INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS MONTERREY

PROGRAMA DE GRADUADOS EN TECNOLOGÍAS DE INFORMACIÓN Y ELECTRÓNICA

On Reachability of Autonomous Systems Based on Announcements Changes

THESIS

Presented as a partial fulfillment of the requirements for the degree of
Master of Science in Electronic Engineering
Major in Telecommunications

Jorge Alberto León Castelán

Monterrey, N.L. July 2007

INSTITUTO TECNOLÓGICO DE ESTUDIOS SUPERIORES DE MONTERREY

DIVISIÓN DE TECNOLOGÍAS DE INFORMACIÓN Y ELECTRÓNICA

PROGRAMA DE GRADUADOS EN TECNOLOGÍAS DE INFORMACIÓN Y ELECTRÓNICA

The members of the thesis committee recommended the acceptance of the thesis of Jorge Alberto León Castelán as a partial fulfillment of the requirements for the degree of Master of Science in:

Electronic Engineering

Major in Telecommunications

Thesis Committee:

_____

César Vargas Rosales, Ph.D.
Advisor

_____

José Ramón Rodríguez Cruz, Ph.D.
Synodal

_____

Carlos Mex Perera, Ph.D.
Synodal

_____

Graciano Dieck Assad, Ph.D.
Director of the Graduate Program

July, 2007

On Reachability of Autonomous Systems Based on Announcements Changes

Jorge Alberto León Castelán

THESIS

Presented as a partial fulfillment of the requirements for the degree of
Master of Science in Electronic Engineering
Major in Telecommunications

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS
SUPERIORES DE MONTERREY

July, 2007

To my parents who showed me that all you have to do is believe in yourself.

I love you.

Thanks, for everything

Acknowledgments

To my family, for all the love and support.

To my friends, for being there when I needed them the most and for never letting me quit.

To my thesis advisor, for his enthusiasm and guidance through the completion of this work.

Thank you

Abstract

In this work data obtained from interdomain routing protocol (IDR) is used to obtain statistical analysis of Autonomous Systems using the protocol. With the statistical description we observe the reachability of three AS belonging to the Internet core and its three most active links. We also calculate the updates and withdrawals to monitor the changes observed in the three AS from the interdomain protocol. This monitoring led us to find a ratio around 1 for updates and withdrawals, this could be taken into consideration for detecting anomalies in the interdomain routing protocol.

The creation of solutions for such anomalies is out of the scope of this work.

Table of contents

List of graphics

Chapter 1

Introduction

In a few years, the internet has rapidly evolved from research networking serving a few users to a huge interconnection of about 350 million hosts (June 2005). In this way, the internet is the largest distributed system ever built. The internet is organized in a multitude of administratively independent networks called domains or Autonomous Systems (ASes). For example, an AS can be an Internet Service Provider (ISP), a University or a corporate network. Over this huge infrastructure there is growing trend to deploy new applications such as the transmission of Voice or Video over IP and new services such as Virtual Private Networks (VPNs). These new applications and services require better or strict guarantees of quality while the Internet has been designed to provide a best-effort service.

Network engineers rely on Traffic Engineering (TE) to adapt the configuration of their network in order to support the evolution of the traffic demand. Traffic Engineering is defined by the IETF TE Working group as the process of evaluating and enhancing operational IP networks performance. The objective of Traffic Engineering can be summarized in avoiding congestion, providing resilience and supporting Quality of service (QoS). Most of the Traffic Engineering complexity comes from hop-by-hop destination based IP forwarding, i.e. each router on the path selects the next router to forward the packet based on the packet destination only. There is no way to explicitly determine the path followed by the packets to reach their destination. Moreover, the routing decisions are taken in a distributed manner by each hop along the path. One of the main difficulties of Traffic Engineering comes thus from routing. Inside a single domain routing is done thanks to link-state protocols such as IS-IS or OSPF. From the perspective of Traffic Engineering these protocols have two twofold advantage of propagating information on the whole topology and optimizing a single global objective: least cost path. Intradomain Traffic Engineering is a well understood problem and solutions exist.

In contrast when Traffic Engineering has to be performed over the boundaries of multiple domains things are far more difficult. The central problem is the Internet routing system itself. Internet routing is currently built around the Border Gateway

Protocol (BGP). BGP is a path vector protocol that propagates only limited view of the topology. A BGP router will advertise to its neighbors a single route per reachable destination.

One characteristic of BGP is that each domain is administered independently. For this reason BGP in each domain is configured to optimize local objectives. The objectives of one domain might be very different from those of another one. Moreover, each domain is allowed to filter the routes advertised to other domains.

The limited view of the topology due to the path vector nature of BGP and due to the local routing policies decrease the diversity of interdomain paths and subsequently the freedom of an AS to direct traffic along alternative paths. An AS is not eager to let other AS control the routing in its own network. BGP provides very limited control on the routing decisions taken by other domains.

Due to the limitations of BGP, we propose a measurement of three nodes, three Autonomous Systems considered to be part of the Internet core, [22]

1.1 Problem definition

Failures at the BGP level can have significant impact on the overall Internet. Understanding the behavior of BGP is thus both an important practical challenge and an interesting research problem. Security incidents such as, system cracking, DDoS attacks, worms and misconfigurations have an adverse impact not only on the end systems, but also on the Internet routing, resulting in many out of reach prefixes. Intuitively, security accidents exert a bad effect on the Internet routing, which is crucial to the reliability of the Internet. Certain ASes become partially or totally unreachable during such incidents, this because their routes or prefixes are being withdrawn from the interdomain. Since the AS cares more about updates than from withdrawals it may not realize that the number of withdraw routes is exceeding the number of updated routes leading to degrading its reachability. If we take into account the number of updates and withdrawals we will be able to foresee when the ratio between these two parameters is affecting the AS reachability.

To understand the true dynamics and to avoid having to deal and interpret the multiple gigabytes of a BGP log data, we parsed the data given by Link Rank [32]. This

data contains significant information such as number of routes updated number of routes withdrawals to and from ASes. The result of this parsing gives weighs to the links between Autonomous Systems by the number of routing prefixes going through each link.

There is a particularity of BGP protocol that has to be taken in consideration regarding the time between updates, particularity why we choose time between updates to perform some of our analysis. To control the number of updates and reduce the processing in routers, it is recommended that BGP enabled routers set a timer called MinRouteAdver (minimum route advertisement) to 30 seconds. This timer establishes the minimum time a router has to wait before sending BGP updates to its neighbor regarding the same destination. Therefore, if a route is unstable and it is changing constantly in a short period of time, instead of flooding its neighbors with updates regarding the same routing change, the router needs to wait 30 seconds between them. Reducing the processing in routers improves the BGP convergence. Instability, in the form of wide-scale cascading failures can occur when a number of routes repeatedly timeout due to router reboots, link congestion or physically link intermittent failures. Instability, in the form of delayed convergence (up to several minutes) can also occur upon routing or policy changes due to the MinRouteAdver timer because ASes speaking BGP explore alternate paths to reach their destination. As there's no way of knowing if BGP neighbors have this timer set, the AS would accept announcements that may be lead it to present BGP delay convergence by admitting unstable routes. If we figure out a way of recognize ASes without the MinRouteAdver timer set we would be able to reject announcements and improve BGP convergence by reducing the processing in routers.

Since we chose three important ASes to perform our analysis we expect them to have this timer set to 30 seconds, otherwise they would be responsible for wide-scale cascading failures because they are part of the Internet backbone. We visualize the fluctuation of updates and withdrawals along the three months chosen for this study graphically. With this, we estimate the scope of routing changes and reveal important routing dynamics in the presence of BGP update and withdraw routes in BGP sessions establish between Autonomous Systems.

1.2 Hypothesis

Since routers speaking BGP interchange information of updates and withdrawals to announce new ways of getting in touch with other routers speaking BGP,

the statistical characterization of number or routes updated and withdrawn in established BGP sessions will reveal us the changes in the topology of our network by telling us how many ASes are no longer suitable to send BGP traffic and which ASes are now the appropriate to. If we can visualize them in a graphic and translate these announcements into a plot of reachability of the Autonomous System, we would be able to say if determined AS loosing or gaining routes in a way that is not normal.

1.3 Objectives

As a way to quantify AS reachability we propose the following metrics: the number of updates and withdrawals along with the number of prefixes added or retreated from the link and the time elapsed between updates. We define updates as the number of prefixes added to certain link and withdrawals as the number of routes or prefixes taken away from the link between Autonomous Systems. The principal objectives of this thesis are:

- To obtain a statistical relationship between the number of updates and the numbers of withdrawals received by an AS.

- To obtain a graphic that reveals us the behavior of reachability based on the announcements received by the AS. This graphics should be seen from the entire AS and from its three most active links point of view.

- To obtain a graphic that shows us how updates and withdrawals occurred during the three months chosen for this analysis.

- To determine if ASes in the Internet core contribute to Internet routing instability by announcing unstable routes causing BGP delayed convergence.

Even though our analysis was done off line our methodology can be applied on line and used in real time to measure reachability.

1.4 Final product

A tool to visualize Internet routing changes at AS and links scale which can visually capture the number of routes carried over the Internet and the changes in link weights. With this tool we can easily observe important routing changes from massive

amount of real routing data, discover routing problems, understand the impact of topological events and infer root causes of observed routing changes.

A way of determine if a AS is contributing to Internet routing instability by looking at its MinRouteAdver timer even though this information is not explicitly given by any AS due to its policies.

1.5 Contributions

- We show that there is a relationship between the number of updates and withdrawals received by an AS and its links.
- A tool to visually take notice of Internet routing behavior.
- We prove that three of the most important AS in Internet have its MinRouteAdver timer set to 30 seconds which is the time recommended to prevent BGP low convergence.
- We show that, as expected, the time between updates shows a power law behavior.

1.6 Summary

Chapter 2 encloses the necessary background to understand the work proposed here, concepts like: BGP protocol, Internet organization, Internet routing focusing on interdomain and intradomain routing, BGP dynamics and power law relationships in Internet topology.

Chapter 3 contains the methodology and results of the reachability visual tool and the analysis of BGP dynamics.

Chapter 4 includes the conclusions of this work, in particular the possible causes of the relationships between updates and withdrawals and why we saw no abnormal behavior in any of the three Autonomous Systems or in any of its links.

Chapter 2

Background

This chapter is organized as follows. We first give an overview of the Internet organization in section 2.1 in order to understand how the Internet is deployed, we introduce themes such as Autonomous System, relationship between them, intradomain and interdomain these concepts are essential to understand the context of the thesis. Secondly, in section 2.2, we introduce the BGP protocol which is the de facto standard interdomain routing protocol. In this section we explain how BGP routers exchange information to establish connection between them by sending a series of messages explained in later subsections. Then, we describe in section 2.3 how well does internet routing work today. In section 2.4 we give a special analysis on difficulties in understanding internet routing focusing on intradomain routing in section 2.4.1 and interdomain routing in section 2.4.2. We next study BGP dynamics in section 2.5 since we analyze some of them in chapter 3, and we give an example of BGP dynamics in section 2.5.1. We conclude in section 2.6 we give an overview of power law in internet topology, given that we found this behavior in chapter 3.

In order to learn routes towards destination located outside their own domain, the routers run the Border Gateway Protocol (BGP), [16, 17, 18]. BGP is the de facto standard routing protocol for the selection of the interdomain paths. The rationale behind the design of BGP was to provide reachability among domains and the ability for any domain to enforce its own routing policies, i.e. controlling what traffic enters and leaves the domain, and where.

BGP routers exchange routing information by means of BGP sessions. Each BGP session is established between a pair of routers over a TCP connection. External BGP (eBGP) sessions are established over the edge links while internal BGP (iBGP) sessions are established between the routers of the AS. There is a full-mesh of iBGP sessions between the routers of the AS.

A route advertisement indicates the reachability of a network. A route advertisement contains the prefix of the destination network as well as the complete

interdomain path that the route follows. The interdomain path is the list of all the ASes that must be crossed in order to reach the AS of the destination.

The Internet today is owned by no single administrative entity, but instead consists of thousands of networks. Each has its own routing polices. A network belonging to a single administrative entity is considered a unit of routing policy, also known as an Autonomous System.

One administrative entity, however, can have more than one Autonomous System. The Internet is a large decentralized network that already connected about 350 million hosts in June 2005, [1]. Furthermore, these hosts are organized in about 21,000 distinct domains, [2], a domain corresponding roughly to a company, an Internet Service Provider (ISP) or a campus network. All these domains are interconnected to form the global Internet. Over this large interconnection of networks, ISPs run two different families of routing protocols. Intradomain routing protocols are used within the ISP while an interdomain routing protocol is used across the ISP boundaries

The Internet is expected to grow more complex with increasing number of users, as more people get online, more places get wired, and more stub networks are connected or multihomed to multiple upstream providers for increased redundancy. Thus, a better understanding of how to make the Internet robust and fault-tolerant is increasingly important. To achieve these goals, it becomes critically important to have better insight into the run-time behavior of BGP: to answer why routes take so long to converge, some destinations are unreachable, and packets flow along an unexpected path deviating from the routing information. Answers to these questions today are not easy to obtain due to lack of visibility.

The initial research Internet was designed with a best-effort service in mind where connectivity was the most important issue. Today, connectivity is considered to be granted, but the architecture initially designed to provide a best-effort service is used for more demanding applications, and sometimes with Service Level Agreements (SLAs), [3]. To meet the requirements of these applications or to ensure the Quality of Service (QoS) required by SLAs, several ISPs rely on a process called Traffic Engineering (TE), [4].

Traffic Engineering covers the evaluation and the improvement of the performance of operational IP networks. However, if performing Traffic Engineering

inside a single AS is a well understood problem, it is far more difficult when performed across the boundaries of multiple ASes. The main limitation of interdomain Traffic Engineering comes from the current Internet routing organization.

In the next section we introduce the Internet organization in order to understand the limitations to perform routing across different ASes.

2.1 Internet organization

The Internet is a network composed of a huge collection of smaller networks, themselves containing numerous end systems and routers. The end systems are hosts such as personal computers or servers. They are usually the sources or sinks of data packets transiting on a network. The routers are the intermediate systems that intervene in the transport of data from an end system to another. Since the many networks that form the Internet are operated by a lot of independent institutions, the Internet is organized in two levels.

The first level is the intradomain level. A set of routers that is under a single administrative authority form a domain. A domain can be the network of a company, an Internet Service Provider (ISP) or a single campus network. An example ISP is represented in Figure 2.1. The routers of a domain are usually interconnected using multiple Synchronous Optical Networking links (SONET/SDH) and/or Ethernet. We distinguish the core links that interconnect the routers within the domain and the edge links that cross the domain boundaries. Since the edge links connect to routers lying outside of its network, a domain only manages one side of the edge links.

Through the edge links, the domain is connected to different kinds of neighbor networks. On one side, the access links mainly connect to customer networks. For example, an access link could connect DSL users or a university or corporate campus network. On the other side, the peering links connect to other domains. For example, peering links could connect to neighboring ISPs. The routers where edge links are terminated are called the domain’s border routers. The different geographical locations of border routers and access routers are usually called the Points of Presence (PoPs) of the domain.

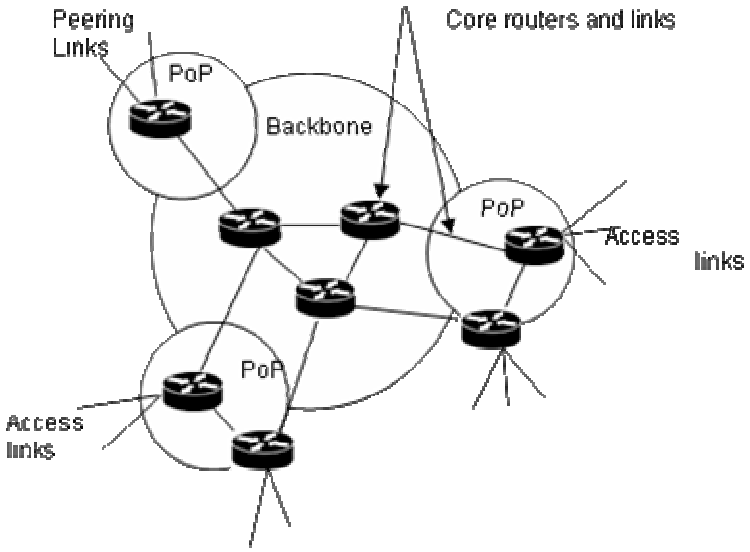

Figure 2.1 Topology of an ISP network

The second level of the Internet is the interdomain level. It designates the interconnections between the different domains. In the Internet, a domain is also called an Autonomous System (AS). Most ASes are uniquely identified by an Autonomous System Number (ASN). Note that all domains need not to have a public ASN. This is usually the case for small to medium size university or corporate campus networks that buy connectivity from a single ISP. We show in Figure 2.2 the sketch of a small imaginary Internet composed of 8 different AS domains: Carrier&Wireless, Level3, Belnet, Janet, Geant, Google, ISPx and ISPy. In addition, there are 4 customer networks that do not have their own AS: UCL.be and UCL.uk are campus networks of universities while apple.com and m$.com is corporate networks.

In the example Internet of Figure 2.2, not all domains play an equal role. They can first be distinguished based on their connectivity. In, [5, 6] Huston and Gao have shown that there are two major types of interconnections between distinct domains: the customer-provider and the peer-to-peer relationships.

In the customer-provider relationship, customer domain purchases connectivity from a larger domain, called the provider. In this case, the provider agrees to forward the packets received from the customer to any destination. It also agrees to forward the packets destined to the customer. In Figure 2.2, ISPx and ISPy are examples of customer ASes that buy connectivity from Level3.

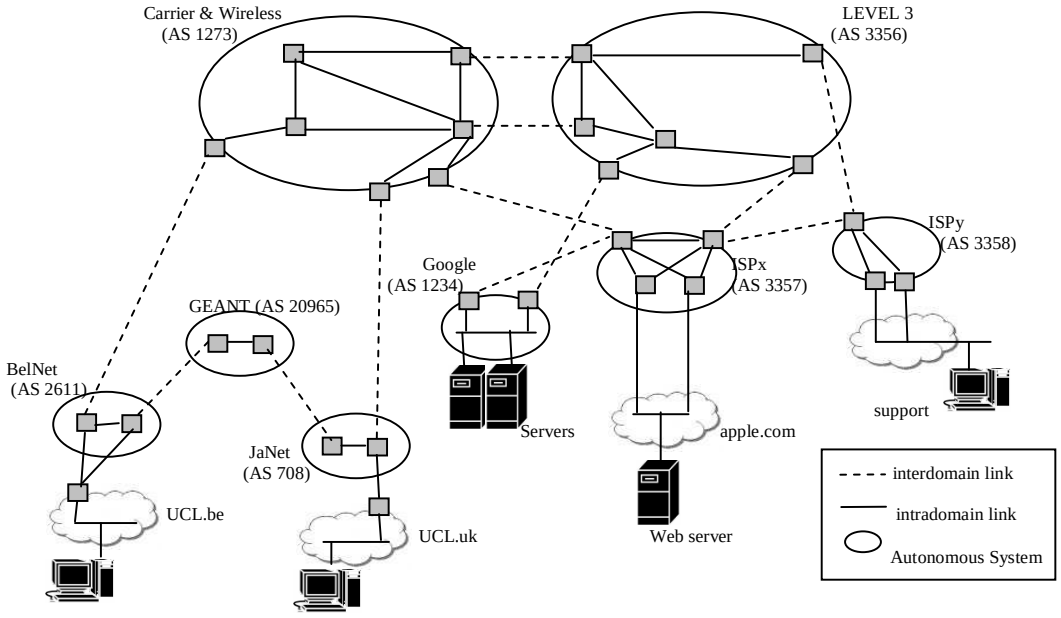

Figure 2.2 Sketch of the Internet architecture

According to a study performed by Subramanian in 2002, [7] the customer-provider relationship was used for about 95 % of the domains interconnections in the Internet. The classification of interdomain relationships in customer-provider and peer-to-peer leads to an interesting view of the Internet as a graph where money is ascending along customer-provider links (Figure 2.3), [5].

Relying on this classification of interdomain relationships, Subramanian, [7] made a first characterization of domains. There are basically two types of domain: transit domains and stub domains. Transit domains constitute the core of the Internet and their purpose is mainly to carry packets from a neighbor domain to another. In the example of Figure 2.2, Carrier & Wireless and Level 3 are example of large transit ASes. According to, [7] the core corresponds to about 15 % of the domains in the Internet and can be divided in three different subtypes (dense, transit and outer core depending on the connectivity of each domain).

On the other hand, stub domains are regional ISPs or customer networks that do not provide transit. Stub domains correspond to 85 % of the Internet and they maintain only a few customer-provider relationships with domains in the core and some peer-to-peer relationships with other small domains. In the example of Figure 2.2, BelNet, JaNet, Google, ISPx and ISPy are stub ASes.

In addition, domains can also be distinguished based on the type of service they provide to their customers. This is interesting mostly for stub domains. For instance, a stub domain can be a small regional ISP providing Internet access to Small/Medium Enterprises (SME) and/or dialup/xDSL/ users. In this case, it will often receive more traffic than it sends. We call this kind of domain a content-consumer. In Figure 2.2, BelNet and JaNet are examples of such domains since they only provide Internet connectivity to universities campus networks. In contrast, a stub domain that hosts video streaming servers or the web servers of a large company will often have more outgoing traffic than incoming traffic. This kind of domain is called a content-provider. An example of such domain in Figure 2.2 is Google who hosts a farm of servers containing a lot of information accessed from everywhere in the Internet.

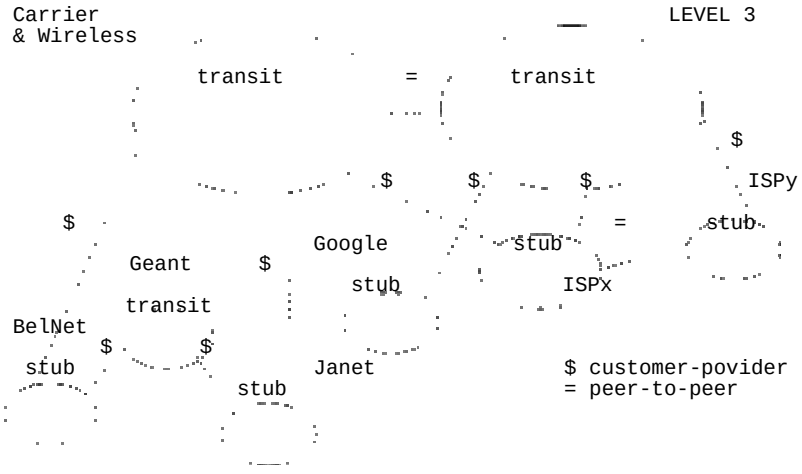

Figure 2.3 Internet business relationships

Both interdomain and intradomain routing protocols dynamically adapt to failures and attempt to route around them. Due to the commercial nature of the Internet, the dynamic behavior of BGP which is essentially policy-based routing has been rather difficult to understand due to lack of policy and topology information.

2.2 The BGP Protocol

BGP is a Path-Vector protocol. It uses TCP as its transport protocol (via port 179). Therefore, all transport reliability is taken care by TCP.

Routers that run a BGP routing process are often referred to as BGP speakers. Two BGP speakers that form a TCP connection between one another to exchange route information are referred to as neighbors or peers.

2.2.1 BGP neighbor negotiation

Neighbor negotiation is based on the successful completion of a TCP transport connection, the successful processing of the OPEN message, and periodic detection of the UPDATE or KEEPALIVE messages.

The BGP negotiation proceeds through different states before the connection is fully established. These are the key states, [16]:

1. Idle. BGP is waiting for a Start event, which is initiated by an operator or the BGP system. After that event, BGP initializes its resources, resets a ConnectRetry timer, initiates a TCP connection, and starts listening for a connection from a remote peer.

