter 7.2.1).

*[1] http://www.blackhat.com/us-15/briefings.html#bgp-stream

Chapter 4

Taxonomy

Our classification of BGP update messages is based on the effects of messages on routers' RIBs. For example, (1) a BGP message may provide a new path to reach a known IP prefix or (2) signal a new routed prefix to be added in the RIB. For the first case the RIB is updated with a new path whereas for the second case a new entry is added to the RIB.

We have identified 17 different classes of update message and organized them as a tree, with four level of details (see Figure 4.1). Classes close to the root of the tree are very generic and the leaves stand for the most descriptive classes. These classes are not exclusive, a BGP message may result in multiple changes in the RIB. Therefore, a message may correspond to multiple classes in the taxonomy.

4.1 Change Size

Starting from the left hand side of our hierarchical taxonomy (Figure 4.1) the first generic class is *Change Size*. This class represents all update messages that affect the growth of the RIB. These messages are either increasing or decreasing the size of the RIB which are represented by two different sub-classes:

Remove Prefix stands for BGP messages that discard entries in the RIB, thus decrease its size. These BGP messages are explicit withdrawals for routes that are registered in the RIB. Withdrawals for IP prefixes that are not registered in the RIB are not classified as *Remove Prefix* (see the description for *Duplicate Withdrawal* below).

New Prefix stands for BGP messages that result in new entries in the RIB, thus increase its size. These BGP messages signal the reachability to a new IP prefix or the fragmentation of known IP prefixes into smaller prefixes [11, 14].

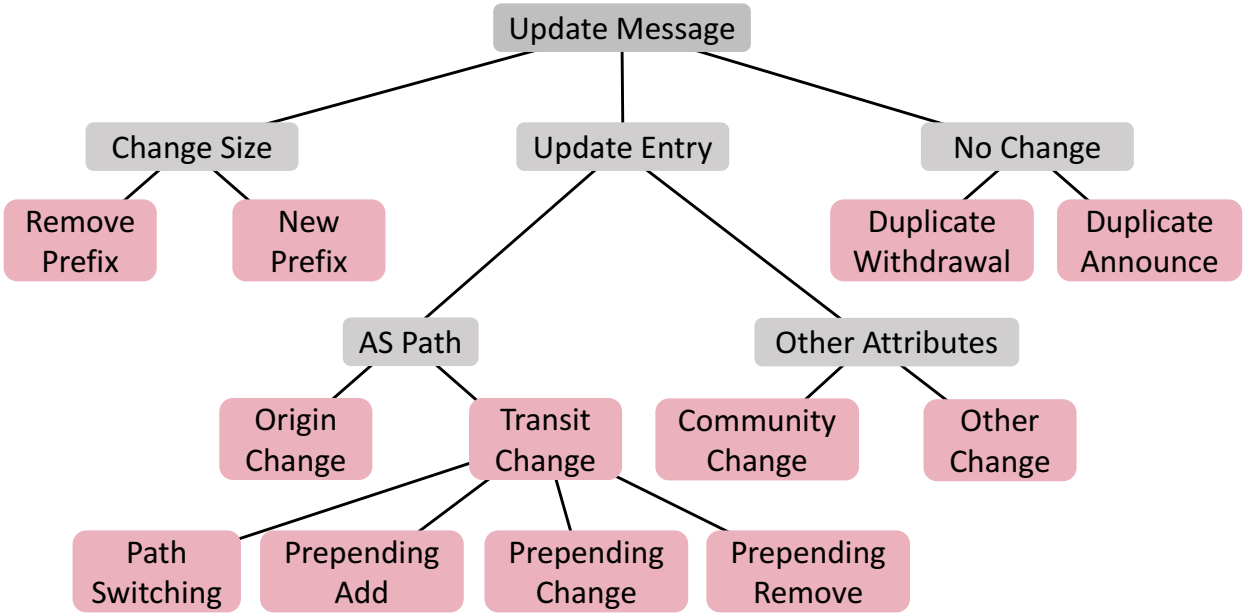

Fig. 4.1: **Hierarchical taxonomy for BGP update messages.** Classes are based on the differences between a BGP message and the corresponding entry in the RIB. The classes are not mutually exclusive, several classes can be assigned to a single update message. Pink nodes represents labels reported by BLT.

4.2 Update Entry

The second generic class is *Update Entry*. This class represents all update messages that modify BGP attributes stored in the RIB. Since RIBs holds multiple attributes for each IP prefix, this class is further decomposed in multiple sub-classes.

4.2.1 AS Path

The AS path is probably the most important attribute in BGP, changes to the AS path have a direct impact on the way traffic is routed. It also discloses a lot of information related to the ASes on the path, for example, AS business relationships [15] and traffic engineering [14].

The *AS Path* class represents any path change observed for IP prefixes registered in the RIB. We further categorize these changes into two sub-classes: *Transit Change* and *Origin Change*.

Transit Change represents any modification made to the AS path except the origin AS, namely the last AS in the path. This class is composed of four sub-classes.

Path Switching represents messages that advertise an AS path that is different from the one registered in the RIB but is the same as the one previously registered in the RIB. These type of

messages are mainly revealing route flaps due to hardware or software problems [1].
Prepending Add/Change/Remove exhibit all changes related to AS path prepending. AS path prepending consists in adding multiple times the same AS in the AS path so that the path seems longer hence less preferable in the path selection process. This is a common traffic engineering technique to setup backup links or avoid a certain path.
Origin Change stands for messages that advertise an AS path where the origin AS (i.e. the last AS in the path) is different than the one stored in the RIB. This class of message signals IP prefixes migrating to a different AS. It also can be a sign of unintentional or malicious prefix hijacks [12].

4.2.2 Other Attributes

Entry updates that are not changing the AS path are classified as *Other Attributes*. Here we essentially distinguish between BGP communities updates and other changes.
Community Change represents messages with BGP communities that differ from the ones registered in the corresponding RIB entry. BGP communities increase greatly the information carried by an update message. For example, a recent study leverages BGP communities to pinpoint peering facilities traversed by an advertised AS path [13].
Other Change stands for any attribute change except for the AS path and community attribute. We group changes made to attributes other than the AS path and BGP communities because they represent only a very small fraction of observed messages and are usually irrelevant to the analysis of Internet routing.

4.3 No Change

Update messages that advertise the same attributes as the ones found in the corresponding RIB entries are classified in the generic class *No Change*. These superfluous messages are detrimental to routers as they contribute to BGP churn [2]. We further divide this class into two sub-classes:
Duplicate Withdrawal represents messages signaling withdraw for a prefix that is absent from the RIB.
Duplicate Announce represents messages whose attributes are all already registered in the RIB.

