BGP Anomaly Detection Techniques: A Survey

Article *in* IEEE Communications Surveys & Tutorials · October 2016

DOI: 10.1109/COMST.2016.2622240

CITATIONS
110

READS
8,921

3 authors, including:

Bahaa Qasim Musawi
University Of Kufa
31 PUBLICATIONS **229** CITATIONS
SEE PROFILE

Philip Branch
Swinburne University of Technology
113 PUBLICATIONS **2,136** CITATIONS
SEE PROFILE

BGP Anomaly Detection Techniques: A Survey

Bahaa Al-Musawi, Philip Branch, and Grenville Armitage

***Abstract*—The Border Gateway Protocol (BGP) is the Internet's default inter-domain routing protocol that manages connectivity among Autonomous Systems (ASes). Over the past two decades many anomalies of BGP have been identified that threaten its stability and reliability. This paper discusses and classifies these anomalies and discusses the 20 most significant techniques used to identify them. Our classification is based on the broad category of approach, BGP features used to identify the anomaly, effectiveness in identifying the anomaly and effectiveness in identifying which AS was the location of the event that caused the anomaly. We also discuss a number of key requirements for the next generation of BGP anomaly detection techniques.**

***Index Terms*—BGP, inter-domain routing, routing, BGP anomaly, BGP stability, anomaly detection, Internet security.**

I. Introduction

TODAY the Internet provides the communication infrastructure for modern commerce, education, entertainment and health services. Because of society's increasing reliance on the Internet, its reliability and security are of critical concern. The Internet has been subjected to many types of attacks such as Denial of Service (DoS), hijacking of hosts and servers, and threats to routing protocols [1]. The Border Gateway Protocol (BGP) is the most widely used inter-domain routing protocol. The topic of this paper is the detection of BGP anomalies that prevent the successful exchange of network reachability information.

BGP is a path vector protocol responsible for managing Network Reachability Information (NRI) between Autonomous Systems (ASes) with guarantees of avoiding routing loops [2]. An AS is a set of routers under a single administrative authority. ASes are not bound by physical relationships but reflect business and organizational relationships. Internet Service Providers (ISPs) apply routing policies to implement their relationships. ISPs may also apply traffic engineering to control the direction and load balance of traffic through path prepending [3]. In this environment it can be difficult to define what is meant by an anomaly. To obtain a satisfactory definition, it is necessary to consider the purpose of BGP and how specific BGP activity does or does not contribute to that purpose. BGP's purpose is to further the business goals of an organisation in providing its NRI to other organisations. Any BGP activity that does not contribute to those business goals or undermines them can be considered anomalous. Unfortunately, it can be very difficult to determine whether or not particular activity is or is not furthering those goals. For example, BGP updates that do not reflect underlying topology changes may be anomalous. They might be the consequence of route flapping where routes are repeatedly announced and then soon after withdrawn. Such activity is anomalous. However, changes that do not reflect underlying topology changes might also be a consequence of traffic engineering where some routes are preferred over others because they use under-utilized link capacity. Such activity is not anomalous even though it may not reflect underlying topology changes.

Even when BGP activity does not contribute to the business goals of the AS not all such activity can be regarded as being of equal significance. There is a spectrum of anomalous behaviour from relatively harmless to highly harmful. For example, route flapping, although it consumes router and link resources, is relatively harmless. Further along the harm spectrum might be path announcements that add unnecessary delay to routed packets. Further still might be path announcements whose purpose is directing traffic via nodes where it can be collected and exploited for surveillance or intelligence purposes. At the far end of the spectrum might be where routes are announced that direct the traffic to a destination where it is dropped ("blackhole-ing").

In this paper we differentiate between anomalous behaviour that does not threaten BGP's ability to disseminate accurate NRI and harmful anomalies that do. We define BGP traffic generated by the first type as an instability. We refer to the second type as an anomaly and its consequences of BGP traffic as anomalous traffic. For example, route flapping may cause long term instabilities while traffic engineering may result in short term instabilities. Neither is a direct threat to the ability of BGP to communicate reachability information. However, a misconfiguration by BGP router operators can result in announcing used and/or unused prefixes which is a threat. The process of differentiating between BGP anomaly and instability is a challenge.

BGP is vulnerable to anomalies caused by hijacking, misconfiguration, and DoS attacks. The consequences of these anomalies can range from a single to thousands of anomalous BGP updates. These consequences have threatened the Internet performance and reliability [4]. Recent statistics show approximately 20% of the hijacking and misconfigurations lasted less than 10 minutes but were able to pollute 90% of the Internet in less than 2 minutes [5].

BGP was developed at a time when information provided by an AS could be assumed to be accurate. Consequently, BGP did not provide any authentication measures for advertising routes [6], [7], [8]. Several methods and proposals have since been introduced to improve the security of BGP. Using our

Bahaa Al-Musawi is with the Faculty of Engineering, University of Kufa, Al-Najaf, Iraq. He is also with the Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia, e-mail: balmusawi@swin.edu.au.

Philip Branch is with the Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia, e-mail: pbranch@swin.edu.au.

Grenville Armitage is with the Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia, e-mail: garmitage@swin.edu.au.

definition of anomaly and instability these can be classified into four broad categories: cryptographic based prevention, anomaly mitigation, mitigation of unstable route propagation, and anomaly detection. Cryptographic approaches [9], [10] use Public Key Infrastructure (PKI) to ensure the authentication of routing announcements to minimize the risk of hijacking. Anomaly mitigation approaches [11], [12] propose ignoring or delaying suspicious route updates by the operators after detecting them. None of these approaches, however, offer a combination of suitable performance, adequate security, and deployable support infrastructure [8]. Moreover, these proposals are not able to mitigate BGP misconfiguration and some forms of hijack (explained in Section III-A) [13], [14]. Mitigating propagation of unstable routes such as described in [15] and [16] has been proposed as a way of limiting the propagation of unstable BGP routing information. Finally, anomaly detection approaches such as [4], [5], [17] aim to discover anomalous information or behaviour in BGP traffic and raise an alarm or take other action.

In the years since it was widely deployed, many types of anomalies have been recorded, including hijacking, misconfiguration by operators, link failure and DoS attacks. It is worth noting that it is not just direct attacks on BGP that can cause anomalies and instability. Although malware such as Nimda and Slammer were directed at web servers, BGP routing was also affected during these attacks [18], [19]. An example of misconfiguration is the Pakistan Telecom incident. In response to a censorship order from its government, the major ISP in Pakistan advertised an unauthorised YouTube prefix causing many ASes to lose access to the site [20]. The panix.com domain incident is an example of hijacking. On 22 January 2006 AS27506 hijacked the panix.com domain causing loss of connectivity for several hours [21]. Other hijacks have continued for long periods without detection. An example is the Link Telecom incident, when an attacker obtained control of the company's prefixes for approximately 6 months and used them to send spam e-mails [22].

In addition to reported events, many events are unreported or even unnoticed [23]. A recent statistical analysis on BGP performance for a period of 10 years beginning from 2001 shows that the huge growth in the size of the Internet was leading towards increased instability [24]. Shi et al. [5] presented statistics and trends of bogus routes in the Internet over a period of 1 year from May 2012. During this period, around 40k bogus routes were detected. Among the causes of these bogus routes, there were 193 BGP hijacks and 27 misconfigurations.

There are many surveys of BGP security (such as [7] and [8]) and anomaly detection in a broad range of disciplines (such as [25]). But as far as we know this is the first survey of BGP anomaly detection techniques. The contribution of this survey is in four areas. Firstly, it classifies BGP anomalies into four main categories, these being direct intended anomaly, direct unintended anomaly, indirect anomaly, and link failure. Secondly, it classifies BGP data sources for detection BGP anomaly into three main categories: BGP raw data and route registry database as well as other less commonly used sources. Thirdly, it explores BGP anomaly detection techniques in term of their approaches, type of BGP data and features used, ability to identify different types of anomalies and their source causes, the network from which the anomaly originated. Finally, it discusses a number of key requirements for a next generation of BGP anomaly detection techniques, these being real-time (in seconds) detection, differentiation between types of BGP anomalies, and identification of the source network of the BGP anomaly.

The rest of this paper is organized as follows. Section II presents a brief overview of BGP. Section III provides a detailed summary of different types of BGP anomalies while section IV discusses different BGP data sources and features. In section V, we review the major approaches to detecting BGP anomalies. In section VI, we discuss a number of key requirements for the next generation of BGP anomaly detection techniques and show a summary of strengths and weakness of current BGP anomaly detection techniques. The paper concludes with section VII.

II. Border Gateway Protocol (BGP)

A. The Architecture of BGP

The Internet is a decentralized global network comprised of tens of thousands of Autonomous Systems (ASes). An AS is a set of routers under a single technical administration using an Interior Gateway Protocol (IGP) such as Open Shortest Path First (OSPF) to communicate with other routers within the AS and an Exterior Gateway Protocol (EGP) such as Border Gateway Protocol (BGP) to communicate with other ASes. Routing protocols are classified into three main types based on their algorithm: link state such as OSPF, distance vector such as Routing Information Protocol (RIP), and path vector such as BGP. BGP has two forms: Internal BGP (IBGP), running between BGP routers within an AS, and External BGP (EBGP), running between BGP routers within different ASes. However, ASes are peered together through a dedicated connection between peers or by a third party such as Internet Exchange Point (IXP). BGP has undergone a number of revisions and refinements over the years. The current version of BGP is version 4 documented in RFC4271 [2].

BGP is the Internet's default EGP. It maintains and exchanges network reachability information between ASes which are organized in a hierarchical fashion. As with IP addresses, each AS has a unique identifier called the AS number, taken from either public or private AS number space [26]. Original AS numbers were 2-bytes and ranged from 0 to 65535. Due to growth in demand, 4-byte AS numbers were subsequently introduced ranging from 0 to 4294967295 [27]. The Internet Assigned Number Authority (IANA) has reserved, for private use, the last 1023 numbers of 2-byte AS numbers, namely 64512-65534, and the last 94967295 numbers of 4-byte AS numbers, namely 4200000000-4294967294 [28]. Each AS has got a range of IP addresses identified by a prefix. For example, the IPv4 address prefix 192.2.2.0/24 refers to all addresses in the range 192.2.2.0-192.2.2.255 while the IPv6 address prefix 2001::/19 refers to all addresses in the range 2001:: to 2001:1fff:ffff:ffff:ffff:ffff:ffff:ffff. BGP provides a set of mechanisms for supporting Classless Inter-Domain Routing

(CIDR) described in RFC4632 [29]. These mechanisms include aggregation support of routes with their AS-PATH (a BGP's attribute described later in Section II-D) and advertising support for a set of destinations as a prefix. Aggregation is the process of combining the characteristics of several routes with common addresses into a single route. This helps reduce the amount of routing messages as well as the number of advertised routes.

IANA manages a variety of activities such as domain names, prefixes, and AS numbers. IANA delegates allocation of prefixes and AS numbers to five Regional Internet Registries (RIRs). These are Réseaux IP Européens (RIPE) for assigning prefixes and AS numbers for Europe, the American Registry for Internet Numbers (ARIN) manages the IP addresses assignments for North America, the Asia-Pacific Network Information Center (APNIC) assigns IP addresses in Asia and the Pacific Rim, Latin American and Caribbean Internet Address Registry (LACNIC) manages addresses space through the Latin American and Caribbean regions, and the African Internet Numbers Registry (AfriNIC) serves the African region. In some cases, RIRs provide services such as domain names, prefixes, and AS numbers through National Internet Registries (NIR). Figure 1 shows the structure of IANA where RIR such as APNIC assigns blocks of IP addresses and AS numbers to NIRs and ISPs.

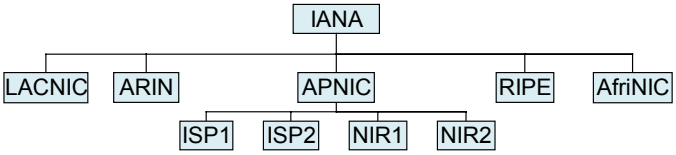

Figure 1. Address distribution hierarchy for the Internet

B. BGP Messages (Update and RIB Table)

BGP is an incremental protocol where after a complete exchange of routing table or Routing Information Base (RIB), only changes to the routing table information are exchanged through announcement messages, withdrawal messages or an update of existing route attributes. RIB for a BGP speaker (a router or a device that runs BGP) consists of Adj-RIBs-In, Adj-RIBs-Out, and Loc-RIB. Adj-RIBs-In refers to routing information that is learned from (adjacent) neighbors. Adj-RIBs-Out refers to routing information that is ready for advertisement to (adjacent) peers while Loc-RIB refers to the routes that will be used by the local BGP speaker based on its local policies and Adj-RIBs-In received [2].

BGP uses the Transmission Control Protocol (TCP) with TCP port number 179 [2]. Using TCP as a transport protocol avoids the need for BGP to manage message delivery and flow control between its peers and eliminates extra data used to confirm connection reliability. The size of BGP messages ranges from 19 octets, containing only a BGP header, to 4096 octets. Regardless of type, each message has a fixed size header as shown in Figure 2.

The first 16 octets are all ones to mark the start of a message. While the length field represents the total message length, the type field refers to one of four possibilities: OPEN, UPDATE, NOTIFICATION, and KEEPALIVE. OPEN message is the first message sent after establishing a TCP connection between two peers. When the other side accepts this message, KEEPALIVEs are periodically transmitted to confirm the connection. Figure 3 shows BGP OPEN message format for a 2-byte AS number. A NOTIFICATION message supplies information regarding a terminated session.

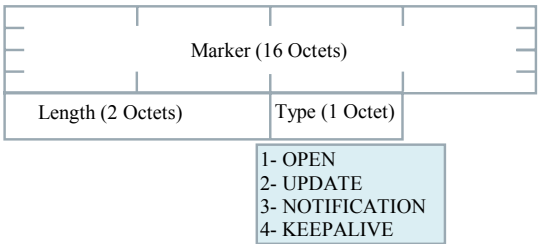

Figure 2. BGP common message header format

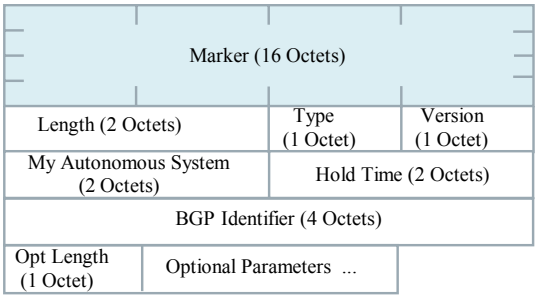

Figure 3. BGP open message format

The most important message is the UPDATE message which is used to announce a new route, withdraw a route that was advertised previously, or update an existing route with new parameters. An AS can withdraw an announced route if and only if that AS previously advertised it. Also, an AS can announce or withdraw multiple routes that have the same path attributes.

