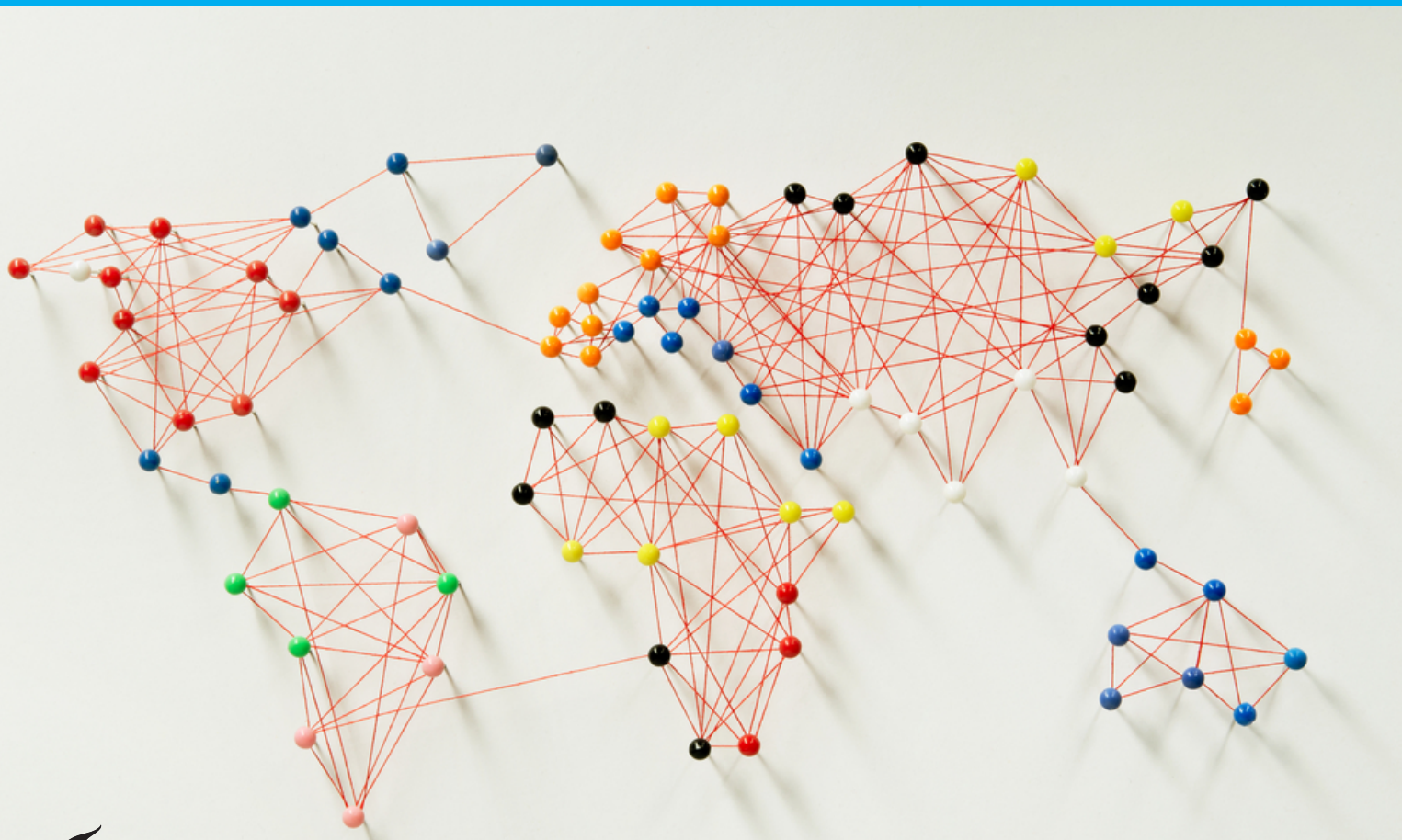
Placement Strategies to Monitor the Inter-Autonomous System Routing Information

Anant Semwal



TU Delft

Placement Strategies to Monitor the Inter-Autonomous System Routing Information

by

Anant Semwal

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on May 27, 2019.

Student number:	4630734		
Project duration:	February 01, 2018– May 16, 2019		
Thesis committee:	Dr. ir. F. A. Kuipers,	TU Delft, Associate Professor	Chair
	Dr. C. Doerr,	TU Delft, Assistant Professor	Supervisor
	Dr. S. Picek,	TU Delft, Assistant Professor	

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Acknowledgements

I would like to acknowledge many individuals without whom this thesis would not have materialized. I would like to begin by expressing my sincere gratitude to my supervisor, dr. Christian Doerr, for trusting me with the freedom to explore various ideas and showing dedicated interest in my progress. I would like to thank him for his patience, encouragement, and supervision.

My master's journey in Delft, for the most part, has been shaped by my friends. I would like to thank Isidora Radevic without whose support my work definitely would not have transpired. I would like to thank Shambhuraj Sawant, Anurag Kulkarni and Abbas Jhabuawala for the fun times we had together. I will always cherish the discussions we had that helped me in getting to the core of my thesis work.

Working in a project group was an exciting experience. I would like to extend my gratitude towards Simone van Veen, Frits Kastelein and Christian Veenman for all the discussions we had regarding the thesis.

Finally, and most importantly, I would like to thank my parents and my brothers for their consistent support and encouragement throughout my entire academic pursuit. They have stood by me in my lows and highs and have always tried to inspire me to do better.

Anant Semwal
Delft, May 16, 2019

Contents

List of Figures

List of Tables

List of Acronyms

Acronym	*Expansion*
AS	Autonomous System
BGP	Border Gateway Protocol
c2p	Customer-to-Provider
EGP	Exterior Gateway Protocol
ICMP	Internet Control Message Protocol
IP	Internet Protocol
IGP	Interior Gateway Protocol
OSPF	Open Shortest Path First
p2c	Provider-to-Customer
p2p	Peer-to-Peer
RC	Route Collector
rrc	Ripe Route Collector
RIB	Routing Information Base
RCC	Regional Co-ordination Centre
s2s	Sibling-to-Sibling
TTL	Time-to-live
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
VP	Vantage Point

1 Introduction

Inter-Autonomous System (AS) route monitoring is the process of collecting the inter-AS routing information. This information flows on the Internet in the form of BGP UPDATE messages, and the BGP data are the messages obtained by the monitors. BGP data are highly valuable from commercial and academic point-of-view and has enabled the study of various aspects of the Internet such as:

1. ***Internet evolution:*** [16, 23, 50] have contributed to studying the aspects of Internet evolution. Evolution studies enable research in directions of capacity planning and management and pave-ways of improving network architecture to improve overall Internet experience.

2. ***Inferring AS relationships:*** [25, 26, 37] have studied BGP data to infer AS relationships as a static property of the Internet and an indicator of AS-level business interactions.

3. ***Stability of the Internet:*** [27] studied aspects of the stability of the Internet. BGP is an unstable protocol and does not respond well to changes. BGP data help in assessing the impact and understanding the causes of the instabilities in the inter-AS routing.

4. ***Security:*** [20, 41, 44] have performed studies on BGP security aspects. BGP hijacks are commonly occurring incidents on the Internet, and unlike normal hijacks, where the affected user base is limited to a single user or a few users, BGP hijacks have an impact on a significant number of users. The impact on a significant number of users is because BGP hijacks potentially target the entire network of AS.

5. ***Commercial use:*** [2, 3] are examples of commercially available BGP analytic tools that provide an insight into the routing health of an ISP's network.

Inter-AS routing is the backbone of Internet routing and enables the routing of Internet traffic between Internet Service Providers (ISPs) and large enterprises. ISPs and large enterprises operate multiple networks under a single "technical administration," and in BGP terminology the "technical administration" is known as the Autonomous Systems (ASes). Therefore, an AS is a collection of networks which are within the technical administration of a single entity. The knowledge of the available paths is used to exchange Internet traffic from source AS to destination AS. A feasible path is a sequence of ASes through which a packet traverse before reaching the destination AS. BGP allows ASes to exchange available paths for all IP addresses with the neighboring ASes.

1.1. Inter-AS route monitoring and challenges

The Inter-AS route monitoring is possible by capturing the BGP UPDATE messages or extracting routing information from a BGP capable routers. Several organizations are performing such monitoring for public and commercial use [2, 3, 9, 11]. In this thesis, while making the selection of monitoring locations (also referred to as monitoring ASes or vantage points) the prior knowledge of the network topology and certainty in BGP information flow on the network will be exploited. Performance monitoring within an Autonomous System

(AS) has been quite popular and is performed to collect performance statistics of the network and to deliver quality services to the end-users, continuous performance monitoring is essential. The network performance statistics collected, are also used in resolving issues between neighboring ASes and ensuring that their neighboring AS meets the agreed service level agreements (SLAs). These statistics provide evidence in case of any disputes arising due to degradation in network services by neighboring AS.

Inter-AS route monitoring differs from network level performance monitoring. While the objective of performance monitoring is to collect data related to parameters like bandwidth, quality-of-service, and speed, the purpose of collecting inter-AS routes is to gain insights into the routing decisions that are made by the ASes, and to determine AS-level properties like prefix's true origin AS and infer AS-relationships.

ASes which provide the Internet access to edge ASes either directly or by purchasing transit from a larger AS (and becoming their customers) are called providers of the edge ASes. ASes exchange a list of IP prefixes (or simply prefixes), reachable through their infrastructure with the neighboring ASes using the Border Gateway Protocol (BGP) and this information is later used by the neighboring ASes to generate a routing information base (RIB), which acts as the reference to route packets towards any destination.

Inter-AS routes (or AS paths) are used to route IP packets from a source AS to the destination AS. The BGP decision process helps in selecting which links are used to form the AS path to route the traffic from their AS. The BGP decision process runs in real-time and communicates the changes in the routing information to the neighboring ASes. The distribution of the updated routing information across the Internet takes place through a series of exchange between the adjacent pairs (of ASes).

Unlike, the internal routes within an AS, the AS paths (or the inter-AS routes) are highly dynamic. The BGP decision process allows an AS to implement path preferences in the form of policies and these policies evolve with time. The changes in local preferences of the ASes impacts the outcome of the routing decision process, and hence transitions in preferred paths are observed. This behavior makes the AS paths highly dynamic. Moreover, network abnormalities, like a faulty network component or a software bug can also force changes in the inter-AS routes.

The relationship with neighboring AS also plays an essential role in influencing the outcome of the routing decision process. The profitability is ensured by preferring the peer routes over the provider routes as the peer routes are the settlement-free arrangement, whereas the provider routes may be a fixed price or pay per use based settlement. Gao reported four types of relations between a pair of neighboring ASes, namely, provider-to-customer, customer-to-provider, peer-to-peer and sibling-to-sibling [25].

1.2. Consequences of the BGP vulnerabilities and inadequate monitoring

Anomalies in the global routing pose a severe threat to security and availability of the Internet. Malicious routing information can compromise the user's sensitive data and induce instability in the global routing, leading to disruption of services. The Internet is a medium to exchange sensitive information from one host to another and has provided businesses with an opportunity to sell/buy products, exchange ideas and collect consumer feedback. Thus any malicious activity or disruption in the Internet services would directly influence the user's and business' interests and to ensure the success of any reactive mechanism to prevent malicious activities on the Internet, proper monitoring capabilities are required. The relevant observations from the monitoring system would enable successful deployment reactive-mechanisms to counteract the impact of malicious behavior.

The foundation of any IT service lies within the confines of CIA, i.e. confidentiality, integrity and availability and malicious activity in inter-AS routing can breach confidentiality or affect availability. Prefix hijack, path hijack, route-leaks, and route-flapping influence confidentiality, and availability of information transported on the Internet. For instance, a small AS may leak a route to larger AS and attract enormous volumes of traffic towards it. If the small AS is not designed to handle such large volumes of traffic, it may disrupt the Internet services in parts of the Internet.

As recent as April 2018, BGP was used maliciously to steal crypto-currencies [10]. The case gained special attention as the impact of this attack was financial. BGP hijacks occur frequently, and implementing exhaustive filtering on the imported/exported routes from/to neighbors is the one of the mechanisms to prevent this. Alternatively, BGP monitoring can be improved to provide information to reactive mechanisms to counteract malicious activities.

BGP offers a mechanism to implement filtering on the imported/exported routes from/to neighbors and requires information regarding the prefix-origin association. The requirement for updated information regarding prefix-origin association, in most scenarios, is not met. There are consistent efforts from the network operators community, the Internet Assigned Numbers Authority (IANA) and regional coordination centers (RCC) to keep the information updated, and is made available by accessing whois database [14]. In the analysis of the whois data, using Unix whois command and for /24 prefixes announced on 4th July 2018, only 137982 out of 457020 found a match in the whois database. The situation is better when it comes to information regarding ownership of an AS, but that information is again not maintained frequently, thus filtering of routes is a challenging mechanism.

RIPE RIS [9] and University of Oregon's RouteViews [11] are the two major systems which collect BGP data. The collectors (or monitors) are deployed worldwide to collect the BGP data. The collectors are BGP capable routers which only receive BGP updates and do not participate in manipulating the routing of the Internet. An AS may connect to the collector by configuring a BGP session with the preferred collector. An AS that establishes a peering session with a collector is called a collector peer, monitor AS, or vantage point (VP) [52]. BGP data collected by these systems is archived and made available publicly for researchers and is called the Public View [22].

Traceroute is a tool widely used by network engineers. While troubleshooting an incident, engineers rely on the information provided by traceroute to connect with concerned party/parties to resolve the incident. CAIDA Ark [4] project is a traceroute-based tool and collects traceroute data from 319 monitors (as of 8th May 2018). This project also offers real-time on-demand measurements as well. The data has been used to map AS-level topology along with BGP data [37]. However, inferring topology through IP information is not so trivial, and a study related to IP-to-AS mapping by Mao et al. reported that about 10% of the prefixes are mapped to multiple origin ASes, thus making the IP-to-AS mapping challenging [38].

The ARK data can discover AS links which are not discovered using BGP data. This shows that BGP data is provides an incomplete view of the Internet. The incompleteness in the link related information regarding the Internet leads to misleading conclusions regarding the state of the Internet and prevents useful advancements towards the improvement of the network architecture and network protocols.

1.3. Organization of the report

The report contains six chapters. Chapter 1 introduces the topic and discusses the inter-AS monitoring, motivation for the thesis and states the research questions. In chapter 2 the background of the inter-AS routing is explained and includes a discussion on the Border Gateway Protocol, message type and other relevant technical terms, necessary for understanding the report. In chapter 3, state-of-the-art is discussed for inter-AS route monitoring infrastructure and includes the importance of inter-AS route monitoring, architectural vulnerabilities and proposed solutions for the BGP vulnerabilities. In chapter 4, we discuss the graph motifs of the Internet and the Internet model used for evaluation of a novel algorithm presented in the thesis — finally, the presentation of algorithms for monitor selection. The discussions regarding the evaluation of the algorithm is present in chapter 5. The performance of various monitor set selection (used interchangeably with VP set selection or VP selection) schemes is done based on link-coverage, and the impact of real-network conditions on the VP-selection scheme proposed. Chapter 6, presents concluding remarks and highlights the contributions of the work and ends with a note on future research directions.

1.4. Motivation

The nature of BGP announcement is predictable with a certain level of confidence by carefully examining the relationships between the neighboring ASes and the source of BGP UPDATE. In [22], Chen et al. revealed that the adjacent ASes that act as monitoring points are the source of redundant information in the BGP data. AS-relationships between two peers lead to redundant observations regarding inter-AS routes. The redundant observations artificially inflate the size of the BGP data. Additionally, several ASes are sources of pathological or multiple duplicate UPDATE messages and are another reason for an increase in the size of BGP data [29]. Thus a researcher is forced to process entire BGP data to extract necessary information or, in some instances, arbitrary decisions are made to reduce the size of BGP data. For example, [16] used rrc00 collector data, which is one of the collectors of RIPE project [9] and based the decision on the fact that rrc00 collector is in

the default-free zone and must have the complete view of the AS-level internet.

While addressing issues related to pathological updates and security of the inter-AS routing remain outside the scope of this thesis, the main focus is to discuss monitor placement schemes for collecting inter-AS routing information. The monitor placement aims to reduce the size of redundant BGP data collected as well as improve the quality of the BGP data. The studies have revealed that there is a redundancy in the BGP data which degrades the quality of the BGP data. Additionally, not much work is done to improve the ways to tackle the BGP data redundancy in BGP monitoring projects. In [52], Zhang et al. have discussed (1) Random based, (2) Degree based, (3) Greedy Link Based, and (4) Address block-based algorithms for BGP monitor selection. However, these algorithms do not consider the possibility of expanding the monitoring network, and no study is performed to understand the benefits of these algorithms in terms of link coverage when used to expand the monitor network.

1.5. Research question

Studies have shown that there is a redundancy in the BGP data. Additionally, BGP data fails to capture at least one-thirds of inter-AS links. In this thesis, we focus on developing strategies for selecting the BGP monitors to improve the quality of BGP data. The following hypothesis forms the basis of this thesis:

"Monitors for inter-AS route monitoring if selected based on the number of peer-to-peer links would provide a link coverage better than the simple degree-based monitor selection. And, the number of monitors can be removed if BGP message flow is considered."

The objectives of this thesis revolve around the hypothesis, and the formulates the following research questions:

1. Can peering-degree based monitor selection scheme outperform monitor placement schemes discussed in [52]?
2. What would be an appropriate topology generator that can be used to evaluate monitor set performance?
3. The global view of the Internet requires consolidation of the local views of multiple ASes. What would be the minimum number of monitoring ASes required to achieve 100% link coverage (complete visibility)?
4. How does the monitor placement algorithm perform under real Internet condition of limited visibility and incorrect relationship inferences?

2

Theoretical background

In this chapter technical preliminaries required for understanding the discussions in the following chapters are presented and includes a presentation of the topological model of the Internet, explanation of inter-AS routing and the BGP protocol. Additionally, discuss the selected topics from graph theory which would find their relevance in the following chapters.

2.1. The Internet topology

The Internet topology is studied at three different levels (figure 2.1) which are as follows:

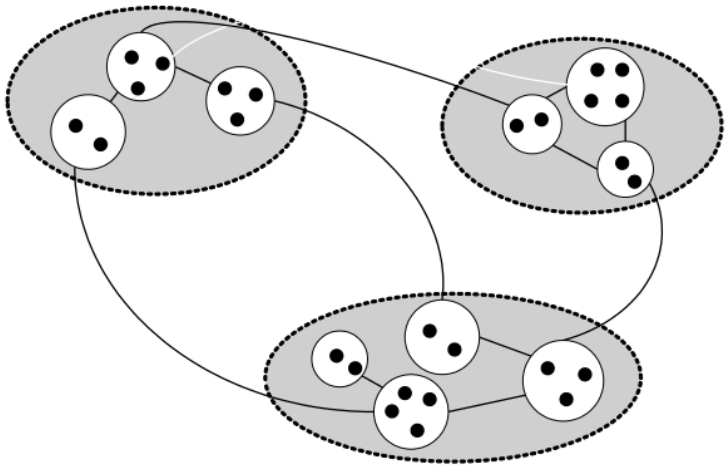


Figure 2.1: The Internet topology [42]

1. ***IP-level:*** In this model, network interface cards are the nodes, and the links that join these network interface cards are the edges and is the microscopic view of the Internet. To obtain this view, network specifications are required as there are no tools which can help discover this detail [43]. Black dots within white circles are representing interface cards of a router in figure 2.1.

2. ***Router-level:*** In this model, routers act as nodes and the physical links between two routers act as edges of the graph. Multiple traceroutes are generated and collected to generate this model. [43]. Routers are represented as white circles in figure 2.1.

   In figure 2.2 the list of IP addresses traveled by the IP packets which are destined to IP address 108.177.119.102 are shown, and the result is generated using the traceroute tool. Every router to which the probe packets generated by traceroute reach responds with its IP address. Asterisk indicates that the probe packet was

```
asemwal@asemwal-Inspiron-5558:~$ traceroute  131.180.77.82
traceroute to 131.180.77.82 (131.180.77.82), 30 hops max, 60 byte packets
 1  compalhub.home (192.168.178.1)  3.088 ms  3.829 ms  5.383 ms
 2  * * *
 3  nl-ams02a-rc2-lag-8-0.aorta.net (84.116.130.98)  24.014 ms  31.042 ms  34.179 ms
 4  nl-ams04a-ri3-ae8-0.aorta.net (84.116.130.97)  35.502 ms  34.886 ms  35.622 ms
 5  xe-2-1-5.jnr01.asd001a.surf.net (145.145.166.89)  34.684 ms  43.336 ms  46.845 ms
 6  ae1.500.asd002a-jnx-01.surf.net (145.145.176.3)  47.141 ms  27.192 ms  27.355 ms
 7  tudelft-router.customer.surf.net (145.145.26.98)  27.642 ms  29.414 ms  30.174 ms
 8  * * *
 9  * * *
10  scd.tudelft.nl (131.180.77.82)  33.985 ms  33.972 ms  30.679 ms
asemwal@asemwal-Inspiron-5558:~$
```

Figure 2.2: Traceroute output.

dropped at the respective hop and the packet may be dropped either due to performance constraints at the respective router or excessive traffic at the router.

3. ***AS-level:*** In this model, Autonomous Systems (ASes) act as the nodes of the graph and the edge indicates interconnection among the ASes [43]. The data to generate this model is gathered using the BGP UPDATE messages, or by collecting the routing information base (RIB) from the routers. The gray area in figure 2.1 represents the ASes.

```
Username: rviews
route-views>show ip bgp
BGP table version is 12519865, local router ID is 128.223.51.103
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *   1.0.0.0/24       198.58.198.254                         0 1403 13335 i
 *                    198.58.198.255                         0 1403 13335 i
 *                    212.66.96.126                          0 20912 13335 i
 *                    193.0.0.56                             0 3333 1257 13335 i
 *                    194.85.40.15             0             0 3267 13335 i
 *                    202.93.8.242                           0 24441 13335 i
 *                    140.192.8.16                           0 54728 20130 6939 13335 i
 *                    162.251.163.2                          0 53767 13335 i
 *                    195.208.112.161                        0 3277 3267 13335 i
 *                    4.69.184.193             0             0 3356 2914 13335 i
 *                    208.51.134.254           0             0 3549 3356 1299 13335 i
 *                    89.149.178.10           10             0 3257 13335 i
 *                    132.198.255.253                        0 1351 10578 13335 i
 *                    144.228.241.130         80             0 1239 6453 13335 i
 *                    206.24.210.80                          0 3561 209 1299 13335 i
 *                    203.62.252.83                          0 1221 13335 i
 *                    208.74.64.40                           0 19214 174 13335 i
 *                    137.39.3.55                            0 701 2914 13335 i
 *                    64.71.137.241                          0 6939 13335 i
 *                    207.172.6.20             0             0 6079 13335 i
 *                    207.172.6.1              0             0 6079 13335 i
 *                    202.232.0.2                            0 2497 13335 i
 *                    37.139.139.0                           0 57866 1299 13335 i
 *                    173.205.57.234                         0 53364 3257 13335 i
 *                    217.192.89.50                          0 3303 13335 i
 *                    114.31.199.1                           0 4826 13335 i
 *                    12.0.1.63                              0 7018 6453 13335 i
 *>                   91.218.184.60                          0 49788 13335 i
```

Figure 2.3: Local-RIB of a route collector [11].

The figure 2.3 shows a part of RIB from a BGP speaker. RIB is used to resolve the next hop (IP address of the next hop/router) for traffic destined to a host belonging to prefix 1.0.0.0/24 and multiple paths are maintained for every prefix to ensure that there is no loss in connectivity, however only the best-path is used to route traffic. AS-level topology is generated using the values in the "path" column of the table, and prefix-origin association can be determined using the last AS in the corresponding path for the respective prefix.

2.2. Inter-Autonomous System routing

An Autonomous System is a group of routers within the technical administration of a single Internet Service Provider (ISP) or an enterprise. The routers of the same AS, communicate the routing information using Interior Gateway Protocols (IGP), such as OSPF (Open-Shortest Path First). Also, they determine routes using standard metrics and exchange routing information with routers of a different AS using an Exterior Gateway Protocol (EGP), such as BGP (Border Gateway Protocol) [28].

The traffic on the Internet flows in the form of IP packets and these packets travel in a hop-by-hop manner from one router to another until the packet can be delivered to the intended recipient. Local-RIB is used to

determine the next hop at every router and path attribute from the local-RIB can be used to generate an AS-level topology. The topology is the router's interpretation (visualization) of the AS-level interconnections and is called as the local view of the Internet. Moreover, by definition of AS, the routers of the same AS, will have the same local view to ensure consistency in routing decisions and thus for an AS's local view the local-RIB from single BGP speaker of an AS is sufficient.

The global view of the Internet requires consolidation of the local views of multiple ASes. Theoretically, the global or complete view of the Internet can be obtained using the local-RIBs of the ASes from DFZ (default-free zone). DFZ refers to the collection of all the Internet ASes that do not require a default route to transport a packet to any destination. Conceptually, DFZ routers have a **complete** BGP table, sometimes referred to as the Internet routing table, global routing table or global BGP table. However, the widespread use of route filtering and the rapid rate of change in Internet routing ensure that no router anywhere has a complete view of all routes. Also, any such routing table would look different from the perspective of different routers, even if it achieves a stable view. Thus there is a need to determine the minimum number of ASes from which RIBs must be collected to generate the complete AS-level topology of the Internet, and this thesis addresses this question.

2.3. Border Gateway Protocol

The BGP is a dynamic routing protocol which is capable of propagating inter-AS routing information efficiently across the Internet, and a router which supports BGP is known as the BGP speaker. BGP is an Inter-AS Routing Protocol and could serve as EGP and IGP. BGP is a multipurpose application that is capable of:

1. Determining the best-routes to every reachable IP address.
2. Exchanging the routing information with neighboring routers efficiently.
3. Detecting unavailability of primary-path to dynamically switch to the next best path.
4. Communicating only incremental changes and reduces network overheads due to BGP.

Routing Information Base or the RIB

Routing information base (RIB) is the critical database which acts as a bridge between the two logical separations of the router, namely the control plane and the forwarding plane [51]. There are three sub-classification of RIB, and they are as follows:

1. ***Adj-RIB-In:*** is the database that contains unprocessed routing information that has been advertised to the local BGP speaker by its neighbors and includes the incoming routing information along with the tag indicating which neighboring AS announced the information.
2. ***Local-RIB:*** The local routing information base stores the resulted information from processing the RIBs-In database's information. These are the routes that are used locally after applying BGP policies and decision process[1].
3. ***Adj-RIB-Out:*** is the database that contains the routes for advertisement to specific neighbors through the local speaker's UPDATE messages and includes the outgoing routing information.

BGP peering

BGP peering is the process of establishing a BGP session with the neighboring BGP speaker. It is sub-classified into Interior BGP peering (iBGP) and Exterior BGP peering (eBGP) depending upon the AS which contains the respective BGP speakers. The figure 2.4 shows iBGP and eBGP peering and their description are as follows:

1. ***iBGP peering:*** iBGP or Interior Border Gateway protocol peering is the BGP session established between two BGP capable speakers which are part of the same AS and allows routers within an AS to

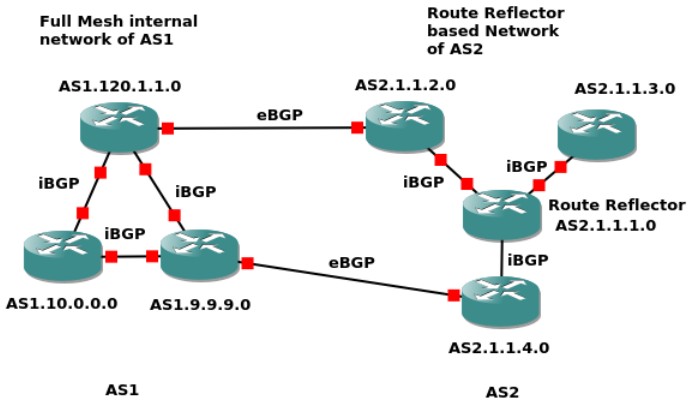


Figure 2.4: BGP peering

exchange reachability information with one another. iBGP requires routers within the same AS to be connected in full mesh. This requirement is fulfilled using *route reflectors* [15]. Multiple BGP routers can peer with a central point, the RR- acting as a route reflector (RR) server, rather than peer with every other router in a full mesh. All the other iBGP routers become route reflector clients. The figure 2.4 shows the two possible configurations of iBGP interconnections.

2. ***eBGP peering:*** eBGP or Exterior Border Gateway protocol peering is the BGP session established between BGP capable routers that belong to different AS and allows routers within an AS to exchange reachability information with routers in different ASes.

Point-of-peering

Point-of-peering or PoP is the location where ASes peer with each other and maybe a private peering location or an IXP (Internet Exchange Point). An Internet Exchange Point is where multiple ASes come together to peer and exchange traffic between their networks. IXPs are not Internet service providers and only provide infrastructure service point for multiple ASes. Private peering is a point of Internet exchange which is accessible to limited ASes.

Route server[13]

A route server provides a look into the IP routing tables of the autonomous system in which the server resides. The concept of a route server has its origins in the old Unix-based route servers that used to be located in the Network Access Points (NAP) during the early days of the Internet. These Unix machines provide a custom routing software *routed*, explicitly designed to make best-path calculations and distribute a routing table to the routing devices forming the backbone of the Internet at these significant peering points.

A route server is used most frequently by network operations engineers while trying to determine the cause of connectivity failures their customers are experiencing between two endpoints. Typically, when the network is running, and both sites are reachable from the NOC engineer's location, but the customer is unable to access the desired resource. A route server provides a view into other AS's network and enables the NOC engineer to trace the problem of a lousy route advertisement, aggregation or pre-pending errors.

BGP messages

TCP connections are used to send BGP messages, usually on port 179, and the maximum size of the BGP message is 4096 octets. The processing of a message is done once it is received entirely. The different types of BGP messages are as follows:

1. ***OPEN*** message is used to establish a BGP session with a BGP speaker and is used to send a connection request to a BGP speaker who is usually one hop away.

2. ***UPDATE*** message is used to exchange incremental changes in the routing information with a BGP neighbor. As a one-time activity the complete RIB is exchanged when the BGP connection is initialized.

3. ***NOTIFICATION*** message is used to communicate errors encountered in a BGP session.

4. ***KEEPALIVE*** messages are exchanged to avoid termination of a BGP session due to connection timeout and lets BGP neighbor know if the session is still active.

5. ***ROUTE-REFRESH [21]*** is used to request BGP neighbor to send its RIB on an existing BGP session thereby eliminating the need for reestablishing the BGP session to get complete RIB from neighboring AS.