Chapter 5

BLT: BGP-Labeling Tool

Using the above taxonomy we developed a BGP message classification tool, named BLT. It classifies BGP update messages so that network operators, or researchers, can filter irrelevant messages and dedicate their efforts only to a certain type of messages. Our implementation of BLT is made publicly available*[1].

5.1 Data Source

BLT is designed as an extension of the BGP framework from CAIDA, BGPStream [16]. It retrieves BGP data using BGPStream and output labeled BGP messages according to the taxonomy presented in Chapter 4.

5.2 Implementation

The classification process consists of four steps illustrated in Figure 5.1.

1) Initialization: BLT retrieves the RIB data corresponding to the BGP collector and timestamp selected by the user. These RIBs are loaded in memory and will be used to compute BGP messages labels.

2) Attributes comparison: BLT retrieves BGP update messages for a selected time frame. The messages are handled in sequential order, the attributes of a message are compared to the attributes of the corresponding entry in a RIB. The differences between the message and the entry are then sent to update the RIB and to the classification step.

3) RIB update: The differences obtained in the previous step represent a change propagated by the routing infrastructure. To classify subsequent BGP messages we update the loaded RIBs with this new piece of information.

4) Classification: The differences between the last update message and the RIBs are also used

*[1] https://github.com/romain-fontugne/blt

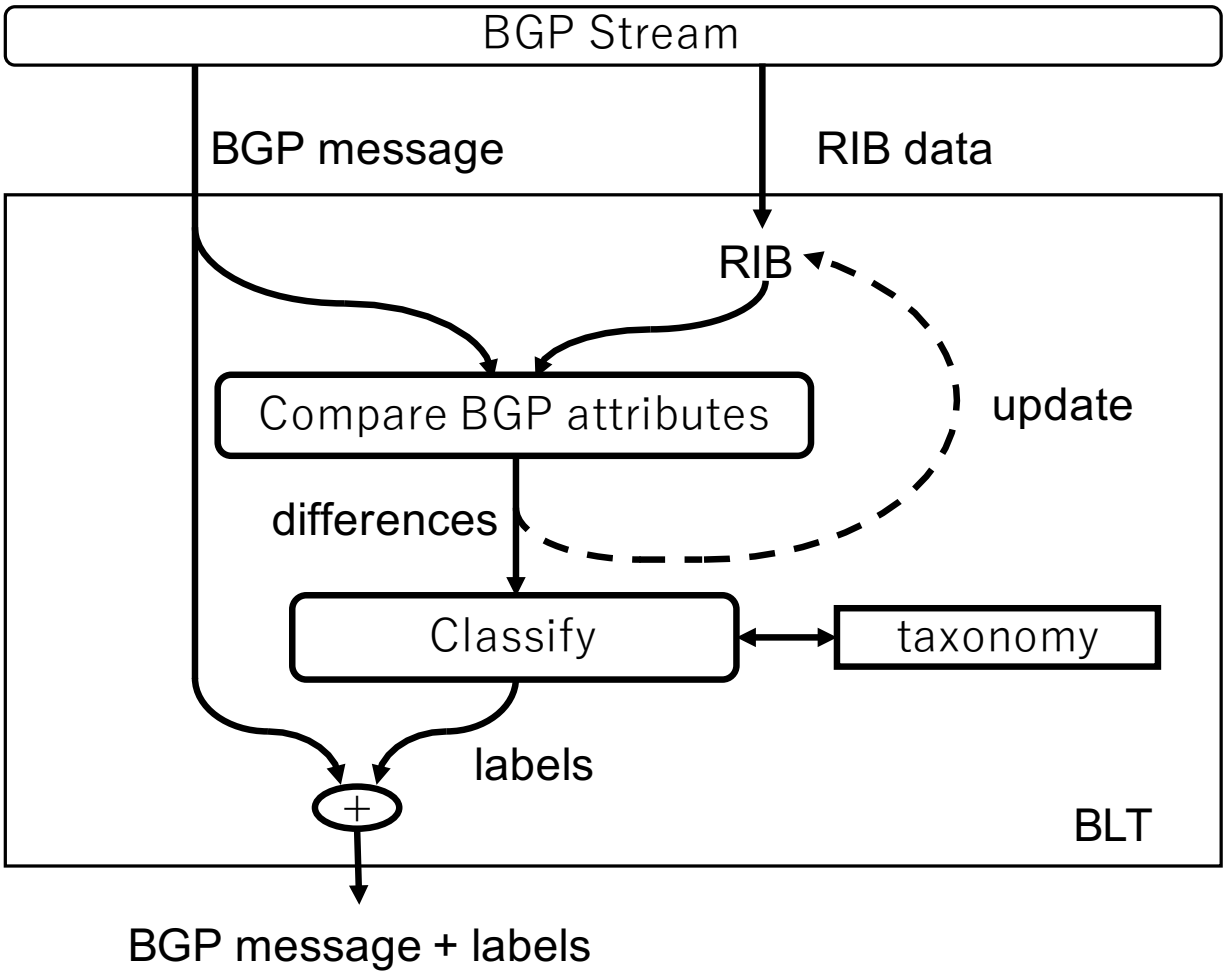

Fig. 5.1: **BLT overview.** Obtain BGP data from BGPStream, classify BGP update messages based on their differences with the local RIBs, and output both BGP messages and labels.

to classify that message. This step is essentially traversing the taxonomy tree (Figure 4.1) and finding nodes that match the observed differences. Only the most specific nodes are reported (i.e. pink nodes in Figure 4.1). For example, if a message signal a prefix advertised from a new origin AS then only the label *Origin Change* is reported (not *AS path* nor *Update Entry*).

Finally, BLT outputs the original BGP update messages retrieved from BGPStream along with the computed labels.

5.3 Usage

The tool can be run only in a python2.x environment. If you want to know how to install python2.x, check the official web site. Use of this tool needs to install following libraries.

- py-radix[*2]
- BGPReader[*3]

You can type below command to annotate BGP messages.

$$python\ bltReader.py\ -v\ [ip\ version]\ -s\ [start\ time]\ -e\ [end\ time]\ -c\ [collector]\ -o\ [path\ of\ output]$$

5.3.1 Arguments

- **-v** The version of Internet Protocol. {-v 4} or {-v 6}.
- **-s** The start time of analysis. The format is {%Y%m%d}. %Y is a year with centry as a dicimal number. %m is a month as a zero-padded decimal number. %d is a day of the month as a zero-padded decimal number.
- **-e** The end time of analysis. The format is the same as that of the start time.
- **-c** The collector you want to use. Available collectors are introduced here[*4].
- **-o** The path you want to output.