Two identities for BGP speaker are represented in the OPEN message: "My Autonomous System" refers to AS number of the sender and BGP identifier described in [30], a unique identifier within an AS where its value is determined on startup and is the same for every local interface and BGP peer.

C. BGP Policies

Routing policy can be defined as how routing decisions are made. It is the exchange of routing information between ASes, where ASes are the unit of routing policy in BGP as stated in RFC1930 [31]. ASes interconnect with each other by different relationships. In general, there are three types of relationships: customer-provider, peer-to-peer, and sibling-to-sibling [32], [33]. In customer-provider relationships, customers pay a fee to their providers for transiting traffic. ASes in peer-to-peer relationships exchange traffic without paying each other. Nevertheless, only traffic originating to or from the peered AS or their downstream customers is accepted. Traffic from their providers or other peers is not accepted. The sibling-to-sibling relationships, which is a rare case, refers

to the relationship between two ASes belonging to the same organization. None of these three relationships are restricted by a physical relationship; they are business and organizational relationships. BGP routing policies are classified into four main classes: business relationship, traffic engineering, scalability, and security-related policies [34]. ISP operators need to configure their BGP routers taking into consideration the four types of policies to enforce their relationships with other ASes. For example, an AS may need to configure its policy so that it does not provide transit services between its providers.

BGP routing policies are based on different BGP attributes; where there is no BGP policy specified, BGP will select a route with a minimum AS-PATH length. However, configuring BGP policies is not an easy task since the number of configuration lines in a single BGP router can range from hundreds to thousands lines [35]. A fault in configuration of a BGP router could produce a local impact or even a global impact. For example, TTNet, an ISP in Turkey, announced more than 100,000 incorrect routes to its peers causing a large number of Internet users to lose connectivity to a large number of domains for several hours [36].

Changing BGP policies can lead to route flapping [37]. To improve Internet stability at Internet edges, Minimum Route Advertisement Interval (MRAI) and Route Flap Damping (RFD) mechanisms have been developed to limit propagation of unstable routes. MRAI specifies a minimum time between which the speaker can send successive update messages. If an announcement is withdrawn within the MRAI neither is forwarded. RFD works on the receiving side using a larger time scale than MRAI. For each peer, RFD monitors the frequency of BGP updates for a given prefix. When the update rate between two BGP peers exceeds a set threshold, the update related to this prefix is suppressed. RFD was introduced in [38] and [39] to improve Internet stability by reducing sustained routing oscillation on network edges. RFD is a technique that has been widely implemented. However, in 2006 RIPE's routing working group recommended against using RFD [40]. Recently there are recommendations to use it but with adjustments to algorithm constants [41], [42]. It is worth noting that between the dates of RFD being introduced and it being mostly deactivated (1998-2006), RFD was not able to mitigate the propagation of unstable routes caused by indirect anomalies (described later in Section III-C).

D. BGP Attributes

BGP attributes are a set of properties carried in a BGP update and used to determine the best route among many possible paths to a specific destination. These attributes are mainly classified into four types: well-known mandatory (should be included in all BGP updates and all BGP speakers can recognise them), well-known discretionary (could be included in a BGP update and all BGP speakers can recognise them), optional transitive (can be recognised by some BGP speakers. They should be accepted and sent to peers even if it is not recognized by BGP peers) and optional non-transitive attributes (can be recognised by some BGP speakers. They can be ignored and not advertised to peers). The most well-known and widely used attributes are: Origin, AS-PATH, LOCAL-PREF, AGGREGATOR, and Multi Exit Discriminator (MED) [43], [2].

Origin is a well-known mandatory attribute created by the BGP speaker that generates the related routing information. It refers to the type of an originated update with three possibilities: 0 refers to an update originating from IGP, 1 refers to an update originating from EGP, and 2 for INCOMPLETE, when a route originates from another routing protocol instead of BGP such as static route.

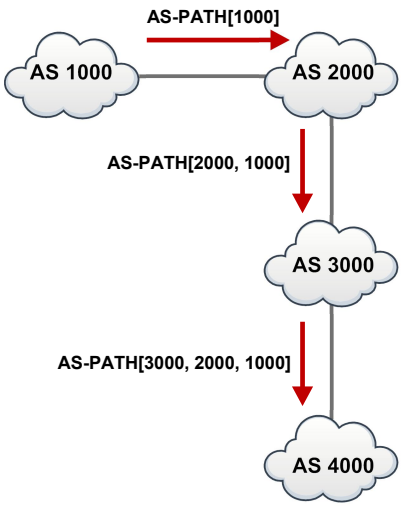

Figure 4. An example of BGP AS-PATH attribute

AS-PATH is a well-known mandatory attribute which identifies a list of ASes that have had an update message passing through their prefixes. The components of this list can be AS-SETs or AS-SEQUENCEs. AS-SET refers to an unordered set of ASes while AS-SEQUENCE refers to an ordered set of ASes. BGP is a path vector protocol where each BGP speaker adds its own AS number in the path of a BGP update before passing it to an EBGP peer. This attribute prevents routing loops between BGP speakers. Figure 4 shows an example of how the AS-PATH attribute works. Each AS inserts its own AS number before propagating a BGP update to its peers. When AS1000 sends a route to AS4000, it adds its own AS number to the beginning of the path. AS2000 receives the update and appends its AS number before passing it to AS3000. Finally, AS3000 receives the update, and inserts its own AS number to send it to AS4000. BGP is a path vector protocol where [3000,2000,1000] shows the full path for an update sent by AS1000 to AS4000.

LOCAL-PREF is a well-known discretionary attribute. LOCAL-PREF represents a degree of preference for a network operator for a route between multiple routes within an AS. A high value of this attribute shows a strong preference for a particular route. For example, in a business relationship ISPs will usually prefer routes learned from their customers over routes learned from a peer; therefore, a high value of LOCAL-PREF in range 99-90 could be assigned for customers, 89-80 for peers, and 79-70 for providers [34]. This attribute was used by PGBGP [12] to mitigate the propagation of suspicious

routes through assigning them with low LOCAL-PREF. This attribute, however, should not be used with external peers except for the BGP confederation case described in RFC5065 [44].

AGGREGATOR is an optional transitive attribute. It contains information about the BGP speaker that aggregate the route. Although the aggregation helps to reduce the number of advertising routes, it can hide AS-PATH and other attributes of the aggregated prefixes. Figure 5 shows an example for route aggregation. In this example, AS1 and AS2 advertise 10.10.3.0/24 and 10.10.4.0/24 respectively to AS3. AS3 aggregates these prefixes by sending the single prefix 10.10.2.0/23. The value of AS-PATH for the single prefix is based on the aggregation configuration at AS3. AS3 can hide the paths to AS1 and AS2 and send the prefix 10.10.2.0/23 with AS-PATH=[3]. This can cause a blackhole if any of the prefixes advertised by AS1 or AS2 are withdrawn. AS3 can also configure the aggregation to include both of the originating ASes as AS-SET, in this case AS4 will receive the prefix 10.10.2.0/23 with AS-PATH=[3,{1,2}].

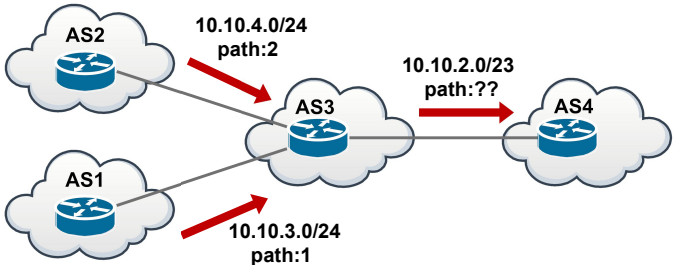

Figure 5. An example of BGP route aggregation

MED is an optional non-transitive attribute which provides a mechanism to influence external neighbors about the preferred path into an AS that has multiple entry points. The MED with the lower metric is preferred as an exit point.

Among these attributes, a BGP router follows a sequence of comparisons to find its best route among various routes based on their attributes. Table I shows the sequence of comparisons.

Table I
BGP PATH SELECTION PRIORITY

Priority	Policy Attribute
1.	Highest LOCAL-PREF value
2.	Lowest AS-PATH length
3.	Lowest Origin Type
4.	Lowest MED value
5.	EBGP learned over IBGP learned
6.	Lowest IGP cost
7.	Lowest Router ID

BGP messages are sent to reflect changes in ASes topology and policy. When a BGP router receives a BGP message that changes its routing table it will propagate that message to all or a group of its neighbors based on its local policies. Otherwise, the message will be terminated. Figure 6 shows the stages of convergence to a new prefix announced by AS1. Firstly, AS1 announces the new prefix with its AS number (10.10.0.0/16 and path:1). When AS2 receives the new announcement, it will check the announcement with its entire RIB table and because this entry is new, AS2 will add and send it to all its neighbors[1]. Secondly, AS3 and AS4 will receive the new announcement with path (2,1) and add it to their RIB as it does not exist in their RIB. In this stage, AS3 and AS4 will propagate the new announcements to their neighbors. BGP can guarantee routes are loop free. Therefore, any BGP updates received by an AS which contains its own AS number will be ignored.

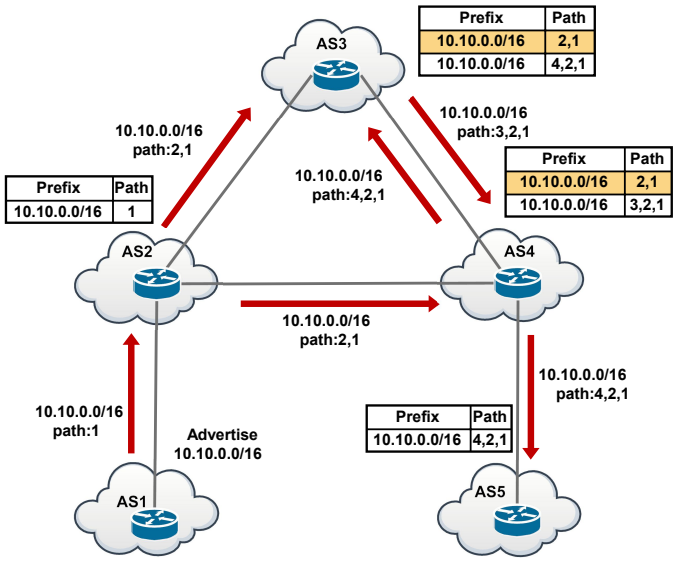

Figure 6. Announcing a new prefix

At the third stage, AS5 will receive the full path to the new prefix from AS4. Meanwhile, AS4 and AS3 receive new updates from AS3 and AS4 for the new prefix with paths (3,2,1) and (4,2,1) respectively but these updates will not be forwarded because they are not the best routes. The highlighted path (2,1) at AS3 and AS4 represents the best route.

III. BGP ANOMALIES

In this paper we refer to harmful changes in BGP behaviour as an anomaly. The consequences of BGP anomalies can range from single to thousands of anomalous BGP updates. A single BGP update is classified as an anomaly if it contains an invalid AS number, invalid or reserved IP prefixes, a prefix announced by an illegitimate AS, AS-PATH without a physical equivalent or which does not match a common routing policy [45]. A set of BGP updates can be classified as an anomaly if its characteristics show a rapid change in the number of BGP updates, containing longest and shortest paths, or changes in the behaviour of total BGP traffic over time [4], [17].

Detecting BGP anomalies enables network operators to protect their network from the worst consequence of the anomalous behaviour. In this survey, we construct a taxonomy of BGP anomalies with four main categories as follows:

1) Direct intended anomaly.
2) Direct unintended anomaly.

[1]Routes are chosen based on the AS-PATH attribute only where other attributes such as MED and LOCAL-PREF are not considered in our examples.

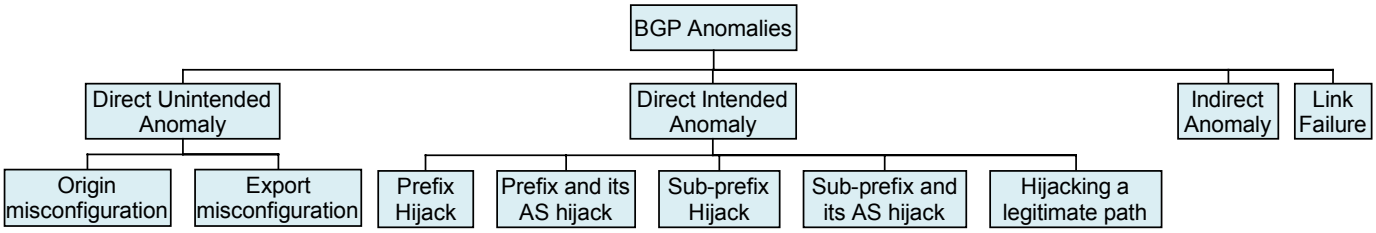

Figure 7. Taxonomy of BGP anomalies

3) Indirect anomaly.
4) Link failure.

These categories can be further classified into subcategories as shown in Figure 7 and discussed in the next section. Each type of anomaly can produce different consequences. For example, a misconfiguration by an ISP can result in announcement of used and/or unused prefixes with a consequent significant increase in BGP volume and a significant change in number of hops in paths to specific prefixes. Table II shows possible consequences for different types of BGP anomalies.