2. Connect. BGP is waiting for the completion of the TCP connection. If successful, transitions to OpenSent. If unsuccessful, transitions to Active. If the ConnectRetry timer expires it returns to the Connect state.

3. Active. BGP tries to acquire a peer by initiating a TCP connection. If successful, transitions to OpenSent. If the ConnectRetry timer expires it returns to the Connect state. In general, a neighbor state that is oscillating between Connect and Active indicates that something is wrong with the TCP connection.

4. OpenSent. BGP is waiting for an OPEN message from its peer. In case of errors, sends a NOTIFICATION message and returns to Idle. If there are no errors, BGP starts sending KEEPALIVE messages and resets the KEEPALIVE timer.

5. OpenConfirm. BGP waits for a KEEPALIVE message. If received goes to Established and the negotiation is complete. If there are errors, returns to Idle.

6. Established. BGP starts exchanging UPDATE packets with its peers.

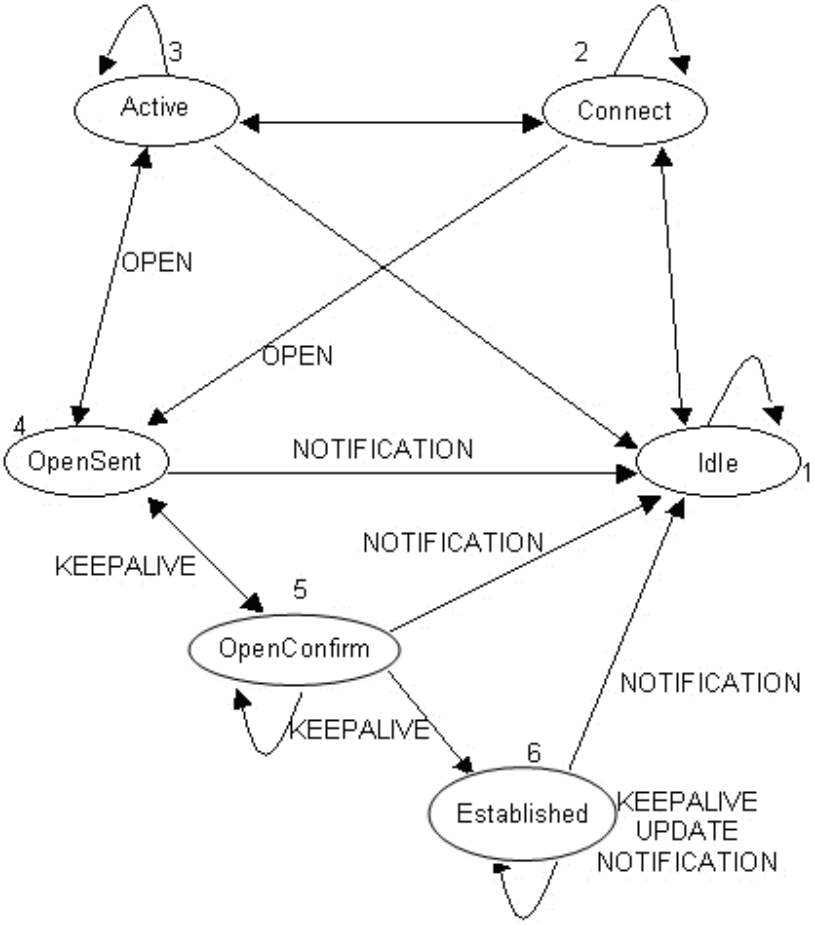

Figure 2.4 BGP neighbor negotiation

2.2.2 BGP Messages

The NOTIFICATION message contains an error code, an error subcode and a data field. The error code indicates the type of notification, the error subcode provides more information and the data field contains data relevant to the error.

The KEEPALIVE messages are periodic and are exchanged between peers to determine reachability.

The UPDATE message is the core of the BGP protocol. It contains the routing updates that are the necessary information that BGP uses to construct a loop-free picture of the network. Its basic blocks are the Network Layer Reachability Information (NLRI), Path attributes and Unfeasible routes.

The NLRI indicates the networks being advertised in the form of prefixes. The unfeasible routes are a list of unreachable or withdrawn routes. An UPDATE message can withdraw multiple routes at once, but can only announce a single route.

The path attributes are a set of parameters used to keep track of route specific information such as path information, degree of preference, NEXT_HOP, and aggregation information.

Path attributes fall into four categories: well-know mandatory, well-know discretionary, optional transitive, and optional nontransitive, [16].

2.2.3 The Routing Process

BGP is a fairly simple protocol. Routes are exchanged between peers by UPDATE messages. BGP routers receive those, run some policies or filters on them and then pass the routes to other BGP peers. BGP picks the best route and sends it. A BGP router can pass along EBGP routes from peers, IBGP routes from route reflector clients or advertise internal networks from its own AS. Valid local routes and the best routes received from peers are then installed in the IP routing table, which is the final routing decision and is used to populate the forwarding table.

2.2.4 BGP Routes: Advertisement and Storage

As specified in RFC 1771, [17]:

For purposes of this protocol a route is defined as a unit of information that pairs a destination with the attributes of a path to that destination:

- Routes are advertised between a pair of BGP speakers in UPDATE messages. The destination are the systems whose IP addresses are reported in the Network Layer Reachability Information (NLRI) field, and the path is the information reported in the path attributes fields of the same UPDATE message.

- Routes are stored in the Routing Information Bases (RIBs): namely, the Adj-RIBs-In, the Loc-RIB, and the Adj-RIBs-Out. Routes that will be advertised to other BGP speakers must be present in the Adj-RIB-Out; routes that will be used by the local BGP speaker must be present in the Loc-RIB, and the next hop for

each of these routes must be present in the local BGP speaker's forwarding information base; and routes that are received from other BGP speakers are present in the Adj-RIBs-In.

If a BGP speaker chooses to advertise the route, it may add to or modify the path attributes of the route before advertising it to a peer.

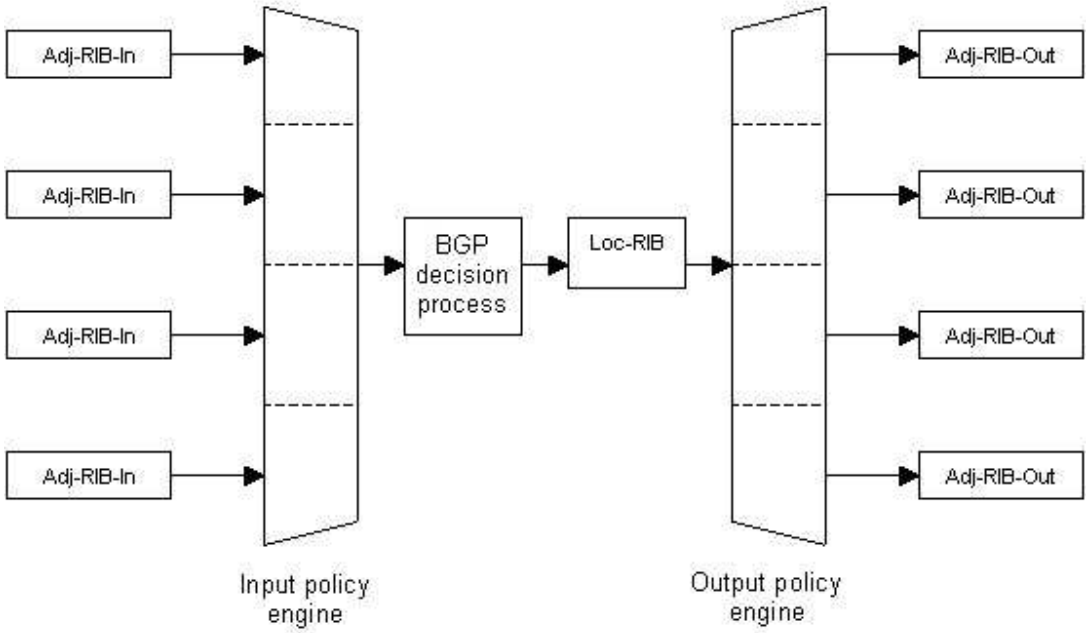

Figure 2.5BGP Routing Information Bases (RIBs)

Although in the figure there's a distinction between Adj-RIBs-In, Loc-RIB, and Adj-RIBs-Out in reality, most implementations store one copy of the information with pointers in order to conserve memory, [16].

2.3 How Well Does Internet Routing Work Today?

We would first like to understand how well Internet routing works today. We first turn to understanding how to evaluate Internet routing.

2.3.1 Routing in the Internet

When a customer complains about routing problems either in terms of reachability or poor performance, it typically is in the context of some applications. It is

not easy to understand the root cause of such problems, especially when they are caused by suboptimal routing decisions. There are many reasons why performance degrades. Network operators, for instance, can install filters in their routers to determine which to accept in calculating the best forwarding path. Packet filters at the routers are much more flexible in the sense that they determine which packets are accepted for forwarding based on attributes of the packets, e.g., port numbers and protocol types. Given a route in one's routing table received by one's upstream provider, there is no guarantee that all application traffic can reach the destination due to the presence of packet filters. Some networks, for instance, perform port-based filtering to protect against known worm traffic. When debugging routing problems, one needs the application's view to understand which type of application traffic is being correctly forwarded.

Thus, it is difficult to judge the quality of Internet routing because of the inability to determine if application degradation is due to routing problems. Network operators have very limited tools to debug routing problems. Only primitive tools like traceroute and ping are used to determine existing routing behavior. However, such probes may not reproduce problems experienced by applications. They require support from routers, which is not universally available. Moreover, there is little visibility into the routing behavior of other ASes from a given AS's perspective, making it even more difficult to identify the source of any routing anomalies. Thus, it is difficult to predict the impact of routing policy changes on global routing behavior. Nevertheless, important classes of routing problems are relatively easy to detect. They involve the lack of reachability for destination prefixes caused by the unavailability of certain routes in the routing tables. Another class of more easily identifiable routing problems involves forwarding loops either within or between ASes. Such routing problems can be detected based on routing table information and routing announcements can reveal the source of the problems.

To be uniquely identified in the Internet, each end system and router receives one or more Internet Protocol (IP) addresses. In the current version of the IP protocol (IPv4), an IP address is a 32-bits integer number. It is usually represented in the dotted format A.B.C.D. An example of IP address is 66.249.93.99. Each AS in the Internet is often being allocated blocks of contiguous IP addresses that they can use for their own network or delegate to their customers. Such a block is usually refereed as network prefix. A network prefix represents the set of IP addresses that start with the same first bits. For example, the IP address 66.249.93.99 belongs to the network prefix 66.249.64.0/19 since its 19 most significant bits are equal to those of the prefix. In this

case, we say that the IP address 66.249.93.99 matches the prefix 66.249.64.0/19. Network prefixes are also sometimes referred to as subnets.

The physical topology of the Internet defines the feasible paths that can be used to cross the network. The role of routing consists in determining for a given Internet device the path to be used to reach a destination IP address. In order to determine these paths, all the routers in the Internet usually exchange information about the network topology. These exchanges are supported by a routing protocol. In the Internet, routing is handled by two distinct protocols with different objectives. An intradomain routing protocol is used inside each domain and a single interdomain routing protocol is used between domains.

There are three main reasons for this division. The first one is the need for scalability. An intradomain routing protocol usually has a very detailed knowledge of the whole domain topology. It handles routes towards any destination within the domain. To the contrary, an interdomain routing protocol has a limited view of the Internet topology, restricted to the interconnection between domains. An interdomain routing protocol also handles routes towards large aggregates of IP addresses. This avoids having to handle routes towards any destination. The second reason for having two distinct routing protocols is the Independence of domains.

Each domain is allowed to setup its intradomain routing in an independent manner. Each domain is also allowed to perform policy routing. For example, a domain can refuse to serve as a transit domain for another domain.

2.4 Difficulties of Internet Routing

We now summarize the difficulties in understanding Internet Routing.

- Unknown information: local policies and internal topologies of ASes are considered private information and not revealed globally. Various Internet mapping and policy inference efforts exist, [13, 14]; however, they are hard validate. The routing behavior heavily depends on both the policy and topology information; therefore, it is rather difficult to do root cause analysis given a feed of BGP updates today, [15].

- Ambiguous specifications: the protocol specification for BGP as defined in RFC 1771 intentionally left out some details by giving the freedom to the vendors on defining things such as whether MinRouteAdvertisement Timer is applied per prefix or per peer. Thus the behavior the protocol depends on the router implementation variants.

- Operational realities differ from specifications (RFCs): similarly to the above point, vendors may deviate from the router specifications, suiting to their own router architecture design. Suggested default timer values, for example, are often not followed.

- Large distributed systems: Internet is a large distributed system, consisting of thousands of autonomous systems, each can have hundreds of thousands of routers. The dynamics of such a system can be very difficult to reason.

- Local changes may or may not propagate globally depending on the the policies: causal analysis of BGP updates is extremely hard, as there is often insufficient information on whether a local routing change can affect other ASes' best route selections.

2.4.1 Intradomain routing

Inside its network, an AS runs an Interior Gateway Protocol (IGP) such as OSPF or IS-IS, [21] in order to compute the interior paths from any AS's router towards the AS's other routers and prefixes. The IGP is typically a link-state protocol, that is, it floods information about the state of the adjacencies between all routers in the whole AS. The objective of the intradomain routing is to find the shortest paths according to a selected metric assigned by the network administrator.

ISPs usually use a metric that is proportional to the propagation delay along the path or to the bandwidth. Many network operators use the Cisco default metric, which is one over the bandwidth, [16]. Some large ASes use a hierarchical IGP, where the AS is divided into different areas. Inside an area, all the adjacency information is flooded. Between areas, only aggregated information is exchanged.

In addition to the IGP, an AS sometimes uses static routing. Static routes are often used on the edge links since routers on both side of these links are not operated by the same authority. Static routes are also used to setup access to small customers that do not have their own AS.

2.4.2 Interdomain routing

In order to learn routes towards destination located outside their own domain, the routers run the Border Gateway Protocol (BGP), [16, 17, 18]. BGP is the de facto standard routing protocol for the selection of the interdomain paths. The rationale behind the design of BGP was to provide reachability among domains and the ability for any domain to enforce its own routing policies, i.e. controlling what traffic enters and leaves the domain, and where. To the contrary of the intradomain routing protocol, BGP does not optimize a single global metric but relies on a decision process composed of a sequence of rules.

BGP is a path-vector protocol that works by sending route advertisements.

BGP routers exchange routing information by means of BGP sessions. Each BGP session is established between a pair of routers over a TCP connection. External BGP (eBGP) sessions are established over the edge links while internal BGP (iBGP) sessions are established between the routers of the AS. There is a full-mesh (a clique) of iBGP sessions between the routers of the AS.

A route advertisement indicates the reachability of a network. A route advertisement contains the prefix of the destination network as well as the complete interdomain path that the route follows. The interdomain path is the list of all the ASes that must be crossed in order to reach the AS of the destination.

This list is called the AS-Path of the route. The AS-Path is used to avoid interdomain level routing loops. In addition to the AS-Path, a route contains a next-hop attribute. The next-hop of the route is the IP address of the router to which packets must be sent in order to reach the destination network. The route also contains several additional attributes.

A router sends a route advertisement for a network if this network belongs to the same AS as the advertising router or if this network is reachable from the router

through a neighboring AS. An important point to note about BGP is that if a BGP router A in ASx sends to a BGP router B in ASy a route advertisement for a network N, this implies that ASx accepts to forward the IP packets to destination N on behalf of ASy.

To better understand the operation of BGP, it is useful to consider a simplified view of a BGP router as shown in Figure 2.6. The router is composed of 4 main components. First, route input and output filters can be configured for each BGP session. The role of a route filter is to deny the routes received or sent by the router or to manipulate their attributes. An example filter would be to only accept the routes with an AS-Path containing a set of trusted ASes. The route filters are configured by the network operator. The second component of a BGP router is the BGP routing table. This routing table contains all the routes received by the router and accepted by the input filters. The attributes of the routes stored in the routing table may have been updated by the input filters. The third component of a BGP router is its decision process. It is responsible for selecting among the routes stored in the routing table a single best route for each destination prefix. When a route is selected as best, it is installed in the forwarding table and it is sent to the neighboring routers. The forwarding table is the fourth component of the router. Each time a packet is received, this table is looked up and it indicates the outgoing interface that must be used to forward the packet to the destination.

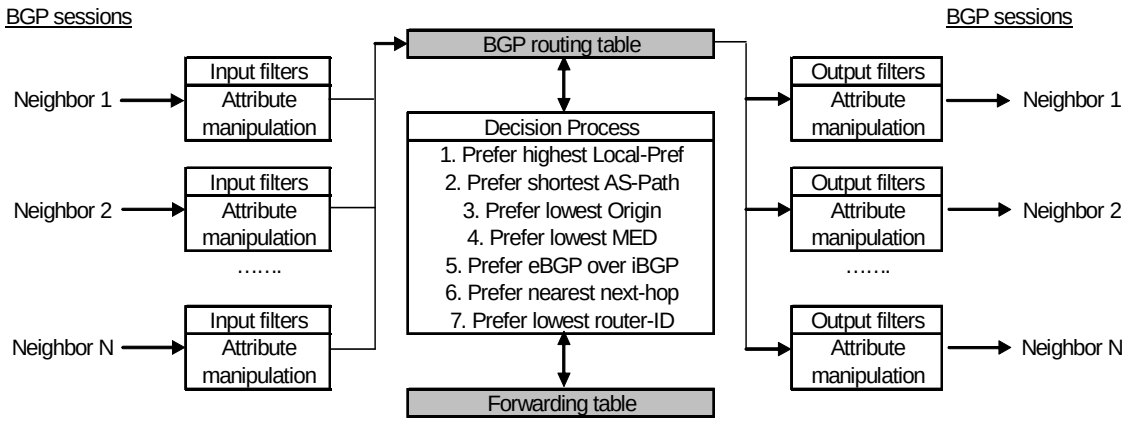

Figure 2.6 Sketch of a BGP Router

Through its BGP sessions, each router receives BGP routes towards destination prefixes. Since there might be multiple routes towards the same destination prefix, a choice must be made. Each router uses its decision process on a per-prefix basis to select the routes it will use. The BGP decision process is a sequence of rules applied to a set of routes towards the same destination prefix to select a single route called the best route

towards this prefix. Basically, the BGP decision process ranks the routes according to their attributes. Each rule of the decision process keeps the routes that it prefers. The surviving routes are then submitted to the next rule, until a single route remains. A summary of the BGP decision process is shown in Figure 2.6.

The BGP decision process considers several of the BGP route's attributes. The first attribute is the Local-Pref which corresponds to a local ranking of the route. It is usually attached to the route upon reception by a border router and it is never propagated outside the AS. The decision process prefers the routes having the highest value of the Local-Pref attribute. The second attribute is the AS-Path. The AS-Path contains the sequence of ASes that the route crossed to reach the local AS.

In the decision process, the AS-Path is used as a distance metric in AS hops. The decision process prefers the routes with the shortest AS-Path. The third attribute is the Multi-Exit-Discriminator (in short, the MED). This attribute is used to rank routes received from the same neighboring AS. Usually, the MED attribute is set by the neighbor AS to indicate the preferred peering link to use (based on the IGP cost in the neighboring AS for instance). The decision process prefers the routes with the smallest value of the MED.

If there are still more than a single route at this step, the decision process will consider the BGP next-hop attribute of the route. The BGP next-hop is often called the egress of the route, i.e. the exit point of the AS. Note that the BGP next-hop may be different from the immediate IP next-hop. When a BGP router receives a route, it first checks that the next-hop is reachable before considering it in the decision process. The decision process uses the IGP cost of the intradomain path towards the next-hop to rank the routes. It prefers the routes with the smallest IGP distance to the next-hop. This rule implements hot-potato routing, [19]. Its aim is to hand over packets to a neighboring AS as quickly as possible in order to consume as few network resources as possible in the local AS. In addition, it automatically adapts routing to topology changes that affect the IGP distance to the egress points inside the AS. This step within the BGP decision process is where the IGP and BGP protocols interact.

Finally, if there are multiple routes remaining, the decision process will break the ties by preferring the route announced by the neighbor router that has the lowest router-ID. The router-ID is the highest IP address of the router. Another tiebreaking rule that is sometimes deployed in BGP routers consists in preferring the older route, [20].

2.5 BGP dynamics

Internet routing is dynamic in nature. Caused by the regular or irregular exchange of routing updates between routers, routing dynamics has always been a major concern of the Internet engineering community. Irregular dynamics can not only cause high bandwidth and processing overhead on routers, but may also lead to packet forwarding failures, including packet delay, jitter, drop, reordering, duplication, or other difficulties in reaching destinations.

With the Internet being indispensable to modern communications and the economy, it is critical to understand the characteristics of routing dynamics.

The most comprehensive study, [10] of Internet routing dynamics is from nearly a decade ago, a substantial period for the fast-evolving Internet. These factors include routing protocol implementation by vendors, network engineering practices, and Internet topologies.

Engineering practices also constantly brought changes. Multi-homing, load balancing, and address fragmentation, have led to substantial increases in BGP routing table sizes.

The Internet itself has been growing at a staggering rate. The number of ASes has been increasing at a linear rate, and both the number of prefixes and the size of BGP routing tables have been growing almost exponentially, resulting in more complicated traffic engineering and routing policies. Moreover, while BGP updates are carried over the same link as the ordinary traffic, the usage patterns of the Internet change over time and new applications and traffic patterns could have an impact on BGP as well.

Essentially, the constant growth and changing features of the Internet create the necessity of revisiting the topic of BGP routing dynamics in order to capture new statistics and trends. In chapter three, we investigate characteristics of the BGP routing dynamics on the Internet using recent BGP data from January 2007 to March 2007.

Our primary focus in this work is to observe and understand how BGP dynamics look nowadays.

2.5.1 An example of the complexity of BGP dynamics

We now describe a simple example, where even if given the precise topology and policy information, it is still extremely difficult to infer the root cause of a routing update.

We examine a simple topology in Figure 2.7 used by many previous works, [15]. Here we have a 5-node topology; each node denotes an AS for simplification. In this figure AS5 wants to communicate with AS1. Outside every node there is a number or series of numbers, these digits indicate the preferred route of the nodes to reach AS1; take a look at the numbers of AS4: its preferred route to reach AS1 is through AS2 (AS4-AS2-AS1 [4-2-1]). AS4 can also reach AS1 through AS3 [4-3-1] but it has established in its routing policies that it prefers the first link over the second one to reach AS1.

In steady state (no update or withdrawal of route is send among the ASes) AS5 sends it packets to AS1 through the path [5 3 1]. But when does AS5 prefer its alternatives routes to reach AS1? There can be several possibilities. We describe a subset of them below.

1. AS3 can't reach AS1 anymore, so the link between AS1 and AS3 is down. Once AS3 detects this it communicate this to AS5 causing AS5 to withdraw its preferred route to reach AS1 [5 3 1].

2. AS3 also communicates to AS4 that it can reach AS1, this causes AS4 to subsequently communicate to AS5 that AS4 can reach no longer AS1 through AS3[4 3 1] causing AS5 to eliminate its second preferred route [5 4 3 1].

3. AS2 withdraws its route to AS1, triggering AS4 to announce [4 3 1] to AS5 causing AS5 to prefer [5 4 3 1] to [5 3 1].

AS5 has to perform traffic engineering in each case listed above, its routing preference changes preferring the longer route over the previously announced shorter route.