BGP connection mechanism

A BGP speaker initiates the BGP session by sending an OPEN message to the BGP speaker it wishes to peer. Upon receiving the OPEN message, the neighboring BGP speaker responds with a 19-octet KEEPALIVE message to acknowledge that the connection is successfully established or a NOTIFICATION message with the relevant error code to indicate the reason for declining the connection request. A connection request may be unacceptable to the BGP speaker if the capability negotiations fail; for example, version is unsupported or more than maximum prefixes configured is announced. The BGP speaker retires the connection after it has addressed the reason for which the error code was received.

Upon successfully establishing the BGP session between the BGP speakers, the local-RIB is exchanged between BGP speakers using the UPDATE messages. If there are no routes to be exchanged, a BGP peer must exchange KEEPALIVE message after the connection is idle for a maximum idle time. At present this value is set to 60 seconds. If either peer does not send KEEPALIVE message, the hold timer expires. In this situation, the peer for which the hold timer expired, would send a NOTIFICATION message to BGP peer with the respective error code and terminate the BGP session.

AS identifier

An AS is a network or a collection of networks, in which the routing decisions are uniform. To enforce the uniformity in routing decisions routing policies are applied on the routes chosen by each router which lies within the jurisdiction of the respective AS. A 32-bit identifier known as the AS number or AS identifier is assigned to each BGP speaker (BGP enabled router) within an AS.

Network Layer Reachability Information

Network Layer Reachability Information (NLRI) is a tuple of length and prefix of the network for which the neighboring BGP speaker provides next hop IP and feasible AS path. The components of the NLRI are as follows:

1. ***Length and prefix:*** CIDR (Classless Inter-Domain Routing) is a memory efficient way of representing IP addresses. In this, continuous IP addresses are identified as a single-block and are described by the length and prefix. The length in CIDR represents the size of the network, and a larger number indicates a smaller network size. The prefix is ideally the first IP address within the aggregated IP address space. CIDR representation of IP addresses is shown in table 2.1.

Start IP	**End IP**	**CIDR notation**		**Comments**
		Prefix	**Length**	
0.0.0.0	255.255.255.255	0.0.0.0	/0	Entire IPv4 range
12.0.15.0	12.0.15.255	12.0.15.0	/24	255 hosts
150.0.0.0	150.255.255.255	150.0.0.0	/8	16581375 hosts

Table 2.1: CIDR notation and IP ranges.

Representation of IP addresses in CIDR representation reduces the storage requirement at a router for maintaining the table for next-hop information at least by 255 times (size of smallest /24 prefix permissible by BGP). The next-hop information is maintained in a router based on the CIDR representation of IP addresses and next-hop IP address (see figure 2.3). The destination IP address of a packet is first resolved into the largest length, and the next hop IP is derived to route packets.

2. ***AS Path or Path:*** Each BGP speaker exchanges feasible path with its neighbors to indicate the distance to the destination prefix. The AS-path is a string of ASes through which a packet would traverse before reaching the destination (see figure 2.3). The information from the AS path is also used to prevent routing loops. In a case in the AS path, AS identifier of the receiving AS is already present, the respective AS path is ignored to avoid routing loops.

3. ***Next Hop IP or Next hop:*** An IP address identifies every routing component on the Internet. Thus it is essential for any router to know the IP address of next routing device to forward IP packets (see figure 2.3) to successfully deliver the IP packet to the intended recipient. This IP address is known as the next hop IP. In most cases, next hop IP is the same as the IP of the neighboring BGP speaker which provides the routing information.

AS relationships

The neighboring ASes of an AS may have different business agreements for exchanging Internet traffic. These arrangements are either paid or settlement-free. Depending upon the type of arrangement, AS relationships may be classified into the following four categories as defined by Gao in [25]:

1. Provider-to-customer: is a directed link from a provider AS to a customer-AS. Provider AS provides transit services to the customer AS and charges for the same.

2. Customer-to-provider: is a directed link that is in the opposite direction of the provider-to-customer link.

3. Peer-to-peer: is the symmetric link between two ASes that have a settlement-free arrangement to exchange traffic.

4. Sibling-to-sibling: emerges as a result of corporate mergers and acquisitions and is also the symmetric link.

The AS relationship with the source from which routing information is received would determine which neighboring ASes would receive the respective routing information. The same is summarized in table 2.2.

Origin-Source relationship	Source-Neighbour relationship	Sent
SELF	ANY	yes
customer-to-provider	ANY	yes
sibling-to-sibling	ANY	yes
provider-to-customer	provider-to-customer	yes
provider-to-customer	sibling-to-sibling	yes
provider-to-customer	peer-to-peer	no
provider-to-customer	customer-to-provider	no
peer-to-peer	provider-to-customer	yes
peer-to-peer	sibling-to-sibling	yes
peer-to-peer	peer-to-peer	no
peer-to-peer	customer-to-provider	no

Table 2.2: BGP routing information flow rules based on AS relationships.

The cone

The Internet is accessible to any AS on the Internet using three possible options:

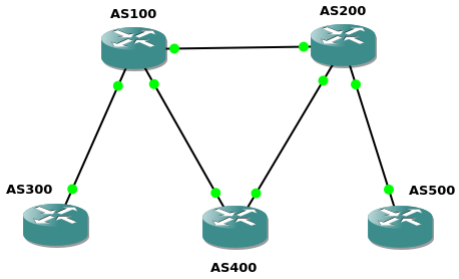
1. ***Self network or customer network:*** This is the primary choice of the network operator as this selection generates revenue and ensures faster delivery. This set of ASes is also defined as the customer cone [5].

2. ***Peer network:*** This is the second choice of the network operator as this selection is economical when compared with the provider network. The peering is a settlement-free business arrangement between the network operators. Like, customer cone, this set of ASes can be defined as the peer-customer cone.

3. ***Provider network:*** This is the ultimate choice. The provider of an AS charges the customer AS for the volumes of data that is routed through the provider's infrastructure and thus provider is responsible for providing access to the global internet to the customer AS either through its network or through peering networks. It is also possible that a provider may be a customer of another AS which would enable global internet services for such provider AS and its customers.

Inter-AS routing information exchange

In figure 2.5a, a system of five BGP routers belonging to different ASes is shown. AS100 is the provider of AS300 and AS400, AS200 is the provider AS of AS400 and AS500, and AS100 is a peer of AS200. The table lists prefixes assigned to each AS and the router IP for each router is X.X.X.0 for all ASes (see table 2.5b).



(a) AS-graph for Toy Internet

AS	Router IP	Prefix Originated
AS100	12.0.0.0	-
AS200	20.0.0.0	-
AS300	30.0.0.0	30.0.0.0/24
AS400	40.0.0.0	40.0.0.0/24
AS500	50.0.0.0	50.0.0.0/24

(b) Configuration of the toy-Internet

Figure 2.5: Toy-Graph for discussing inter-AS routing mechanism

As soon as the BGP session is initialized between neighboring ASes, each AS would share routing information of the prefixes they own. In this case, AS100 and AS200 would not announce any prefix since they own none. AS300 and AS400 would announce their prefixes to AS100 which in this case is their provider AS. Simultaneously, AS400 and AS500 would announce their prefixes to AS200. Based on the routing information received by AS100 and AS200 from their respective customers, they would create their routing information. In table 2.3 routing information of AS100 and AS200 are shown.

AS Identifier	Prefix	Next Hop	Path
AS100	30.0.0.0/24	30.0.0.0	AS300
	40.0.0.0/24	40.0.0.0	AS400
AS200	40.0.0.0/24	40.0.0.0	AS400
	50.0.0.0/24	50.0.0.0	AS500

Table 2.3: RIB for AS100 and AS200

Once the routing information is generated by AS100 and AS200, based on their outgoing routing policies, they would prepare announcements for their neighboring ASes. Adj-RIB-In for the neighbors of AS100 and AS200 is shown in table 2.4 and would be used by the recipients to determine routing information for

remaining prefixes. For instance, now AS300 can reach prefix 40.0.0.0/24 using path (AS100 AS400) by choosing 12.0.0.0 as next hop. It is important to note, that AS200 and AS100 have received a new path to prefix 40.0.0.0/24 from AS100 and AS200 respectively. These new paths would be stored as backup paths in the RIBs of respective ASes and are because the BGP decision process prefers shorter paths to a prefix. Moreover, since there was no change in routing decision by AS100 and AS200 to reach prefix 40.0.0.0/24, it would not be announced to the neighboring ASes unless primary path becomes unavailable due to link failure or becomes unfeasible for any other reasons. Table 2.5 summarizes the resulting RIBs for all ASes.

AS Identifier	Announcing AS	Prefix	Next Hop	Path
AS300	AS100	40.0.0.0/24	12.0.0.0	AS100 AS400
AS400	AS100	30.0.0.0/24	12.0.0.0	AS100 AS300
AS400	AS200	50.0.0.0/24	20.0.0.0	AS200 AS500
AS500	AS200	40.0.0.0/24	20.0.0.0	AS200 AS400
AS200	AS100	40.0.0.0/24	12.0.0.0	AS100 AS400
		30.0.0.0/24	12.0.0.0	AS100 AS300
AS100	AS200	50.0.0.0/24	20.0.0.0	AS200 AS500
		40.0.0.0/24	20.0.0.0	AS200 AS400

Table 2.4: Adj-RIB-In for the neighbours of AS100 and AS200.

AS Identifier	Prefix	Next Hop	Path
AS100	30.0.0.0/24	30.0.0.0	AS300
	40.0.0.0/24	40.0.0.0	AS400
	40.0.0.0/24	20.0.0.0	AS200 AS400
	50.0.0.0/24	20.0.0.0	AS200 AS500
AS200	30.0.0.0/24	12.0.0.0	AS100 AS300
	40.0.0.0/24	40.0.0.0	AS400
	40.0.0.0/24	12.0.0.0	AS100 AS400
	50.0.0.0/24	50.0.0.0	AS500
AS300	30.0.0.0/24	30.0.0.0	
	40.0.0.0/24	12.0.0.0	AS100 AS400
AS500	50.0.0.0/24	50.0.0.0	
	40.0.0.0/24	20.0.0.0	AS200 AS400
AS400	40.0.0.0/24	40.0.0.0	
	30.0.0.0/24	12.0.0.0	AS100 AS300
	50.0.0.0/24	20.0.0.0	AS200 AS500

Table 2.5: Local RIB after processing announcements from AS100 and AS200.

The RIB of AS100 and AS200 is updated with a path to prefix 50.0.0.0/24 and 30.0.0.0/24 respectively and would trigger AS100 and AS200 to share newly acquired routing information with their neighboring ASes. Thus once again, Adj-RIB-In of ASes 300,400 and 500 will be updated (table 2.6). Upon receiving announcements from neighboring ASes, AS300, AS400, and AS500 would update their routing tables. The final BGP tables at all ASes is shown in table 2.7.

AS Identifier	Announcing AS	Prefix	Next Hop	Path
AS300	AS100	50.0.0.0/24	12.0.0.0	AS100 AS200 AS500
AS400	AS100	50.0.0.0/24	12.0.0.0	AS100 AS200 AS500
AS400	AS200	30.0.0.0/24	20.0.0.0	AS200 AS100 AS300
AS500	AS200	30.0.0.0/24	20.0.0.0	AS200 AS100 AS300

Table 2.6: Adj-RIB-In for ASes

RIB of AS300 and AS500 would accept new paths to 50.0.0.0/24 and 30.0.0.0/24 respectively. Moreover,

AS Identifier	Prefix	Next Hop	Path
AS100	30.0.0.0/24	30.0.0.0	AS300
	40.0.0.0/24	40.0.0.0	AS400
	40.0.0.0/24	20.0.0.0	AS200 AS400
	50.0.0.0/24	20.0.0.0	AS200 AS500
AS200	30.0.0.0/24	12.0.0.0	AS100 AS300
	40.0.0.0/24	40.0.0.0	AS400
	40.0.0.0/24	12.0.0.0	AS100 AS400
	50.0.0.0/24	50.0.0.0	AS500
AS300	30.0.0.0/24	30.0.0.0	
	40.0.0.0/24	12.0.0.0	AS100 AS400
	50.0.0.0/24	12.0.0.0	AS100 AS200 AS500
AS500	50.0.0.0/24	50.0.0.0	
	40.0.0.0/24	20.0.0.0	AS200 AS400
	30.0.0.0/24	20.0.0.0	AS200 AS100 AS300
AS400	40.0.0.0/24	40.0.0.0	
	30.0.0.0/24	12.0.0.0	AS100 AS300
	50.0.0.0/24	20.0.0.0	AS200 AS500

Table 2.7: Stable RIB for ASes

AS400 would continue using existing paths to prefixes 30.0.0.0/24 and 50.0.0.0/24 and store new paths to these prefixes as backup routes and would be used in case the primary route becomes unavailable. It is important to point out, that AS400 did not announce routes learned from AS200 to AS100 and vice-versa and is because AS200 and AS100 are providers of AS400 and the routes received from the provider are not shared with other providers and peers (see table 2.2).

Link failure

To understand the dynamic nature of inter-AS routes let us consider figure 2.5a once again. In case the link between AS200 and AS400 fails for any technical or non-technical reason, AS400 becomes unreachable to AS200 through its primary path. However, in the RIB of AS200, an alternate path to AS400 through AS100 exists. Thus, AS200 would update its primary path as the path "AS100 AS400" and send UPDATE messages to its neighboring ASes. Table 2.8 shows the final RIB of all ASes under AS200-AS400 link failure.

AS Identifier	Prefix	Next Hop	Path
AS100	30.0.0.0/24	30.0.0.0	AS300
	40.0.0.0/24	40.0.0.0	AS400
	50.0.0.0/24	20.0.0.0	AS200 AS500
AS200	30.0.0.0/24	12.0.0.0	AS100 AS300
	40.0.0.0/24	12.0.0.0	AS100 AS400
	50.0.0.0/24	50.0.0.0	AS500
AS300	30.0.0.0/24	30.0.0.0	
	40.0.0.0/24	12.0.0.0	AS100 AS400
	50.0.0.0/24	12.0.0.0	AS100 AS200 AS500
AS500	50.0.0.0/24	50.0.0.0	
	40.0.0.0/24	20.0.0.0	AS200 AS100 AS400
	30.0.0.0/24	20.0.0.0	AS200 AS100 AS300
AS400	40.0.0.0/24	40.0.0.0	
	30.0.0.0/24	12.0.0.0	AS100 AS300
	50.0.0.0/24	20.0.0.0	AS200 AS500

Table 2.8: Stable RIB for ASes under AS200-AS400 link failure

RIB to AS-graph

AS-graph is an alternate term used to describe AS-level topology of the Internet. The AS-graph could be defined as a graph G(V, E) where V is a set of ASes and E is a set of edges connecting ASes. The path to link (edge) conversion can be done using a simple algorithm (algorithm 1). Table 2.9 shows a list of links that can be extracted from each AS and is also termed as AS-Link map.

The RIBs from ASes can be used to generate the complete AS-graph, for instance, if RIBs of AS300 and AS500 are selected then complete graph can be generated using path column from table 2.7. However, upon closer inspection, a partial graph of the toy-Internet (see figure 2.5a) can be generated just by using either AS300 or AS500. There is only one new link which is discovered by additional monitoring AS, also showing that adjacent ASes have high information overlap and by reducing monitoring nodes by 50%, graph visibility is affected by only 20% (for this case).

AS Identifier	**Visible Links using RIB from table 2.7**
AS100	AS100 AS300, AS100 AS400, AS200 AS500, AS100 AS200, AS200 AS400
AS200	AS200 AS100, AS200 AS500, AS100 AS400, AS200 AS400, AS100 AS300
AS300	AS200 AS500, AS100 AS200, AS100 AS400, AS300 AS100
AS400	AS400 AS200, AS200 AS500, AS100 AS300, AS400 AS100
AS500	AS200 AS100, AS100 AS300, AS200 AS400, AS500 AS200

Table 2.9: AS link map generated from the RIBs of all ASes in figure 2.5a

On the one hand, AS300 and AS500 may be used to have the complete-visibility of toy-Internet; on the other hand, AS100 or AS200 can solely, provide the complete-visibility. However, there is no way of determining what combination of ASes would provide better coverage and this thesis answers this question by performing traffic flow analysis on various Internet-like smaller graphs.

To understand the monitor selection process, sub-graphs that are embedded on the Internet require analysis from a BGP traffic perspective. A larger graph comprises of smaller sub-graphs which are interconnected. These are known as graph motifs. Moreover, from the previous discussion, the information from adjacent ASes would be redundant. Thus there is a need to eliminate adjacent ASes from the monitor set selected and could be achieved by looking at the monitor set selection from the perspective of the classic independent-set problem of the graph theory, and is discussed in the next section.

2.4. Graph motifs

Graph motifs are sub-graphs that repeat themselves in a specific network or even among various networks. Each of these sub-graphs, defined by a particular pattern of interactions between vertices, may reflect a framework in which particular functions are achieved efficiently. Indeed, motifs are of notable importance primarily because they may reflect functional properties. They have recently gathered much attention as a useful concept to uncover structural design principles of complex networks. Although graph motifs may provide a deep insight into the network's functional abilities, their detection is computationally challenging. The graph motif exploration is a computationally intensive because:

1. It requires listing all possible combinations of vertices in the graph $\binom{n}{k}$ where n is the number of nodes and k is the size of the motif.

2. It needs to validate that the selected combination generates a connected graph to ensure that BGP messages can flow.

3. To extract the graph motif and count the number of occurrences for each motif or create a new entry when the motif is encountered for the first time.

The idea about the existence of recurring structures within a complex network was presented by Milo. The network motifs are defined as patterns of interconnections occurring in a complex network at numbers that are significantly higher than those in randomized networks [39]. Graph motifs and motifs are other terms that define the idea of network motifs. The figure 2.6 shows graph motifs of size three for a directed graph.

This thesis implements an algorithm motivated by the existing sampling-based approach, and the pseudo-code for the same is presented in algorithm 2. The sampling reduces the memory overhead for listing all combination of ASes by making combinations of *k-size* among the neighbors of an AS. The sampling is done in the neighborhood of an AS because the likelihood of obtaining a connected graph is lower when two ASes which are far apart are selected. The program maintains the set of motifs encountered and increments the counter each time it encounters a similar motif. In the event where the program identifies a new motif, the new motif is appended to the set of motifs.

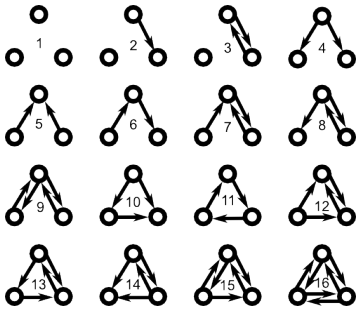


Figure 2.6: Graph Motifs [12].

3

State-of-the-art review

This chapter discusses the architectural vulnerabilities of the Border Gateway Protocol and the state-of-the-art to counteract the malicious activities possible due to these vulnerabilities. The counter-measures are discussed to highlight that inter-AS route monitoring is one of the most promising ways to protect the interests of the ASes. This discussion follows a discussion on existing inter-AS route monitoring infrastructure and a suitable Internet topology generator for evaluation of various monitor placement schemes. The chapter concludes with an explanation on the working of the BGP simulator developed by [31]. This simulator is used to speed up the evaluation of various monitor placement schemes by generating inter-AS routing information through simulations.

As we shall see, existing methods of monitor placement, take into account the AS-level topological information and AS-path information and these methods provide inadequate monitoring capabilities. In this thesis, the monitor placement schemes, which are discussed in the following chapter, take into account the BGP message flow along with the topological information of the network. Thus, the schemes allow a more customized monitor placement for inter-AS route monitoring.

3.1. The Border Gateway Protocol

Vulnerabilities of the BGP

In 1989, Finn published that the methods of dynamic routing used by computer networks, may incorporate elements in their design that allow widespread denial of service. Also, introduced the design requirements for reducing the vulnerability of a dynamic computer network by suggesting that a network should be resistant to direct and indirect attacks and formulated that *an attacker should gain no relative advantage via an indirect versus a direct attack* [24]. In [33], discussed the two sources of threats to secure operation of the routing protocol in a network by pointing out that a subverted router is participating legitimately in routing protocol and an illegal intruder who may illegally attempt to interfere in the routing protocols by masquerading as routers. Smith and Garcia-Luna-Aceves presented their analysis on the security of the Border Gateway Protocol (BGP) and discussed security vulnerabilities in the BGP routing protocol. They also proposed BGP security countermeasures like the use of encryption in BGP link and using sequences and time-stamps in BGP messages to avoid replay attacks [48].

Stability

Route flapping is the state of a BGP speaker in which it demonstrates a pattern of repeated withdrawals and announcement of the same route to a prefix. Furthermore, it is a pathological state in the dynamic network that can lead to instabilities in the network. This behavior can be attributed to hardware failure, configuration error or an unstable link state and may have a far-reaching impact on the stability of the Internet.

A feature known as route flap damping is built into many BGP implementations in an attempt to counteract the effects of route flapping [19]. Without this feature, the excessive activity can cause a heavy processing load on routers and in turn delay updates on other routes, thus affecting the overall routing system stability.

With damping, a route's flapping is exponentially decayed. At the first instance when a route becomes unavailable and quickly reappears, damping does not take effect, to maintain the standard fail-over mechanism of BGP. At the second occurrence, BGP avoids that prefix for a certain length of time, and exponentially times out subsequent occurrences. After the abnormalities have stopped and a suitable length of time has passed for the offending route, prefixes can be reinstated, and its delay timing is reset.

Route leak is a situation where an AS Path violates the valley-free rule. The valley-free rule implies that a valid AS path would consist of one or more customer-to-provider links followed by zero or one peer-to-peer link followed by one or more provider-to-customer links [25]. However, in the present day internet, inter-AS relationships have become more complicated due to the emergence of sibling-to-sibling relations, and it becomes difficult to determine if an observed route-leak is valid. Route-leaks have the potential to destabilize the Internet in cases where a smaller AS leaks the path to a larger AS and then fails to handle enormous traffic redirected towards it destined to the prefixes of larger AS.

BGP and attack objectives

BGP is the de-facto routing protocol of the Internet and is neither secure nor stable and if appropriately misused, can help gain several attack objectives. In this next section, attack objectives which can be achieved by exploiting vulnerabilities of the BGP are discussed. It is assumed that the adversary has successfully compromised one or more BGP speakers or operates an AS with malicious intentions [45].

1. ***Blackholing:*** occurs when a prefix becomes unreachable from large parts of the Internet. One one hand, blackhole routing is done intentionally to enforce private and unallocated IP-ranges and counteract DDoS attacks, and on the other hand, it is done maliciously to attract traffic for specific prefix and then drop it [45].
2. ***Redirection:*** occurs when traffic for a specific prefix is forced to take a different path and reach the incorrect and compromised destination with the intentions to either steal confidential information or generate congestion (through enormous traffic redirection) on the network [45].
3. ***Subversion:*** is a particular case of redirection in which traffic is redirected to compromised routers with the intentions of eavesdropping or modifying the data and then route traffic to the correct destination [45].
4. ***Instability:*** in inter-AS routing can be caused by successive announcements and withdrawals for the same prefix and triggers route-dampening in upstream providers causing connection outages. Alternatively, this generates an enormous amount of BGP traffic and causes longer convergence delays [45].

Mechanisms to exploit BGP security weaknesses

After discussing the attack objectives which can be successfully achieved by exploiting security vulnerabilities of the BGP, let us take a look at the mechanisms in which the same can be performed.

1. ***Prefix hijack*** In this attack, the attacker AS claims to be the origin of the prefix which is owned and announced by the victim AS. Due to a shorter path and other parameters configured in the BGP decision process, the path announced by the attacker is accepted by some ASes. To rely on the shortest path only is leaving things to chances, the impact of the attack can be intensified when combined with the denial-of-service attack on the victim AS and virtually disconnect the victim from the network. Once the victim is unreachable from its neighboring provider and peering ASes, the likelihood of other ASes accepting the path announced by the attacker increases.
2. ***Sub-prefix hijack:*** This is a more common form of attack when compared with Prefix Hijack. BGP path selection process requires routers to select the more specific path whenever available, and the adversary exploits this feature of the BGP to gain control of the victim traffic.
3. ***Path hijack or man-in-the-middle attack:*** In this attack, the attacker AS aims to gain access to the network traffic which is bound to the victim AS by inserting its AS in the legit path to the victim AS amounting to AS path forgery [40].

Proposed architectural improvements for BGP

The existence of vulnerabilities in the protocol motivated researchers to develop a secure protocol or identify ways to integrate security into the protocol. In Secure Border Gateway Protocol (S-BGP), Kent et al. proposed a protocol enhancement. S-BGP consists of four major elements:

1. A public key infrastructure (PKI) that represents the ownership and delegation of address prefixes and AS numbers.
2. Address attestations that the owner of a prefix uses to authorize an AS to originate routes to the prefix.
3. Route attestations that an AS creates to authorize a neighbor to advertise prefixes.
4. IPSec for point-to-point security of BGP traffic transmitted between routers

Another attempt to invent a secure protocol was made by Brian Weis (Cisco Systems). The soBGP (Secure Origin BGP) targets the need to verify the validity of an advertised prefix and advertised path before the updates are forwarded [18].

Karlin et al. introduced another version of routing protocol and called it *Pretty Good BGP* (PGBGP) and compared its usefulness to be at par with soBGP, by giving theoretical proofs for the same. In this work, Karlin et al. also quantified the impact that known exploits can have on the Internet. Further, it identifies its minimum deployment requirement for the effectiveness of the protocol. BGPSec is one of the most popular and most widely researched variant of enhanced routing protocol schemes and is an extension of BGP protocol itself [35]. In this extension path attribute in UPDATE message is cryptographically signed to prevent any tampering efforts.

Mechanisms to prevent hijacks

AS Hijacks are performed to maliciously gain control of Internet traffic for prefixes owned by the victim AS this attack objective is successfully achieved by advertising a more favorable route originating from the adversarial AS. The mechanism to prevent such incidents is done by merely establishing the authenticity of the announcement by establishing the prefix-AS association. Under normal operating conditions, BGP data can be used to extract this information. Other sources of prefix-AS association are whois database [14] and routing registries which allocate prefixes and ASes to the enterprises.

Alternatively, a more secure way of obtaining prefix-AS association is by creating Route Origin Authorization (ROA), which is merely a cryptographically signed information regarding the prefix-AS association and states that an AS is authorized to announce specific prefix. An AS before accepting a route advertisement can validate the ROA and reject in case of any suspicion and is the mechanism of RPKI (Resource Public Key Infrastructure) method of preventing hijacks. Not many ASes are accepting this solution to avoid any unforeseen reachability issues.

Another method of preventing hijacks is based on the credibility of the AS, which is either claiming to own specific prefix or to have a more favorable path through it to the respective prefix. The credibility of ASes is maintained based on their history of malicious behavior [20]. In case an AS behavior is classified as malicious credibility factor drops and other ASes can be suspicious about the future behavior of respective AS.

3.2. Inter-AS route monitoring infrastructure

RIPE RIS & Oregon RouteViews

RIPE Routing Information Service (RIS) [9] and Oregon's RouteViews [11] are public systems which host a network of route collectors. The route collectors are spread across the globe and capture BGP traffic from selected ASes in MRT format [17]. The BGP data captured is used for academic and industrial purposes.

Route collector

A Route Collector or simply collector is the point of peering for the ASes who agree to share their BGP data with collector systems [9, 11]. A BGP session is configured manually at the peering AS to establish the peering

session with a collector. The collector does not participate in global routing, but merely collect the BGP data provided by their peers.

Currently there are 23 and 18 collector nodes of RouteViews [11] and RIPE RIS [9] respectively. These collector nodes are spread geographically and collect BGP traffic generated by the collector peers (monitor ASes) or vantage points. Roughan et al. have concluded that BGP data is not ideal for inferring or mapping the AS-level connectivity of the Internet and that the purpose of BGP is to allow ASes to express and realize routing policies without revealing the AS-internal features.

Vantage points or the monitor ASes

Vantage Point (VP) is an AS which establishes a BGP session with one or more BGP collector nodes. By doing this, VP shares its BGP messages which flow out from its system. Thus the role of a VP is to provide the collector with its RIB. A VP is classified into the following types depending upon the type of information exchanged with a collector :

1. ***Full view:*** These VPs provide entire RIB and incremental changes (UPDATES) in the RIB at a predetermined time interval.
2. ***IPv4:*** These VPs report the incremental changes in the RIB for IPv4 prefixes at a predetermined time interval.
3. ***IPv6:*** These VPs report the incremental changes in the RIB for IPv6 prefixes at a predetermined time interval.

BGP data formats

Multi-threaded Routing Toolkit (MRT) is a format in which the BGP data is stored. [17] describes the specification of the file in which the BGP data is stored. The MRT files are generated at predetermined time intervals, and BGP data is stored differently for RIBS and UPDATES. ***RIBS*** file contains complete routing information or BGP table of VPs and are collected every 120 minutes. The ***UPDATES*** file contains incremental changes in the routing information of VPs and is collected every five minutes.

BGPStream

BGPStream is an open-source framework for live and historical BGP data analysis. It offers a command line interface as well as C++ and python API for processing BGP data collected by RouteViews and RIS data archives. This tool is equipped with capabilities to parse MRT format which is used to store the BGP data.

BGPMon and BGPStream.com

BGPMon and BGPStream.com are commercially available services that provide real-time network health statistics to subscribed customers. These systems have their monitoring network, and not much information regarding their monitoring sources is available in the public domain.

3.3. Internet topology generator: AStop [43]

Several studies have shown that the distribution of the degree of an AS follows power-laws and that its cluster coefficient is higher than one found in random networks [43]. [43] takes into account the power-law behavior of the AS-level Internet and also ensures that the peer-to-peer and provider-to-customer links are labeled at the time of creation. [43] highlights that Internet topology generators such as GE (growing exponent), random, Inet and Barabasi-BA fail to capture the real Internet like properties [43]. The conclusion was based on the following metrics:

1. ***Maximum degree:*** Maximum degree of the graph is defined as the degree of the node with the maximum number of neighbors. On the Internet, maximum degree ASes are found in tier-2 ASes.

2. ***Number of ASes with degree = 1:*** In the AS-level graph of the Internet, there are ASes at the edge of the graph such that they are connected to the core with only one edge. Such ASes are in the lowest tier of the network.

3. ***Complementary cumulative density function (CCDF):*** The CCDF is defined as $F_d = Prob(D \geq d) = \sum f_i$, for $d \leq i \leq inf$, where D is a random variable that indicates the number of incident neighbors upon an AS [43].

