Impact of IP Addresses Localization on the Internet Dynamics Measurement

Tounwendyam Frédéric Ouédraogo
University of Koudougou
UFR-Sciences et Technique
Avenue M. Yameogo BP 376 KDG, Burkina Faso

Tonguim Ferdinand Guinko
Université de Laval
Facult´e des Sciences et de G´enie
1045 avenue de la M´edecine, Qu´ebec
Canada

Abstract—**Many projects have sought to measure the dynamics of the Internet by using end-to-end measurement tools. The RADAR tool has been designed in this context. It consists in periodically tracing the routes from a monitor toward a set of destinations, IP addresses chosen randomly in the Internet. However, the localization of these destinations on the topology has a significant influence on the observed dynamics.**

We study the dynamics observed when the destinations are localized at a country scale. We show that this localization may lead to observe a different dynamics. The local dynamics observed in our case is mainly a routing dynamics whereas the load-balancing dominates the entire Internet dynamics.

Keywords—*Networks; Internet; Dynamics; Measurement; Localization*

I. INTRODUCTION

One of the main challenges in studying the Internet is the topology dynamics. However, understanding Internet dynamics remains crucial for many applications, including measurement tools, network protocols.

Recent studies have revealed important results on the measurement and characterization of the Internet dynamics [19], [13], [22], [13], [6], [5], [25]. Earlier work on the Internet dynamics has concerned mainly the measurement challenge. One of the important results has introduced the ego-centered views of the Internet dynamics and provides measurement tool and data [16].

The ego-centered view approach to study the Internet dynamics at IP-level topology consists in focusing on what a single monitor can see of the entire Internet dynamics. This manner to measure the Internet topology performs periodically end-to-end measurements from the monitor to the destinations and provides a time series of routing trees. This approach to study the dynamics has allowed obtaining important results on the characterization of the Internet dynamics [17], [19].

The set of destinations is comprised of IP addresses chosen randomly among all those which are on the Internet. The observed dynamics with the ego-centered view approach relies on where the destinations are located on the topology.

This paper addresses the issues of the impact of the destinations localization on the observed dynamics. We analyze the observed dynamics when the destinations are in a restricted area of the topology, precisely at the scale of a country.

We have performed, in the same period two kinds of measurements of the Internet dynamics, the one at country scale and the other at global scale. We found different routing topologies and dynamics. Surprisingly, we observe in the local measurement that the routing topology has more branching on the paths towards the destinations and weak dynamics mainly due to the routing changes, whereas in global measurement, the routing topology is filiform with high dynamics due to the load-balancing. The main contribution of this paper to the research community is to highlight the heterogeneity of the routing dynamics at IP level.

The rest of the paper is organized as follows. The section II presents the measurement framework and the dataset. In section III, we make an analysis comparatively of the local and global routing topologies. Section IV presents the results the investigation on the impact of the destinations localization on the observed Internet dynamics through the local and global measurements.

II. DATASET

Our work relies on the data of Internet dynamics obtained from the RADAR measurement. A RADAR measurement consists in periodical TRACETREE measurements done from a single monitor toward a set of destinations. A TRACETREE measurement is the outcome of parallel end-to-end measurements toward the destinations. Then TRACETREE provides a routing tree where the root is the monitor and the leaves are the destinations. The RADAR measurement represents regular snapshots of the routing topology at IP-level around the monitor.

There are RADAR data measurements performed from about hundred monitors mainly from PLanetLab[1] and publicly available[2]. These measurements are done with parameters to whom the relevance has been shown by these authors [16]. Among these parameters there are 3 000 destinations, 10 minutes between two consecutive TRACETREE measurements, random choice of destinations and monitors. Obviously, other parameters may be also relevant.

We performed two kinds of RADAR measurements, in the same period of time and on the same monitor but with different sets of destinations. The first set of destinations composed of

---

[1]https://www.planet-lab.org/
[2]http://data.complexnetworks.fr/Radar/

$3\,000$ IP addresses chosen randomly among all the IP addresses available on the Internet[3], aims to measure the entire topology dynamics. Next, we will refer to this measurement as the global measurement. The second set of destinations also made of $3\,000$ IP addresses but chosen randomly among the IP addresses assigned to a country. The measurement is restricted to the country scale. We will refer to the measurement with the second set of destinations as the local measurement. We conducted the global and local measurements from several monitors located at different countries and we made similar observations.

In this paper, we present our results from a monitor located in France (Paris). The measurement has lasted five months. We obtained from this monitor $15\,375$ rounds of global measurement and $15\,620$ with the local measurement. Each round measurement lasted around 4 minutes, with 10-minute break between two consecutive rounds.