Table II
POSSIBLE CONSEQUENCES OF BGP ANOMALIES

ID	Type of anomaly	Potential consequences of anomaly
1	All types of anomaly except direct intended	A significant change in volume of BGP updates
2	Link failure	Observing rare ASes and long paths in the AS-PATH
3	Direct intended and unintended anomaly	A significant change in number of hops between attended prefixes and monitoring points
4	Direct intended and unintended anomaly	Geographic deviation of intermediate ASes between attended prefixes and monitoring points
5	All types of anomaly	Reachability issues to attended prefixes

A. Direct Intended BGP Anomaly

This type of anomaly refers to all types of BGP hijacking which can appear in different scenarios such as prefix hijack and sub-prefix hijack. Hijacking occurs when an attacker claims to own a prefix or sub-prefix that belongs to another AS causing redirection of routes from the AS to the attacker. Attackers hijack prefixes to produce different malicious activities. For example, the hijacker can blackhole all traffic to the victim causing a DoS for that network. In another scenario, the attacker becomes a man-in-the-middle, intercepting the traffic without affecting victim reachability. Phishing attacks can also be done by hijacking a prefix through redirecting traffic to an incorrect destination. Additionally, the attacker can use stolen IP addresses to send spam [46].

Since BGP was created, many hijacking events have been observed. Notable examples include the following. On the seventh of May 2005 AS174 hijacked one of Google's prefixes causing it to lose connectivity to the google.com domain for nearly an hour [47]. Other events have continued for longer periods of time. For example, in the Link Telecom (AS12812) incident an attacker obtained control of the company's prefixes and AS ownership for approximately 6 months and used them to send spam e-mails. The attacker took advantage of a financial crisis experienced by Link Telecom to send a forged letter of authorization for the AS12812, then started to advertise routes with the hijacked AS and its prefixes [22]. While these incidents are notable as a result of their size and scale, many similar but smaller scale incidents are unreported or even noticed [23].

In addition to illegitimate announcements, some Distributed Denial of Service (DDoS) mitigation services will legitimately advertise sub-prefixes of a particular AS for short periods of time in order to redirect, clean (remove suspicious traffic) and reinject traffic heading towards their customer such as in [48].

The direct intended anomalies are classified into five subtypes: hijacking a prefix, a prefix and its AS, a sub-prefix, a sub-prefix and its AS, and hijacking a legitimate path [49]. To demonstrate these types of hijacking, we use the topology shown in Figure 6.

1) Prefix Hijack: In this type of hijack, an attacker configures its BGP router to announce a prefix belonging to another AS. BGP allows any BGP speaker to announce any route regardless of whether the route actually exists or not [13]; therefore, the attacker's neighbors will adopt it as a new route. Figure 8 shows an example of prefix hijacking. AS4 hijacks the prefix 10.10.0.0/16 belonging to AS1 causing a Multiple Origin AS (MOAS) conflict for other ASes. A MOAS conflict occurs when a particular prefix appears to originate from more than one AS. MOAS conflicts occur legitimately in many cases such as IXP, multi-homing, and anycast. However, identifying a valid MOAS from an attack is difficult [50].

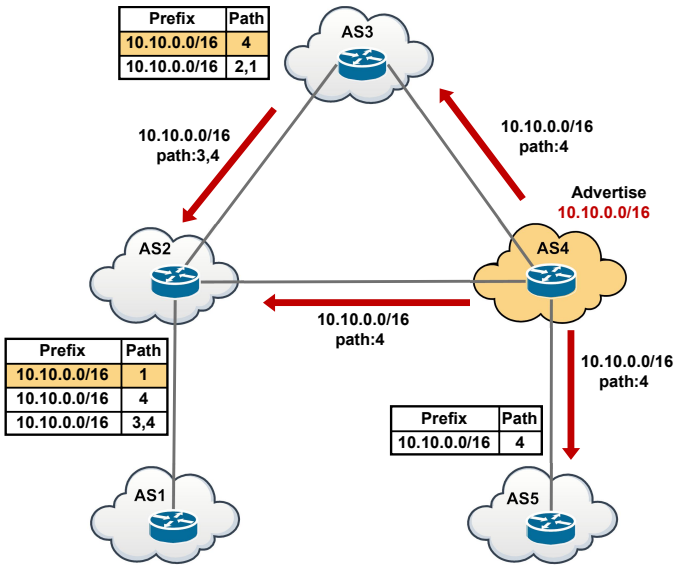

Figure 8. Prefix hijacking

When AS2, AS3 and AS5 receive the AS4 advertisement, they compare the new prefix with their RIB. While AS3 and AS5 update the entry (10.10.0.0/16 with path 4,2,1) to (10.10.0.0/16 with path 4) as they previously received this prefix by AS4 (as discussed in Figure 6), AS2 will add it as a new entry. AS2 will not use it as a best route as it has the same path length. However, in this example, AS5 and AS3 will send all packets that relate to prefix 10.10.0.0/16 to AS4 instead of AS1.

2) Prefix and its AS Hijack: In this scenario an attacker announces that there is a direct connection between its AS and a victim AS causing redirection of routes from the AS to the attacker instead. The attacker tries to avoid a MOAS conflict by sending a fake path with the hijacked prefix. Figure 9 shows an example of hijacking an AS and its prefix. AS4 sends an announcement that it has a connection with AS1. AS2, AS3 and AS5 will receive this update and compare it with their RIB table. While AS5 will use the new announcement as a best route, AS2 and AS3 will not. In this example, only AS5 is affected by the hijack of AS4. AS4 can now carry out malicious activities such as DoS against AS1 and tampering with packets that are sent from AS5 to AS1.

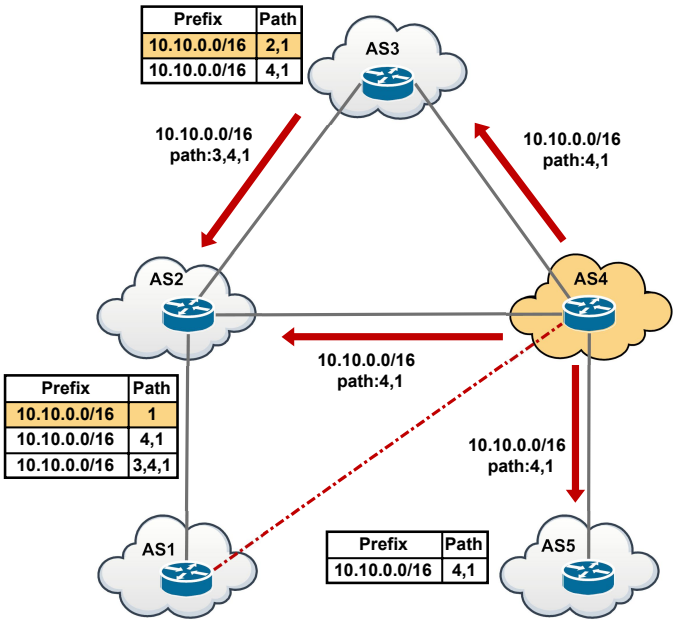

Figure 9. Prefix and its AS hijacking

3) Sub-prefix Hijack: In this scenario an attacker announces a sub-prefix that belongs to a victim AS. BGP selects the most specific address or longest address match. For example, a BGP router will select a specific address such as 10.10.0.0/24 over a more general address such as 10.10.0.0/16. Figure 10 shows an example of this type of hijack where AS4 announces a prefix 10.10.0.0/24 which is a part of the prefix 10.10.0.0/16 owned by AS1. AS2, AS3 and AS5 receive this update and add it as a new entry. Although there is a direct connection between AS1 and AS2, AS2 will send all packets that belong to the prefix 10.10.0.0/24 toward AS4 instead of AS1. This is the most widely propagated type of hijacking since all ASes between the attacker and the victim are affected. Moreover, this type of hijacking can be globally propagated when there is no other advertisement or filtering for this route [49].

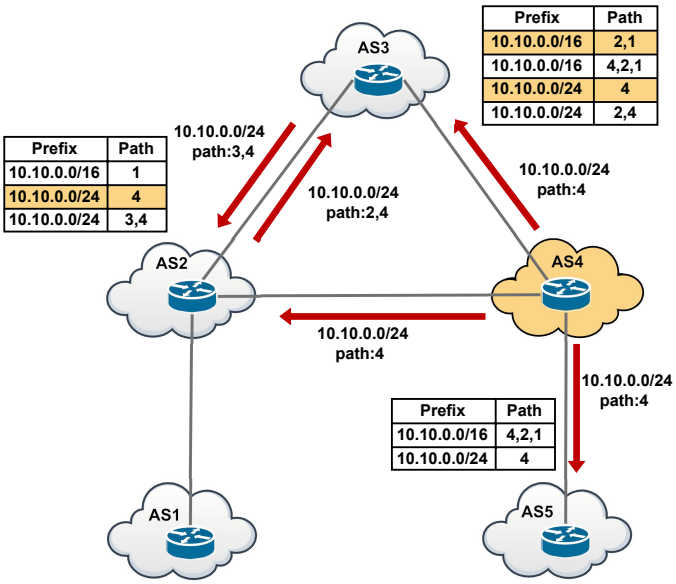

Figure 10. Sub-prefix hijacking

4) Sub-prefix and its AS Hijack: In this scenario, the attacker announces a fake path to a subnet of a target prefix. Using a fake path with sub-prefix hijack represents a critical challenge for detection as the attacker does not claim to own a full prefix length which can be detected using control plane data [51]. Hu and Mao in [49] claim that this type of hijacking is the most difficult to detect.

Figure 11 shows an example of this type of hijacking. AS4 announces it has a path to the prefix 10.10.0.0/24 which is a part of 10.10.0.0/16 owned by AS1. AS2, AS3, and AS5 will add it as a new entry. Although the path length to the address 10.10.0.10 at AS2 is just 1, AS2 uses the longer path (4,1) as the prefix 10.10.0.0/24 is more specific than 10.10.0.0/16.

5) Hijack a Legitimate Path: This type of hijacking does not require any announcements by the attacker. The attacker simply manipulates received updates before propagating them. Figure 12 shows how to accomplish this type of hijacking. AS4 received the update 10.10.0.0/16 with the path (2,1). It will propagate the update 10.10.0.0/16 with the path (4,1) instead of the full path (4,2,1). In this example, only AS5 adopts the manipulated route as a default route. This hijack is similar to the prefix and its AS hijacks but the attacker violates propagation instead of announcing an attractive path. This type of hijacking is one of the key security issues considered by the IETF [52], but it has received little research attention [5].

B. Direct Unintended BGP Anomaly

This type of anomaly refers to BGP misconfiguration by BGP router operators. Faulty configuration of BGP routers can result in announcing used and/or unused prefixes. While announcing used prefixes causes hijacking since the prefixes belong to other ASes, unused prefixes cause leaked routes which may result in an overload or blackhole to other ASes. The effect of announcing used prefixes is similar to 'Prefix Hijack' described earlier but is unintended and can be

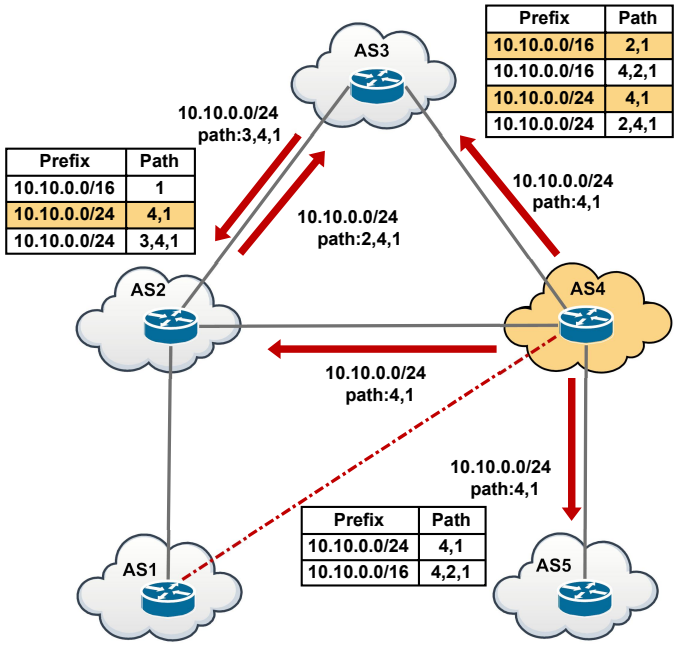

Figure 11. Sub-prefix and its AS hijacking

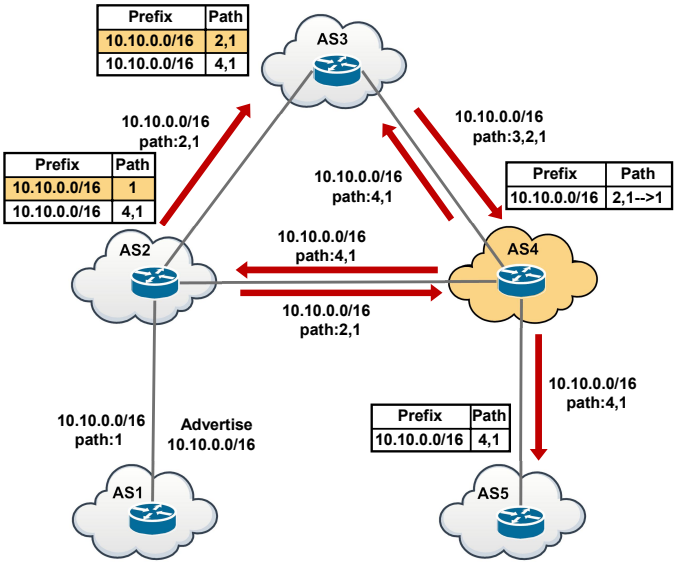

Figure 12. Hijacking a legitimate path

corrected as soon as the operator discovers it. However, the misconfiguration may cause packet loss, unintended paths between hosts, and forwarding loops [35].

Configuring BGP policies is not an easy task as there are many factors that need to be considered such as business relationships, traffic engineering, scalability, and security-related policy [34]. Consequently, this type of anomaly can occur easily. An example is a route leak incident by Dodo, an ISP in Australia, on 23 February 2012. One of Dodo's routers accidentally announced all its internal routes to Telstra, one of the major ISPs in Australia. As Dodo is a Telstra customer, Telstra used the announced routes as its best routes causing loss of internet connectivity in most of Australia for around 45 minutes [53]. Although some tools have been developed to help operators eliminate faults, such as router configuration checker [35], rancid [54], and BGP Visibility Scanner [55], faults in BGP configuration are still frequently seen. Some of these faults have global effects such as two recent incidents by Indosat and Turk Telecom. Indosat, an Indonesian ISP, propagated over 320,000 incorrect routes for more than two hours [56]. Turk Telekom, the major ISP in Turkey, in response to instructions from the government of Turkey to censor twitter.com, accidentally hijacked IP addresses of popular DNS such as 8.8.8.8 and 4.2.2.2 [57].