The AS relationships are an essential property of the AS graph and impact the BGP decision process. Thus, influence the inter-AS routes chosen between a source-destination pair. To generate a graph and then assign AS-relationship to every link may lead to inconsistent relationships and the parts of the graph are likely to become unreachable due to several inter-AS routes violating the valley-free rule [25] and may even cause routing loops in the graph. A routing loop is an AS-graph inconsistency, where a cycle of provider-to-customer is formed. In figure 3.1a, AS1 is a provider of AS2, AS2 is a provider of AS3 and AS3 is a provider of AS1 and is impossible because AS1 can never be a customer of AS3 given that AS3 is a customer of AS2. Thus, AS relationships which lead to the existence of routing loops are inconsistent. In figure 3.1b, if the peer-to-peer link between AS2 and AS5 is removed (assumed not created by topology generator), AS2 cannot reach AS4 since the only possible route AS2-AS3-AS1-AS4 is going to violate the valley-free rule. Similarly, AS5 cannot reach AS3 since AS path AS5-AS1-AS3 would violate the valley-free rule.

(a) Routing loop in AS-graph

(b) Violation of valley-free rule

Figure 3.1: Consequences of inconsistent relationship assignment.

The AS relationships are inferred using the BGP data and to simulate the BGP data in a toy graph, AS relationships are essential. Therefore, a topology generator must assign a relationship to every link as soon as it is created based on the constraints under which the link is selected. [43] based topology generator assigns AS relationships at the time of link creation. Therefore, this model seems promising for the evaluation of the monitor placement schemes.

Algorithm

The algorithm proposed in [43] is a two phase algorithm and requires following inputs:

1. N_o**:** The number of ASes acting as the core of the Internet.
2. N_c**:** The number of ASes acting as the edge-customers of the Internet.
3. L_o**:** The number of links that are used to interconnect the core.
4. L_c**:** The number of links that are used to attach customer ASes to the core.
5. α***-exponent:*** which controls the slope among scales of axis X.
6. β **:** represents the maximum degree an AS-core can have.
7. ***X***: represents the maximum degree an AS-client can have.

In the first phase core of the Internet model is generated with peer-to-peer or sibling-to-sibling links and in the second phase customers are added to the core.

1. ***Phase 1- building the core:*** In this phase N_o ASes are selected to act as the core. These N_o ASes are interconnected using L_o number of links. For each AS_i forming the core, an available degree D_i is assigned to each AS_i in accordance with power-law,

$$D_i = \beta x^{-\alpha} \tag{3.1}$$

where, x is a random number such that $(1 < x < N_0)$ and β represents the maximum degree an AS can have in the model. Moreover, the algorithm requires that the following condition be satisfied,

$$\sum_{i=1}^{N_0} D_i \geq 2L \tag{3.2}$$

where L is total number of links in the graph and is given by,

$$L = L_0 + L_c \tag{3.3}$$

Where, L_c is the number of links that customer ASes use to connect to the core. Then for all links in a preferential manner two ASes: AS_i and AS_j are chosen. The link L_i is then assigned to (AS_i, AS_j) if their respective available degrees are greater than 0, and then their available degrees are reduced by 1. The probability of selection for all ASes with a degree more than 0 is uniform and is updated in every iteration. Then the probability that a given link's endpoint attaches to a particular existing AS is given by the preferential attachment equation,

$$P(d_i, t + \Delta t) = \frac{d_i(t)}{\sum_{j=1}^{N_0} d_j t} \tag{3.4}$$

Duplicated links are not allowed in both phases.

2. ***Phase 2- attaching customer ASes:*** In this phase, N_c customers are attached to the core using L_c links. Initially, one link is assigned to each customer AS so the minimum degree every AS has is 1, and every customer AS gets access to the core. The remaining links, $R = (L_c - N_c)$ are assigned to customer ASes using the following: for 1 to R a client AS(i) is randomly chosen, link $L_c(i)$ is assigned to $AS(i)$ only if its degree is less than X. X is the maximum degree a client AS can have. Thus, the links that a customer AS can have is a random variable $L_c(x)$ delimited by $1 < L_c(x) < X$ and represents its degree. Then the N_c customer ASes chose $L_c(X)$ core ASes according to equation 3.4. The link $L_c(i)$ is assigned to a core AS only if the remaining degree of the AS is greater than 0 (the set of ASes with free credits at iteration t). Then the degree of this core AS is reduced by 1.

3.4. The BGP simulator

Kastelein has developed a BGP simulator which is used to generate the routing information for the Internet-like graphs to study the behavior of inter-AS routing, and evaluation of monitor selection schemes [31]. Routing information or the BGP data is essential to establish the effectiveness of the monitoring network as the objective of monitoring is to observe as many links as possible, and by using RIB from simulations, the link-coverage for a monitor set is estimated. The Internet-like graphs are provided to the simulator, and the simulator generates the routing information for each graph. The results from the simulation are stored in the routing table object and consumed for evaluation.

The core API of the simulator is accessible through interfaces. The simulation process consists of three stages. First, initialize the AS-graph and is done through a graph file. The graph file is a representation of the AS graph which contains the edges along with the relationship assigned to the edge in each line of the file

Start AS	End AS	Relationship
AS93	AS145	customer-to-provider
AS457	AS138	customer-to-provider
AS20	AS140	customer-to-provider
AS21	AS121	customer-to-provider

(a) Graph file for BGP simulator

Prefix	Origin AS
151.164.214.0/24	AS93
37.33.45.0/24	AS457
231.13.112.0/24	AS20
206.223.168.0/24	AS21

(b) Prefix-association file.

Figure 3.2: Sample files provided to the simulator.

(see figure 3.2a). The first stage is deemed complete once the AS-graph is initialized. The second stage creates announcements and is done through a prefix-AS association file, which contains prefix-origin AS association details (see figure 3.2b). The create announcement stage acts as a trigger for UPDATE messages in the simulator. The ASes which are originating a prefix would announce the assigned prefix in the third stage. And, the third stage is to iterate announcements. In this stage, the announcements created in the second stage are propagated across the graph. The simulations process is similar to the one discussed in section 2.3.

Let us look at the first line in figure 3.2a. Here AS93 and AS145 are the two end-points of the link and customer-to-provider is the relationship that is assigned to the link, which indicates that AS93 is a customer of AS145 and conversely, AS145 is a provider of AS93. Figure 3.2b shows a prefix association file using which announcements are created in the BGP simulator. The first line of figure 3.2b indicates that the prefix 80.19.126.0/24 is assigned to AS93. Using such files as input the simulator's stage one and two are completed, and stage 3 is executed to obtain routing information object.

3.5. The monitor selection schemes [52]

[52] studied four deployment schemes for inter-AS route monitors. These schemes are based on the network properties like node degree and link coverage and none of the schemes take into account the predictable nature of BGP information flow as we shall see in the next chapter. The schemes studied by [52] are discussed in detail in the following subsections.

3.5.1. Address based monitor selection scheme

In this scheme, monitor selection is made based on the size of address space which is within the AS's customer network. Top-level ASes tend to aggregate address space of their customer network into larger prefixes to reduce the number of UPDATE messages. During aggregation information regarding prefix's true origin is lost, because the AS-path is also updated in such UPDATE message to show the AS that is aggregating the prefixes, as the true origin of the aggregated prefix and creates a bias for tier-1 ASes and makes monitor selection accordingly. Monitoring multiple tier-1 ASes would lead to redundancy in the BGP data as all tier-1 ASes form a clique. Clique is a term used to describe a complete sub-graph of a graph. [22] identified that the sources of redundant information in BGP data are adjacent monitors. Hence, this is not an ideal monitor selection scheme, and therefore, this scheme has not been considered for further evaluation.

3.5.2. Random based monitor selection scheme

In this scheme, each AS is assigned the equal probability of selection into the monitor set, and k-number of ASes are picked with a uniform probability distribution to be selected as a monitoring node. This scheme can be used to provide a monitoring network of definite size. This can be considered as the most basic form of monitor selection and is used by public route collector system [9]. There are no nomination or selection criteria for an AS to participate in the RIPE project. An AS can create a BGP peering session by sharing the technical details of its BGP speaker.

The main issue with this algorithm is the randomness while selecting the monitors. No mechanism could be used to bias the probability of an AS during selection and could lead to the selection of adjacent monitors, and by selecting adjacent monitors, the monitor set would end up collecting redundant BGP data. Moreover, there is no guarantee that the algorithm would generate a monitor set such that monitors are distributed across all the tiers of the Internet. However, the advantage of this scheme is the speed in which monitor set can be determined, and it may by chance end up selecting the best monitor set possible. The pseudo-code

for this monitor placement scheme is presented in algorithm 3.

Algorithm

The algorithm to determine the monitor set using the random based monitor selection is relatively simple. Each AS in the graph is assigned a probability of selection which is equal to $\frac{1}{N}$ where N is the total number of ASes in the graph, and then k elements are picked from the list of ASes using the assigned probabilities. As the probabilities are equal for each AS, any AS could be selected as the monitor.

Figure 3.3 shows two combinations of monitor set of size three (indicated in red and blue respectively). There are seven ASes and therefore the probability of selection of each AS is $\frac{1}{7}$. The total number of combinations of monitors are $\binom{7}{3}$ which is 35.

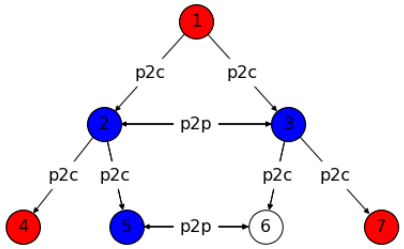


Figure 3.3: Random-based monitor selection

3.5.3. Degree-based monitor selection scheme

In this monitor placement scheme, AS's neighbor degree plays an important role. An AS is surrounded by ASes which may belong to any of these categories; provider, customer, sibling or peer. The total number of AS's neighbors are the determining factor if an AS is selected as a monitor or not. The ASes with the higher neighbor degree would be selected as monitoring AS. This scheme has an advantage over the random-based monitor selection scheme because an AS with more neighbors would receive UPDATES from more ASes and therefore is more likely to have better visibility of the Internet.

The disadvantage of this scheme is that it would become biased towards tier-2 ASes. An AS in tier-2 would have a larger subset of ASes to peer with and almost an equally larger subset of ASes acting as its customer. This is because tier-2 ASes peer with as many ASes as possible to reduce dependency on their tier-1 provider AS. Additionally, tier-3 ASes which operate as ISPs in local regions would become their customers due to their wide geographic presence and reduces the possibility of maintaining a monitoring network distributed across all tiers of the Internet. Additionally, tier-2 ASes peer heavily and thus this scheme is more likely to select adjacent ASes as BGP monitors, which would add on to the redundancy in the BGP data.

Tier-1 ASes are characterized by a low-medium number of peers and medium-high number of customers. This is because tier-1 ASes are not easily accessible to tier-3 and tier-4 ASes which are the majority of the ASes on the Internet. The number of peers with which a tier-1 AS peers is also limited due to a meager number of tier-1 providers. Although there is no formal report on the tier-wise classification of ASes, they are estimated to be limited to 19 ASes (section 4.3). Tier-2 ASes become the customer of tier-1 ASes, and their number is also not very sizable when compared to the tier-3 and tier-4 ASes.

Tier-3 and Tier-4 ASes are less likely to peer due to technical challenges. Additionally, their limited influence over the networks, render these ASes less lucrative for peering. Moreover, ASes in these tiers are either local ISPs which provide the Internet access to end-users like us or are networks managed by large/small enterprises and hence cannot become provider ASes of large parts of the Internet. Hence ASes in these tiers are characterized by a low number of customers and non-existent number of peers.

The degree based selection scheme can be implemented a monitor set of k-size by ranking all ASes based on their neighbor degree and selecting top-k ASes. Algorithm 5 provides the pseudo-code for k-size degree-based monitor selection scheme. Algorithm 4 shows the algorithm to generate a degree map. A degree map

is a representation of AS neighbor count based on the relationship count. If an AS has two providers, three customers, one peer, and one sibling, then the degree map would be {AS: p2p:1, s2s:1, c2p:2, p2c:3}.

Algorithm

The algorithm for degree based monitor selection assigns a rank to each AS in the graph based on their degree (total number of neighbors) and then selects top-k ASes in the monitor set. Figure 3.4 shows three possible combinations of monitor set determined using degree based monitor selection. Each combination includes the ASes shown in blue. ASes shown in blue has a degree of four which makes them highest ranking ASes, followed by ASes shown in red with degree two. Therefore, if a monitor set of size 3 is to be selected, the algorithm can produce one of these combinations.

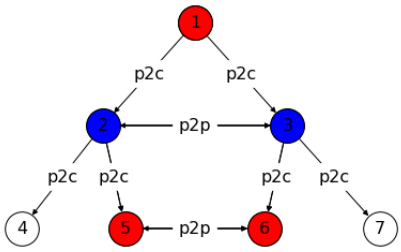


Figure 3.4: Degree-based monitor selection

3.5.4. Greedy-link based monitor selection scheme

[52] proposed greedy-link based monitor selection scheme which reduces the redundancy in the existing monitors. The scheme is based on a set-cover problem heuristic and aims to determine a set of monitors to maximize visibility. The advantage of this scheme over degree-based and random monitor selection schemes is that this monitor selection can ensure that 100% visibility is achieved. However, beyond 80% addition of a new monitor does not improve the visibility significantly. Each AS can see at least 45% of the visible Internet. The term visible is used to remind that the extent to which the Internet is visible is not yet determined.

Although the scheme is the most beneficial of all the schemes presented earlier, suffers a major disadvantage. This scheme requires information regarding inter-AS routes from all ASes to begin monitor selection. This is unavailable beforehand, and without this information, the algorithm is rendered less useful for practical purposes. Moreover, in case the existing BGP data is used to determine the monitor placement- the algorithm would end up reducing the redundancy in the BGP data but fail to identify new monitor set.

This scheme is implemented by providing a set of visible links for each AS and selecting the monitoring ASes by an iterative process of ranking the ASes based on the maximum number of unobserved links and selecting the AS with the maximum number of unobserved links. The greedy-link based selection scheme can be implemented in two ways:

1. ***k*-size based:** In k-size based scheme, a link-coverage set map is used and sorted based on the maximum number of unobserved links. Then algorithm selects highest ranked AS and calculates the new link-coverage set map for all ASes. The algorithm terminates when k-ASes are in the monitoring set. Here k- represents the size of the monitoring network.

2. ***k*%- coverage based:** In k%- coverage based scheme, the monitor selection process terminates where the link-coverage of k%- is obtained. The k-coverage is forced to 100% by using the threshold of 1. As long as there are any new links observed the algorithm must continue. The ties are always broken randomly.

The monitor set produced by the greedy-link monitor selection scheme cannot be considered optimal. However, it does provide an insight into the approximate size of the monitoring network that is required to observe 100% AS-links of the known Internet. The optimization of greedy-link based monitor selection scheme

can be achieved by solving an objective function to maximize link-coverage for a given size of the monitoring network. The solver can explore all possible combinations of ASes of the desired size and provide the one with the best link-coverage. However, the listing of all combinations of *k*-size for a total of *n*-ASes would be computationally intensive. The Internet presently comprises of approximately 62775 ASes and to select k-ASes as monitors would require listing of ${}^{62775}C_k$ possible combinations and this would be an infinitely high value. To reduce the size of the set holding the possible combinations of *k*-size for a total of *62775*-ASes, a sample-based approach is adopted and would increase the options of monitoring sets explored. However, this will also not produce an optimal set as the optimal set may lie in the region of unexplored combinations. Alternatively, *size-n* of the ASes can be reduced by collapsing ASes with no peering and customer ASes into their provider AS, starting from the edge ASes. This is an appropriate step because the inter-AS routes chosen by such ASes would be governed by their provider AS and therefore, by collapsing such ASes into their provider AS can reduce the size of the set of ASes eligible for monitor selection. For the Internet, the size of AS-graph (in this thesis is always measured in terms of the number of ASes) is reduced from 62775 ASes to 20000 ASes, which is again very high. The number of combinations grow exponentially and ${}^{100}C_{20}$ is computed to be 535,983,370,403,809,591,296 using [6] and verifying monitor performance for 535,983,370,403,809,591,296 combinations would require 1.6995921^{13} years if monitor performance of one monitor set is evaluated every second. Therefore, evaluating monitor performance of ${}^{20000}C_{20}$ combinations would be unfeasible as the number of combinations would be a number larger than 535,983,370,403,809,591,296.

Algorithm

The greedy-link based monitor selection scheme requires BGP tables of the ASes from which the monitor set is to be determined and is an approximation solution of the set-cover problem. A set of visible links v_i for each AS is created. Next, a set of visible links V is created in which the links visible using the ASes that have been selected as monitors is created, this set is empty at the start of the algorithm as no monitors are selected. The number of links in the visible link set for each AS is used to rank the ASes and the AS with the highest rank is selected and added to the monitor set, and the links of this ASes are added into the set of visible links of the monitor set. The visible link set is updated to remove links which are already present in the visible link set of the monitors ($v_i = v_i \setminus V$) and the ASes are ranked again based on updated visible link set for each AS. The process continues until no more ASes could be added into the monitor set or *k*- number of monitors are added into the monitor set.

Algorithm 6 provides the pseudo-code for k-size greedy-link based monitor selection scheme.

4

Methodology

In this chapter, we shall analyze the impact of announcements originating from frequently occurring graph motifs on the Internet. This is done by simulating the BGP traffic using the simulator and studying the routing information available at each AS. Kastelein developed a BGP simulator which is used to perform BGP simulations [31]. Next, we discuss the shortfalls of the topology generator proposed by [43] and propose a topology generator based on the well-known behavior of the ASes. The topology generator based on the behavioral model is shown to obey the power-law which is a necessary condition for the Internet topology generators. Finally, we discuss the peering-degree based monitor selection scheme which is the research outcome of this thesis.

4.1. Impact of graph motifs on the Internet

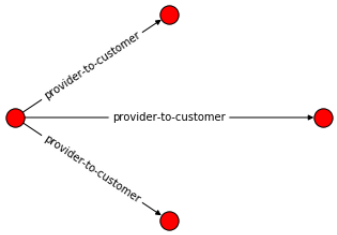
The primary objective of the monitoring network is to obtain higher link-coverage and capture announcements as close as possible to its originating AS. Thus, understanding how an AS selects inter-AS routes is essential to ensure effective monitor selection. The AS level Internet is such that every network is reachable from each AS. This means that an AS can either directly through its customer network, its provider AS or its peer network can access the Internet. This knowledge is exploited to represent the Internet with a two AS representation where one AS acts as a global provider and the other AS is the customer of this AS. To study the impact of announcements originating within a graph motif, the graph motif is connected to this representational system of the Internet through a peer-to-peer link and in a few cases through a provider-to-customer link.

The life-cycle of an announcement is such that it originates within the customer cone of one AS and then traverses up towards the core ASes and then finally flows down to the ASes within the customer cones of different ASes before terminating at the ASes with no customers. Peering links play an important role in manipulating the global routing and act as a by-pass between the two higher tier ASes, providing an alternate path to its customers to reach various prefixes.

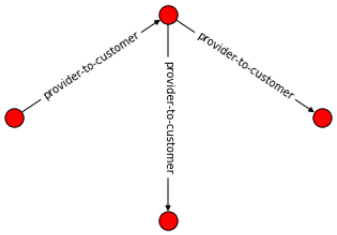
A provider AS ensures global connectivity for the ASes which lie within its customer cone and is achieved by connecting through peer-to-peer links with nearly equal ASes or customer-to-provider links with larger ASes. The inter-AS route selection is influenced by the local preferences assigned to each neighboring AS, which is determined by the business relations that exists with a neighbor. For instance, a peer AS (an AS connected with peer-to-peer link) would have a higher local preference than provider AS and the customer AS would have the highest local preference. In the case where an AS has multiple neighbors with the same relationship, different preferences would be assigned to each neighbor to ensure that network's routing policies are implemented.

BGP data of July 4^{th} 2018 is used to build the Internet topology by extracting the AS paths announced on that day. These AS-paths are then used to infer AS relations by using the heuristic proposed in [25]. Once the AS-relations are available, graph motif exploration is performed using the algorithm 2, by selecting the ASes in the immediate neighborhood of each AS and listing all possible combinations of size 4 and collecting the graph motifs found for different combinations. 134 graph motifs of size 4 were discovered for the

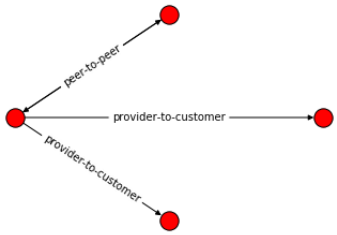
respective AS topology. The frequency of occurrence of each motif varied and four graph motifs with the frequency of occurrence more than 3000 are shown in figure 4.1. In figure 4.1 the direction of the arrowhead in a provider-to-customer link indicates customer and a peer or a sibling in a peer-to-peer or a sibling-to-sibling link respectively.



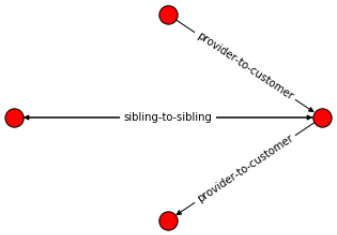
(a) Frequency of occurrence 62907.



(b) Frequency of occurrence 62823.



(c) Frequency of occurrence 3389.



(d) Frequency of occurrence 3302.

Figure 4.1: Graph motifs from the Internet of size 4.

The process described in section 3.4 is used to simulate BGP traffic on each sub-graph. As discussed earlier, a graph file and a prefix association file are provided to the simulator which are used by the simulator to initialize the AS graph and generate the announcements using the prefix association file. The resulting routing information obtained when the simulation is complete is used to discuss the impact of announcements originating under different scenarios.

The routes adopted by edge ASes are restricted by their provider ASes and announcing prefixes from edge ASes would trigger the ASes close to the core to make their path selection. Therefore, instead of announcing prefixes from all ASes, the prefixes are announced from all edge networks (ASes that do not have customers) to reduce the simulation time.

Sub-graph type 1

The thesis is promoting the use of peering degree instead of the degree to determine the BGP monitors. We begin by presenting the most basic graph motif of the Internet, that is, provider-to-customer chains. The provider-to-customer chains exist on the Internet, such that every AS is the provider of exactly one AS and the customer of exactly one AS. This network motif is hard to find using the algorithm 2 as the algorithm explores graph motifs in the immediate neighborhood of an AS. However, finding this motif is relatively easy and by representing every link of the AS paths that are observed in the BGP data by its relationship, this linear chain of provider-to-customer is seen quite often.

In figure 4.2 and 4.3, AS1 is the provider of AS2, AS2 is the provider of AS3, AS3 is the provider of AS4, and AS4 is the provider of AS5. AS6 is provider of AS1 in figure 4.2 and AS6 is peer of AS1 in figure 4.3. The variation in figure 4.2 and 4.3 is due to the fact that the top-level AS (AS1 in this case) of the chain could connect to the rest of the Internet either through a peering link or a customer link.

As explained in section 3.4 the graph files are required by the simulator to initialize the AS graph. The graph files for respective sub-graphs is shown along with their topology. The prefixes are announced from

AS5 in both cases and the resulting routing information objects are presented in table 4.1 and 4.2.

(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS6	AS1	provider-to-customer
AS1	AS2	provider-to-customer
AS2	AS3	provider-to-customer
AS3	AS4	provider-to-customer
AS4	AS5	provider-to-customer

(b) Graph file.

Figure 4.2: Sub-graph type 1a.

(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS6	peer-to-peer
AS1	AS2	provider-to-customer
AS2	AS3	provider-to-customer
AS3	AS4	provider-to-customer
AS4	AS5	provider-to-customer

(b) Graph file.

Figure 4.3: Sub-graph type 1b.

AS	Number of paths	Number of links	Paths
1	1	4	1 2 3 4 5
3	1	2	3 4 5
2	1	3	2 3 4 5
5	0	0	
4	1	1	4 5
6	1	5	6 1 2 3 4 5

Table 4.1: Sub-graph type 1a- total links 5.

AS	Number of paths	Number of links	Paths
1	1	4	1 2 3 4 5
3	1	2	3 4 5
2	1	3	2 3 4 5
5	0	0	
4	1	1	4 5
6	1	5	6 1 2 3 4 5

Table 4.2: Sub-graph type 1b- total links 5.

From table 4.1 and 4.2 it can be concluded, that ASes close to the top-level provider can discover more links as compared to ASes close to the edge (where the prefix is originated). Additionally, AS6 is the provider of

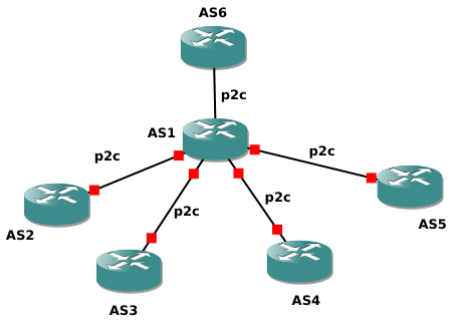
AS1, and it makes AS6 in figure 4.2a the owner of the entire provider-to-customer chain and thus can discover an additional link AS6-AS1. The sub-graph type 1 is the simplest structure of the Internet and explains how the routes are propagated in a linear chain of provider-to-customer links.

The inter-AS routes used by customer ASes in the provider-to-customer chain is governed by routing decisions of the top-level provider or the AS which owns the customer chain because the top-level AS is the only direct connection to the Internet. Additionally, the top-level provider is bound to announce its customer paths to the rest of its neighbors. Therefore, in a provider-to-customer chain, there would be no difference in the link-coverage if the monitoring is done at an AS other than the top-level AS. However, a top-level AS of such graph motif is solely responsible for routing on the Internet and thus, is a suitable choice for monitoring network.

Sub-graph type 2

Another simple yet frequently occurring graph motif is a provider AS with multiple customers (for example see figure D.1a). An AS may act as the provider for multiple ASes as shown in figure 4.4 and 4.5. In the remaining sub-graphs, AS6 and AS5 are representing the Internet and graph motif is represented by AS1, AS2, AS3, and AS4. AS2, AS3, and AS4 are customers of AS1. AS6 is provider of AS1 and AS5 in figure 4.4, and AS6 is peer of AS1 and provider of AS5 in figure 4.5.

Such sub-graphs are frequently occurring when multiple tier-4 ASes prefer a single tier-3 provider AS due to its local influence or multiple tier-3 ASes consume services from a single tier-2 provider or multiple tier-2 ASes act as the customers of a single tier-1 AS. The BGP traffic simulation is performed for the above mentioned sub-graphs using the graph files as shown in figure 4.4 and 4.5. The prefixes are announced from AS2, AS3, AS4, AS5, and AS6, and the resulting routing information objects are presented in table 4.3 and 4.4.

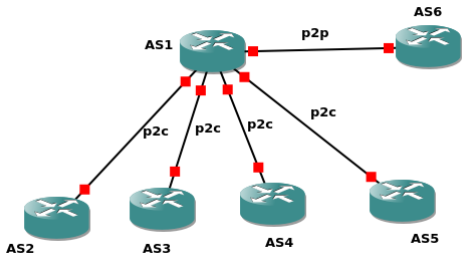


(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS1	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS1	AS5	provider-to-customer
AS6	AS1	provider-to-customer

(b) Graph file.

Figure 4.4: Sub-graph type 2a.



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS1	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS1	AS5	provider-to-customer
AS1	AS6	peer-to-peer

(b) Graph file.

Figure 4.5: Sub-graph type 2b.

From table 4.3 and 4.4 it could be concluded that AS1 maintains five paths. These are corresponding to

AS	Number of paths	Number of links	Paths
1	5	5	1 2, 1 3, 1 6, 1 4, 1 5
3	4	5	3 1 5, 3 1 6, 3 1 4, 3 1 2
2	4	5	2 1 6, 2 1 4, 2 1 3, 2 1 5
5	4	5	5 1 3, 5 1 2, 5 1 4, 5 1 6
4	4	5	4 1 2, 4 1 3, 4 1 6, 4 1 5
6	4	5	6 1 2, 6 1 3, 6 1 5, 6 1 4

Table 4.3: Sub-graph type 2a- total links 5.

AS	Number of paths	Number of links	Paths
1	5	5	1 2, 1 3, 1 6, 1 4, 1 5
3	4	5	3 1 5, 3 1 6, 3 1 4, 3 1 2
2	4	5	2 1 6, 2 1 4, 2 1 3, 2 1 5
5	4	5	5 1 3, 5 1 2, 5 1 4, 5 1 6
4	4	5	4 1 2, 4 1 3, 4 1 6, 4 1 5
6	4	5	6 1 2, 6 1 3, 6 1 5, 6 1 4

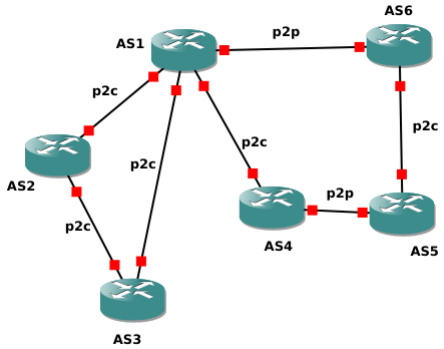
Table 4.4: Sub-graph type 2b- total links 5.

each prefix announced. AS2 through AS6 maintains four paths each for four out of five prefixes announced during the simulation. The fifth path that is not included in the table is the path to the prefix which is originated by them. Additionally, all the ASes discover the equal amount of links and each edge AS possess similar path information which is provided by a common top-level provider AS. Therefore, all ASes can act as a monitoring node in this situation and provide equally good link-coverage.

The sub-graph type 1 and type 2 highlight that in simple sub-graphs it is easier to place a monitor as each AS could provide equally good link-coverage. However, these sub-graphs excluded peer-to-peer relations and multiple provider cases. In the following subsections, the impact of graph motifs with peer-to-peer links would be evaluated.

Sub-graph type 3

An AS may be able to purchase transit from ASes which are on different tiers (figure D.1d). This is done to achieve speed, bandwidth, and cost optimization related objectives. For instance, a tier-3 AS connects with a tier-2 AS and a tier-1 AS. Figure 4.6 depicts such a situation where AS1 is the provider of AS2, AS3, and AS4. Additionally, AS1 is the peer of AS6 whose customer is AS5. AS5 is the peer of AS4, and AS3 is a customer of AS2. The graph file provided to the simulator is shown in figure 4.6b. The prefixes are assigned to AS3, AS4, and AS5, and the resulting routing information object is presented in table 4.5.



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS1	AS3	provider-to-customer
AS2	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS6	AS5	provider-to-customer
AS6	AS1	peer-to-peer
AS5	AS4	peer-to-peer

(b) Graph file.

Figure 4.6: Sub-graph type 3.

AS	Number of paths	Number of links	Paths
1	3	4	1 3, 1 4, 1 6 5
3	2	4	3 1 4, 3 1 6 5
2	3	5	2 1 4, 2 3, 2 1 6 5
5	2	4	5 4, 5 6 1 3
4	2	3	4 5, 4 1 3
6	3	4	6 5, 6 1 3, 6 1 4

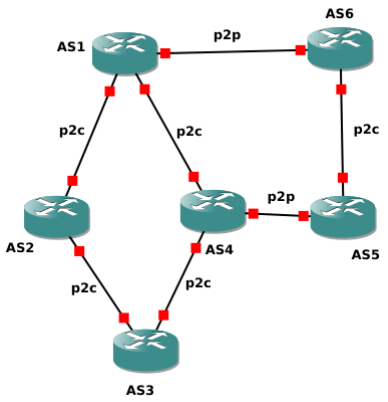
Table 4.5: Sub-graph type 3- total links 7.

