INSTITUTO TECNOLOGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS MONTERREY

PROGRAMA DE GRADUADOS EN TECNOLOGIAS DE INFORMACIÓN Y ELECTRONICA

TECNOLÓGICO DE MONTERREY®

On Power Law Reachability Analysis at an Autonomous System Granularity

THESIS

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

Roberto Enrique Magaña Rodríguez

Monterrey, N.L., July 2007

Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus 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 Roberto Enrique Magaña Rodríguez 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

_____
Gerardo Castañon Ávila, Ph.D.
Synodal

_____
Artemio Aguilar Coutiño, M.C.
Synodal

_____
Graciano Dieck Assad, Ph.D.
Director of the Graduate Program
July, 2007

Dedicatory

To my parents, because they taught me all I know and show me that there are not unreachable goals. They are my inspiration, my support and most of all, the best example in my life.

Thank you for everything, I will love you forever.

Acknowledgments

I want to thank to my parents for all their love and support in my professional, personal and spiritual goals.

I want also to thank to my girlfriend, Luz Helena Salgado Locela, for all her love, support and understanding in my search of new opportunities in personal and professional development. I love you precious.

I can never forget to thank the Salgado Locela family for their affection and support.

I always will be thanked with my thesis advisor, Cesar Vargas Rosales, Ph.D. for his guidance and teaching through my path in the ITESM. Thank you very much Doctor Vargas.

I want to thank my synodals, Artemio Aguilar Coutiño, M.C. and Gerardo Castañon Ávila, Ph.D. for their comments and recommendations to make of this work a better thesis.

Finally but not last, to my friends Jorge A. León Castelán, Alfonso Sánchez De Lucio, José Carrillo Valdés, Francisco Martínez Baltodano, Carlos Barrera Suárez, Luis Peraza Rodríguez, and Zaira Pineda Rico, for their support and friendship. You are my better finding during my stay in Monterrey.

Roberto Enrique Magaña Rodríguez

Instituto Tecnológico de Estudios Superiores de Monterrey
July, 2007.

Abstract

The internet consists of rapidly increasing number of Autonomous Systems(AS) interconnected. Interdomain routing in Internet is coordinated by the Border Gateway Protocol. Which routing table has been growing dramatically. This situation lead us to search a new routing scheme able to support the new requirements. Since each AS administers its reachability information via routing policies, the new routing scheme must consider the reachability of the nodes into the network. These routing policies are constrained by the contractual commercial agreements between ASes. This implies that the relationships among Autonomous System granularity (i.e. Customer-Provider) are an important aspect of Internet structure and affects the reachability of the nodes in the network topology. In this work, we explore the relationship between the times between topology changes in the AS-Level and the reachability over a network based in a customer-provider relationship to characterize the reachability process. We begin with a network generator based in the nodes at 1 hop distance, followed by a routing algorithm to find the shortest path to all available destinations from a particular origin. Afterwards, we run a simulation modifying the reachability state of the nodes over time. We shown that the decrease of reachable Autonomous Systems between topology changes in the nodes obeys a power law distribution that is consequence of the power law observed in the times between topology changes at the interdomain level.

Contents

List of Figures

Chapter 1

Introduction

The social interactions through networking were conceived in a series of memos written August 1962 by J.C.R. Licklider [22], who envisioned a globally interconnected computer through which everyone could quickly access data and programs from any site. DARPA (Defense Advanced Research Projects Agency) was the first one embracing this concept and giving to the networking concept the importance needed to convert this dream in a reality.

L. Kleinrock published his first paper on packet switching theory in 1961, [19]. Leonard was the pioneer in the paradigm change of communications using packets rather than circuits. In the same year Lawrence G. Roberts connected the TX-2 computer in Massachusetts to the Q-32 in California through a low-speed dial-up telephone line [25], creating the first-ever (though small) wide-area computer network. This experiment leads us to the idea that time-sharing computers could work well together, running programs and retrieving data as necessary on remote machines, however, the circuit switched telephone system was inadequate for the job confirming the packet switching theory proposed by Kleinrock.

It was until 1966 when Roberts went to DARPA to develop the network idea creating the ARPANET (Advanced Research Projects Agency Network), published it in 1967, [24]. Taking the Kleinrock switching concept the first switch was installed by Bolt, Beranek and Newman Corp. (BBN) at UCLA (University of California, Los Angeles) and the first host computer was connected. In December 1970, the Network Working Group finished the initial ARPANET host-to-host protocol, called the Network Control Protocol (NCP) which means that network users finally could begin to develop applications.

In late 1972, Robert E. Khan introduced in DARPA the concept of open-architecture networking. Vinton Cerf had been deeply involved in the original NCP

design and development. So, using the Kahn's architectural concept to communications and Cerf's NCP experience the birth of the Transmission Control Protocol/Internet Protocol (TCP/IP) era was beginning. TCP/IP was adopted as a defense standard in 1980.

The Internet was established as a technology supporting a broad community of researchers and developers in 1985. The growth of the Internet resulted in vastly increased attendance at IETF (Internet Engineering Task Force) meetings and as a result the working groups were created. The expanded community also meant that DARPA was no longer the only major player when it came to funding the Internet. In 1985, Dennis Jennings came from Ireland for a year to lead the National Science Foundation's Network (NSFNET) program, this network was an alternative to DARPA. He established that TCP/IP would be mandatory for NSFNET. The need for a wide-area networking infrastructure to support the general academic and research community, as well as the need to develop a strategy for establishing such infrastructure to ultimately be independent of direct federal funding. In addition to NSFNET and the various U.S. and international government-funded activities, interest in the commercial sector was beginning to grow creating an increased concern regarding the standards process. The motivations of making the process open and fair, and the need to win Internet community support eventually led in 1991 to formation of the Internet Society.

Widespread development of LANs (Local Area Networks), PCs and workstations in the 1980s allowed the nascent Internet to flourish. Ethernet technology, developed by Bob Metcalfe at Xerox PARC in 1973, is now probably the dominant network technology in the Internet and PCs and workstations the dominant computers. The result was the definition of three network classes (A,B, and C) to accommodate the range of networks. The Internet growing challenged the capabilities of the routers. As the number of networks exploded, the initial design of a single distributed routing algorithm was replaced by a hierarchical model of routing, with an Interior Gateway Protocol used inside the network and an Exterior Gateway Protocol used between networks. Hence, not only the routing algorithm, but the size of the addressing tables, stressed the capacity of the routers originating the creation of the Classless Interdomain Routing to control the size of router tables.

In spite of, its history is full of difficulties we need to recognize that Internet is the largest distributed network operating these days and also that the Internet Protocol is the standard in the internetworking technology, and its growing has push up to the research community to continue looking for new solutions to make it scalable and ready to face the future generations.

1.1 Problem Definition

The routers contains reachability information of each prefix using the Border Gateway Protocol (BGP), [23]. As a consequence, each BGP routing table entry contains reachability information for a single prefix. The number of prefixes contained in the routing tables has grown in the last years, [17],[18]. This phenomena affects the packet forwarding speed and demand more router memory space.

A partial solution to this problem was the introduction of the Classless Interdomain Routing (CIDR), [12], which reduces the routing table size by enabling more aggressive route aggregation in which a single prefix is used to announce the routes to multiple prefixes. However, when a customer has multiple providers, route aggregation might not be performed, [5].

This shows us the importance of the relationships between Autonomous Systems in an internet routing scheme. The design of a new routing scheme for the internet is not an easy task, since the routing function must find the shortest path to send the packets. In [14], it is defined the shortest path routing scheme as the selection, for any network, of a shortest path routing whose implementation minimizes the number of hops per link. This definition applies for all kind of networks; however, the routing scheme that returns a shortest path routing function that can be implemented using a unique interval per link, is only partial in the sense that there exist networks on which this routing scheme is not defined, [10]. At this point, it is mandatory the design of a new interdomain routing scheme that performs a shortest path with minimal memory requirements in the routing table for scalability issues.

On one hand, The BGP protocol designers did not preview the critical growing in the routing table size and on the other hand, the open environment characteristics of the Internet do not allow us to obtain specific scalability measurements. Therefore, we have to use alternative measurements to approximate the internet scalability measurements in order to create a new generation of Internet routing protocols.

The measurement of network topology and routing information variables are very important. The characterization of the node reachability based in the topology changes and the routing information can help us to gain future insight of the network behavior. A reachability measurement of the nodes into the topology changes at an Autonomous System granularity and how it affects the stretch rate is necessary to create models able to predict the internet behavior and measure the scaling capabilities of a new interdomain routing scheme.

This work proposes a measurement of the node reachability at an Autonomous System granularity into the topology changes and an stretch rate analysis. The analysis is based on a topology generated through the Circle of Trust nodes, a routing function to find the shortest path from origin to destination into this network topology and a simulation in time to observe the topology changes.

The network topology generator will be a function of the nodes Circles of Trust that represents a customer−provider relationship at an Autonomous System level at distance one, [13]. The routing function will use only the Circle of Trust of each node to find the shortest path from origin to destination. We will use the network generator and the routing function to create a simulation in time that allow us to obtained information from the topology changes and node reachability.

The data is obtained through the observation of the topology in the simulation time at every power law time, since the time between topology changes follows a power law distribution, [16]. The reachability node is measured using this information to obtained an statistical model that allow us to find a relationship with the times between topology changes power law distribution, besides the stretch rate is calculated to observe how the node reachability affects it.

1.2 Hypothesis

The statistical behavior in the reachability process of an Autonomous Systems between topology changes of a network based in the customer−provider relationship depends on the statistical description of the laying times between topology changes. Thus, statistical characterization provides the frequency at which is given the Autonomous System reachability regarding to the times between topology changes into the network and its explanation and significance in the routing scheme design.

1.3 Objectives

We divide the main objectives of this thesis as follows:

- Obtain a network based on the interdomain customer-provider relationship between Autonomous Systems considering only the nodes at 1 hop of distance.

- Obtain a routing algorithm able to reach all available destinations in the network, based only on the nodes at 1 hop of distance of each member of the topology.

- Obtain a statistical behavior model of the decrease in the number of reachable Autonomous Systems between topology observations from a network based on an interdomain customer−provider relationship topology.

- Find a relationship among statistical times between topology changes model of the interdomain architecture and the statistical behavior model of the decrease in the number of reachable Autonomous Systems between topology changes over a network based on the customer−provider relationship.

- Observe the implications on scalability issues analysis and interdomain routing scheme proposal related to the behavior model of the decrease in the number of reachable Autonomous Systems over topology changes, based on a network topology that considers the interdomain customer−provider relationship.

1.4 Final Product

A power law behavior statistical reachability process regarding topology changes over a network based in an interdomain customer−provider relationship topology generated through the knowledge of the 1 hop distance nodes. This power law it is related to the laying topology as the ones shown in [16],[26],[8], and consequently it is also related to those power law relationships.

1.5 Contributions

This work uses tools already available. However, there are several contributions derived from them. A list follows:

- The concept of Circle of Trust to define the nodes at 1 hop distance with an interdomain customer provider relationship. Generating a topology based only on the Circle of Trust nodes of a provider.

- A routing scheme to calculate the shortest path as a function of the node reachability and the nodes at 1 hop distance of every topology member.

- A topology generator based on the internet customer-provider relationship.

- A routing scheme analysis based on a directed graph that represents the links among customers and providers to forward packets.

1.6 Summary

The chapter 2 includes the necessary background to understand the work proposed. It involves the routing basics, the kinds of routing protocols, the interdomain issues, the Border Gateway Protocol and the routing schemes in distributed networks.

The chapter 3 is related to the simulation, considering the network generator, the circle of trust definition, and the internet power law time between topology changes. Besides, it states the backbone considerations, and the topology changes.

The chapter 4 contains the conclusions of this work, in particular the possible design of a routing scheme based in the nodes at 1 hop distance that considers a power law reachability process related with the power law times between topology changes.

Chapter 2

Background

In this chapter we are going to develop the principles and concepts that we consider necessary to obtain a complete understanding of this work.

2.1 Routing Basics

At some point in the near future the Internet will require a deployed new version of the Internet protocol. Two factors are driving this: routing and addressing. Global Internet routing based on the 32-bit addresses of IPv4 is becoming increasingly strained. IPv4 addresses do not provide enough flexibility to construct efficient hierarchies that can be aggregated. The deployment of Classless Inter-Domain Routing is extending the lifetime of IPv4 routing by a number of years, but the effort expended to manage the routing will continue to increase. Even if the IPv4 routing can be scaled to support a full IPv4 Internet, the Internet will eventually run out of network numbers. Therefore, a new version of the Internet Protocol, designed as a successor to IP version 4 was developed and it is called IP version 6 (IPv6).

The changes from IPv4 to IPv6 fall primarily into the following categories:

- **Expanded Routing and Addressing Capabilities:** The IP address size increases from 32 bits to 128 bits, in order to support more levels of addressing hierarchy, a much greater number of addressable nodes, and a simpler autoconfiguration of addresses are needed.

- **Header Format Simplification:** Some IPv4 header fields have been dropped or made optional, as a compensation method (to reduce the processing cost of packet and to keep the bandwidth cost of the IPv6), due to that, the addresses

of IPv6 are four times longer that IPv4.

- **Improved Support for Options:** The way IP header options are encoded was changed for more efficient forwarding, less stringent limits on the length of options and greater flexibility (for future options).

- **Quality-of-Service Capabilities:** Enable the labelling of packets belonging to particular traffic "flows" for which, the sender request special handling (such as non-default quality of service or "real-time"service).

- **Authentication and Privacy Capabilities:** IPv6 includes the definition of extensions that provide support for authentication, data integrity, and confidentiality.

In conclusion IPv6 supports large hierarchical addresses, which will allow the Internet to continue to grow and provide new routing capabilities not built into IPv4. It has anycast addresses that can be used for policy route selection and has scoped multicast addresses that provide improved scalability over IPv4 multicast. The address structure of IPv6 was also designed to support carrying the addresses of other Internet protocol suites. Space was allocated in the addressing plan for IPX and NSAP addresses (to facilitate migration of other protocols).

2.2 Routing Protocols

It is very easy to notice that in a network the need to interconnect nodes and share information are mandatory. A way to organize this communication lies on the Open Systems Interconnection Model, where each task related to a network operation is assigned to seven layers, from the Physical Layer to the Application Layer. Following this model, we can find that the organization of sending and receiving packets from an origin to a destination are responsibilities of the Network Layer.

In order to perform its job correctly, this layer needs the topology information of the network and as a consequence, the knowledge of the best paths from any node origin to all the destinations. Also, we have to consider that the Network Layer is responsible of the traffic control and the network intercommunication, even if the control protocols are different.

Since, in most subnets a packet will have to pass through several hops (jumps between nodes) before reaching its destination, it is necessary a routing algorithm to perform this task. Therefore, and since a network can work continuously for years without system wide failures, the robustness property of a routing algorithm is a very important issue. This means, that it should be able to deal with many topology changes (hardware and software failures) and selective traffic control. Today, we can identify two classes of routing algorithms Non-adaptive and Adaptive.

2.2.1 Non-Adaptive Algorithms

These algorithms use static routing. In other words, the calculation of the network paths, and the selection of a route from node i to j is computed in advance (offline) and stored in the routers when the network is booted. In this category lie the Shortest Path Routing, Flooding and Flow-Based Routing.

Shortest Path Routing

Consider that we have a graph representing a network, where each node represents a router and the arcs represents a communication line (link). The algorithm goal is to find the shortest path between two nodes using the available links in the graph. It seems to be a simple task; however, what is the shortest path? It depends on what we are measuring or what policy is the most important for the network administrator. If the metric is the number of hops, two paths totally different may appear to be equally long. However, one of them will be the shortest path depending on our metric interests. In this matter, we can use as a metric: the number of hops, the geographical distance, transmission delay, average traffic, etc. In general, the metric to differentiate the links in our network is a weighting function that considers the distance, bandwidth, communication cost and other factors.

The Dijkstra (1959) algorithm [6] is the perfect example for shortest path calculation. It works relating each node with its distance from the origin node to its destination through the best known path. Initially, the nodes are labeled with infinity because no paths are known, when the algorithm starts these labels are modified to reflect better paths (these paths can be permanent or tentative) until the shortest path is found from origin to destination. When the algorithm finishes the shortest path is static and never changes thereafter, unless a failure occurs.