BGP misconfiguration can be classified into origin misconfiguration and export misconfiguration. Origin misconfiguration occurs when the operator accidentally announces prefix/prefixes that they do not own or fail to filter private ASes. Export misconfiguration occurs when the operators accidentally configure BGP policies, for example, by blocking some authorized routes causing DoS to the blocked prefixes [58], [59]. Each type of misconfigurations has different effects. Deshpande et al. in [59] show that origin misconfiguration causes dangerous fluctuations in BGP routes while export misconfiguration threatens BGP routing convergence.

Direct unintended anomaly also refers to AS number space attribution overlaps between RIRs. In November 2009, it was noted that AS1712 has been used by both Twilight Communications (an organization in Texas assigned by ARIN) and Ecole Nationale Superieure des Telecommunications (an organization in Paris assigned by RIPE) [60]. According to IANA, AS1712 should be assigned by ARIN, not RIPE. This overlap occurred because AS numbers in the 1700's were assigned by RIPE in 1993, before the existence of ARIN, and consequently AS1712 was used by both RIPE and ARIN [61].

C. Indirect Anomaly

This type of anomaly refers to malicious activities directed at Internet components such as web servers. Although BGP is a routing protocol for managing Internet reachability information between ASes, it experienced periods of instability during the Nimda, Code Red II, and Slammer worm attacks. These attacks caused routing overload to the entire Internet through affecting some ASes to send significant numbers of BGP messages [19], [18], [62].

Nimda and Code Red II, two well-known computer worm attacks observed during 2001, were directed at hosts and servers running Microsoft Operating Systems [63]. However, large spikes of BGP messages were also observed during these attacks. During the Nimda attack around 30 times the normal number of BGP updates were observed [19]. Another example of indirect attacks is the Slammer worm attack. Slammer is the fastest computer worm yet seen, infecting more than 90% of vulnerable hosts in around 10 minutes. Although it was not directed at any routing protocol, BGP experienced critical instability during this attack. Lad et al. in [18] show the effect of the Slammer attack on BGP stability. They show a dramatic increasing in BGP update announcements during the attack. The average BGP announcement on some ASes exceeded 4500 updates per prefix compared to an average of 47 updates per prefix on the day before of the Slammer

event. The problem was that no differentiation of BGP routing traffic and normal data traffic was made so that a congested data path led to BGP peer failures since KEEPALIVEs were choked. Today providers usually differentiate between routing/control/management traffic and data traffic to reduce the impact of this class of problem.

D. Link Failure

ASes are peered together by either a private peering, through a dedicated connection between peers, or a public peering by a third party such as an IXP. Failure in one of these connection links (private or public) or one of the Internet core ASes could cause national or global instability for many ASes. The last twenty years have recorded many BGP events caused by link failure. For example, on 25 May 2005 there was a blackout in Moscow causing the MSK-IX, an organization operating Internet Exchange and providing Internet businesses in Moscow and many Russian cities, to be shutdown for several hours. This blackout affected many ISPs in Russia [64]. Other failures have a global effect. For example, in January 2008 a Mediterranean cable break incident caused thousands of networks in more than 20 countries to be unreachable. This outage caused BGP rerouting[2] as a result of losing reachability to these networks [66]. Consequently, thousands of networks were in the situation of sending a high volume of BGP updates to find alternative paths.

IV. BGP DATA SOURCES AND FEATURES

BGP anomaly detection techniques use various sources of BGP data for detection of BGP anomalies. Usually, extracting significant information from BGP data is done during a preliminary stage of anomaly detection. These data sources include BGP raw data, route registries database as well as other types of BGP data sources. Extracting relevant information from BGP data source produces different numbers and types of BGP features which are then used as an input to a BGP anomaly detection technique. In general, these features can be classified into two types related to deviations in number of BGP updates and in the path data contained within the update field AS-PATH. We now discuss BGP data sources and features in more details.

A. BGP Data Sources

As noted, BGP anomaly detection techniques and approaches use various types of BGP data which fall into three main categories: BGP raw data, route registries database, and other less commonly used sources [67]. Below, we describe each of these types of data.

1) BGP Raw Data: There are two types of BGP raw data: control plane, which refers to RIB and/or BGP update messages exchanged between BGP speakers, and data plane, based on the routes that packets use between an observer and the source [68].

a) Control Plane: Control plane data can be obtained from free download repositories such as RouteViews project [69] and Réseaux IP Européens (RIPE) Network Coordinate Centre (NCC) [70] or monitored in a real-time from BGP speakers such as BGPmon [71] in the RouteViews project. The RouteViews and RIPE NCC are the most well-known repositories that provide free download for BGP updates and RIB. RouteViews peers with many sites in north America and had provided BGP data since 2001, while RIPE peers with many sites in Europe and provides BGP data since 1999. The total numbers of collectors and peers change over time as a result of adding/removing some vantage points[3]. The RouteViews repository provides BGP updates every 15 minutes and BGP routing tables every 2 hours. Until June 2003, RIPE was providing offline BGP updates every 15 minutes with BGP routing tables every eight hours. From 2003 it offers BGP updates every 5 minutes. The RouteViews and RIPE have been used in many research efforts such as [17], [72] and [73]. These two well-known repositories provide data in MRT (Multi-Threaded Routing Toolkit) format described in [74]. The MRT format is not a human readable. Software such as bgpdump [75] and pybgpdump [76] are used to convert it to a readable format.

The BGP speakers generate up to a gigabyte of control plane data a day [77]. Unfortunately, as well as being large, there is no direct information to identify the network that triggered the BGP messages [78]. Many tools have been introduced to extract significant information such as [79], speed up processing such as [80], and replay past BGP events such as [81]. A first step of building a scalable BGP tool that enables users querying BGP archived data with some simple analysis and statistics for a given period of time was presented in BGP-Inspect [82]. Although this tool provides much significant information such as shortest and longest path, it does not provide necessary information for operators and researchers. Other tools with a visualization capability were presented to help diagnosis BGP anomalies such as BGPlay [83] and VisTracer [84]. Biersack et al. presented a short survey of BGP visualization tools for monitoring BGP messages and particularly for the identification of prefix hijacks [68].

In contrast to the two offline BGP repositories, BGPmon represents the next step for the RouteViews Project in term of real-time data. The capability of the new monitoring system is not limited to providing real-time monitoring but also uses Extensible Markup Language (XML) format, which is supported by many applications and is human readable [71]. BGPmon, however, does not provide data processing. Where this processing is needed, tools such as Cyclops [85] can be used for this purpose.

b) Data Plane: Data plane is based on the way that packets actually flow between two nodes. This data can be obtained through active probing of available live hosts in the monitored networks. Different techniques of obtaining BGP data plane have been introduced. For example, Schlamp et al. used archived netflow data of Munich's Scientific network

[2]Rerouting can be triggered by different causes such as network faults, misconfiguration, and hijack [65].

[3]For example, as at the 18th of January 2016 there are 18 collectors for the RouteViews project with 588 peers in different locations around the world while RIPE peers with 14 collectors around the world with 566 peers.

[22]. Biersack et at. developed a tool called Spamtracer to monitor routes toward malicious hosts [68]. Others used different types of fingerprints such as host OS properties, IP identifier, TCP time stamp, and ICMP time stamp as an indicator to detect suspicious prefixes [49].

Control plane and data plane sources each have advantages and drawbacks. In general, techniques that use the control plane such as [86] and [85] are easy to deploy, but can be inaccurate while techniques that use the data plane such as [87], [84] and [51] have a better detection accuracy, but suffer from vantage point limitations [51], [5], [87]. Techniques that use a combination of control and data plane can be both accurate and have good vantage points, but their deployment is not easy [68], [22], [5].

2) Route Registries Database: To ensure stability and consistency of Internet-wide routing, Internet Routing Registry (IRR) was established to share information between network operators [88]. The IRR is a distributed routing database where ASes store their routing policies expressed in the Routing Policy Specification Language (RPSL) described in [89]. This data may be used by anyone worldwide to help debug routing problems, configure backbone routers, and engineer Internet routing and addressing. It also enables validation of linkage between a BGP speaker and the networks it announces, such as APNIC's whois database [90]. However, each RIR has its own network information database, part of which is used for routing information. IRR has been used by many researchers to test whether BGP updates originated from valid BGP speakers. For example, Nemecis [91] is a tool to extract and infer information from an IRR database and validate it against a BGP routing table. In addition to IRR databases, there are other sources that provide IP to AS number mapping such as team Cymru [92].

Unfortunately IRR information is incomplete because of lack of maintenance [58], [33]. Siganos and Faloutsos in [91] show that just 28% of IRR information is consistent. Furthermore, using this information is limited to detecting the two direct types of anomalies. In spite of these limitations, route registries database can be used with BGP raw data sources to produce quite robust anomaly detection [67], [5].

3) Other Sources of BGP Data: In addition to the first two types of BGP data sources, there are other sources such as bogon prefixes, the mailing list archive of North American Network Operator's Group (NANOG), and IP geolocation databases. Bogon prefixes are IP addresses that should not appear on the Internet. These are either within private address space [93], in a range of space reserved by IANA, or not allocated by any RIRs [94]. Bogon prefixes are not a static list where IP addresses are regularly added/removed from bogon lists. These lists are regularly updated and published by different sources. For example, CIDR report offers a daily list of bogon prefixes based on the IANA registry files, the RIR stats files, and the RIR whois data [95]. The team Cymru also offers a list of bogon prefixes [96] which is periodically updated. ISP operators do not need only to filter these prefixes and mitigate their propagation, they need a plan for keeping their filters up-to-date. Observation of bogon prefixes may be used as an indicator to detect BGP anomaly [6]. For example, BGP misconfiguration can produce large number of bogus routes (unused prefixes) as well as used prefixes [58]. Bogon prefixes may also be used as an identifier to differentiate hijacking toward intended prefix from misconfiguration [87].

The archives of NANOG mailing list have also been used as a source of BGP data in BGP anomaly detection techniques. The NANOG contains technical information, discussion, operational issues exchanged by network operators. Feamster and Balakrishnan in [35] use the archive of NANOG to identify challenges and common problems of configuring a BGP router to build rrc, a tool that detect BGP configuration faults based on static analysis. However, the archives of NANOG mailing list cannot be used for automated detection or real-time analysis.

IP geolocation databases map an IP address to its geographical location such as MaxMind's and IP2location [97], [98]. MaxMind and IRR databases were used in [99] map prefixes to a particular country. In addition to the described sources, looking glass has also been used [5]. Looking glass are computers on the Internet that provide information relative to backbone routing and network efficiency.

Adoption of different types of BGP data can produce a more reliable BGP anomaly detection approach. For example, various types of data such as RouteViews and RIPE NCC, public route servers, looking glass, and routing registry databases have been used to detect BGP anomalies in [5]. In [99], control plane, data plane, IRR, MaxMind, and active traceroute probing from Ark [100] were used to analyse episodes of BGP anomalies in Egypt and Libya caused by service providers implementing government censorship orders.

B. BGP Features

BGP messages are complex structures and detecting abnormal data in a series of BGP messages is a challenge [101]. Although individual BGP messages that constitute an anomaly provide no direct indication of why and where they originated, analysing a series of BGP messages can give such information. Some researchers have successfully used a single BGP feature to detect BGP anomalies such as [73], where BGP message volume was used as a single BGP feature. BGP message volume refers to the number of announcements and withdrawals sent from an AS or a prefix during a selected time interval. Other researchers have used more than a single BGP feature. For example, [59] and [17] used BGP volume, AS-PATH length, and observation of rare ASes in the AS-PATH to detect BGP anomalies. During instability periods caused by anomalies such as hardware or link failure, BGP path exploration attempts to find possible alternative paths for the unreachable destination. As a result, a large number of long and rare AS-PATHs appear. The topology in Figure 13 explains the effect of link failure on AS-PATH length and how different lengths of AS-PATH and rare ASes appear in alternative paths. In this topology, each node represents a different AS with a single BGP router. When the path between AS3 and AS4 fails, AS3 will send a withdrawal message to its neighbors. AS1 receives notification of alternative paths from its neighbors as a result of lose the connection between AS3 and AS4 such

as (2,3,5,7,4) and (2,3,6,8,4). Observing different AS-PATH lengths and rare AS numbers can help to detect such types of BGP anomalies.

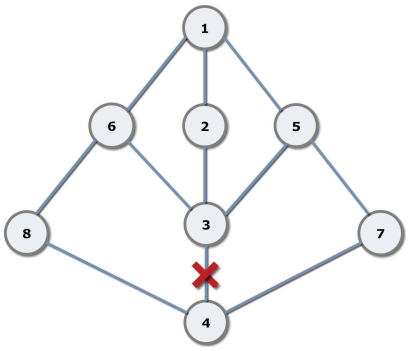

Figure 13. An example for the effect of a link failure

Table III
POSSIBLE BGP FEATURES FROM NUMBER OF BGP UPDATES

ID	Features
1	Total number of announcements/withdrawals/updates per AS
2	Number of announcements/withdrawals/updates per prefix
3	Maximum/Average announcements per prefix
4	Number of duplicate announcements/withdrawals
5	Number of new announcements
6	Number of IGP, EGP, and INCOMPLETE in the Origin attribute
7	Number of re-announcements after a withdrawal
8	Number of withdrawals transmitted to unreachable prefix
9	Number of withdrawals after announcing the same path
10	Number of unique prefixes originated by an AS
11	Concentration ratio (first, second, and third)

Table IV
POSSIBLE BGP FEATURES FROM AS-PATH ATTRIBUTE

ID	Feature
1	AS-PATH length
2	Maximum/Average AS-PATH length
3	Maximum/Average unique AS-PATH length
4	Announcement to longer/shorter path
5	Observation of rare ASes in the path
6	Maximum/Average of rare ASes in the path
7	AS-PATH change according to geographic location
8	Prefix origin change
9	Number of new paths announced after withdrawing an old path
10	Number of new-path announcements

Other researchers have used more specific features extracted from the two main features (deviation in number of BGP updates and AS-PATH) such as number of announced and withdrawn prefixes, average AS-PATH length, and maximum AS-PATH length, and then adopted an algorithm or a technique to find the most-important features which produce highest detection performance. Al-Rousan and Trajkovic extracted 37 features from BGP updates calculated on a 1 minute sliding window [72], then used Fisher score [102] and minimum Redundancy Maximum Relevance (mRMR) [103] to select the most-important features for detection. For this they identified 10 features that they adopted as input for detecting BGP anomaly. de Urbina Cazenave et al. presented new features related to BGP volume called concentration ratios. These features refer to the observation that the update volume is not equally distributed between all ASes and prefixes [104]. Table III and Table IV show possible features used by different BGP anomaly detection techniques related to number of BGP updates and AS-PATH respectively. We will discuss these in more detail later in the next section.