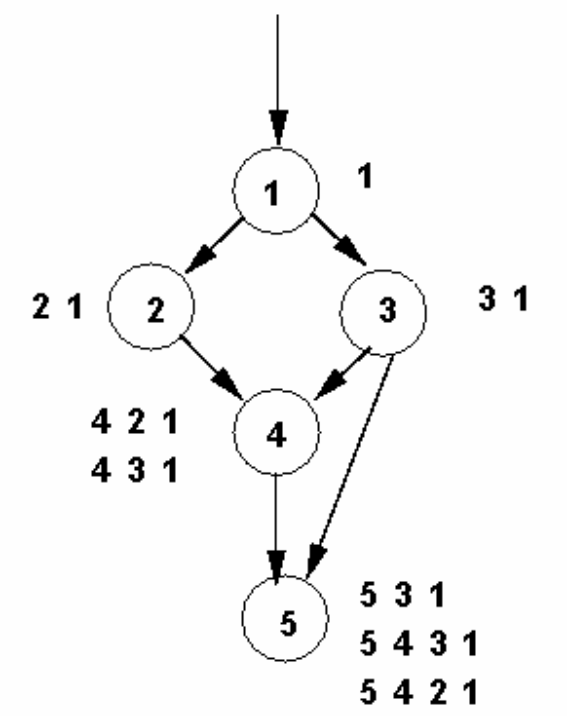

Figure 2.7 A simple topology to illustrate the complexity of BGP dynamics.

This example shows that even knowing the initial routing preference and actual routing topologies, it is difficult to identify the root cause for any routing change. This strongly motivates an experimental infrastructure where the routing change can be precisely controlled to understand its impact on the Internet. In such an infrastructure, it then is possible to measure basic metrics such as convergence delays. It also allows us to understand policies and configuration settings used in remote networks.

2.6 Power Laws in Internet Topology

Many man made and naturally occurring phenomena, including city sizes, incomes, word frequencies, and earthquake magnitudes, are distributed according to a power-law distribution. A power-law implies that small occurrences are extremely common, whereas large instances are extremely rare. This regularity or 'law' is sometimes also referred to as Zipf and sometimes Pareto.

All three terms are used to describe phenomena where large events are rare, but small ones quite common. For example, there are few large earthquakes but many small

ones. There are a few mega-cities, but many small towns. There are few words, such as 'and' and 'the' that occur very frequently, but many which occur rarely.

Zipf's law usually refers to the size y of an occurrence of an event relative to its rank *r*. George Kingsley Zipf, a Harvard linguistics professor, sought to determine the 'size' of the 3rd or 8th or 100th most common word. Size here denotes the frequency of use of the word in English text, and not the length of the word itself. Zipf's law states that the size of the *r*-th largest occurrence of the event is inversely proportional to its rank:

$$y \approx r^{-b}$$

with *b* close to unity.

Pareto was interested in the distribution of income. Instead of asking what the r-*r*-th largest income is, he asked how many people have an income greater than *x*. Pareto's law is given in terms of the cumulative distribution function (CDF), i.e. the number of events larger than x is an inverse power of *x*:

$$P\left[X > x\right] \approx x^{-k}$$

It states that there are a few multi-billionaires, but most people make only a modest income.

What is usually called a power law distribution tells us not how many people had an income greater than x, but the number of people whose income is exactly *x*. It is simply the probability distribution function (PDF) associated with the CDF given by Pareto's Law. This means that

$$P\left[X = x\right] \approx x^{-(k+1)} = x^{-a}$$

That is the exponent of the power law distribution a = 1+k (where *k* is the Pareto distribution shape parameter) [31].

Empirical studies, [28] have shown that Internet topologies exhibit power laws of the form:

$$y \propto x^{\alpha}$$

In fact, several parameters of the network are related through these power laws. As examples, we have relationships such as degree of a node versus rank, number of nodes versus degree, number of node pairs within a neighborhood versus neighborhood size, and eigenvalues of the adjacency (or routing) matrix versus ranks. To further explain these relationships, let us first define three concepts: degree, rank and diameter.

Given a graph, the degree or outdegree (d) of a node is defined as the number of edges incident to the node. If the nodes are ordered in decreasing degree sequence, the rank (r_v) of a node is its index in the sequence. The diameter (δ) of a graph is the maximum distance between two nodes given in hops (h)

Now the power law relationships can be expressed as follows:

Rank exponent: the degree (d_v) of a node (v) is proportional to the rank of the node (r_v) to the power of a constant R, i.e.,

$$d_v \propto r_v^{R}$$

Degree exponent: the frequency (f_d) of a degree (d) is proportional to the degree to the power of a constant O.

$$f_d \propto d^{O}$$

Hop-plot exponent: the total number of pairs of nodes ($P(h)$) between h hops of distance is proportional to the number of the hops to the power of a constant H.

$$P(h) \propto h^{H}$$

If $\delta >> h$

According to, [28] there are four reasons for these behaviors:

1. Preferential connectivity of a new node to existing nodes. New nodes have the tendency of connect to the existing nodes with higher outdegree.

2. Incremental growth. New nodes join the Internet in an incremental way.

3. Geographical distribution of nodes. The space distributions of nodes follow heavy-tailed distributions.

4. Locality of edge connections. New nodes have the tendency of connect to nearby nodes instead of far away nodes.

This behavior is also observed in, [29] where it is explained as a fading wave effect. When a new node joins a network it triggers a chain of restructuring changes. If many new nodes connect to an existing node, it will probably have to increase its connectivity to accommodate the new demand in traffic. Therefore, at any time, the topology is characterized by the same fundamental properties.

In the interdomain level each node represents an autonomous system and each edge is an interdomain connection. It is shown in, [30] that the power law relationships also hold for the interdomain topology.

In this chapter we studied the BGP protocol which is de facto routing protocol in Internet, we showed how routing is done in Internet (intradomain and interdomain), introduced the concept of BGP dynamics and power law. These concepts will allow us to better understand our analysis performed in the next chapter. In chapter 3 we show the analysis did on BGP logs corresponding to January, February and March of 2007. We show the reachability of 3 ASes and of its 3 most active links graphically, we found that the updates and withdrawals are correlated in time and that the time between updates exhibits a power law behavior.

Chapter 3

Analysis and results

We finished the last chapter with a brief insight of what the dynamics in BGP are. In this chapter we study four dynamics. The first one is the reachability. We aboard it from two points of view that of an entire Autonomous System and from that of its three most active links. We use the word reachability to indicate the number of routes that the AS has stored in its routing table, these routes will be announced to several ASes surrounding it to communicate them of new or withdrawn routes in the AS routing table.

We study the statistics of three ASes in three months. The study of the reachability in Section 3.4.1 will give us an insight of the behavior in time of the ASes. We show a visual tool in which we can see the number of reachable routes that the AS and three most active links have through the three month period. This is important since we can detect any anomaly in the routing process; we can detect, for example, if an AS is removing more routes than it is inserting in its routing table or viceversa, this would produce that the AS fill its routing table (causing delay of convergence) or that the AS lost connection with its neighbors or maybe an indicator that the AS is being attacked, causing routing instability. The importance of this study is that if we are able to detect an anomalous behavior in the AS routing we can take actions to detect the root of the problem and neutralize it. In order to do this examination, we recollected information of the ASes from several vantage points on the internet to extract the parameters of interest. The result is a graphic showing the number of reachable routes from the AS and three most active links.

In this thesis the first task was to pick the ASes on which we will perform our analysis. We decided to choose AS 701 (UUnet Technologies, prefix 157.130.10.233), AS 7018 (AT&T, prefix 12.0.1.63) and AS 1239 (Sprint, prefix 144.228.241.81) because they can reach most of the internet core and they do it in the minimum number of hops, [22] and because of its number of routes and the number of origin ASes they have, [23].

We also study the number of updates and withdrawals announced to the AS. When we say updates we are referring to all newly installed routes in the AS routing table, new ways of getting in touch with its neighbors, this information is transmitted to the AS neighbors to let them know which routes can be reached trough the AS; the

withdrawals contain all new unfeasible routes in the AS routing table. We also show in Section 3.4.1 that there is a correlation in time between the number of updates and withdrawals; this means that the number of updates and the number of withdrawals are likely to be similar in time. We present two graphics for this study: in one we plot the number of updates registered by the autonomous system along with the number of withdrawals recorded by the AS; this figure reveals the similarity between the number of updates and withdrawals in time. In the second graphic we show the correlation existing between these two parameters. This relationship is important since it lets us know that the AS keeps similar number of ways to reach all its destinations. We use the correlation to show similar are updates and withdrawals, the figure shows the result of this comparison.

We also present a figure where we plot the number of routes seen by the AS along the three month period. In this figure we combine the number of updated and withdrawn routes letting us see the number of total routes reachable once the Autonomous System has taken notice of these announcements. Studying the number of updates and withdrawals gives us the opportunity of foresee abnormalities in the reachability of the AS.

We also study the time between updates from an AS point of view and from the three most active links point of view. We plot the difference between the times the AS registered one update until it receives another. This figure will reveal how many time elapses once the AS has changed its routing table until a new update is revealed to it forcing the AS to change its routing table again. There is a particularity of BGP protocol that has to be taken in consideration regarding the time between updates. To control the number of updates and reduce the processing in routers, it is recommended that BGP enabled routers set a timer called MinRouteAdver (minimum route advertisement) to 30 seconds. This timer establishes the minimum time a router has to wait before sending BGP updates to its neighbor regarding the same destination. Therefore, if a route is unstable and it is changing constantly in a short period of time, instead of flooding its neighbors with updates regarding the same routing change, the router needs to be waiting 30 seconds between them. Reducing the processing in routers improves the BGP convergence. Our study reveals that the ASes selected in our study do have this timer set. Several studies, [8-12] have examined the dynamic behavior of interdomain routing and have highlighted the negative impact of unstable routes.

Recently, attention has turned to the internet which seems to display quite a number of power-law distributions: the number of visits to a site, the number of pages

within a site, and the number of links to a page, to name a few. If the distribution in a plot is so extreme that the distribution curve be a perfect L shape and the same plot, but on a log-log scale the same distribution showed itself to be linear this is the characteristic signature of a power-law.

We use the term reachability throughout this thesis to denote the number of routes the AS has in its routing table at a given moment of time, giving it the opportunity of getting in touch with another ASes.

Throughout the rest of this thesis we'll use the term announcement to make reference to both updates and withdrawals.

3.1 Link Rank Data

In order to monitor global reachability, analyze BGP dynamics, and promptly detect routing failures, many individual ASes monitor their local view of the BGP routing system; some ASes also provide public access to local routing table snapshots, referred to as looking glasses. A few passive BGP monitoring sites, such as RIPE and RouteViews, have also been established to collect BGP update data. These monitoring sites peer with BGP routers of various ASes to passively collect the updates from the peering routers, and make the resulting BGP log data publicly available.

Although publicly accessible looking glasses and BGP update logs provide potentially useful information into the operations of the Internet routing infrastructure, it is not easy to make effective use of the available data. First, due to routing policies, each vantage point tends to have a different view of BGP reachability and routing activities, and observations at a particular router does not reflect the state for the rest of the Internet in general. Second, there is no easy or clear way to combine these individual views into a coherent picture of the global routing changes. Third, the data volume is large, millions of routing updates are generated daily and there is no easy way to extract information about most important or most relevant routing changes.

.

That's why we decided to take the processed routing information of Oregon Route Views from Link Rank to conduct our work since it contains the parameters we want to study: AS origin, AS destination, updates, withdrawals and number of routes seen by the AS at the time it captures an announcement from the interdomain. Link Rank

extracts the total number of routes carried over individual links in the Internet topology, called link weight, and measures the changes in the number of routes on each link as a way to capture aggregate routing changes. To reduce the data size to a comprehensible level, Link Rank uses an input-filter to extract the most important or relevant routing changes from the large amount of routing data.

In order to illustrate the idea of how Link Rank works lets take a look at Figure 3.1

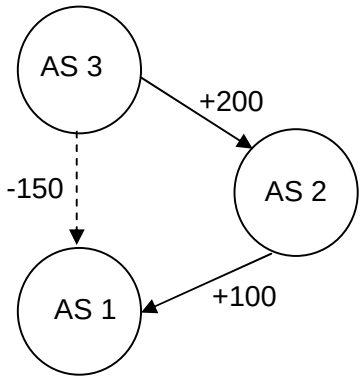

Figure 3.1 A BGP dynamics captured by Link Rank

In Figure 3.1 we have three ASes interconnected among them. AS3 decides to withdraw 150 routes from AS1, at the same instant that updates AS2 routing table by adding 200 routes to its routing table. This could be a result of routing policies or just because the link between AS3 and AS1 became unreachable. AS2 routing table is filled in with the updates received by AS3, this means AS2 has learned or gained 200 routes to reach AS3. But these are not the only routes AS2 has to proclaim to AS1 along with these new routes learned from AS3 it broadcast 100 routes to AS1 that maybe routes learned from another AS or just 100 new ways in which AS1 can reach AS2.

Of course this process doesn't occur every time that a AS learns new routes to get in touch with others ASes, if so the routing tables would experience massive burst of updates and the time to deal with them would affect the BGP convergence time. To control the number of announcements propagated in interdomain routing BGP encloses a particularity, a timer called MinRouteAdver which can be set in all routers running the BGP protocol, this timer is recommended (not obligatory) to be set in 30 seconds, forcing BGP routers to wait 30 seconds to announce new routes to its neighbors, this because in some period of time a router could receive duplicated paths to reach the same AS.

3.2 Selection of ASes and period of time

The first task once we know what the data obtained tells us, is to pick the ASes on which we will perform our analysis. We decided to choose AS 701 (UUnet Technologies, prefix 157.130.10.233), AS 7018 (AT&T, prefix 12.0.1.63) and AS 1239 (Sprint, prefix 144.228.241.81) because they can reach most of the internet core and they do it in the minimum number of hops, [22], and because of their number of routes and the number of origin ASes they have, [23].

We believe that taking three of the most active ASes in the Internet will be enough to tell something about the dynamics happening in the core of Internet, but for how long do we have to observe the behavior of these ASes?

We decided to take a three months frame as they did on, [24]. We choose the months of January, February and March of 2007, the months were chosen in a total arbitrary fashion. Link Rank logs offered information from January 2004 to June 2007 until the redaction of this dissertation.

Deciding the ASes and months for our investigation it's just the first step of the way. As soon as we decided these limitations the next step is to download information referring to these bounds. The method followed to do this is described next:

1. We access the processed data RV data on link rank. This information gathers information from various points of view in the Internet belonging to the Oregon Route views.

2. The information showed in the directory is given away in a year/month fashion. We can select the period of our interest; first we pick 2007/01.

3. Once we opt for the interval of interest we are ready to retrieve information referent to our three ASes of interest. First, we download the information concerning to the prefix 12.0.1.63 (AT&T) for that month in particular, once we have this information we carry on downloading the information for the prefixes 157.130.10.233 (UUnet) and 144.228.241.81 (Sprint).

4. When we are done getting the information for January 2007 we repeat the step two but instead of selecting this month we select February and repeat step three. We repeat step four for March 2007.

3.3 Parsing and ordering the data

The data obtained from these dumps contains information unnecessary for fulfilling our analysis such as prefixes (we only need the AS number of the prefix), duplicated routes, AS-paths, and some more parameters that are not of our interest. So there's need to parse the dumps in order to leave only the statistics we want to observe which are:

- Origin AS number.
- Destination AS number.
- Time of the change in the interdomain routing table.
- Number of updates and withdrawals.
- Number of routes reachable by the AS when an update or withdrawal takes place.

The result of parsing the data gives is illustrated in Figure 3.2

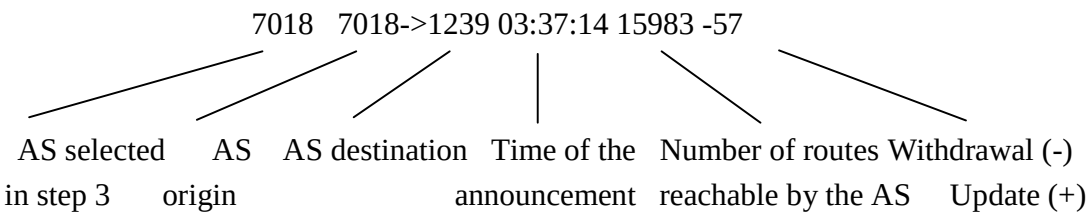

Figure 3.2 Structure of the parsed data

We have already explained these attributes. As soon as we have all of our data parsed we continue to join together the three month information regarding to each of our ASes previously indicated. We had three files in each AS; each one contained the information recollected for a specific month: January, February and March. We joined up these three files into one ordered chronologically. The reason of this sorting is that if we arranged the data in this way we are able to observe and study how many routes were updated or withdrawn from the AS from the first announcement captured in January until the last announcement registered in March, giving us the opportunity of conduct our study in a sequentially manner. After leaving our numbers parsed and ordered we imported the data to MATLAB to begin with the study of the recollected records.

3.4 Analysis

The way we organized our data leave us in position of start our analysis. We import the data from a .txt file and organize it in vectors. We will have six vectors with significant information, each vector will store the data showed in Figure 3.2: vector 1 will have the AS selected for the study, vector 2 will stack AS origin and so on. One problem we faced was to give continuity to our data in vector 4. This was because the logs from Oregon Route Views contains the information as we said in a year/month way. The registers contain the data organized by days of the month, to try to explain ourselves better we offer Figure 3.3

7018 3549->22822 02:03:05 3 -98
7018 7018->3549 02:03:05 7662 -102
7018 3549->22822 02:06:36 66 +63
7018 7018->3549 02:06:36 7726 +64
7018 7018->1239 03:37:14 15983 -57
7018 7018->10888 04:21:16 53 +53
Figure 3.3 An example of a parsed log

Figure 3.3 shows a segment of the parsed data corresponding to January first 2007 from the AS 7018 (as we can see in the first column), in the figure we count 6 lines corresponding to 6 announcements revealed to the AS 7018. What the firs row indicates us is that AS 7018 received an announcement from AS 3569 towards AS 22822 at 02:03:05 Greenwich Mean Time (GMT). The next number in the row (vector 5) specify the number of routes or ways that AS 7018 knows to reach AS 22822 through AS 3549 after this last one had withdrawn 98 routes from AS 7018 routing table. We had to convert this GMT standard time to a decimal standard time so we were able to sum the quantities and give continuity to our framework, otherwise we would’ve had several times repeated.

After solving this issue we can proceed to make our first analysis.

3.4.1 AS 7018.

We are interested in two BGP dynamics: time between announcements and fluctuation of reachability along time. In this section we will examine the statistics referent to the AS 7018 and its dynamics. In this part of the analysis we took all the

information of the AS, all the links established between the ASes, all announcements and routes that AS 7018 receives along the three months period.

First we wanted to know the behavior of the AS reachability and announcements in the three months period. How it looked like in a plot of Reachability and Announcements vs time. We took our data and plotted our information in the next figure.

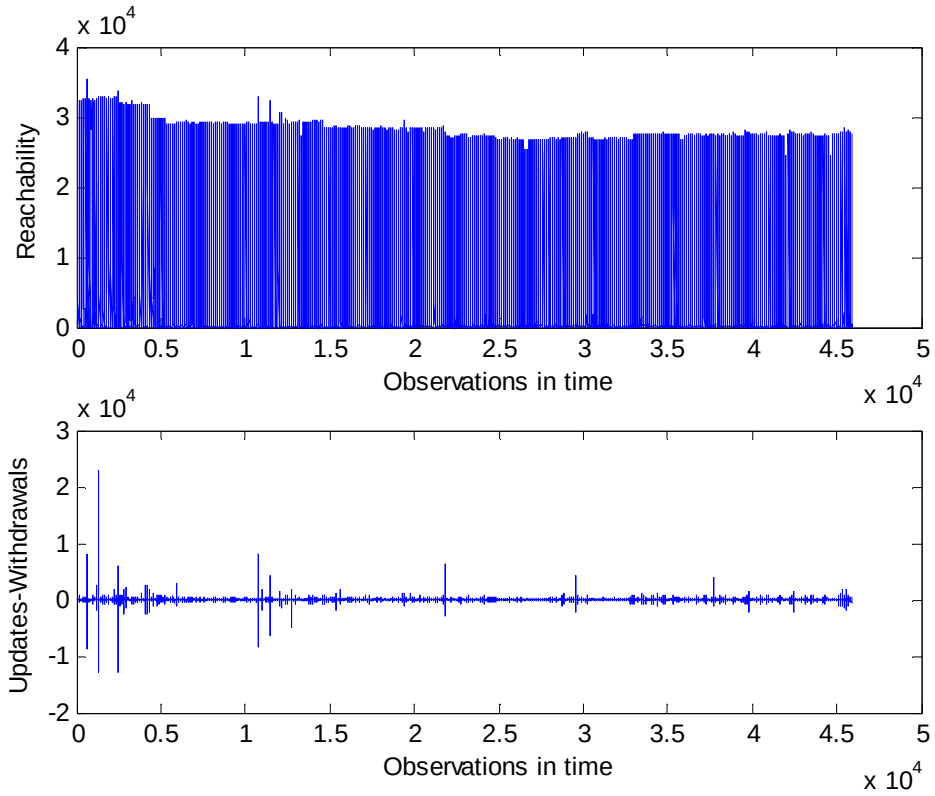

Figure 3.4 (a) Reachability (b) Updates-Withdrawals from AS 7018

Notice how our simulation starts with reachability above zero, this is because when we started our simulation this AS had already known certain amount of routes to get in touch with other ASes. As we explained in section 3.2 we took January to start our study. This means that we recollected data from the first announcement of that month, this announcement had information updated from the last announcement of December, at that point the AS had announcements from the interdomain and routes to get in touch with other ASes.

Along the three months of study the AS 7018 received several announcements from the interdomain, every time the AS received an announcement we took it to

represent it in our plots. That's why on the 'x' axis we have observations in time and not time itself.

AS 7018 reported 351,028 announcements through the three months period. In Figure 3.4 (a) we observe the total routes the AS7018 knows of, time zero means the beginning of our analysis, look how the AS in that moment already had certain number of routes learned from before we initialize our investigation. We had to make an adjustment to our time vector to plot this figure correctly. As we saw in Figure 3.3 the AS can receive more than one announcement at the same time. In Figure 3.3 we observe that at 02:03:05 AS 7018 received two announcements: one indicating it that AS 3549 had removed 98 routes to reach AS 22822 and the other specifying 102 withdrawals from AS 7018 to AS 3549. These withdrawals happened at the same time, so, if we had taken these events as separated ones maybe the reachability in Figure 3.4 (a) would be half than it's showed in the plot but the time it's plotted against would be double than it's.

In Figure 3.4 (b) it's shown the number of updates and withdrawals that occurred along the time of our three months of study. The lines going up of the horizontal axis are the updates and the ones going down are the withdrawals. Look how the reachability of the AS is directly affected each time that occur an announcement, if we look carefully we can see that the peaks that the reachability presents along time correspond to triggers of updates or withdrawals that took place during in that time. If Figure 3.4 (b) maintains a "stable" tendency the reachability doesn't seem to have discontinuities. We can perceive that in time the AS collects updates and withdrawals from the interdomain (Figure 3.4 (b)) but the number of updates is not more significant than the number of withdrawals to cause a crest in Figure 3.4 (a). If the number of updates by far exceeded the number of withdrawals we would observe two main changes in our plot. In figure 3.4 (a) there would be a peak or peaks where the number of updates surpassed the number of withdrawals or there would be a valley where the number of withdrawals surpassed the number of updates. We would observe something similar in figure 3.4 (b). Which is one of the main reasons we are interested in studying the updates and withdrawals: if we can characterize this parameters we will be able to tell if the AS is suffering anomalies on its routing, events that harm its reachability such as attacks, [25] misconfigurations, [26] or power failures, [27] incidents that can lead to anomalous or pathological routing behavior and that can affect the global routing infrastructure.

Figure 3.4 (a) tells us that even though the AS 7018 gain and losses routes, it always finds a way out to the routing packets crossing it.