Flooding

The best example for this kind of routing is the hot potato. This means that every incoming packet is duplicated and sent to every outgoing link available in a node (except the incoming link). As we can notice, an infinite number of packets will be flooding the network unless some measures are taken to control it. One of them is a counter to zero, where each packet has a counter initialized to the length of the path from origin to destination (or in the worst case, it is initialized with the full diameter of the subnet) decreasing it every time a node sends the packet. Another flooding control technique, is to keep track of the packets through identification in order to avoid sending them a second time. Also, there is an algorithm called selective flooding that only sends the packet through those lines that are approximately in the right direction.

Flow-Based Routing

The main difference of this technique, is the consideration of the traffic load and the topology to take a routing decision. If we know the capacity and the average flow of a line, then we can find out its mean packet delay, extend this analysis for all the lines to calculate the flow-weighted average to get the mean packet delay of the whole network. Hence, we can base the routing decisions on finding the path that produces the minimum average delay for the network. However, as we can see, it is strictly necessary to know in advance the full topology, the traffic matrix and the capacity matrix of the network. Then, we can apply a routing algorithm.

2.2.2 Adaptive Algorithms

In this case, the topology and traffic changes affect directly the routing decisions. This information is obtained mainly from the routers in the network. The algorithms that belong to this group are: Distance Vector Routing and Link State Routing. Since the routing tables are updated exchanging information from the neighboring nodes to reflect topology changes.

Distance Vector Routing

This dynamic routing uses the Bellman-Ford algorithm ([4], [9]), necessarily implies that each node in the network maintains a table giving the best known distance to each destination, and which link to use to get there. Therefore, the routing table has two parts: the preferred outgoing link for a specific destination and an estimate of the time or distance to that destination. As we have seen, the metric used in the algorithm

depends on the network administrator and its policies. Although, this algorithm works in theory, in the practice its convergence may be slow (the count to infinity problem and bouncing effect).

Link State Routing

This routing algorithm is based on the node neighborhood knowledge and can be stated in five parts as follows:

1. Each node must identify its neighbors and learn their names.
2. Measure the cost (or delay) to each neighboring node.
3. Construct a packet that contains a list of the names and cost to each neighboring node.
4. Send this packet to all other nodes in the network.
5. Compute the shortest path to every node (Dijkstra Algorithm can be used).

2.3 Internet Organization

The internet is a distributed network formed by Autonomous Systems. However, the members that form this network must follow the Internet standards to participate in it. One of the most important concepts to understand about the Internet is the domain. An Autonomous System may have assigned more than one IP address as an identifier, this address allows to other ASes recognized the owner of the AS. In other words, each AS has an address that represents its domain over the Internet.

The administration of the Internet is performed by the Internet Assigned Numbers Authority(IANA) organization. It is responsible of the management of the Internet Protocol addresses and domain names at upper levels. This responsibility is delegated to the following global organizations:

- Réseux Internet Protocol Européens Network Coordinator Center (RIPE NCC).
- American Registry for Internet Numbers (ARIN).
- Asia Pacific Network Information Center (APNIC).
- Latin America and Caribbean Internet Protocol address Regional Registry (LACNIC).

- African Regional Registry for Internet Number Resources (AfriNIC).

They, as well, delegate the IP addresses administration to other regional organizations. Another important organization for the Internet operation is the Internet Engineering Planning Group(IEPG), they are in charge of the global coordination of the Internet Service Providers(ISPs). The IEPG delegates its responsibilities to other regional representatives allowing the dissemination of technical information in networking technologies and operational practices. The representatives of IEPG supports in particular operational issues that involves backbones, Internet Exchange Points and network connectivity. A list is shown with the organizations related to the IEPG around the globe:

- Asia Pacific Network Information Center(ARIN).
- North American Network Operators Group(NANOG).
- European Operators Forum Working Group(EOF).
- Asia Pacific Networking Group(APNG).

2.4 Routing in the Internet

Since the Internet network is formed by Autonomous Systems, we can identify two ways of routing information. The interior routing given within Autonomous Systems and exterior routing given between Autonomous Systems

2.4.1 Interior Routing

All interior routing protocols perform the same basic functions. They determine a route to each destination, and distributes routing information among the systems on a network. The procedures to perform these functions, an the decision process to select a particular route in the network is what makes routing protocols different from each other. The most used protocols for interior routing are:

- The Routing Information Protocol(RIP): This protocol is adequate for local area networks. It selects the route with the lowest hop count as a metric for the best route. The hop count represents the number of gateways through which data must pass to reach its destination; assuming that the best route is the one that uses the fewest gateways. This approach to route choice is called a distance-vector algorithm.

- The Open Shortest Path First(OSPF): This protocol is based on the concept of link-state routing algorithm. Which means that every node has a routing table with the shortest paths to all the destinations in the network.

2.4.2 Exterior Routing

Since this routing is performed between Autonomous Systems, the routing information passed between them is called reachability information. Reachability information is simply information about which networks can be reached through a specific autonomous system. There are two protocols for exterior routing are:

- The Exterior Gateway Protocol(EGP): This protocol announces that it can reach networks that are part of its autonomous system, but it does not announce that it can reach networks outside its autonomous system. Since a routing structure that depends on a centrally controlled group of gateways does not scale well, the EGP protocol was replaced by the Border Gateway Protocol.

- The Border Gateway Protocol(BGP): This is the leading exterior routing protocol of the Internet, because it supports policy-based routing in its routing decision process. This gives to an Autonomous the privilege to choose between routes and to implement routing policies without relying on a central routing authority.

2.5 The Interdomain

The interconnection of multiple networks, named domains or Autonomous Systems (ASes), forming the Internet are called interdomain. There are two types of ASes:

- **Stub:** This ASes only forwards packets for which it is either the source or the destination. Typically smalls Internet Service Providers, Enterprises and Campus are stub ASes.

- **Transit:** This ASes agrees to forward packets for which it is neither the source nor the destination. Most Internet Service Provider networks are Transit ASes.

In order to ensure that a host in a specific domain can reach another host, the routing information must be exchanged between domains. This is the role of the interdomain routing protocol called the Border Gateway Protocol (BGP). The first

version of BGP appeared in RFC 1105 [23] published in 1989. Since then, BGP has evolved to provide better performance and to better suit the needs of network operators. New usages of BGP have also been found. Initially, BGP was used to distribute IPv4 reachability information. Today, BGP is used to distribute IPv6 prefixes, multicast routes, and reachability information in various types of virtual private networks.

What makes the interdomain routing protocol so different from other routing protocols? First and foremost, scalability matters. The interdomain routing system is probably one of the largest distributed systems today. Second, the interdomain routing protocol must support routing policies. Each domain should be able to define its own set of criteria to select (and distribute to its peers) its best path to reach each destination. Finally, BGP is the "glue" that connects all domains that form the Internet together. This makes BGP one of the most critical services in the Internet.

BGP is a path vector protocol, it is similar to a distance vector protocol, but instead of being told a distance to a destination, the path to the destination is used. BGP devices exchange routing information upon initial data exchange and after incremental updates. When a router first connects to the network, BGP routers exchange their entire BGP routing tables. Similarly, when the routing table suffers changes, the routers send the portion of their routing table that has changed. BGP routers do not send regularly scheduled routing updates, and BGP routing updates advertise only the optimal path to a network.

BGP uses a single routing metric to determine the best path to a given network. This metric consists of an arbitrary unit number that specifies the degree of preference of a particular link. The BGP metric typically is assigned to each link by the network administrator. The value assigned to a link can be based on any number of criteria, including the number of Autonomous Systems (through which the path passes), stability, speed, delay, or cost.

2.6 The BGP Protocol

The most important job as an exterior gateway protocol (EGP) is routing information between Internet Service Providers (ISP), because the customer networks usually use an interior gateway protocol (IGP) such as Routing Information Protocol (RIP) or Open Shortest Path List (OSPF) to route information within their networks.

The BGP uses a TCP connection to exchange routing tables with its peers, allowing the AS to know, if there have been changes in the routing information. When these changes occur, the protocol shares with their neighbors only the routes that suffered any changes. We have to keep clear that the routing updates are not periodic and are only necessary to advertise the preferred path to a destination.

The policies are established according to the needs of each organization using the BGP Attributes. They are very important to determine the best route when there are multiple paths to a particular destination. In the route selection process to determine the best path, BGP goes through several attributes such as:

- **Weight Attribute:** This parameter is local, and it is not advertised to the neighbors. If the router learns several paths to a destination, the route with the highest weight will be selected and installed in the IP routing table.

- **Local Preference Attribute:** This attribute is propagated throughout the local AS and it is used to prefer an exit point from the local AS. This means that the local preference will aid the protocol to decide the exit point for a specific route when multiple exit points exists.

- **Multi-Exit Discriminator Attribute:** This attribute is used to suggest to an external AS a route selection. We have to enforce the idea that the external AS that is receiving the attribute metric may or may not take in consideration the suggestion, depending on its route selection policies.

- **Origin Attribute:** A router can learn a new route via IGP, EGP or Incomplete (the origin is unknown or learned some other way). A tag attached to every route (origin attribute) indicating how the router learned the path to a destination.

- **AS-Path Attribute:** This attribute prevents the loop generation in the route advertisements, because every router maintains in the AS-Path, a list of AS's numbers where the route advertisement has been generated. This way, a specific AS number will recognize a route advertisement that it has already been sent throughout and will accept or reject it.

- **Next-Hop Attribute:** This is the attribute that allows the router to know its neighbors that are sending route advertisements and also, is the IP address of the connection between peers.

- **Community Attribute:** As we can infer, this attribute allows group destinations (called communities), and sets the restrictions of the routing updates between them. Predefined community attributes are:

  - **No export:** The route is only advertised between the community members only.

  - **No Advertise:** The route is exclusively for the knowledge of a specific member of the community.

  - **Internet:** All routers in the network belong to this community. Therefore the route advertisement is for common knowledge.

2.6.1 The BGP Path Selection

To select a path for a destination, before registering it in the IP routing table and propagating it to its neighbors, BGP uses the following criteria (in the order presented):

1. If the path specifies a next hop that is inaccessible, drop the update.
2. The path with the largest weight will be preferred.
3. In case the weights are the same, prefer the path that was originated by BGP running on this router.
4. If no route was originated, prefer the route that has the shortest AS-Path
5. If all paths have the same length in the AS-Path attribute, then prefer the path with the origin type. In order of preference will be: IGP, EGP and Incomplete.
6. If the origin attributes are the same, prefer the path with the lowest Multi-Exit Discriminator (MED) attribute.
7. If the paths have the same MED, prefer the external path over the internal path.

8. If the paths are still the same, prefer the path through the closest IGP neighbor.
9. If everything fails, prefer the path with the lowest IP address, as specified by the BGP router ID.

2.7 Routing in Distributed Networks

Any distributed communication network has, as a basic activity, the task of sending packets from one processor to another. The routing scheme is the mechanism used to achieve this goal. This algorithm can be invoked at any source node, and be required to deliver a message to some destination node, [14].

A routing algorithm is a function that for each packet arriving at a node determines the path on which the packet has to go through in order to reach its destination.

2.7.1 Stretch Factor

The stretch factor is the ratio between an alternative path and the shortest path taken to a destination. This rate, allow us to measure the performance of a routing algorithm and has become, a parameter to calculate the storage capacity needed in a network topology.

As we have seen, the most important specification that we are studying in the routing schemes is the path to reach a destination from a specific origin. According to Gavoille in [11], there exists a relationship between the length of the routes and the size of the local data structures used by a routing scheme.

Furthermore, the stretch factor has many other interests, where one of the most important is the measurement of the impact of the routing algorithm. But, mainly the stretch factor measures the tradeoff between the needs of collecting information of the network and the length of the paths.

This factor can be defined as follows. Let R be a routing scheme on a graph G. The stretch factor of R on G is denoted by $s(R, G)$ and satisfies, [11], [7],

$$s(R, G) = \max_{x \neq y} \frac{d_{R(x,y)}}{d_{G(x,y)}} \tag{2.1}$$

where $d_{R(x,y)}$ is the path length between nodes x and y given by the result of the routing algorithm came out on graph G. $d_{G(x,y)}$ is the shortest distance in G. This

means that a routing scheme of stretch factor 1 is termed a shortest path routing scheme. As the reader can clearly see, the optimal stretch value in a routing scheme is 1, but with significantly reduced memory requirements.

2.7.2 Autonomous System Granularity

In this section we are going to explain the granularity at an AS level and the topology of the internet, in order to have a better understanding of the concepts developed in this chapter.

The internet is organized by the Internet Assigned Numbers Authority (IANA), which is divided in several bodies that assists in the Internet administration. This organizations group several Internet Service Providers that may have more than one Autonomous System, (see Figure 2.1) .

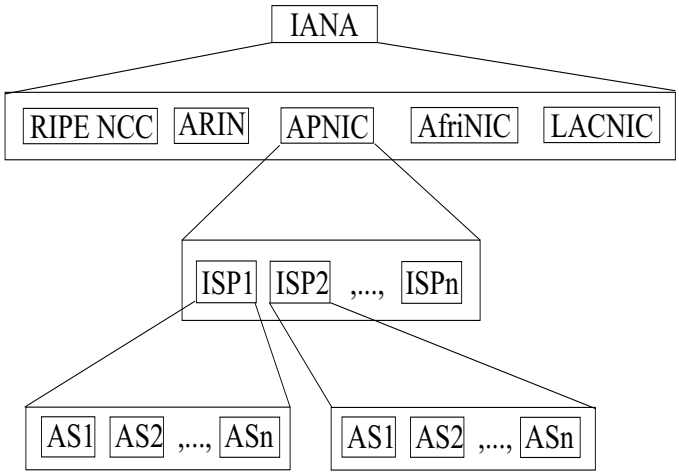

Figure 2.1: Representation of the organizations in charge of the Internet administration.

The routers within in the same AS run a common routing algorithm and have information about each other. The AS need to be interconnected, thus one or more routers in an AS will have to route packets with in the AS and to outside. This router is called a gateway router and its routing process is know as interdomain routing. The Interior Gateway Protocol is the generic name of the intra-autonomous system routing protocol and the Exterior Gateway Protocol is the inter-autonomous routing protocol performed by the Border Gateway Protocol (see Figure 2.2).

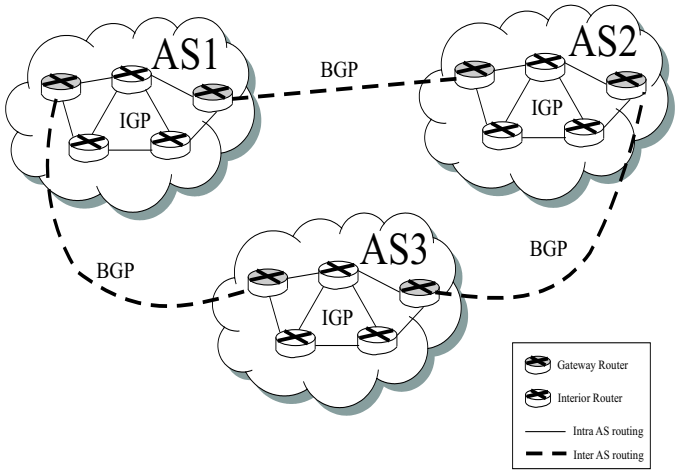

Figure 2.2: Routing information exchange between Autonomous System.

The granularity at an Autonomous System level is defined by the relationship among them. In the interdomain there three types of relationships: Stub AS, Multihomed Non-transit AS and Multihomed Transit AS.

Stub AS

This type of relationship means that an Autonomous System only has one exit point to communicate with networks outside its domain (the Stub AS relation is also called single-homed from its provider). For routing purposes, it could be regarded as a simple extension of the other AS. In fact, an Stub AS does not need to learn the Internet, because the provider is announced and the Stub AS obtains Internet access through its unique link with the provider(see Figure 2.3).

Multihomed Non-transit AS

The main characteristic of this relationship lays in that the Autonomous System has more than one connection with different Internet Service Providers. However, in the Non-transit type of Multihomed, the Autonomous System only will announce its own routes (this means that the routes learned from each ISP connection will not be know by others AS). Therefore, the traffic flowing from the AS to the each ISP will be the corresponding to the routes learned from every ISP in specific. In the Figure 2.4, the AS4 learns routes r1 and r2 from ISP1, in the same way learns routes r3 and r4 from ISP2. However, the AS only will propagate its own routes r5 and r6.

