

## Internet Infrastructure Topology Assessment

A. Kajackas, R. Rainys

Telecommunications Engineering Department, Vilnius Gediminas Technical University,

Naugarduko str. 41, LT-03227, phone: +370 5 2744976; e-mails: algimantas.kajackas@el.vgtu.lt, rytis.rainys@elst.vgtu.lt

### Introduction

Since the Internet technological and functional basis allows for unlimited expansion of the network and connecting new nodes and sub-networks, there are lots of Internet service providers (ISP) the number of which is continuously growing. New technological possibilities, new services, growing number of consumers and their generated data flows have essential influence on the development of the Internet, which is also accompanied with the growth of the responsibility of the Internet and the requirements for reliability, vitality and security of data, therefore the modern Internet is a subject for continuous scientific studies.

Analyzing the evolution of the Internet networks in different regions and states, a common trait can be noted – each ISP develops its own network elements. One could say that a regional Internet network consists of individual connected sub-networks therefore it is not known whether the transmission capacities of network nodes and connecting channels would be sufficient in case of faults of individual nodes or channels, incidental increase of data flows or cyber-attacks. It is important emphasize that the infrastructure of the Internet networks in regions forms stochastically, without a systematic design of the regional network and without a comprehensive analysis of data flows and transmission capacities.

The study presents the direction of the performed research, i. e. the analysis and examination of the Internet network of a region or a state, formed of stochastically interconnected sub-networks, as one system. We are searching for the ways to identify the indicators of support of reliability, continuity of functioning and quality of services of such system. In order to reach mentioned goals, the following first tasks are being resolved: drawing a topological scheme of the Internet network and selection of models and tools for network analysis. With this article we are presenting selected methods for Internet network topology assessment and use that to investigate Lithuanian Internet network infrastructure in order to evaluate its vitality.

### Internet network model – multigraph

The infrastructure of the Internet network has been examined in lots of studies [1–3, 5–7]. The most frequently applied model is the one where the Internet is presented as a network, consisting of interconnected autonomous systems (AS) [4]. In a real network the Board Gateway Protocol (BGP), which employs the AS with unique numbers for drawing the dynamic connection tables, is used for communication between individual sub-networks and the AS plays the essential role in the routing of data flows inside the Internet network. Upon taking into consideration the fact that an abstract model of the Internet network, formed of interconnected AS, is a widely spread subject for scientific researches, the model should be based upon when reaching for the goals of the study. The model of the fragment of the Internet network, formed of interconnected AS, is provided in Fig. 1.

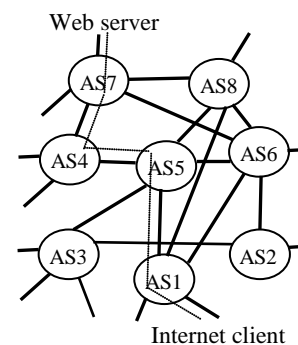


Fig. 1. Internet AS topology fragment

For the purpose of a further analysis each  $AS_i$  is formally described by graph  $G_i(V, U)$ , comprising a set  $V$  of vertices together with a set  $U$  of edges. The vertices correspond to the nodes of a real AS and the edges correspond to the connecting communication lines. According to the network structure the initial data of the examined network and the  $AS_i$  are identified as:  $N_i$  – number of nodes,  $M_i$  – number of edges.

A complex graph can be divided into simpler ones. The most simple AS, consisting of one node, is shown by a graph with one node and vice versa, when joining several AS with the characterizing graphs, a more complex AS and more complex graph are obtained. The infrastructure of the entire network of the region, summarized in such a way may be presented as a multigraph  $G_{net}$  or a totality of graphs  $\{G_i(V, U)\}_{i=1}^l$ . The graphs may be used for performing such actions as removing the node or discontinuation of an edge, which is useful for connectivity analysis. In addition, a graph can be analyzed when examining the shortest, fastest, etc., routes (Fig. 1).