So we took our vector containing updates and withdrawals and we split it in two. One resulting vector will enclose the updates and the other will have the withdrawals. When we separated this vector we made sure that the column of updates and withdrawals conserved all the information of the row they corresponded to. The reason of doing this was to observe updates and withdrawals in one unified plot on the same vertical axis and try to see if there was some type of relationship between them. The result of this process is shown in Figure 3.5

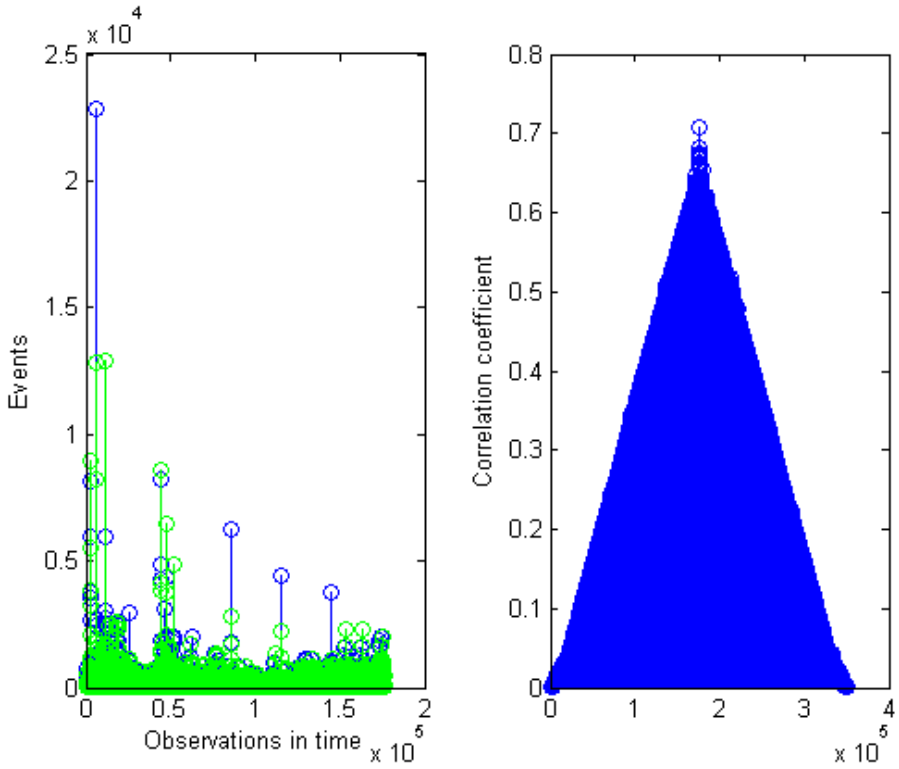

Figure 3.5 (a) Updates and Withdrawals (b) Correlation

We perceive that the number of updates is almost the same as the number of withdrawals in time, except in some part of the plot where there are obvious disparities. The most noticeable is at the beginning of the plot, where the number of updates surpasses all the values of the rest of the plot. This means that in that instant the AS 7018 received the highest number of updates of routes from the interdomain. In Figure 3.5 (b) there's a plot of the correlation between the updates and withdrawals of the AS 7018, we can see that there is correlation between the vectors of updates and withdrawals since the

resulting shape of a bell. This means that our vectors of updates and withdrawals have the similar shape or values.

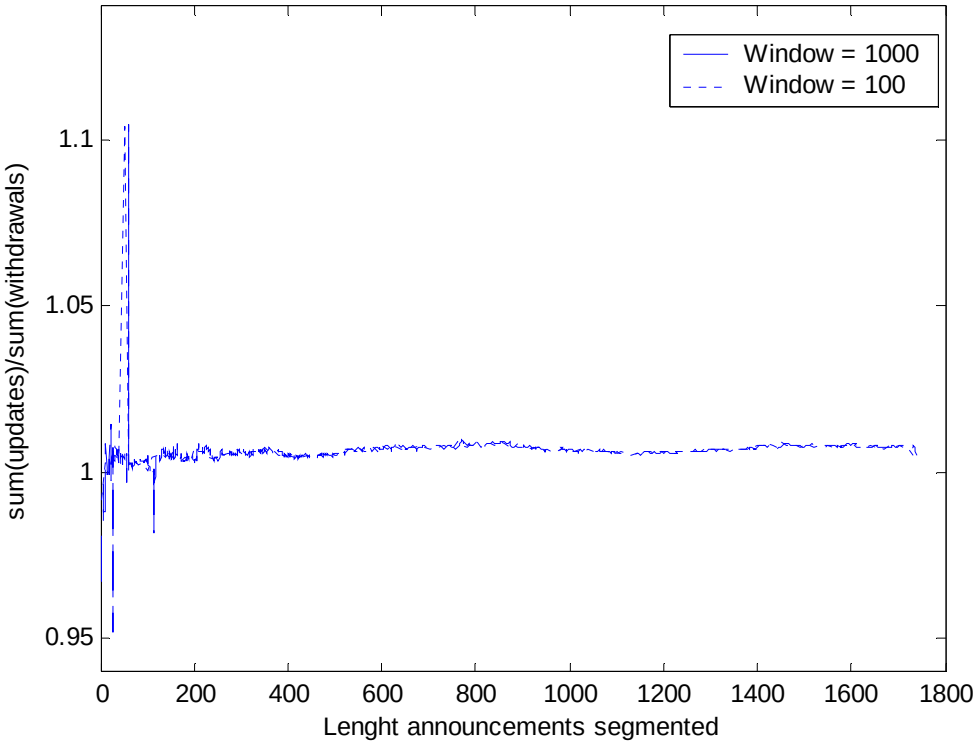

Figure 3.6 Curves of ratio updates/withdrawals

We show in figure 3.6 the ratio between the sum of updates and the sum of withdrawals. We took our vector of announcements and we split it with windows of 100 and 1000. This means that every time the length of the vector reached 100 or 1000 we took sum the updates and the withdrawals to divide them, we notice that this ratio is around 1 except at the beginning of the plot where as saw in figure 3.4 (b) occurs the highest number of updates and withdrawals.

After doing this observation we decided to calculate the frequency of our updates and withdrawals. To do this we draw in the same picture the histogram of these vectors. This histogram is the graphical version of the vectors of updates and withdrawals and it reveals us the distribution of data values within the vectors. By doing this we will be able to observe the values incidence of our vectors.

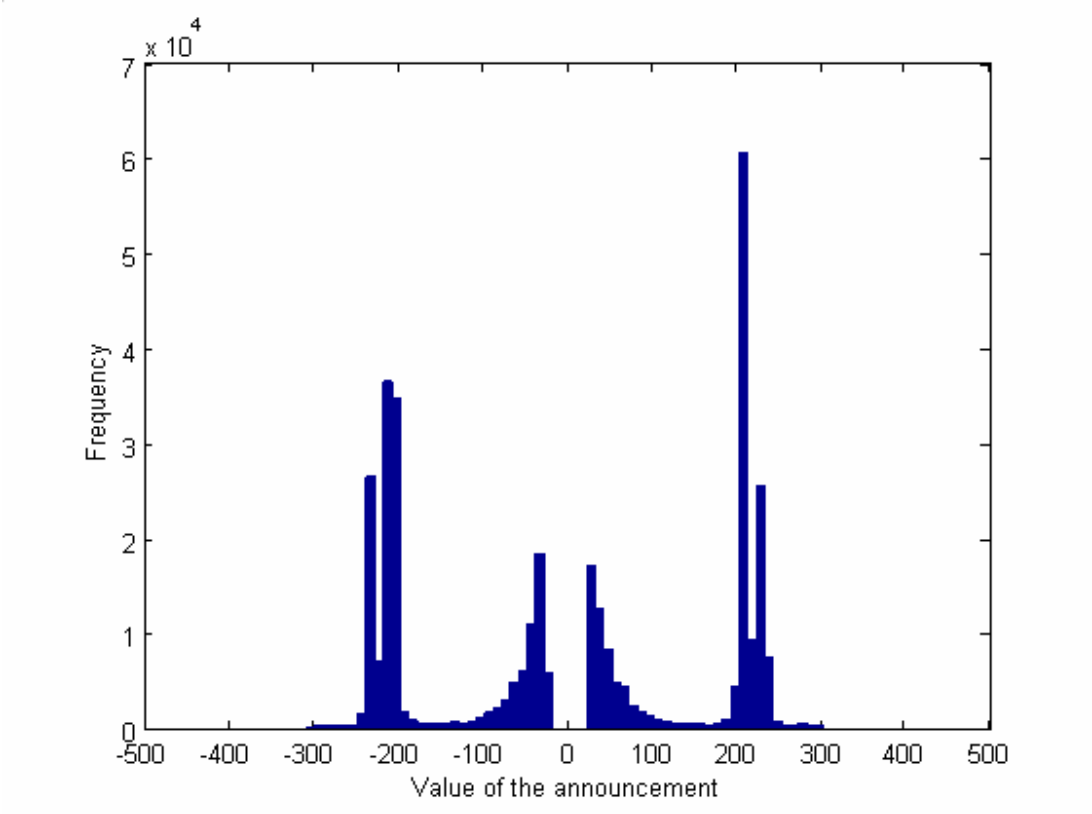

Figure 3.7 Histogram of Updates and Withdrawals

In Figure 3.7 we can see the frequency of number of updates and withdrawals. The withdrawals are plotted on the left side of the picture (negative values of x) and the updates on the right side of the picture. These are not the only values registered in our vectors of announcements but they are the ones that have the most occurrences of all. In the picture we can notice that the AS 7018 received almost with the same frequency the numbers of updates and withdrawals. If the AS aggregated 100 routes to its routing table it was the same number of routes that it withdrawn from it. Look how the frequencies of updates and withdrawals are pretty similar in Figure 3.7, this leads us to think of some relationship between the number of updates and withdrawals in the interdomain.

We have shown the behavior of several parameters in the interdomain: reachability, updates, withdrawals and their histograms. With these parameters covered we can carry on to study the time between announcements and fluctuation of reachability along time.

Next we illustrate the time between updates and time between withdrawals from the AS 7018.

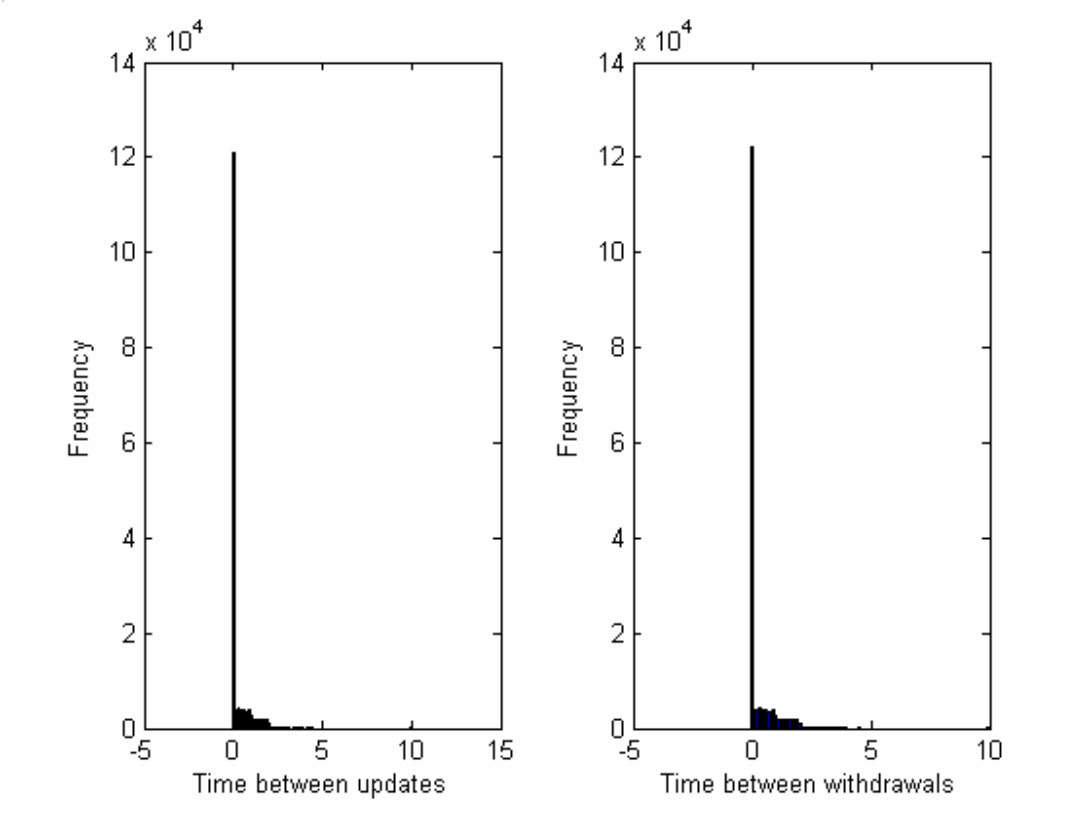

Figure 3.8 Histogram of time between (a) Updates and (b) Withdrawals

We recall that we recollected 351,028 announcements in our period of study. These announcements where both updates and withdrawals but we don't know how many updates or withdrawals correspond to that quantity of announcements. Since we want to study these elements not as a single issue we need to separate them. The number of updates was 174,979 and the withdrawals were 176,049. Although the number of withdrawals exceeds the number of updates it's not a significant number (1,070). This doesn't mean that the AS lost connectivity at all in any period of the study. We can explain this disparity with Figure 3.4 (a); all along the period of study we noted that the AS reachability suffered little changes depending on the number of updates or withdrawals but the AS never suffered considerable or abrupt changes in its reachability. Therefore this dissimilarity between the numbers of announcements isn't enough to decree an abnormal phenomenon in the AS reachability.

In Figure 3.8 we illustrate the histogram time between updates and withdrawals.

In the image we can notice that the time between announcements is zero. The reason why this happens is maybe because in this part of our study we are considering all

the announcements received by the AS. If we remember the explanation given above for Figure 3.3 we have announcements happening at the same time. We expect this behavior to change once we make our analysis on the three most active links of the AS.

We plot our histograms of figure 3.8 on a loglog plot since doing so has the virtue of all the data is plotted, even the tiny one.

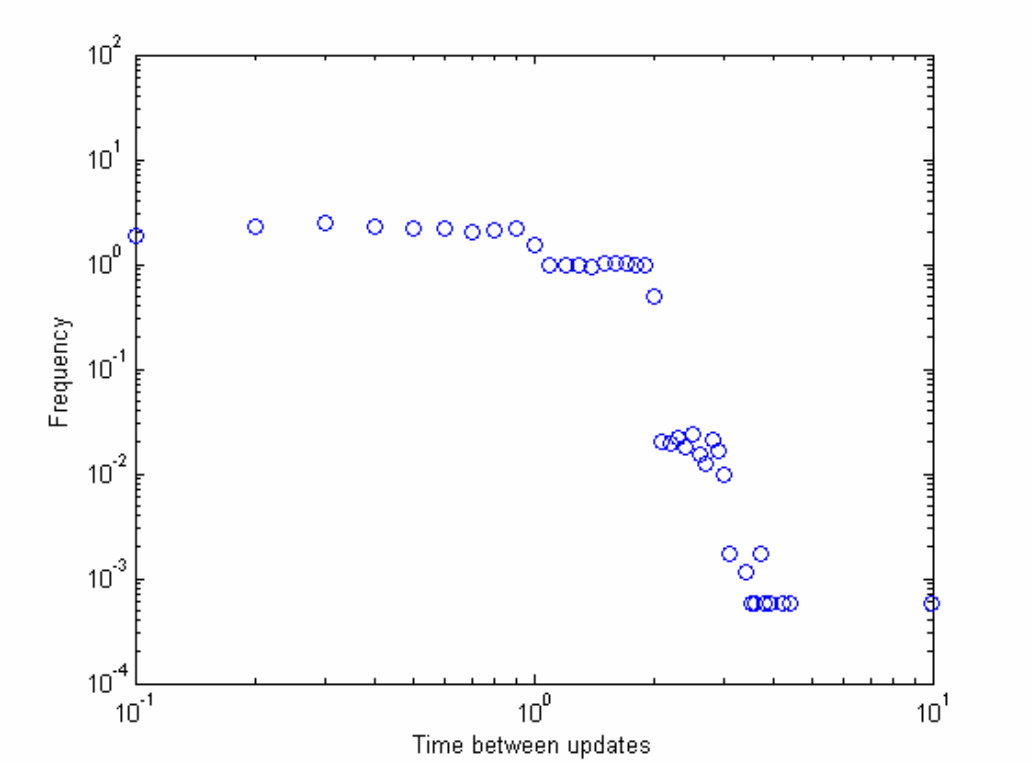

Figure 3.9 Loglog plot time between Updates

By taking a look at Figures 3.9 and 3.10 we can see that not only the time between updates and withdrawals is near zero, but also that as mentioned above since we are taking all the announcements from the interdomain and several of them may arrive at the same time we took them as one causing this famine of data in our plots.

We have studied the behavior of one the BGP dynamics we're interested into: time between updates and withdrawals. At this moment this lack of data doesn't tell us anything about this parameter. In the next section we'll study the same dynamic from a link point of view.

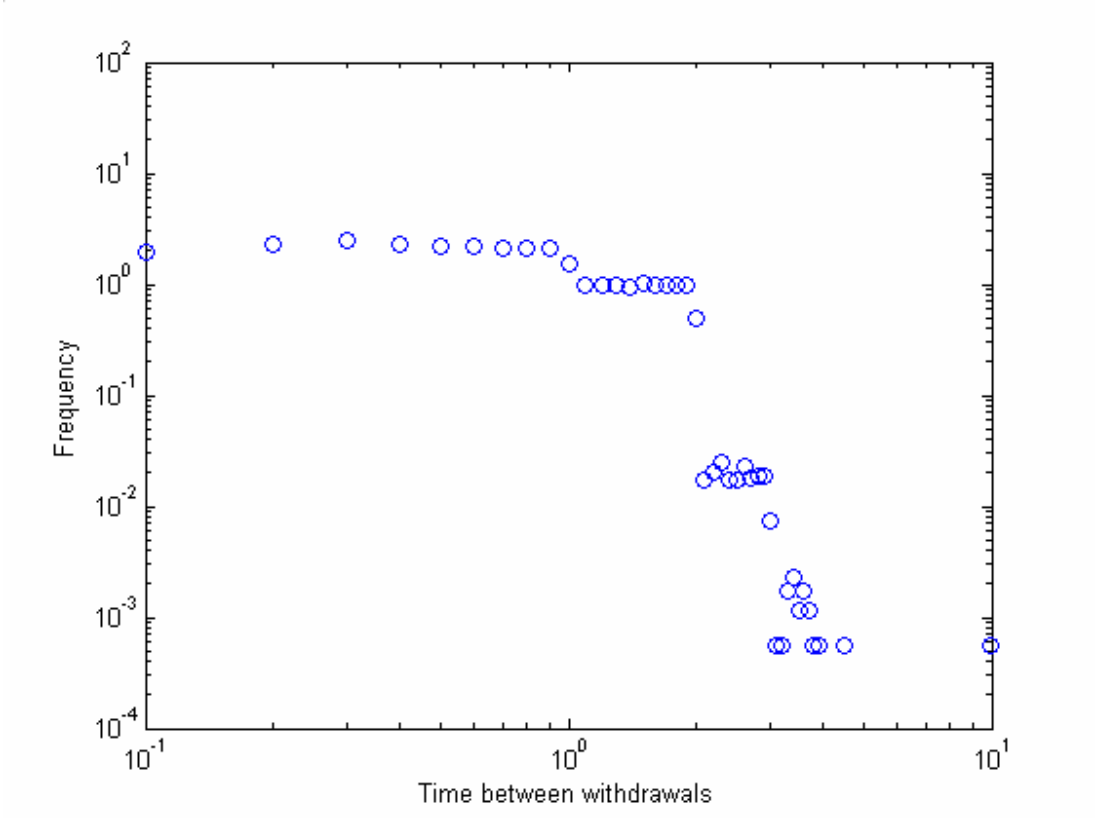

Figure 3.10 Loglog plot time between Withdrawals

Once again we remark the similarity between updates and withdrawals; previous to this analysis of time between updates and withdrawals we showed association connecting these parameters.

The other dynamic we wish to evaluate is the difference between routes once an update or withdrawal has occurred in the interdomain. Notwithstanding we took the number of announcements arriving at the same time as one and sum all the values of routes corresponding at that instant of time generate Figure 3.4, we cannot take the same approach when it comes to routes reachable through time since each announcement represents a change in the interdomain topology. If we keep the approach taken to calculate the time between announcements we would be loosing noteworthy data from the fluctuation between updates and withdrawals.

With this basis we next study the frequency between number of updates and withdrawals in the interdomain.

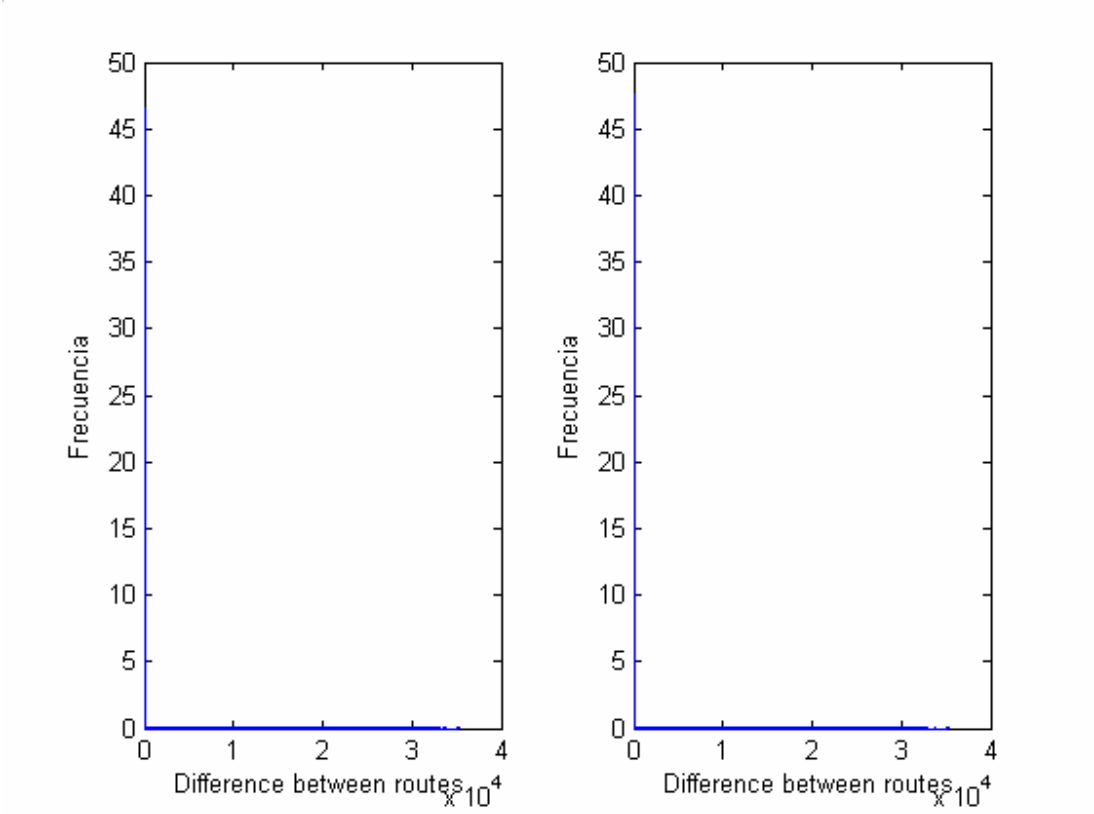

Figure 3.11 Histogram of difference between (a) updates and (b) withdrawals

In Figure 3.11 we take notice that almost all the events occur near zero, this plot is similar to Figure 3.8 were we displayed the time between updates and withdrawals with the variation that the distribution is so extreme that the curve is almost a perfect L shape which is the characteristic signature of a power-law.