From table 4.5, it is evident that AS2 can discover more links than top-level provider AS1. This is because AS1 would reach AS3 directly by provider-to-customer link and links AS1-AS2 and AS2-AS3 may remain undiscovered. However, if the observations are prolonged at AS1, undiscovered provider-to-customer links can become visible should the link AS1-AS3 fail or any traffic engineering is performed by AS3. Link AS4-AS5 would remain undiscovered to all ASes but AS4 and AS5 which are the end-points of the peering link. Thus, it is essential also to monitor AS4 and AS5 to discover the peering link.

ASes with peer-to-peer links use the peer routes to reach several prefixes and hence ignore the provider routes. Thus if such ASes are not monitored, the peering links would remain undiscovered and would impact the link-coverage. In this case, there is only a single peering link, but on the Internet, the peering links are five-sixths of the total links. Therefore, in addition to top-level providers it is essential to monitor either end-points of a peering link to capture peering links.

Sub-graph type 4

An AS may purchase transit from two providers who share a common provider. Consider a sub-graph is shown in figure 4.7 where AS4-AS5 and AS1-AS6 are connected via a peer-to-peer link. AS1 is the provider of AS2 and AS4, AS6 is the provider of AS5, and AS3 is a customer of AS2 and AS4. The graph file provided to the BGP simulator is shown in figure 4.7b. AS3 and AS5 announce the prefixes, and the resulting routing information object is presented in table 4.6.



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS2	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS4	AS3	provider-to-customer
AS6	AS5	provider-to-customer
AS6	AS1	peer-to-peer
AS5	AS4	peer-to-peer

(b) Graph file.

Figure 4.7: Sub-graph type 4.

Table 4.6 suggests that only AS4 and AS5 can capture the peering link. AS1 which is the top provider can see all links except the peering link between AS4 and AS5. AS6 can only see a single path to AS3 because AS1 announced the best path only. AS path "AS1 AS2 AS3" and "AS1 AS4 AS3" are equal in path length which is 3. However, AS1 preferred AS2 over AS4 to route packets towards AS3. In the simulations, this is randomly decided and thus in a different simulation "AS1 AS4 AS3" could be chosen by AS1. This would not change the link-coverage of AS6. AS3, AS4, and AS5 have the worst link-coverage. They only capture two links each. This is because the customer of AS with peer-to-peer link would be forced to reach destinations within their

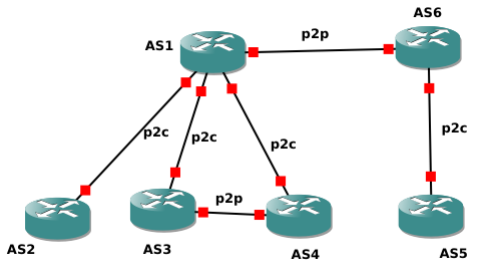
AS	Number of paths	Number of links	Paths
1	3	6	1 4 3, 1 2 3, 1 6 5
3	1	2	3 4 5
2	2	4	2 3, 2 1 6 5
5	1	2	5 4 3
4	2	2	4 5, 4 3
6	2	4	6 5, 6 1 2 3

Table 4.6: Sub-graph type 4- total links 7.

provider's peer-customer cone through peer-to-peer links. The peering complicates the BGP route monitoring when compared with sub-graphs type 1 and 2. AS3 received two routes to AS5, one from AS2 and another from AS4. However, the route provided by AS4 is shorter and hence chosen as the best route by AS3. The peering-degree based monitor selection would be able to produce higher link-coverage as it would capture both provider routes (routes announced by provider) and peer routes (routes announced by peer) as well.

Sub-graph type 5

ASes within the customer cone of a single provider AS can form a peer-to-peer relation. The graph motif shown in figure D.1a is used, and a peering link is added between two ASes within the customer cone. In figure 4.8, AS1 is the provider of AS2, AS3, and AS4, and AS6 is the provider of AS5. AS1 and AS3 are peers of AS6 and AS4 respectively. The graph file used in the simulation for this sub-graph is shown in figure 4.8b and the prefixes are announced from AS2, AS3, AS4, and AS5. The resulting routing information object is presented in table 4.7.



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS1	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS6	AS5	provider-to-customer
AS6	AS1	peer-to-peer
AS3	AS4	peer-to-peer

(b) Graph file.

Figure 4.8: Sub-graph type 5.

AS	Number of paths	Number of links	Paths
1	4	5	1 2, 1 3, 1 4, 1 6 5
3	3	5	3 4, 3 1 2, 3 1 6 5
2	3	5	2 1 6 5, 2 1 4, 2 1 3
5	3	5	5 6 1 4, 5 6 1 3, 5 6 1 2
4	3	5	4 1 6 5, 4 3, 4 1 2
6	4	5	6 1 2, 6 1 3, 6 5, 6 1 4

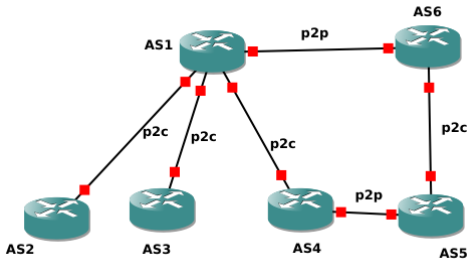
Table 4.7: Sub-graph type 5- total links 6.

Table 4.7 suggests that the peering link AS3-AS4 is invisible to every AS, but AS3 and AS4. AS3 and AS4 use provider routes to reach AS2 and AS5, and peering link to reach AS4 and AS3 respectively. AS2 and AS5 do not have any peers and thus rely on the routes announced by their providers. AS1 does not see the peering link AS3-AS4 since the peer routes are not exported to the provider AS and thus peering link remain undiscovered at every AS but AS3 and AS4. link-coverage offered by each AS is equally good, however to achieve 100% link-

coverage, ASes with peering links require monitoring, and by biasing the monitor selection process based on peering-degree ensures that this requirement is fulfilled.

Sub-graph type 6

An AS within the customer cone of an AS can form a peer-to-peer AS, which belongs to a different AS. To simulate this behavior graph motif shown in figure D.1a is used again, and a peering link is added between the ASes from different customer cones. Figure 4.9 is similar to figure 4.8 except AS4-AS5 are peering instead of AS3-AS4. The graph file used in the simulation is shown in figure 4.9b and the prefixes are announced by AS2, AS3, AS4, and AS5. At the end of simulation, a routing information object is obtained, which is presented in table 4.8.



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS1	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS6	AS5	provider-to-customer
AS6	AS1	peer-to-peer
AS5	AS4	peer-to-peer

(b) Graph file.

Figure 4.9: Sub-graph type 6.

AS	Number of paths	Number of links	Paths
1	4	5	1 2, 1 3, 1 4, 1 6 5
3	3	5	3 1 4, 3 1 2, 3 1 6 5
2	3	5	2 1 4, 2 1 3, 2 1 6 5
5	3	5	5 4, 5 6 1 3, 5 6 1 2
4	3	4	4 1 2, 4 1 3, 4 5
6	4	5	6 1 2, 6 1 3, 6 5, 6 1 4

Table 4.8: Sub-graph type 6- total links 6.

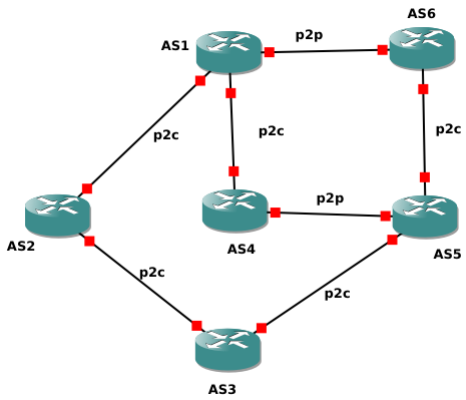
Table 4.8 shows that the peering link between AS4 and AS5 can be discovered by monitoring AS4 or AS5. Remaining ASes can discover five out of six links and, we can observe the importance of ASes with peering links. Therefore, peering-degree based monitor selection could provide an improvement in the link-coverage.

Sub-graph type 7

An AS may buy transit from ASes which belong to different customer cones as shown in figure 4.10a. Here, AS3 is a customer of AS2 and AS5 that belong to different customer cones of AS1 and AS6 respectively. AS4 is the peer of AS5, and customer of AS1 and AS1 and AS6 are peers. Such interconnection may be the result of the business strategy to meet traffic engineering requirements to optimize speed, bandwidth and cost objectives. Moreover, this may even happen for ASes with a diverse geographical presence which can purchase transit services from different providers in different regions.

The graph file used for simulating the BGP traffic in this sub-graph is shown in figure 4.7b and the prefixes are announced from AS3 and AS5. The resulting routing information object generated at the end of the simulation is presented in table 4.9.

Table 4.9 shows that monitoring only the top-level provider with larger customer cone discovers most links. However, to achieve 100% link-coverage, it is essential to capture the peer-to-peer link between AS4



(a) AS-topology of sub-graph.

Start AS	End AS	Relationship
AS1	AS2	provider-to-customer
AS2	AS3	provider-to-customer
AS1	AS4	provider-to-customer
AS5	AS3	provider-to-customer
AS6	AS5	provider-to-customer
AS6	AS1	peer-to-peer
AS5	AS4	peer-to-peer

(b) Graph file.

Figure 4.10: Sub-graph type 7.

AS	Number of paths	Number of links	Paths
1	2	4	1 2 3, 1 6 5
3	1	1	3 5
2	2	4	2 3, 2 1 6 5
5	1	1	5 3
4	2	2	4 5, 4 5 3
6	2	2	6 5, 6 5 3

Table 4.9: Sub-graph type 7- total links 6.

and AS5 and to capture peer-to-peer link AS4, or AS5 needs to be monitored. This selection is possible by peering-degree based monitor selection scheme.

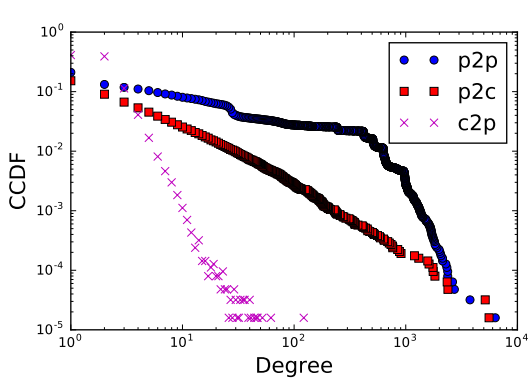
Of the total links present on the Internet, five-sixth links are peering links and as shown in section 4.1 peering links remain undiscovered if monitoring is performed at an AS which does not have a peering link or does not belong to the customer cone of the ASes with peering links. The peering links discovered using the BGP data accounts for only one-fifth of the total peering links and the remaining are inferred using the alternate data sources like ARK data and Looking Glass as indicated in [5]. Therefore, it can be confirmed that peering-degree based monitor selection would achieve higher link-coverage by capturing multiple peer-to-peer links and could be implemented for further evaluation.

4.2. The Internet model

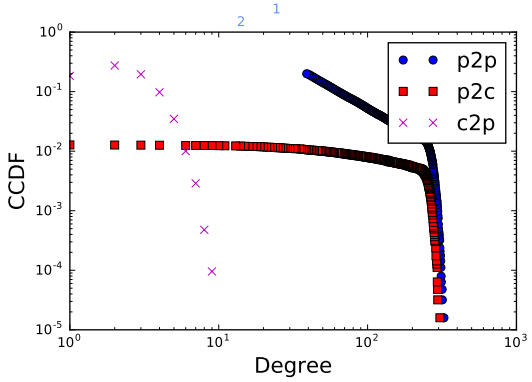
The Internet is a complex network with 62773 nodes, 123606 provider-to-customer links and 568054 peer-to-peer links (as recorded in AS relationship data from 01-October-2018[5]). To perform simulations on a network of such dimensions is a computationally challenging task and would require infinite simulation time. A model of the AS level topology of the Internet is required such that the graph size can be reduced and the BGP traffic simulation could be performed effectively to evaluate the monitor set resulting from different algorithms for monitor placement.

In section 3.3, the model proposed by Nieto-Hipólito et al. is discussed. [43] have highlighted that power-law based graph topology generators are not sufficient to model the internet and proposed an approach to generate the AS level graphs of the Internet that assigns AS relationships at the time of link creation [43]. Additionally, [43] follows the CCDF of the AS topology closely as reported in [43]. The plot of CCDF for a graph generated using [43] and a graph using topology collected by [5] is shown in figure 4.11.

CCDF is a widely accepted measure of the Internet topology. It is defined as the probability of an AS to have a degree greater than the specified degree, i.e. $F_i = P(d \geq D)$ and is plotted on a log-log scale. Figure 4.11a is the log-log plot of CCDF against degree for the AS graph generated using the topology information



(a) CCDF of the Internet as captured in [5] data on 01-10-2018.



(b) CCDF of the topology generated using [43] for α = 0.69, β = 8000, N_o = 12755, N_c= 50218, L_o= 568054 and L_c= 123606.

Figure 4.11: CCDF of AS graph using [5] and [43].

collected by [5] data on 01-10-2018. The plot shows that the probability of an AS to have higher degree reduces exponentially. The legends p2c, p2p, and c2p represents provider-to-customer. Peer-to-peer and customer-to-provider degrees respectively. Additionally, from the plot, the maximum degree of provider-to-customer, peer-to-peer and customer-to-provider link can be estimated to be 6000, 6000 and 110 respectively.

The slope of provider-to-customer and customer-to-provider degrees can be seen to be fairly linear on the log-log scale and that the slope of customer-to-provider link is steeper than that of provider-to-customer degrees. This is because an AS can have multiple customers with a higher probability when compared to the probability of an AS having multiple providers. This is true for the Internet as the tier-1 and tier-2 ASes can have multiple customers and hence have a higher value of the provider-to-customer degree. An AS does not purchase transit from multiple providers unless necessary. Therefore, the slope of customer-to-provider CCDF is steeper. However, ASes in tier-2 and tier-3 try to purchase transit from multiple ASes to ensure that their network is resilient to failures at their provider ASes and back-up options are made available to optimize operational requirements.

The CCDF of the peer-to-peer link in the plot shown in figure 4.11a behaves differently. The slope varies in three different phases. In the first phase between degree 10^0 and $3*10^1$, the slope is less steep (as compared to the third phase), and the CCDF reduces at a slow rate. This suggests that the peering is being adopted by a few smaller networks and are trying to exploit peering to their advantage by peering with willing network operators. The second phase between degree $3*10^1$ and $8*10^2$ shows that the tier-1 ASes peer with not only ASes that belong to tier-1 but also from other tiers, for instance, AS286 (a well known tier-1 operator) and AS6939 (a customer of AS1299, a tier-1 AS) are inferred as peering ASes. Thus the global behavior of tier-1 ASes is visible. And finally, the behavior of tier-2 ASes which peer heavily with a majority of networks and can be observed beyond degree $8*10^2$.

Parameter	Value	Description
α	0.69	Slope along x-axis
β	6000	Maximum degree an AS can have in the topology
N_o	12755	Number of ASes that form the core
L_o	568054	Number of edges that form the core
N_c	50218	Number of client ASes
L_c	123606	Number of links client ASes use to connect with the core
X	150	Maximum providers a client AS can have

Table 4.10: Parameters for AStop- topology generator [43].

[43] requires seven parameters as the input to generate a topology resembling the Internet. Using the plot of real Internet CCDF shown in figure 4.11a, these parameters are calculated and presented in the table 4.10.

The model is shown to be closely replicating the CCDF of the Internet in [43], however suffers disadvantages and fails to replicate the present topology with above mentioned parameters. In phase 1, assignment of available degrees to each node i where $D_i = \beta x^{-\alpha}$, is done by choosing a random number x, where $1 < x < N_o$ and as the value of N_o increases it becomes impossible to satisfy $\sum_{i=0}^{N} D_i \geq 2L$, which is a necessary condition for AStop and the topology generator would fail. To generate the topology with parameters shown in table 4.10, the constraint on x is relaxed such that $1 < x < 2000$.

By relaxing the constraint on the random number x, allowed the topology generator to satisfy $\sum_{i=0}^{N} D_i \geq 2L$ and proceed with the topology generation. However, there is a more important disadvantage of this model. The model fails to replicate the hierarchical property of the Internet. The Internet is a multiple-tier hierarchy and [43] based topology generator generates only two. In the first phase, [43] generates a core with N_o number of ASes and connects them with peer-to-peer or sibling-to-sibling links by using the preferential link assignment as explained in section 3.3 and later attaching customers to the ASes of the core. The assignment of only peer-to-peer links in the core would create unreachable source-destination pairs. A core network's customer routes would become unreachable to the core ASes which do not peer with it. Therefore, there is a need to discuss an Internet model which can capture the hierarchical characteristics of the Internet.

4.3. Behavioural model of the Internet

[43] based model does not implement the hierarchical nature of the Internet or discusses the relationship assignment mechanism for provider-to-customer links within the core. This thesis proposes a model for AS topology generation which incorporates the hierarchical nature of the Internet. Additionally, since the objective of this topology generator is to create AS-graphs with up to 600 ASes, the preferential neighbor attachment is replaced with random neighbor attachment by applying well-known rules of the AS interactions. This does not have any impact while performing studies on inter-AS routing as the primary objective of the topology generator for BGP simulation is consistent relationship assignment and ensuring every AS is assigned a provider. Additionally, even though the preferential attachment is replaced with random neighbor attachment, the graphs generated by this model obeys the power-law which is confirmed by the CCDF against degree plots for the graphs generated using the proposed method as the degree of each AS is driven by exponentially controlled probability.

The tier-1 core is a complete graph and is capable of transporting traffic globally. Tier-1 ASes form a complete graph and are interconnected through peer-to-peer links with other tier-1 ASes. These include ASes like KPN, Telia, Sprint. Tier-1 ASes form the backbone of the Internet and act as the providers for tier-2 ASes. The AS can become part of the tier-1 core with very low probability. The identification of tier-1 ASes could be made by creating a sub-graph of the AS graph by using the list of ASes with zero customer-to-provider link. The absence of customer-to-provider link indicates that the AS can access the entire Internet through its peer and own network. This list is used to create a sub-graph of the AS graph generated using [5], and the largest clique is identified. There are 19 tier-1 ASes which are determined using the property of tier-1 core on the AS graph generated using the [5] data from 01-10-2018. This implies that only 0.0003 fractions of ASes on the Internet are acting as tier-1 ASes. Thus, in the graph model, the number of ASes in the tier-1 is computed by multiplying this fraction to the total number of nodes in the graph. Since the simulation of BGP traffic is done on network sizes of up-to 600 ASes, it is necessary to select a minimum number of tier-1 ASes. To ensure that there are at least three customer cones of tier-1 ASes, the number of tier-1 ASes is determined by $max(3, 0.0003 * 600)$.

The tier-2 ASes have regional influence. These ASes connect with tier-1 AS through customer-to-provider links to achieve global connectivity. Peering is done on this tier based on accessibility of peering location, point-of-presence and the size of tier-2 AS with which AS wishes to peer. Peering must be mutually beneficial to both ASes, and this can happen only when there are sufficiently high volumes of traffic that can be exchanged between the respective ASes. In the case of peering, where one AS is much smaller than the other AS, the smaller AS would get immense benefits and hence such peering would not be lucrative for the larger AS. Thus, the larger AS would prefer smaller AS to be a customer instead of a peer. The tier-2 ASes are characterized by a high number of peers and large customer network. Tier-2 ASes can be found by selecting the customers of tier-1 ASes which are the ASes connected to tier-1 ASes through customer-to-provider links, and their peering degree is greater than 25 and ASes are not found in tier-3 and tier-4. 472 tier-2 ASes are found by using this heuristic, which comprises to be 0.007 fractions of the total ASes. To ensure that there are at least ten customer cones of tier-2 ASes, the number of tier-2 ASes is determined by $max(10, 0.007 * N)$ where N is

the total number of ASes required in the model Internet.

The tier-3 ASes have local influence. These ASes would consume services from tier-2 ASes to gain access to the global Internet. Peering is less likely to happen as the infrastructure overhead would play a significant role. Also, there would be not enough traffic exchange between two tier-3 ASes to compensate for the overwhelming cost of infrastructure required to peer and therefore peering would remain ineffective. Tier-3 ASes provide internet access to individual consumers and internet services to tier-4 ASes. To determine tier-3 ASes, the set of tier-2 ASes is used, and the neighbors of tier-2 ASes which are connected through a provider-to-customer link are added to tier-3 ASes. 6105 tier-3 ASes are identified, which constitute 0.097 fractions of ASes in the AS-graph.

The tier-4 ASes are stub ASes, and they only accept traffic which is destined to their networks. These are large enterprises with their own AS number and IP range purchased either from an upstream provider or IANA [8] directly. It is possible for such ASes to connect with two or more tier-3 or tier-2 network providers depending upon the traffic volumes and preferences. However, having multiple providers would have a financial impact; this is less likely to happen. These ASes do not have peers. The ASes of the AS graph which are not present in tier-1, tier-2 or tier-3 are considered as tier-4 ASes.

In the above paragraphs, peering links and transit links and hierarchy of the Internet was discussed for the Internet topology generator. Due to mergers of the network providers, certain peering or transit links become sibling links or new links emerge as sibling links. These are rarely occurring, and their behavior is similar to a bi-directional provider-to-customer link. Due to their limited occurrences, these links are ignored from the model. Finally, prefixes are assigned to the edge ASes. An AS can announce multiple prefixes on the Internet. However, announcing more prefixes per AS in simulations would increase computational load. Therefore only single prefix is announced per edge AS. On the Internet, AS paths from source to destination are unique and thus by announcing multiple prefixes from same AS would not provide any additional information during simulation. The prefixes announced may or may not be aggregated by the top-level service provider. The distribution of address space was inefficient in the early days of the Internet, which implies that it is hard for a top-level provider to aggregate customer prefixes due to unavailability of continuous address space within its customer cone. Thus the effect of aggregation has been ignored.

The algorithm 7 generates the toy graph for the experiments runs in five stages and aims to incorporate the following properties of the AS-level Internet:

1. AS-level is hierarchical, and every AS is assigned to a tier. This is done through the tier assignment process where every AS of the graph is assigned to a tier from 1 through 4.
2. An AS of higher tier acts as a provider to an AS in lower tiers.
3. Peering is possible within a tier with a certain probability.
4. An AS may have multiple providers depending on their business requirements and is restricted to a maximum of 3 providers per AS due to the size of the graph.

(2) and (3) items from the above list are achieved using neighbor assignment procedure (see algorithm 8) which accepts a list of ASes to which neighbors are to be added, along with a list of ASes from which neighbors are added. The third list is used to determine the number of neighbors that an AS can have and is selected randomly by assigning probability $exp(-1 * k * i)$ to every element i of the third list where k is the scaling factor controlling the probability of assigning multiple neighbors.

Algorithm

The algorithm requires following inputs:

1. Total number of ASes in the graph.
2. A tuple of four elements indicating the fraction of ASes in each tier.
3. The tuple of two elements indicating minimum and maximum number of providers from respective tier for each tier.
4. The tuple of two elements indicating minimum and maximum number of peers within each tier.

5. A tuple of four elements probability scaling factor such that each element in range of minimum and maximum number of providers can be chosen accordingly.

6. A tuple of four elements probability scaling factor such that each element in range of minimum and maximum number of peers can be chosen accordingly.

7. A tuple of two elements indicating minimum and maximum number of ASes that can announce a single prefix.

8. A probability scaling factor for assigning probabilities to each element between the range of minimum and maximum number of ASes that can announce a single prefix.

9. (Optional) if sibling relations are to be added then, the fraction of sibling links desired in the final model.

In the first stage, ASes are created based on the required size of the graph and the ASes are distributed in four tiers using input (2). In the second stage, the algorithm creates peering links within a tier. By using the minimum and maximum number of peers an AS in tier j can have a list of numbers between this range is generated, and every element is assigned a value based on $exp(-1 * k_j^{peer} * i)$, where i is the i^{th} element of the list and k_j^{peer} is the peering neighbors scaling factor for the tier j. In the third stage, each the providers are assigned to each AS, lower the tier-number higher the AS in hierarchy. By using the minimum and maximum number of provider an AS in tier j can have from the ASes of tier k, a list of numbers between this range is generated and every element is assigned a value based on $exp(-1 * k_{jk}^{provider} * i)$, where i is the i^{th} element of the list and $k_{jk}^{provider}$ is the provider scaling factor for the tier j to tier k. In the fourth stage (optional), some links are converted into sibling relations by making a random selection. The fraction of sibling-to-sibling links on the Internet is under 0.01 and thus ignored for small graphs. In the final stage, prefixes are assigned to all ASes without any customer AS. Again, by using the minimum and maximum number of ASes that can announce a single prefix is generated and each element of the list is assigned a probability of selection based on $exp(-1 * k_{jk}^{provider} * i)$ where i is the i^{th} element of the list and k_{prefix} is the probability scaling factor. Prefixes are attached to edge-ASes only to avoid redundant path calculations for prefixes originating near the core. Edge ASes can only use the inter-AS routes selected by their providers, and these decisions remain unchanged and thus by announcing prefixes from the edge sufficient to explore inter-AS routes adopted by ASes close to the core. This completes the algorithm used for the generation of the Internet like toy graphs for the experiments.

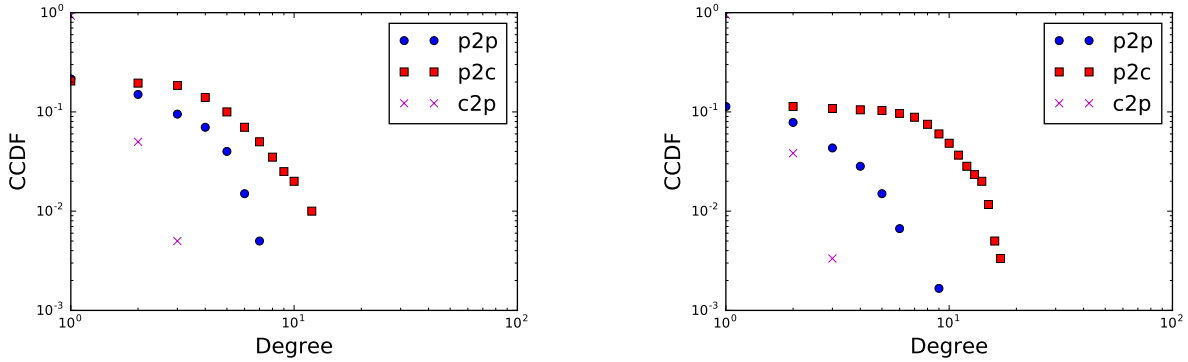
Illustration

To illustrate how the algorithm is generating the graph, let us consider that we need to create a graph of 100 ASes. In order to distribute ASes into different tiers input (2) is provided to the algorithm as $(0.0003, 0.007, 0.0927)$. As we can see that the fraction of ASes in tier-1 and tier-2 are tiny and a product with 100 would produce zero elements in tier-1 and tier-2, therefore an additional fail-safe is implemented which would force at least 3, 10 and 30 ASes in tier-1, tier-2, and tier-3 respectively. Remaining ASes would be assigned to tier-4.

In the next step, peering links are added to within each tier. Peering is 100% in tier-1 and therefore minimum, and maximum values of peers in tier-1 would be 2, and the probability scaling would not be required. This would allow the algorithm to create a tier-1 clique. An AS in tier-2 peers with as many ASes as possible and the minimum and the maximum number of tier-2 peers an AS could have is determined by allowing a tier-2 AS to peer with at most 50% of the ASes within tier-2. The probability scaling factor would assign probability to each element of the list $[0, 1, 2, 3, 4]$ as $(0.0117, 0.0317, 0.0861, 0.2341, 0.6364)$. Then for each AS of tier-2 a number x is picked from the range of minimum and the maximum number of peers an AS could have, and x peers are added to the AS. The selection of x peers is made by assigning equal probability to each AS which is computed using $\frac{1}{(tier\_2)-1}$. The process is similar for adding peers in tier-3. The maximum number of peers that a tier-3 network could have is 30% of all tier-3 ASes, and the scaling factor used is $k_3^{peer} = 3$.

After the peer assignment, providers are assigned to an AS from a lower tier from an AS of a higher tier. For instance, a tier-3 AS may have a provider from tier-2 or tier-1 ASes. Again, the minimum and the maximum number of providers an AS from tier-j can have from tier-k is generated and probabilities are assigned to bias selection according to the behavior of ASes reflected on the Internet. Since the graph size up to 600 are used in the thesis, the maximum number of providers an AS can have is limited to 5, 3 and 2 for tier-4, tier-3 and

tier-2 ASes respectively, The probability scaling factor $k_{jk}^{provider}$ is used as 5, 2, and 1 for tier-4 to tier-3, tier-3 to tier-2 and tier-2 to tier-1 respectively. Finally, prefixes are generated for all tier-4 ASes and assigned to each AS.



(a) CCDF of the topology generated for 200 ASes.

(b) CCDF of the topology generated for 600 ASes.

Figure 4.12: CCDF of AS graph using the behavioural Internet model proposed in this thesis.

Figure 4.12 shows the CCDF against degree log-log plot for two graphs with 200 and 600 ASes generated using the algorithm proposed in section 4.3. The plots confirm that the model is complying with the power-law: $d_i \propto \beta x^{-\alpha}$. The hierarchy of the topology is enforced as the model is based on the hierarchical properties of the AS-level Internet. Additionally, the links are assigned relationship at the time of end-point assignment and thus the possibility of inconsistent relationship assignment is avoided (see section 4.2).

4.4. Challenges of inter-AS route monitoring

BGP monitoring is a challenging task, and the important challenges encountered while monitoring the BGP traffic are as follows:

1. The size of decompressed data produced each day amounts to 300GB. Due to transients on the Internet, As soon as there is a change in RIB of an AS, the AS communicates the change to neighboring ASes using UPDATE messages. A single UPDATE message creates a chain reaction and every AS receiving the UPDATE may accept the change and modify its RIB and generate an UPDATE message for its neighbors as explained in section 2.3.

2. Another reason for a high volume of BGP traffic is pathological UPDATES [34]. Due to faulty router software, an AS may generate duplicate UPDATE messages and if these are not suppressed would increase the volumes of UPDATE messages generated on the network.

3. BGP traffic monitoring does not have any business incentive for the AS. Transporting customer AS traffic is the source of revenue for an AS. Hence many ASes are unwilling to share network bandwidth and computational power to support monitoring projects.