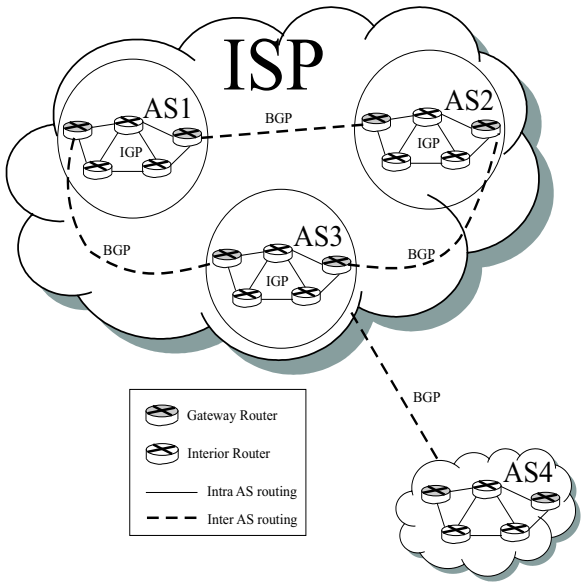

Figure 2.3: An example of the Stub AS relationship.

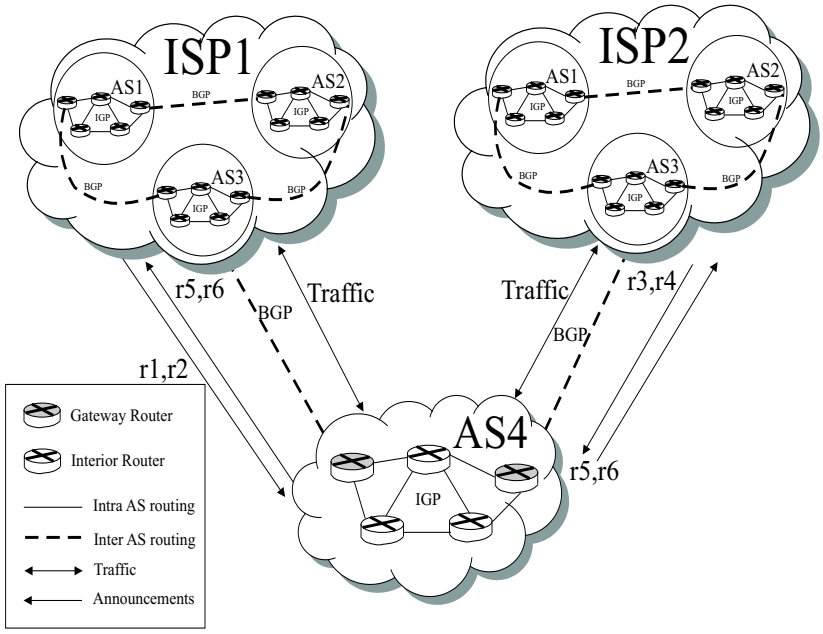

Figure 2.4: An example of the Multihomed non-transit AS relationship.

Multihomed Transit AS

When an AS has multiple connections with different ISPs and this connections are used to propagate traffic from other AS, the AS is called a Multihomed Transit AS. It will announce the routes learned from other AS to all the AS known, sending traffic to an AS that does not belongs to it. The BGP connections between gateway routers into the AS are know as internal BGP (IBGP), whereas connections between routers from different AS are known as external BGP (EBGP). In Figure 2.5, the AS4 learns routes r1 and r2 from the ISP1, in the same way it learns routes r4 and r5 from ISP2, and it announces all the known routes to all its connections with the ISPs. Therefore, ISP1 uses the transit AS4 to reach r3 and r4 into ISP2, in the same way ISP2 uses AS4 to reach r1 and r2 into ISP1.

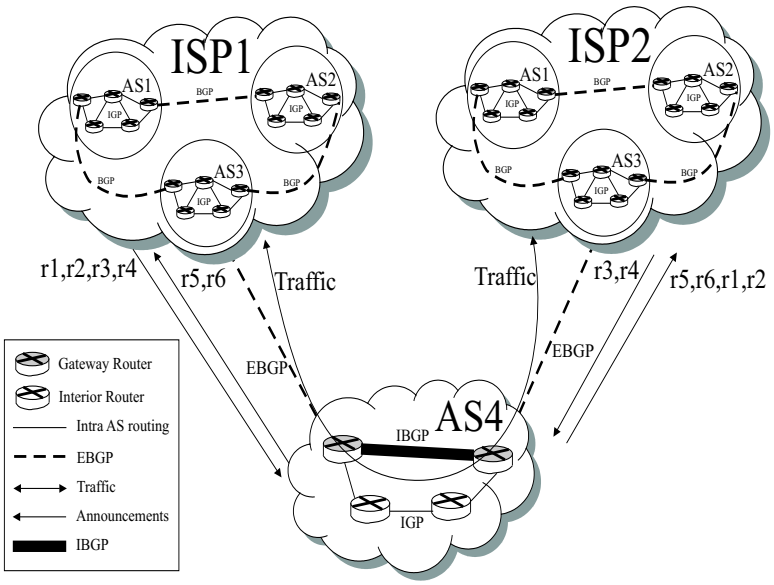

Figure 2.5: An example of the Multihomed transit AS relationship.

In this work we are interested in the Multihomed relationships between Autonomous System. Therefore from this moment an Autonomous System with a customer−provider relationship can be also represents a node in the network (see Figure 2.6).

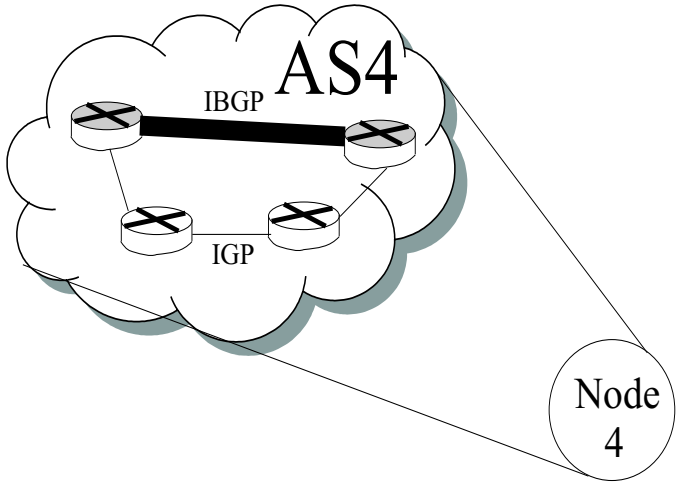

Figure 2.6: An autonomous system as a network node representation.

Chapter 3

Simulation

In this work, we are making an analysis of the internet routing under a new perspective, generating a network that can be very useful to emulate the Internet topology. These days, the researchers have been proposing several power law topologies to evaluate the internet, [1],[2],[26],[8]. However, the problem of modeling the internet topology and creating realistic topologies still an important open problem. Therefore, our simulation only can be considered an approach to the internet topology based in the results of the researchers before mentioned.

3.1 Network Generator

The granularity of the internet (i.e. see Figure 3.1) is given by the different levels that we can identify in it. At a global level the internet is divided into regions encompassing several countries, which also group different Internet Service Providers(ISPs). The ISPs may have one or more Autonomous Systems representing domains, which also group border and interior routers that will give internet access to the end users.

In this work, we will focus our attention in the granularity at an Autonomous System level given by the relationships among them defined by [13]. These relationships between Autonomous Systems allows other AS to send traffic, updates or withdrawals. The most common relationship is the commercial one; we support the idea that the trustable relationship between two AS's is the customer-provider (as shown in Figure 3.2.), since there is a monetary compensation for the service.

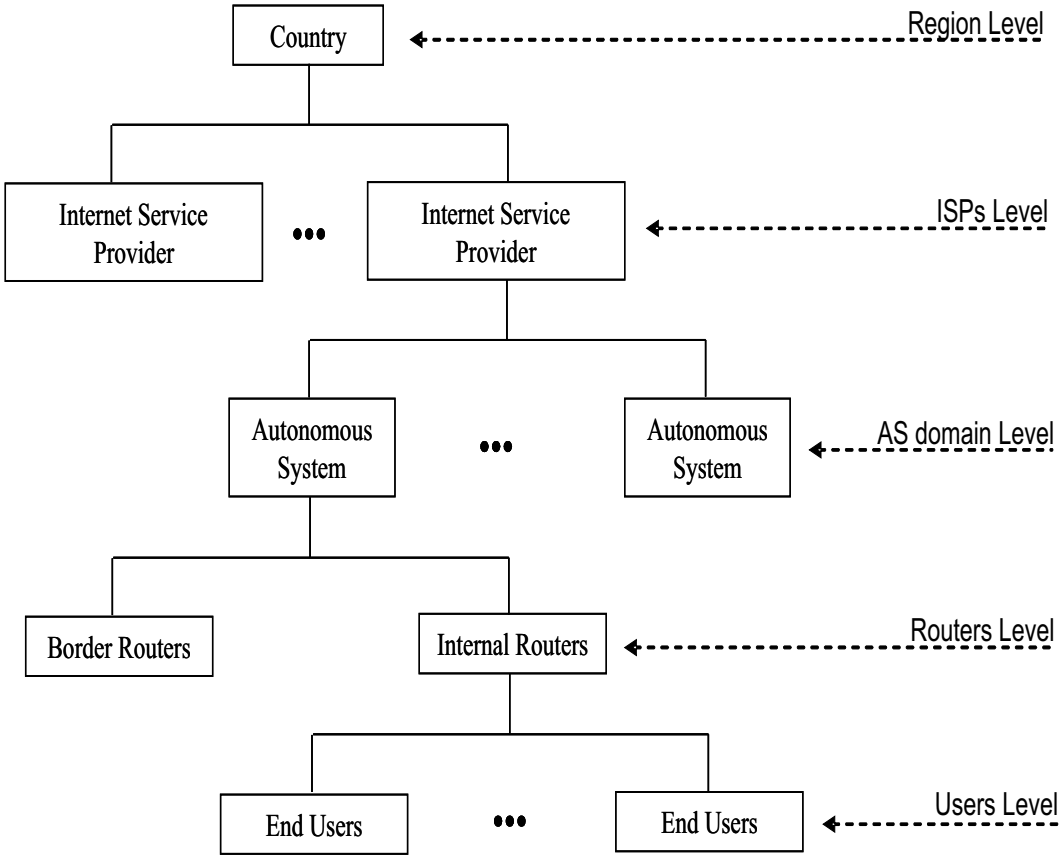

Figure 3.1: Internet granularity at a region level representation.

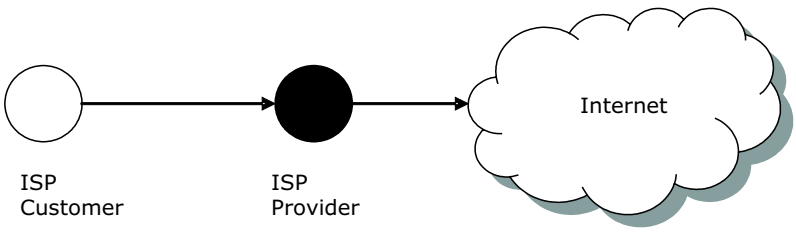

Figure 3.2: Customer-Provider relationship between Autonomous Systems, [13].

Following the idea of a trustable relationship between ASes, we can conceive a network where, every AS only considers the links in its circle of trust. The circle of trust is defined by the ASes that are reachable at one hop (assuring that they are not spurious) and therefore, they are always going to let the traffic flow forward. Under this assumption, there is not a routing table with the defined paths to reach every destination in the network, the only knowledge that they have, are the links to the nearest neighbors to one hop of distance. This idea is better shown in the Figure 3.3, where the circle of trust of the white node is represented by its neighboring black nodes.

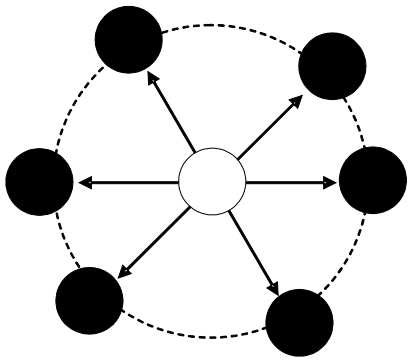

Figure 3.3: Neighbors Autonomous Systems at distance 1.

The commercial agreements between pairs of administrative domains can be classified into customer-provider, peering, mutual-transit, and mutual-backup agreements, [17],[18]. A customer pays to another Internet Service Provider (ISP) for connectivity to the rest of the Internet (The provider does transit traffic for its customers). It is important to underline, that a customer does not let traffic through between two of its providers.

The peering relationship is basically an agreement to exchange traffic between their respective customers free of charge. A mutual-transit relationship allows two ISPes (typically between small ISPes who are located close to each other and cannot afford better connectivity Internet services) of administrative domains to provide connectivity to the rest of the Internet for each other. Between ISPes is also common to provide backup connectivity to the Internet for each other in the event that one administrative domain's connection to its provider fails.

At this point, it is very important to clarify the trust relationship beyond the commercial relationship. The trust chain idea can be illustrated by using an example of our daily life. Suppose that we are going to open a bank account; the account

Table 3.1: Definitions and Symbols

Symbol	Definitions
E	Set of links
N	Set of nodes
$G(N, E)$	A directed graph
n	Number of nodes in the graph
κ	Number of trustable nodes in each vertex
δ	The maximum number of links from one vertex to another
$\delta(\upsilon, \nu)$	Number of hops from node υ to node ν

executive will ask us to fill out an application where, we have to include our address, full name, income, etc. This is classified information for us, however we do not hesitate to give this information to the bank, because they have our trust, but this does not mean that the bank trusts in us.

This is the kind of relationship we have in our network, the trust relationship is only forward (as you can see in Figure 3.4), the fact that we are part of a circle of trust (CoT) into an AS does not imply that this AS will be part of our circle of trust.

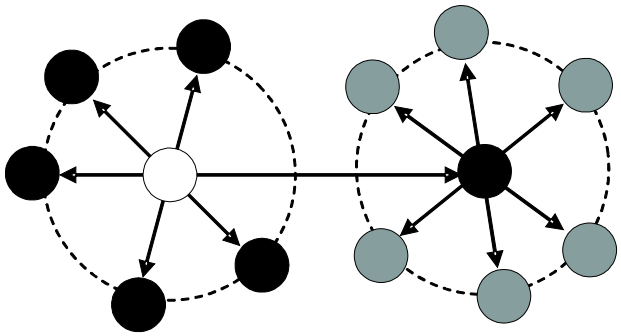

Figure 3.4: Forward trust relationship between Autonomous Systems.

The relationships between ASes may be represented as a directed graph G where each vertex is an AS, the edges are the trustable ASes links and the degree of the vertex is the number of ASes in the CoT. In order to develop this idea, we have to consider the definitions in Table 3.1.

Using this definitions, we can say that an AS ν that belongs to the N nodes can create a CoT where $1 \leq \kappa \leq n-1$. More formally, the number of distinct circles of trust q that can choose a node ν with κ trustable ASes from the $n-1$ nodes will be the result of a binomial coefficient,i.e.,

$$q(\nu;\kappa) = \binom{N-1}{\kappa}, \quad \forall \quad \nu \in N \quad and \quad 1 \leq \kappa \leq n-1. \tag{3.1}$$

as we can see, the maximum number of CoTs that our graph can have is n. Therefore, we can define that the probability of a ν vertex to choose κ different nodes from $n-1$ will be defined by

$$p(\nu;\kappa) = \prod_{x=1}^{N} \binom{N-1}{\kappa}, \quad \forall \quad \nu \in N \quad and \quad 1 \leq \kappa \leq n-1. \tag{3.2}$$

However, how can we be sure that there will not be concentrations of nodes forming a big star network topology? Well, it is not an easy question. In order to answer this questions we can find the conditional probability of having the same node in the circles of trust $c_i, i = 1, \ldots, n$ that will be given by

$$p(v_i \in c_i \mid v_i \in c_j) = \prod_{x=1}^{n} \left(1 - \frac{1}{n-1}\right), \quad \forall \quad i \neq j. \tag{3.3}$$

This means that when $n \to \infty$ the probability of attraction of one specific node as part of a trust circle will tend to one.