The general model of the Internet as a network, formed from AS of different complexity is convenient, since when analyzing the macro-level characteristics, quite frequently there is no need to examine the full topological schemes of all the AS, one can limit only to certain external AS parameters. As shown by the related studies [6, 7] practice, when examining the reliability of the network, a fragmentary examination of certain Internet AS is possible by distinguishing the territorially limited structural units or by assigning them to a specific organization.

### Principles of assessment

The study practically examines the model of the Internet network, covering all the sub-networks, limited by the IP address range and allocated to the undertakings, operating in the region. Some of ISPs are running through the AS, while smaller ISPs don't, usually they are connected to larger AS by purchasing communication services. Therefore it could be argued that such ISPs do not make significant impact for the region Internet network infrastructure vitality. Taking the aforementioned circumstances and the examples of studies [6, 7] into consideration, the study of vitality of Lithuanian Internet infrastructure is performed mostly on the AS level. AS topology model must be verified by reflecting the existing hierarchic structure of the Internet and highlighting the importance and influence of the individual sub-systems. It is popular Internet hierarchical structure [3], where three distinct types can be identified: Stub, Transit, Multihomed.

A Stub AS (or Customer AS) is only connected to one other AS. For routing purposes, it could be regarded as a simple extension of the other AS. In fact, most networks with a single Internet connection don't have a unique AS number assigned, and their network addresses are treated as part of the parent AS. In our case this type of AS will correspond to small ISPs.

A Transit AS has connections to more than one other AS and allows itself to be used as a transit traffic node between other AS's. Most large ISP's are Transit AS's.

A Multihomed AS has connections to more than one other AS. Multihoming is a technique to increase the reliability for an IP network with redundant Internet connections (for example AS duplication).

We suggest introducing a new National Tier-1 AS type, substantiating such choice by the fact that there are exceptional interconnections with foreign AS within the Internet network of the country. This kind of

communication ensures the availability of the global Internet for the domestic Internet users. When analyzing the real network the AS, interconnected with other countries' networks, are attributed to the National Tier-1 AS group. Usually, such interconnections are related with the so-called Tier-1 level AS (a network that can reach every other network on the Internet without purchasing IP transit or paying settlement [9]), managing the main data flows of the global Internet.

The description of the model of the Internet network will not be complete without identification of the communications between the AS. We distinguish two types of interconnection, important for description of the AS topology: Peering and Transit.

Peering connection type means voluntary interconnection of administratively separate AS networks for the purpose of exchanging traffic between the customers of each network [9]. That is a common way of interconnection however the communications of the peering type have their limitations, since they are intended only for exchange of flow between two specific AS users. The limitation has great influence on the AS topology on the national level.

The communication of the Transit type is dedicated lines, where the receiving AS ensures the transfer of data towards the recipient. The greater part of the data within the Internet network is transferred via the communication of the Transit type.

### AS metrics

The initial analysis of the topological structure of the Internet can be performed by employing the well-known graphs metrics: node degree and node degree distribution [10, 11].

**Node degree.** In graph theory, the degree of a vertex of a graph is the number of edges incident to the vertex. The degree of a vertex  $v$  is denoted  $\deg(v)$ . The degree  $\deg(V_i^k)$  of vertex  $V_i^k$  of the graph  $G_i(V_i^k, U_i^l)$  is the number  $k$  of the communication lines between AS<sub>*i*</sub> and other AS. In the common case  $k=1, \dots, K$ .

When characterizing the nodes, connected with other regions' networks the concept of the degree of the node is divided into two parts: for domestic  $\deg_{in}(V_i^k)$  and international interconnections  $\deg_{out}(V_i^k)$ , i. e. the Tier-1.

We suggest a conception that AS (in our case National Tier-1 AS) managing connections to the upper (we mean international) level networks has responsibility to the regional network respectively AS (in our case Transit AS) managing connections down to the Customer AS is important for the reliability of the regional network.

**Node degree distribution.** It is the probability distribution of nodes degrees over the whole network [9]. The degree distribution  $P(k)$  of a network is then defined to be the fraction of nodes in the network with degree  $k$ . Thus if there are  $n$  nodes in total in a network and  $n_k$  of them have degree  $k$ , we have  $P(k)=n_k/n$ .