V. A REVIEW OF BGP ANOMALY DETECTION APPROACHES

In the BGP domain, many methods, systems, and approaches have been presented to detect BGP anomalies such as in [49] and [5] or locate the source cause of anomaly after detection such as in [101] and [105]. To simplify the comparison between BGP anomaly detection approaches and techniques, we build a taxonomy of algorithms, methods, and techniques into five main classes. These classes are: time series analysis, machine learning, statistical pattern recognition, validation of BGP updates based on historical BGP data, and reachability checks. Figure 14 shows our taxonomy of BGP anomaly detection in term of approaches, BGP data sources and features.

In this section, we explore approaches of BGP anomaly detection based on five aspects.

1) Technique used for detection.
2) Ability to identify different types of anomalies.
3) Used data sources.
4) Observed BGP features.
5) Ability to identify the source cause of anomalies.

We now discuss BGP anomaly detection approaches in more detail.

A. Time Series Analysis Approaches

One of the earliest efforts of identifying BGP anomalies was by Labovitz et al. [106]. The authors applied the Fast Fourier Transform (FFT) [107] to routing update rates. They used BGP data collected from five IXPs in the USA for a period of 9 months. The authors adopted five BGP features to detect BGP instability as listed in Table V. Although the technique did not provide a way to identify the cause or source of routing instability, it demonstrated that rapid changes in routing updates are correlated with instability.

Another time series analysis method to detect BGP anomalies made use of the Wavelet Transform [109]. Mai et al. introduced a new framework to detect BGP anomalies called BAlet [73], an extension to the work described in [110]. The BAlet uses Daubchies5 (db5) Wavelet transform to detect BGP anomaly and Single-Linkage [111] as a clustering algorithm to identify possible networks that originate anomaly. The BAlet is based on the BGP control plane where RouteViews and RIPE NCC is used to extract BGP message volume as a single BGP feature. To evaluate the BAlet, BGP traffic during the Slammer attack and 6 months of monitoring BGP log files at the AS12 were used to detect BGP anomalies. Although the BAlet is able to identify possible location from which the

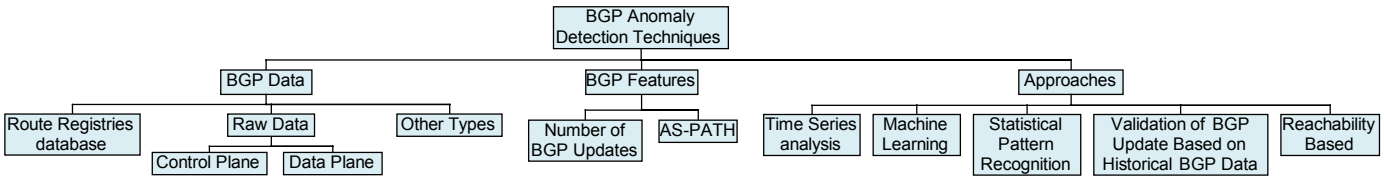

Figure 14. Taxonomy of BGP anomaly detection

Table V
SUMMARY OF APPROACHES BASED ON TIME SERIES ANALYSIS

N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Used Data Source	Observed BGP Features	Technique	Types of Anomaly	Identify Source Cause
Labovitz et al. [106]	Control plane	• Number of new paths announced after withdrawing an old path • Number of new-path announcements • Number of re-announcements after withdrawing the same path • Number of duplicate announcements • Number of withdrawals transmitted to unreachable prefix	FFT	Tested with 9 months of BGP data. No testing for specific types	N2
Mai et al. [73]	Control plane	• BGP message volume	db5 Wavelet transform and Single-Linkage	Tested with indirect anomaly and 6 months of BGP data	N1
Prakash et al. [108]	Control plane	• Number of announcements and withdrawals per a prefix	Haar Wavelet transform and Median filtering	Tested with 2 years of BGP data. No testing for specific types	N1
Al-Musawi et al. [4]	Control plane	• BGP volume and average length of AS-PATH	RQA	Tested with direct unintended anomaly	N3

anomaly originated, it is slow, typically requiring 20 minutes of data.

The Wavelet transform was also adopted in the BGP-lens [108], a tool to analyze BGP data and detect anomalies. The BGP-lens is based on using the Haar Wavelet transform and median filtering approach [112]. Its goal was to identify what characterized normal BGP data and how to detect BGP anomalies. This tool was evaluated with Abilene data (an academic research network) for a period of two years through investigation of BGP updates (announcements and withdrawals) per prefix. The BGP-lens offers three levels of alarm and a range of periods to check. However, this work did not address time delay of detection and appears not capable to detect anomalies in real-time (in seconds).

Al-Musawi et al. in [4] showed that BGP updates sent from BGP routers have the characteristics of determinism, recurrence, and non-linearity. They used these characteristics to present an approach to BGP anomaly detection based on Recurrence Quantification Analysis (RQA). RQA is an advanced non-linear analysis technique based on a phase plane trajectory [113]. The approach uses BGP volume and average length of AS-PATH as BGP features extracted every second. The authors evaluated their approach using one of the recent BGP incidents by Telekom Malaysia (TMnet) which caused significant network problems for the global routing system [114]. RQA is able to rapidly detect BGP anomalies caused by a high volume of BGP updates as well as hidden abnormal behaviour that may otherwise pass without observation. However, the work did not address the problem of identifying the location which caused the anomaly but is possibly capable of doing so.

Table V shows a summary of BGP anomaly detection techniques based on time series analysis. It is noted that all approaches based on time series analysis in [106], [73], [108] and [4] have not been tested to detect direct intended anomaly.

B. Machine Learning Based Approaches

Li et al. presented an Internet Routing Forensic (IRF) framework to detect BGP anomalies based on using a machine learning algorithm [77]. The IRF framework was based on the control plane where the RouteViews and RIPE NCC were used to extract 35 features every 1 minute. The framework applies the C4.5 algorithm [116] to build a decision tree. The authors evaluated their framework by using two different cases: worm attacks (indirect BGP anomaly), comprising the CodeRed and Nimda, and electricity failure, East Coast and Florida blackout [117]. The IRF is able to detect BGP events based on learned rules of past BGP events. For example, using the rules learned from CodeRed and Nimda worms attack to detect Slammer. However, the framework did not address the problem of identifying the location which caused the anomaly and appears not capable of doing so. The IRF, furthermore, was not evaluated to detect direct types of BGP anomalies. A similar approach to IRF framework was introduced by de Urbina Cazenave et al. [104]. This framework has the ability of using different data mining algorithms such as decision tree, Naive Bayes, and Support Vector Machine (SVM) [116]. This framework consists of two main parts: an advance features

Table VI
SUMMARY OF APPROACHES BASED ON MACHINE LEARNING

N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Used Data Source	Observed BGP Features	Technique	Types of Anomaly	Identify Source Cause
Li et al. [77]	Control plane	• Number of announcements and updates for an AS and prefix • Number of new-path announcements • Re-announcements after withdrawing the same path • Number of paths announced after withdrawing an old path • Number of withdrawals after announcing the same path Selected from 35 features extracted every 1 minute	C4.5	Tested with indirect anomaly and link failure	N4
de Urbina Cazenave et al. [104]	Control plane	• Number of announcements, withdrawals, and updates for an AS and prefix • Maximum and average announcements for a prefix • Maximum and average AS-PATH length • Maximum and average unique AS-PATH length • Announcements to longer and shorter path • Concentration ratio (first, second, and third) order Extracted every 30 and 60 seconds	Decision tree, Naive Bayes, and SVM (best results)	Tested with direct unintended anomaly, indirect anomaly, and link failure	N4
Al-Rousan and Trajkovic [72]	Control plane	• Average rare ASes in an AS-PATH • Maximum AS-PATH length • Number of duplicate withdrawals • Number of withdrawals, incomplete packets and duplicate announcements • Packet size Selected from 37 features extracted every 1 minute using Fisher and mRMR	SVM and HMMs	Tested with indirect anomaly	N4
Lutu et al. [115]	Control plane	• Average number and its standard deviation for LVP generated by same origin AS • Average proportion and its standard deviation of active monitors detecting the LVP • Average and standard deviation of absolute visibility degree for the LVP • Prefix length of the LVP Selected from 9 features extracted every two weeks using classification and regression trees	Winnowing algorithm	Tested with direct unintended anomaly	N4

extraction for 15 BGP features extracted every 30 seconds from BGP raw data downloaded from the RouteViews and RIPE NCC and data mining algorithms to classify BGP events. The authors show that SVM produces better performance than decision tree and Naive Bayes. Although the new proposal shows its ability to detect a wide range of BGP anomalies such as misconfiguration, blackout, and worm attacks, it was not tested to detect direct intended anomaly and did not address the problem of identifying the location that caused the anomalies and appears not capable of doing so.

Another example of using machine learning to detect BGP anomalies is [72]. The new mechanism consists of two main phases, an advance features extraction from BGP updates and a classifier to classify BGP updates as normal or abnormal. In the first phase, 37 features are extracted every 1 minute from BGP raw data downloaded from the two well-known BGP control plane repositories, then Fisher [102] and minimum Redundancy Maximum Relevance (mRMR) [103] scoring algorithms are used to select the 10 most-important features that have highest performance for detection. However, the authors show that volume features are more important features to detect anomaly behavior than AS-PATH features. The second phase uses the SVM and Hidden Markov Models (HMMs) [118] to detect BGP anomalies. The suggested mechanism was evaluated using three indirect anomaly events: Slammer, Nimda, and Code Red I, other types of BGP anomalies such as direct and indirect anomalies have been tested. This mechanism, however, did not address how to identify the location which caused the anomalies and appears not capable of doing so.

Lutu et al. in [115] presented a system to detect BGP anomalies at an early stage based on prefix visibility at the inter-domain level. Prefix visibility is the occurrence of a prefix in the global routing table at every sampling moment. They classified prefix visibility into Limited-Visibility Prefix (LVP) and High-visibility prefix (HVP), where LVP is a stable long-lived internet route with a prefix visibility less than 95% of all routing table extracted from the RouteViews and RIPE NCC. HVP refers to prefix visibility with more than 95% of all the extracted routing table. The authors used the BGP Visibility Scanner [55], a tool for identifying limited visibility for stable prefixes in the Internet, to detect LVP and HVP. They also used a machine learning winnowing algorithm to classify whether LVP resulted as a misconfiguration by a BGP router operator or was a natural expression of global routing policies in the Internet. "Winnowing" is based on boosted classification trees described in [119]. Among 9 features based on visibility of prefixes, the 7 most important features were selected using

classification and regression trees [116]. Although the proposal is able to detect anomalies which are still undetected by many other tools, it is limited to detect direct unintended anomaly. In addition, it does not show its ability to identify the location which caused the anomaly and appears not capable of doing so.

Table VI shows a summary of BGP anomaly detection techniques based on machine learning. It is noted that none of these approaches [77], [104], [72], [115] address detecting BGP hijacking (direct intended anomaly) or are able to identify the source cause of anomaly.

C. Statistical Pattern Recognition Based Approaches

Huang et al. [120] introduced a technique to detect BGP node, link, and peer failure. This technique uses a Principal Component Analysis (PCA) based subspace method [116] to detect and differentiate between the three failures. They used three types of data source: BGP updates, operational mailing list, and routing configuration for the Abilene network, where BGP update volume was used as a single BGP feature extracted every 10 minutes with window size of 200 minutes. The authors used the NANOG operational mailing list to validate the ground truth of detection. Although the approach is able to detect, identify, and differentiate between BGP node, link, and peer failure, it requires information of router configuration and is unsuitable for real-time detection since it takes from 9--96 minutes.

Deshpande et al. in [17] presented a BGP anomaly detection approach based on the Generalized Likelihood Ratio Test (GLRT), a standard statistical technique used in hypothesis testing, deployed on a single BGP router. Three BGP features (BGP volume, AS-PATH length, and rare ASes) were extracted every 5 minutes from BGP updates downloaded from the RIPE NCC. The most relevant features among the three features were selected using Fisher score [123]. The authors showed that using AS-PATH and rare AS in AS-PATH features with message volume improved the false positive rate compared with using message volume alone. The authors compared their proposal with the adaptive Exponentially Weighted Moving Average (EWMA) scheme in [124], PCA based scheme used in [120], and Wavelet based scheme used in [110]. The new approach was evaluated with ten well-known events (such as worm attacks, misconfiguration, equipment failure, and hijacking) and produced a high level of detection accuracy and low computation cost compared with the other mechanisms. However, this detection system is slow, typically needing around an hour to detect anomalies.

Ganiz et al. [121] presented a Higher-order path analysis (HOPA) to detect BGP anomalies and able to differentiate between indirect BGP anomaly and link failure. HOPA is a data-mining approach used to produce relevant information from BGP updates, downloaded from the RouteViews project, every 6 minutes, then classify them using Student's t-test [116], a statistical hypothesis analysis. Although their approach is able to differentiate between two types of BGP anomalies in 360 seconds, it was not evaluated with the most common types of BGP anomalies (the two direct anomalies). In addition, HOPA did not address how to identify the location which caused the anomalies but it possibly capable of doing so.