3.2 Circle of Trust Size (CoTS)

In this section we are going to observe the data of some Autonomous Systems to obtain an approximation to the number of nodes in the circle of trust for our simulation, based in the calculation of the mean value in the one hop distance statistics.

One of the most important parameters in our network is the κ size of the CoTs. Observing the interdomain and taking the statistics shown by the UUNet Canada through the AS701, the Reach Network Border Node AS4637 and the Telstra AS1221 we have the following analysis.

The statistical data of AS4637 is shown in, Figure 3.5, calculating the mean value from the number of hops at 1 hop distance; we found that 11.12% of the entire ASes present in the Internet are at 1 hop distance from AS4637 with a variance of 0.5%.

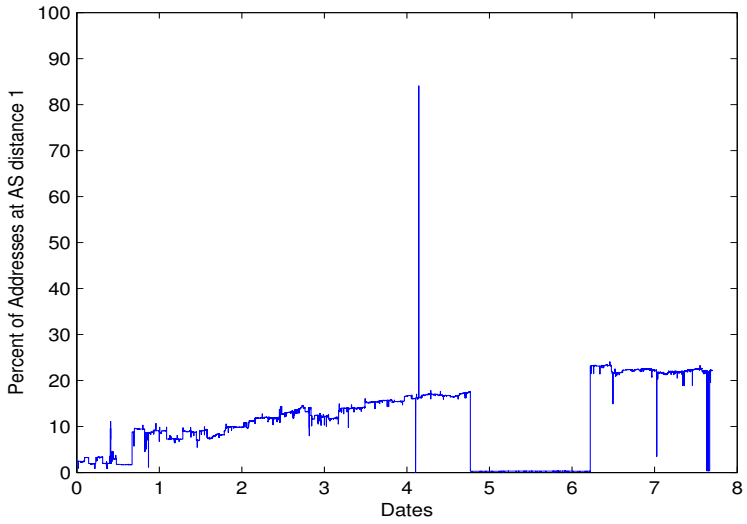

Figure 3.5: AS4637 Statistics from year 2000 to 2007.

In the same way, we can observe the statistics of AS701 shown in Figure 3.6 where the mean value of the ASes addresses at distance 1 from it is 18.2% of the entire ASes present in the Internet and the variance is 0.02%.

Following the same procedure, in the observation of AS1221 statistics shown in Figure 3.7 we can say that the mean value of the ASes addresses at distance 1 from the AS1221 is 10.49% of the entire ASes present in the Internet and the variance is 3.29%.

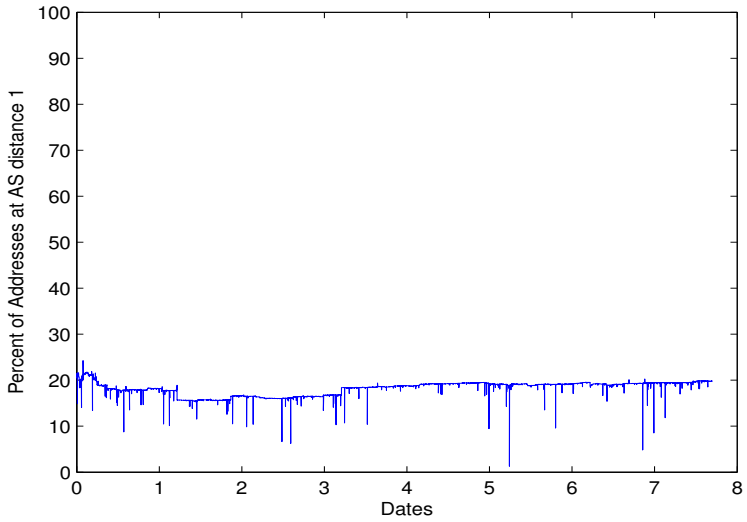

Figure 3.6: AS701 Statistics from year 2000 to 2007.

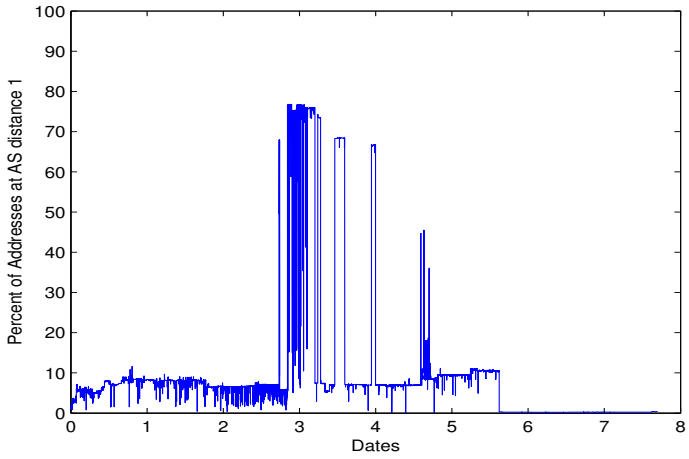

Figure 3.7: AS1221 Statistics from year 2000 to 2007.

Supporting us in these statistics, we decide to use a $\kappa \geq (n-1) * 10\%$ since this is the minimum mean value observed in the three Autonomous Systems shown in Figure 3.8, which means that the range of values corresponding to the number of ASes belonging to a specific CoT will be given by $0.1 * (n-1) \leq \kappa \leq n-1$.

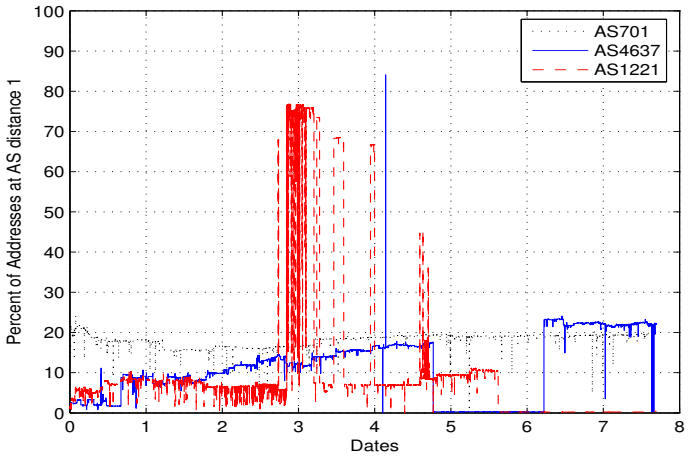

Figure 3.8: AS1221,AS701 and AS4637 Statistics from year 2000 to 2007.

3.3 Simulation Time

Time is an important concept in any performance. We maintain a simulation clock whose value is the current time in the algorithm. This simulation time is distinctly different than the computer time. Simulation time starts at zero and then advances in discrete jumps of one (the clock is an integer variable). Therefore, it is not possible to make time move backwards during a simulation run.

In this work a power law time is a time duration given by a power law random variable generated in the simulation. Therefore, in our simulation we are going to observe the network topology every power law time; based in the results of [16], where states that the times between topology changes follows a power law distribution. In other words, the algorithm will collect the network topology information as long as the time duration between observations will be a power law time.

The simulation will observed the network every power law time. This is a very important issue, due to the time between UPDATES messages (topology changes) observed in the BGP protocol follows a power law distribution [16] as we will analyze in this section.

Modelling the Internet topology has been in the last days focus of attention of many researchers. Siganos and Faloutsos [26], observed three power laws (Rank Exponent, Degree Exponent and Eigen Exponent) in the interdomain level, where each node represents an Autonomous System and each edge is an interdomain connection. In this

work the authors shown that the power law relationships also hold for the interdomain topology.

3.3.1 Power Law Times Between Topology Changes

The changes into the Internet topology can be seen through the BGP protocol via the UPDATE messages. The delay between changes can be determined from these messages as well as the time between changes. An approach to the determination of the times between the changes in the Internet topology is analyzed in [16] where, the direct conclusion is that these times observed are also described by a power law distribution.

When a node is added to a network, it triggers a chain of restructuring changes. Every time a route is advertised or withdrawn, an UPDATE message notifies this event. Hence, as in [16], this is considered as a topology change. This change affects the network nodes in a wave propagation mode.

In [16], The author obtained original data from AS701, AS1239, AS7018, AS3130, AS1239, and AS7575 and approximate the parameters of a power law via the Min-Square method, resulting the following power law, [16],

$$y = e^{\beta} x^{-\alpha}. \tag{3.4}$$

The power law functions obtain by [16], are shown in the loglog plot, Figure 3.9. The author states that the correlation with the original data was of 93% and that the power law characterizes better the behavior for the lowest interarrival times, while the exponential is better in the characterization for the bigger interarrival times.

3.4 Simulating Topology

The internet is dominated by a handful of large ISPes called as tier one providers, and they account for the majority of routes and bandwidth that comprise the public Internet. These public exchange points are considered the core of the Internet where providers peered, or exchange routing information and traffic, [20].

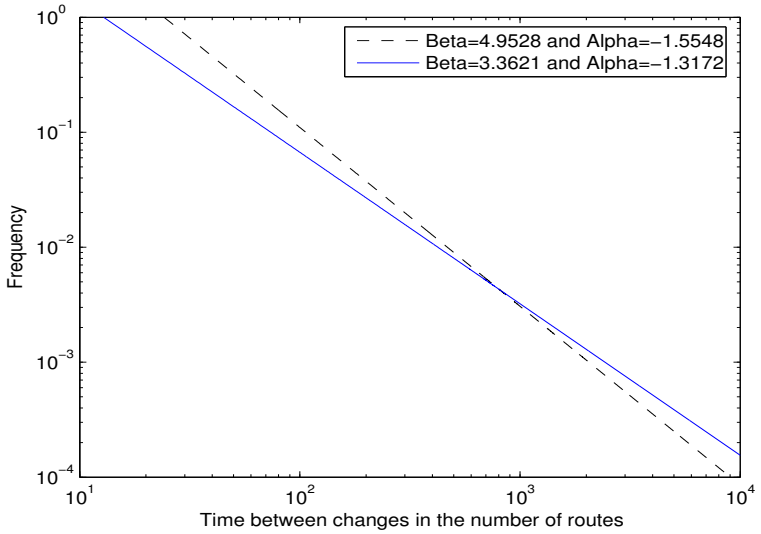

Figure 3.9: Power Law times between topology changes approximations, [16].

3.4.1 Network Backbone

In [3] the authors, state that even if the internet topology obeys a power law distribution, the internet core connectivity does not. The authors also defined the customer−provider relationship based in the concept of outdegree and indegree; where the first one, represents the directed customer links that are connected to a provider node (the links that send traffic into the provider) and the second one, represents the directed provider links that a customer may have to send traffic out of its node.

Using heuristic in [13], the authors obtained a directed provider to customer graph defining indegree and outdegree adjacency matrices to infer AS relationships. Where, the indegree adjacency matrix A^I has components $\{a^I_{ij}\} = 1$ when node i is a customer of node j or node i is a peer of node j and $\{a^I_{ij}\} = 0$ otherwise. The outdegree adjacency matrix A^O has components $\{a^O_{ij}\} = 1$ when node i is a provider of node j or node i is a peer of node j and $\{a^I_{ij}\} = 0$ otherwise.

The observed results of this research, state that around 79.7% of the ASes are end customers with 61.6% of the links and the 9% of the ASes are small provider ISPes with 8% of the relationships. These ASes act as transit between the end customers and the core of the Internet, but they do not have any peer to peer relationship among themselves.

Just like in the interdomain, our simulation includes a backbone that will help to maintain always a degree of connectivity among the nodes. The number of nodes that belongs to the backbone represents the 10% of the network nodes. This value is determined as an approximation to the 11.18% of ASes included in the internet core according to [3].

3.4.2 Topology Changes

The changes in the topology are given every time a node starts or stops working. Furthermore, we can say that a node in the network has two possible states (reachable or unreachable). In order to simulate this process we will resort to the indicator function to characterize the reachability event.

More formally. Let A be an event related to the state of a node in the network. The indicator function for A is defined by

$$I_A(\zeta) = \begin{cases} 0, & \text{if } \zeta \text{ not in } A, \\ 1, & \text{if } \zeta \text{ in } A, \end{cases}$$

that is, $I_A(\zeta)$ equals one if the node is reachable, and zero if it is unreachable. I_A is a discrete random variable with range $S_x = \{0, 1\}$, and its probability mass function is that of Bernoulli trials with success probability p, i.e.,

$$p_0 = 1 - p, \quad \textit{and} \quad p_1 = p, \tag{3.5}$$

3.5 Network Generated with a Deterministic CoTS

In order to get a better understanding of the way how our network generator works, we perform an example that includes all the issues developed in this chapter.

Let $G = (V, E)$ be a directed graph with $N = 100$. Hence, as a consequence we have that $\kappa = 0.1 * N = 10$ and the network backbone will include also 10 nodes. Besides, we have to consider the following assumptions:

1. In a cold start all nodes in the network are reachable.
2. The backbone nodes are considered reachable over simulation time, since represents the internet core.
3. The customer access the network through its provider, since the customer have a forward relationship with its provider.

Table 3.2: AS Number of Nodes in the Network

AS90	AS9	AS67	AS18	AS62	AS34	AS54	AS154	AS105	AS121
AS99	AS123	AS78	AS128	AS139	AS50	AS38	AS94	AS153	AS77
AS101	AS132	AS87	AS143	AS95	AS117	AS125	AS40	AS25	AS14
AS142	AS81	AS145	AS135	AS80	AS119	AS29	AS10	AS5	AS26
AS8	AS13	AS74	AS55	AS156	AS82	AS6	AS70	AS37	AS42
AS148	AS147	AS68	AS100	AS76	AS157	AS152	AS23	AS58	AS133
AS47	AS120	AS71	AS66	AS144	AS111	AS11	AS63	AS45	AS140
AS112	AS17	AS106	AS150	AS27	AS97	AS85	AS130	AS114	AS116
AS19	AS115	AS158	AS136	AS92	AS59	AS36	AS52	AS46	AS109
AS48	AS2	AS43	AS56	AS33	AS53	AS79	AS151	AS51	AS146

4. The two nodes that belong to the relationship customer−provider are always reachable.

5. The largest path length from an origin to a destination will be $n - 2$.

6. All the nodes into an specific CoT will be different among themselves.

Once the assumptions above are satisfied, we can say that we have Table 3.2 of AS nodes in the network (the Table does not have any particular order).

Based in a uniform random variable we obtained a backbone conformed by, AS152, AS10, AS13, AS150, AS46, AS33, AS29, AS116, AS23 and AS117. Furthermore, let AS117 be the customer and AS10 the provider. Therefore, node AS117 will reach the network through AS10. In Figure 3.10, the links among backbone nodes are shown.

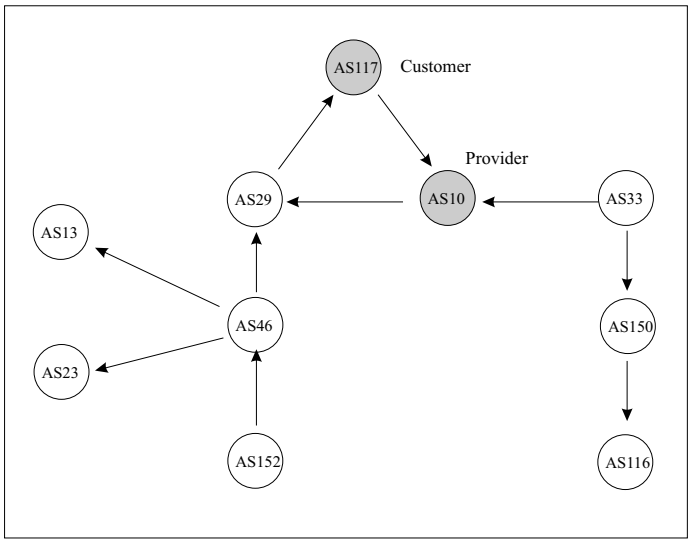

Figure 3.10: Backbone connections and Customer Provider Relationship.