When dividing the constituents of the network into the hierarchical structure, there is a need for initial data on each ISP and the graph  $G_i(V, U)$ , characterizing its network, specifying  $N_i$  and  $M_i$  and evaluating the parameters of the

lines, connecting the AS. In addition, it is necessary to evaluate the real rules for arranging interconnections, based by the BGP routing tables.

### Lithuanian region Internet network assessment

Totally, approximately 100 ISPs of different level are registered in the Lithuania. It has been identified that 40 ISPs (of 64 of which data were collected), which performed their activities through created AS. Data related to Lithuanian Internet network AS topology collected directly surveying the ISPs and additionally examined with the European resource of IP (RIPE) database data. Using collected data and above mentioned AS metrics, some relevant results achieved.

Upon classifying the AS, it was identified that most AS within the topology of Lithuanian Internet network belong to the Customer type (81 %), while there are only 12 (19 %) of the Transit type AS (Table 1), which can be explained by the internal economical and network arrangement relating peculiarities of the ISP.

**Table 1.** AS types distribution

AS type	Quantity	Percentage, %
Customer	52	81
Transit	12	19
Multihomed	4	6
Tier-1	11	17

In terms of reliability of the Internet, the small number of Transit AS can result in a situation where damages to AS, for instance, in case of a directed DDoS attack, will bring significant disruption of the operation of region Internet network. The 11 identified AS of the National Tier-1 type are also a limited resource for ensuring the availability of the global Internet to Lithuanian consumers.

The classification of types of communication showed that connections of Peering type between the AS (76 %) prevail (Table 2). However it is necessary to take into consideration that the majority of data communications is performed by employing the Transit type communications, which amount only to 24 % of the network infrastructure. When analyzing in terms of network reliability, it is desirable that the part of communications of the Transit type should become predominant.

**Table 2.** Connections types

Connection type	Quantity	Percentage, %
Peering	373	76
Transit	117	24

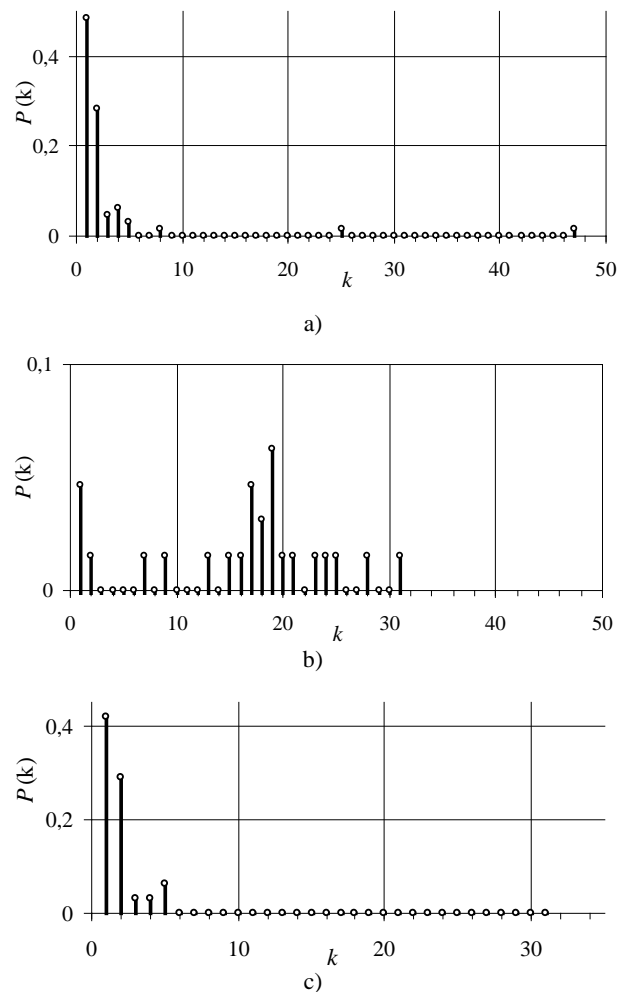
The described hierarchical network structure should be supplemented by evaluating the criteria of importance and significance of the AS, by assigning the top priority to the AS, directly relating with other regions' networks.