[*2] https://pypi.python.org/pypi/py-radix
[*3] https://bgpstream.caida.org/docs/tools/bgpreader
[*4] https://bgpstream.caida.org/data

Chapter 6

Anomaly Detection

6.1 Implementation

To illustrate the benefits of BLT for monitoring the Internet routing infrastructure, we developed a routing anomaly detection method based on BLT results. This application demonstrates the relevance of BLT labels to Internet routing activities and the practical use of BLT for network operators.

This application monitors the proportion of message labels and reports periods of time when the number of message for a certain class is abnormally high. The cause of the detected anomalies differ depending on the reported label. For example, an excessive number of messages labeled as *Duplicate Announce* might reveal noisy BGP messages that might be due to BGP session resets, whereas the surge of messages classified as *New Prefix* might reveal an accidental leak of internal prefixes and more specific prefixes [9].

The principles of the proposed anomaly detector are fairly simple. First, we use BLT to retrieve BGP messages and corresponding labels for a selected time frame and BGP collector. Second, for each message class we model the usual number of messages and report time periods when the data significantly deviates from this computed reference. The reference is obtained from the median number of messages and the median absolute deviation (MAD). These two operators are robust to outlier values [17] and have been extensively employed for anomaly detection [18, 10].

Formally, let $X_l(t)$ be the number of messages classified with label l at the time bin t. Then we define as anomalous a time bin t that satisfies the following equation:

$$X_l(t) > median(X_l) + \tau MAD(X_l)$$

where τ is the sensitivity parameter, and, $median(X_l)$ and $MAD(X_l)$ are, respectively, the median and MAD values for all time bins. In our experiments we set the bin size to ten minutes and the sensitivity parameter $\tau = 10$.

We also make the source code of this anomaly detector publicly available[*1].

6.2 Usage

To use anomaly detector, the conversion from messages annotated to pickle data is needed. convert_blt_to_pickle.py is responsible for this. You can type following command to convert it.

$$python\ convert\_blt\_to\_pickle.py\ [blt\ file\ path]$$

After that, run following command.

$$python\ anomay\_detector.py\ [pickle\ file\ path]$$

Then, you can get figure as I show following chapter.

[*1] https://github.com/tktbtk/BLT-tools

Chapter 7

Evaluation

In this chapter we present several case studies that demonstrate the values of BLT and the proposed anomaly detector to monitor different types of routing anomalies. Chapter 7.2 illustrates results obtained by monitoring BGP update messages for all ASes on the Internet therefore large-scale routing anomalies. On the other hand, in Chapter 7.3 we monitor small set of prefixes and events that affect these prefixes.

7.1 Dataset

The RIBs and BGP update messages analyzed for these case studies are all from the Route Views project [19] which is an archive of BGP data maintained by the University of Oregon. Route Views consists of multiple data sources, in this work we are only analyzing the data collected at the LINX collector. In 2017 this collector contains data from 25 full-feed BGP peers that provide a good representation of Internet AS paths diversity [20]. For each study case we analyze 24 hours of data, namely BLT retrieves the RIB for each BGP peer and the BGP update messages collected in the following 24 hours.

7.2 Monitoring Internet-wide events

To monitor the entire Internet routing infrastructure one can fetch all BGP messages from a set of BGP peers and classify these messages with BLT. We illustrate this, by looking at events that had a global impact on the Internet. The three following case studies are BGP route leaks that happened in 2016 and 2017.

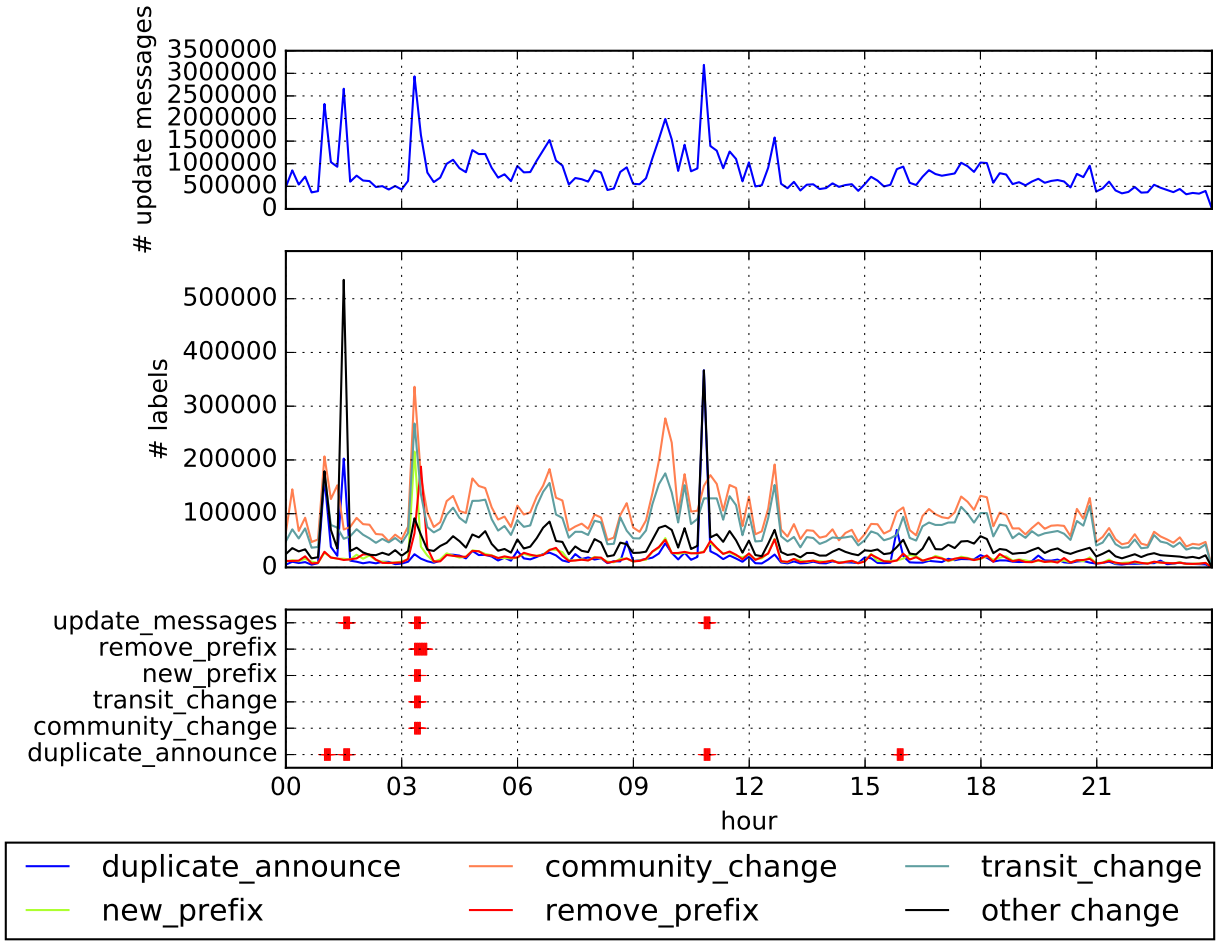

Fig. 7.1: **BGP route leak from Google.** Number of BGP messages observed on August 25th 2017 (top plot), the number of corresponding labels found by BLT (middle plot), and detected anomalies (bottom plot).