Table 3.3: CoT of the nodes involved in the shortest path from AS10 to AS148

AS10	AS29	AS147	AS119	AS90	AS125	AS76	AS112	AS80	AS136	AS153
AS29	AS100	AS76	AS11	AS26	AS9	AS19	AS117	AS55	AS135	AS45
AS45	AS9	AS54	AS115	AS11	AS2	AS77	AS10	AS17	AS117	AS148
AS148	AS54	AS147	AS114	AS48	AS10	AS51	AS25	AS9	AS71	AS139

We must remark, that in our network the most important thing is the nodes that belong to our CoT. Therefore, from this information we can find the path length to all the destination nodes in the network through the AS10.

In order to perform this task, we need to run a recursive algorithm to find the shortest path length to every node in the network (from AS10 to every reachable node, excluding AS117).

For instance, suppose that we want to know the path length from the origin nodes AS10 and AS117 to the destination node AS148. For the corresponding path, we need to know the ASes adjacent to the AS 117 and follow the information until we reach the AS148 (assuming that its state is reachable). To help us in this task, we need the Table 3.3 with the CoT to the nodes involved in the route.

In the Table, the first column includes the origin ASes and the rest of them the nodes at distance 1 into the CoT of its corresponding AS. If we consider the AS10 as the origin and lookup the into its CoT nodes we can find that AS10 shortest path to AS148 is $AS117 \rightarrow AS10 \rightarrow AS29 \rightarrow AS45 \rightarrow AS148$ only takes 3 hops from AS10 to AS 148. Using this procedure for every destination from the origin AS10, we can find the shortest paths for all AS in the network.

Once we have a table with the shortest paths from the origin to all the destinations, a table with the CoTes for all the ASes in the network and a defined backbone. We will begin to take snapshots of the network every power law time, checking the reachability of every AS (the state of the nodes will be random depending of a Bernoulli variable) and therefore recalculating the shortest path to every reachable destination from the origin node AS10.

Chapter 4

Results

In this chapter, we introduced the methodology used in the creation of products of this thesis. We also explain the stages needed to obtained the results. In this work we will use the following steps to analyze the data obtained: first we will use the autocorrelation, since we have to be sure that the nodes reachability events are independent; second we will use the periodogram, since we have to analyze the periodicity of the data to be sure that we can find a knowledge distribution for our information; third we will compare the estimated distribution function of the data obtained with the normal, poisson, exponential and power law distributions using the cumulative distribution function, quantile and probability comparisons to characterize the reachability process.

4.1 Network Generator

As we explained before, our network generator is not based on the full topology previous knowledge, but in the relationship between the nodes and therefore, in the nodes that belongs to its circles of Trust. In order to generate the results, we have to go through two stages: The Initialization and The Simulation.

4.1.1 The Initialization Stage

This stage allows us to generate relationships between nodes, according to the parameters of the network that we want to analyze. The variables to define in this stage are: Number of nodes, the CoT size, the Customer-Provider nodes and the nodes that will belong to the backbone. The Figure 4.1 shows the procedure to follow, in order to initialized our network generator.

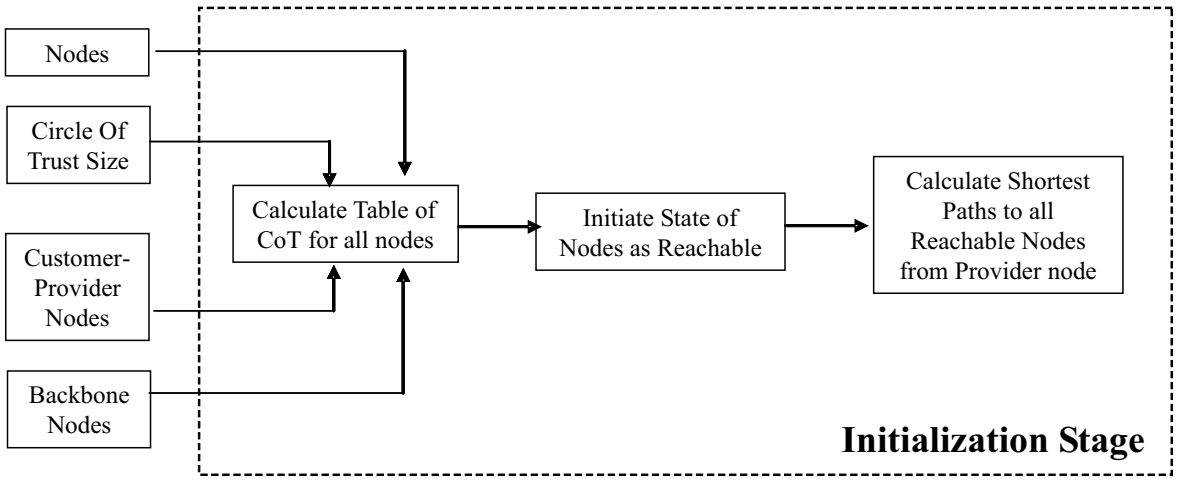

Figure 4.1: Flow Chart of the Initialization Stage of the Network Generator.

Once we have set the variables with its corresponding values explained in Chapter 3, we can begin to run the initialization of the next procedures.

Table of CoT for Nodes

Before obtaining this table, we have to consider some rules that the CoT must follow. Let i be the node and C_i the corresponding CoT. Hence, the following characteristics must be accomplished.

1. The maximum size of C_i will be N-1 and the minimal will be 1 node.
2. All the nodes that belong to C_i must be different to i.
3. All the nodes in the C_i must be different among them.
4. Every node in C_i it is to one hop of distance from the i node.

Initializing Node States

As we explained before, in a cold start, the state of all the nodes in the network must be marked as reachable. This means that at this moment, it is possible to reach all the nodes in the network and therefore, use all the available links included in the CoT table of the nodes.

Shortest Path Table

At this point, we are able to calculate the Shortest Path Table. Due to in the cold start we have all the nodes reachable, we assume these as the best case of our network simulation and as a consequence, at these moment we can obtain the optimal Shortest Path Table to reach all the destinations (we should remind you, that this table is only valid for an specific provider node variable chosen as origin).

The singularity of this network generator resides, in the ignorance of the network topology (They do not have a table with all the paths to the destination nodes). Therefore, in order to find the shortest path to a destination we construct a recursive algorithm (see Figure 4.2) that will find all the paths from provider node to the destinations based only in the CoT table.

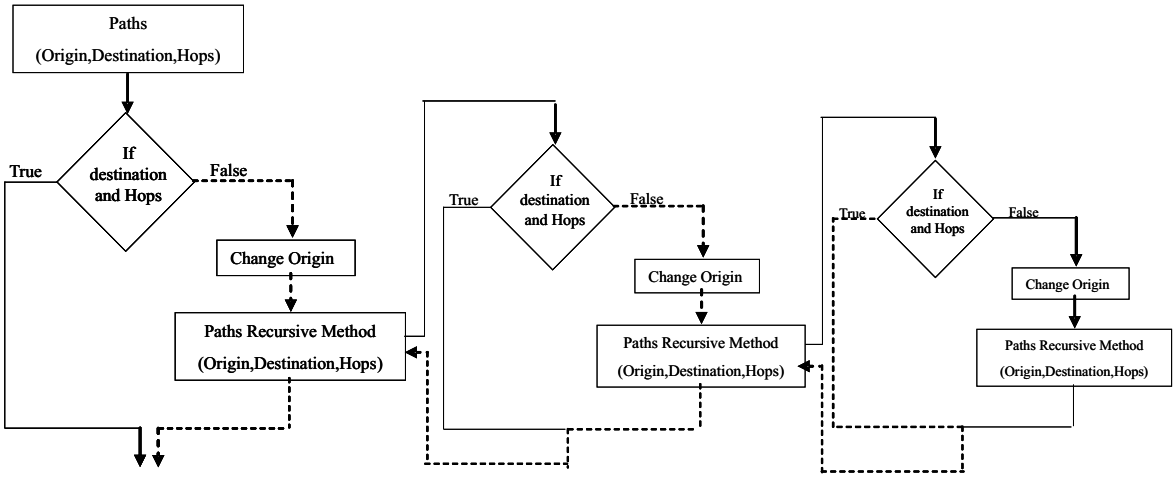

Figure 4.2: Recursive flow of control to find the paths using CoT information.

In each recursive step, the algorithm will replace the origin given in the past evaluation with a node in the CoT of the past origin, in order to lookup for the destination in the next CoT. Once we have all the available paths to a destination node we collect one of the paths which length is minimum.

4.1.2 The Simulation Stage

This stage will allows us to simulate the changes in the network topology, given by the reachability changes in the nodes (We should remind you, that every time a node changes from reachable to unreachable -or viceversa- we assume that the network topology has changed). Furthermore, at this point we are able to collect information of the network into an specific time.

As the reader can to infer, the changes in the topology will also change the Shortest Path table, giving us the opportunity to analyze not only the reachability process, but also, the dynamism in the stretch value rate through the simulation time.

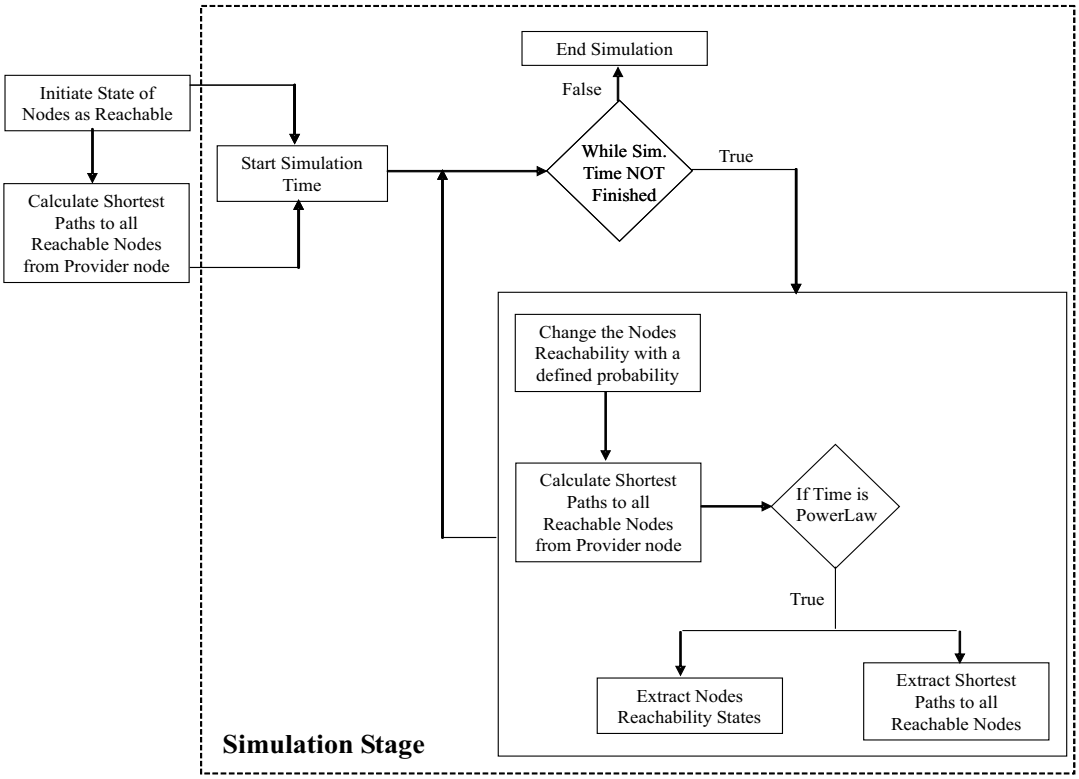

Figure 4.3: Flow Chart of the Initialization Stage of the Network Generator.

Topology Changes

The simulation is based in two logical conditions:

1. The simulation will stop only when the time expires.
2. The information of reachability node states and shortest path will be collected only if the time corresponds to a random power law value [16].

The procedure into the while loop, will update (at every simulation time) the reachability state of all the nodes that not belong to the backbone or to the customer-provider nodes will change with a Bernoulli random variable. Afterwards, the shortest paths will be recalculated from origin to all the reachable destinations.

Once we have the new reachability information, and the new shortest paths, the algorithm will proceed to evaluate the second logical condition. If the condition is satisfied it will collect the information in two matrices.

The reachability matrix row corresponds, to the simulation time iteration and the columns, to the information about the nodes state. The shortest path matrix, will calculate the stretch rate using the optimal shortest path table vs the new shortest path; again, the row corresponds to the simulation time iteration and the columns to the nodes stretch.

4.2 Reachability Analysis

In this work, reachability is defined as the property of a node to be connected to the network (at any time) of being used as a possible destination or as a transit node, in order to send traffic from an origin at a specific time.

In Figure 4.4, it is shown the reachability process of a node over simulation time (the time was cut for observation purposes). It is not rare to see the reachability as a discrete process, since we are using Bernoulli trials to affect the reachability state of the nodes. We should remind you that since zero values represents the simulation times where the node state was marked as unreachable; we can say that most of the simulation time the node was reachable.

However, in [16], states that the network topology remains with out changes in the power law time duration. This leads us to consider the reachability as a continuous process. In above Figure 4.5, we can see the resulting continuous process, where we can appreciate that the power law time duration when the node remain unreachable is very low in comparison with the simulation time.

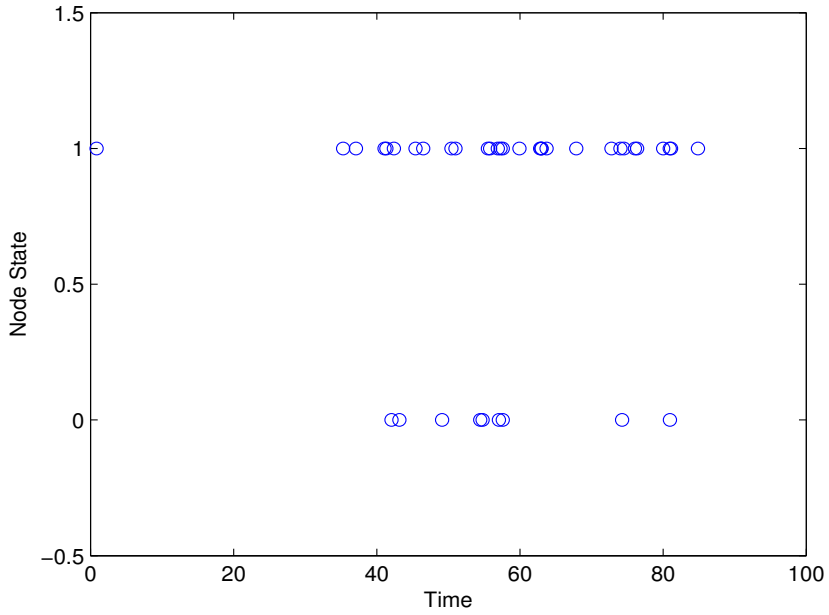

Figure 4.4: Reachability states of a node during the first 100 simulation times.

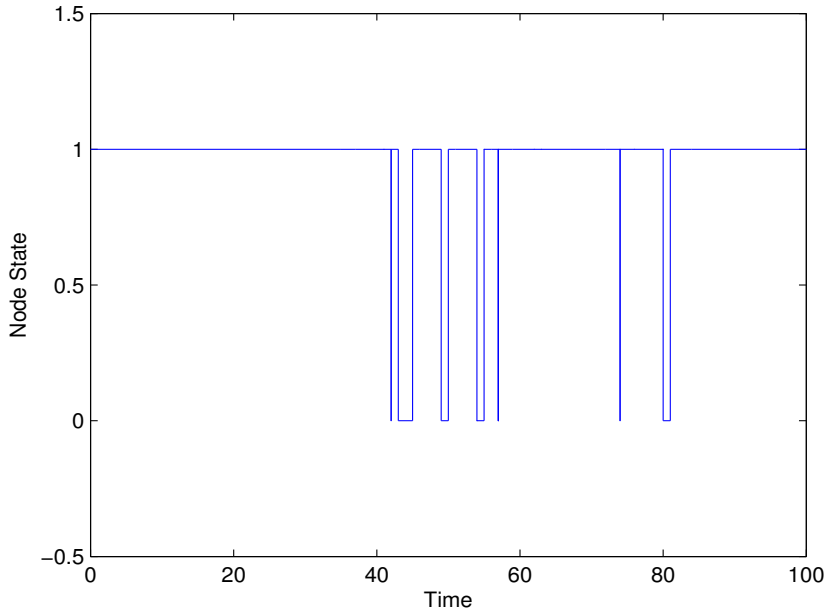

Figure 4.5: Reachability states of a node without changes between simulation times.