**Table 3.** Nodes degree for Transit type connections

AS number	$\text{deg}_{\text{in}}(V_i^k)$	$\text{deg}_{\text{out}}(V_i^k)$
AS1	44	5
AS2	10	2
AS3	25	2
...	...	...
AS <sub>i</sub>	$K_{\text{in}}$	$K_{\text{out}}$

In the AS topology of the regional network the communication with the international AS (of the National Tier-1 type) play an important role, therefore it should be evaluated separately from the internal connections inside the region. In Table 3 we present the principle of distribution of the degrees of the nodes against both inside and outside connections of the Transit type. It is expected that internal connections amount is predominant within the infrastructure. Analogical data can also be presented with regard to the connections of the Peering.

In the Fig. 2 we present the nodes degree distribution within Lithuanian Internet network  $P(k)$  in relation to  $k$ , where  $k$  presents communications lines from AS<sub>i</sub> to other AS ( $k$  starts from  $k=1$  as AS<sub>i</sub> so called "last mile" lines to network subscribers are not included in that survey).



**Fig. 2.** Nodes degree distribution for types of connections: a – transit; b – peering; c – connecting National Tier-1 AS

### Conclusions

The identification of the region Internet AS topology is the necessary approximation towards the evaluation of vitality of the network infrastructure. When describing the hierarchical structure of the Lithuanian Internet network, the Customer type AS identified as most spread (81 %). The classification of types of communications has shown that Peering type connections between the AS prevail in the infrastructure (76 %).

Nodes degree distribution in Fig. 2a shows that 48 % of nodes has only 1 Transit type connection ( $P(1)=0,48$ ) and 28 % has 2 ( $P(2)=0,28$ ). We observed two outstanding nodes with 25 and 47 Transit type connections. It allows for the conclusion that the AS topology of Lithuanian Internet network contains small number of nodes with sufficient connections of the Transit type. We assume that the node with dominant connections amount has great influence on the network topology and has impact to the reliability of the region internal interconnection.

Nodes degree distribution in Fig. 2b is more even as only 4,7 % of nodes has 1 Peering type connection ( $P(1)=0,047$ ). Also it was observed that network has some nodes with large amount of Peering type connections: 4,7 % of nodes has 17 connections ( $P(17)=0,047$ ) and 6,2 % has 19 ( $P(19)=0,062$ ).

Nodes degree distribution for connections to National Tier-1 AS (Fig. 2c) shows that 42 % of nodes has 1 connection National Tier-1 AS ( $P(1)=0,48$ ) and 29 % has 2 ( $P(2)=0,29$ ). In terms of vitality of the network, nodes with National Tier-1 connections could have a strong impact to the international Internet accessibility if it false. We assume that those nodes have responsibility to the rest of the regional network.

When analyzing the network, the attention should be drawn to the fact that AS topologies are dynamic [2], new AS and relating lines between them come into existence, therefore the data, used for characterization of the AS topology should be continuously updated. In the ideal case all the updates should be monitored in real time.

Our final conclusion is that Lithuanian Internet network infrastructure is segmented out of proportion in relation to AS and AS connection types, which means that interconnection nodes of special importance with crucial influence onto the functioning of the Internet network come into existence. Those nodes will be evaluated in our future works and monitoring system for that purpose will

be implemented. In order to evaluate network vitality substantially, criteria should be defined and that will be done in next assessment stage.