In Figures 3.12 and 3.13 we'll show the same plot in but on a loglog scale.

Figure 3.12 represents the difference between updates in the course of January to March. Even though it shows some linearity at the beginning of the plot there's a discontinuity in the plot that leads us to think twice about this plot having a power-law behavior. On the other hand Figure 3.13 doesn't show this disparity and although it has messy tail end of the distribution, this linearity at the beginning of the plot is a symptom of a power-law behavior.

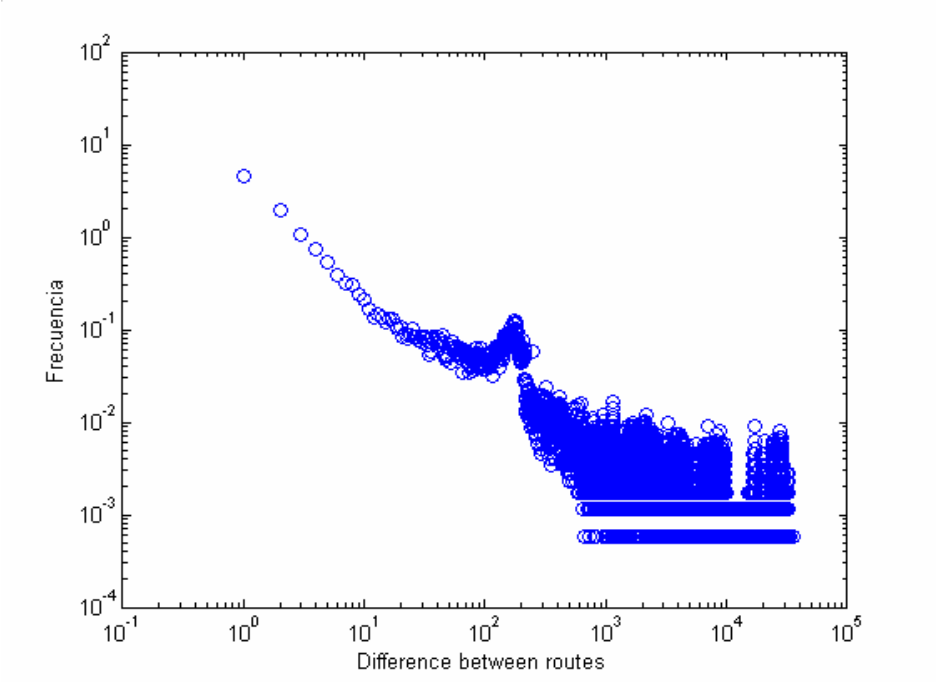

Figure 3.12 Loglog plot difference between Updates

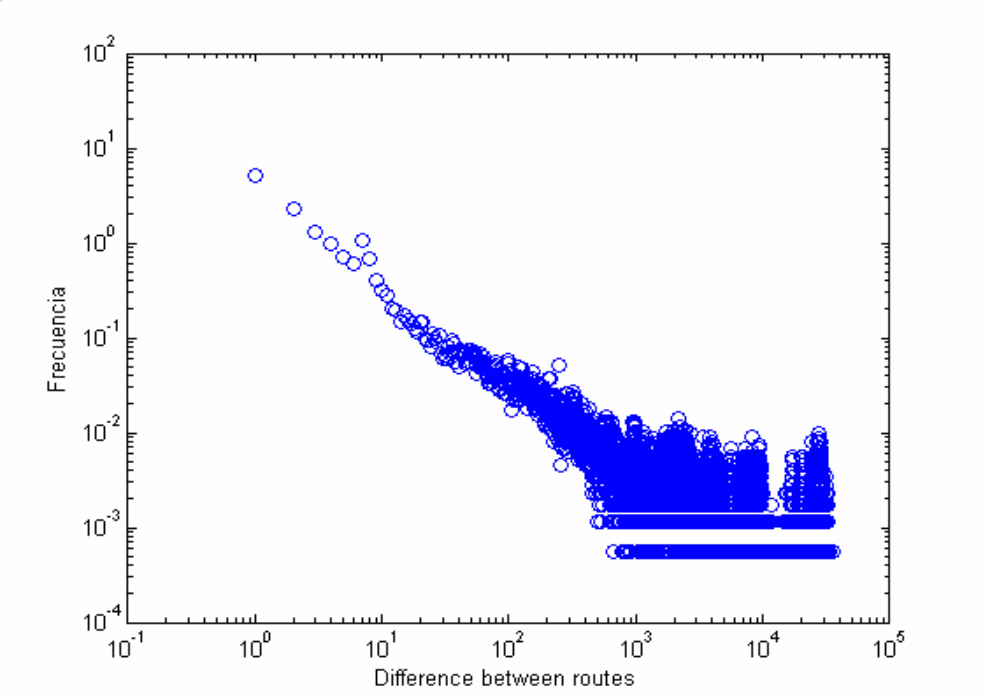

Figure 3.13 Loglog plot difference between Withdrawals

3.4.1.1 Study of the three most active links in AS 7018.

We have seen the behavior of BGP dynamics from a AS point of view. Unfortunately some problems arose when we study time between updates and withdrawals (lack of data) and some discrepancy when we plotted our loglog figure corresponding to the difference of routes reachable by the AS.

The intention of studying the three most active links in AS 7018 is to try to find out if we should investigate BGP dynamics in the interdomain from an AS or from a link point of view.

We'll study the performance of link 7018-1239 (4232 announcements), 7018-3356 (2503 announcements) and link 7018-701 (22,086 announcements). We'll study their parameters in the same order as we did in section 3.5.1.

If we take a look to Figures 3.14-.16 we'll see a difference from Figure 3.4. In these plots we do not observe the reachability of routes to be constant. Here's where we can notice the change in the reachability of routes. Look at the variation of reachability. If we look carefully we can notice that the number of withdrawals is greater than the number of updates. Justifying what we stated earlier that the number of withdrawals exceeded the number of updates by 1,070.

The correlation of updates with withdrawals holds for this link point of view, the result is pretty similar to Figure 3.5, we omitted their display to show how the difference between updates and withdrawals behave for this tactic of studying dynamics form a link point of view.

Coincidentally the study for time between changes and difference of routes have the same outcomes, this is because we anchor the AS origin so every time an announcement is presented the number of routes fluctuates with the same proportion.

In figures 3.17-.22 we show the time between changes in these three links. Dissimilar to figure 3.12 the figures plotting the difference between changes doesn't have any disparity.