III. SNAPSHOTS OF THE TOPOLOGY

The ego-centered views approach to measure the Internet allows to map in short time the topology and provides a snapshot of the routing topology around the monitor as routing tree. In this section, we present how the localization of the destinations influences the routing tree.

Figure 1 shows a large difference between local and global IP addresses although they use the same number of destinations. Moreover the number of destinations reached is roughly the same with local and global measurements. These observations allow saying that the local and global topologies around the same monitor are different. But notice that two rounds measurement may have the same number of IP addresses with large difference of their topologies.

The local measurement obtains approximately the same number of IP addresses. It is not the case of the global where we observe more fluctuations. This fluctuation shows some dynamics of the topology with the global measurement that are not observed with the local, see section IV.

Before going further to analyze the difference of their topologies, let us show that the distances of the destinations can explain the difference on the number of IP addresses observed with the local and global measurements. The distance D_{sd} from the monitor s to a given destination d is the path length with the extremities s and d.

$$D_{sd} = \sum_{i=1}^{k}(n_i, n_{i+1}) \text{ where } n_1 = s, \ n_{k+1} = d.$$

We compute the distance average $\overline{D} = \frac{1}{N}\sum_j D_{sd_j}$ of distances D_{sd_j} of destinations d_j.

Figure 2 shows the average distances of the local and global measurements. The global distance is larger than the local distance over time.

The difference of average distances means that the global destinations are farther from the monitor than the local destinations, therefore routes from the monitor to the global destinations gather more IP addresses. The routes towards
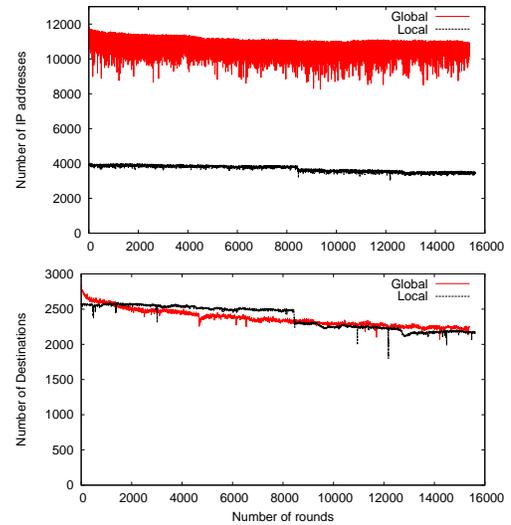

Fig. 1. Top: Evolution of the number of IP addresses seen at each round of local and global measurements. Bottom: Evolution of the number of destinations reached at each round of local and global measurements
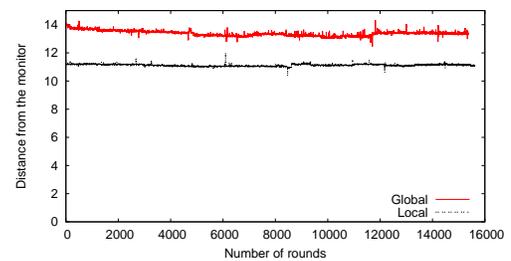

Fig. 2. Average distance of the destinations reached over time.

local destinations in the routing tree are less disjoint and they have a large part of their paths in common. Obviously, most of them are more close to the monitor. A rough estimation of the number of IP addresses induced by the difference of the route lengths (approximately 3), proves the fact that the global measurement observes more IP addresses than the local measurement.

The degree of vertices[4] is an important feature of the topology of trees and more generally complex networks. We use the degree of vertices to define a quantity that characterize the routing topology of the local and global. We denote $T = (V, E)$ a tree, V the set of vertices and E the set of edges. We define a function $\mathcal{F}_T$ to measure the filiform level of the tree T, characterized by vertices of degree 2 *i.e.* vertices having a single successor in the tree. We denote $e_{v_i^2 v_j^2}$ an edge where extremities v_1 and v_2 have degree less or equal to 2,

$$\mathcal{F}_T = \sum_{v_i, v_j \in V} e_{v_i^2 v_j^2}$$

. The more there are many edges $e_{v_i^2 v_j^2}$, the more the tree is filiform.

---

[3]IP addresses that reply to PING request.

[4]A degree of a vertex v is the number of edges connected to v.