## References

1. **Dimitropoulos X., Krioukov D., Riley G.** Revealing the Autonomous System Taxonomy // The Machine Learning Approach, PAM workshop paper source. – 2006.
2. **Chang H., Govindan R., Jamin S., Shenker S. J., Willinger W.** Towards capturing representative AS-level Internet topologie // Elsevier Computer Networks Journal. – 2004. – No. 44(6). – P. 737–755.
3. **Subramanian L., Agarwal S., Rexford J., Katz R. H.** Characterizing the Internet hierarchy from multiple vantage points // IEEE Infocom paper source. – 2002.
4. **Hawkinson J., Bates T.** Guidelines for creation, selection, and registration of an Autonomous System // RFC 1930. – 1996.
5. **Garšva E., Skudutis J.** Secure Computer System Design // Electronics and Electrical Engineering. – Kaunas: Technologija, 2003. – No. 6(55). – P.43–48.
6. **Dohoon K.** Reliability and Risk Analysis of the Internet: Case of Korean Internet Interconnections // IEEE SoSE, 2008.
7. **Dolev D., Sugih J., Osnat M., Shavitt Y.** Internet resiliency to attacks and failures under BGP policy routing // Computer Networks. – 2006. – Vol. 50. – Iss. 16. – P. 3183–3196.
8. **Dimitropoulos X., Krioukov D., Vahdat A., Riley G.** Graph Annotations in Modeling Complex Network Topologies // Tomacs. – 2009. – Vol. 19. – No. 4. – Article 17
9. **Norton W. B.** The Art of Peering: The Peering Playbook [online <http://arneil-py.sacramento.ca.us/ipv6mh/playbook.pdf>]. – 2002.
10. **Albert R., Barabasi A. L.** Statistical mechanics of complex networks // Reviews of Modern Physics 74. – 2002. – P. 47–97.
11. **Brohée S., Faust K., van Helden J.** Network Analysis Tools (NeAT) Tutorial. – RSAT, 2009.

Received 2010 03 06

**A. Kajackas, R. Rainys. Internet Infrastructure Topology Assessment // Electronics and Electrical Engineering. – Kaunas: Technologija, 2010. – № 7(103). – P. 91–94.**

The study presents the performed research on the analysis of the Internet network of a region formed of stochastically interconnected sub-networks, as one system and searching for the ways to identify the indicators of support of reliability, continuity of functioning and quality of services of such system. Different types of AS and its interconnections identified and graphs metrics, node degree and node degree distributions, proposed to be used for the analysis of network topology. Assessment of Lithuanian Internet network identified that regional network infrastructure is segmented out of proportion in relation to AS and AS connection types which means that interconnection nodes of special importance with crucial influence onto the functioning of the Internet network come into existence. Ill. 2, bibl. 11, tabl. 3 (in English; abstracts in English, Russian and Lithuanian).

**A. Каяцкас, Р. Райнис. Исследование топологии инфраструктуры интернета // Электроника и электротехника. – Каунас: Технология, 2010. – № 7(103). – С. 91–94.**

В статье представлены результаты исследований, направленных на поиск путей определения показателей надежности, непрерывности функционирования и качества услуг в региональных сетях интернета, образовавшихся путем стохастического соединения независимых подсетей. Для анализа общих свойств этих сетей предложена определенная классификация различных АС и метрика, позаимствованная из теории графов, – степень узла и распределение степеней узлов. Анализ литовской сети интернета показал, что совокупность АС сегментирована неравномерно по отношению к АС и типам связи. К примеру в сети имеются слабо защищенные АС и узлы с решающим влиянием на функционирование сети. Илл. 2, библи. 11, табл. 3 (на английском языке; рефераты на английском, русском и литовском яз.).

**A. Kajackas, R. Rainys. Interneto infrastruktūros topologijos tyrimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2010. – Nr. 7(103). – P. 91–94.**

Šio darbo tikslai – analizuoti iš stochastiškai sujungtų potinklų susiformavusį valstybės ar regiono interneto tinklą kaip bendrą sistemą ir ieškoti būdų, kaip nustatyti tokios sistemos patikimumo, veiklos nenutrūkstamumo bei paslaugų kokybės palaikymo rodiklius. Analizuojant tinklo topologiją, pasiūlytas AS ir jų sujungimo linijų klasifikavimas bei pasirinkta panaudoti grafų metriką su mazgų laipsniais ir jų pasiskirstymais. Lietuvos interneto infrastruktūros tyrimas atskleidė, kad regiono tinklas, vertinant AS ir jų sujungimus, yra neproporcingai segmentuotas, turi ypač svarbių tarpinklinių sujungimų mazgų, kurių funkcionalumas lemia tinklo savybes. Il. 2, bibl. 11, lent. 3 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).