Computer Communications 119 (2018) 15–28

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Radiography of internet autonomous systems interconnection in Latin America and the Caribbean

Sofía Silva Berenguer*, Francisco Valera Pintor

University Carlos III of Madrid Av. Universidad, 30, Leganés (Madrid) 28911, Spain

ARTICLE INFO

Keywords:
Interconnection
Autonomous system
Internet topology
Graph
Latin America

ABSTRACT

Lots of studies about the Internet Autonomous System (AS) level topology have been carried out during the last twenty years, most of them analyzing this topology on a world-wide scale, a lot of them based on routing information from the Border Gateway Protocol (BGP). However, studies focusing on a specific region and making comparisons between regions are not that popular and in fact, most world-wide studies are not valid in some particular regions. This work is targeting this particular problem of the regional or country topology analysis by enhancing regular AS-level graphs where to apply different connectivity metrics. The focus is set on Latin America and the Caribbean (the LAC region) which exhibits appropriate conditions for this type of analysis and where we show that a basic metric comparison may not be good enough so as to realize that there is a connectivity problem in the region. After concluding that the situation in the LAC region in terms of interconnection is even worse than expected, we perform some country-level studies finding correlations between graph characteristics and some socioeconomic indicators. We then use these correlations to identify countries in which it would be worth pushing for the deployment of an Internet Exchange Point (IXP), as simulating the creation of an IXP there has a great impact on the interconnection level and on the robustness of the regional Internet.

1. Introduction

The Internet has continuously been studied from every angle and perspective since it was created. Its own evolution in terms of size, traffic patterns, applications, hardware improvements, etc. makes it necessary to keep on performing regular measurements to understand this hugemongous network. The authors of [1] make a good prospective of the different studies performed so far during the last 20 years. As it is explained in this survey, one of the perspectives that has been commonly adopted to understand the Internet connectivity is the Autonomous System (AS) level topology because the performance of the Internet highly depends on the quality of the existing paths between these pieces that together form the Internet jigsaw puzzle. There are other different alternatives but, as it is described in [1], they are more focused on the physical topology than in the logical topology (connectivity) which is the main purpose if this article.

The earliest studies on AS topology are from the late nineties [2,3]. All these studies are normally based on the use of graphs, representing ASes as nodes and the relationships between them as edges and they are usually analyzing the AS-level topology of the Internet on a world-wide scale based on BGP routing information (e.g. [3]). Works involving regional-level analysis (country level or even continental) or comparisons amongst regions are not that popular and in fact, most global Internet topology studies are not valid in some particular regions.

The main reason for this is that in general, this type of global Internet studies tends to use routing information from projects like RIPE NCC's Routing Information Service (RIS) [4] or University of Oregon's RouteViews (RV) [5], which collect and make publicly available BGP routing data from several locations around the world. However, the goal of these projects is not to provide information to infer AS-level topologies, because the BGP protocol was not designed with this purpose in mind and the topologies that can be obtained from this information are usually not complete [6,7]. It is a fact for instance that a number of ASes relationships, in particular peer to peer ones, are not revealed in this topology map because in many occasions they are not announced as the rest of the relationships (at least 35% according to [8]). One of the main drawbacks of the topological studies is that they require a considerable amount of vantage points (collectors) properly located so as to be able to infer a reasonable AS-level topology and this is normally not the case when the focus is set on particular regions where the connectivity between ASes is not completely developed instead of keeping it at a world-wide scope. Latin America, Africa or even Asia are the type of regions whose AS topological maps would probably

* Corresponding author.
E-mail addresses: ssilva@it.uc3m.es (S. Silva Berenguer), fvalera@it.uc3m.es (F. Valera Pintor).

https://doi.org/10.1016/j.comcom.2018.01.008
Received 6 July 2016; Received in revised form 25 June 2017; Accepted 15 January 2018
Available online 31 January 2018
0140-3664/

be incomplete because of the commented reasons (see [9] for instance for a particular regional study about the African region).

In this article we provide evidences to show this effect in the Latin American and Caribbean region (LAC region), although it will be seen that the methodology, the considered metrics and the conclusions can be easily applied to any region.

The LAC region is quite heterogeneous with regards to Internet infrastructure and to the connectivity between networks. It is usually mentioned that the LAC region is a poorly interconnected region with low local traffic exchange [10]. One of the reasons for this is the presence of very few cables in the region due to its geographic characteristics [11], which leads to the dependence of the LAC region on the infrastructure of the operators from North America. And it is also widely known that there are not many collectors in the LAC region, making it even harder to have a realistic view of this region [12].

Because of these characteristics, the LAC region is a good choice to perform our analysis that will initially get a basic graph of the region using common routing information. This graph will be later on enriched using looking glasses in the region or directly requesting data from local Internet Exchange Points (IXPs) or Internet Service Providers (ISPs), as it is typically required for an AS level topological study in order to better approach the real topology and as expected, it will be shown that many peering relationships are now exposed. These hidden peering relationships have been particularly analyzed in [13] for Bolivia and some general conclusions were also obtained for the continent and whole world.