4.2.1 Statistical Reachability Analysis

The statistical analysis of the data obtained in the simulation, can help us to get a better understanding of the process as a inherent phenomena that has not been studied in deep under these conditions. This work is a first approach to the characterization of the reachability process considering power law times between topology changes.

The scenario used in this simulation to obtain the information is based on a network formed by 1000 nodes, running 6000 iterations. In [21], the authors conclude that when an Autonomous System is under an attack its reachability is approximately of 50.49%. Therefore, in our simulation we set the reachability probability of the autonomous system in 50% to study the reachability process under an attack situation.

From this simulation, we obtained 1000 observations of the topology changes. Using this information we calculate the number of nodes reachable for every observation to calculate the differences in the number of reachable nodes between observations.

Taking the data obtained from the differences, we can construct a histogram that will help us get an approximation to the reachability process distribution. The result is the histogram shown in Figure 4.6.

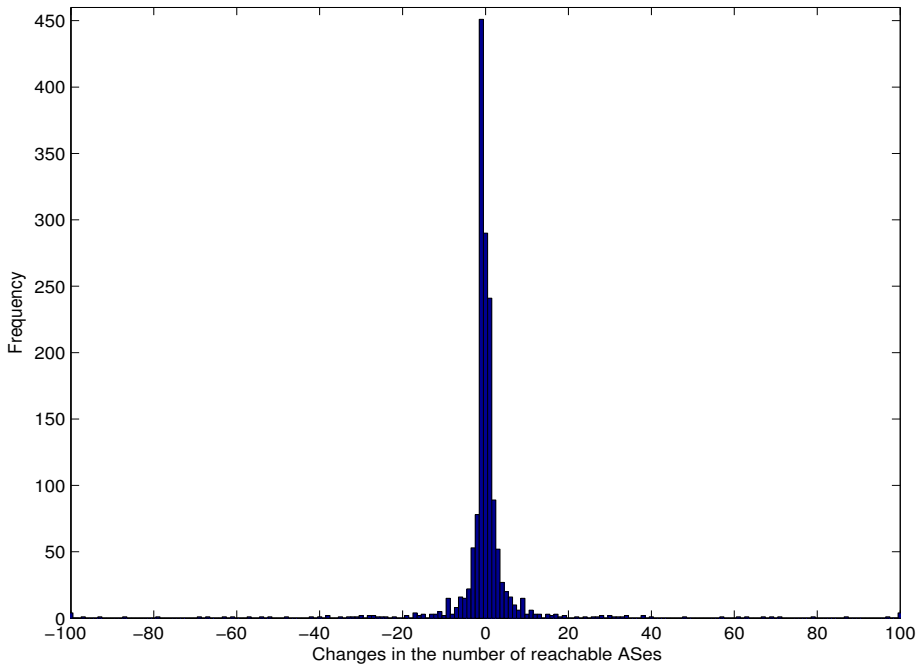

Figure 4.6: Changes in the number of reachable ASes between topology observations.

The above Figure 4.6 shows us the changes in the number of reachable ASes observed between topology changes. The positive side corresponds to the decrease in the number of reachable nodes observed by an AS, this side represents the reachability frequency of an AS regarding the network; while the negative side represents the increase in the number of reachable nodes, in other words the growing of the network between topology observations. In this work we are interested in the reachability process, therefore our analysis will be focus on the decrease of the number of ASes observed between topology changes (see Figure 4.6).

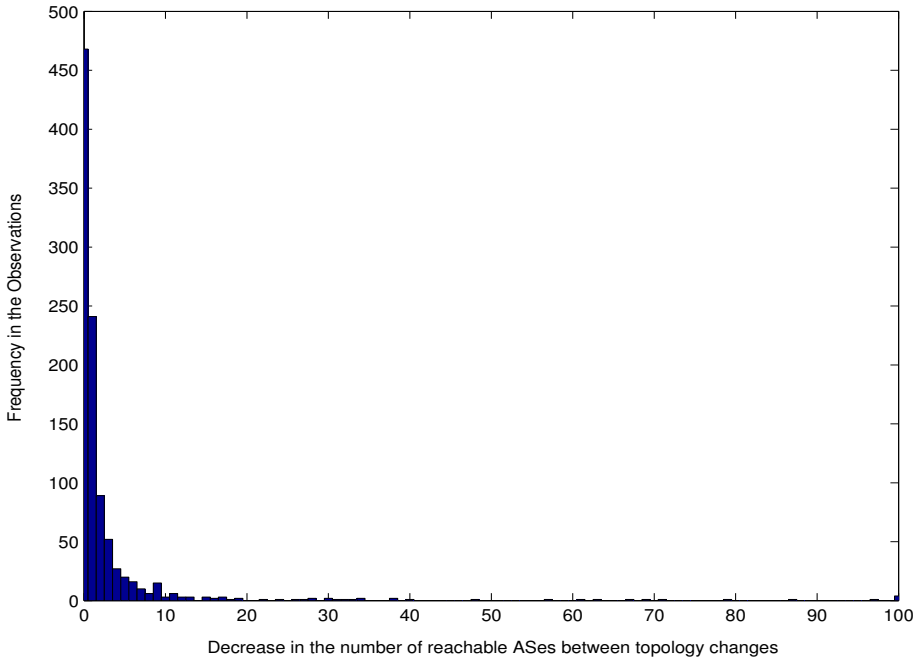

Figure 4.7: Histogram of an Autonomous System reachability over topology changes.

From Figure 4.7, we can say that there is a high frequency in the little changes on the number of reachable nodes and a low frequency of founding big changes in the number of reachable nodes over the network. Therefore, it is more probable to found little changes in the number of reachable nodes, despite of the topology changes in the network. This means that with a high probability the node will be able to reach the network, although there is an attack situation that affects the network topology.

In Figure 4.8, we show the autocorrelation plot of the reachability process. Since, the plot result does not shown a triangle form, we can say that the changes in the number of reachable nodes over the topology changes observations are independent and uncorrelated among them.

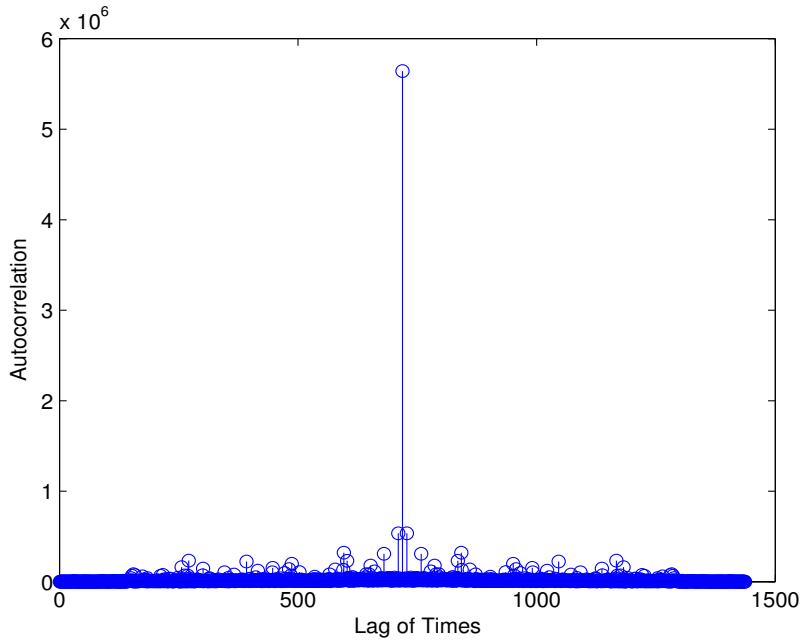

Figure 4.8: Autocorrelation of the reachability process.

Therefore, we can say that the changes in the number of reachable nodes in the topology change i is independent from the number of reachable nodes in the topology change j for the t observations. Hence, changes in the number of reachable nodes over the network are independent.

From the histogram, we can obtain the estimated probability density function -pdf- (see Figure 4.9). This estimated function will assist us in the characterization of the reachability process, since allow us to compare our statistical data with other distribution functions.

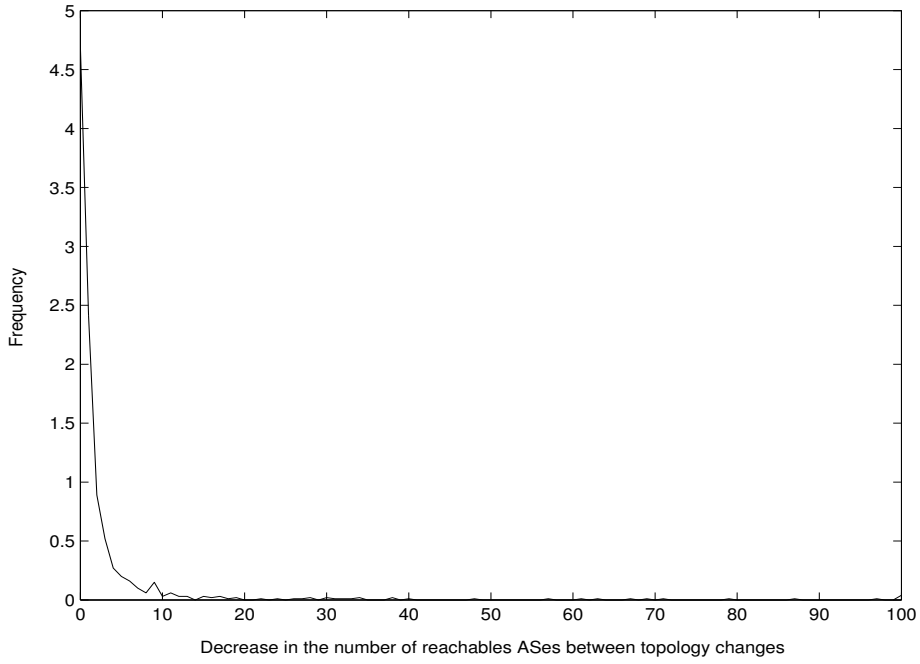

Figure 4.9: Estimated probability density function(pdf) of the reachability process.

Before continuing with the characterization of the reachability process, we need to analyze our data in the frequency domain. In Figure 4.10, is shown the periodogram plot of the data in the frequency domain.

The power spectral density amplitudes in Figure 4.10, do not have periodical repetitions that can show us a periodicity in our reachability process. Besides, there is not a convergence of the power spectral density amplitudes to a unique value that formed a defined peak for all the data information. Therefore, we can say that our data information does not have a chaotic distribution. Hence, we can find a known distribution function to characterize our reachability process.

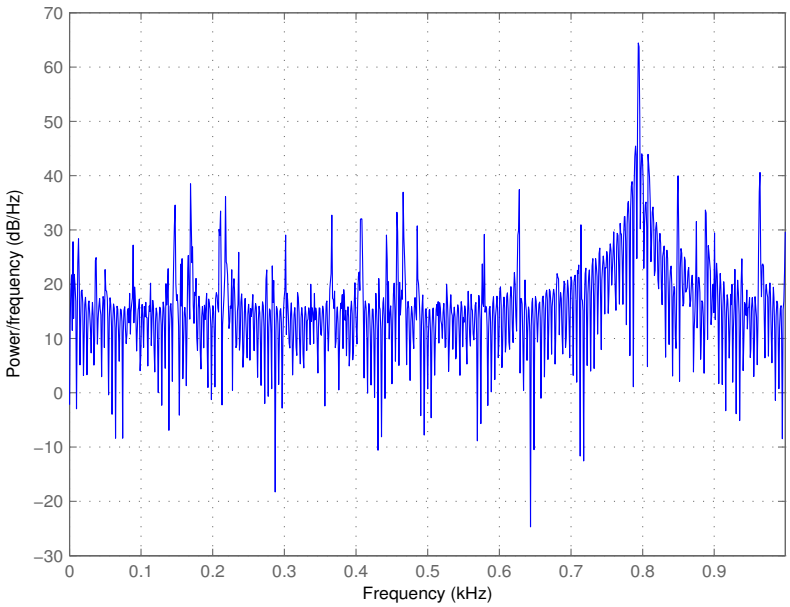

Figure 4.10: Periodogram power density function estimated of the reachability process.

4.2.2 Comparisons with Other Distributions

Since we are sure that our data is not periodic and we do not have a chaos distribution, it is time to compare them with other distributions in order to characterized the reachability process.

Estimated pdf vs Normal Distribution

The comparison with the Normal distribution is shown in Figure 4.11, we use three different values $\lambda = 0.50, 0.75, 1.00$ and as we can appreciate, in spite of when the λ value is increasing the center of the distribution it begins to fit the estimated pdf, the frequency of the gaussian distribution begins to tend to zero.

Therefore, we can say that the Gaussian Distribution is not the best way to characterized the reachability process. This conclusion can be reinforced observing a comparison through the quantile−quantile plot (see Figure 4.12). As we can clearly see the samples clearly are not from the normal distribution.

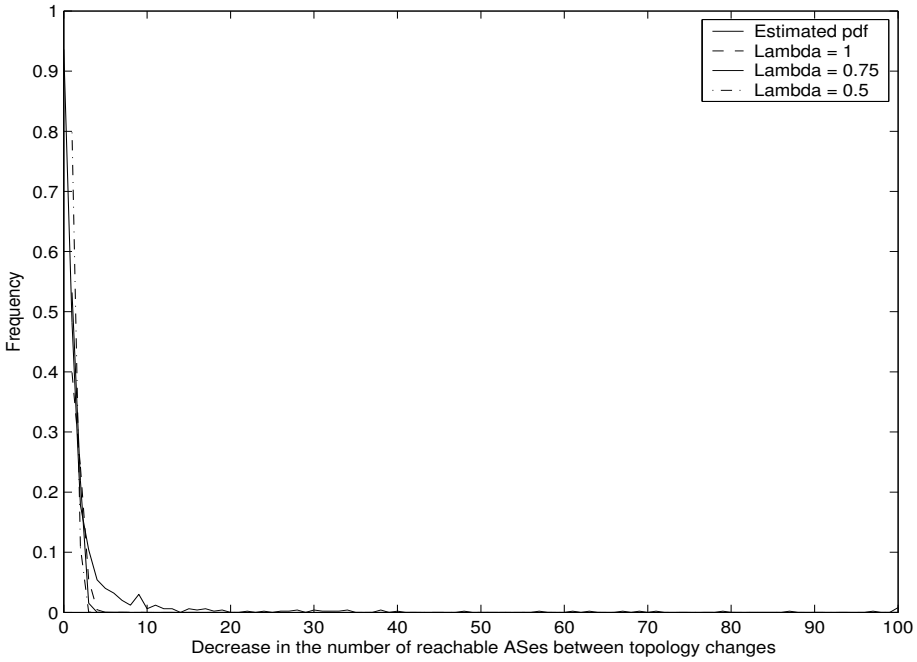

Figure 4.11: Estimated pdf of the reachability process vs the Normal Distribution.

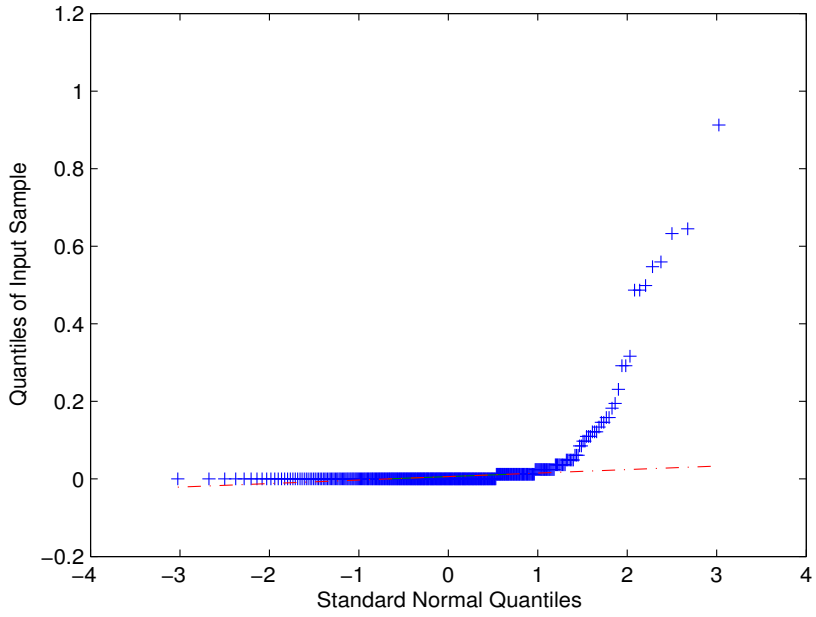

Figure 4.12: QQ Plot of Sample Data versus Standard Normal.