7.2.1 BGP route leak from Google

On August 25th 2017 around 3:22 UTC, Google (AS15169) advertised over 150k routes for small prefixes that were presumably used for their internal traffic engineering*[1]. Because these prefixes were longer than corresponding prefixes found in routing tables, numerous ASes have preferred the leaked paths and routed their traffic towards Google's network. This has affected the reachability to the origin ASes of the leaked prefixes and in particular a major access network in Japan, NTT OCN (AS4713).

Using BLT we retrieved the BGP messages received from the Route Views LINX collector on August 25th. Figure 7.1 depicts the total number of messages observed on that day (top plot), the

*[1] https://dyn.com/blog/large-bgp-leak-by-google-disrupts-internet-in-japan/

number of labels assigned to the messages (middle plot) and the results of the anomaly detector (bottom plot). Usually we observe around 300 thousand BGP messages per 10-minute bin for this collector, but the average number of BGP messages per bin is markedly over 800 thousand messages three times during the day (Fig. 7.1 top plot).

The labels obtained with BLT and the results of the anomaly detector (respectively the middle and bottom plot of Figure 7.1) provide a lot more insights into the collected messages. First, through out the entire day the vast majority of the messages are classified as *Community Change* and *Transit Change*. But the three peaks going over 800 thousand messages are mainly due to different types of messages.

The peak around 1:30 and the one around 11:00 are both due to a surge of messages classified as *Duplicate Announce* and *Other Change*. Both peaks are due to a lots of duplicate and change of the next hop attribute from a single BGP peer, this is likely due to an unstable link in that AS. We found this type of events for all the analyzed case studies. Apart from increasing BGP churn, these events are not particularly appealing. They represent no changes on the inter-domain routing infrastructure and can be easily filter out with BLT.

The peak at 3:20 is composed of different classes of messages. This event is first characterized by the outbreak of numerous new prefixes which is due to Google's BGP route leak. Along with these new prefixes we observe the emergence of multiple BGP messages classified as *Transit Changes* and *Community Changes* that reveal messages exchanged during BGP convergence. These events are then followed by numerous withdrawals that correspond to Google's response to mitigate the route leak.

This example clearly illustrates the small number of alerts reported by our detector and its capacity to pinpoint the BGP leak although we are monitoring millions of messages.

7.2.2 BGP route leak from Level(3)

The other BGP leak we look at was initiated by Level(3) on November 6th 2017. Around 17:47 UTC, Level(3) advertised numerous routes that were used for Level(3) internal routing. Similar to Google's leak, these prefixes were longer than previously advertised prefixes so numerous ASes have preferred the paths leaked by Level(3). Comcast connectivity was particularly impacted by this event because a lot of their prefixes had been leaked.

Figure 7.2 illustrates BLT results for the BGP messages gathered by the LINX collector on November 6th 2017. The total number of messages (top plot) shows a few times during the day when the total number of BGP update messages was abnormally high (> 800 thousand messages). BLT labels and the anomaly detector, however, reveal that most of these events are caused by duplicate messages and other changes that are assimilated to BGP noise and flapping routes.

Since the Level(3) BGP leak generated an abnormal number of new prefixes, this event is clearly identified by the anomaly detector (see new_prefix alarms in the bottom plot of Figure 7.2). We