Theodoridis et al. [122] introduced an unsupervised mechanism to detect BGP hijacking using control plane BGP raw data. This mechanism is based on observing the geographic changes of intermediate AS in the AS-PATH between the competing routes. Frequency of AS appearance in the path and geographic deviation of intermediate AS were introduced as a two BGP features. These two features have also been used by BGPfuse [125], a visualization tool to detect BGP hijacking. The unsupervised proposal uses Z-score calculation, a statistical measure for a particular value of the number of standard deviations [116], to rank how much the intermediate AS are suspicious. The authors assumed that the attackers usually carried out their activity on remote locations to avoid any legal actions and reduce the possibility of identification. This mechanism was evaluated using the Link Telecom incident and showed it was able to detect the deviation of intermediate AS during the hijack. This mechanism did not address how to identify the location which caused the anomaly nor real time detection but it possibly capable of doing so. However, the authors did not evaluate their mechanism to detect other types of BGP anomalies such as direct unintended and indirect anomalies.

Table VII shows a summary of techniques based on statistical pattern recognition to detect BGP anomalies. It is noted that approaches based on statistical pattern recognition show their ability to detect different types of BGP anomaly and identify the source cause.

D. Validation of BGP Updates Based on Historical BGP Data

This approach to BGP anomaly detection uses a history of RIB table and/or BGP updates to validate new BGP updates, assuming that Internet topology does not frequently change. Pretty Good BGP (PGBGP) is a detection and mitigation system against BGP attacks [12]. PGBGP uses history of both RIB and BGP updates downloaded from the RouteViews project to validate new updates. It uses a prefix and origin AS pair (prefix origin change feature) from both RIB and update history over the previous 10 days. It also eliminates routes that are no longer active and older than 10 days. When a new route is received and its pair (prefix and origin AS) is not recorded in the history period, it considers the update suspicious and will be propagated with a low LOCAL-PREF. PGBGP did not address the problem of identifying the location which caused the anomaly but it possibly capable of doing so. Although PGBGP is able to detect prefix and sub-prefix hijacking, it has some limitation related to using the history period (10 days). For example, MOAS may last for a few days and may not be observed for months. An enhanced history-based algorithm was introduced by Sriram et al. [67] that successfully overcame PGBGP drawbacks through using a longer period (e.g, months).

Lad et al. [86] presented PHAS, a prefix hijacking detecting system with the capability of sending alarms to the real prefix owners. PHAS analyzes BGP data in real-time to detect when a prefix hijacking event occurred or was resolved. PHAS

Table VII
SUMMARY OF APPROACHES BASED ON STATISTICAL PATTERN RECOGNITION

N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Used Data Source	Observed BGP Features	Technique	Types of Anomaly	Identify Source Cause
Huang et al. [120]	Control plane, operational mailing list, and routing configuration	• BGP volume extracted every 10 minutes with a window size of 200 minutes	PCA-based on subspace method	Tested with link failure	N1
Deshpande et al. [17]	Control plane	• BGP volume • AS-PATH length • Rare AS in a path. Extracted every 5 minutes	EWMA, PCA, and GLRT (better performance)	Tested with all types of anomalies	N1
Ganiz et al. [121]	Control plane	• Number of announcements, withdrawals, and updates for an AS and prefix	HOPA	Tested with indirect anomaly and link failure	N3
Theodoridis et al. [122]	Control plane	• Frequency of AS appearance in the path • Geographic deviation of intermediate AS	Z-score calculation	Tested with direct intended anomaly	N3

Table VIII
SUMMARY OF APPROACHES BASED ON USING HISTORICAL BGP DATA

N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Used Data Source	Observed BGP Features	Technique	Types of Anomaly	Identify Source Cause
Karlin et al. [12]	control plane	Prefix origin change	Validate new BGP update based on 10 days of BGP history	Tested with direct intended anomaly	N3
Lad et al. [86]	control plane	Prefix origin change	Validate new BGP updates based on prefix registration by owners	Tested with direct intended anomaly	N3
Haeberlen et al. [13]	control plane	Prefix origin change	Validate new BGP updates based on 1 years of BGP and set of rules	Tested with a control testbed	N1
Shi et al. [5]	control plane, data plane, and IRR	Prefix origin change	Classify new BGP updates based on 2 months of BGP history	Tested with direct intended and unintended anomalies	N1

Table IX
SUMMARY OF APPROACHES BASED ON REACHABILITY CHECK

N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Used Data Source	Observed BGP Features	Technique	Types of Anomaly	Identify Source Cause
Zheng et al. [87]	Data plane	Number of hops toward a suspicious prefix	Combination of ping, traceroute, and iplane	Tested with direct intended anomaly	N1
Hu and Mao [49]	Control plane and data plane	Set of fingerprints such as OS properties, IP identifier, TCP time stamp, and ICMP time stamp	Combination of hping and Nmap	Tested with direct intended anomaly	N3
Tahara et al. [126]	Data plane	Reachability from different vantage points	Ping test	Tested with direct intended anomaly	N3
Zhang et al. [51]	Data plane	Reachability of a prefix from transit ASes	Combination of iTraceroute, ping, and TCP ping	Tested with direct intended anomaly and link failure	N2

requires a registration process from prefix owners who want to use it. The registration is used as a base point to monitor origin changes of that prefix. The system uses an adaptive window scheme with 1 hour as initial size window. The adaptive scheme increases the window size when there are many changes in the origin of a prefix and decreases it in case of a small number of changes. The authors believe that only the prefix owners can distinguish between legitimate changes and hijacking changes for their prefixes; therefore, this approach offers a filter facility to prefix owners to add which AS can use a specified prefix. PHAS uses the RouteViews repository as a BGP control plane and prefix origin changes as a single BGP feature. Although the system does not require any router reconfiguration, it requires registration and has consequent issues related to authenticating legitimate ownership as there is no secure mechanism to differentiate between a legitimate owner and an attacker.

Haeberlen et al. [13] presented a prototype to detect BGP faults at the AS level called NetReview. This prototype uses a tamper-evident log, containing BGP messages to and from AS neighbors, to detect BGP faults, where BGP faults include BGP router and link failure, misconfiguration, policy violations, and attacks. NetReview requires each BGP speaker to maintain a history log file of around one year of data and a set of rules that describe its best practices and routing polices, then other ASes can use these information to audit how the rules are

followed. NetReview requires each AS to log a public pledge to show its ownership of AS and prefixes. NetReview has been tested under a control testbed to detect different types of BGP anomalies. Although NetReview can detect in real-time different types of BGP anomalies and identify their source cause, it requires each AS to reveal information related to its policy configuration and suffers from scalability problems related to the size of storing log files especially for large ISPs.

Argus [127], [5] is a system to detect prefix hijacking and identify the attacker in real-time. Argus uses the control plane to detect bogus routes and the data plane to verify anomalies through checking their reachability. This system uses more than 2 months of historical BGP data to classify new BGP updates as normal or suspicious, then checks the reachability of prefixes to verify the suspicious updates through using tools such as iplane [128] and CAIDA's Ark [100]. In addition, the IRR data is also used to improve the false positive rate. Although Argus can detect BGP hijacking and identify the source cause in real-time, it cannot detect sub-prefix hijacking nor other types of anomalies such as indirect anomaly and link failure.

Table VIII shows a summary of work in detecting BGP anomalies based on historical RIB and/or BGP updates. It is noted that all approaches based on historical BGP data use prefix origin change as a single BGP feature.

E. Reachability Check

This type of technique uses the BGP data plane to check reachability to a certain prefix using different types of tools such as hping [129], Nmap [130], traceroute [131], iTraceroute [132], and Paris traceroute [133]. One of the earliest works to detect BGP hijacking based on data plane was by Zheng et al. [87]. The authors assumed that the network location for a prefix remains unchanged over time so a significant change in network distance (hop count between source and destination) from a selected vantage point to a certain IP may be used as an indicator of hijacking. They used a combination of ping, traceroute, and iplane [128] to count the number of hops towards a certain prefix. This proposal can detect prefix hijacking in real-time, but it is not able to detect sub-prefix hijacking. Furthermore, it was not tested to detect other types of anomalies.

In [49], the control plane is used to detect suspicious BGP messages, then data plane probing is launched to verify if the suspicious data is an anomaly or not. In general, attackers use dissimilar OS or configure OS to open some ports when compared with legitimate users; thus, tools such as Nmap [130] can identify the OS fingerprint of the attacker. The authors used a set of fingerprints such as host OS properties, IP identifier, TCP time stamp, and ICMP time stamp to identify the attackers. For example, the IP identifier is designed to be unique for each IP datagram to help IP fragment reassembly. The identifier is incremented for every outgoing packet regardless of its destination. The authors used this identifier to send probe packets simultaneously to the same suspicious IP from two locations to check if they arrived at the same destination. The authors did not address the problem of identifying the location which caused the anomaly but it possibly capable of doing so. However, this system is difficult to deploy as it relies on complicated probing and requires installation of customized software at its vantage points [5].

Tahara et al. [126] proposed a method to detect prefix hijacking based on data plane by using a ping test. When an attacker hijacks a prefix, packets traverse toward the attacker instead of the real prefix owner. However, during a hijacking attack, not all AS experience the effect of the attackers (as shown in Figure 8); therefore, checking reachability to a suspicious prefix from different vantage points can be used by an observer to detect a prefix hijack. The authors, however, did not address time delay of detection nor identification of the location which caused the anomalies but the method is possibly capable of doing so.

Zhang et al. presented iSPY [51], a tool for detecting prefix hijack in real-time based on the observation that connectivity to victim hosts is lost during hijacking attempts. iSPY is able to distinguish between a real attack and a link failure or network congestion. It uses the BGP data plane where a combination of iTraceroute, ping, and TCP ping, a ping over TCP port such as that available in Nmap [130], are used to check the reachability to a certain prefix. Deploying iSPY on a network requires collecting in advance a set of live IPs using active probing, an IP-to-AS mapping, and a continuous update to its database. However, iSPY's ability is limited to detecting regular prefix hijacking only. Other types of hijacking such as sub-prefix hijacking cannot be detected.

Table IX shows a summary of work for detecting BGP anomalies based on reachability check. This table shows that all approached based on reachability check use data plane to detect or verify the exist of anomaly. However, none of these approach tested to detect indirect BGP anomaly.

VI. Key Requirements for Next Generation of BGP Anomaly Detection

We have presented many systems, approaches, and tools with different capabilities for detecting BGP anomalies. A summary comparison between BGP anomalies detection techniques in term of real-time detection, ability to detect different types of BGP anomalies, differentiate between types of BGP anomalies, and identify the location which caused the anomaly is shown in Table X. This table shows that none of these works offers a combination of adequate real-time detection and identification for all types of anomalies as well as locating the cause of anomalies. A recent analysis [5] shows that some hijackings proceed in less than ten minutes and can affect 90% of the Internet within less than two minutes. Thus, detection of BGP anomalies in real-time (in seconds) is required to mitigate their propagation between BGP speakers. However, detecting BGP anomaly in real-time is not enough to stop propagation of anomalies without requiring an action from the operators. To accomplish that, the operators need extra information to ensure that an alarm is raised by serious threats. This can be done through identifying the type of anomaly as well as the location which caused the anomaly. For example, the action taken by an operator when dealing with a direct intended anomaly is

Table X
REVIEWED WORKS IN LIGHT OF FOUR CHARACTERISTICS PROPERTIES

C1=Has not been tested with a specific type of anomaly , C2= Direct intended anomaly, C3=Direct unintended anomaly, C4= Indirect anomaly, C5=Link failure, N1= Addressed and capable, N2= Addressed but not capable, N3= Not addressed but possibly capable, N4= Not addressed and appears not capable

Work	Real-time Detection	BGP Detected Anomalies	Differentiate Between Anomalies	Identify Source Cause
Labovitz et al. [106]	N4	C1	N4	N2
Mai et al. [73]	N2	C4	N4	N1
Prakash et al. [108]	N4	C1	N4	N1
Al-Musawi et al. [4]	N1	C3	N4	N3
Li et al. [77]	N2	C4 and C5	N4	N4
de Urbina Cazenave et al. [104]	N1	C3, C4, and C5	N4	N4
Al-Rousan and Trajkovic [72]	N4	C4	N4	N4
Lutu et al. [115]	N4	C3	N4	N4
Huang et al. [120]	N2	C5	N4	N1
Deshpande et al. in [17]	N2	C2, C3, C4, and C5	N4	N1
Ganiz et al. [121]	N1	C4 and C5	between C4 and C5	N3
Theodoridis et al. [122]	N3	C2	N4	N3
Karlin et al. [12]	N2	C2	N4	N3
Lad et al. [86]	N2	C2	N4	N3
Haeberlen et al. [13]	N1	C1	N4	N1
Shi et al. [5]	N1	C2 and C3	N4	N1
Zheng et al. [87]	N1	C2	N4	N1
Hu and Mao [49]	N1	C2	N4	N3
Tahara et al. [126]	N3	C2	N4	N3
Zhang et al. [51]	N1	C2 and C5	between C2 and C5	N2

different from that taken when dealing with direct unintended anomaly, where the unintended anomaly may stop as soon as the operator discovers the fault or is informed of it by its neighbors. However, distinguishing direct and intended from unintended anomaly has not been resolved [23]. The ability to locate the attackers is critical for mitigating anomaly effects through introducing an early recognition mechanism to stop the propagation of attacks.

VII. Conclusions

BGP is the Internet's default inter-domain routing protocol. It was developed at a time when information provided by an AS could be assumed to be accurate. BGP has been threatened by different types of anomalies that affect its stability and performance. During the past twenty years many different types of anomalies have affected BGP stability and performance. These can be mainly classified into four main categories: direct intended anomaly, direct unintended anomaly, indirect anomaly, and link failure.

This paper surveys 20 significant works in the field of BGP anomaly detection during the period of 1998 to late 2015. It examines these works in terms of BGP data sources and features, detection technique, ability to detect different type of BGP anomalies and locate the source cause of anomalies. Time series analysis, machine learning, statistical pattern recognition, validation of BGP updates based on history log, and reachability check are the main techniques that have been used to detect BGP anomalies. This survey paper also classifies BGP data sources into three main categories: raw data (control plane and data plane) and route registry database as well as other types of BGP data sources.

There is still much to be done in the field. None of these significant works offers a combination of detecting in real-time for all types of anomalies, differentiating between them, and identifying the source cause of the anomaly. This combination is needed to enable operators to mitigate the propagation of anomalies, protect their network, and help to understand the inter-domain routing protocol.