Estimated pdf vs Poisson Distribution

The comparison with the Poisson distribution is shown in Figure 4.13, we use three different values $\lambda = 1.00, 2.00, 3.00$ and as we can appreciate, whenever the λ value is increasing the poisson distribution become away from the estimated data distribution.

Therefore, we can say that the Poisson Distribution is not the best way to characterized the reachability process. This conclusion can be reinforced observing a comparison through the cumulative distribution functions plot (see Figure 4.14). As we can see the samples clearly are not from the poisson distribution.

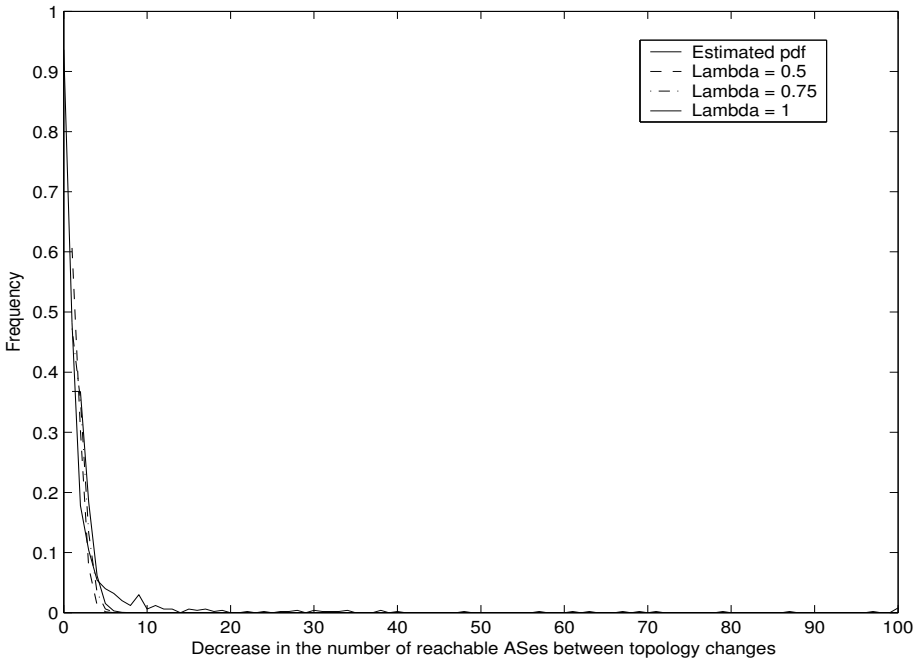

Figure 4.13: Estimated pdf of the reachability process vs the Poisson Distribution.

Estimated pdf vs Exponential Distribution

The comparison with the Exponential distribution is shown in Figure 4.15, we use three different values $\beta = 0.50, 0.75, 1.00$ and as we can appreciate, whenever the β value is increasing the exponential distribution become more likely the estimated distribution.

Due to the seemed, it is more difficult to discard this distribution. Nonetheless, we can see in the comparison probability−probability plot (see Figure 4.16) that the Exponential Distribution is not enough to characterized the reachability process. Now, we can say that the samples clearly are not from the Exponential Distribution.

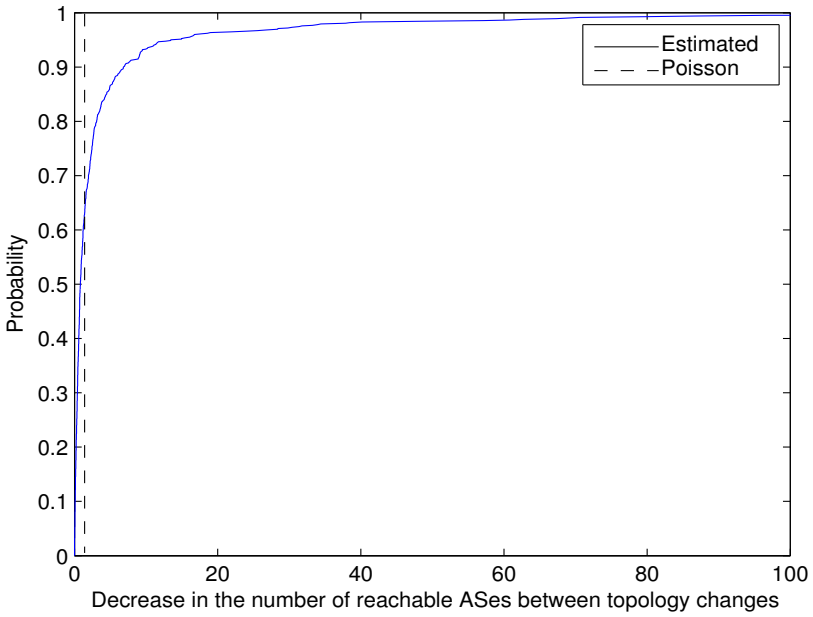

Figure 4.14: Cumulative Comparison of Sample Data vs the Poisson Distribution.

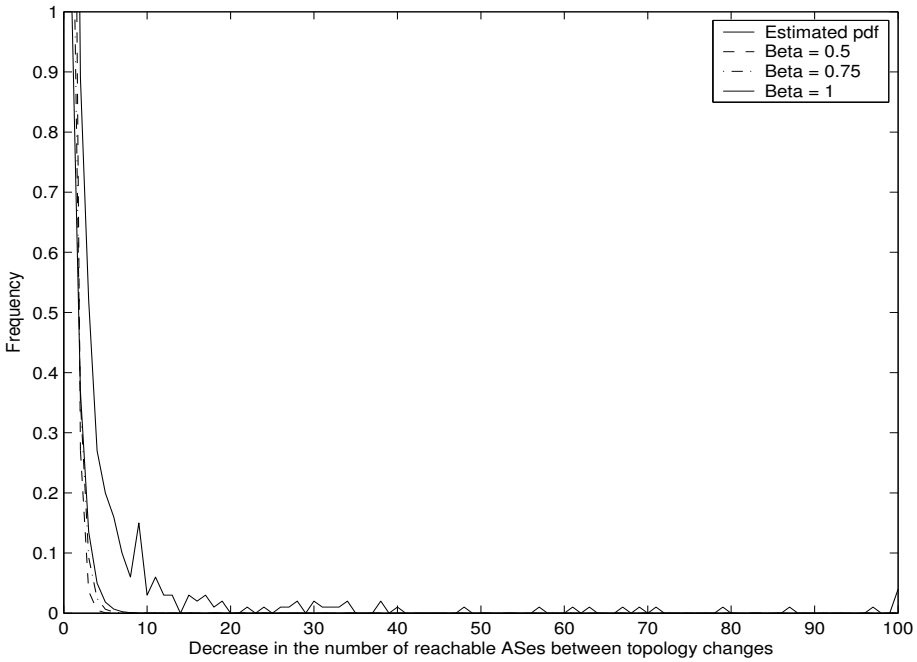

Figure 4.15: Estimated pdf of the reachability process vs the Exponential Distribution.

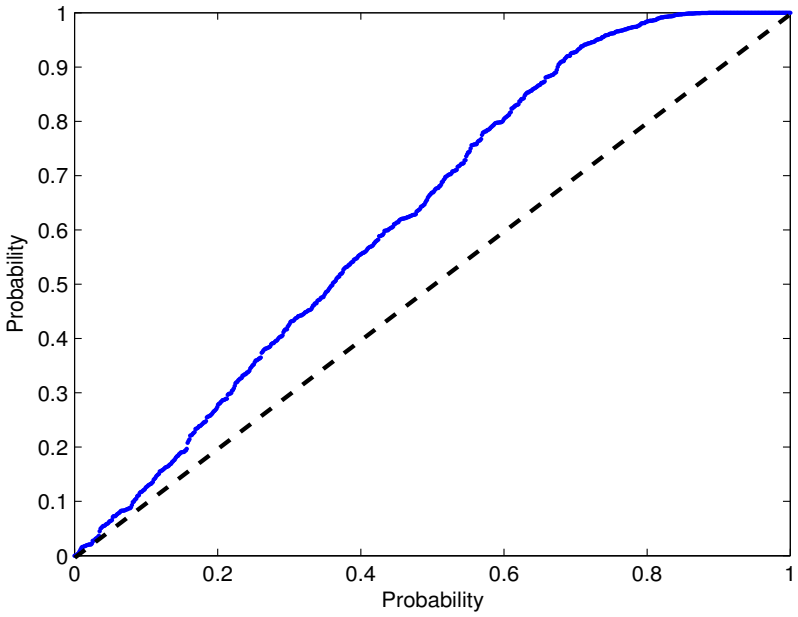

Figure 4.16: Probability - Probability Comparison of Sample Data vs the Exponential Distribution.

Estimated Data vs Power Law Function

The comparison with the Power Law Function is shown in Figure 4.17, we use three different values $\alpha = 0.90, 1.1, 1.3$ and as we can appreciate, the growing of the β value give us a better approximation to the form of the estimated data distribution.

The similarity of the Power Law Function make us very difficult affirm that the form of the estimated distribution is a Power Law function. However, analyzing the data through the cumulative comparison we can clearly observe (see Figure 4.18) that in effect the cumulative distribution of the Power Law function fits very well with the cumulative distribution of the statistical data. Now, we can say that the reachability process can be approximated using a Power Law function.

From the results obtained, we can infer that the probability of reachability in the ASes that belongs to the network can be characterized by the increment of the alpha parameter of the power law function. Besides, if we consider an internet core as ASes always reachable the power law function of the reachability may begin from the number of ASes that formed the backbone of the network. There am the importance of the statistical study of power law functions with two parameters (i.e. Pareto distribution).

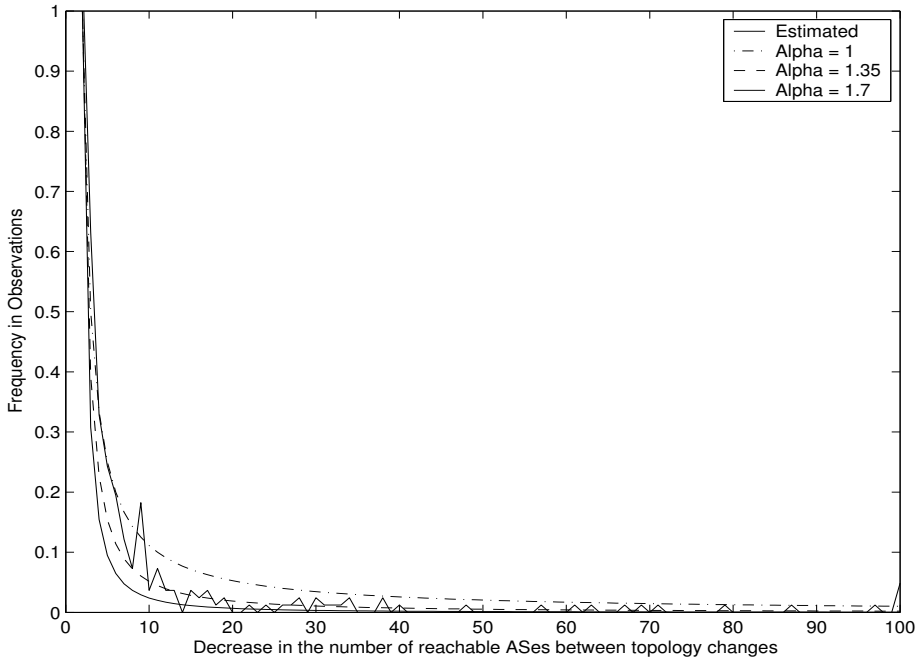

Figure 4.17: Estimated pdf of the reachability process vs the Power Law Function.

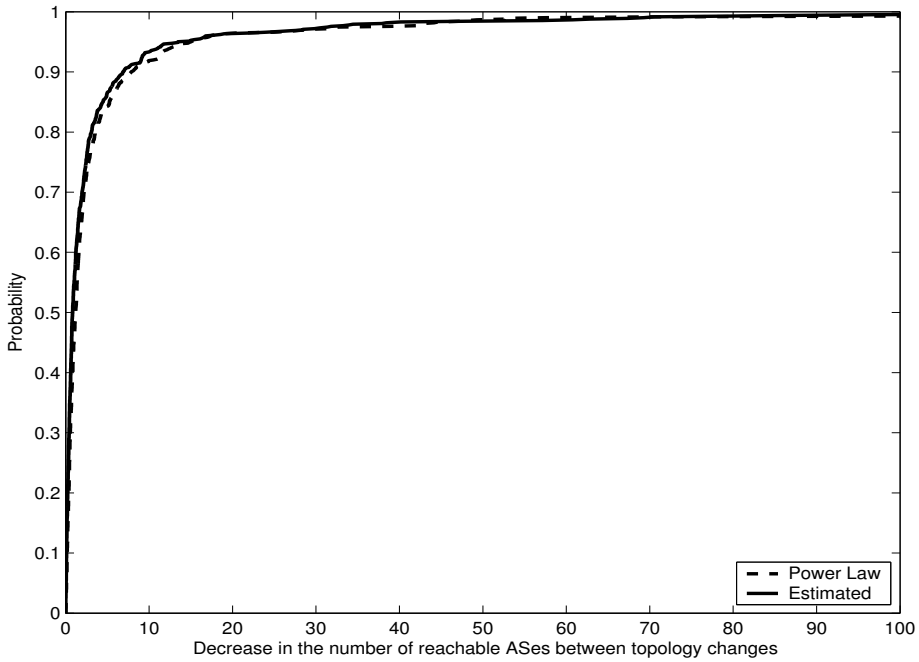

Figure 4.18: Cumulative Comparison of Sample Data vs the Power Law Function.

4.3 Stretch Analysis

The stretch analysis of the collected data, will be divided in three cases related to the probability of each node to become reachable or not reachable:

1. The reachability probability will be set to 0.25 (Lowest Bound).
2. The reachability probability will be set to 0.5 (Middle Bound)
3. The reachability probability will be set to 0.75 (Highest Bound).

This analysis will help us to observe the dynamism of the shortest paths founded through the simulation snapshots of the network topology. The reachability probability will be done in the simulation stage at the first block into the while loop (see Figure 4.19).

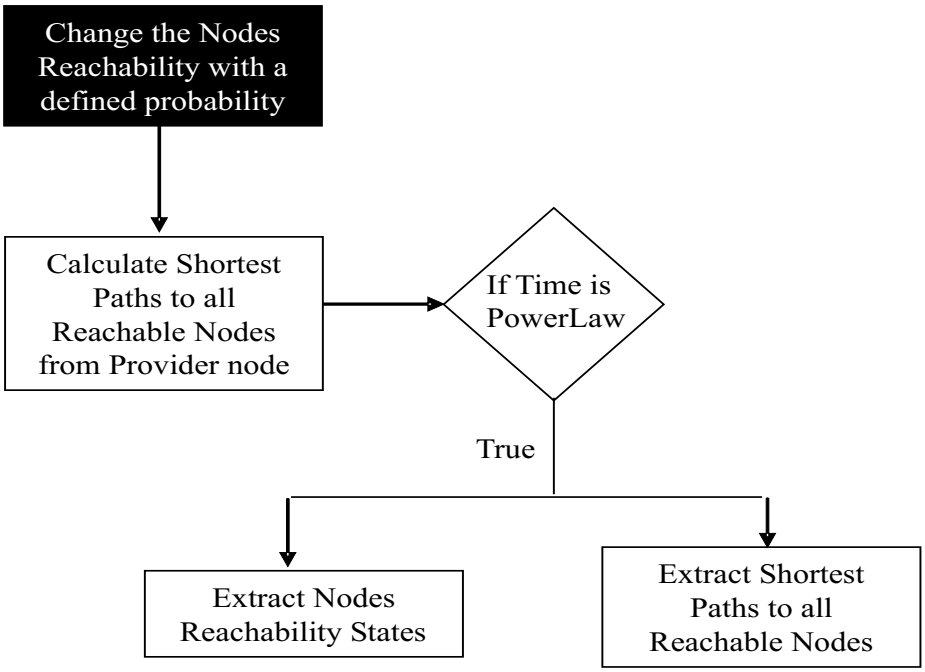

Figure 4.19: The Simulation Stage into the While Loop.