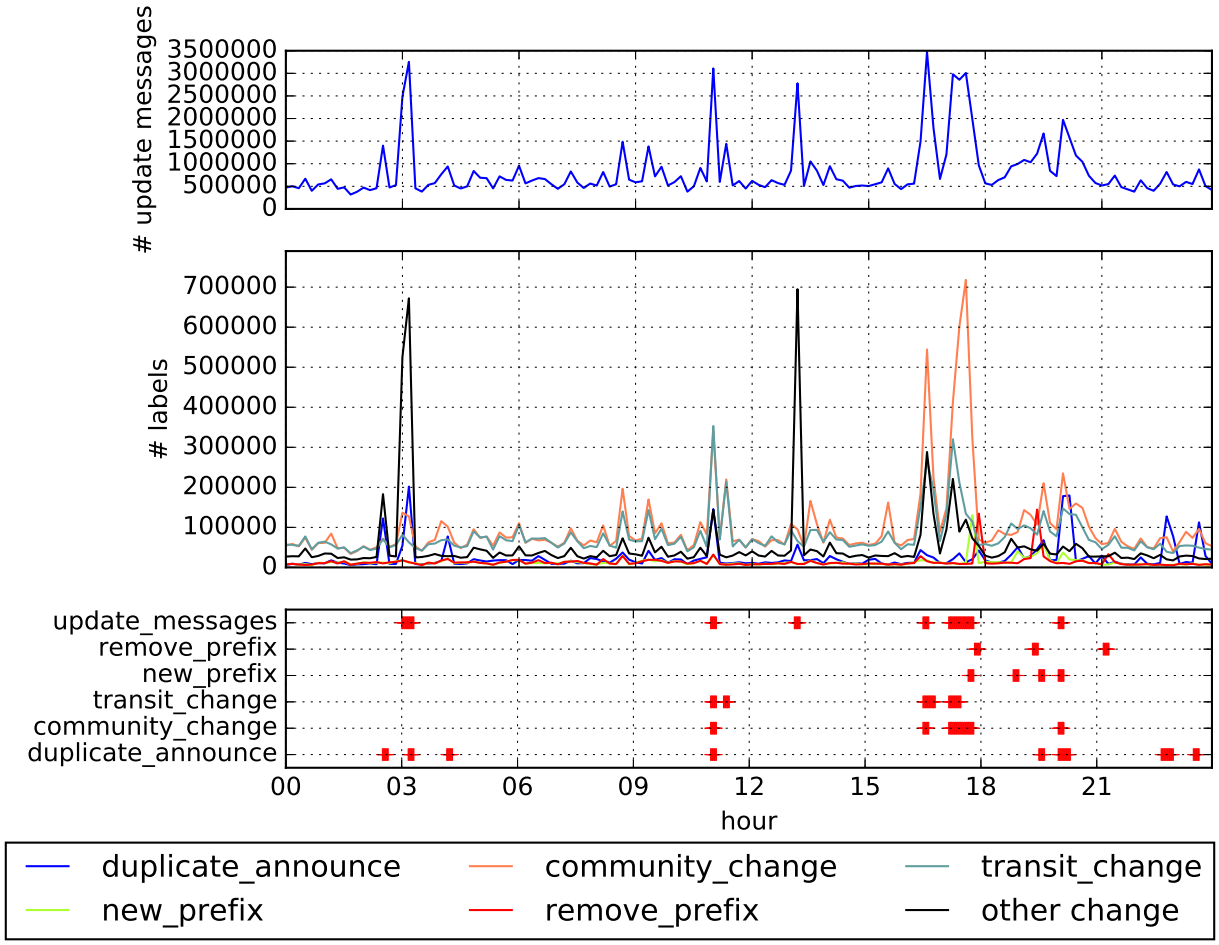

Fig. 7.2: **BGP route leak from Level(3).** Number of BGP messages observed on November 6th 2017 (top plot), the number of corresponding labels found by BLT (middle plot), and detected anomalies (bottom plot).

also observe attempts to mitigate the problem afterwards, just before 18:00 UTC numerous prefixes are withdrawn and again around 19:30 when the problem seemed to have been fixed*[2]. At 21:15 we also found a lot of withdrawn prefixes but only from a single BGP peer so we suppose that event is not related to the BGP leak. After the Level(3) BGP route leak we also observe numerous ASNs advertising smaller prefixes to mitigate the impact of the outage or circumvent impacted ASNs.

7.2.3 Prefix Hijack by Innofield AG

The last Internet-wide case study is a different type of BGP leak. Here the leaking AS is seen as the origin of prefixes that actually belong to other ASes. On April 22nd 2016 at 17:09 a large scale

*[2] `https://blog.thousandeyes.com/comcast-outage-level-3-route-leak/`

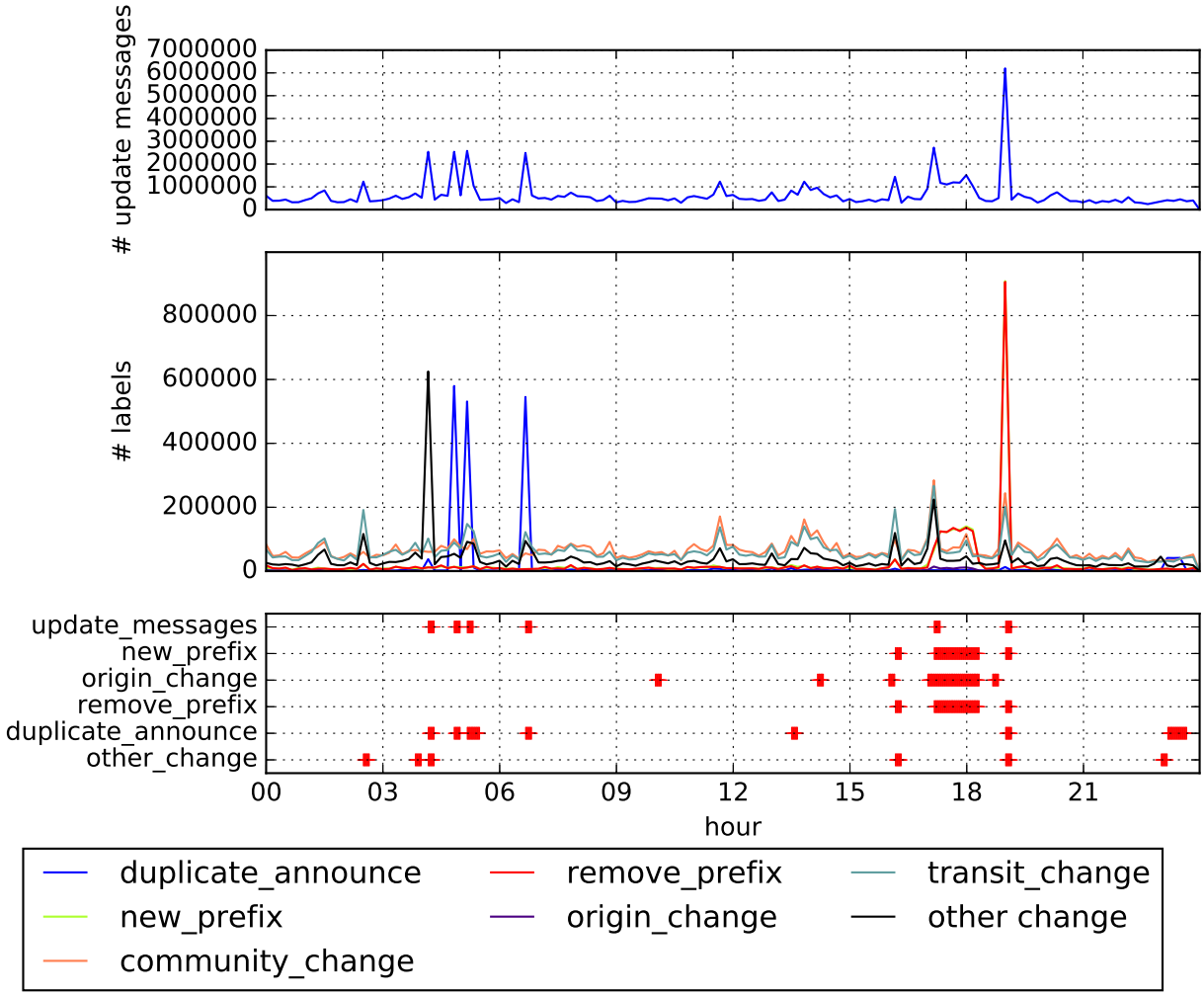

Fig. 7.3: **BGP Hijack of Innofield AG.** Number of BGP messages observed on April 22nd 2016 (top plot), the number of corresponding labels found by BLT (middle plot), and detected anomalies (bottom plot).

routing incident was caused by the Swiss provider Innofield AG. Innofield usually advertises only one IPv4 and one IPv6 prefix but during the incident this AS and its private sibling AS became the origin of 3431 prefixes that are usually announced by 576 other ASes including popular networks like, Google, Amazon, and Facebook*[3].