References

[1] A. Barbir, S. Murphy, and Y. Yang, "Generic Threats to Routing Protocols," RFC 4593 (Informational), Internet Engineering Task Force, October 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4593.txt

[2] Y. Rekhter, T. Li, and S. Hares, "RFC 4271: A Border Gateway Protocol 4 (BGP-4)," RFC 4271 (Proposed Standard), Internet Engineering Task Force, January 2006. [Online]. Available: http://tools.ietf.org/html/rfc4271

[3] S. Secci, K. Liu, and B. Jabbari, "Efficient inter-domain traffic engineering with transit-edge hierarchical routing," *Computer Networks*, vol. 57, no. 4, pp. 976–989, 2013.

[4] B. Al-Musawi, P. Branch, and G. Armitage, "Detecting BGP instability using Recurrence Quantification Analysis (RQA)," in *2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC)*, Dec 2015, pp. 1–8.

[5] X. Shi, Y. Xiang, Z. Wang, X. Yin, and J. Wu, "Detecting Prefix Hijackings in the Internet with Argus," in *Proceedings of the 2012 ACM Conference on Internet Measurement Conference*, ser. IMC '12. New York, NY, USA: ACM, 2012, pp. 15–28.

[6] J. Qiu, L. Gao, S. Ranjan, and A. Nucci, "Detecting bogus BGP route information: Going beyond prefix hijacking," in *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*, Sept 2007, pp. 381–390.

[7] K. Butler, T. Farley, P. McDaniel, and J. Rexford, "A Survey of BGP Security Issues and Solutions," *Proceedings of the IEEE*, vol. 98, no. 1, pp. 100–122, Jan 2010.

[8] G. Huston, M. Rossi, and G. Armitage, "Securing BGP - A Literature Survey," *Communications Surveys Tutorials, IEEE*, vol. 13, no. 2, pp. 199–222, Second 2011.

[9] S. Kent, C. Lynn, and K. Seo, "Secure Border Gateway Protocol (S-BGP)," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 4, pp. 582–592, April 2000.

[10] G. Huston and G. Michaelson, "RFC 6483: Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs)," RFC 6483 (Informational), Internet Engineering Task Force, February 2012. [Online]. Available: http://tools.ietf.org/html/rfc6483

[11] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. Katz, "Listen and whisper: Security mechanisms for BGP," in *Proc. First Symposium on Networked Systems Design and Implementation (NSDI)*. Citeseer, 2004, p. 11.

[12] J. Karlin, S. Forrest, and J. Rexford, "Pretty Good BGP: Improving BGP by Cautiously Adopting Routes," in *Network Protocols, 2006. ICNP '06. Proceedings of the 2006 14th IEEE International Conference on*, Nov 2006, pp. 290–299.

[13] A. Haeberlen, I. Avramopoulos, J. Rexford, and P. Druschel, "NetReview: Detecting when Interdomain Routing Goes Wrong," in *Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation*, ser. NSDI'09. Berkeley, CA, USA: USENIX Association, 2009, pp. 437–452.

[14] S. Goldberg, M. Schapira, P. Hummon, and J. Rexford, "How secure are secure interdomain routing protocols," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 87–98, Aug. 2010.

[15] J. Chandrashekar, Z. Duan, Z.-L. Zhang, and J. Krasky, "Limiting path exploration in BGP," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 4, March 2005, pp. 2337–2348 vol. 4.

[16] G. Huston, M. Rossi, and G. Armitage, "A Technique for Reducing BGP Update Announcements through Path Exploration Damping," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 8, pp. 1271–1286, October 2010.

[17] S. Deshpande, M. Thottan, T. K. Ho, and B. Sikdar, "An Online Mechanism for BGP Instability Detection and Analysis," *Computers, IEEE Transactions on*, vol. 58, no. 11, pp. 1470–1484, Nov 2009.

[18] M. Lad, X. Zhao, B. Zhang, D. Massey, and L. Zhang, "Analysis of BGP Update Surge during Slammer Worm Attack," in *Distributed Computing - IWDC 2003*, ser. Lecture Notes in Computer Science, S. Das and S. Das, Eds. Springer Berlin Heidelberg, 2003, vol. 2918, pp. 66–79.

[19] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Observation and Analysis of BGP Behavior Under Stress," in *Proceedings of the 2Nd ACM SIGCOMM Workshop on Internet Measurment*, ser. IMW '02. New York, NY, USA: ACM, 2002, pp. 183–195.

[20] M. A. Brown, "Pakistan Hijacks YouTube," Renesys Blog, February 2008. [Online]. Available: http://www.renesys.com/2008/02/pakistan-hijacks-youtube-1/

[21] T. L. Simon, "oof.Panix Sidelined by Incompetence. . . Again," North American Network Operators Group, January 2006. [Online]. Available: https://www.nanog.org/mailinglist/mailarchives/old_archive/2006-01/msg00483.html

[22] J. Schlamp, G. Carle, and E. W. Biersack, "A Forensic Case Study on As Hijacking: The Attacker's Perspective," *SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 2, pp. 5–12, Apr. 2013.

[23] M. Wählisch, O. Maennel, and T. C. Schmidt, "Towards Detecting BGP Route Hijacking Using the RPKI," *SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 103–104, Aug. 2012.

[24] Y. Zhang and M. Tatipamula, "A Comprehensive Long-Term Evaluation on BGP Performance," in *Communications (ICC), 2011 IEEE International Conference on*, June 2011, pp. 1–6.

[25] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009.

[26] J. Mitchell, "Autonomous System (AS) Reservation for Private Use," RFC 6996 (Best Current Practice), Internet Engineering Task Force, July 2013. [Online]. Available: http://tools.ietf.org/html/rfc6996

[27] Q. Vohra and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space," RFC 6793 (Proposed Standard), Internet Engineering Task Force, December 2012. [Online]. Available: http://www.ietf.org/rfc/rfc6793.txt

[28] Internet Assinged Number Authority (IANA), "Autonomous System (AS) Numbers," July 2014. [Online]. Available: http://www.iana.org/assignments/as-numbers/as-numbers.xhtml

[29] V. Fuller and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," RFC 4632 (Best Current Practice), Internet Engineering Task Force, August 2006. [Online]. Available: http://tools.ietf.org/html/rfc4632

[30] E. Chen and J. Yuan, "Autonomous-System-Wide Unique BGP Identifier for BGP-4," RFC 6286 (Proposed Standard), Internet Engineering Task Force, June 2011. [Online]. Available: http://tools.ietf.org/html/rfc6286

[31] J. Hawkinson and T. Bates, " Guidelines for creation, selection, and registration of an Autonomous System (AS)," RFC 1930 (Best Current Practice), Internet Engineering Task Force, March 1996. [Online]. Available: http://tools.ietf.org/html/rfc1930

[32] L. Gao, "On Inferring Autonomous System Relationships in the Internet," *IEEE/ACM Trans. Netw.*, vol. 9, no. 6, pp. 733–745, Dec. 2001.

[33] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "The (In)Completeness of the Observed Internet AS-level Structure," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 1, pp. 109–122, Feb 2010.

[34] M. Caesar and J. Rexford, "BGP routing policies in ISP networks," *Network, IEEE*, vol. 19, no. 6, pp. 5–11, Nov 2005.

[35] N. Feamster and H. Balakrishnan, "Detecting BGP Configuration Faults with Static Analysis," in *Proceedings of the 2Nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2*, ser. NSDI'05. Berkeley, CA, USA: USENIX Association, 2005, pp. 43–56.

[36] T. Underwood, "Internet-Wide Catastrophe-Last Year," Renesys Blog, December 2005. [Online]. Available: http://www.renesys.com/2005/12/internetwide-nearcatastrophela/

[37] C. Pelsser, O. Maennel, P. Mohapatra, R. Bush, and K. Patel, "Route flap damping made usable," in *Passive and Active Measurement*. Springer, 2011, pp. 143–152.

[38] C. Villamizar, R. Chandra, and R. Govindan, "BGP Route Flap Damping," RFC 2439 (Standards Track), Internet Engineering Task Force, November 1998. [Online]. Available: http://www.ietf.org/rfc/rfc2439.txt

[39] T. Barber, S. Doran, D. Karrenberg, C. Panigl, and J. Schmitz, "RIPE Routing-WG Recommendation for coordinated route-flap damping parameters," ripe-178, Februray 1998, obsoleted. [Online]. Available: http://www.ripe.net/ripe/docs/ripe-178

[40] P. Smith and C. Panigl, "RIPE Routing Working Group Recommendations On Route-flap Damping," ripe-378, May 2006, obsoleted. [Online]. Available: http://www.ripe.net/ripe/docs/ripe-378

[41] C. Pelsser, R. Bush, K. Patel, P. Mohapatra, and O. Maennel, "Making Route Flap Damping Usable," RFC 7196 (Proposed Standard), Internet Engineering Task Force, May 2014. [Online]. Available: http://www.ietf.org/rfc/rfc7196.txt

[42] R. Bush, C. Pelsser, M. Kuhne, O. Maennel, P. Mohapatra, K. Patel, R. Evans and Janet, "RIPE Routing Working Group Recommendations On Route-flap Damping," ripe-580, January 2013, obsoletes: ripe-378. [Online]. Available: http://www.ripe.net/ripe/docs/ripe-580

[43] R. White, D. McPherson, and S. Sangli, *Practical BGP*. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc., 2004.

[44] P. Traina, D. McPherson, and J. Scudder, "Autonomous System Confederations for BGP," RFC 5065 (Standards Track), Internet Engineering Task Force, August 2007. [Online]. Available: http://tools.ietf.org/html/rfc5065

[45] M. Wubbeling, M. Meier, and T. Elsner, "Inter-AS routing anomalies: Improved detection and classification," in *Cyber Conflict (CyCon 2014), 2014 6th International Conference On*. IEEE, 2014, pp. 223–238.

[46] H. Ballani, P. Francis, and X. Zhang, "A Study of Prefix Hijacking and Interception in the Internet," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 4, pp. 265–276, Aug. 2007.

[47] T. Wan and P. Van Oorschot, "Analysis of BGP prefix origins during Google's May 2005 outage," in *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, April 2006, pp. 8–pp.

[48] Micron21 Datacentre, "Micron21 DDoS Soak and Scrub as a Service." [Online]. Available: http://www.micron21.com/ddos-soak-scrub.php

[49] X. Hu and Z. Mao, "Accurate Real-time Identification of IP Prefix Hijacking," in *Security and Privacy, 2007. SP '07. IEEE Symposium on*, May 2007, pp. 3–17.

[50] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "An Analysis of BGP Multiple Origin AS (MOAS) Conflicts," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, ser. IMW '01. New York, NY, USA: ACM, 2001, pp. 31–35.

[51] Z. Zhang, Y. Zhang, Y. Hu, Z. Mao, and R. Bush, "iSPY: Detecting IP Prefix Hijacking on My Own," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 6, pp. 1815–1828, Dec 2010.

[52] IETF, "Charter of the IETF Secure Inter-Domain Routing Working Group." [Online]. Available: http://tools.ietf.org/wg/sidr/charters

[53] G. Huston, "Leaking Routes," March 2012. [Online]. Available: http://www.potaroo.net/ispcol/2012-03/leaks.html

[54] I. Shrubbery Networks, "RANCID - Really Awesome New Cisco confIg Differ," 2004. [Online]. Available: http://www.shrubbery.net/rancid/

[55] A. Lutu, M. Bagnulo, and O. Maennel, "The BGP Visibility Scanner," in *Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on*, April 2013, pp. 115–120.

[56] E. Zmijewski, "Indonesia Hijacks the World," Renesys Blog, April 2014. [Online]. Available: http://www.renesys.com/2014/04/indonesia-hijacks-world/

[57] A. Toonk, "Turkey Hijacking IP addresses for popular Global DNS providers," March 2014. [Online]. Available: http://www.bgpmon.net/turkey-hijacking-ip-addresses-for-popular-global-dns-providers/

[58] R. Mahajan, D. Wetherall, and T. Anderson, "Understanding BGP Misconfiguration," *SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 4, pp. 3–16, Aug. 2002.

[59] S. Deshpande, M. Thottan, and B. Sikdar, "An online scheme for the isolation of BGP misconfiguration errors," *Network and Service Management, IEEE Transactions on*, vol. 5, no. 2, pp. 78–90, June 2008.

[60] S. Bortzmeyer, "Who has AS 1712?" North American Network Operators Group, November 2009. [Online]. Available: http://seclists.org/nanog/2009/Nov/647

[61] D. Madory, "Bonjour, Y'all! ASN Split Personalities," Dyn Research, December 2009. [Online]. Available: http://research.dyn.com/2009/12/bonjour-yall-asn-split-persona/

[62] S. Deshpande, M. Thottan, and B. Sikdar, "Early detection of BGP instabilities resulting from Internet worm attacks," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 4, Nov 2004, pp. 2266–2270 Vol.4.

[63] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," in *Proceedings of the 11th USENIX Security Symposium*. Berkeley, CA, USA: USENIX Association, 2002, pp. 149–167.

[64] A. Roudnev, "Re: More on Moscow power failure(was RE: Moscow: global power outage)," North American Network Operators Group, May 2005. [Online]. Available: https://www.nanog.org/mailinglist/mailarchives/old_archive/2005-05/msg00767.html

[65] Y. Liu, X. Luo, R. K. Chang, and J. Su, "Characterizing inter-domain rerouting by betweenness centrality after disruptive events," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 6, pp. 1147–1157, 2013.

[66] T. Bilski, "Disaster's impact on internet performance–case study," in *Computer Networks*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2009, vol. 39, pp. 210–217.

[67] K. Sriram, O. Borchert, O. Kim, P. Gleichmann, and D. Montgomery, "A Comparative Analysis of BGP Anomaly Detection and Robustness Algorithms," in *Conference For Homeland Security, 2009. CATCH '09. Cybersecurity Applications Technology*, March 2009, pp. 25–38.

[68] E. Biersack, Q. Jacquemart, F. Fischer, J. Fuchs, O. Thonnard, G. Theodoridis, D. Tzovaras, and P. Vervier, "Visual analytics for BGP monitoring and prefix hijacking identification," *Network, IEEE*, vol. 26, no. 6, pp. 33–39, November 2012.