4. The optimization of BGP monitor selection is computationally unfeasible due to the size of the Internet. 62775 ASes are participating in the global routing and contributing to the BGP traffic. Selecting *k-monitors* from a set of *n-ASes* requires evaluation of $\binom{n}{k}$ possible combinations of monitor sets. Even if somehow all combinations of the monitoring ASes are retrieved, the only way to know the link-coverage at each monitor is through respective RIB. Again, collecting RIB for all ASes would require extensive co-operation from all ASes, which is unlikely to happen. Alternatively, the BGP traffic can be simulated in a simulator, but it would be computationally intensive and may take years to obtain relevant results thus rendering the activity irrelevant.

5. Lastly, optimal criteria for monitor selection is not yet defined for the BGP monitoring system. This makes the evaluation of results from each scheme based on link-coverage only.

During the study, it is identified, that significant contribution to the size of the BGP data is the frequent collection of the RIB. This information is dumped into the collector database every 120 minutes from selected VPs (or monitors). RIBs can be used as an instantaneous snapshot of the Internet. However such snapshots can be generated at regular intervals using a single RIB dump and incrementally processing the BGP UPDATE messages for the desired interval. Thus a reduction in the size of BGP data can be achieved, eliminating the concerns regarding data management and storage.

The business objective of the network operator is to transport high volumes of user-data to maximize their revenue. Since transporting BGP traffic towards a collector node would limit the bandwidth available for routing the user-data, network operators perceive this as a business loss. If the estimated cost of transporting one kilo-bit of data per second is $ 0.0001, and assuming that an AS generates 100 BGP announcements every second. An AS could estimate a loss of $ 12,240 in revenue, considering the minimum size of BGP UPDATE message is 37 octets. Frequently generating monitoring data would consume existing computing resources. The lack of awareness of the benefits of the inter-domain route monitoring makes network operators reluctant towards participation. The attention of the network operators has been drawn towards the benefits of BGP monitoring, by providing them with solutions for stability issues and hijacks using the inter-AS routing information. However, due to insufficient monitoring infrastructure, these solutions remain less effective, leaving the Internet, vulnerable to such conditions and fail to draw the attention of a larger audience.

The Internet is a complex network with a hierarchical composition of ASes. The core of the Internet comprises of tier-1 network providers, which form a clique. There are 19 tier-1 operators, and each provides global transit to tier-2 network providers. Tier-2 networks buy transit services of tier-1 providers. For profitability reasons, tier-2 networks also peer with other tier-2 networks to reduce dependency on their tier-1 providers. Peering is a settlement-free arrangement made with ASes of similar size measured in terms of customer-cone [37]. Peering creates additional paths within a customer cone or between two customer cones and hence creating additional paths between source and destination.

The benchmarking of the monitoring system can be done by comparing the number of visible links. By obtaining RIB from a single AS, all prefixes can be observed. However, the paths to respective prefixes would be governed by the internal policies of the intermediate ASes that are traversed. This would lead to incompleteness in the number of ASes that can be observed through the data obtained from a single AS and also provide limited visibility in terms of links. Therefore, multiple points of monitoring are essential to reveal all possible inter-AS routes. In the literature, it is concluded, that only BGP data is insufficient to observe all BGP links. There is conclusive evidence that BGP data does not reveal all links and there are a fraction of links that are obtained by making use of alternate AS link discovery techniques which are based on traceroute tool [36, 37].

4.5. Formulation of monitor selection problem and related assumptions

The objectives of this thesis are as follows:

1. To evaluate the effectiveness of monitor placement schemes proposed by [52] on the Internet-like graphs.
2. To produce a monitor set such that the information overlap between monitors is reduced.
3. Estimating the size of monitoring system required to achieve 100% link-coverage.
4. Identify the optimal scheme for monitor placement from set of available monitor placement schemes.

A scheme for monitor selection would require information regarding the network topology and knowledge of other constraints which influence the behavior of monitoring data, which in this case is inter-AS routing information. The schemes discussed in this thesis would make the following assumptions while selecting monitors:

1. The topology of the Internet is known.
2. AS relationships are known.

3. During simulations, uncertainties arising due to local preferences are ignored and cases where LOCAL_PREF of two provider ASes is same, the ties are broken randomly.

4. Information regarding the RIB for an AS is not available at the time of monitor selection.

4.6. Peering-degree based monitor placement scheme- basic version

Let us now move towards development of peering-degree based monitor selection schemes, the novel algorithms proposed in this thesis. Three versions of the algorithm are presented based on the complexity of the search performed in each scheme. In section 3.5, we saw the schemes proposed by [52]. The author evaluated the performance of those schemes by plotting the number of observed links for each scheme against the number of vantage points and it was reported that greedy-link based monitor selection scheme outperformed remaining three approaches (address block-based, random-based and degree-based) [52]. The greedy-link based monitor selection algorithm cannot be used to identify new monitors as the BGP data is unavailable for such ASes. BGP data acts as an input for the algorithm and is used to extract the list of AS paths announced by the monitoring ASes, and because the information would be unavailable for the new monitors, this scheme cannot be implemented for practical purposes.

Next to greedy-link based monitor selection scheme is the performance of degree-based monitor selection scheme. As we shall see in section 5, degree-based monitor selection scheme shows a linear rise in the number of observed links, but the gradient decreases as more monitors are added. This is because the degree-based monitor selection scheme could select adjacent ASes and the number of newly observed links would diminish as more monitors are added.

Adjacent ASes also introduce redundancy in the BGP data. Redundancy in BGP data reduces the effectiveness of the monitor system. The greedy-link based algorithm avoids redundant monitoring sources and therefore, performs the best. Without the knowledge of BGP data for every AS, greedy-link based scheme cannot be used for practical purposes to deploy new monitors. The performance of degree based scheme is next to greedy link based scheme but this scheme does not address the redundancy in BGP data, and *the question is could something be done about this?*

[22] highlighted that adjacent peer ASes that provide BGP data have higher redundancy in the data. Thus to begin with degree-based monitor placement scheme can be improved to reduce the redundancy in BGP data by either avoiding the selection of adjacent ASes or by eliminating the adjacent ASes at a later stage of the algorithm. Thus, by eliminating adjacent peering ASes, BGP data redundancy is addressed. However, *can something else be done to improve the link-coverage of the monitoring network further?*

In section 4.1 we performed BGP analysis on the sub-graphs generated by attaching frequently occurring graph motifs to a representative Internet. There were some interesting observations which could help further in monitor placement. Firstly, in the absence of the peer-to-peer links in the graph, all the links could be discovered at the top-level AS which in figures 4.2a and 4.3a are AS1 and AS6 (see table 4.1 and 4.2). Secondly, in the presence of the peer-to-peer links in the sub-graph top-level AS could still discover most links but peer-to-peer links within its customer cone. In figure 4.7a, the peer-to-peer link between AS4 and AS5 is only observed at AS4 and AS5 which are the end-points of the peering link (see table 4.6).

Top-level ASes tend to observe most links and, yet the performance of address-based monitor placement which tends to select top-level ASes is worst or similar to that of random-based monitor selection [52]. This is because on the Internet, peering is extensively done at tier-1 and tier-2 and to a certain extent on tier-3 as well and the peering links at the lower tier ASes remain invisible to the higher tier networks. From figure 4.6a through 4.10a, we could see that peer-to-peer links were discovered only at the ASes which form the end-points of the peer-to-peer links or at an AS that is within the customer cone of the ASes that form the endpoints of peering link. Thus, it could be favorable to select monitors based on their peering-degree rather than neighbor degree as described in section 3.5.

Algorithm

The peering-degree based monitor selection scheme begins by accepting the AS graph as the input. The AS graph is a directed graph in which the nodes represent the ASes and the edges define the business relationship between the two ASes. The output of the algorithm would be the monitor set and contain a set of nodes (or

ASes) from which BGP data should be collected. The algorithm works in two phases. In the first phase it searches for possible monitor candidates, and in the second phase, it reduces the possibility of including the adjacent peering ASes.

The search phase identifies the ASes that have peer-to-peer links and adds them into a list of possible monitors. ASes with peer-to-peer links are added to the monitoring network to ensure that the endpoints of a peering link are included in the monitor set as peering links cannot be observed by the ASes which are not present in the customer cone of peering ASes. The phenomena of peering is more common on higher tiers of the Internet, and this list of possible monitors would contain the ASes from tier-1, tier-2, and tier-3. In the future, if peering becomes common for lower tier ASes, the algorithm would select the ASes from lower tiers as well, thus ensuring that the monitors are distributed across multiple tiers of the Internet.

In the second phase, a sub-graph of the AS graph is created using the list of possible monitors. Next, from this sub-graph, all but peer-to-peer links are removed, and all the nodes of the sub-graph are ranked based on the number of edges (peering degree) each node posses. The top-ranking AS is selected for the final monitoring set and all the peering edges of AS selected into the final monitor set are removed. This would affect the peering degree of the neighboring nodes of the AS. The process of ranking and elimination is repeated until all the peering edges are removed from the graph.

It is possible that multiple nodes have the same number of edges (peer-to-peer links) and that their ranks would be equal. In such cases, the tie could be broken on four different levels. The first level of check is based on the total number of neighbors the ASes have in the original AS graph. The customer networks commonly originate the announcements and propagated by their provider ASes across the Internet through a series of BGP UPDATE exchange with neighboring ASes. An AS with a higher number of neighbors would potentially influence the routing of more ASes. Thus, the first level of the tie is broken in favor of an AS with a higher degree.

In the case of the number of neighbors of multiple ASes are equal, the second level tie is broken based on the number of providers. This is because, an announcement which is distributed to a large number of provider ASes would travel faster on the Internet, as compared to an announcement distributed to a large number of customers. In the event, where tie could not be broken based on the number of providers, the next level of the tie is broken based on the number of customers.

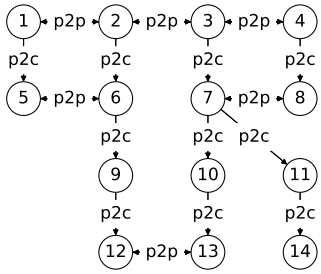
If multiple ASes have an equal degree and the equal number of providers, the tie is broken in favor of the AS with most customers. An AS with a large number of customers would generate higher volumes of announcement towards the core of the Internet. As the number of providers for both the ASes is equal, the potential of originating a malicious activity is higher within the customer cone of the AS with a higher number of customers. Therefore, to reduce the time interval between the instance at which a malicious announcement is generated by a customer network and the instance at which the monitor reports the announcement, it is important to break ties in favor of an AS with a large number of customers.

In case the number of customers for multiple ASes are also equal, the next level of tie-breaking cannot be done based on the number of sibling ASes as this would also be equal. Until now, we have compared the number of peers, total number of neighbors, number of providers and the number of customers and as all these parameters are equal for the ASes under tie-breaking mechanism, the number of sibling ASes would also be equal. Therefore, the final tie-breaking is done based on the AS number. AS numbers are unique and therefore would break the tie in favor of one AS.

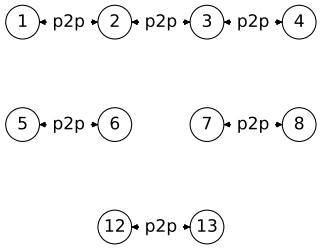
Illustration of algorithm

Let us consider a simple AS graph as shown in figure 4.13a and the arrowhead in a p2c link indicates customer. The search phase of the algorithm selects AS1, AS2, AS3, AS4, AS5, AS6, AS7, AS8, AS12 and AS13 as the list of possible monitors because each of these ASes posses at least one peer. The elimination phase would generate a sub-graph of the AS graph shown in figure 4.13a and remove all edges but peer-to-peer edges which would result in the sub-graph as shown in figure 4.13b. Then, the algorithm would rank the ASes based on their peering degree. In this case, AS2 and AS3 are ranked highest as they both have two peers each and the tie would be broken in favor of AS2, as it's AS identifier is lower because the remaining parameters are equal for both the ASes. The peering edges of AS2 are removed from the sub-graph.

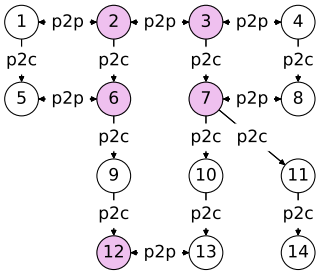
The process of elimination would again rank the ASes based on the peering degree. Now, all the remaining ASes have the peering degree of one, and therefore the tie-breaking would be performed. The ties are



(a) Toy graph used to illustrate the monitor selection process.



(b) The sub-graph generated during the elimination phase of basic version.



(c) Final monitor set generated by the basic version of the peering-degree based algorithm.

Figure 4.13: Illustration of the basic version of the peering-degree based monitor selection scheme.

broken on four levels based on total degree, number of providers, number of customers and finally on the AS identifier. Therefore, the first tie would be broken in favor of AS7 as it has the highest degree in the original graph and all the peering edges with one endpoint as AS7 would be removed from the sub-graph.

Again, the remaining ASes AS3, AS5, AS6, AS12 and AS13 would be ranked equal, but the degree of AS3 and AS6 is highest. Therefore the tie would be evaluated between AS6 and AS3, and AS6 would be selected as its provider degree is more than the provider degree of AS3. After addition of AS6 into the monitor set, AS5 is no longer eligible for peering degree based ranking because it had only one peering neighbor.

Next AS3, AS4, AS12, and AS13 have a peering degree of one and their degree but the degree of AS3 is highest in the original graph, and therefore AS3 is added into the monitor set. After addition of AS3 into the monitor set, AS4 is no longer eligible for peering degree based ranking because it had only one peering neighbor. Finally, AS12 and AS13 remain and because the AS identifier of AS12 is lower and it would be added to the final monitor set. The final monitor set is shown in figure 4.13c and the monitors are the colored nodes.

Limitations and complexity

The algorithm is fairly simple and performs multiple sorting and selection. The time complexity of presently available sorting algorithms is $n \log n$ and to select m-monitors in a graph with n-ASes would require:

$$\sum_{i=n-m+1}^{n} i \log i \approx \mathcal{O}(n \log n)$$

This simplicity comes at the cost of the number of monitors selected. The algorithm produces a monitor set in which there is a possibility of further reducing the number of monitors. As we can see in figure 4.13c, information available at AS2 and AS3 could be captured effectively at AS6 and AS7 respectively. Therefore, there is a scope of refinement in the algorithm, and we can improve the basic algorithm to consider the situation where a provider of a monitor is present in the monitor set, the provider is eliminated from the monitor set. This refinement could be achieved by increasing the complexity of the monitor selection scheme by avoiding the selection of an AS whose direct customers are in the monitor set.

4.7. Peering-degree based monitor placement scheme- intermediate version

In the previous section, we saw a peering-degree based monitor selection scheme in its basic form. It relied on the fact that the peering link can only be seen by the ASes that form the endpoints of the peering link. However, the peer provided paths can also be captured by the ASes which lie within the customer cone of the ASes. In the intermediate version of the algorithm, this fact would be exploited to reduce the size of the monitoring set further.

The algorithm would again work in two phases as it was in the basic version of the algorithm. The task of the search phase would be to identify the ASes with peering degree greater than zero and provide it to the elimination phase, which is the second phase of the algorithm. In the improved version of the algorithm, the elimination phase would attempt to search within the direct customers of the ASes with peering links to determine if the monitoring is possible at a customer AS. Again, it is important for this customer AS to have peering links, because in the absence of peering links it would be challenging to determine which customer to pick and the size of monitor would remain unchanged if the monitoring is performed at the provider AS or a customer network.

Algorithm

As indicated above, the search phase of the algorithm is similar to the basic version. The input AS graph is processed to generate a list of possible monitors by adding the ASes with peering degree higher than one. Using this list, the elimination phase generates a sub-graph of the AS graph. This reduces the search space in which the exploration of final monitors is performed. In the intermediate version of the algorithm, the elimination phase does not remove any links as it was required for the elimination phase of the basic version.

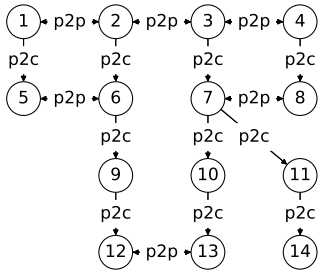
The elimination phase ranks the ASes based on their peering degree and selects the AS with the highest rank. In the case where multiple ASes have been ranked equal, the tie-breaking mechanism is similar to that of the basic version. Once, the AS with the highest rank is determined, the algorithm begins exploring all the ASes that can be reached through customer-to-provider links from the highest ranked AS and adds them into a provider list. Furthermore, for each AS of the provider list, the algorithm searches for the peer-to-peer links and creates a provider-peer-edge list. Now, we have obtained a set of peering links that could be observed by monitoring the highest ranked AS and all the ASes which are reachable through a series of customer-to-provider links. The algorithm would then remove all the edges added in the provider-peer-edge list and all the nodes corresponding to the ASes that belong to provider list from the sub-graph and add the highest ranking AS into the final monitor list.

Once the AS is added into the list of final monitors, the algorithm performs an additional check and searches for a direct customer with the peering links for this monitor. If the direct customers with peering links are not found, the algorithm allows this AS to remain in the list of final monitors and proceeds with the ranking of the remaining ASes with peers. However, if a direct customer with peering link is found, then the algorithm removes the highest ranked AS from the list of final monitors and creates a provider list and provider-peer-edge list for this customer AS and then removes all the ASes of the provider list and all the edges of the provider-peer-edge list from the sub-graph and adds this customer AS into the final list of monitors.

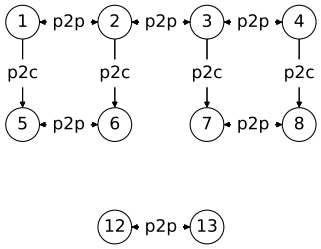
In case, the highest ranking AS has multiple direct customers with peering links, the customer ASes are ranked and contested as per the elimination process. The algorithm, in an iterative manner, looks for a possible monitor within the customer cone of the AS which is about to be added into the list of final monitors. This iterative search allows monitor placement algorithm to produce a monitor set adaptively and reduce the size of the monitor set.

Illustration

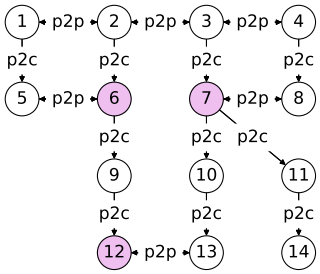
Let us consider a simple AS graph as shown in figure 4.14a and the arrowhead in a p2c link indicates customer. The search phase of the algorithm selects AS1, AS2, AS3, AS4, AS5, AS6, AS7, AS8, AS12 and AS13 as the list of possible monitors as each of these ASes posses at least one peer. The elimination phase would generate a sub-graph of the AS graph as shown in figure 4.14b and rank the ASes based on their peering degree. In this case, AS2 and AS3 are ranked highest as they both have two peers each and the tie would be broken in favor of AS2, as its AS identifier is lower.



(a) Toy graph used to illustrate the monitor selection process.



(b) The sub-graph generated during the elimination phase of intermediate version.



(c) The monitor set generated by the intermediate version of the peering-degree based algorithm.

Figure 4.14: Illustration of a intermediate version of the peering-degree based monitor selection scheme.

Now the algorithm would generate a list of providers for AS2, which in this case is an empty set and the list of provider-peer-edges would contain AS1-AS2 and AS2-AS3 edges. The edges in the provider-peer-edges list would be removed from the sub-graph, and the algorithm would be ready to add AS2 into the list of final monitors. However, AS2 has a customer who has a peering degree greater than zero and therefore, the customer AS (AS6) of AS2 would be added to the final monitor list instead. Before adding AS6 into the final monitor set, the provider-list and provider-peer-edge list for AS6 would be generated, which would contain (AS2) and (AS5-AS6) respectively. AS2 would be removed from the sub-graph, and the edge AS5-AS6 would also be removed from the sub-graph.

AS6 does not have a customer with a peering degree greater than zero, and therefore AS6 is finally added into the monitor set. The same process of ranking-elimination-re-evaluation would be executed until the monitor set is complete. The resulting monitor-set, in this case, would be AS6, AS7, and AS12. This is a reduction of 40% when compared with the monitor set generated by the basic version. The final monitor placement is shown in figure 4.14c where the monitoring nodes are colored.

Limitations and complexity

This algorithm increases the complexity of finding a possible monitor in the direct neighborhood of the AS which is selected as the possible monitoring candidate. This allows the algorithm to reduce the size of the monitor set but not entirely. As we can see from figure 4.14c, AS6 could also be removed from the monitor set. This is because AS12 is in the customer cone of AS6 and would capture all routing information generated at AS6. This happened because the elimination phase of the intermediate algorithm did not adapt to a situation where the monitor is present deep within the customer cone of an AS. This is the basis of the advanced version of the algorithm, where the existence of the predetermined monitors is also contested when a new monitor is added.

4.8. Peering-degree based monitor selection scheme- advanced version

The monitor set generated for figure 4.14a using the intermediate version of the algorithm, could be reduced further, if during the elimination phase the complete graph is used instead of the sub-graph. This is because,

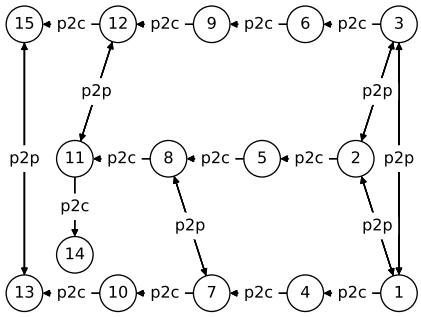


Figure 4.15: Illustration of similarity in conflict information.

AS12 and AS13 are within the customer cone of the ASes which are a part of the monitor set and therefore, if the exploration of provider list is allowed to follow link AS9-AS6 (a customer-to-provider link), then the monitor set can be reduced further, by eliminating AS6 from the monitor set. The advanced version of the algorithm takes this into account and allows provider-list and provider-peer-edge list to be prepared by considering the complete graph. This would increase the time in which the monitors are selected, but is more likely to generate the smallest possible monitor set.

Algorithm

In this scheme, the search phase is not required, and the complete graph information is used to explore the monitor set. This implies that we can begin directly with the elimination phase. The elimination phase is designed by because the peering links can be captured at an AS which is within the customer cone of a peering AS. To explore possible monitors within the customer cone of an AS is a difficult task as the number of customers would be large within an AS and selecting one of the customers would become challenging in the absence of any technical criteria to filter out an AS which could be used as a monitor. To counter this challenge, the algorithm only looks up in the hierarchy for ASes that can be removed from the monitor set due to the addition of a new monitor which lies within the customer cone of a previously selected monitor AS. Additionally, the algorithm looks for the peers of the provider AS and eliminates them as well from the monitor set if their removal does not affect the link-coverage of the peer-to-peer links.

The elimination phase would rank the ASes based on their peering degree and create a list of providers and provider-peer-edge list for the AS that is ranked highest. In the case where multiple ASes have been ranked the same, the tie-breaker criteria would be similar to that used in the basic and intermediate versions of the algorithm. Once, the provider list and provider-peer-edge list is available, and the algorithm would remove the providers from the graph. Moreover, if any AS which is in the provider list, is also present in the monitor set, it would be removed from the monitor set. This is because all the information at the provider would flow to this newly identified monitor AS which is within the customer cone of the previously determined monitor. Furthermore, all edges which are present in the provider-peer-edge list are removed from the graph, and the ranking and elimination process is repeated until all peering links are removed from the AS graph and the final monitor set is determined.

In the case where the peer of an AS from the provider list is present in the monitor set, the scheme would generate a conflict status along with the conflict information and would restart monitor selection using this conflict information. The conflict information would provide the algorithm first-level tie-breaking information. The conflict information would be a two-tuple element where the first element would be a monitor that peers with the provider AS within whose customer cone a new monitor is identified. Thus, the algorithm in its new iteration break ties in favor of the second element in the conflict information and continue this process.

It is possible that the same pair of ASes conflict in different stages of monitor selection. For instance, in the graph, as shown in figure 4.15, there are three conflicting pairs (AS1, AS2), (AS2, AS1) and (AS7, AS8). This multi-level conflict information would confuse the scheme and stall the processing by restarting each time a conflict-pair is detected. Thus, we require a mechanism to allow one conflict pair to restart the algorithm only once.

Illustration

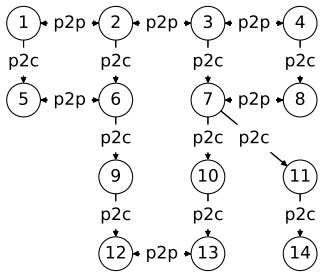
Let us consider a simple AS graph as shown in figure 4.16a and the arrowhead in a p2c link indicates customer. The search phase is not required, and the algorithm would begin the elimination phase directly. The elimination phase would begin by ranking the ASes based on their peering degree. In this case, AS2 and AS3 are ranked highest as they both have two peers each and the tie would be broken in favor of AS2, as its AS identifier is lower, since the degree, the number of providers and the number of customers is same for both.

Now the algorithm would generate a list of providers for AS2, which in this case is an empty set and the list of provider-peer-edges would contain AS1-AS2 and AS2-AS3 edges. The edges in the provider-peer-edges list would be removed from the AS graph, and the algorithm would be ready to add AS2 into the list of final monitors. However, AS2 has a customer who has a peering degree greater than zero and therefore, the customer (AS6) of AS2 would be added to the final monitor list instead. Before adding AS6 into the final monitor set, the provider-list and provider-peer-edge list for AS6 would be generated, which would contain (AS2) and (AS5-AS6) respectively. AS2 would be removed from the AS graph, and the edge AS5-AS6 would also be removed from the AS graph. This is the same as in the illustration of the intermediate version.

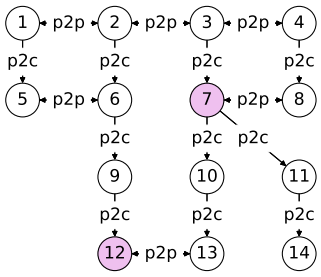
Now, that AS6 has been added into the final monitor list because the customer of AS6 does not have a peering link, the elimination phase would not remove AS6 from the monitor set and rank the remaining ASes based on their peering degree. AS3, AS4, AS7, AS8, AS12, and AS13 have the same peering degree, and the tie would be broken in favor of AS7 since the degree of AS7 is highest in the original graph. Now the provider-list would contain AS3, and the provider-peer-edge list would contain (AS3-AS4, AS7, AS8). The AS3 would be removed from the graph, and the edges in the provider-peer-edge list would also be removed, and AS7 would be added into the monitor set. Again, the direct customers of AS7 do not have a peering link, and there would be no contest between the customers of AS7 and AS7 to enter into the final monitor set.

The algorithm would again rank the remaining ASes with peering degree. AS3 is removed from the graph and is no longer eligible for selection. AS4 and AS8 are also excluded from the ranking process because their peering degree is zero. AS12 and AS13 are ranked equal, and the ties are broken in favor of AS12 (lower AS identifier). The provider list generated for AS12 would contain (AS9, AS6) and the provider-peer-edge list would contain (AS12-AS13). The final monitor set at this point is (AS6, AS7). Since the provider list of AS12 contains AS6, it implies that AS12 is within the customer cone of AS6 and therefore all information available at AS6 would be available at AS12. The algorithm would remove AS6 from the list of final monitors and add AS12. Thus, the final monitor list would contain (AS7, AS12). This shows, how the algorithm has matured and generated an adaptive monitor set. Now no more ASes with peering links are left, and the algorithm would terminate. The final monitors are shown in figure 4.16b.

The advanced version of the algorithm achieved a 60% lower monitor set size when compared with the basic version and does not compromise the link-coverage.



(a) Toy graph used to illustrate the monitor selection process



(b) Final monitor set generated by the advanced version of the peering-degree based algorithm.

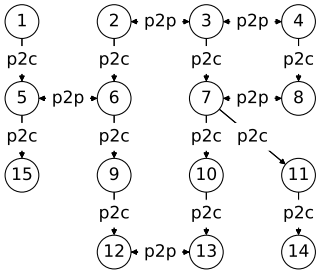
Figure 4.16: Illustration of the advanced version of the peering-degree based monitor selection scheme.

Use of conflict information

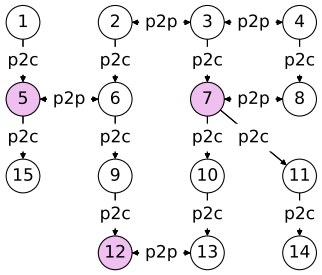
The advanced version of the algorithm attempts to place a monitor deep within the customer cone of any peering AS. This is possible because the algorithm eliminates an AS from the monitor set if a new monitor is

found within the customer cone of the AS. This allows the algorithm to reduce the number of monitors required for the inter-AS route monitoring. Moreover, when a new monitor is determined such that its provider list contains an AS which peers with a previously identified monitor, then the algorithm generates conflict information. For instance, the monitor selection for figure 4.17a would result in (AS5, AS7, AS12) as shown in figure 4.17b. In this monitor set, we can eliminate the peering monitor because the information regarding the peering link would be available at AS12 as well.

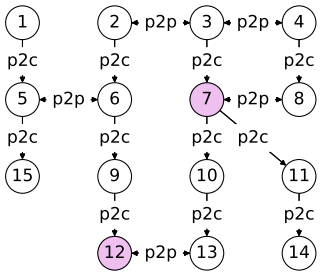
To remove an existing monitor would require restoring the changes in the graph made by its addition and would require maintaining of additional information regarding the edges and the ASes which are eliminated due to the selection of an AS and it would impact the run-time memory of the program. A simpler approach, as adopted in the advanced version of the algorithm and is achieved by raising a conflict flag and restarting the monitor selection process with conflict information indicating the algorithm to prioritize accordingly when breaking ties between the two ASes which form the conflict information. The conflict information would contain a two-element tuple in which one element is the peer monitor which was included into the monitor list at the time of conflict, and the other element is the AS within whose customer cone a possible monitor is identified which is responsible for generating the conflict. The conflict information would be used while breaking the tie between the conflicting ASes. The ties would be broken in favor of the AS which was earlier excluded from the monitor set.



(a) Toy graph.



(b) The monitor set generated by the advanced version of the peering-degree based algorithm.



(c) The monitor set generated by implementing the conflict flag variation.

Figure 4.17: Illustration of the limitations of the advanced version of the peering-degree based monitor selection scheme.

Let us consider figure 4.17a and perform the monitor selection based on raising a conflict flag. The first AS which would be added into the monitor set would be AS7 with peering degree of one and is the customer of AS3 with the highest peering degree. Next, the algorithm determines AS5 as a possible monitor since it has the lowest AS identifier, and finally, AS12 is selected. The provider list of AS12 would contain (AS9, AS6, AS2) and the provider-peer-edge list would contain (AS5-AS6, AS12-AS13) respectively. The monitor set determined up to this point is (AS5, AS7). The algorithm detects that AS5 is a peer of AS6 and would raise the conflict flag and force the algorithm to restart with conflict information (AS5, AS6). The conflict information would inform the algorithm to select AS6 over AS5 should there be a tie-breaking between them and would allow the algorithm to select AS6 instead of AS5 and later allow the algorithm to eliminate AS6 from the monitor set while adding AS12. The resulting monitor set is shown in figure 4.17c.