Figure 7.3 shows the total number of messages counted on that day (top plot), the number of labels reported by BLT (middle plot) and the results of the anomaly detector (bottom plot). The detector reveals surges of *Origin Change*, *New prefix* and *Remove Prefix* messages from 17:00 to 18:20 and around 18:40. The peak of *Origin Change* is caused by Innofield's BGP route leak. Although this event contains much less prefixes that the two previous case studies, this is easily

*[3] https://bgpmon.net/large-hijack-affects-reachability-of-high-traffic-destinations/

identified with BLT as a significant surge of *Origin Change*.

On that day, we also observed three other surges of *Origin Change* around 10:00, 14:10 and 16:00. These three events represent IP prefixes that have moved among the numerous ASes own by the United States Department of Defense and we believe these changes are not related to the Innofield issue.

7.3 Monitoring local routing changes

In this chapter we look at smaller-scale events. These examples illustrate how an operator can leverage BLT to monitor a certain set of prefixes. The following case studies are two outages in 2017, one in Puerto Rico and one in Syria. For monitoring only networks from this countries we retrieve only the BGP messages corresponding to the prefixes originated by these countries. To find the prefixes of a country we rely on the `http://geoinfo.bgpmon.io` service [21].

7.3.1 Outage in Puerto Rico

Hurricane Maria which is recognized as the worst natural disaster in Puerto Rico was originated from tropical wave and caused massive damage on Dominica and Puerto Rico. When making landfall on Puerto Rico, the hurricane caused significant infrastructure damages and disrupted multiple communication lines. On September 20th 2017 about three-quarters of the prefixes in Puerto Rico became unreachable due to hurricane Maria.

Figure 7.4 shows the total number of the messages only for Puerto Rican prefixes on September 20th 2017(top plot), the corresponding labels obtained with BLT (middle plot) and results of the anomaly detector (bottom plot).

Hurricane Maria made landfall in Puerto Rico around 10:15 UTC but we observe first disappearing prefixes from 5:30 UTC, then another set of disappearing prefixes around 8:30, 10:00 and most prefixes around 11:30 (see remove_prefix, Fig. 7.4 bottom plot). In addition to vanishing prefixes, the damages caused a significant number of network changes identified by the anomaly detector as peaks of *Transit Change*. Our manual inspection of the data validate this results as about 50% of prefixes originated from Puerto Rico at 8:30 disappeared by 12:00.

7.3.2 Outage in Syria

The last case study is an outage in Syria that coincide with national examination in that country. There is a few reports on the Syrian government shutting down Internet for the entire country in order to prevent students from cheating[*4]. We believe the following event is also related to the national examinations in Syria.

[*4] `https://motherboard.vice.com/en_us/article/xygv7d/syrian-internet-outages-correspond-exactly-to-nati`

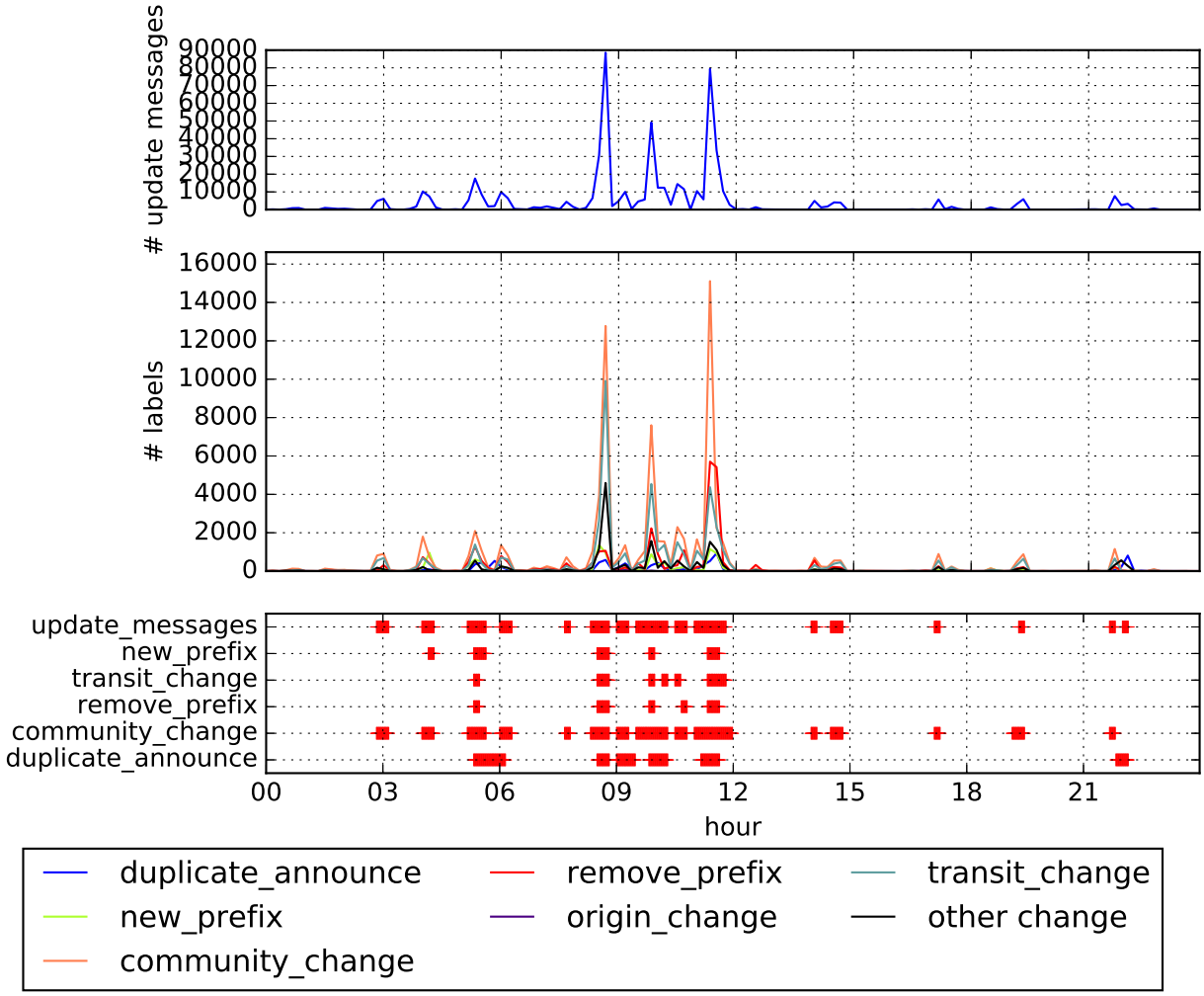

Fig. 7.4: **Outage in Puerto Rico.** Number of BGP messages observed on September 20th 2017 (top plot), the number of corresponding labels found by BLT (middle plot), and detected anomalies (bottom plot).