Conversely, a tree with many branches (vertex with degree more than 2) is less filiform. We define the number of triples as a quantity that measures the branch-level in the tree. A triple is defined as a couple of edges having a same extremity in the tree. The more there are triples, the more there are branches in the tree.
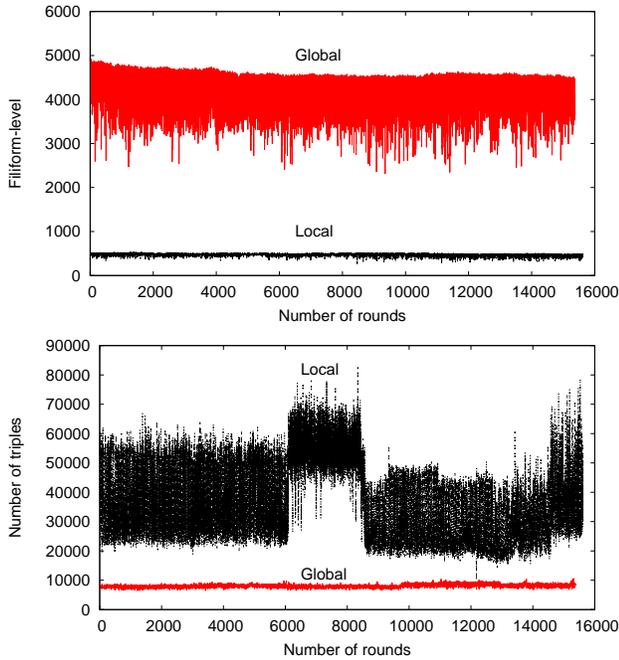

Fig. 3. Top: Filiform $\mathcal{F}_T$ of the routing trees of the local and global over time. Bottom: Number of triples in the routing trees of the local and global over time.

Figure 3 shows the evolution of the two quantities, the filiform function and the number of triples of the local and global routing trees. It seems obvious that when the number of triples in the tree is high, the filiform becomes low because they are opposed. However, we note that the routing tree of global measurement is quite different from the local routing tree. The global routing tree has a high level of filiform and has small number of triples while conversely, the local routing tree has low level of filiform and a high number of triples. We observe a clear fluctuation of the global filiform and not in the local because the scale effect flattened the variations of the local filiform small value. it is the same for the number of triples.

The high number of triples of the local routing tree means that most routers have several paths leading to the destinations and most leaves are near these routers, just few hops. Therefore, the paths towards these destinations have a large part in common that reduces the possibility to observe more IP addresses. This explains why the local measurement gets less IP addresses than global measurement.

The high level of filiform of the global routing tree shows that the routes from the monitor toward the destinations diverge rapidly. Most of the destinations have a large part of the paths isolated.

In summary, the local and global sets of destinations lead to different routing topologies around the monitor. This result is particularly important in the case of ego-centered views approach that is inherent local measurement.

IV. DYNAMICS CHARACTERIZATION

There are two main properties that characterize the observed dynamics at IP-level topology. The first property is the high pace of discovery of the IP addresses. Indeed, new IP addresses appear until the end of the measurement with a more sustained pace than expected [17]. The second property concerns the dynamics pattern of the topology at IP-level. The dynamics of the IP addresses observed around the monitor reveals a parabolic shape. The dynamics pattern relies on the occurrences of the IP addresses observed during measurement [19].

A. Pace of appearance of IP addresses

This property of the dynamics has been observed in many measurements of the Internet topology. *New* IP addresses appear sustainably until the end of measurement. How many of these IP addresses existed on Internet before being observed by the measurement? Investigate on this issue is out of the scope of this paper. In this work we consider as new, all the IP addresses which appear for the first time.
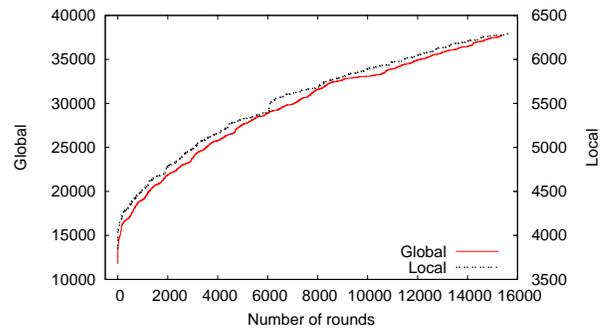

Fig. 4. Evolution of the number of new IP addresses observed during the measurement. Left y-axis: global measurement. Right y-axis: local measurement.