[69] University of Oregon, "University of Oregon Route Views Project." [Online]. Available: http://www.routeviews.org/

[70] Reseaux IP Europeens Network Coordination Center. [Online]. Available: http://www.ripe.net/

[71] H. Yan, R. Oliveira, K. Burnett, D. Matthews, L. Zhang, and D. Massey, "BGPmon: A Real-Time, Scalable, Extensible Monitoring System," in *Conference For Homeland Security, 2009. CATCH '09. Cybersecurity Applications Technology*, March 2009, pp. 212–223.

[72] N. M. Al-Rousan and L. Trajkovic, "Machine learning models for classification of BGP anomalies," in *High Performance Switching and Routing (HPSR), 2012 IEEE 13th International Conference on*. IEEE, 2012, pp. 103–108.

[73] J. Mai, L. Yuan, and C.-N. Chuah, "Detecting BGP anomalies with wavelet," in *Network Operations and Management Symposium, 2008. NOMS 2008. IEEE*, April 2008, pp. 465–472.

[74] L. Blunk, M. Karir, and C. Labovitz, "RFC 6396: Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format," RFC 6396 (Standards Track), Internet Engineering Task Force, October 2011. [Online]. Available: http://tools.ietf.org/html/rfc6396

[75] RIPE NCC RIS Projec, "bgpdump." [Online]. Available: https://bitbucket.org/ripencc/bgpdump/wiki/Home

[76] J. Oberheide, "pybgpdump." [Online]. Available: https://jon.oberheide.org/pybgpdump/

[77] J. Li, D. Dou, Z. Wu, S. Kim, and V. Agarwal, "An Internet Routing Forensics Framework for Discovering Rules of Abnormal BGP Events," *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 5, pp. 55–66, Oct. 2005.

[78] A. Sapegin and S. Uhlig, "On the extent of correlation in BGP updates in the Internet and what it tells us about locality of BGP routing events," *Computer Communications*, vol. 36, no. 15, pp. 1592–1605, 2013.

[79] J. Wu, Z. M. Mao, J. Rexford, and J. Wang, "Finding a Needle in a Haystack: Pinpointing Significant BGP Routing Changes in an IP Network," in *Proceedings of the 2Nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2*, ser. NSDI'05. Berkeley, CA, USA: USENIX Association, 2005, pp. 1–14.

[80] M. Rossi, "MRT dump file manipulation toolkit (MDFMT) - version 0.2," Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia, Tech. Rep. 090730B, 30 July 2009. [Online]. Available: http://caia.swin.edu.au/reports/090730B/CAIA-TR-090730B.pdf

[81] B. Al-Musawi, P. Branch, and G. Armitage, "BGP Replay Tool (BRT) v0.1," Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia, Tech. Rep. 160304A, 04 March 2016. [Online]. Available: http://caia.swin.edu.au/reports/160304A/CAIA-TR-160304A.pdf

[82] D. Blazakis, M. Karir, and J. Baras, "BGP-Inspect - Extracting Information from Raw BGP Data," in *Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP*, April 2006, pp. 174–185.

[83] L. Colitti, G. Di Battista, F. Mariani, M. Patrignani, and M. Pizzonia, "Visualizing Interdomain Routing with BGPlay." *J. Graph Algorithms Appl.*, vol. 9, no. 1, pp. 117–148, 2005.

[84] F. Fischer, J. Fuchs, P.-A. Vervier, F. Mansmann, and O. Thonnard, "VisTracer: A Visual Analytics Tool to Investigate Routing Anomalies in Traceroutes," in *Proceedings of the Ninth International Symposium on Visualization for Cyber Security*, ser. VizSec '12. New York, NY, USA: ACM, 2012, pp. 80–87.

[85] Y.-J. Chi, R. Oliveira, and L. Zhang, "Cyclops: The AS-level Connectivity Observatory," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 5, pp. 5–16, Sep. 2008.

[86] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang, "PHAS: A Prefix Hijack Alert System," in *Proceedings of the 15th Conference on USENIX Security Symposium - Volume 15*, ser. USENIX-SS'06. Berkeley, CA, USA: USENIX Association, 2006.

[87] C. Zheng, L. Ji, D. Pei, J. Wang, and P. Francis, "A Light-weight Distributed Scheme for Detecting Ip Prefix Hijacks in Real-time," in *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, ser. SIGCOMM '07. New York, NY, USA: ACM, 2007, pp. 277–288.

[88] Internet Routing Registry. [Online]. Available: http://www.irr.net/

[89] L. Blunk, J. Damas, F. Parent, and A. Robachevsky, "Routing Policy Specification Language (RPSL)," RFC 4012(Standards Track), Internet Engineering Task Force, March 2005. [Online]. Available: http://tools.ietf.org/html/rfc4012

[90] APNIC Whois Database. [Online]. Available: http://wq.apnic.net/apnic-bin/whois.pl

[91] G. Siganos and M. Faloutsos, "Analyzing BGP policies: methodology and tool," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 3, March 2004, pp. 1640–1651 vol.3.

[92] The Team Cymru Route-server, "IP TO ASN MAPPING." [Online]. Available: http://www.team-cymru.org/IP-ASN-mapping.html

[93] M. Cotton, L. Vegoda, R. Bonica, and A. B. Haberman, "Special-Purpose IP Address Registries," RFC 6890 (Best Current Practice), Internet Engineering Task Force, April 2013. [Online]. Available: http://tools.ietf.org/html/rfc6890

[94] Team Cymru Community Services, "Bogon Route Server Project (Bogons via BGP)." [Online]. Available: http://www.team-cymru.org/Services/Bogons/bgp.html

[95] Bogon Report. [Online]. Available: http://www.cidr-report.org/bogons/

[96] The Team Cymru Route-server, "TEAM CYMRU-Bogon Route Announcements." [Online]. Available: http://www.cymru.com/BGP/bogons.html

[97] MaxMind GeoLite Country: Open Source IP Address to Country Database. [Online]. Available: http://dev.maxmind.com/geoip/legacy/geolite/

[98] IP2location database. [Online]. Available: http://www.ip2location.com/

[99] A. Dainotti, C. Squarcella, E. Aben, K. C. Claffy, M. Chiesa, M. Russo, and A. Pescapé, "Analysis of Country-wide Internet Outages Caused by

Censorship," *IEEE/ACM Trans. Netw.*, vol. 22, no. 6, pp. 1964–1977, Dec. 2014.
[100] Center for Applied Internet Data Analysis (CAIDA), "Archipelago Measurement Infrastructure." [Online]. Available: http://www.caida.org/projects/ark/
[101] A. Feldmann, O. Maennel, Z. M. Mao, A. Berger, and B. Maggs, "Locating Internet Routing Instabilities," *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, pp. 205–218, Aug. 2004.
[102] Q. Gu, Z. Li, and J. Han, "Generalized fisher score for feature selection," *arXiv preprint arXiv:1202.3725*, 2012.
[103] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 8, pp. 1226–1238, Aug 2005.
[104] I. de Urbina Cazenave, E. Kosluk, and M. Ganiz, "An anomaly detection framework for BGP," in *Innovations in Intelligent Systems and Applications (INISTA), 2011 International Symposium on*, June 2011, pp. 107–111.
[105] T. Qiu, L. Ji, D. Pei, J. Wang, J. Xu, and H. Ballani, "Locating Prefix Hijackers Using LOCK," in *Proceedings of the 18th Conference on USENIX Security Symposium*, ser. SSYM'09. Berkeley, CA, USA: USENIX Association, 2009, pp. 135–150.
[106] C. Labovitz, G. R. Malan, and F. Jahanian, "Internet Routing Instability," *IEEE/ACM Trans. Netw.*, vol. 6, no. 5, pp. 515–528, Oct. 1998.
[107] P. Bloomfield, *Fourier analysis of time series: an introduction*. John Wiley & Sons, 2004.
[108] B. A. Prakash, N. Valler, D. Andersen, M. Faloutsos, and C. Faloutsos, "BGP-lens: Patterns and Anomalies in Internet Routing Updates," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '09. New York, NY, USA: ACM, 2009, pp. 1315–1324.
[109] P. Abry and D. Veitch, "Wavelet analysis of long-range-dependent traffic," *Information Theory, IEEE Transactions on*, vol. 44, no. 1, pp. 2–15, Jan 1998.
[110] J. Zhang, J. Rexford, and J. Feigenbaum, "Learning-based Anomaly Detection in BGP Updates," in *Proceedings of the 2005 ACM SIGCOMM Workshop on Mining Network Data*, ser. MineNet '05. New York, NY, USA: ACM, 2005, pp. 219–220.
[111] Y. Xie, H.-A. Kim, D. R. O'Hallaron, M. K. Reiter, and H. Zhang, "Seurat: A Pointillist Approach to Anomaly Detection," in *Recent Advances in Intrusion Detection*, ser. Lecture Notes in Computer Science, E. Jonsson, A. Valdes, and M. Almgren, Eds. Springer Berlin Heidelberg, 2004, vol. 3224, pp. 238–257.
[112] D. Vernon, *Machine Vision: Automated Visual Inspection and Robot Vision*. Prentice Hall, 1991.
[113] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5, pp. 237–329, 2007.
[114] A. Toonk, "Massive route leak causes Internet slowdown," BGPMON, June 2015. [Online]. Available: http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/
[115] A. Lutu, M. Bagnulo, J. Cid-Sueiro, and O. Maennel, "Separating wheat from chaff: Winnowing unintended prefixes using machine learning," in *INFOCOM, 2014 Proceedings IEEE*, April 2014, pp. 943–951.
[116] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
[117] J. Cowie, A. T. Ogielski, B. Premore, E. Smith, and T. Underwood, "Impact of the 2003 blackouts on internet communications," *Preliminary Report, Renesys Corporation (updated March 1, 2004)*, 2003.
[118] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models*. Springer, 1994.
[119] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in *Computational learning theory*. Springer, 1995, pp. 23–37.
[120] Y. Huang, N. Feamster, A. Lakhina, and J. J. Xu, "Diagnosing Network Disruptions with Network-wide Analysis," *SIGMETRICS Perform. Eval. Rev.*, vol. 35, no. 1, pp. 61–72, Jun. 2007.
[121] M. Ganiz, S. Kanitkar, M.-C. Chuah, and W. Pottenger, "Detection of Interdomain Routing Anomalies Based on Higher-Order Path Analysis," in *Data Mining, 2006. ICDM '06. Sixth International Conference on*, Dec 2006, pp. 874–879.
[122] G. Theodoridis, O. Tsigkas, and D. Tzovaras, "A Novel Unsupervised Method for Securing BGP Against Routing Hijacks," in *Computer and Information Sciences III*, E. Gelenbe and R. Lent, Eds. Springer London, 2013, pp. 21–29.
[123] J. Wang, X. Chen, and W. Gao, "Online selecting discriminative tracking features using particle filter," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, June 2005, pp. 1037–1042 vol. 2.
[124] M. Roughan, T. Griffin, M. Mao, A. Greenberg, and B. Freeman, "Combining Routing and Traffic Data for Detection of IP Forwarding Anomalies," in *Proceedings of the Joint International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '04/Performance '04. New York, NY, USA: ACM, 2004, pp. 416–417.
[125] S. Papadopoulos, G. Theodoridis, and D. Tzovaras, "BGPfuse: Using Visual Feature Fusion for the Detection and Attribution of BGP Anomalies," in *Proceedings of the Tenth Workshop on Visualization for Cyber Security*, ser. VizSec '13. New York, NY, USA: ACM, 2013, pp. 57–64.
[126] M. Tahara, N. Tateishi, T. Oimatsu, and S. Majima, "A method to detect prefix hijacking by using ping tests," in *Challenges for Next Generation Network Operations and Service Management*, ser. Lecture Notes in Computer Science, Y. Ma, D. Choi, and S. Ata, Eds. Springer Berlin Heidelberg, 2008, vol. 5297, pp. 390–398.
[127] Y. Xiang, Z. Wang, X. Yin, and J. Wu, "Argus: An accurate and agile system to detecting IP prefix hijacking," in *Network Protocols (ICNP), 2011 19th IEEE International Conference on*, Oct 2011, pp. 43–48.
[128] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane: An Information Plane for Distributed Services," in *Proceedings of the 7th Symposium on Operating Systems Design and Implementation*, ser. OSDI '06. Berkeley, CA, USA: USENIX Association, 2006, pp. 367–380.
[129] S. Sanfilippo, "hping," 2006. [Online]. Available: http://www.hping.org/
[130] G. Fyodor, "Nmap," 2006. [Online]. Available: http://www.nmap.org/
[131] S. Branigan, H. Burch, B. Cheswick, and F. Wojcik, "What can you do with traceroute?" *Internet Computing, IEEE*, vol. 5, no. 5, p. 96, 2001.
[132] Y. C. Hu, "iTraceroute," Purdue University, West Lafayette, 2009. [Online]. Available: https://engineering.purdue.edu/~ychu/itraceroute/
[133] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding traceroute anomalies with Paris traceroute," in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. ACM, 2006, pp. 153–158.

Bahaa Al-Musawi received a B.Sc. and M.Sc. in computer and control engineering from the University of Technology, Iraq in 2003 and 2005 respectively. He is a lecturer at the University of Kufa, Iraq since 2006. Currently, he is a Ph.D. candidate at the Centre for Advanced Internet Architectures, Swinburne University of Technology, Melbourne, Australia. His research interests include routing and network security, network traffic classification, and anomaly detection.

Philip Branch received a Ph.D. in Engineering from Monash University, Victoria, Australia in 2000. Since 2003 he has been an associate professor in Telecommunications Engineering at Swinburne University of Technology, conducting research within the Centre for Advanced Internet Architectures. His research interests are in game traffic, network security and lawful interception. He is a co-author of Networking and Online Games Understanding and Engineering Multiplayer Internet Games (John Wiley and Sons, UK, April 2006).

Grenville Armitage earned a B.Eng. in electrical engineering (Hons) in 1988 and a Ph.D. in electronic engineering in 1994, both from the University of Melbourne, Australia. He is a full professor of telecommunications engineering and founding director of the Center for Advanced Internet Architectures at Swinburne University of Technology. He authored Quality of Service In IP Networks: Foundations for a Multi-Service Internet (Macmillan, April 2000) and co authored Networking and Online Games Understanding and Engineering Multi player Internet Games (Wiley, April 2006). He is also a member of ACM and ACM SIGCOMM.