4.3.1 Stretch in the Lowest Bound Case

We notice that the Stretch rate is < 2 and it fluctuates between 1 and 4/3 values as in [15]. We observe that the 70.7% of the paths founded in the simulation are shortest

and the rest of them are only at 1/3 hops away from the optimal shortest paths (see Figure 4.20).

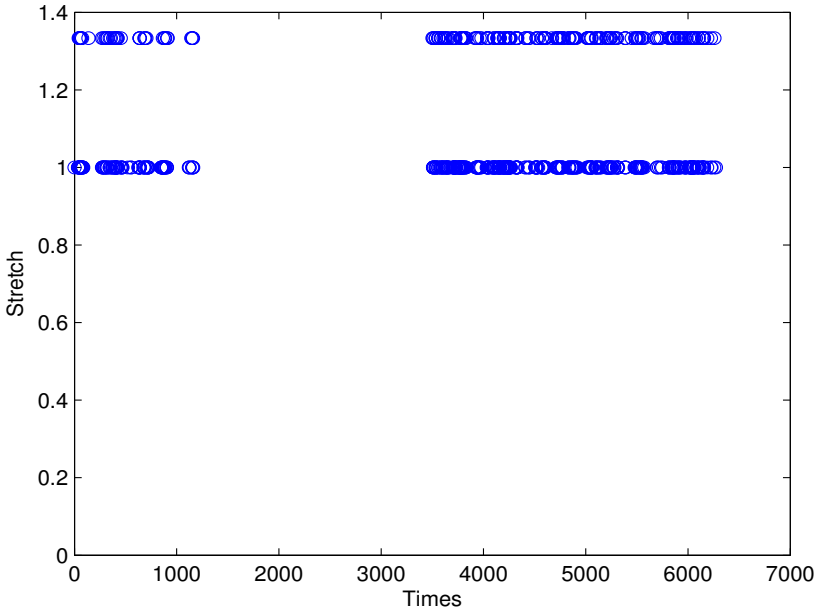

Figure 4.20: The Stretch Rate through Simulation Time (Reachability = 0.25).

4.3.2 Stretch in the Middle Bound Case

As in the Lowest Bound Case, We found that the Stretch rate is also < 2 and it fluctuates between 1 and 4/3 values as in [15]. However, in this case the 75.9% of the paths founded in the simulation are shortest and consequently, the rest are only at 1/3 hops away from the optimal shortest paths (see Figure 4.21).

4.3.3 Stretch in the Highest Bound Case

As in the other two cases, We found that the Stretch rate is again < 2 and it fluctuates between 1 and 4/3 values as in [15]. However, the percentage increase in a 15% more than the Lowest Bound Case, reaching that the 86% of the simulation snapshots the nodes achieve the optimal number of hops according to the optimal shortest path and consequently, the rest are only at 1/3 hops away from it (see Figure 4.22).

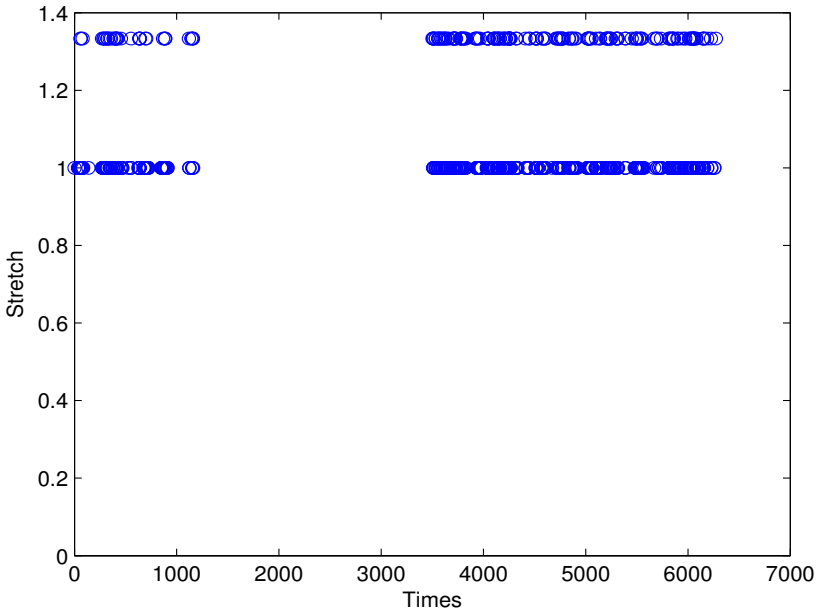

Figure 4.21: The Stretch Rate through Simulation Time (Reachability = 0.50).

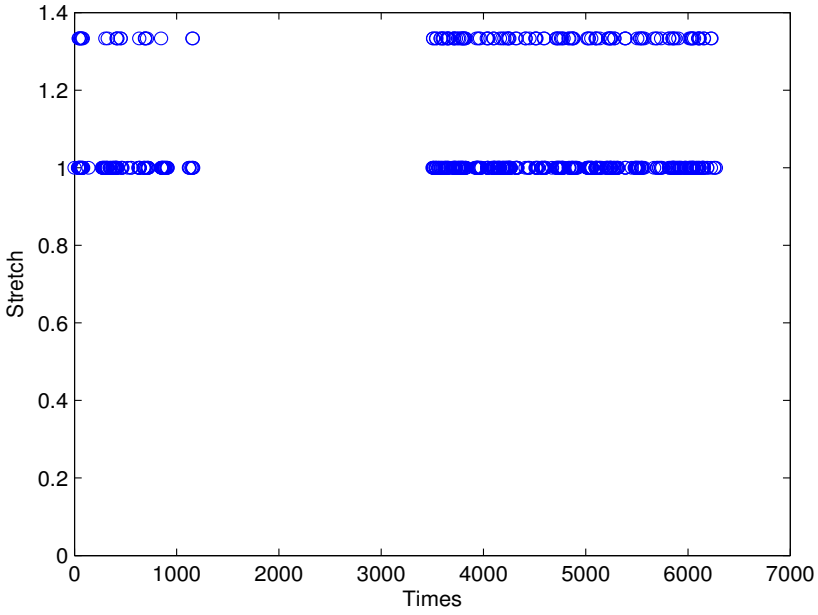

Figure 4.22: The Stretch Rate through Simulation Time (Reachability = 0.75).

4.3.4 Stretch Comparison

As we can to infer, there is a close relationship between the reachability of the nodes that belong to the network topology, and the probability to found the optimal shortest path to a destination (see Figure 4.23).

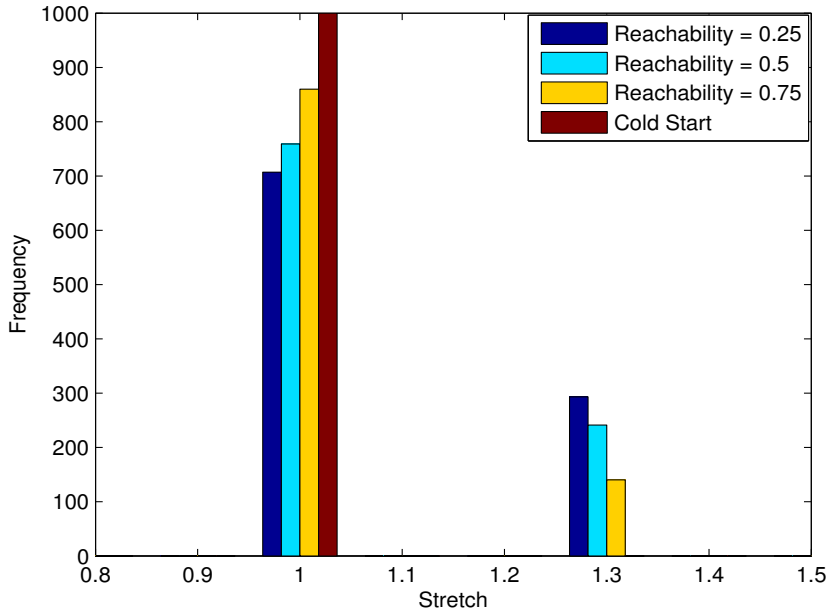

Figure 4.23: The Three Stretch Cases Comparison.

Pierre Fraigniaud and Cyrill Gavoille in [11], proved that for any stretch factor $1 \leq s < 2$ any shortest path routing scheme uses a total of $\Omega(n^2)$ memory bits and locally use at least $\Omega(n)$ memory bits on a router. Due to this, we can say that our routing scheme is a valid approach in Distributed Network, considering that we do not keep a table with all the shortest paths to the destinations, but it is constructed using the CoT table and the Reachability of the nodes and therefore.

Chapter 5

Conclusions and Future Work

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

5.1 Conclusions

In this work we have analyzed routing data from a proposal network topology based in the interdomain customer-provider relationship and power law relationships among the reachability changes of the nodes and the times between topology changes in the interdomain level.

From this analysis, we can conclude the following:

- It is possible to obtain a connected network topology using only the hops at 1 distance from each node to represent an approximation to the interdomain.

- The customer provider interdomain relationship can be introduced in the network topology generated using only the forward links of one node to another without affecting the range of the stretch rate of a universal scheme for distributed networks.

- A routing function to obtain all the paths from an origin to all destinations can be created based only in the node reachability and the knowledge of hops at 1 distance.

- A power law relationship was observed in the frequency of the decrease in the number of reachable nodes between topology changes over a network based in a customer-provider relationship and the times between topology changes in the interdomain level.

- The stretch rate resulting related to the states of the nodes in time shows not only that the network topology is valid, but the routing scheme where it is based is also valid for the representation of a general routing scheme for distributed networks.

- The statistical behavior model of the reachability process based in a customer-provider relationship is consequence of the power law previously observed in the times between topology changes in the interdomain level.

- From the stretch analysis, we can be attempted to conclude that with respect to the router memory requirements there is no a significant changes from the interdomain routing protocol proposed. However, considering that our router scheme only takes as a reference (to find the paths to the destination), the nodes at 1 hop of distance in its circle of trust based in the interdomain customer-provider relationship, we can say that the routing scheme proposal reduces the memory requirements for the routers. Nonetheless, there are other interdomain relationships and interdomain characteristics that must be considered for scalability issues in the routing scheme design.

- The knowledge of the power law behavior in the reachability process based in a customer-provider relationship represents an important issue that can be considered in scalability issues of a interdomain routing scheme design, to prevent anomalies in the network reachability that may assist in the routing decision process.

5.2 Future Work

According to the work presented in this thesis, the following suggestions to continue this line of research are shown as follows:

- It needs to observe an interdomain topology map and collect the reachability information based in the BGP messages, analyzing the data to compare the times between changes of the real data vs the network generator data, in order to validate the routing scheme proposed.

- The analysis of the processing time of the routing function to calculate the paths from the origin to the destinations, it is an important issue to improve the algorithm. It is mandatory compare the processing time of the routing scheme proposed vs the processing time of the BGP routing tables.

- A connectivity analysis of the network topology generated with the customer provider relationships and the nodes at 1 hop is an important issue that will aid to consider other factors that may be affecting the performance of the routing scheme proposed.

- Extend the research of the routing scheme proposed to more than one interdomain relationships, to observe the changes in the processing time, stretch values and reachability behavior.

- Develop a mathematical model of this work that allow a better understanding of the phenomena created by the interdomain relationships to find boundaries and measurements that will lead to the routing scheme proposal to new factors that may be consider in scalability matters.

- Study the relationship between the reachability process and the Internet growing, in order to find measurements able to helps to obtain a better understanding of the Internet behavior in scalability issues.

- In this work an AS in the reachability process is static if we consider that when it becomes reachable always will be connected to the same nodes at 1 hop distance. Therefore, it is important the study of the reachability process considering random connections of an AS that becomes reachable to any node at 1 hop distance from it.

Bibliography

[1] Aiello, W., Chung, F., and Lu, L. "A random graph model for massive graphs". In *Proceedings of the thirty-second annual ACM symposium on Theory of computing,* pp. 171-180, 2000.

[2] Alderson, D., Li, L., Willinger, W., and Doyle, J.C. "Understanding internet topology: Principles, models and validation". In *IEEE/ACM Transaction on Networking,* Vol. 13, No.6., 2005.

[3] Barceló, J.M., Nieto-Hipólito, J.I., and García Vidal, J. "Study of internet autonomous system interconnectivity from bgp routing tables". In *Computer Networks,* Vol. 45, 2004.

[4] Bellman, R. "On a routing problem". In *Quarterly of Applied Mathematics,* Vol. 16, pp. 87-90, 1958.

[5] Bu, T., Gao, L., and Towsley, D. "On characterizing BGP routing table growth". In *Global Telecommunications Conference, GLOBECOMM '02. IEEE,* Vol. 3, pp. 2185-2189, 2002.

[6] Dijkstra, E.W. "A note on two problems in connexion with graphs". In *Numerische Mathematik,* Vol. 1, pp. 269-271, 1959.

[7] Eilam, T., Gavoille, C., and Peleg, D. "Average stretch analysis of compact routing schemes". In *Discrete Applied Mathemtics,* Vol. 155, 2007.

[8] Faloutsos, M., Faloutsos, P., and Faloutsos, C. "On power law relationships of the internet topology". In *SIGCOMM '99: Proceedings*

of the conference on Applications, technologies, architectures, and protocols for computer communication, pp. 251-262, 1999.

[9] Ford, L.R., and Fulkerson, D.R. *Flows in networks,* Princeton University Press, 1962.

[10] Fraigniaud, P., and Gavoille, C. "Optimal interval routing". In *CONPAR 94-VAPP VI: Proceedings of the Third Joint International Conference on Vector and Parallel Processing,* pp. 785-796, 1994.

[11] Fraigniaud, P., and Gavoille, C. "Memory requirement for universal routing schemes". In *Proceedings of the fourteenth annual ACM symposium on Principles of distributed computing*, 1995.

[12] Fuller, V., Li, T., Yu, J., and Varadhan, K. "Classless Inter-Domain Routing (CIDR): an address assignment and aggregation strategy". In *Request For Comments 1519*, 2002.

[13] Gao, L. "On inferring autonomous system relationships in the internet". In *IEEE/ACM Transactions on Networking,* Vol. 9, No. 6, 2001.

[14] Gavoille, C. "Routing in distributed networks: Overview and open problems". In *ACM Press-SIGACT News,* Vol. 32, No. 1, 2001.

[15] Gavoille, C., and Perennes, S. "Memory requirements for routing in distributed networks". In *Proceedings of the 15th Annual ACM Simposium on Principles of Distributed Computing*, 1996.

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

[17] Huston, G. "Interconnection, peering and settlements - part 1". In *The Internet Protocol Journal*, 1999.

[18] Huston, G. "Interconnection, peering and settlements - part 2". In *The Internet Protocol Journal*, 1999.

[19] Kleinrock, L. "Information flow in large communication nets". In *RLE Quarterly Progress Report*, 1961.

[20] Labovitz, G., Malan, G.R., and Jahanian, F. "Origins of internet routing instability". In *Conference Proceeding of the IEEE Computer and Communications Societies.* Vol. I, 1999.

[21] Lee, D.K., Moon, S., Choi, T., and Jeong, T. "Forensic Analysis of Autonomous System Reachability". In *ACM Proceedings,* pp. 335-340, 2006.

[22] Licklider, J.C.R., and Clark, W. "On-line man-computer communication". In *AFIPS Conference Proceedings,* Vol. 21, pp. 113-128, 1962.

[23] Lougheed, K., and Rekhter, Y. "A border gateway protocol". In *Request For Comments 1105*, 1989.

[24] Roberts, L. "Multiple computer networks and intercomputer communication". In *ACM Gatlinburg Conference Proceedings*, 1967.

[25] Roberts, L., and Merrill, T. "Toward a cooperative network of time shared computers". In *AFIPS Conference Proceedings*, 1966.

[26] Siganos, G., Faloutsos, M., Faloutsos, P., and Faloutsos, C. "Power laws and the AS-level internet topology". In *IEEE/ACM Transactions on Networking,* Vol. 11, No. 4, 2003.