Figure 4 shows the number of new IP addresses of the local and global measurements. We do not observe significant difference between the local and global measurements. In both case new IP addresses are discovered with a sustained pace, except some fluctuations corresponding to rounds of measurement where new IP addresses did not appear. The local measurement observes less IP addresses than the global measurement but proportionally they discover the new IP addresses at the same pace. For example, in the last day of measurement the local measurement sees 10 IP addresses and the global measurement 108 IP addresses.

As the routes of the global destinations are longer than those of the local destinations, they are probably less stable [5], thus allows discovering more IP addresses. The local destinations are more close to the monitor and the dynamics captured is mainly near the monitor. However, the appearance

---

[5]A route in Internet is stable when it does not change over time. The same sequence of IP addresses is discovered between the monitor and destination.

of new IP addresses also remains sustained until the end of the measurement.

There is no significant difference between the local and global measurements regarding the pace of appearance of new IP addresses.

B. Dynamics pattern

The pattern of the dynamics around a single monitor is given by the correlation between the number of occurrences and the number of blocks of consecutive occurrences of the IP addresses. The number of occurrences of an IP address is the number of rounds in which it appears. If it was observed in all rounds, there is one block of consecutive occurrences. Otherwise the number of consecutive occurrences is the number of blocks, as illustrated in Figure 5.
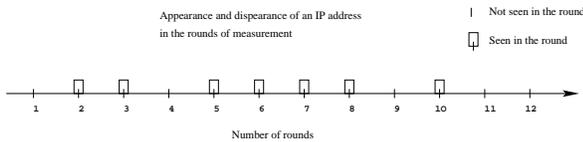

Fig. 5. Representation of the number of occurrence and the number of blocks of an IP address in 12 rounds measurement. The IP address has 7 occurrences (it appears in $2, 3, 5, 6, 7, 8$ and 10) There are 3 blocks of consecutive occurrences. The first block is rounds 2 and 3, the second block is rounds 5, 6, 7 and 8 and the last block is round 10.

Figure 6 shows the pattern produced by the correlation between the numbers of occurrences and the numbers of blocks of the IP addresses of the local and global measurements. This pattern gives a view of the dynamics produced by the appearance and disappearance of the IP addresses.

It appears a parabolic shape which limits for the points inside. The parabolic curve has two tangent lines related to the definition of blocks and occurrences. The first tangent line $(y = x)$ means that the number of blocks of an IP address cannot exceed its occurrences number, see Figure 5. The points closed to this tangent are the IP addresses which tend not to be seen in two consecutive rounds. This blinking trend of IP addresses in the measurement is due to the load-balancing routers [6] [1], [2].

The second tangent line $(y = M - x)$ (M the number of rounds) means that the number of occurrences is less than the total number of rounds. The IP addresses close to this tangent tend to be observed consecutively, without interruption throughout the measurement.

The distribution of points is not uniform. It clearly appears many clusters, in the global measurement, see Figure 6. We note that the most significant is the cluster around the line $y = x/12$ and to a lesser extent around the line $y = x/2$.

As already explained in [19], a load-balancing router spreads the traffic among c paths, each IP address belonging to any of these paths has a probability $p = 1/c$ of being seen at each round, leading to the number of occurrences equal to np approximately.

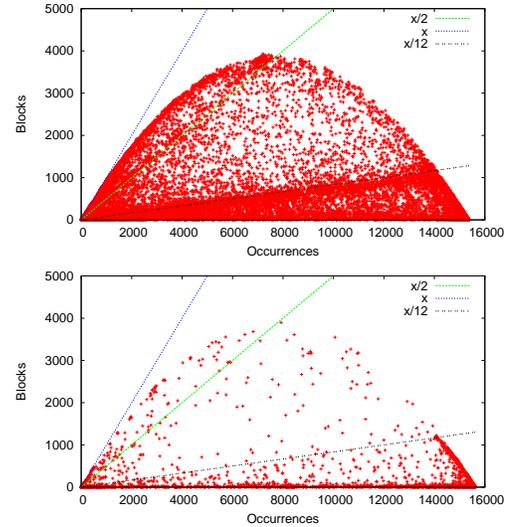[6]A load-balancing router spreads the traffic among several paths, per-packet , per-destination

Fig. 6. Dynamics pattern. Each point represents an IP address, obtained by the number of occurrences on the x-axis and the number of blocks on the y-axis. Top: global measurement. Bottom: local measurement.

An IP address belonging to load-balanced paths can have any probability p of being seen. Therefore, the cluster around the line $y = x/12$ corresponds to IP addresses having a probability $p = 1/12$ of being seen. These IP addresses are observed after load-balancing routers that spread the traffic between 12 paths. Similarly, the cluster around the line $y = x/2$ to IP addresses observed after load-balancing routers that spread the traffic between 2 paths.