Later on the article focuses on the application of different metrics to the LAC graphs and diverse comparisons with the graphs of the other regions. This type of regional AS topology analysis will lead to a better understanding of the Internet performance in a certain area and will provide criteria to determine which countries of the region have greater needs to enhance their interconnection degree, the deployment status of the Internet in the region and the situation of the Internet infrastructure in the region. Network operators may get the incentives to enhance the interconnection level, thus improving the robustness, the security and the performance of the Internet regionally and even creating new business opportunities by promoting content innovation [14].

In addition, the analysis may be used to derive consistent criteria to make decisions about the creation of IXPs by the governments and organizations involved in these decisions and thus be able to better plan the locations for new IXPs. The creation of new IXPs helps reducing transit costs, improves Internet service quality and promotes infrastructure investment in smaller markets [15].

The rest of the article is organized around two main sections. Section 2 introduces the methodology that has been followed, including the data sources that have been used, the different metrics considered and related work. Section 3 presents all the results and explains their implications. This section is also including an analysis on the implications of adding an IXP in particular countries that are identified applying socioeconomic trends. Section 4 shows the main conclusions of the work.

2. Methodology

The development of this work was divided into five stages: (1) Data collection; (2) Topologies construction; (3) Topology Enhancement; (4) Metrics Computation and (5) Topologies Analysis. This process is shown in Fig. 1. This section describes the different tasks performed during each of these stages, mentioning the methodology usually applied by similar projects and commenting on the differences with the methodology that we used.

2.1. Data collection

In order to complete the stages of this project, various information was collected: routing data, ASes assignment and geolocation information, information about existence of IXPs in the countries in the LAC region, information about presence of ASes in those IXPs and economic, transport and tourism indicators for the countries in the LAC region. All this information was collected between April 2015 and February 2016. In general, as AS-level topologies represent business relationships between ASes and these relationships are rather static, we can assume that the inferred topology has not changed much since the moment we performed the data collection.

There are some studies focused on the temporal evolution of the AS-level topologies of interest [16], and therefore they collect routing data periodically. In this case, we wanted to take a snapshot of the Internet AS-level topology and taking into account that the local routing information for the LAC region required manual work to be done in order to collect it, we did not perform a periodic collection of routing data. We used routing data from RIPE NCC's RIS [4], University of Oregon's RouteViews (RV) [5] and Packet Clearing House (PCH) [17] collectors. Additionally, routing information from Looking Glasses (LGs) in the LAC region (from CABASE[1], PTT Metro,[2] NAP Chile[3] and Orange Chile[4]) and some *show ip bgp* outputs provided by operators in the region (Access Haiti, GTD Internet (Chile) and LACNIC) was also used.

Assignment information about the ASes was obtained from Team Cymru's WHOIS service and from IANA's web page,[5] while geolocation information for ASes (information about the countries where an AS is active) was obtained from the API offered by the RIPEstat project [18].

Information about the existence of IXPs in the different countries of the LAC region and about the presence of ASes in these IXPs was obtained by doing manual research (performing Google searches and visiting IXP's web pages).

Finally, economic, transport and tourism indicators for the countries in the LAC region were downloaded from The World Bank's web page [19] and from the CIA's World Factbook [20]. The most recent information provided by The World Bank at the time of the data collection process corresponded to the year 2014. The information provided by the CIA's World Factbook at the time of the data collection process corresponded to the year 2013.

2.2. Graphs construction

In order to generate a world graph, CAIDA's AS Relationship Inference algorithm [21,22] was executed, using routing data from the different collectors of the RIPE NCC's RIS [4] and University of Oregon's RouteViews (RV) [5] projects as an input to infer relationships between the different ASes. This graph represents the way in which the ASes are interconnected. In this sense, the nodes of the graph represent ASes and the edges of the graph represent the relationships between two ASes. These relationships can be Peer to Peer (P2P) or Provider to Customer (P2C) relationships. This classification is also based on the output from CAIDA's algorithm. For most of the analysis performed as part of this project, no distinction is made between these two types of relationships and therefore, undirected graphs are used.

AS topology studies usually consider a world graph as a whole and do not represent region-level or country-level topologies, therefore no reference of how the world graph should be filtered in order to obtain region-level or country-level graphs was found. That is why we had to design some criteria specifically for this work. Two different approaches were considered:

- **Criterion 1 - Active edges:** only the relationships that are active in the area of interest are included in the graph, i.e. those relationships

[1] http://www.looking.cabase.org.ar.
[2] http://www.ix.br.
[3] http://www.lg.nap.cl.
[4] http://www.pit.orange-business.cl/lg.
[5] http://www.iana.org.