The conflict approach seems to solve the problem to a certain extent but raises new problems. Firstly, there could be a situation where AS5 and AS6 are never ranked equally, and therefore it would be impossible to bias the selection according to the conflict information. Secondly, there could be the same AS pair which conflict in the different stages of the algorithm and would require more effective management of conflict information generated in various iterations. One solution to this would be to capture all conflicting pairs and store them in memory to ensure that one conflict pair interferes with the monitor selection process only once.

Limitations and complexity

This variation of the algorithm is highly complex and includes multiple stages of sorting and searching. The selection of a monitor would require $\mathcal{O}(n \log n)$ time to sort the ASes based on their peering degree. Each time a monitor is selected, a list of providers and their peers is created using the well-known depth-first search algorithm with time complexity of $\mathcal{O}(V + E)$, where V is the number of vertices and E is the number of edges present in the graph. Additionally, removing ASes from the monitor set which are part of this list would require searching in the list of monitors which would further add $\mathcal{O}(n\ m)$ where n is the number of monitors present in the monitor set and m is the number of elements present in the provider-peer list. This added complexity is the trade-off to obtain the smallest monitor set without affecting the link-coverage.

To summarize, the thesis proposes three schemes to determine the monitor set for collecting the BGP data where the use-cases could determine the choice of the scheme to be used. The basic version is the simplest and the fastest scheme to determine the monitor set. This could be useful when the monitor selection is to be done for the entire Internet and when large monitor set is not a concern. The intermediate version is moderately complex and could be used when selecting monitors on the entire Internet, and it is desired to obtain as few monitors as possible. Lastly, the advanced version could be used when only regional monitoring is required under the strict requirement of reduced monitor set size.

5

Evaluation of resulting monitor sets

To fully understand the effectiveness of each monitor placement scheme, a static property defined as the link-coverage is compared. The link-coverage is the percentage of links observed by the monitor set produced by the algorithm. The comparison of the link-coverage of various monitor placement schemes is studied by generating the same number of monitors from each algorithm. BGP traffic is simulated in the simulator developed by [31] and the simulation process followed is explained in section 3.4.

The simulation process requires the topology information, prefix-origin AS association, and the relationship between two ASes (section 4.1). Providing the topology of the real Internet to simulate the BGP traffic would push the simulator to its limits. The Internet has 62773 ASes and 691660 edges [5] and the simulation process on a graph of such dimensions would complete in a non-deterministic time and, to restrict the simulation time within the observable interval the graphs with 100 to 600 ASes are used. [43] proposed an AS-level Internet topology generator and suffers several disadvantages as discussed in section 4.2. Therefore, a behavioral model of the ASes is developed (section 4.3) and used to generate the Internet-like graphs. The model described in section 4.3 requires parameters which include:

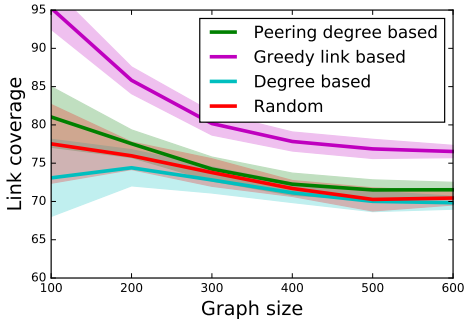
1. Total number of ASes in the graph.
2. A tuple of four elements indicating the fraction of ASes in each tier, $(a_1, a_2, a_3, 1-(a_1+a_2+a_3))$.
3. The tuple of two elements indicating minimum and maximum number of providers from respective tier for each tier, $(b^{i,j}_{min}, b^{i,j}_{max})$ where an AS from tier-i becomes a customer of the ASes from tier-j.
4. The tuple of two elements indicating minimum and maximum number of peers within each tier, $(c^{i,j}_{min}, c^{i,j}_{max})$ where an AS from tier-i becomes a peer of the ASes from tier-j.
5. A tuple of four elements probability scaling factor for each tier such that each element in range of minimum and maximum number of providers can be chosen accordingly, $(k_i^{provider})$.
6. A tuple of four elements probability scaling factor such that each element in range of minimum and maximum number of peers can be chosen accordingly, (k_i^{peer})..
7. A tuple of two elements indicating minimum and maximum number of ASes that can announce a single prefix, (p^i_{min}, p^i_{max}).
8. A probability scaling factor for assigning probabilities to each element between the range of minimum and maximum number of ASes that can announce a single prefix, k_i^{moas}.
9. (Optional) if sibling relations are to be added then, the fraction of sibling links desired in the final model.

The behavior of ASes from each tier as described in the section 4.3 and analysis of the AS topology collected by [5] on 01-10-2018 are used to determine the parameters for the Internet topology generator (algorithm 7) and are presented in appendix B.

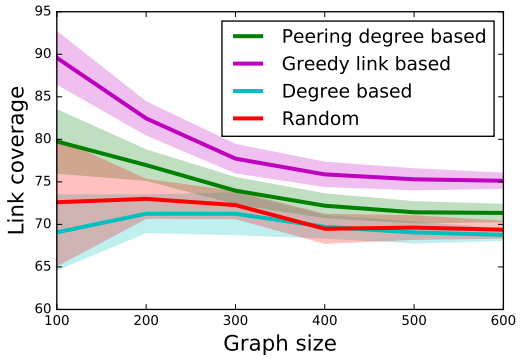
5.1. Evaluation of link-coverage for monitor placement schemes

BGP simulations are performed on the Internet-like graphs generated by implementing the method described in section 4.3. The parameters used to generate the graphs is presented in appendix B. At the end of the simulation, the simulator generates a file with routing information of each AS which could be translated into a table that resembles table 2.8 on page 13. The routing information for each AS comprises of AS paths for every prefix that is announced and using these AS paths a list of links observed by that AS is extracted for each AS and resembles table 2.9 on page 14. This list of links for every AS is called as AS-AS-link set or simply AS-Link set or VP-Link set. The resemblance of tables is implied to structure, and the contents would vary for each experiment.

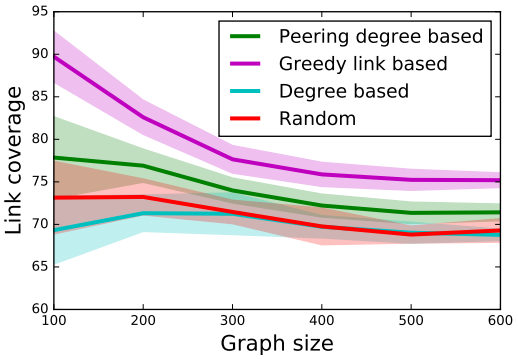
Now that we have AS-link set for every graph, monitor sets can be evaluated. The monitor sets are determined for various graph instances using the methods described in section 3.5 and is compared with the three versions of the peering-degree based monitor placement scheme proposed in this thesis. Figure 5.1 comprises of three plots where the y-axis is representing link-coverage in percentage, and the x-axis is representing the size of the graph which is the number of ASes in a graph. The different graphs are plotted for each version of the peering-degree based monitor selection scheme and shows the comparison of the link-coverage of a peering degree scheme with existing monitor selection schemes for the same number of monitors.



(a) Link-coverage vs the graph size for the monitor set generated by the basic algorithm.



(b) Link-coverage vs the graph size for the monitor set generated by the intermediate algorithm.



(c) Link-coverage vs the graph size for the monitor set generated by the advanced algorithm.

Figure 5.1: Link-coverage of various monitor-set selection schemes compared with the three peering based algorithms proposed in the the thesis.

The plots shown in figure 5.1 reveals that the greedy-link based monitor selection outperforms the remaining schemes as the link-coverage achieved is highest for the same number of monitors. This is as expected. Zhang et al. in their analysis concluded that greedy-link based monitor selection is the best based on the link-coverage it offers when compared with the link-coverage provided by the monitor sets of the same size determined by various algorithms. However, this scheme is only useful for monitor placement when the routing information (BGP data) is known in advance. When new monitors are to be placed, this is not feasi-

ble because the routing information would not be available for an AS which is not yet included in monitoring network, and hence we look at the next best option. Next to greedy-link based algorithm is the performance of the peering-degree based monitor selection scheme in all plots. This is anticipated. Peering-degree based monitor selection algorithm would perform better than the degree-based algorithm (discussed on page 23) for two reasons. Firstly, peering-degree based scheme selects ASes with peer-to-peer links, and peer-to-peer links are the links which can only be observed at either the ASes which peer or at an AS which lies within the customer cone of such ASes and increases the link-coverage. And secondly, the peering-degree based algorithm eliminates the possibility of adjacent ASes acting as the monitors. This elimination reduces the redundancy in the BGP data and the number of monitors required to monitor the inter-AS routing. In the intermediate and advanced versions of the peering-degree based scheme, a more evolved version of elimination is performed and are more effective in reducing the size of monitor set.

Surprisingly, link-coverage of the monitor set produced by the random-based algorithm is slightly better than that of the monitor set determined by the degree-based scheme. [52] showed that the degree-based scheme is better than the random-based scheme and requires further investigation.

The graphs on which simulations are performed are generated using a topology generator, multiple graph instances of the same sizes are evaluated to study the variations in the link-coverage of various monitor placement schemes. link-coverage for various monitor selection schemes was extracted from multiple graphs of the same size, and the deviation of link-coverage from its average value for different graph instance is also shown in figure 5.1 as the shaded region around the average values. This variation is due to the randomness of the Internet-like graphs used for the experiments. The variation in the performance of various monitor selection schemes stabilized and reduced as the graph size (the number of ASes in the graph) is increased and indicates that the impact of randomness reduces as the graph sizes increase.

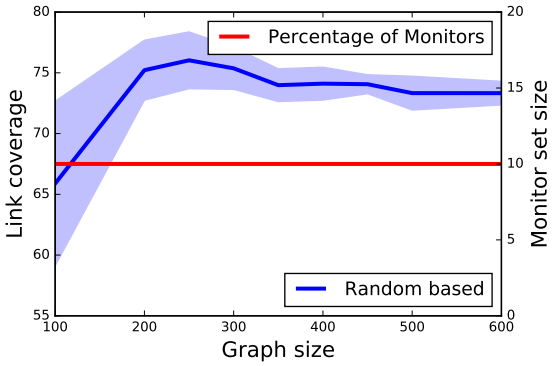


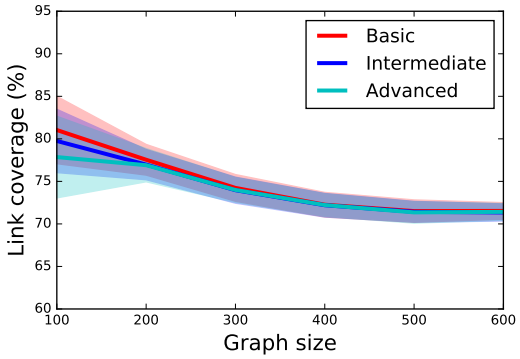
Figure 5.2: Link-coverage of the monitor-set determined by the random selection schemes.

Random-based monitor selection is random, and there is an uncertainty associated with the performance of monitor set generated by this scheme. To measure the degree of variations in the link-coverage of randomly determined monitor sets, multiple monitor sets are generated for the same graph instance using the random-based monitor selection scheme, and link-coverage is evaluated for different monitor sets. The monitor set size during this evaluation is limited to 10% for all graph sizes. For instance, for a graph with 100 ASes, 10 monitors are chosen, and for a graph of 600 ASes, 60 monitors are chosen. The red line in the plot shown in figure 5.2 shows the size of the monitor set for the graphs of different sizes.

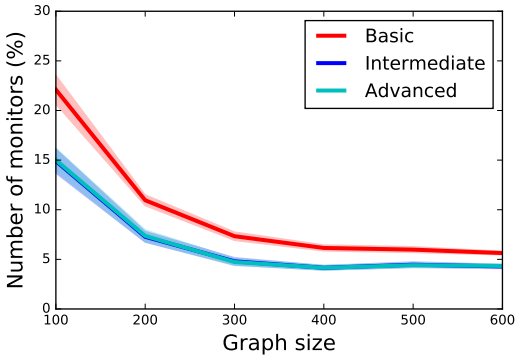
The plot of link-coverage against the graph size is shown in figure 5.2 which is a dual-axis plot. The primary y-axis is representing link-coverage in percentage, the secondary y-axis is representing monitor set size in percentage, and the x-axis is representing the size of the graph. The deviation of link-coverage from the average value is shown in the shaded region along the average value in the plots. The average value of link-coverage is 65% for the graph with 100 ASes and 74% for the graph with 600 ASes with 10% of the ASes acting as the monitors in both cases.

5.2. Comparison of the different versions of peering-degree based scheme

In this thesis, three different versions based on the peering degree are proposed, and in this section, the comparison of the performance of the basic, intermediate and advanced peering-degree based scheme is evaluated. It is anticipated that the link-coverage of the monitor sets determined by the three versions of peering-degree based algorithm would be similar, and thus, each version is allowed to generate their complete monitor set, and then the comparison is made for the number of monitors determined for different graph sizes using the three schemes.



(a) Link-coverage vs the graph size for the monitor set generated by the three versions of the peering-degree based monitor selection scheme.



(b) Number of monitors vs the graph size for the monitor set generated by the three versions of the peering-degree based monitor selection scheme.

Figure 5.3: Comparison of the three peering-degree based monitor selection scheme proposed in this thesis.

The plot shown in 5.3a is a plot where the y-axis is the link-coverage, and the x-axis is the size of the graph, and the average values are plotted for the link-coverage obtained from the monitor sets generated for the three schemes. From the plot, it could be concluded that the link-coverage of each version of the peering-degree based monitor selection is similar and confirming the assumption. The objective of the intermediate and advanced version of peering-degree based scheme is to reduce the size of monitor set without adversely affecting the link-coverage, and it could be concluded that these versions are effective in reducing the monitor set size, given that the link-coverage is similar for all versions. To measure the effectiveness of the intermediate and advanced version of the monitor placement schemes, let us consider the size of a monitor set produced by each version.

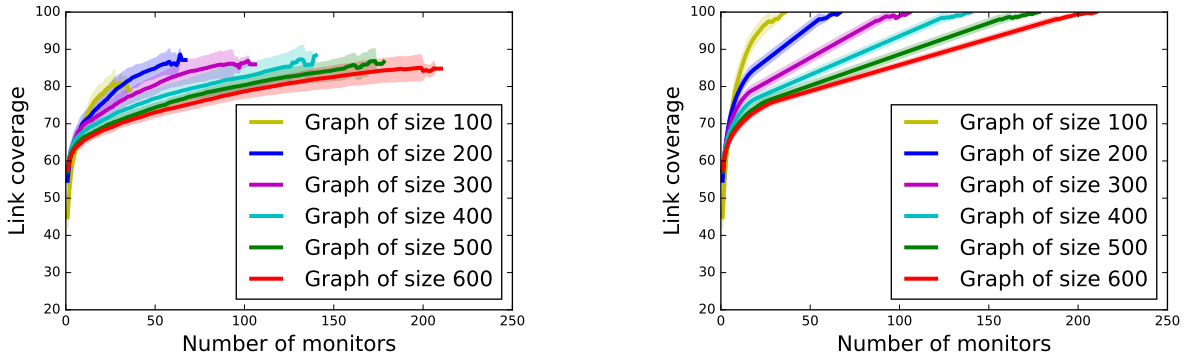
The plot shown in 5.3b is a plot of the number of monitors in percentage on the y-axis, and the size of the graph on the x-axis and the average values of the number of monitors determined by each version are plotted. It could be seen, that the size of the monitor set determined by the intermediate and advanced version of peering-degree based scheme is smaller than that produced by the basic scheme, and yet, their link-coverage is nearly similar. Additionally, the growth of the Internet happens at the edge, and the core remains relatively stable (in the number of ASes) and the number of monitor ASes that could be eliminated due to the selection of a monitor within the customer cone of such ASes is limited. Thus, the increase in graph size results in the reduction of difference in the number of monitors selected by the basic, and intermediate and advanced versions. The intermediate and advanced versions of the peering-degree based scheme are better than the basic version of the algorithm and reduce the size of the monitor set.

The above experiments are conducted on ten graphs of different sizes, and the deviation from the average values are plotted as the shaded region in figure 5.3. The deviation reduces as the graph size is increased because a change in the number of monitors or the number of links when normalized to a smaller number would be magnified more than when normalized to a larger number.

5.3. Complete coverage

In the earlier discussion, we uncovered that the link-coverage of the degree-based monitor set is surprisingly lower than that of the random-based monitor set. Intuitively, this could be attributed to the size of the monitor set, and typically more monitors would mean better coverage. To quantify "more," the impact of monitor set size on link-coverage is analyzed in this section.

The results from the previous section are reused to determine the link-coverage. The size of monitor set is increased by adding one monitor at a time, and the link-coverage is recorded with the addition of every new monitor determined by the degree-based scheme, and then the plot of link-coverage against the number of monitors is plotted for different graph sizes. The results are presented in figure 5.4a. The figure shows the plot of link-coverage (on the y-axis) against the number of monitors (on the x-axis) and the different lines are plotted for different graph sizes. The maximum number of monitors for the respective graph size was selected as 33%. Terzija et al. quoted in [49] that to achieve 100% observability of a power network, it is necessary to monitor about one-third of network buses and, gives a ballpark figure as to how many monitors should be sufficient to have a 100% link-coverage.



(a) Degree based monitor selection

(b) Greedy-link based monitor selection

Figure 5.4: Link-coverage vs Number of Monitors

Figure 5.4a shows that for each graph size link-coverage increases sharply for up to 25 monitors and then decays and eventually flattens out as more and more monitors are added. This observation can be attributed to the fact that the ASes with higher degree contributes towards the addition of most links and as the number of monitors increase, the redundancy comes into play, and a lesser number of new links are discovered with each monitor addition. This experiment failed to provide the number of monitors (selected using the degree-based scheme) required to achieve 100%, and further investigation could be performed for the same. However, the result could be used to explain the reason why the random-based algorithm performed better than the degree-based algorithm.

The maximum coverage possible by selecting 30% of the ASes as monitors are in the range of 80% to 90%. For approximately 10% of the ASes acting as the BGP monitors, the monitors selected by random based scheme provide a link-coverage of over 75%, whereas the monitors selected by degree based scheme provide up to 70% link-coverage. Therefore, it is possible that the size of the monitor set is a factor which is influencing the link-coverage of the degree-based monitor set. Additionally, due to smaller graph sizes the likelihood of random-based algorithm selecting a better combination is high and, thus the random-based monitor selection performed better than the degree based selection. However, for the Internet, the degree-based algorithm should perform better than that of the random-based algorithm as there would be a higher number of bad monitor combinations which the random-based scheme would pick with higher probabilities.

Greedy-link based algorithm is always the best choice for monitor selection if the routing information for each AS is available. This fact is exploited again to discover the monitor set size to achieve 100% link-coverage. The methodology explained above to create a plot of link-coverage against the monitor set size for the degree-based scheme is reused. Additionally, the constraint of selecting at most 33% ASes as monitors is relaxed, and the evaluation is stopped only when 100% link-coverage is achieved. Figure 5.4b shows a plot of link-coverage against the monitor set size for the greedy-link based scheme. The lines are plotted to show the average link-coverage various graph sizes, and the deviation from average value is shown in the shaded region.

Figure 5.4b shows that the link-coverage rises sharply and then the gradient of the line decreases but not as abruptly as degree based (figure 5.4a). This means that the monitors determined using the greedy-link based scheme are continuously discovering new links, and with every new monitor added to the monitor set, the number of newer links discovered are higher than that discovered by degree-based monitor selection. There is not much deviation in the values from different graph instances of the same size as the deviation shown in the shaded region is quite small and would imply that the greedy-link based monitor selection adapts to the topological information and selects the AS with higher number of new links as a monitor. The maximum number of monitors required to achieve 100% link-coverage for a graph with 600 ASes is between 200 and 250, and the maximum monitor size is between 35% and 40%.

The results from the experiments indicate that at least 35% of the ASes must be chosen as monitors to achieve 100% link-coverage. This is only possible through the greedy-link based monitor selection scheme as the degree-based monitor selection scheme failed to achieve 100% link-coverage even with the monitor set size of 33% and the trend of the link-coverage gradient for degree-based monitor selection does not seem

promising. For the recent Internet size of 62775 ASes, approximately 18832 ASes must be selected as the monitoring nodes. Degree-based monitor selection scheme could not perform better than greedy-link based monitor selection scheme, and for a monitor set size same as that selected by greedy-link based algorithm, degree-based monitor selection could provide only 80% link-coverage with a nearly similar number of monitors.

5.4. Impact of real-world conditions on peering-degree based monitor selection

In this section, the impact of limited network visibility or the completeness of the graph topology information and the impact of incorrectness in the AS relationship on peering-degree based monitor placement scheme are discussed. Section 4.5 made assumptions under which the experiments are performed. The assumptions are valid for the toy graphs, but there is a possibility that assumptions related to the topology of the Internet, may be violated by limited visibility of the Internet (number of AS links observed) or incorrectness in the inferred AS relationships.

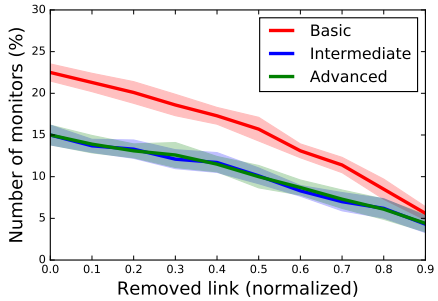
The complete topological information of the Internet is not available through BGP data, and multiple AS links are discovered using the alternate measurement sources like ARK data [4]. The ARK data is collected by performing traceroute measurements from merely 319 locations, and therefore, it would be safe to assume that only a fraction of AS links are discovered using the ARK data and many links are yet to be discovered. Thus, there is a possibility of violation of the assumption which states that complete information about the AS graph is available.

Their business relationship determines the AS relationship between any two neighboring ASes, and this information is not readily available in the public domain. Co-operation from the network operators to make the information available would prove to be highly beneficial. However, collecting this information from ASes individually is a tedious task, and the ASes may not disclose such information, considering any unforeseen negative impact on their business interests. To infer AS relationships, we rely on the heuristics proposed in [25, 37], and the heuristics are shown to be accurately inferring the relationships. However, since the relationships are inferred, it is possible that the inferences may be inaccurate.

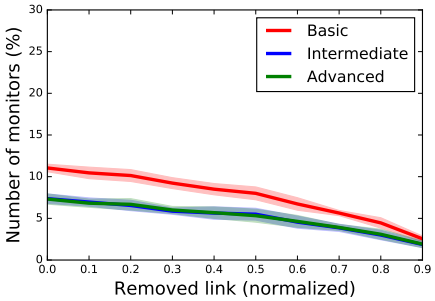
The monitor selection algorithm relies on the correctness of this information and therefore, it is essential to evaluate different versions of the algorithm under limitations as mentioned earlier.

Impact of limited visibility

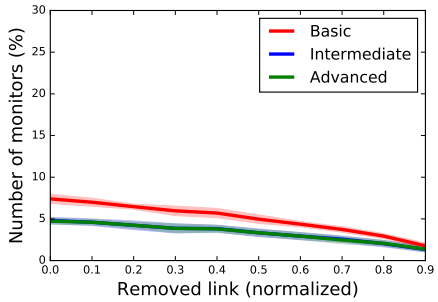
To evaluate the impact of limited visibility the monitor sets are determined by using different versions of the peering-degree based monitor selection after removing a fraction of links. The effect of limited visibility due to provider-to-customer links and peer-to-peer links are studied separately. Firstly, an AS graph with the desired number of ASes is created, and then a list of edges with peer-to-peer links is extracted. Next, we assign an equal probability of selection to all peering links and select a fraction of links from the peer-to-peer links. The fraction of peering links that are selected are removed from the AS graph, and the monitor sets are determined using the three versions of peering-degree based monitor selection.



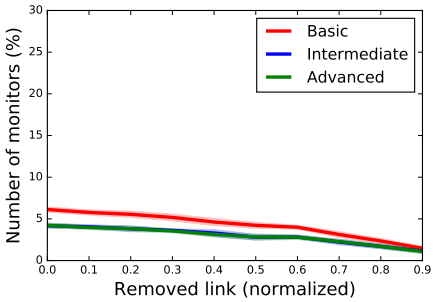
(a) The graph with 100 ASes.



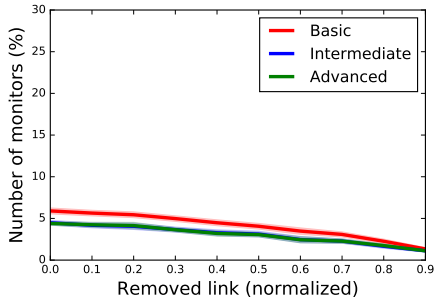
(b) The graph with 200 ASes.



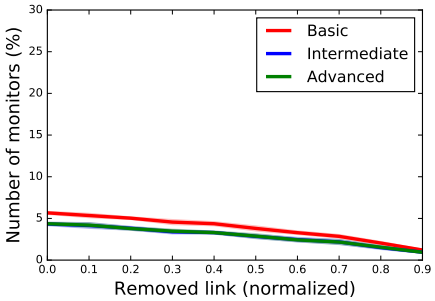
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



(e) The graph with 500 ASes.

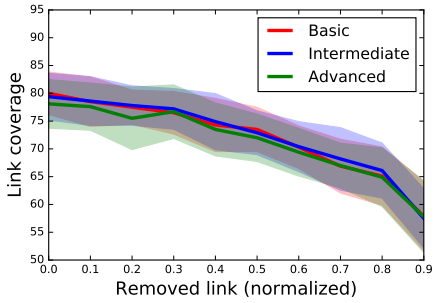


(f) The graph with 600 ASes.

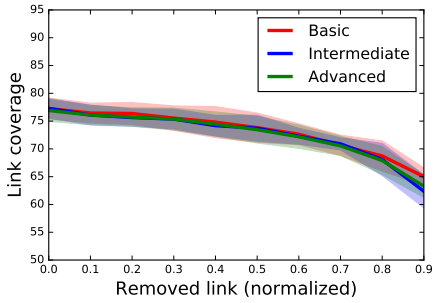
Figure 5.5: Impact on the number of monitors determined by the three versions of the peering-degree based schemes developed in this thesis under limited visibility of peer-to-peer links for various graph sizes.

The plots shown in figure 5.5 are the plots generated for number of monitors (in %) on y-axis, against the fraction of removed links (peer-to-peer) on x-axis. The plots show that as more peering links are removed, the size of the monitor set is reduced. In the search phase of the algorithm, ASes with peer-to-peer links are added into the list of possible monitors. As more and more peering links are removed during the monitor selection, less number of ASes are eligible for monitor selection, and there is a decline in the total number of monitors produced.

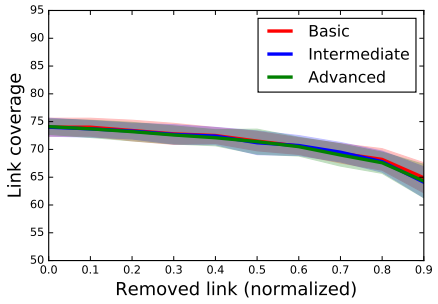
The link-coverage is dependent on two aspects of monitor placement- the location of the monitors and the number of monitors. Under the influence of limited visibility of peering links, the number of monitors and the location of monitors is affected and therefore, link-coverage must deteriorate as more number of peering links are removed during monitor selection, and the plots shown in figure 5.6 confirms the same. The plots shown in figure 5.6 are the plots of link-coverage against the fraction of links (peer-to-peer) removed. The link-coverage of the three versions of peering-degree based monitor selection is similar even though the number of monitors determined by intermediate and advanced versions is less. This confirms the superiority of intermediate and advanced version over the basic version in terms of monitor set size, which is achieved at the expense of computational complexity.



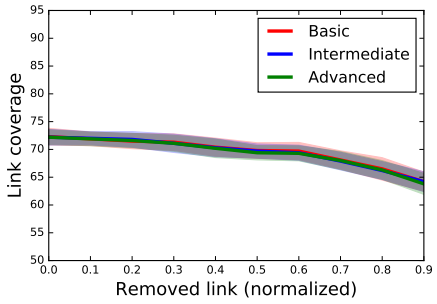
(a) The graph with 100 ASes.



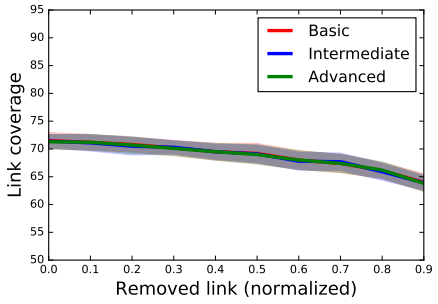
(b) The graph with 200 ASes.



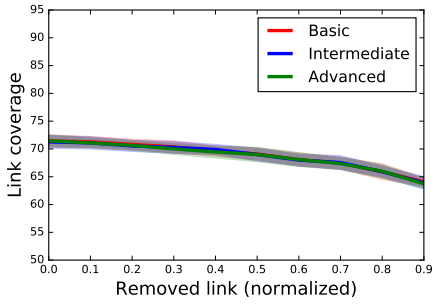
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



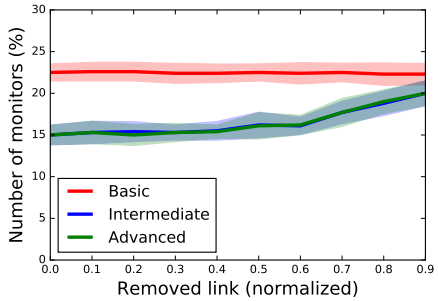
(e) The graph with 500 ASes.



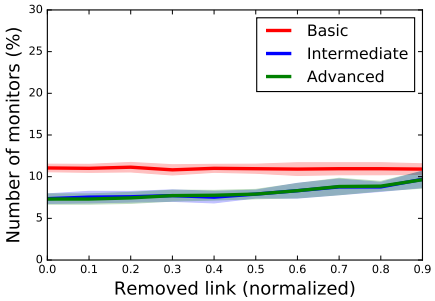
(f) The graph with 600 ASes.

Figure 5.6: Impact on the link-coverage of the monitors determined by the three versions of the peering-degree based schemes developed in this thesis under limited visibility of peer-to-peer links for various graph sizes.