Figure 7.5 shows the total number of messages on June 1st 2017 (top plot), the number of labels obtained by BLT (middle plot) and the output of the anomaly detector (bottom plot).

We observe only two large peaks of messages, one around 01:00 and another at 5:30. For the first one, a lot of *New Prefix*, *Transit Change*, *Remove Prefix* and *Community Change* messages occur at the same time. This correspond to Syrian prefixes vanishing from routers' RIBs (*Remove Prefix*) and corresponding churn caused by path hunting [8].

The second peak, around 5:30, occurs when disappeared prefixes are re-announced on BGP. This peak is composed mainly of *New Prefix*, *Transit Change* and, *Community Change* messages. *New Prefix* messages simply correspond to the first messages announcing the reappearance of the Syrian prefixes. When these prefixes are re-announced, BGP also seeks for the best paths to these prefixes. The convergence phase of BGP is characterized by numerous *Transit Change* and

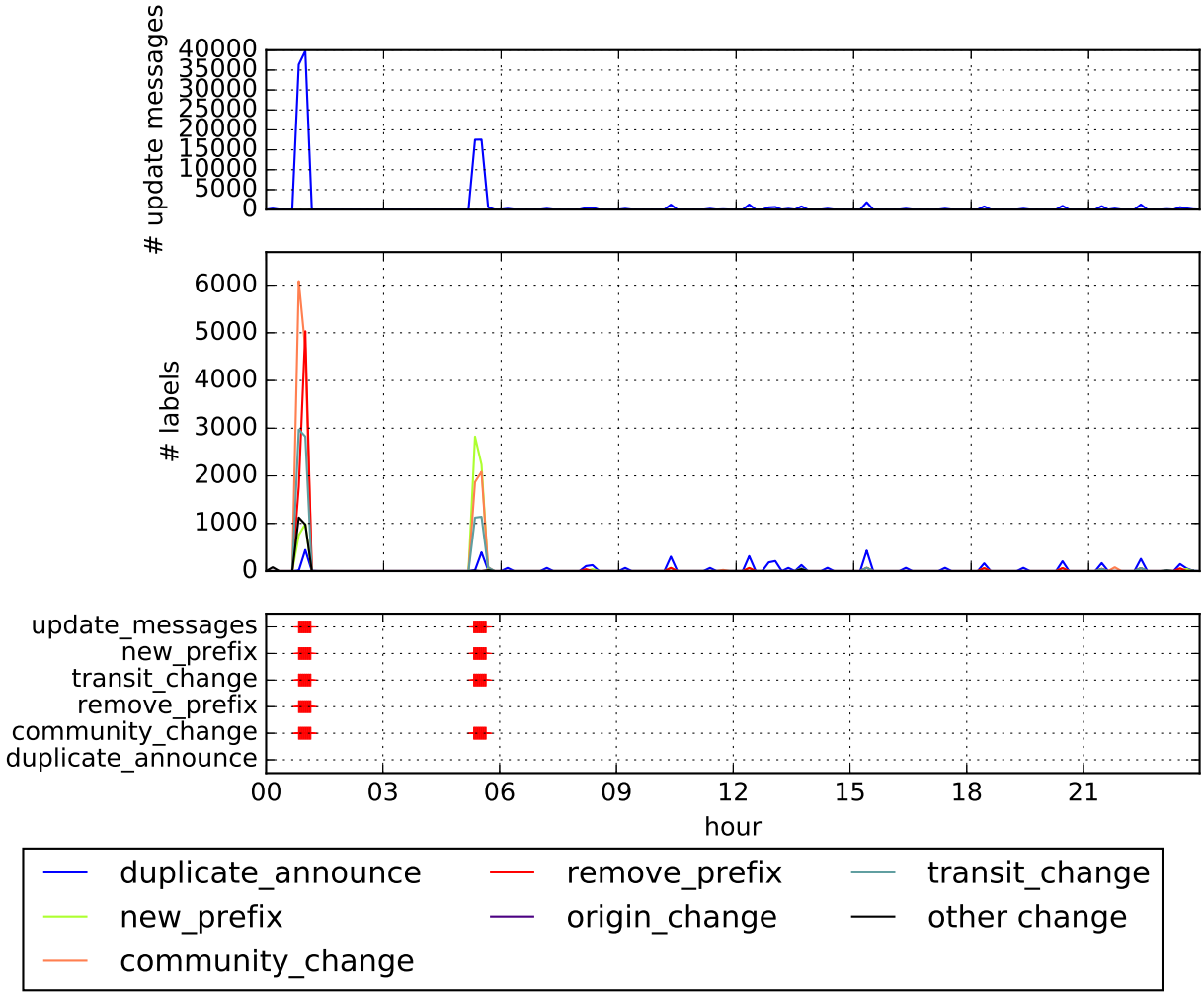

Fig. 7.5: **outage in Syria** Number of BGP messages observed on June 1st 2017 (top plot), the number of corresponding labels found by BLT (middle plot), and detected anomalies (bottom plot).

Community Change messages appearing synchronously with the emergence of new prefixes.

7.4 Monitoring 1 week

To find much more benefits of our taxonomy and BLT, We investigated how much unappealing alerts which are not related with incidentes directly are issued by the detector based on merely counting BGP messages conparing with our anomaly detector based on the classification. Surprisingly, they account for 81% of the total by counting them for 1 week from August 22, 2017 to August 28, 2017. That mean the detector simply couting the messages has a lot of false positives, and our detector based on the taxonomy can pinpoint the incidents although there are millions of messages.

Chapter 8

Discussion

Applying our taxonomy to an anomaly detector enables to clarify details of peaks of BGP update messages. The results of the anomaly detector show that most of the alerts issued by the detector on the basis of merely couting BGP messages are caused by messages which are not particularly appealing, such as *duplicate annouces*, *community changes* and *attribute changes*. The important information to detect incidents is not in such redundant messages but messages such as *transit change*, *new prefix*, *origin change* and *remove prefix*, which can be found in the results of the anomaly detection. Analysis of anomaly type sequences based on the taxonomy may decrease the number of false positives of incident detection. This is one of the future works of this study.