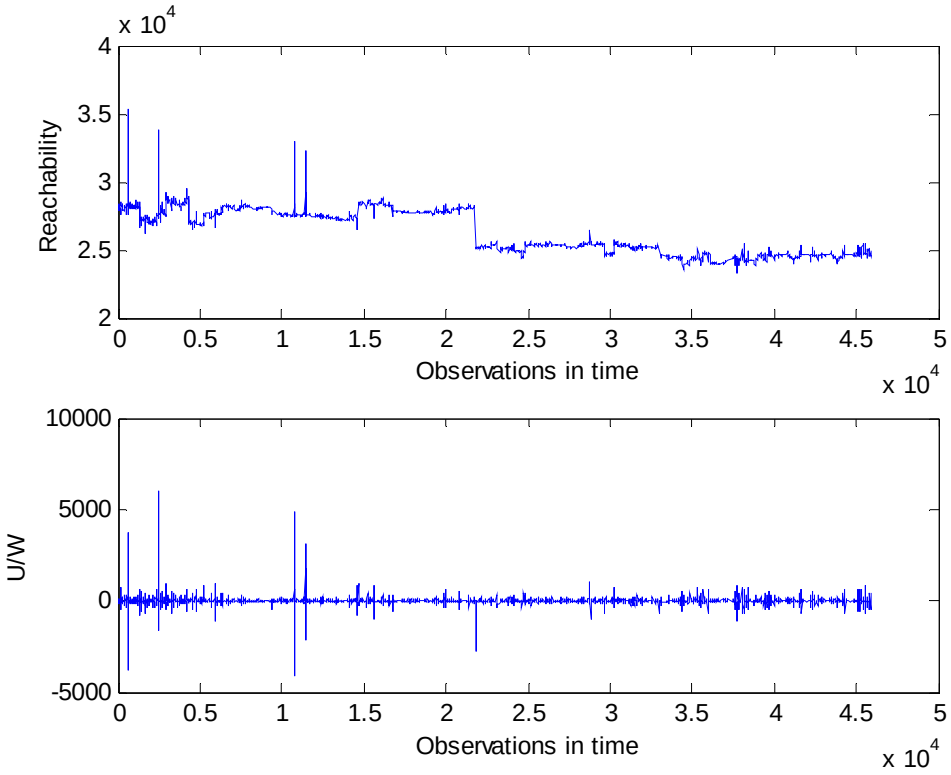

Figure 3.14 (a) Reachability (b) Updates-Withdrawals from link 7018-1239

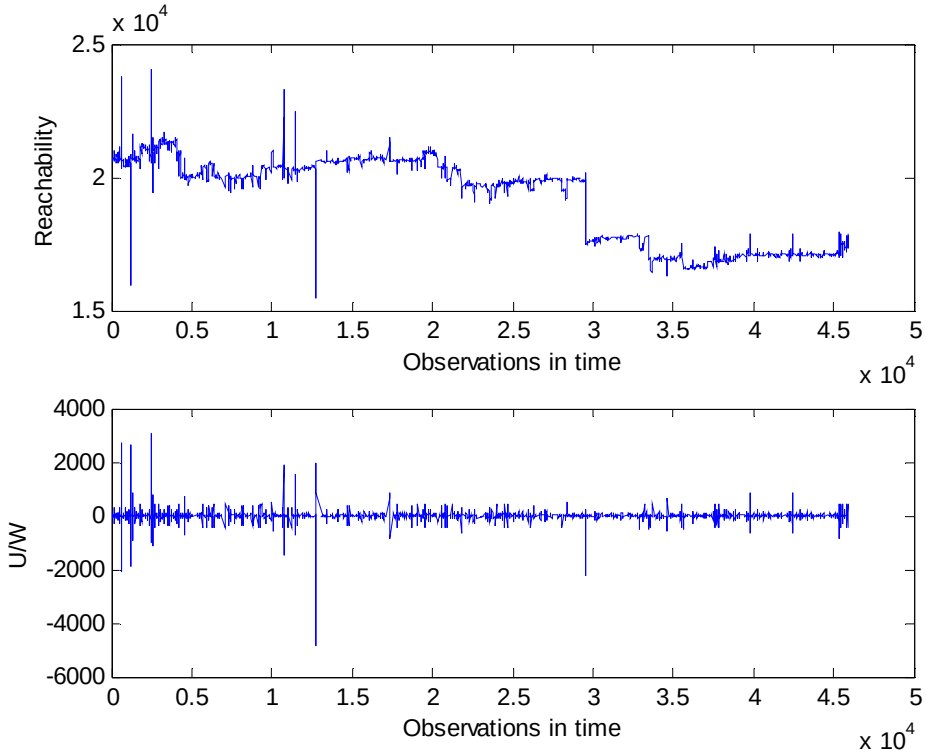

Figure 3.15 (a)Reachability (b) Updates-Withdrawals from link 7018-3356

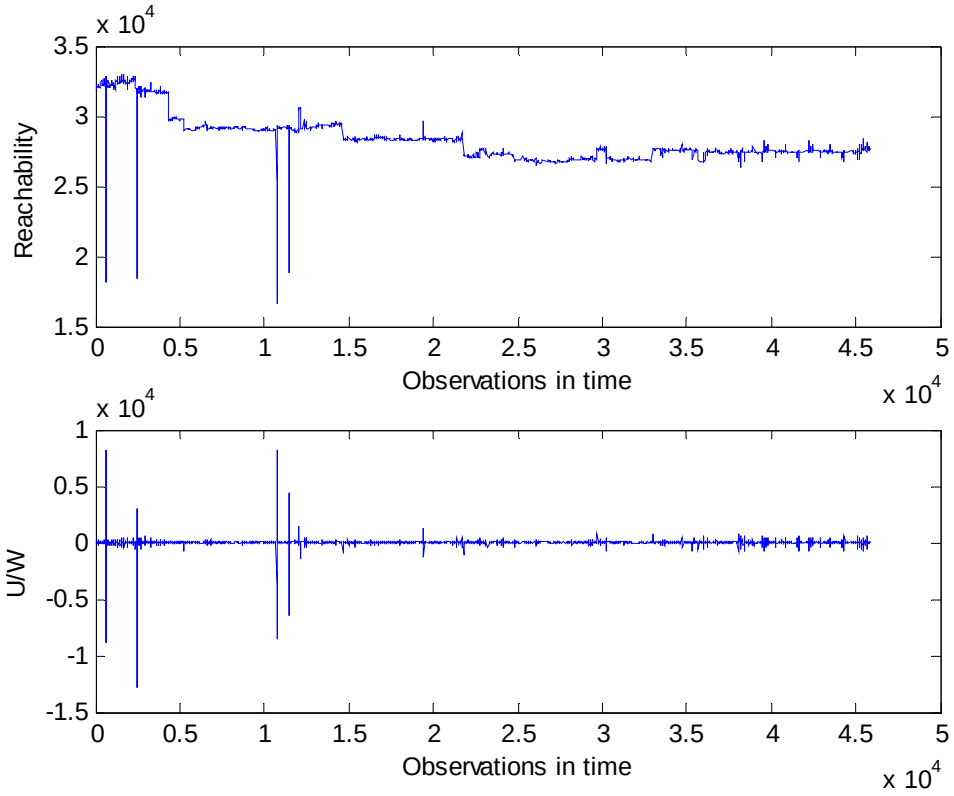

Figure 3.16 (a)Reachability (b) Updates-Withdrawals from link 7018-701

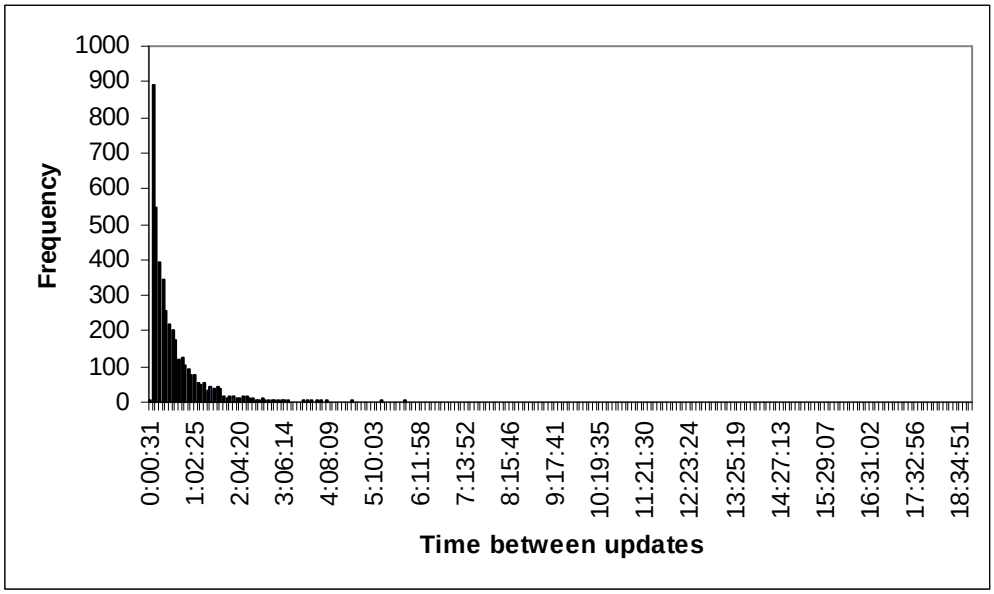

Figure 3.17 Histogram of time between Updates 7018-1239

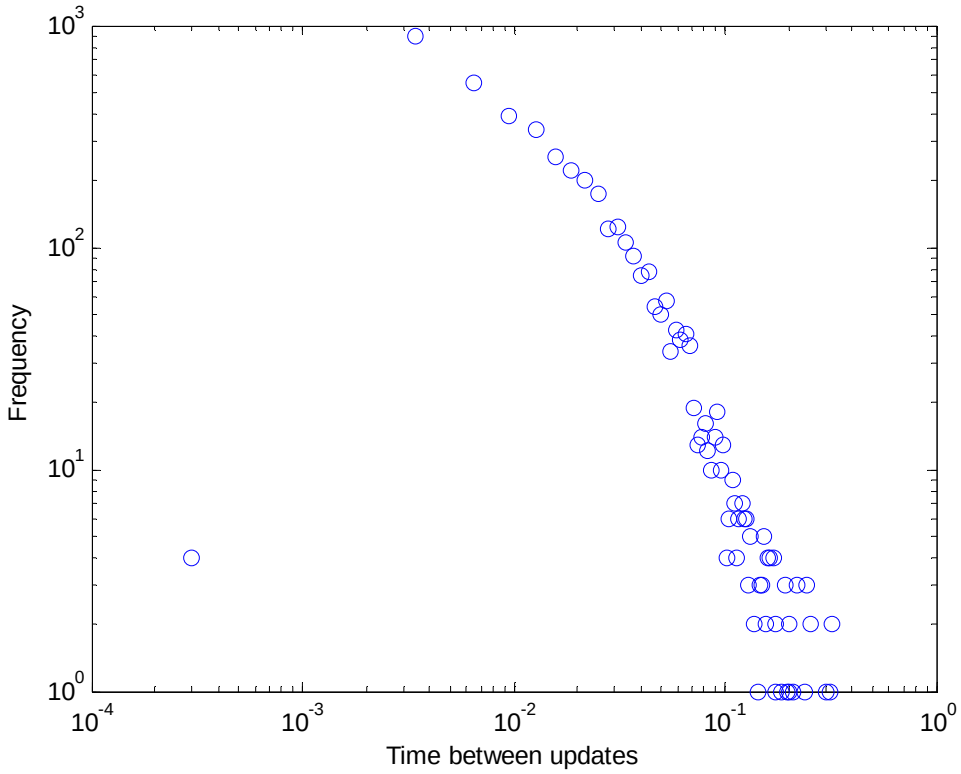

Figure 3.18 Loglog plot time between Updates

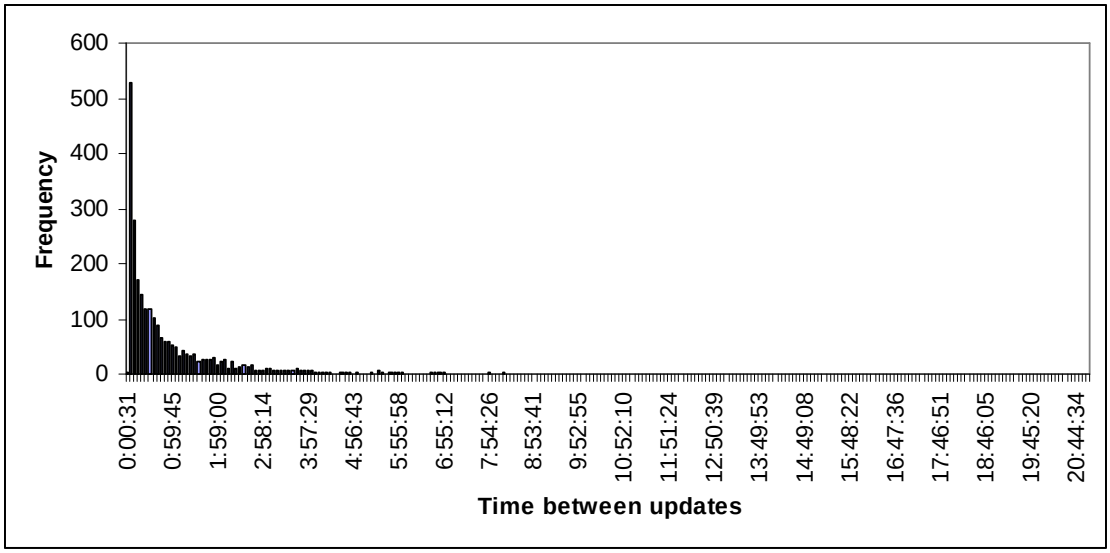

Figure 3.19 Histogram of time between Updates 7018-3356

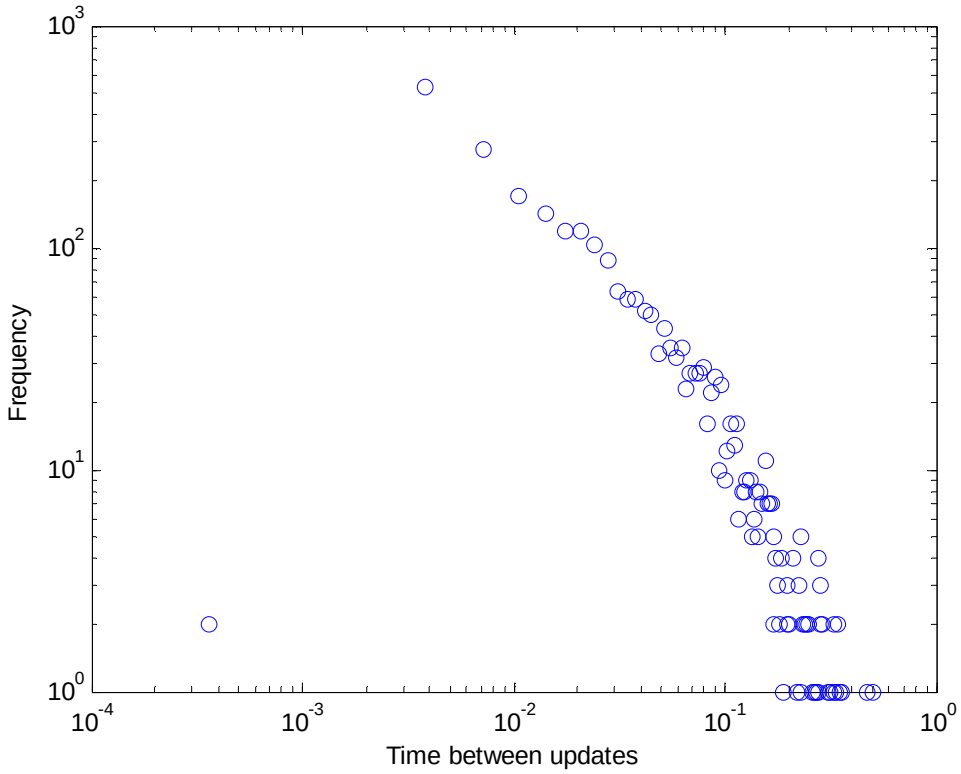

Figure 3.20 Loglog plot time between Updates 7018-1239

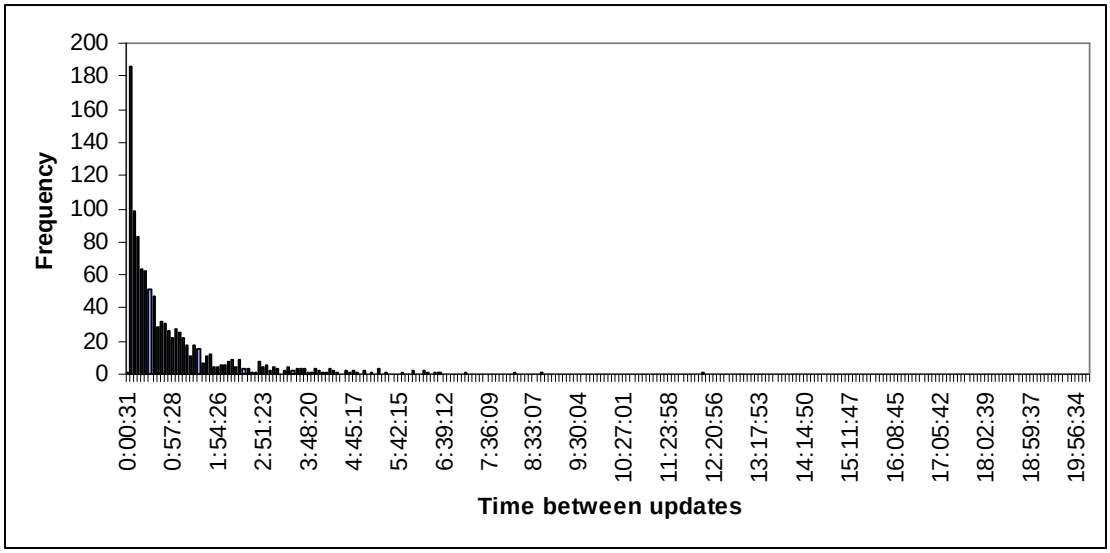

Figure 3.21 Histogram time between Updates 7018-701

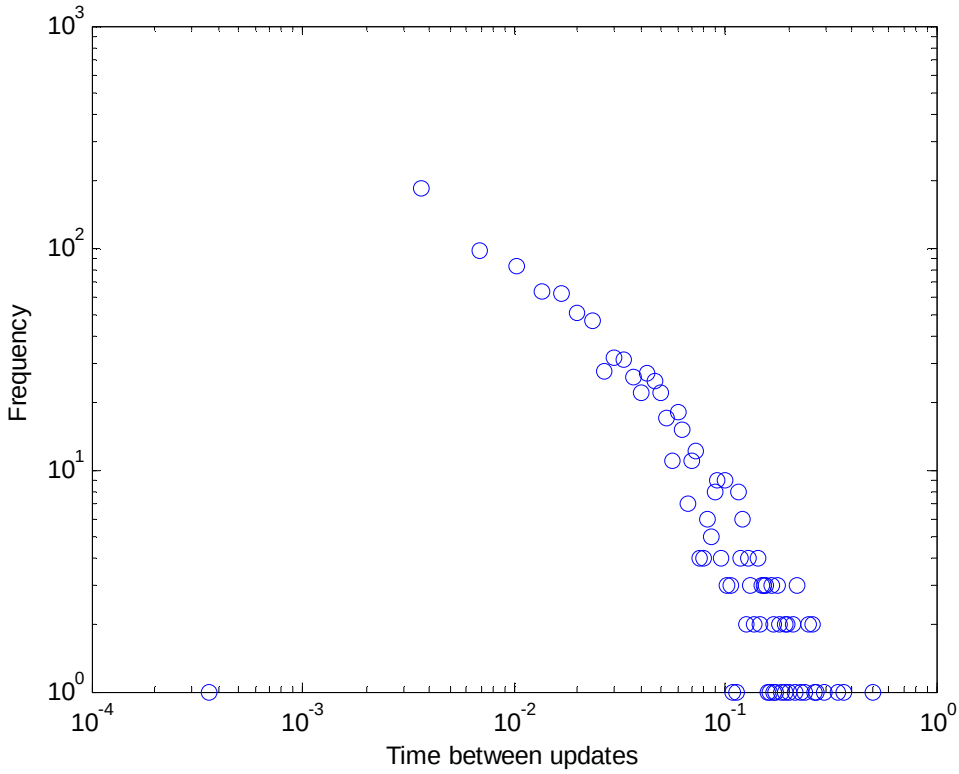

Figure 3.22 Loglog plot time between (a)Updates and (b)Withdrawals 7018-701

3.5 Internet Routing Dynamics Methodology

In this section we show and explain the flowchart of the methodology we followed to implement our tool to visualize Internet routing changes and detect ASes without the MinRouteAdver timer set.

3.5.1 Data recollection

We decided to take the processed routing information of Oregon Route Views from Link Rank to conduct our work since it contains the parameters we want to study: AS origin, AS destination, updates, withdrawals and number of routes seen by the AS at the time it captures an announcement from the interdomain. Link Rank extracts the total number of routes carried over individual links in the Internet topology, called link weight, and measures the changes in the number of routes on each link as a way to capture aggregate routing changes. To reduce the data size to a comprehensible level, Link Rank uses an input-filter to extract the most important or relevant routing changes from the large amount of routing data. In this step of the methodology we need to:

1. We access the processed data RV data on link rank. This information gathers information from various points of view in the Internet belonging to the Oregon Route views.

2. The information showed in the directory is given away in a year/month fashion. We can select the period of our interest.

3. Once we opt for the interval of interest we are ready to retrieve information referent to the AS of interest.

In this work the period of time was set to three months because we thought this was time enough to recollect meaningful data about BGP dynamics in the three ASes we analyzed in this work. However, this window of time can be set by the interest of network operators to study and analyze their dynamics, we recommend, tough, that this period would be more than 30 seconds to meet the time necessary to have new advertisements from the interdomain.

3.5.1.1 Parsing and ordering the data

The data obtained from the dumps contains information unnecessary for fulfilling our analysis such as prefixes (we only need the AS number of the prefix), duplicated routes, AS-paths, and some more parameters that are not of our interest. So there's need to parse the dumps in order to leave only the statistics we want to observe which are:

- Origin AS number.
- Destination AS number.
- Time of the change in the interdomain routing table.
- Number of updates and withdrawals.
- Number of routes reachable by the AS when an update or withdrawal takes place.

3.5.1.1.1 Updates-Withdrawals

We obtained the number of updates and withdrawals from our parsing. We acquired the number of updates for the entire AS and for the three most active links. The number of routes within the updates and withdrawals was almost the same.

We found the ratio between updates and withdrawals to be flanked by 0.985894 and 1.03089. If the ratio between updates and withdrawals overpasses 1.5 or .5 we would say that we are in presence of some anomaly, if so we would report this incident to a BGP anomaly resolver and it will take the proper action, if the ratio is between the normal parameter we can study another interest period of time.

3.5.1.1.2 Time between updates

In this step we only take the updates from our parsing and calculate the difference between the time the AS received an update and got it actualized. We study the updates and not the withdrawals because the MinRouteAdver timer only considers the updates in the interdomain routing. If the time between updates is below to 30 seconds this means that the AS announcing the updates doesn't have the timer set and consequently it has the tendency to announce unstable routes.

Chapter 4

Conclusions and future work

In this chapter we present the conclusions generated from our work and the future work needed to continue this line of research.

4.1 Conclusions

In this work we have analyzed real data from the Interdomain routing in order to obtain the similarities between the number of updates and withdrawals in time, this reveals us the number of changes on the routes seen by the ASes chosen to do this research. We also have proved that the time between of updates follows a power law behaviour.

From the analysis done in this work we can conclude the following:

- The number of updates and withdrawals in the Interdomain are highly correlated. This means that these numbers are pretty similar. We can conclude that, if the AS receives new routes to reach its destinations it also receives the notifications that it has to remove its previous routes.

- The time between updates presents a power law relationship. This power law relationship is consequence of the power law relationships previously observed in the AS level topology in other works [24].

- We cannot take the time between updates of the entire AS to try to show the power law relationship of times between updates, because it may receive several updates at the same time and as we proved in figure 3.7 this plot doesn't show a power law behavior.

- All the three ASes taken into consideration in this thesis, AS7018, AS701 and AS1239 as well as their three most active links have the MinRouteAdver timer set to 30 seconds as the standard recommends this timer provides both a rate-limiter

on BGP updates as well as a window in which BGP updates with common attributes may be bundled into a single update for greater protocol efficiency.

- Even tough the number of updates and withdrawals are highly correlated there are not the same exact number of updates and withdrawals in our statistical analysis, this doesn't affect the reachability of the system. As showed in figures plotting reachability of the ASes or of the links we notice that the reachability maintains certain level.

4.2 Future work

Even though our analysis was done off line it can be implemented in real time. In order to continue this line of work we propose the next suggestions:

- It can be taken a well know BGP event of attack and compare the data obtained from that period of time against this analysis and take notice of the differences.

- The data can be put into a data processor in order to train it with normal and abnormal BGP events.

- Once the data has been trained it can be possible to decide whether the BGP data on an AS is normal or if it is presenting abnormalities, is so the networks administrators can decide what action has to be taken in order to prevent routing failures.

Appendix A
AS 701

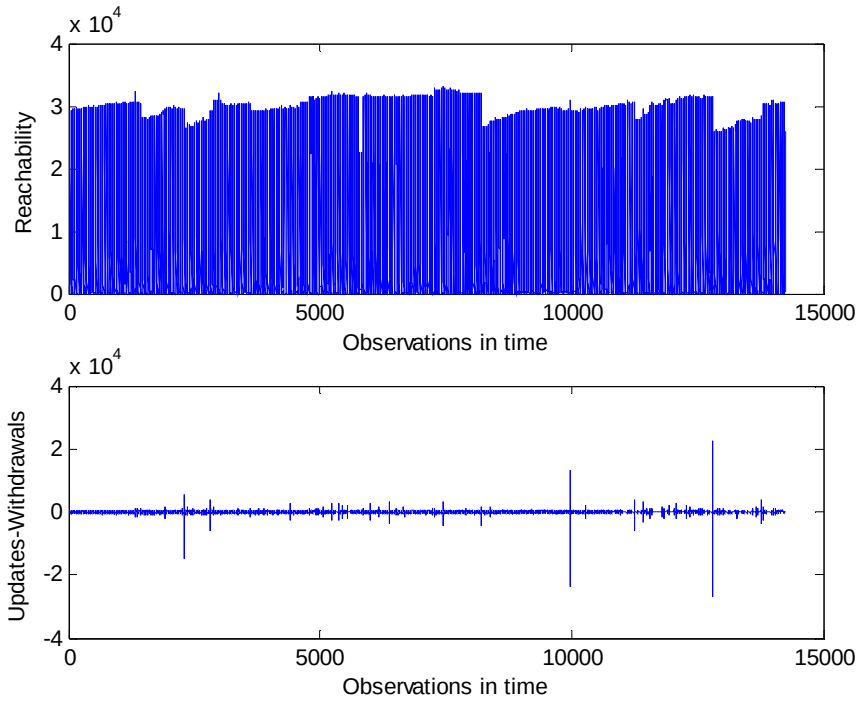

Figure A.1 (a) Reachability (b) Updates-Withdrawals from AS 701

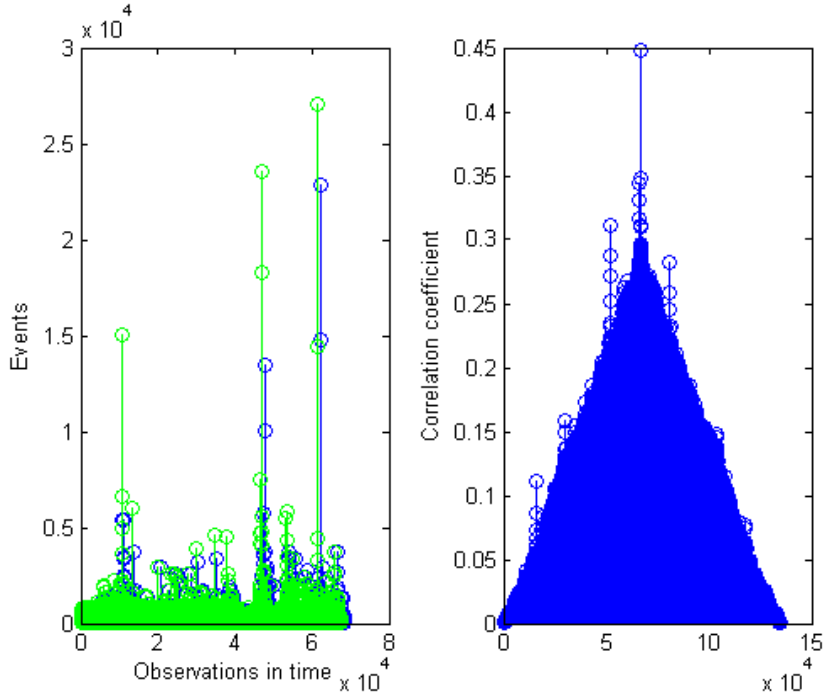

Figure A.2 (a) Updates and Withdrawals (b) Correlation AS 701

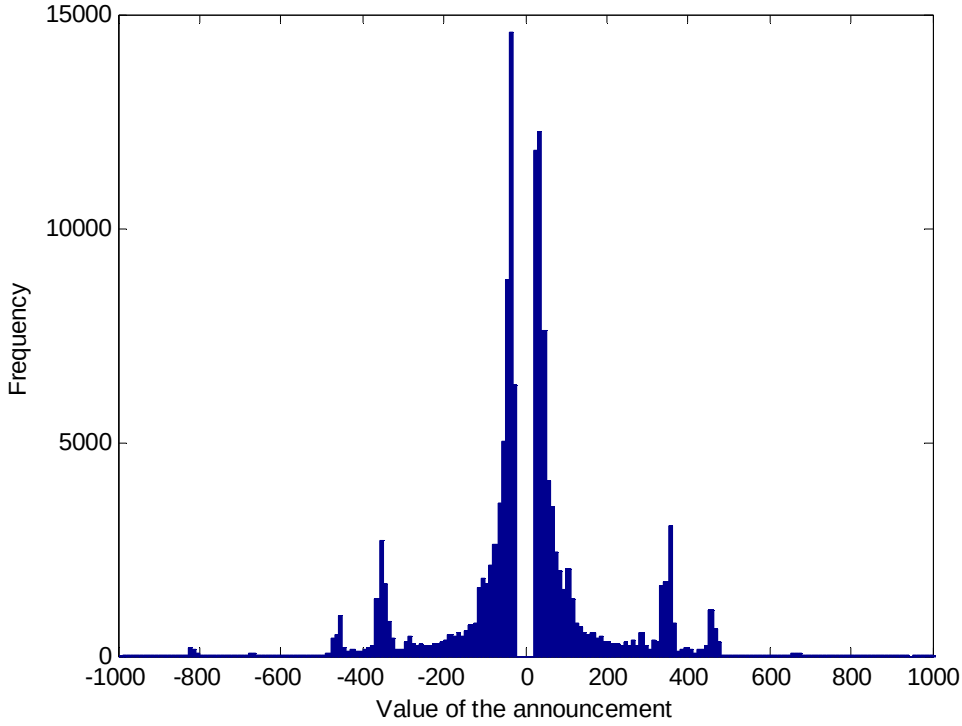

Figure A.3 Histogram of Updates and Withdrawals AS 701

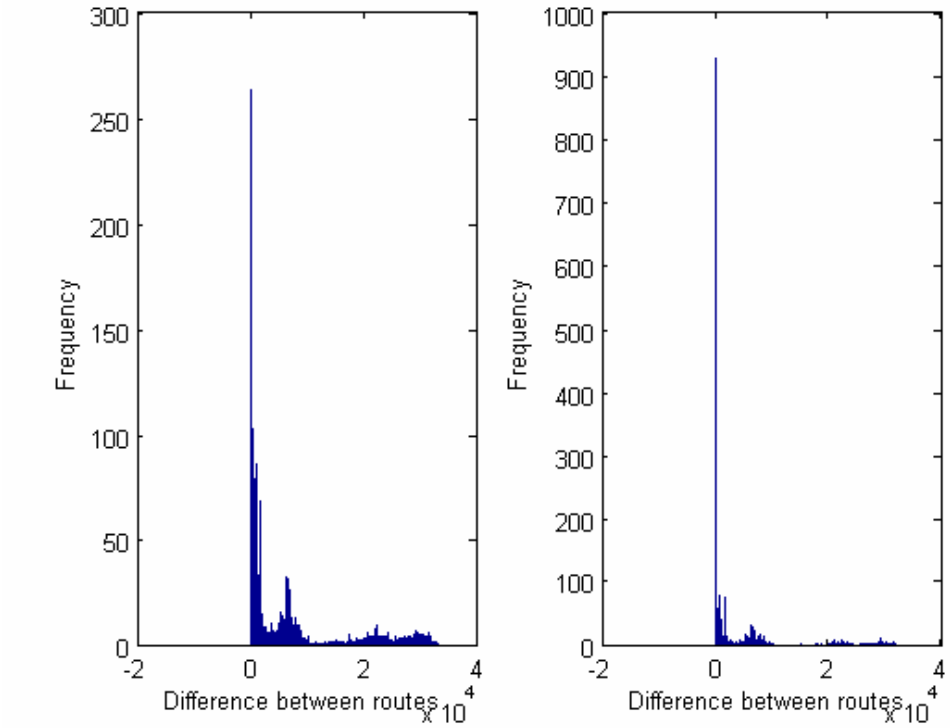

Figure A.4 Histogram of difference between (a) updates and (b) withdrawals AS 701