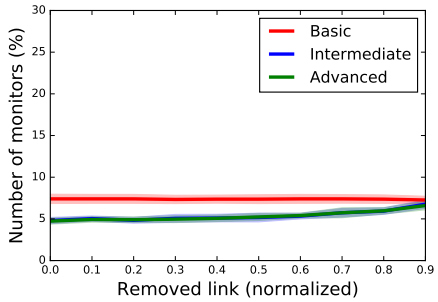
Now, we look at the impact of limited visibility of provider-to-customer links on the three versions of peering-degree based monitor selection scheme. The monitor sets are determined for the three versions of the algorithm proposed in this thesis after removing a fraction of provider-to-customer links. An AS graph with the desired number of ASes is generated, and then a list of provider-to-customer links is extracted for the AS graph. Each link is assigned the equal probability of selection, and a fraction of links are selected from this list which are then removed before performing the monitor selection on the modified AS graph.



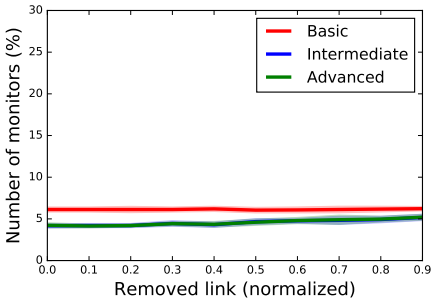
(a) The graph with 100 ASes.



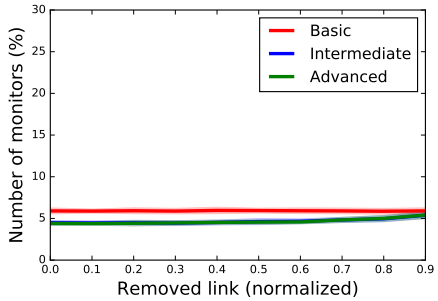
(b) The graph with 200 ASes.



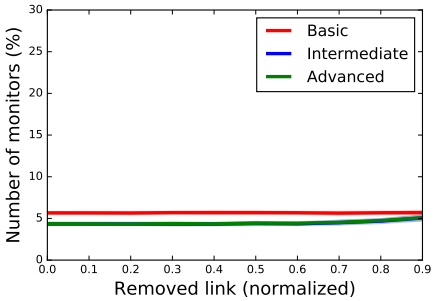
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



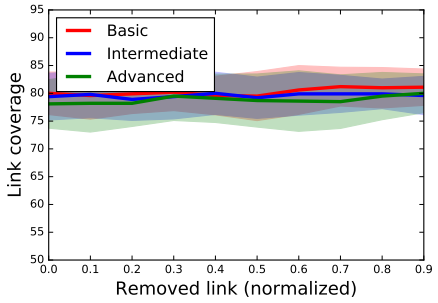
(e) The graph with 500 ASes.



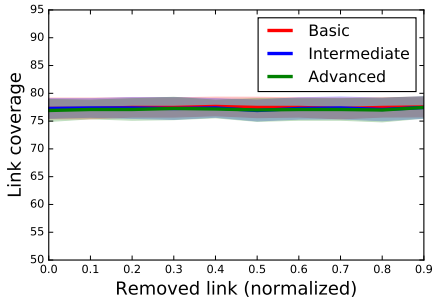
(f) The graph with 600 ASes.

Figure 5.7: Impact on the number of monitors determined by the three versions of the peering-degree based schemes developed in this thesis under limited visibility of provider-to-customer links for various graph sizes.

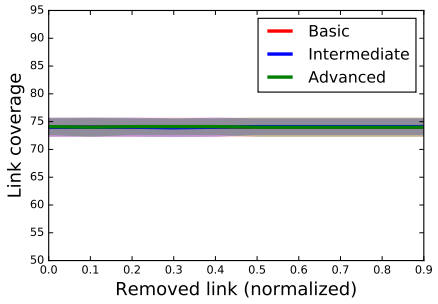
The plots shown in figure 5.7 are the plots showing the number of monitors on the y-axis and the fraction of removed links (provider-to-customer) on the x-axis. The number of monitors determined by the intermediate and advanced versions of the proposed algorithm increases as more provider-to-customer links are removed. This is because the provider-to-customer links play an important role during the elimination process. In the intermediate and advanced version of the proposed algorithm, provider-to-customer links are used to further reduce the number of monitors by placing a monitor deeper within the customer cone of an AS with peer-to-peer links, and in the absence of provider-to-customer links such elimination becomes less effective and thus we observe an increase in the number of monitors. In the basic version, provider-to-customer links are ignored entirely and thus, their elimination play no role in determining the monitor set.



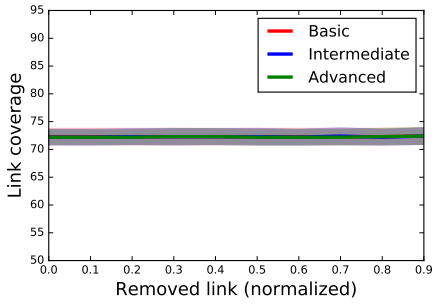
(a) The graph with 100 ASes.



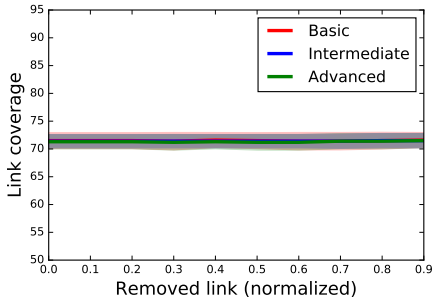
(b) The graph with 200 ASes.



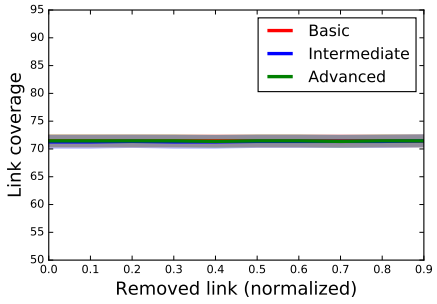
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



(e) The graph with 500 ASes.



(f) The graph with 600 ASes.

Figure 5.8: Impact on the link-coverage of the monitors determined by the three versions of the peering-degree based schemes developed in this thesis under limited visibility of provider-to-customer links for various graph sizes.

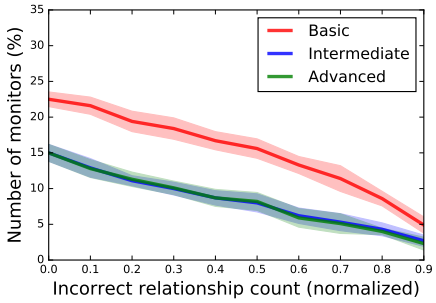
The plots shown in figure 5.8 show that the link-coverage remains unaffected by visibility limitations of provider-to-customer links on the larger graph sizes. Under the influence of limited visibility of provider-to-customer links, the number of monitors has increased for the intermediate and advanced version. This only shows that the schemes have failed to reduce the monitor set size and is as expected.

Impact of incorrectness in AS relationships

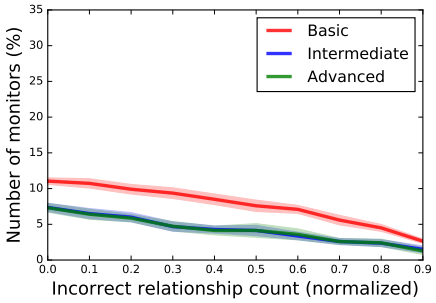
In this section, we discuss the impact of incorrectness in the AS relationships on the monitor selection process. It is possible that a peer-to-peer link is incorrectly inferred as a provider-to-customer link and a provider-to-customer link is inferred as a peer-to-peer link. To study this effect, an AS graph is produced with the desired number of ASes and a list of peer-to-peer links is generated for the AS graph. Next, all the peering links are assigned the equal probability of selection and a fraction of peering links are selected. The relationship for these links is updated to provider-to-customer and the monitor sets are determined using the three versions of the monitor selection algorithm proposed in this thesis.

The plots shown in figure 5.9 are the plots where the y-axis is representing the number of monitors and the x-axis is representing the fraction of peering links converted into provider-to-customer links during the monitor selection. The plots show that as more peering links are converted into provider-to-customer links,

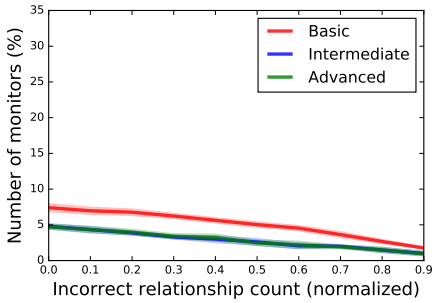
the number of monitors selected is reduced. This is because the search and the elimination phases of the algorithm are impacted by this action. As more peering links are converted into provider-to-customer links, the ASes with a lower peering degree would lose their peering degree ranks first and are not selected by the search phase. Moreover, the elimination phase is able to eliminate more monitor ASes by identifying them as providers of different ASes, which under true relationship would have remained in the monitor set. Thus, the size of the monitor set decreases with an increase in the conversion of peering links into provider-to-customer links.



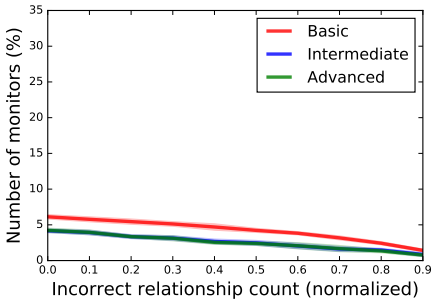
(a) The graph with 100 ASes.



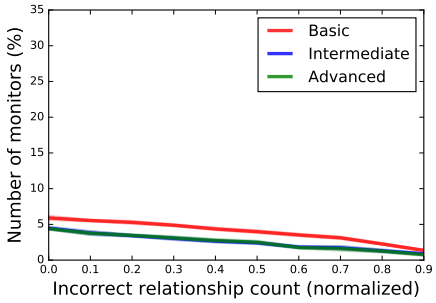
(b) The graph with 200 ASes.



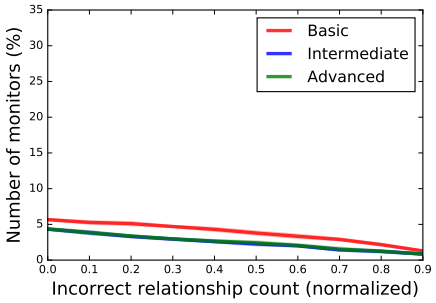
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.

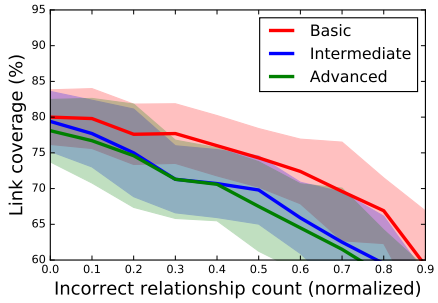


(e) The graph with 500 ASes.

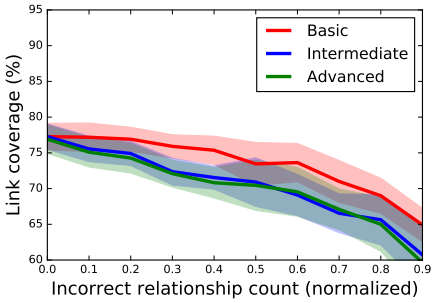


(f) The graph with 600 ASes.

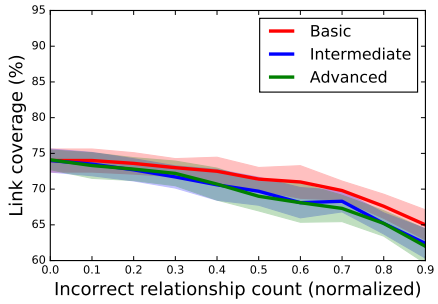
Figure 5.9: Impact on the number of monitors determined by the three versions of the peering-degree based schemes developed in this thesis when peer-to-peer links are incorrectly inferred as provider-to-customer links.



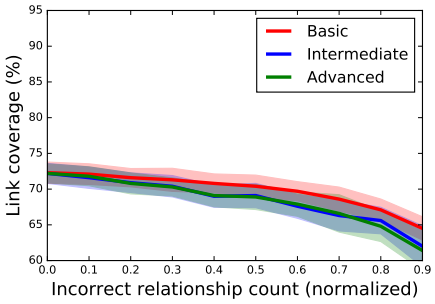
(a) The graph with 100 ASes.



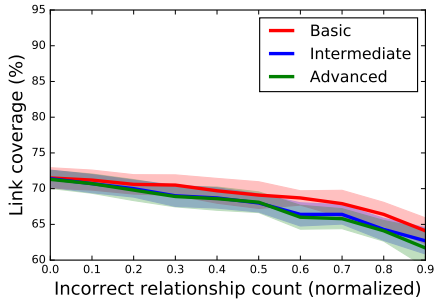
(b) The graph with 200 ASes.



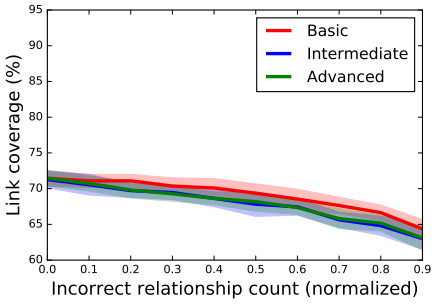
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



(e) The graph with 500 ASes.

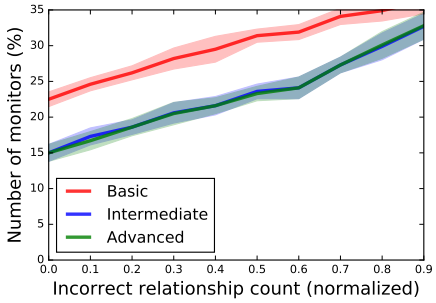


(f) The graph with 600 ASes.

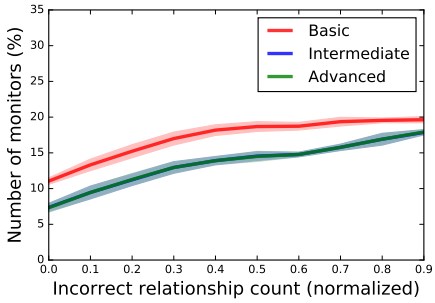
Figure 5.10: Impact on the link-coverage of the monitors determined by the three versions of the peering-degree based schemes developed in this thesis when peer-to-peer links are incorrectly inferred as provider-to-customer links.

The link-coverage of newly determined monitor sets is also impacted due to incorrectness in AS relationships. The plots shown in figure 5.10 are the plots of link-coverage against the fraction of incorrectly assigned peer-to-peer relations. These plots show that the link-coverage is deteriorating as more peer-to-peer links are converted into provider-to-customer links during the monitor selection process. This reduction in link-coverage could be attributed to the fact that the size of the monitor set is reduced. Additionally, by incorrectly tagging a peer-to-peer link as the provider-to-customer link could potentially eliminate an AS from the monitor set by incorrectly identifying the AS as a provider of other monitors. Thus a reduction in link-coverage is due to smaller monitor sets and ineffective monitor placement.

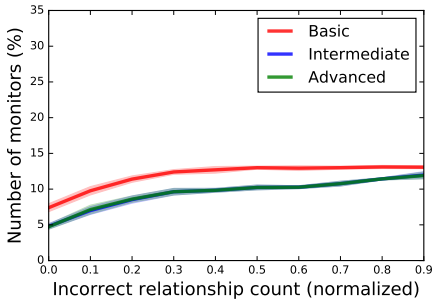
Next, we discuss the impact of incorrectly inferring a provider-to-customer link as a peer-to-peer link. The process of elimination is similar to the one discussed before. An AS graph with the desired number of ASes is produced, and a list of provider-to-customer links is generated. Each link in this list is assigned an equal probability of selection, and a fraction of provider-to-customer links are selected and converted into peer-to-peer links. The monitor selection is performed using the three versions of peering-degree based monitor selection scheme and the monitor sets produced are evaluated for their size and link-coverage.



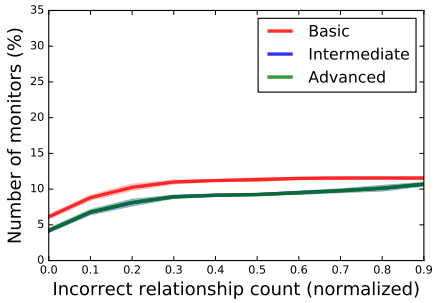
(a) The graph with 100 ASes.



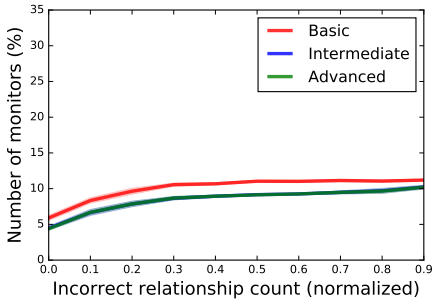
(b) The graph with 200 ASes.



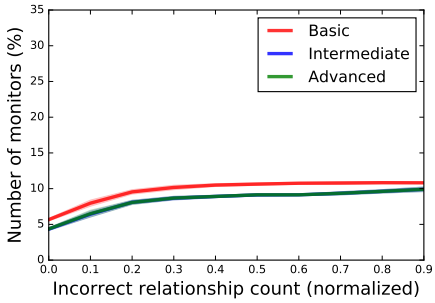
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



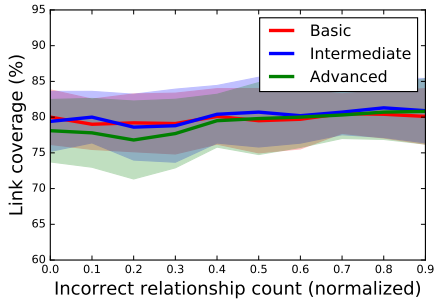
(e) The graph with 500 ASes.



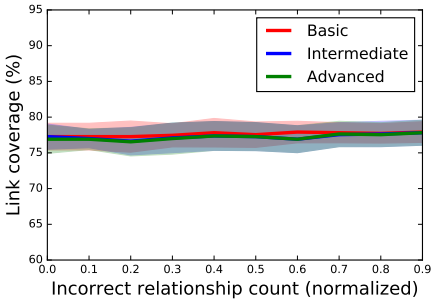
(f) The graph with 600 ASes.

Figure 5.11: Impact on the number of monitors determined by the three versions of the peering-degree based schemes developed in this thesis when provider-to-customer links are incorrectly inferred as peer-to-peer links.

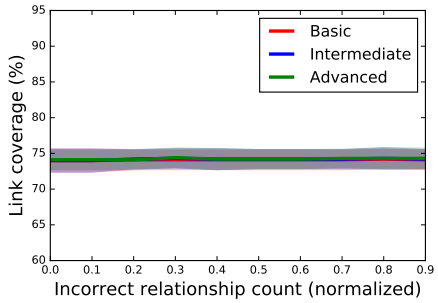
The plots shown in figure 5.11 are the plots of the number of monitors against the incorrect relationship count for provider-to-customer links. The plots show that as more provider-to-customer links are converted into peering links, the size of the monitor set increases. This is because, when a provider-to-customer link is converted into a peer-to-peer link, then the ASes which are the endpoints of such links are selected in the search phase. Also, by converting a provider-to-customer link into peer-to-peer links, the elimination phase becomes less effective as it would fail to determine correctly if an existing monitor is a provider of a newly identified monitor in the intermediate and advanced versions of the scheme. Thus, the size of the monitor set increases with an increase in the number of provider-to-customer links that are tagged as peer-to-peer links.



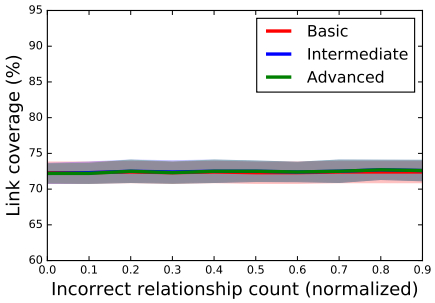
(a) The graph with 100 ASes.



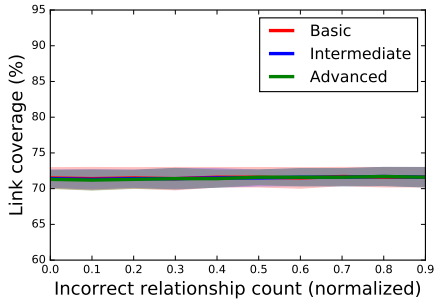
(b) The graph with 200 ASes.



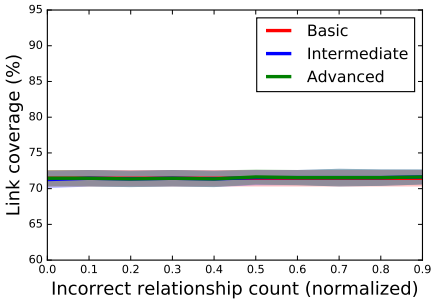
(c) The graph with 300 ASes.



(d) The graph with 400 ASes.



(e) The graph with 500 ASes.



(f) The graph with 600 ASes.

Figure 5.12: Impact on the link-coverage of the monitors determined by the three versions of the peering-degree based schemes developed in this thesis when provider-to-customer links are incorrectly inferred as peer-to-peer links.

The increase in the size of the monitor set would imply that the link-coverage under the influence of incorrectly identifying a provider-to-customer link as a peer-to-peer link should not be affected. The plots shown in figure 5.12 are the plots of link-coverage against the incorrect relationship count (provider-to-customer) plotted for the graphs of various sizes. These plots confirm that the link-coverage remains unchanged as more number of provider-to-customer links are incorrectly inferred to as the peer-to-peer links. This is because the larger monitor set contains the optimal monitor set as well and includes ASes within the monitor set, which could have been eliminated successfully without reducing the link-coverage.

The plots in figure 5.5, 5.7, 5.9 and 5.11 reveals two key aspects regarding the impact of the graph size on the number of monitors determined by the three versions of peering-degree based monitor selection scheme. Firstly, there is a reduction in the difference between the number of monitors selected by the basic and intermediate schemes of the algorithm reduces as the size of graph increases, and secondly, the number of monitors determined by each algorithm becomes similar when the topological information is nearly 90% inaccurate. Both aspects could be explained by the fact that the underlying mechanism of the intermediate and advanced versions of the scheme is similar to the basic version with an additional possibility of elimination of monitors whose customer ASes are in the monitor set. The growth of the Internet happens at the edge, and the core remains fairly stable (in the number of ASes) and the number of monitor ASes that could be eliminated due to the selection of a monitor within the customer cone of such ASes is limited. Thus, the

increase in graph size results in the reduction of difference in the number of monitors selected by the basic, and intermediate and advanced versions. Also, under the influence of incorrect topological information, the advantage of intermediate and advanced versions is lost, and they behave like the basic version only and thus the resulting monitor set sizes become equal as the incorrectness in topological information is increased.

The incorrectness in topological information is induced by a random link selection, and ten experiments are conducted on different graph sizes, and the deviation from average values is also presented in the plots shown as the shaded region in figure 5.5 to 5.12. It could be seen that there is not much deviation in the plots showing the number of monitors for different graph sizes. However, in the plots that show the link-coverage against graph sizes, the deviation is prominent for smaller graph sizes and reduces with the graph size. This is because of two reasons, firstly, in the small graph sizes, the possibility of selecting a link with high impact is higher in smaller graph sizes, and thus there is a huge deviation. Secondly, the link-coverage values are normalized to the total number of links in the graph, and as the graph size increases, the number of links also increases, and thus a change in link-coverage in smaller graphs would appear to be magnified. In the best-case scenario peering-degree based monitor selection scheme would provide a link-coverage of 83% with 12% of the ASes acting as the monitors and in the worst-case scenario of incorrect topological information would provide at least 55% link-coverage with a monitor set containing 5% ASes. The intermediate and advanced versions of peering-degree based monitor selection scheme would always be the best choice if the size of monitor set is a constraint because whenever possible it would ensure that the number of monitors is equal or less than that determined by the basic version with the link-coverage that is similar to that of the basic version.

6

Conclusion

In this thesis, schemes for monitor placement for monitoring BGP traffic are evaluated, and a new scheme-peering-degree based for monitor placement, is discussed. From the experiments, it is concluded, that while the greedy-link based algorithm can achieve 100% link coverage with 35% of the ASes acting as monitoring nodes, degree-based could only reach up to 80% with same monitor set sizes.

The peering-degree based monitor selection scheme, proposed in this thesis, when tested for smaller Internet-like networks successfully achieved,

1. Reduction in the number of monitoring nodes. When compared with existing state-of-the-art algorithms for monitor set selection, peering-degree based approach requires only 4% to 8% of the ASes to be selected as monitoring nodes and provides highest link coverage (next to greedy-link based algorithm).

2. Reduction in the information overlap between the two monitors. The source of overlapping information is the adjacent ASes acting as monitors. Since the algorithm eliminates adjacent ASes, thus reducing redundancy Chen et al..

The peering-degree based monitor placement scheme proposed in this thesis outperforms degree-based and random monitor selection schemes. This is because the algorithm targets ASes with peer-to-peer relationships, and peer-to-peer relations usually remain hidden to top-level ASes and can only be observed by collecting inter-AS routes either at one of the ASes that form the end-points of a peering relation or at an AS which lies within the customer cone of respective ASes. Another way of discovering peer-to-peer links is to generate traceroutes between various source-destination pairs. However, measuring the Internet globally using traceroute is computationally hard.

The smallest prefix of IPv4 is /24 which is a collection of 255 hosts. Even if a single host is selected from /24 prefixes of the complete IPv4 address space, there would be 16581375 hosts and to measure traceroute for each host from every host would consume a considerable amount of time. The scheme for monitor placement proposed in this thesis eliminates adjacent ASes and provider ASes. This serves two functions; firstly it tries to reduce the monitor set size and secondly, eliminates the sources of redundant information and thereby reducing the size of redundant BGP data collected.

Greedy-link based monitor selection scheme can determine the number of monitoring ASes required to achieve 100% link-coverage. This is because the greedy link algorithm assumes that inter-AS routes from all ASes are available at the time of monitor selection and hence can discover new links with each monitor it picks. This is only possible in smaller Internet-like graphs as RIBs can be simulated. For the Internet at large, this is challenging, because it is difficult to foresee link coverage for individual ASes at the time of monitor selection. RIB dump from each AS on the Internet is required by greedy-link based algorithm either through simulation or directly from ASes.

Co-operation is required from the ASes to obtain RIBs and ASes are not enthusiastic about sharing this information. BGP data can be generated using simulators; however, there are various factors like local preferences, AS-relationships, which influence the BGP decision process and the resulting BGP tables generated

through simulations may not reflect the appropriate snapshot of the Internet. Moreover, due to the size of the Internet, memory and performance constraints associated with the simulator would make it unfeasible to simulate BGP traffic on the Internet.

Approximately, 45% of the AS links are visible to every AS, however, due to the limited number of paths available on the Internet between any source-destination pair, there would always be a certain level of AS link information overlap between any two ASes, and thus the performance of random and degree-based schemes is also not entirely worse. The algorithm proposed in this thesis is an extension of the degree-based algorithm itself. The algorithm acknowledges that peer-to-peer links are more valuable than provider-to-customer links and monitoring at customer AS is more beneficial as the information obtained over a peer-to-peer link is only sent to customer ASes and has exploited this fact.

Inter-AS route monitoring has various advantages for all ASes. BGP data would help researchers understand the present inter-AS routing behavior and enable them to propose improvements in network architectures for the Internet and advance BGP protocol as well. Additionally, BGP data would improve the effectiveness of the reactive mechanisms to counteract the malicious activities which can lead to security breaches and instabilities on the Internet. However, not all ASes can become the monitoring ASes, and the ASes which would have the potential to act as monitoring points would have unnecessary overhead regarding network bandwidth and router's processing load, as their routers and communication media would be intermittently occupied with generating and transmitting the routing information to collectors.

The business objective of an AS is to provide network access to the customer ASes that pay for consuming their services and would not be willing to participate in the monitoring projects. Thus, incentives must be provided to the ASes that join the monitoring network by providing them with improved mechanisms to secure their interests. In the era of NSFNet, the predecessor of the Internet, mechanisms were in place that would raise the alarm in case any prefix was announced from an unexpected location and discard the suspicious announcement. The same can be implemented for all ASes, and corrective announcements are generated whenever any breach is suspected, and the BGP data can be used to generate such information [47]. In the absence of central authority which can govern the Internet, all ASes must come together to collaborate towards making the Internet safer for all users.

6.1. Contributions

The main contribution of this thesis are as follows:

1. Proposed a peering-degree based monitor selection scheme which gives over 70% link coverage with 4% to 8% of the ASes acting as monitoring ASes. Considering that 35% of the ASes must act as monitors to achieve 100% link coverage, this number is acceptable.

2. Determined that 35% of the ASes must act as inter-AS route monitors to achieve 100% link coverage.

3. Evaluation of peering-degree based monitor selection scheme under uncertainties of the real-world Internet such as incompleteness in graph information or incorrectness in AS relationship inferences.

4. Identified the short-fall of [43] based topology generator and proposed a behavior-based topology generation scheme which is shown to abide with power-law.

6.2. Future work

The monitor placement schemes discussed in this thesis can be used to select new monitors for collecting inter-AS routing information for real-world monitoring systems. The existing BGP data can be used to determine AS-relationships which in turn can be used to select new monitors. Also, the inter-AS routes are highly dynamic, and AS-relationships evolve with time. Hence there is a need for continuous review of placed monitors to ensure better results.

The advance version of the algorithm uses a naive approach to handle conflict information. A more sophisticated mechanism could be developed by associating the conflict information generated for each monitor selection and treating the information in order each time the conflict flag is raised.

100% link coverage is possible with nearly 35% of the ASes acting as monitoring points. Therefore, there is a need for more efficient methods of collecting and managing the BGP data. The focus must be given to machine learning based traceroute tools which can reduce the time in which traceroute data is collected which would aid BGP monitoring systems. Moreover, since the Internet is self-governing, behavioral studies can be performed to understand how the co-operation from ASes can be encouraged and develop a simplified process of acquiring BGP data. Moreover, as the quality of BGP data improves, it would assist in acquiring a deeper understanding of inter-AS interactions, and localization of instability, and hijack events and ways to localize events must be explored.