Another future work is an online realtime detection. In this work, we developed the anomaly detector whose threshold is calculated from the data on whole day from 00:00 to 23:59 as an evaluation of the taxonomy and obtained benefits of taxonomy from the anomaly detector. However, what is inherently sought for the use of an anomaly detector is immediate alerting. The online realtime detection which always monitors some collectors using BLT and our anomaly detector enables to report anomalies of labels soon. This could provide intuitive understanding of time series of BGP messages for operators or researchers in real time.

We focused on how BGP messages affect a routing table. As a next step, another taxonomy can be composed from the viewpoint of relations of prefixes. For example, *deaggregated* can be defined as messages that are specific prefixes of the *top* prefix. If route leaks are occured, there will be peaks with *new prefixes* and *deaggregated* because routes of prefixes in a private AS are propagated when occuring route leaks. This would improve the accuracy of the incident detection.

Chapter 9

Conclusion

In this work we presented a general framework to monitor the large number of BGP update messages exchanged by routers. First we introduced a hierarchical taxonomy of BGP messages based on the effects of messages on router's RIBs. Then we developed BLT, a classification tool based on our taxonomy, which enabels network operators or researchers to filter irrelevant messages and concentrate few type of messages. And finally we proposed a simple anomaly detector to monitor significant events in the data as one of the applications of the taxonomy and BLT. We illustrated the benefits of this framework with five case studies. The classification of messages allows one to filter out superfluous messages and focus only on relevant ones.

References

[1] Ahmed Elmokashfi, Amund Kvalbein, and Constantine Dovrolis. Bgp churn evolution: a perspective from the core. *IEEE/ACM Transactions on Networking (ToN)*, 20(2):571–584, 2012.

[2] Ahmed Elmokashfi and Amogh Dhamdhere. Revisiting bgp churn growth. *ACM SIGCOMM Computer Communication Review*, 44(1):5–12, 2013.

[3] Yakov Rekhter, Susan Hares, and Dr. Tony Li. A Border Gateway Protocol 4 (BGP-4). RFC 4271, January 2006.

[4] C. Labovitz, G. R. Malan, and F. Jahanian. Internet routing instability. *IEEE/ACM Transactions on Networking*, 6(5):515–528, Oct 1998.

[5] Jun Li, Michael Guidero, Zhen Wu, Eric Purpus, and Toby Ehrenkranz. Bgp routing dynamics revisited. *ACM SIGCOMM Computer Communication Review*, 37(2):5–16, 2007.

[6] Jong Han Park, Dan Jen, Mohit Lad, Shane Amante, Danny McPherson, and Lixia Zhang. Investigating occurrence of duplicate updates in bgp announcements. In Arvind Krishnamurthy and Bernhard Plattner, editors, *Passive and Active Measurement*, pages 11–20, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[7] David Hauweele, Bruno Quoitin, Cristel Pelsser, and Randy Bush. The Origin of BGP Duplicates. In *CoRes 2016*, Bayonne, France, 2016.

[8] Geoff Huston, Mattia Rossi, and Grenville Armitage. A technique for reducing bgp update announcements through path exploration damping. *IEEE Journal on Selected Areas in Communications*, 28(8):1271–1286, 2010.

[9] Kotikalapudi Sriram, Doug Montgomery, Danny R. McPherson, Eric Osterweil, and Brian Dickson. Problem Definition and Classification of BGP Route Leaks. RFC 7908, June 2016.

[10] Romain Fontugne, Cristel Pelsser, Emile Aben, and Randy Bush. Pinpointing delay and forwarding anomalies using large-scale traceroute measurements. In *Proceedings of the 2017 Internet Measurement Conference*, IMC '17, pages 15–28, New York, NY, USA, 2017. ACM.

[11] Luca Cittadini, Wolfgang MuÌĹhlbauer, Steve Uhlig, Randy Bush, Pierre François, and Olaf Maennel. Evolution of internet address space deaggregation: Myths and reality. *IEEE JSAC*, 8(28):1238–1249, 2010.

[12] Xingang Shi, Yang Xiang, Zhiliang Wang, Xia Yin, and Jianping Wu. Detecting prefix

hijackings in the internet with argus. In *Proceedings of the 2012 ACM conference on Internet measurement conference*, pages 15–28. ACM, 2012.

[13] Vasileios Giotsas, Christoph Dietzel, Georgios Smaragdakis, Anja Feldmann, Arthur Berger, and Emile Aben. Detecting peering infrastructure outages in the wild. In *ACM SIGCOMM'17*, pages 446–459. ACM, 2017.

[14] Julien Gamba, Romain Fontugne, Cristel Pelsser, Randy Bush, and Emile Aben. Bgp table fragmentation: what & who? In *CoRes*, 2017.

[15] Lixin Gao. On inferring autonomous system relationships in the internet. *IEEE/ACM Transactions on Networking (ToN)*, 9(6):733–745, 2001.

[16] Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas, and Alberto Dainotti. Bgpstream: a software framework for live and historical bgp data analysis. In *Proceedings of the 2016 ACM on Internet Measurement Conference*, pages 429–444. ACM, 2016.

[17] Rand R Wilcox. *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. Springer Science & Business Media, 2010.

[18] Romain Fontugne, Patrice Abry, Kensuke Fukuda, Pierre Borgnat, Johan Mazel, Herwig Wendt, and Darryl Veitch. Random projection and multiscale wavelet leader based anomaly detection and address identification in internet traffic. In *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, pages 5530–5534. IEEE, 2015.

[19] The RouteViews project. *http://www.routeviews.org/*.

[20] Romain Fontugne, Anant Shah, and Emile Aben. AS hegemony: A robust metric for AS centrality. In *SIGCOMM Posters and Demos*, pages 48–50. ACM, 2017.

[21] Anant Shah, Romain Fontugne, and Christos Papadopoulos. Towards characterizing international routing detours. In *Proceedings of the 12th Asian Internet Engineering Conference*, pages 17–24. ACM, 2016.

Acknowledgements

I would first like to express my profound gratitude to Professor Hiroshi Esaki for providing me a precious oportunity of study as a master student in his laboratory. I would like to thank appreciation to Associate Professor Hideya Ochiai, Dr.Tsukada. They taught me a lot of things, such as what is a research and how to compose thesis. I especially would like to express my deepest appreciation to Romain Fontugne at Internet Initiative Japan Innovation Institute for leading my work to success and giving me a lot of beneficial advices. I also deeply grateful to to Satoru Kobayashi to have a discussion about my research and giving me a lot of advices. I would like to thank all of our lab members, especially Takeru Kishimoto, Yu Komohara, Hiroki Sakamoto, Akira Sonobe for being good friends of mine, and making my lab life joyful. Lastly, my heartfelt appreciation goes to my family for their moral support and warm encouragements.