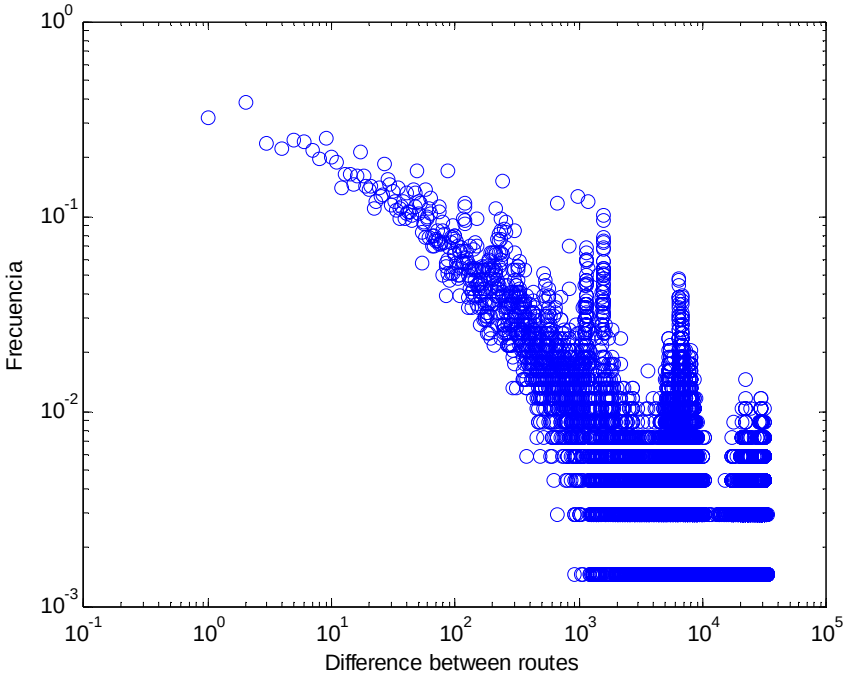

Figure A.5 Loglog plot difference between Updates AS 701

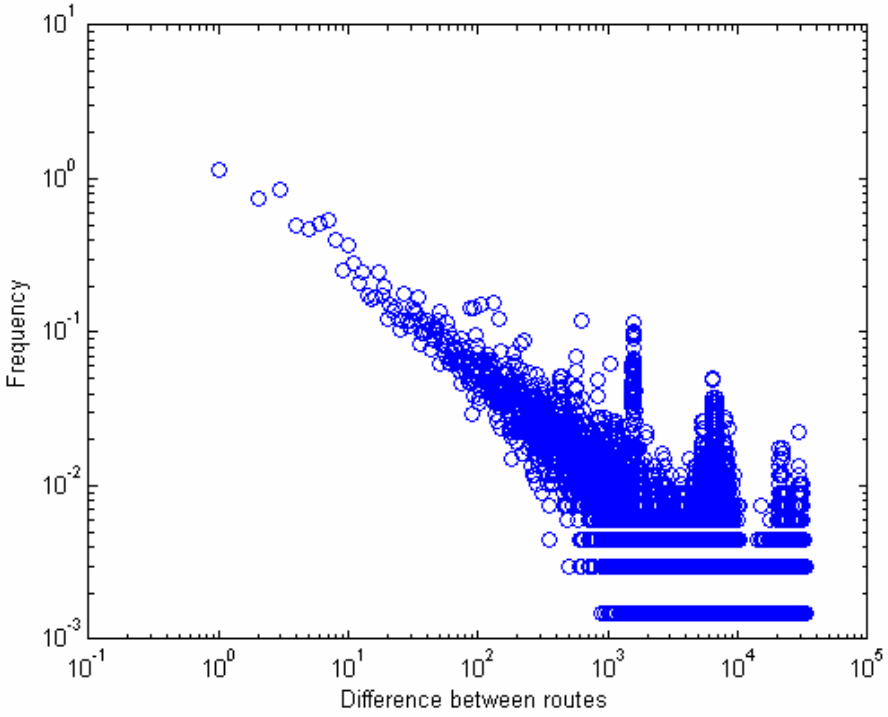

Figure A.6 Loglog plot difference between Withdrawals AS 701

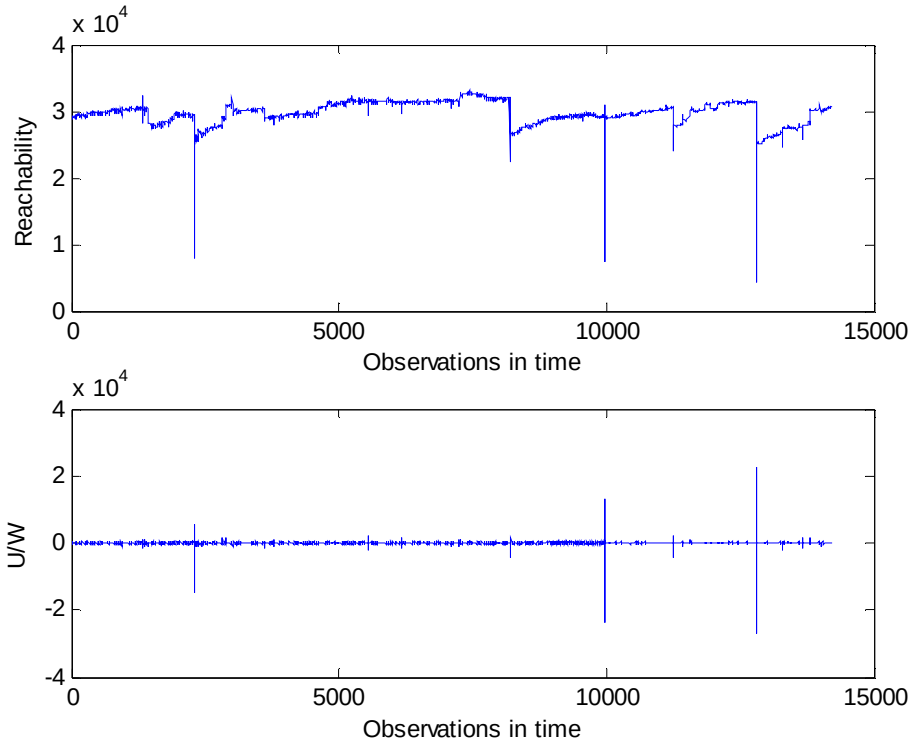

Figure A.7 (a) Reachability (b) Updates-Withdrawals from link 701-1239

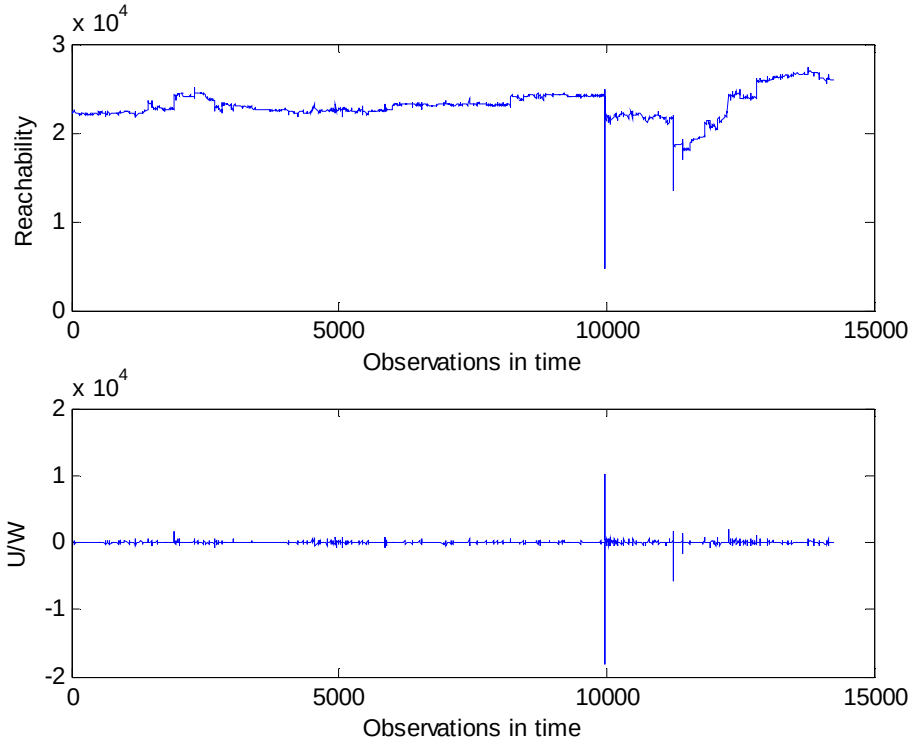

Figure A.8 (a) Reachability (b) Updates-Withdrawals from link 701-3356

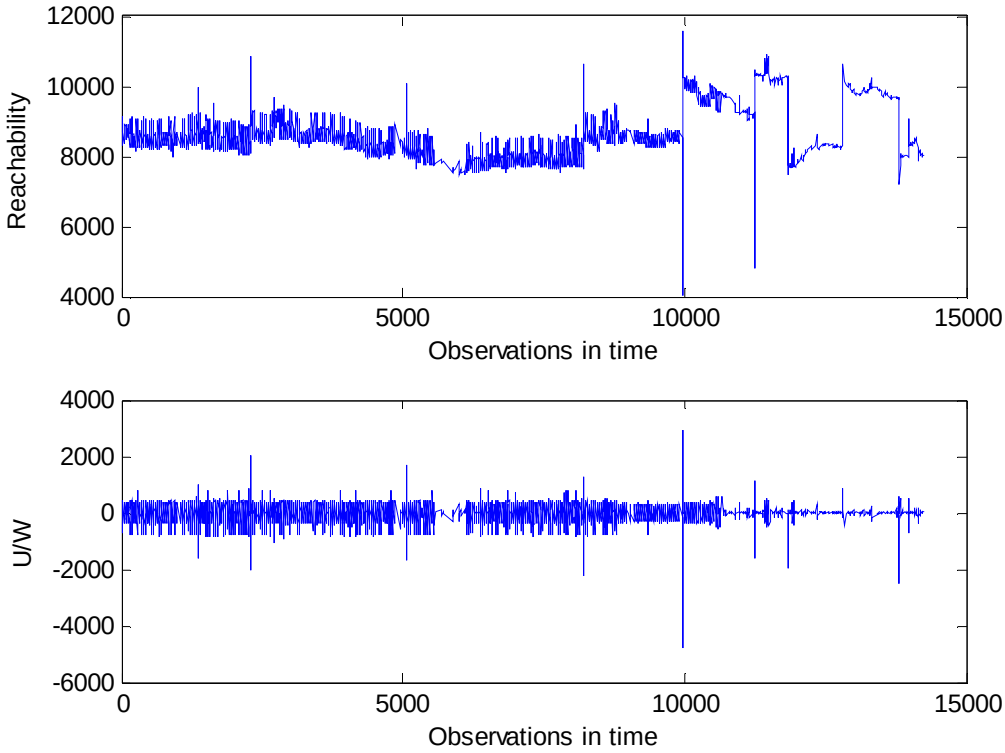

Figure A.9 (a) Reachability (b) Updates-Withdrawals from link 701-3561

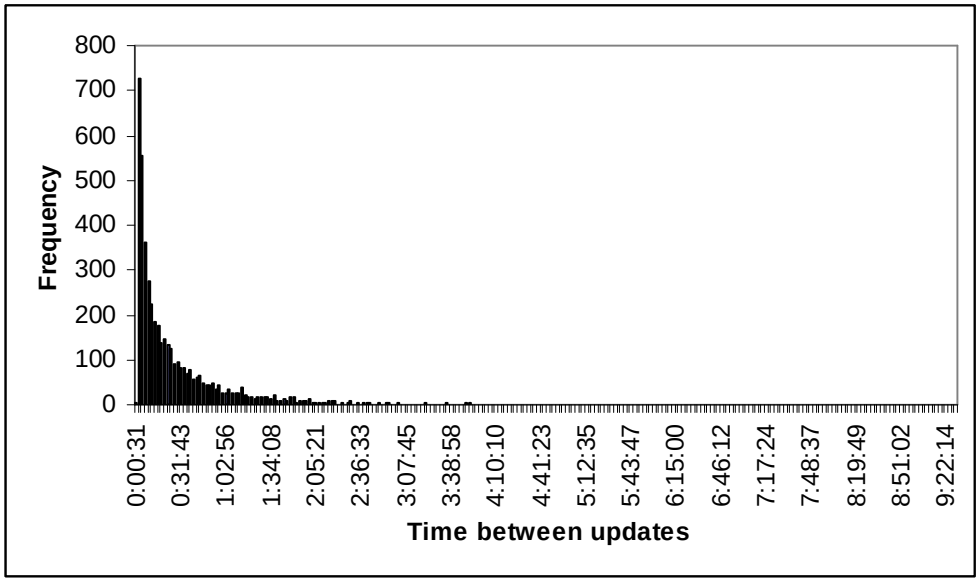

Figure A.10 Histogram time between Updates 701-1239

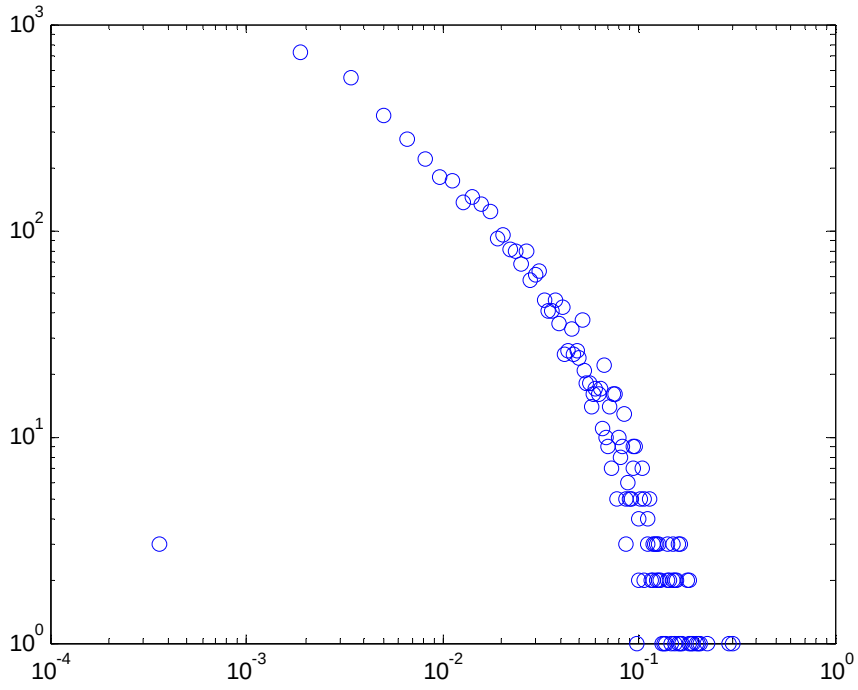

Figure A.11 Loglog plot time between Updates 701-1239

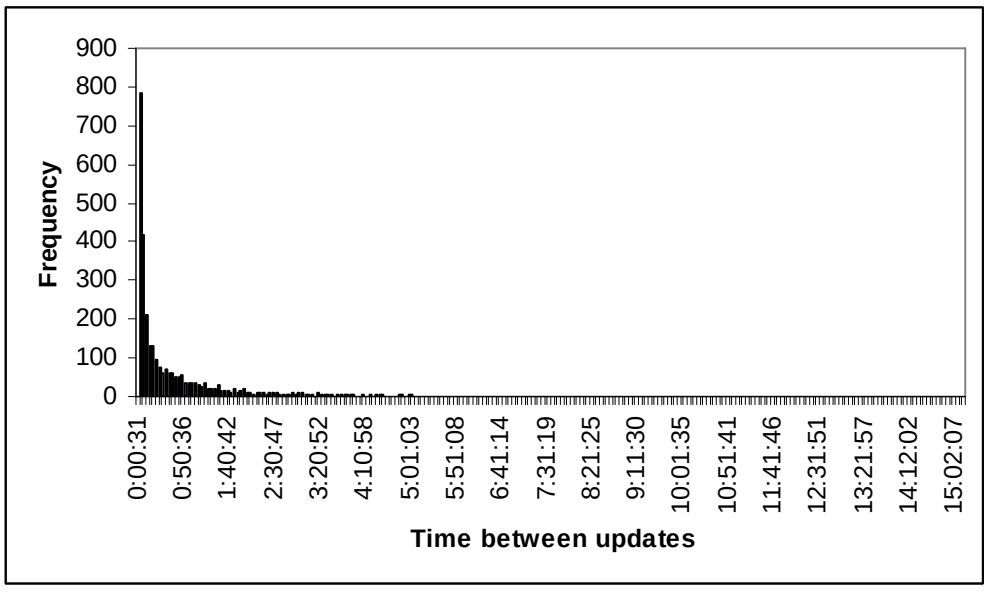

Figure A.12 Histogram time between Updates 701-3356

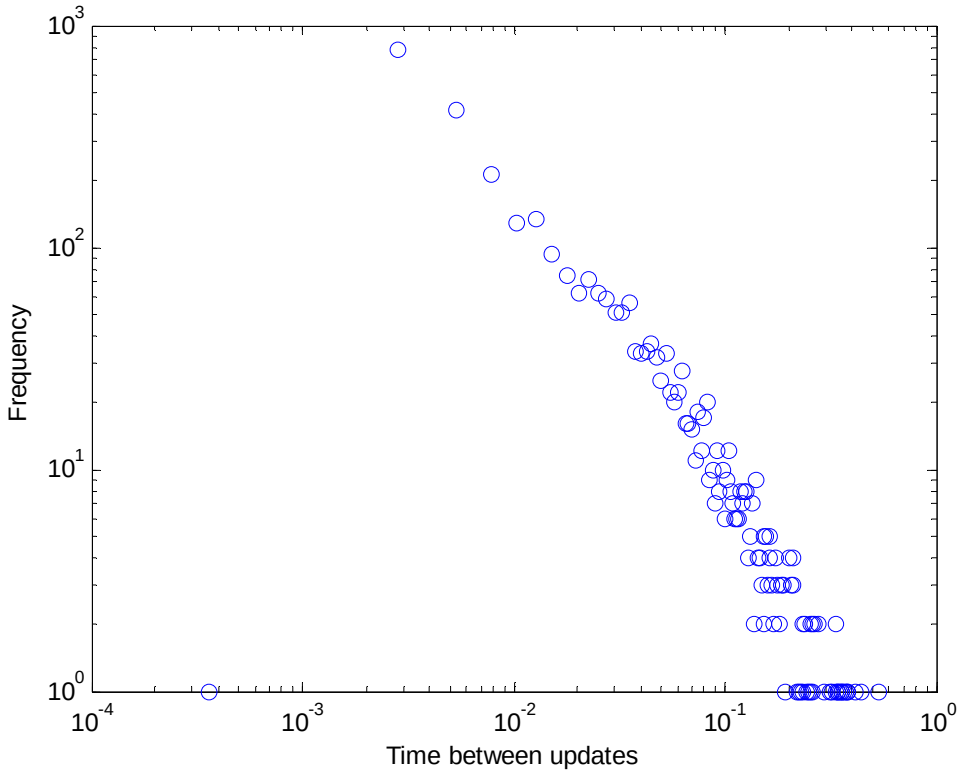

Figure A.13 Loglog plot time between Updates 701-3356

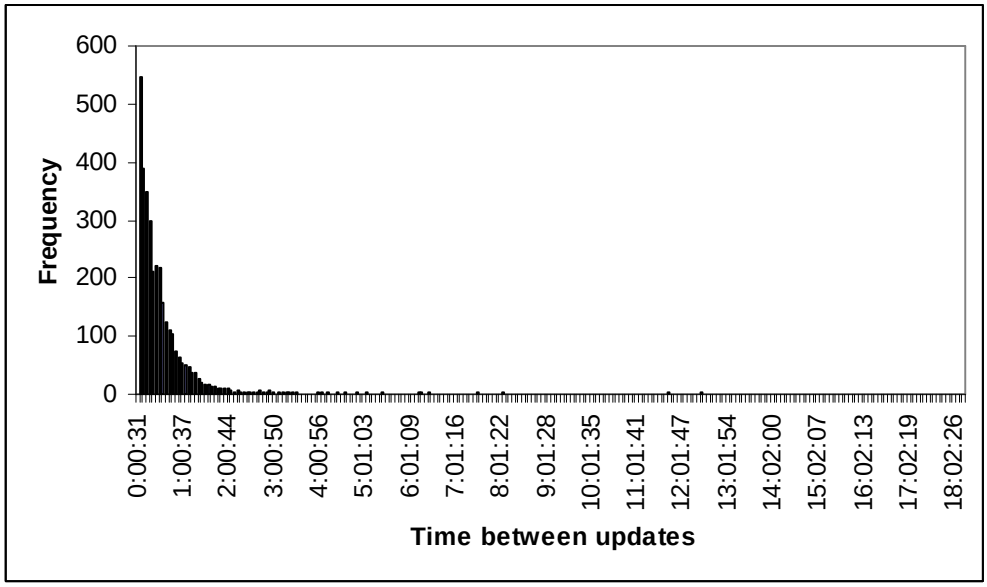

Figure A.14 Histogram time between Updates 701-3561

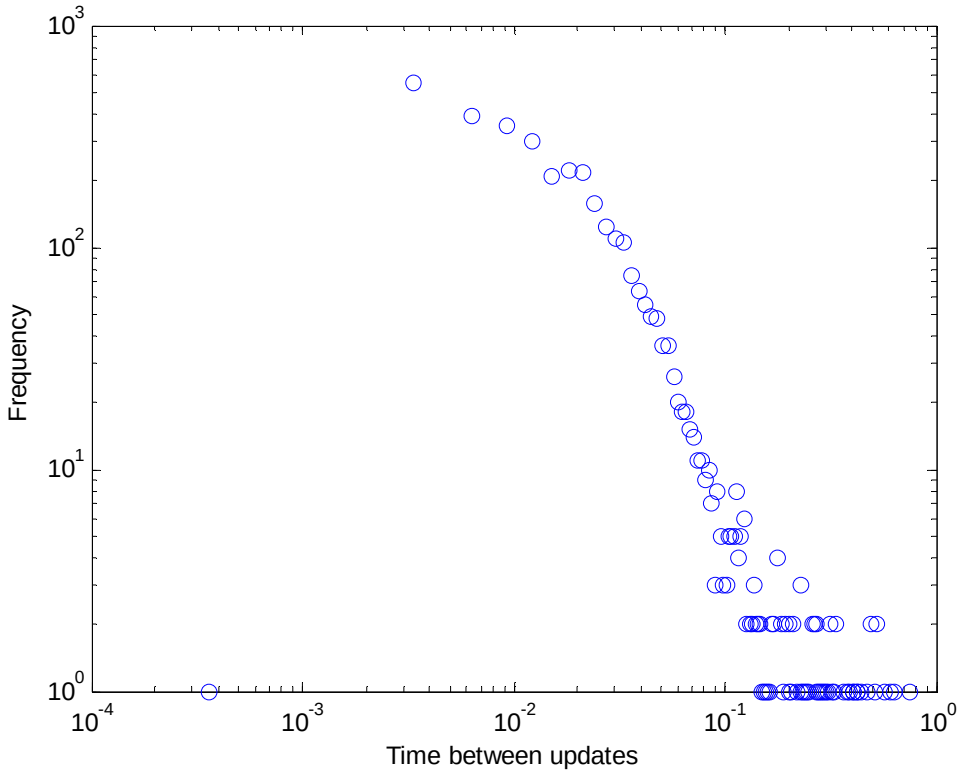

Figure A.15 Loglog plot time between Updates 701-3561

AS 1239

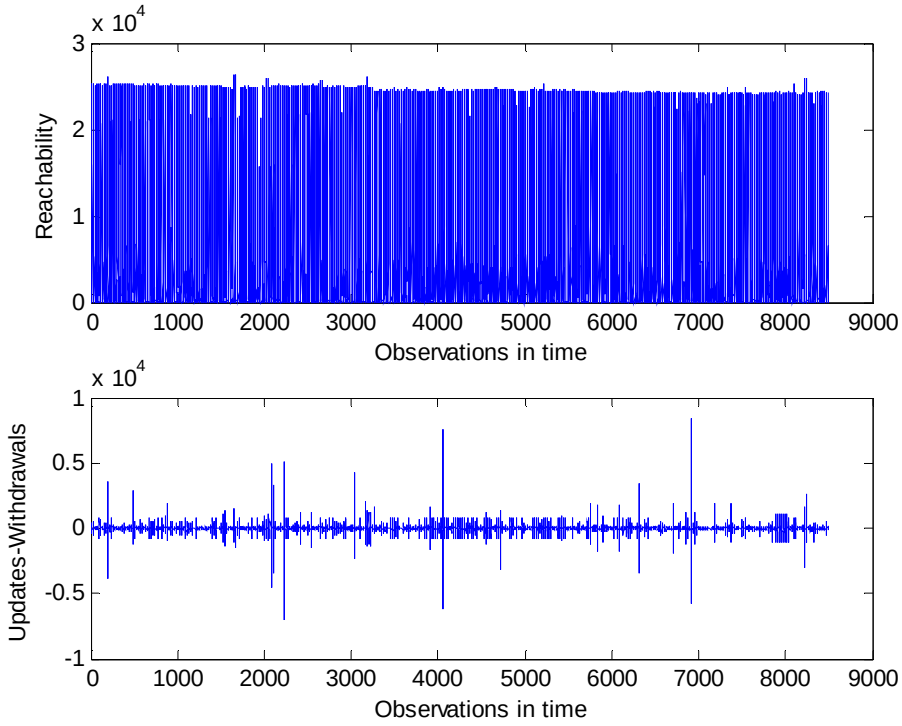

Figure A.16 (a) Reachability (b) Updates-Withdrawals from AS 1239

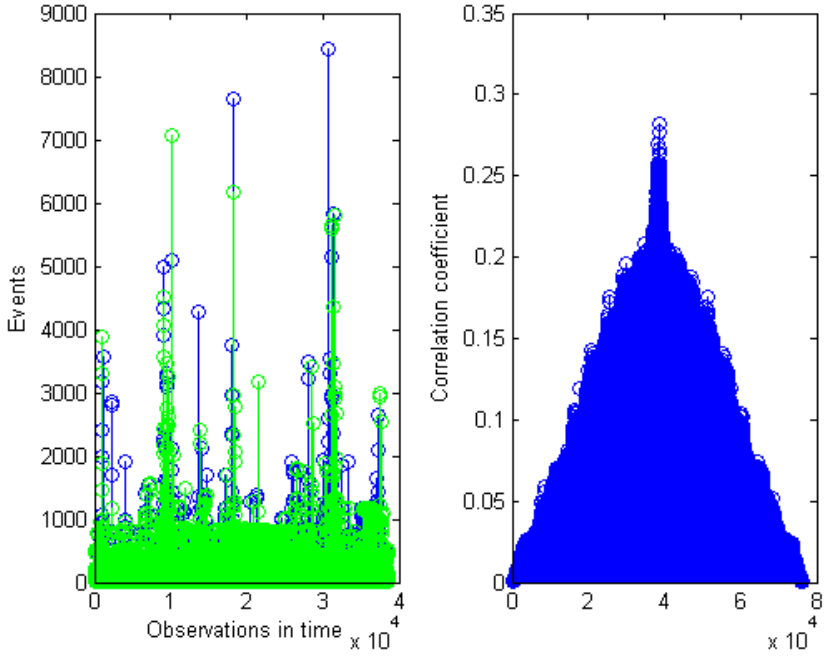

Figure A.17 (a) Updates and Withdrawals (b) Correlation AS 1239

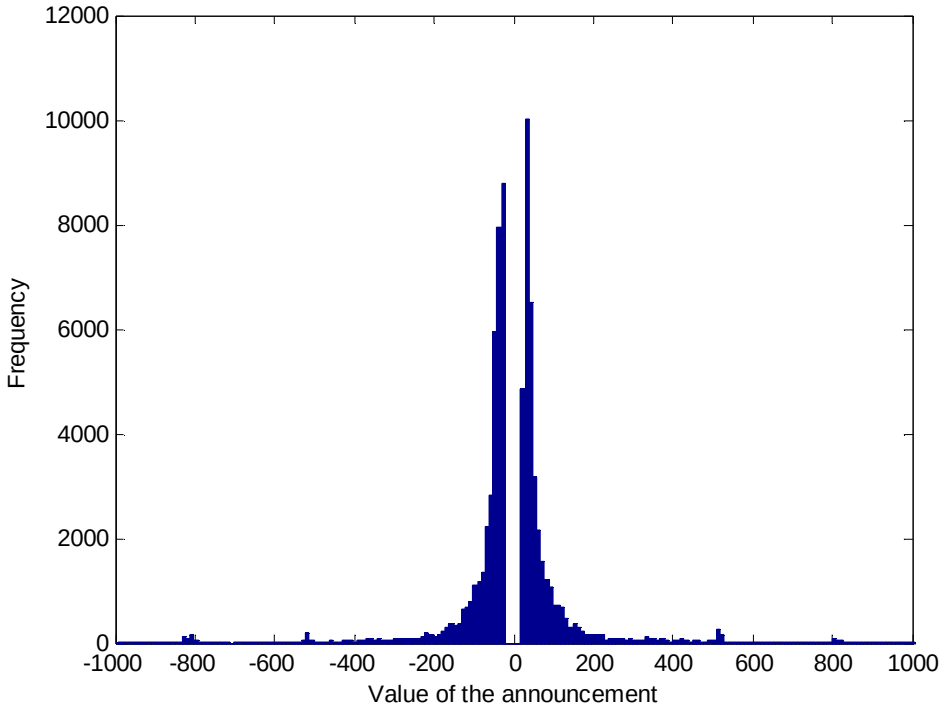

Figure A.18 Histogram of Updates and Withdrawals AS 701

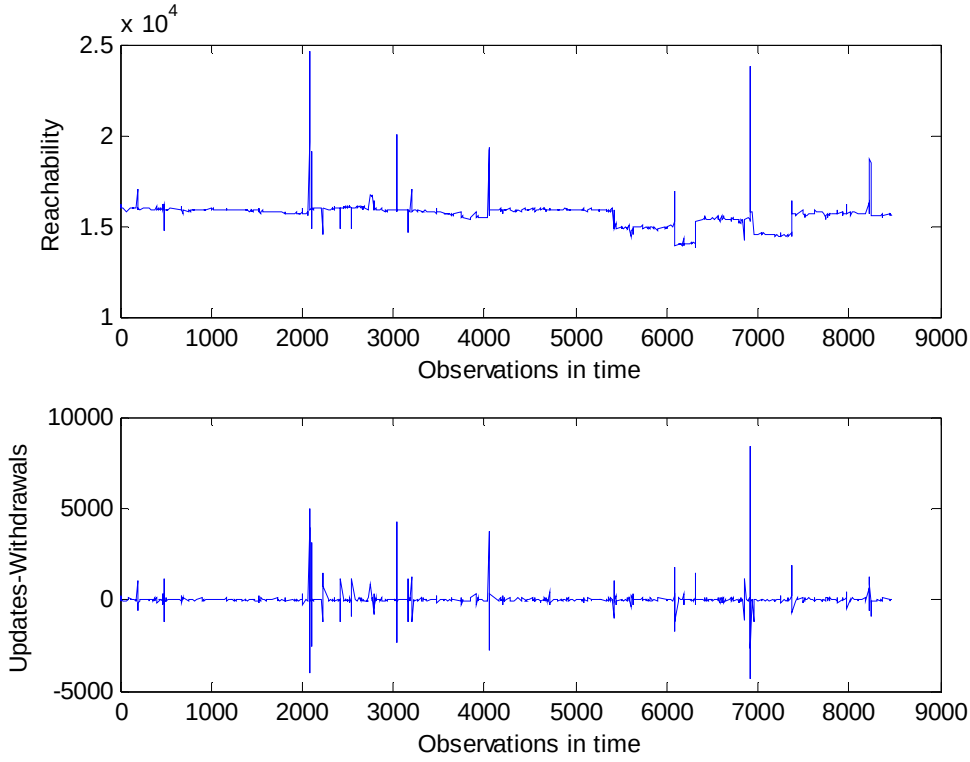

Figure A.19 (a) Reachability (b) Updates-Withdrawals from link 1239-7018

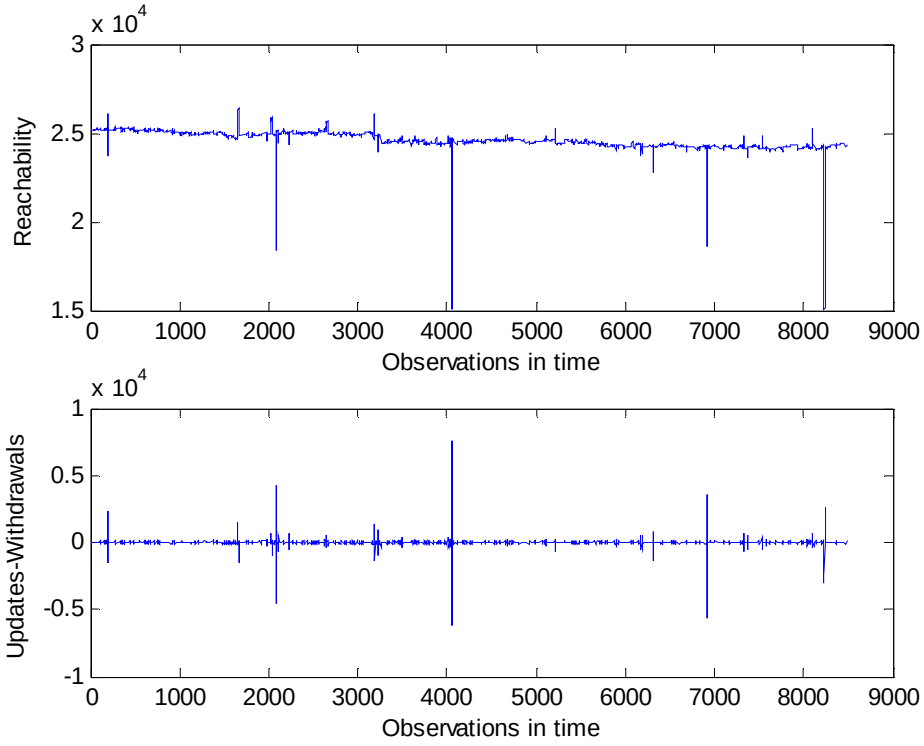

Figure A.20 (a) Reachability (b) Updates-Withdrawals from link 1239-701

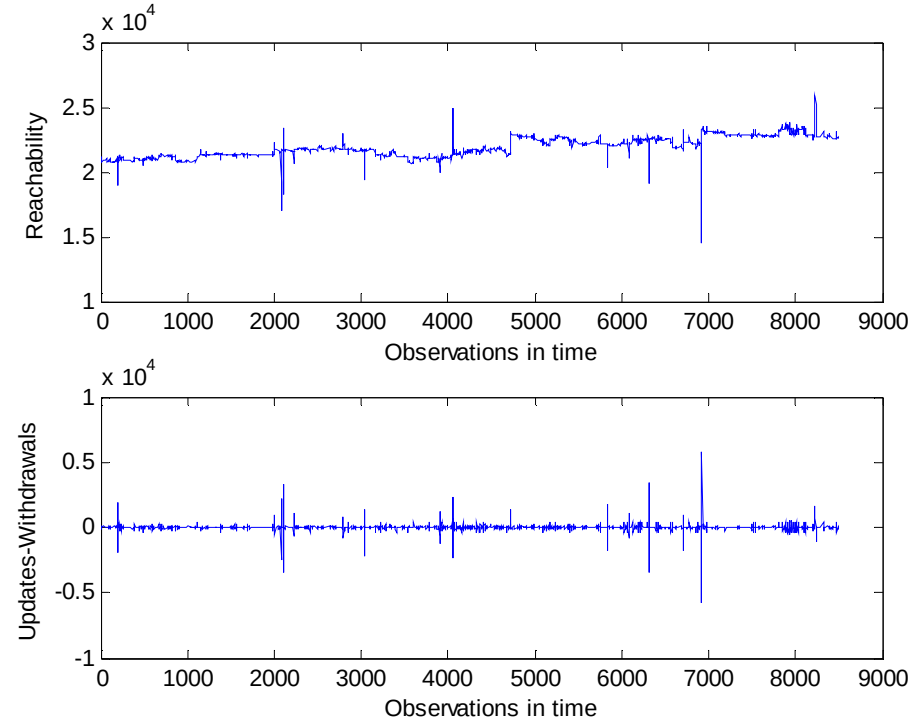

Figure A.21 (a) Reachability (b) Updates-Withdrawals from link 1239-3356

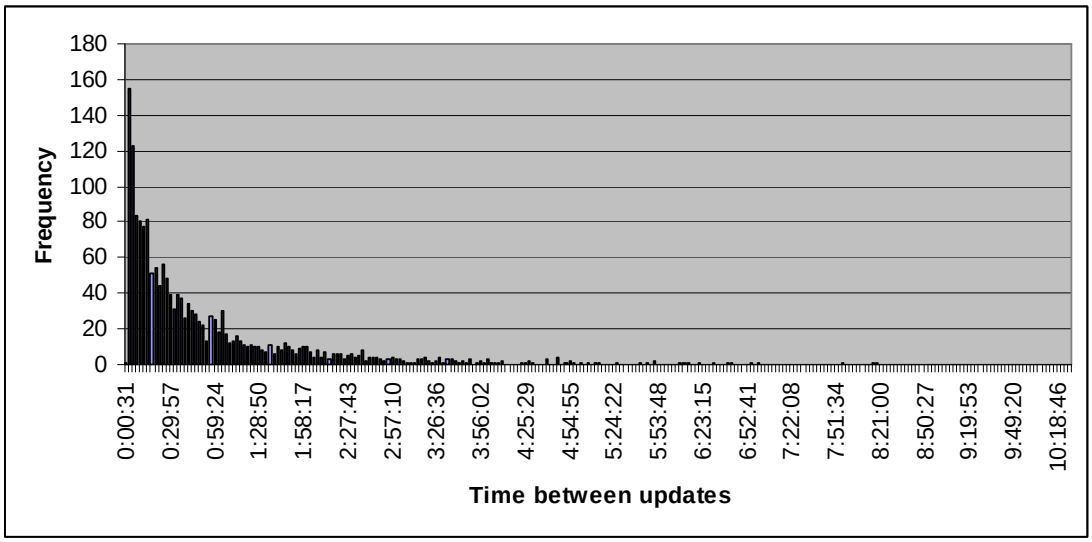

Figure A.22 Histogram time between Updates 1239-701

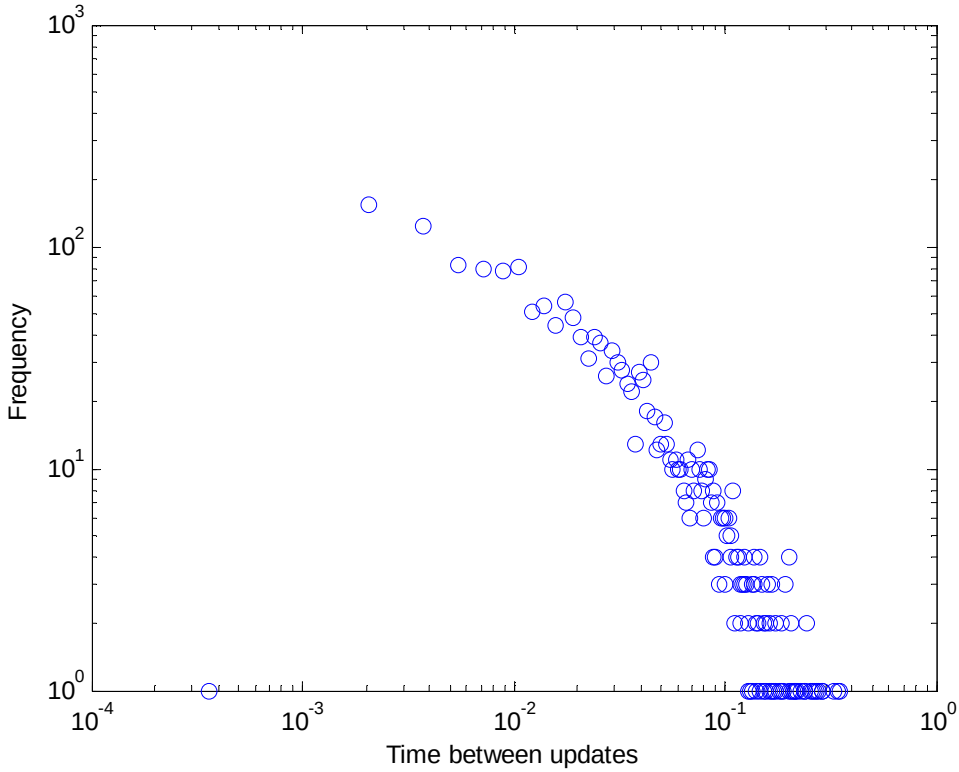

Figure A.23 Loglog plot time between Updates 1239-701

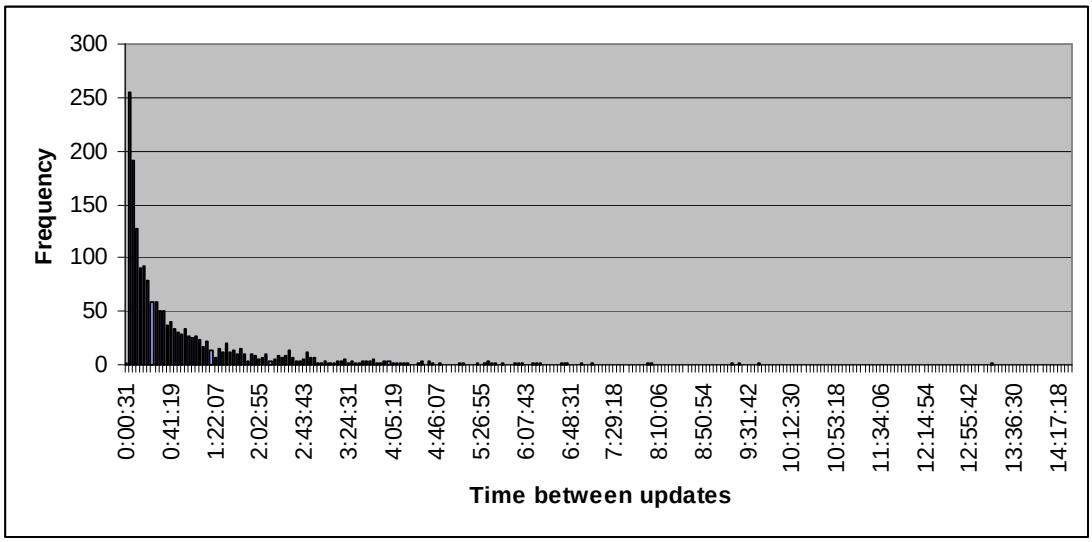

Figure A.24 Histogram time between Updates 1239-3356

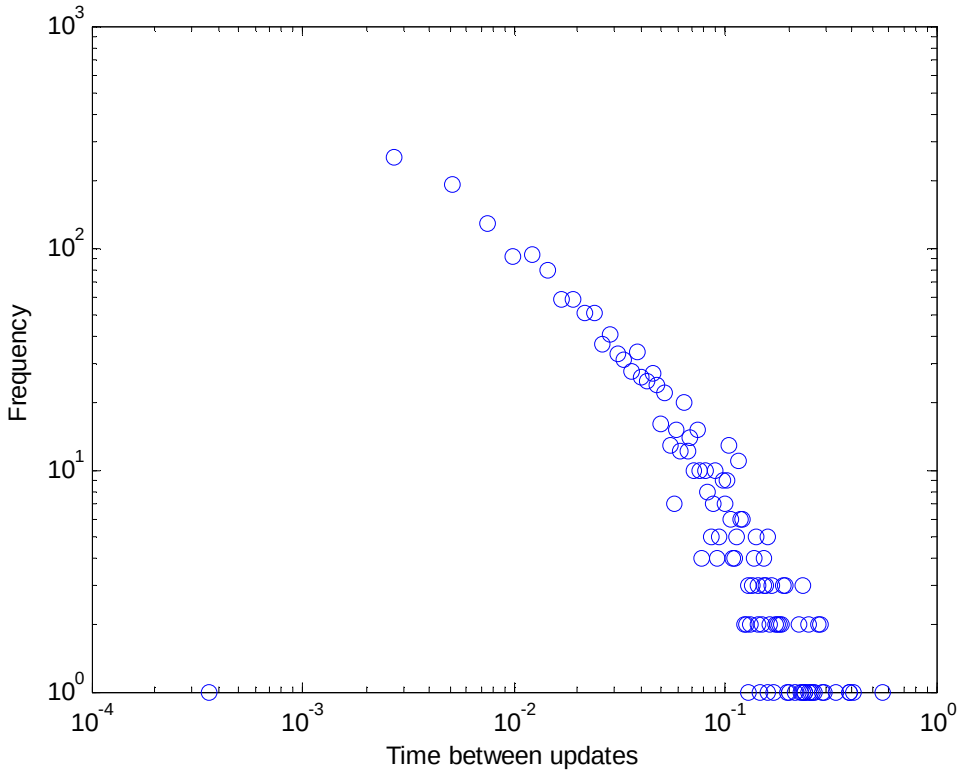

Figure A.25 Loglog plot time between Updates 1239-3356

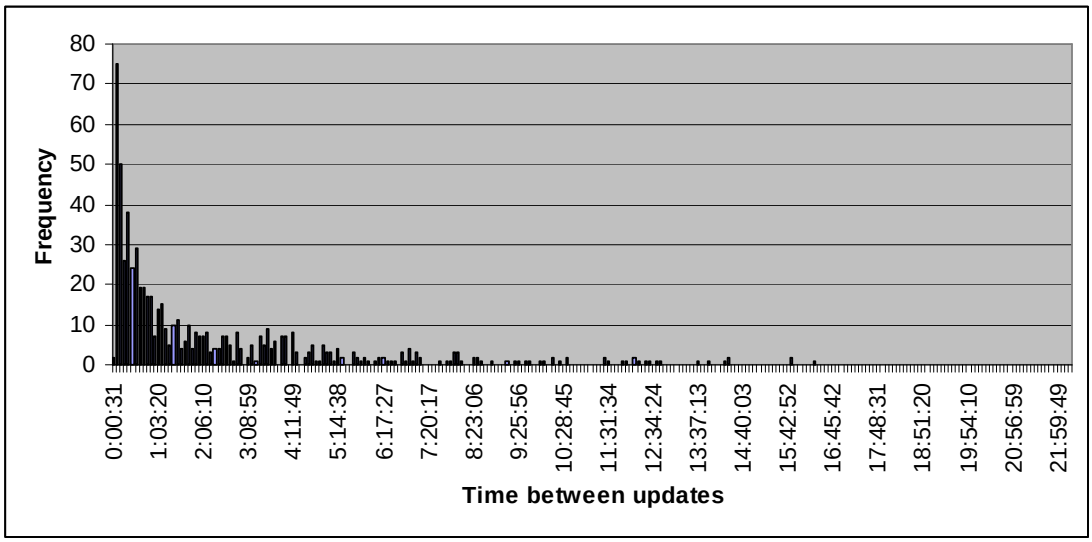

Figure A.26 Histogram time between Updates 1239-7018

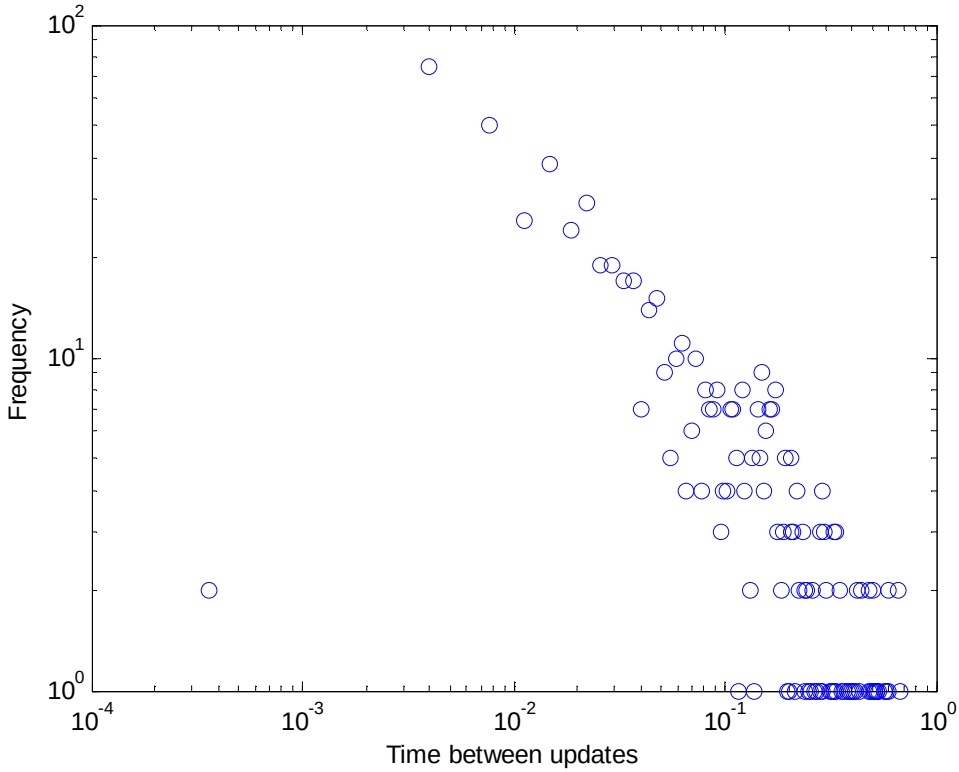

Figure A.27 Loglog plot time between Updates 1239-7018

Bibliography

[1] ISC. Internet Domain Survey, June 2005. http://www.isc.org/index.pl?/ops/ds

[2] G. Huston. The CIDR Report, January 2006. http://www.cidr-report.org/

[3] C. Filsfils and J. Evans. Deploying Diffserv in Backbone Networksfor Tight SLA Control. IEEE Internet Computing, January-February 2004

[4] D. O. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao. Overview and Principles of Internet Traffic Engineering. Internet Engineering Task Force, RFC3272, May 2002

[5] G. Huston. The Politics and Economics of Peering and Interconnection. In Proceedings of the Internet Society INET Conference. Available from http://www.isoc.org/inet99/proceedings/1e/1e_1.htm, June 1999

[6] L. Gao. On Inferring Autonomous System Relationships in the Internet. IEEE Global Internet, November 2000

[7] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz. Characterizing the Internet Hierarchy fromMultiple Vantage Points. In Proceedings of INFOCOM, June 2002