A

Algorithms

Algorithm 1 AS path to AS link conversion

```
1:  procedure PathToLinks(Path, AS)                    ▷ Path is AS Path as obtained from RIB of AS
2:      ASes ← Path.split(" ")                         ▷ Create a set of ASes by splitting the path
3:      ASes ← ASes.insert(0, AS)                      ▷ Inserting AS of which RIB is being used.
4:      while l < 2 do
5:          return Set()▷ Return empty set since no link can be extracted if number of ASes in AS path are less
    than 2
6:      l ← length(ASes)
7:      counter ← 1
8:      links ← Set()
9:      while counter < l do                           ▷
10:         links ← links⋃(ASes[counter − 1], ASes[counter])
11:         counter ← counter + 1
12:     return links                                   ▷ Set of links extracted from AS path
```

Algorithm 2 Graph Motif Exploration for AS-Graph

```
1: procedure GRAPHMOTIFEXPLORATION(G, k)                ▷ G is an AS-graph and k is the size of motif.
2:     motif ← Map(< Motif : Count >)
3:     nodes ← G.nodes()                                ▷ Create a set of nodes of the graph G
4:     while length(nodes) > 0 do                       ▷ Indefinite while loop
5:         node ← nodes.pop()
6:         neighbours ← G.neighbours(node)              ▷ Neighbours of node in graph G
7:         neighbours ← neighbours ∪ node
8:         combinations = getcombinations(neighbours, k)   ▷ getcombinations returns all possible
   combinations of elements in set neighbours of size k
9:         while length(combinations) > 0 do
10:            c ← combinations.pop()
11:            g ← G.sub − graph(c)                     ▷ G.sub-graph returns sub-graph of G containing nodes in set c
12:            while g.isconnected() == True do         ▷ execute if sub-graph is connected
13:                while motif.keys().find(g) == True do   ▷ searches is sub-graph is already identified
14:                    motif[g] ← motif[g] + 1          ▷ Increment count of motif g
15:                    break
16:                while motif.keys().find(g) == False do   ▷ searches is sub-graph is already identified
17:                    motif[g] ← 1                     ▷ Adding motif g to the motif map
18:                    break
19:        return motif                                 ▷ returns map of motif<motif:count>
20:
```

Algorithm 3 Random Based Monitor Selection Scheme

```
1: procedure RANDOMMONITORSELECTION(ASes, k)            ▷ List of ASes and size of monitor set required k
2:     l ← length(ASes)                                 ▷ Number of element in set ASes
3:     counter ← 0
4:     prob ← List()
5:     while counter < l do                             ▷ Create Probability table for each AS
6:         prob[counter] ← 1.0
7:         sum ← sum + 1.0
8:         counter ← counter + 1
9:     counter ← 0
10:    while counter < l do
11:        prob[counter] ← prob[counter]/sum            ▷ Normalizing Probability
12:        counter ← counter + 1
13:    monset = RandomChoice(ASes, prob , k) ▷ RandomChoice select k elements from set of ASes based
   on probability list prob
14:    return monset                                    ▷ Set of Randomly Selected Monitors of size k
   =0
```

Algorithm 4 Obtaining Degree Map for an AS Graph

```
1: procedure GETDEGREEMAP(ASGraph, r)  ▷ ASGraph is a directed graph with labelled AS-relations and r is the relationship for which degree map is required
2:     degreemap = Dict()
3:     counter ← 0
4:     nodes ← ASGraph.nodes()
5:     while counter < length(nodes) do
6:         neighbours ← ASGraph.neighbours(nodes[counter])
7:         degreemap[nodes[counter]]['provider-to-customer'] ← 0
8:         degreemap[nodes[counter]]['customer-to-provider'] ← 0
9:         degreemap[nodes[counter]]['sibling-to-sibling'] ← 0
10:        degreemap[nodes[counter]]['peer-to-peer'] ← 0
11:        degreemap[nodes[counter]]['ALL'] ← length(neighbours)
12:        counter2 ← 0
13:        while counter2 < length(neighbours) do
14:            relation ← ASGraph.getEdgeData(nodes[counter], neighbours[counter2], 'relationship')
15:            degreemap[nodes[counter]][relation] += 1
16:            counter2 ← counter2 + 1
17:        counter ← counter + 1
18:        degreemapfinal ← dict()
19:    counter ← 0
20:    while counter < length(nodes) do
21:        degreemapfinal[nodes[counter]][r] ← degreemap[nodes[counter]][r]
22:        counter ← counter + 1
23:        return degreemapfinal  ▷ Returns out-degree for each AS for relationship r
24:
```

Algorithm 5 Degree Based Monitor Selection Scheme

```
1: procedure DEGREEMONITORSELECTION(ASGraph, k)  ▷ ASGraph is a directed graph with labelled AS-relations and size of monitor set required k
2:     monset = Set()
3:     counter ← 0
4:     DegreeMap ← getDegreeMap(ASGraph, 'ALL')
5:     DegreeMap ← Sort(DegreeMap)
6:     while counter < k do
7:         monset ← monset ⋃ DegreeMap[counter]
8:         counter ← counter + 1
9:     return monset  ▷ Set of Degree Based Scheme Selected Monitors of size k =0
```

Algorithm 6 Greedy-link based Monitor Selection Scheme

1: **procedure** GREEDYLINKMONITORSELECTION($ASLinkMap, k$) ▷ AS link map is a sorted list of tuple (AS, set of links observed at the AS) and size of monitor set required k
2: $l \leftarrow length(ASLinkMap)$
3: $monset = Set()$
4: $counter \leftarrow 0$
5: $VisibleLinks \leftarrow \phi$ ▷ Empty set
6: $map \leftarrow ASLinkMap$
7: **while** $counter < k$ **do**
8: $monset \leftarrow monset \bigcup ASLinkMap[0]['AS']$
9: $counter \leftarrow counter + 1$
10: $VisibleLinks \leftarrow VisibleLinks \bigcup ASLinkMap[0]['Links']$
11: $ASLinkMap.pop(0)$ ▷ removing AS from ASLinkMap which has been included in the monitor set
12: $i \leftarrow 0$
13: **while** $i < length(ASLinkMap)$ **do**
14: $ASLinkMap[i]['Links'] \leftarrow ASLinkMap[i]['Links']VisibleLinks$ ▷ Removing visible links for respective AS in ASLinkMap
16: $i \leftarrow i + 1$
17: $ASLinkMap \leftarrow Sort(ASLinkMap)$ ▷ ASLinkMap is sorted based on the number of links
18: **return** $monset$ ▷ Set of Degree Based Scheme Selected Monitors of size k

Algorithm 7 Graph Generator

```
1: procedure GRAPHGENERATOR(n)                                   ▷ n is the number of ASes
2:     nodes ← Set()
3:     counter ← 0
4:     while counter < n do
5:         counter ← counter + 1
6:         nodes ← nodes ⋃ counter
7:     t1 = max(3, integer(n * 0.0003)
8:     t2 = max(10, integer(n * 0.007)
9:     t3 = max(30, integer(n * 0.0927)
10:    t4 = n − (t1 + t2 + t3)
11:    tier1 = nodes[0 : t1 : 1]                                 ▷ Stage 1- Tier Assignment
12:    tier2 = nodes[t1 : t1 + t2 : 1]
13:    tier3 = nodes[t1 + t2 : t1 + t2 + t3 : 1]
14:    tier4 = nodes[t1 + t2 + t3 : n : 1]
15:    G ← DirectedGraph()
16:    G.addnodesfrom(nodes)
17:                                                              ▷ Stage 2- Adding Neighbouring Peers
18:    AddNeighbours(G, tier1, tier1, list(range(0 to length(tier1) − 1)), k = −1, relationship =′
       peer − to − peer′)
19:    AddNeighbours(G, tier2, tier2, list(range(0 to (int(0.6 * length(tier2)) − 1))), k =
       −1, relationship =′ peer − to − peer′)
20:    AddNeighbours(G, tier3, tier3, list(range(1 to (int(0.3 * length(tier3)) − 1))), k =
       3, relationship =′ peer − to − peer′)
21:                                                              ▷ Stage 3- Adding Neighbouring Providers
22:    AddNeighbours(G, tier2, tier1, [1,2], k = 1, relation =′ provider − to − customer′)   ▷
       AddNeighbours adds neighbours to the graph G
23:    AddNeighbours(G, tier3, tier1, [0,1], k = 4, relation =′ provider − to − customer′)
24:    AddNeighbours(G, tier3, tier2, [1,2,3], k = 2, relation =′ provider − to − customer′)
25:    AddNeighbours(G, tier4, tier2, [0,1], k = 4, relation =′ provider − to − customer′)
26:    AddNeighbours(G, tier4, tier3, [1,2,3], k = 5, relation =′ provider − to − customer′)
27:    counter ← 1
28:    while counter ≤ n do
29:        while length(G.getneighbours(counter) == 0 do
30:            G.removenode(counter)
31:            break
32:        counter ← counter + 1
33:    prefixASAssoiation ← prefixes(tier4)      ▷ Prefix Assignment. prefixes returns a dictionary of
       prefix-AS association
34:    return G                                  ▷ Returns AS-graph with link labels of size k
```

Algorithm 8 Add Neighbours

```
1: procedure ADDNEIGHBOURS(G, list1, list2, list3, k, r)        ▷ G is a directed graph
2:                                  ▷ list1 contains list of ASes to which neighbours will be added from list2
3:                                  ▷ list3 will determine number of neighbours an AS can have
4:     ▷ k is used to control the probability with which selection of how many neighbours can be added is done
5:                                  ▷ r is the relationship that will be assigned to the link between any two ASes.
6:   multiProvider ← List()
7:   counter ← 0
8:   probSum ← 0
9:   while counter < length(list3) do
10:    prob ← exp(−1 ∗ k ∗ list3[counter])
11:    probSum ← probSum + prob
12:    multiProvider ← multiProvider.append(prob)
13:    counter ← counter + 1
14:  counter ← 0
15:  while counter < length(list3) do
16:    multiProvider[counter] ← multiProvider[counter]/probSum
17:    counter ← counter + 1
18:  counter ← 0
19:  while counter < length(list1) do
20:    providerCount ← RandomChoice(list3, multiProvider, 1)        ▷ Returns a list of length 1 of elements from list3 based on probability provided in multiProvider
21:    counter ← counter + 1
22:    index ← 0
23:    removed ← False
24:    customer ← tier1[counter]
25:    while list2.find(customer) == True do
26:      removed ← True
27:      index ← list2.indexof(customer)
28:      list2.remove(customer)
29:      break
30:    neighbourCount ← length(list2)
31:    neighbourProb ← Set()
32:    while neighbourCount > 0 do
33:      prob ← 1.0
34:      neighbourCount ← neighbourCount − 1
35:      probSum ← probSum + prob
36:      neighbourProb ← neighbourProb.insert(0, prob)
37:    neighbourCount ← length(list2)
38:    while neighbourCount > 0 do
39:      neighbourCount ← neighbourCount − 1
40:      neighbourProb[neighbourCount] ← neighbourProb[neighbourCount]/probSum
41:    neighbours ← RandomChoice(list2, neighbourProb, neighbourcount)        ▷ Returns a list of length neighbourcount of elements from list2 based on probability provided in neighbourProb
42:    counter2 ← 0
43:    while counter2 < length(neighbours) do
44:      G.addlink(neighbours[counter2], customer, r)
45:      G.addlink(customer, neighbours[counter2], InverseRelation(r))
46:      counter2 ← counter2 − 1
47:    while removed == True do
48:      list2.insert(index, customer)
49:      break
50:    counter ← counter − 1
51:    return G
52:
```

B

Complete Parameter list for Topology Generator Presented in Section 4.3

Tier	**Graph Size**					
	100	200	300	400	500	600
1	3	3	3	3	3	3
2	10	10	10	10	10	10
3	30	30	30	37	46	55
4	57	157	257	350	441	532

Table B.1: Number of ASes in each tier.

N	**Graph Size**					
	100	200	300	400	500	600
0	0	0	0	0	0	0
1	0.9933	0.9933	0.9933	0.9933	0.9933	0.9933
2	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0

Table B.2: Probability of having n-providers for a tier-4 AS from tier-3 ASes ($P_N(k = 5)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0.982	0.982	0.982	0.982	0.982	0.982
1	0.018	0.018	0.018	0.018	0.018	0.018

Table B.3: Probability of having n-providers for a tier-4 AS from tier-2 ASes ($P_N(k = 4)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0	0	0	0	0	0
1	0.8668	0.8668	0.8668	0.8668	0.8668	0.8668
2	0.1173	0.1173	0.1173	0.1173	0.1173	0.1173
3	0.0159	0.0159	0.0159	0.0159	0.0159	0.0159

Table B.4: Probability of having n-providers for a tier-3 AS from tier-2 ASes ($P_N(k=2)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0.982	0.982	0.982	0.982	0.982	0.982
1	0.018	0.018	0.018	0.018	0.018	0.018

Table B.5: Probability of having n-providers for a tier-3 AS from tier-1 ASes($P_N(k=4)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0	0	0	0	0	0
1	0.7311	0.7311	0.7311	0.7311	0.7311	0.7311
2	0.2689	0.2689	0.2689	0.2689	0.2689	0.2689

Table B.6: Probability of having n-providers for a tier-2 AS from tier-1 ASes ($P_N(k=1)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0	0	0	0	0	0
1	0.9502	0.9502	0.9502	0.9502	0.9502	0.9502
2	0.0473	0.0473	0.0473	0.0473	0.0473	0.0473
3	0.0024	0.0024	0.0024	0.0024	0.0024	0.0024
4	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
5	0	0	0	0	0	0

Table B.7: Probability of having n-peers for tier-3 ASes ($P_N(k=3)$)

N	**Graph Size**					
	100	200	300	400	500	600
0	0.0117	0.0117	0.0117	0.0117	0.0117	0.0117
1	0.0317	0.0317	0.0317	0.0317	0.0317	0.0317
2	0.0861	0.0861	0.0861	0.0861	0.0861	0.0861
3	0.2341	0.2341	0.2341	0.2341	0.2341	0.2341
4	0.6364	0.6364	0.6364	0.6364	0.6364	0.6364

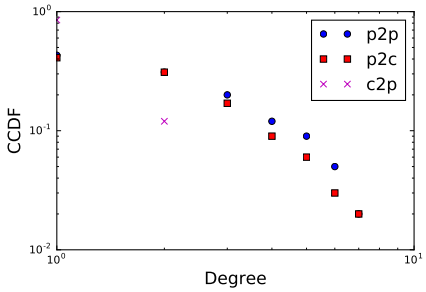
Table B.8: Probability of having n-peers for tier-2 ASes ($\mathbf{N}P_N(k=-1)$)

N	**Graph Size**					
	100	200	300	400	500	600
1	0.2689	0.2689	0.2689	0.2689	0.2689	0.2689
2	0.7311	0.7311	0.7311	0.7311	0.7311	0.7311

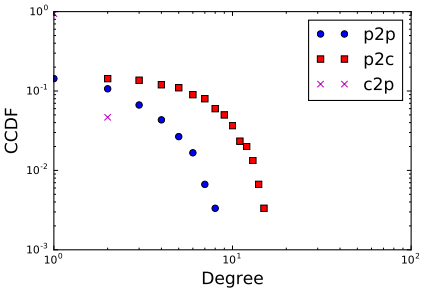
Table B.9: Probability of having n-peers for tier-1 ASes ($P_N(k=-1)$)

C

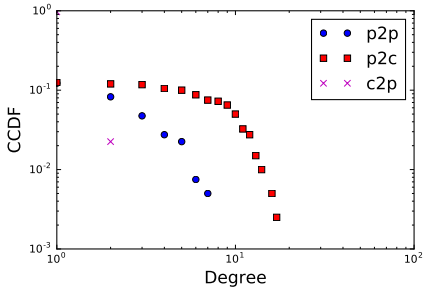
CCDF against Degree plots for few graphs used in Chapter 5



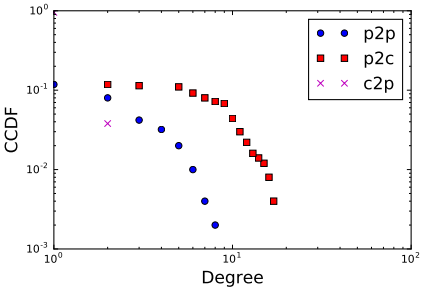
(a) CCDF of the topology generated for 100 ASes



(b) CCDF of the topology generated for 300 ASes



(c) CCDF of the topology generated for 400 ASes

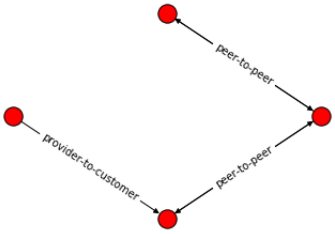


(d) CCDF of the topology generated for 500 ASes

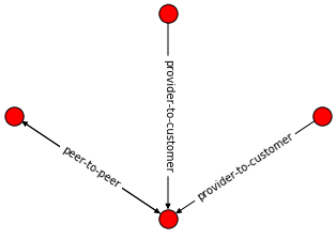
Figure C.1: CCDF of few AS graphs used in evaluation of monitor placement scheme

D

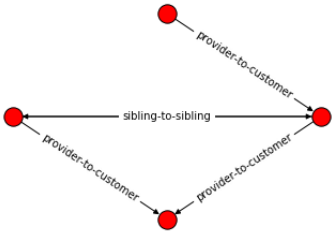
Additional Graph Motifs of the Internet Discovered using Algorithm 2



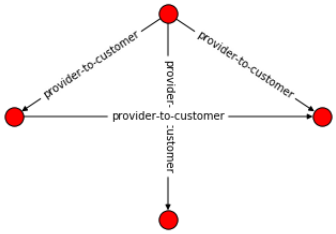
(a) Frequency of Occurrence 205



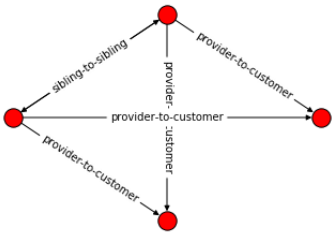
(b) Frequency of Occurrence 276



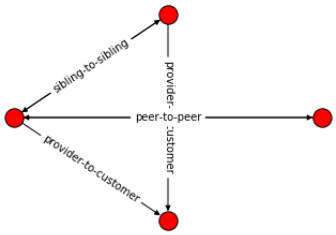
(c) Frequency of Occurrence 58



(d) Frequency of Occurrence 556

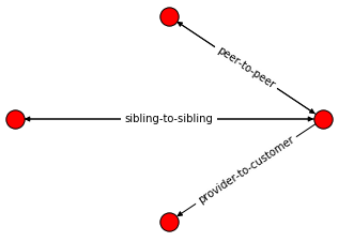


(e) Frequency of Occurrence 1537

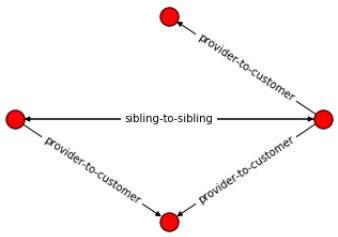


(f) Frequency of Occurrence 114

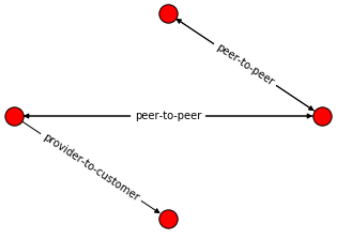
Figure D.1: Additional Graph Motifs from the Internet of size 4



(a) Frequency of Occurrence 116



(b) Frequency of Occurrence 232



(c) Frequency of Occurrence 189

Figure D.2: Additional Graph Motifs from the Internet of size 4 continuation

Bibliography

[1] Cisco Community, . URL https://community.cisco.com/t5/routing/bgp-rib/td-p/2580731.

[2] BGPMon, . URL https://www.bgpmon.io/.

[3] BGPStream.com, . URL https://bgpstream.com/.

[4] Archipelago (Ark) Measurement Infrastructure , . URL http://www.caida.org/projects/ark/.

[5] The CAIDA AS Relationships Dataset 01-10-2018, . URL http://www.caida.org/data/active/as-relationships/.

[6] MATHCelebrity.com. URL https://www.mathcelebrity.com/permutation.php?num=20000&den=20&pl=Combinations.

[7] Cover Picture. URL https://www.google.com/imgres?imgurl=https%3A%2F%2Fmedia.threatpost.com%2Fwp-content%2Fuploads%2Fsites%2F103%2F2015%2F07%2F07003346%2Fshutterstock_168037637.jpg&imgrefurl=https%3A%2F%2Fthreatpost.com%2Fbgp-security-alerts-coming-to-twitter%2F113843%2F&docid=vm4XOwFl1kZ_dM&tbnid=mUgnaF8lanqKfM%3A&vet=12ahUKEwiopKvq9s7eAhVJ1BoKHUo5DIA4rAIQMyhUMFR6BAgBEFU..i&w=1000&h=662&client=ubuntu&bih=630&biw=1301&q=BGP%20monitoring%20images&ved=2ahUKEwiopKvq9s7eAhVJ1BoKHUo5DIA4rAIQMyhUMFR6BAgBEFU&iact=mrc&uact=8.

[8] Internet Assigned Numbers Authority. URL https://www.iana.org/.

[9] RIPE- Routing Information Service (RIS). URL https://www.ripe.net/.

[10] BGP leaks and cryptocurrencies, . URL https://blog.cloudflare.com/bgp-leaks-and-crypto-currencies/.

[11] University of Oregon Route Views Project, . URL http://www.routeviews.org/routeviews/.

[12] Three Node Graph Motif. URL https://mathinsight.org/evidence_additional_structure_real_networks.

[13] Route Servers. URL https://www.inetdaemon.com/tutorials/troubleshooting/tools/route_servers.shtml.

[14] CAIDA WHOIS Map. URL https://www.caida.org/research/id-consumption/whois-map/.

[15] T. Bates, E. Chen, and R. Chandra. BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP). 4456(rfc), 2006.

[16] Tim Blankers, Benno Overeinder, and Nlnet Labs. BGP Evolution Analysis. pages 1–51, 2014.

[17] L. Blunk, M. Karir, and C. Labovitz. Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format. *Development*, 6396(rfc), 2011. URL https://tools.ietf.org/pdf/rfc6396.pdf.

[18] Brian Weis (Cisco Systems). Secure Origin BGP (soBGP) Certificates. *rfc*, pages 1–38, 2003.

[19] R Chandra, C. Villamizar, and R. Govindan. BGP Route Flap Damping This. 2439(rfc), 1998.

[20] Jian Chang, Krishna K. Venkatasubramanian, Andrew G. West, Sampath Kannan, Insup Lee, Boon Thau Loo, and Oleg Sokolsky. AS-CRED: Reputation and alert service for interdomain routing. *IEEE Systems Journal*, 7(3):396–409, 2013. ISSN 19328184. doi: 10.1109/JSYST.2012.2221856.

[21] E. Chen. Route Refresh Capability for BGP-4 This. rfc(2918), 2000. URL https://tools.ietf.org/pdf/rfc2918.pdf.

[22] Kai Chen, Chengchen Hu, Wenwen Zhang, Yan Chen, and Bin Liu. On the eyeshots of BGP vantage points. *GLOBECOM - IEEE Global Telecommunications Conference*, 2009. ISSN 1930-529X. doi: 10.1109/GLOCOM.2009.5425389.

[23] Amogh Dhamdhere and Constantine Dovrolis. Twelve years in the evolution of the internet ecosystem. *IEEE/ACM Transactions on Networking*, 19(5):1420–1433, 2011. ISSN 10636692. doi: 10.1109/TNET.2011.2119327.

[24] Gregory G Finn. Reducing the Vulnerability of Dynamic Computer Networks. page 78, 1989.

[25] Lixin Gao. On Inferring Automonous System Relationships in the Internet. *{ACM/IEEE} Transactions on Networking*, 9(6):733–745, 2001. ISSN 10636692. doi: 10.1109/90.974527. URL http://doi.acm.org/10.1145/504611.504616.

[26] Phillipa Gill, Michael Schapira, and Sharon Goldberg. A survey of interdomain routing policies. *ACM SIGCOMM Computer Communication Review*, 44(1):28–34, 2013. ISSN 01464833. doi: 10.1145/2567561.2567566. URL http://dl.acm.org/citation.cfm?doid=2567561.2567566.

[27] R. Govindan and A. Reddy. An analysis of Internet inter-domain topology and route stability. *Proceedings of INFOCOM '97*, 2:850–857, 1997. ISSN 0743-166X. doi: 10.1109/INFCOM.1997.644557. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=644557.

[28] Y. Rekhter (IBM), T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). *Request for Comments*, -(January):1–104, 2006. doi: 10.1109/ICDCS.2017.47.

[29] Paul Koznek John Armington Purdy Ho and Richard Martinez Hewlett. *Lecture Notes in Computer Science*. Number December 2016. 2011. ISBN 9783642231964. doi: 10.1007/978-3-642-23196-4.

[30] Josh Karlin, Stephanie Forrest, and Jennifer Rexford. Autonomous security for autonomous systems. *Computer Networks*, 52(15):2908–2923, 2008. ISSN 13891286. doi: 10.1016/j.comnet.2008.06.012.

[31] Frits Kastelein. Inferring relationship types and simulating bgp traffic between autonomous systems using the valley-free constraint. Master's thesis, TU Delft, Cyber Security Group, 11 2018. BGP Simulator, Relationship Improver.

[32] S. Kent, C. Lynn, and K. Seo. Design and analysis of the Secure Border Gateway Protocol (S-BGP). *Proceedings - DARPA Information Survivability Conference and Exposition, DISCEX 2000*, 1(4):18–33, 2000. ISSN 07338716. doi: 10.1109/DISCEX.2000.824939.

[33] Brijesh Kumar. Integration of security in network routing protocols. *ACM SIGSAC Review*, 11(2):18–25, 1993. ISSN 0277920X. doi: 10.1145/153949.153953. URL http://portal.acm.org/citation.cfm?doid=153949.153953.

[34] Craig Labovitz, G Robert Malan, and Farnam Jahanian. Internet Routing Instability - Networking, IEEE/ACM Transactions on. 6(5):515–527, 1998. ISSN 1063-6692.

[35] M. Lepinski and K. Sriram. BGPsec Protocol Specification. pages 1–45, 2017. URL https://tools.ietf.org/pdf/rfc8205.pdf.

[36] Matthew Luckie. Scamper: a scalable and extensible packet prober for active measurement of the internet. *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 239–245, 2010. doi: 10.1145/1879141.1879171.

[37] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and Kc Claffy. AS relationships, customer cones, and validation. *Proceedings of the 2013 conference on Internet measurement conference - IMC '13*, pages 243–256, 2013. doi: 10.1145/2504730.2504735. URL http://dl.acm.org/citation.cfm?doid=2504730.2504735.

[38] Zhuoqing Morley Mao, Jennifer Rexford, Jia Wang, and Randy H. Katz. Towards an accurate AS-level traceroute tool. *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '03*, page 365, 2003. ISSN 01464833. doi: 10.1145/863993.863996. URL http://portal.acm.org/citation.cfm?doid=863955.863996.

[39] R Milo. Network {Motifs}: {Simple} {Building} {Blocks} of {Complex} {Networks}. *Science*, 298(5594): 824–827, 2002. ISSN 00368075, 10959203. doi: 10.1126/science.298.5594.824. URL http://www.sciencemag.org/cgi/doi/10.1126/science.298.5594.824.

[40] Asya Mitseva, Andriy Panchenko, and Thomas Engel. The State of Affairs in BGP Security: A Survey of Attacks and Defenses. *Computer Communications*, 124(February):45–60, 2018. ISSN 0140-3664. doi: https://doi.org/10.1016/j.comcom.2018.04.013. URL http://www.sciencedirect.com/science/article/pii/S014036641731068X.

[41] Martin O Nicholes and Student Member. A Survey of Security Techniques for the Border. pages 1–22, 2008.

[42] J Nieto. Analysis of the Internet topology and Generation Models, 2005.

[43] Juan Iván Nieto-Hipólito, José María Barceló-Ordinas, Oscar Iván Lepe-Aldama, José Antonio Michel-Macarty, Juan De Dios Sánchez-López, and Haydeé Melendez-Guillén. AStop: A new topology generator at the autonomous systems level. *Proceedings - Electronics, Robotics and Automotive Mechanics Conference, CERMA 2006*, 2:349–360, 2006. doi: 10.1109/CERMA.2006.105.

[44] Hu Ning, Zhu Peidong, and Zou Peng. Reputation mechanism for inter-domain routing security management. *Proceedings - IEEE 9th International Conference on Computer and Information Technology, CIT 2009*, 2:98–103, 2009. doi: 10.1109/CIT.2009.123.

[45] Ola Nordström and Constantinos Dovrolis. Beware of BGP attacks. *ACM SIGCOMM Computer Communication Review*, 34(2):1, 2004. ISSN 01464833. doi: 10.1145/997150.997152. URL http://portal.acm.org/citation.cfm?doid=997150.997152.

[46] Matthew Roughan, Walter Willinger, Olaf Maennel, Debbie Perouli, and Randy Bush. 10 Lessons From 10 Years of Measuring and Modeling the Internet'S Autonomous Systems. *IEEE Journal on Selected Areas in Communications*, 29(9):1810–1821, 2011. ISSN 07338716. doi: 10.1109/JSAC.2011.111006.

[47] Pavlos Sermpezis, Vasileios Kotronis, Petros Gigis, Xenofontas Dimitropoulos, Danilo Cicalese, Alistair King, and Alberto Dainotti. ARTEMIS: Neutralizing BGP Hijacking within a Minute. pages 1–15, 2018. URL http://arxiv.org/abs/1801.01085.

[48] B R Smith and J J Garcia-Luna-Aceves. Securing the border gateway routing protocol. *Global Telecommunications Conference, 1996. GLOBECOM '96. 'Communications: The Key to Global Prosperity*, pages 81–85, 1996. ISSN 1535-7228. doi: 10.1109/GLOCOM.1996.586129.

[49] Vladimir Terzija, Gustavo Valverde, Devu Cai, Pawel Regulski, Vahid Madani, John Fitch, Srdjan Skok, Miroslav M. Begovic, and Arun Phadke. Wide-area monitoring, protection, and control of future electric power networks. *Proceedings of the IEEE*, 99(1):80–93, 2011. ISSN 00189219. doi: 10.1109/JPROC.2010.2060450.

[50] Beichuan Zhang, Raymond Liu, Daniel Massey, and Lixia Zhang. Collecting the internet AS-level topology. *ACM SIGCOMM Computer Communication Review*, 35(1):53, 2005. ISSN 01464833. doi: 10.1145/1052812.1052825. URL http://portal.acm.org/citation.cfm?doid=1052812.1052825.

[51] Randy. Zhang and Micah. T A T T Bartell. BGP design and implementation LK - https://tudelft.on.worldcat.org/oclc/560270451, 2004. URL http://www.books24x7.com/marc.asp?bookid=35331http://www.myilibrary.com?id=267502http://proquest.safaribooksonline.com/640http://proquest.safaribooksonline.com/9781587058646http://proquest.safaribooksonline.com/9781587058646http://proquest.tech.safar.

[52] Ying Zhang, Zheng Zhang, Zhuoqing Morley Mao, Charlie Hu, and Bruce MacDowell Maggs. On the impact of route monitor selection. *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement - IMC '07*, page 215, 2007. doi: 10.1145/1298306.1298336. URL http://portal.acm.org/citation.cfm?doid=1298306.1298336.

Cover image adopted from Google search[7]