[8] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed Internet Routing Convergence. In Proc. ACM SIGCOMM, 2000

[9] C. Labovitz, A. Ahuja, and F. Jahanian. Experimental Study of Internet Stability and Wide-Area Network Failures. In Proceedings of FTCS, 1999

[10] C. Labovitz, R. Malan, and F. Jahanian. Internet Routing Stability. In Proc. ACMSIGCOMM, 1997

[11] C. Labovitz, R. Malan, and F. Jahanian. Origins of Internet Routing Instability. In Proc. IEEE INFOCOM, 1999

[12] C. Labovitz, R. Wattenhofer, S. Venkatachary, and A. Ahuja. The Impact of Internet Policy and Topology on Delayed Routing Convergence. In Proceedings of INFOCOM 2001, 2001

[13] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson. Inferring Link Weights using End-to-End Measurements. In Proc. Internet Measurement Workshop, November 2002

[14] Neil Spring, Ratul Mahajan, and David Wetherall. Measuring ISP topologies with Rocket-fuel. In Proc. ACM SIGCOMM, August 2002

[15] Tim Griffin. What is the sound of one route flapping? Network Modeling and Simulation Summer Workshop at Dartmouth, July 2002

[16] B. Halabi and D. Mc Pherson. Internet Routing Architectures (2nd Edition). Cisco Press, January 2000

[17] Y. Rekhter and T. Li. A Border Gateway Protocol 4 (BGP-4). Internet2 draft, draft-ietf-idr-bgp4-26.txt, work in progress, October 2004

[18] J. Stewart. BGP4 : interdomain routing in the Internet. Addison Wesley, 1999

[19] R. Teixeira, A. Shaikh, T. Griffin, and J. Rexford. Dynamics of Hot-Potato Routing in IP Networks. In Proceedings of ACM SIGMETRICS, June 2004

[20] E. Chen and S. R. Sangli. Avoid BGP Best Path Transition from One External to Another. Internet draft, draft-chen-bgp-avoid-transition-04.txt, work in progress, December 2005

[21] Radia Perlman. Bridges, routers, switches and interworking protocols

[22] http://www.caida.org/home/

[23] Hyunseok Chang, Ramesh Govindan, Sugih Jamin, Scott J. Shenker. Towards capturing representative AS-level Internet topologies. Computer Networks, Volume 44, Issue 6, 22 April 2004, Science Direct

[24] Gomez Zamorano, J.R. On power law relationships of inter-domainrouting changes". In Instituto Tecnologico de Estudios Superiores Monterrey, 2006

[25] L. Wang, X Zhao, D. Pei, R.Bush, D. Massey, A. Mankin, S. Wu, and L. Zhang. Observation and analysis of BGP behavior under strees. In Proceedings of Internet Measurament Workshop, November 2002

[26] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP misconfigurations. In Proceedings of ACM SIGCOMM, August 2002

[27] Z. Wu, E. S. Purpus, and J. Li. BGP behaviour analysis during the August 2003 blackout. In International Symposium on Integrated Network Management, 2005

[28] Medina, I. Matta and J. Byers, “On the Origin of Power Laws in Internet Topologies”, *ACM SIGCOMM Computer Communication Review*, Vol. 30, No. 2, April 2000

[29] M. Faloutsos, P. Faloutsos and C. Faloutsos, “On Power-Law Relationships of the Internet Topologies”, ACM SIGCOMM Computer Communication Review, Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication SIGCOMM ’99, Vol. 29, No. 4, August 1999

[30] G. Siganos, M. Faloutsos, P. Faloutsos and C. Faloutsos, “Power Laws and the AS-level Internet Topology”, IEEE/ACM Transactions on Networking, Vol. 11, No. 4, August 2003

[31] V. Pareto, Cours d'économie politique. Reprinted as a volume of Oeuvres Complètes (Droz, Geneva, 1896-1965)

[32] http://linkrank.cs.ucla.edu/data/rv/