These clusters do not appear in the local dynamics, see Figure 6 (Bottom). This means that the local measurement has less load-balancing routers than the global measurement.

We are resuming the explanations of the parabola of these authors [19]. Considering a given round as the first of a consecutive blocks of occurrences of an IP address, with the probability p that this IP address was seen in this round, multiplied by the probability $1 - p$ that it was not seen in the previous round. Multiplying these two probability by n, gives the expected number of blocks $np(1 - p)$, which is the equation of parabola. The explanation is simple and formal. But in some cases, an IP address may be on the path of several load-balancing routers.

The IP addresses at the end of the parabola tend to be seen in consecutive rounds. This kind of dynamics is not due to the load-balancing but to the routing changes. The parabolic shape does not appear clearly in the local dynamics, except the beginning and the end of the parabola. There are many points close to the x-axis. These points correspond to the IP addresses observed in consecutive rounds.

The dynamics observed at IP-level shows different patterns with regard to the destinations localization. In our case, the localization of the destinations at a country scale shows that the load-balancing effect is not observed around the monitor. The local dynamics is mainly due to the routing changes.

V. RELATED WORK

The Internet dynamics measurement passes by an efficient and fast method of mapping the topology. Many works have been done to provide methods to map the Internet topology [23], [18], [3], [9], [20], [14]. Most of these methods rely on the end-to-end measurement and the pertinence of some of them has been studied. For instance in [12], [21] the authors study the relevance of vast exploration distributed of the Internet topology. Similar work [15] has addressed the relevance of the properties of Internet and other complex networks obtained by these methods of measurement.

Measuring the dynamics of Internet require fast mapping tool of the topology. Over years, important improvement have been made to get fast and an efficient dynamics measurement tool [16], [8], [20], [6], [5].

Among the early contributions, there is [16]. Their authors proposed a tool RADAR to measure efficiently around single monitor. Recent work [6], [5] studies the dynamics of end-to-end path in the Internet considering the presence of load-balancing effect on the measurement and proposed a new tool DTRACK able to measure and anticipate the path dynamics.

A relevant choice of source, destination (vantage points) is important to efficent end-to-end measurement [10], [24], [4]. Particularly, this work [10] addressed the problem of better selection of destinations to the Internet measurement and they propose an automatic generation of the hit list, an efficient set of destinations.

Our work addresses the relevance of destinations to the observed Internet dynamics. Recent contribution on the Internet dynamics has shown some characteristics. The Internet dynamics is faster than unexpected and the occurrences of IP addresses around the monitor fit to a pattern [19], [17]. On the same stream of this study, we focus on the influence of the destinations localization on the characteristics of the IP-level dynamics observed around a monitor.

VI. CONCLUSION

In order to study the Internet dynamics, previous work proposed an approach to measure the dynamics from a single monitor. The monitor performs TRACEROUTE-like measurements towards a set of destinations. In the end-to-end measurement, the destinations are an important parameter. In this paper, we showed the influence of the destinations localization on the observed dynamics.

We performed two kinds of measurement, local and global and showed that their routing trees are different. We also show how this difference on the routing topologies may influence the observed dynamics through the two main characteristics of the dynamics at IP-level topology: the sustained discovery of IP addresses and the parabolic shape of the dynamics pattern.

We found that the number of new discoveries of IP addresses remains sustainable whatever with local or global, but it is not the same with the dynamics pattern. The load-balancing is the main dynamics in the global measurement whereas the routing changes are the most observed in the local measurement. The load-balancing is the main dynamics in the global measurement whereas the routing changes are the most observed dynamics in the local measurement.

Many works remain to be done on the dynamics characterization and modeling. Precisely, formally identified the different types of dynamics (load-balancing and routing dynamics) is crucial for several applications. For instance, to design efficient measurement tool. Relying on the dynamics characteristics established, it will be possible to propose more realistic models of the Internet dynamics.

REFERENCES

[1] B. Augustin, X. Cuvellier, F. Orgogozo, Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira. Traceroute anomalies : Detection and prevention in internet graphs. *Computer Networks*, 52:998–1018, 2008.

[2] B. Augustin, T. Friedman, and R. Teixeira. Multipath Tracing with Paris Traceroute. In *Proc. Workshop on End-to-End Monitoring, E2EMON*, May 2007.

[3] Robert Beverly, Arthur Berger, and Geoffrey G. Xie. Primitives for active internet topology mapping: Toward high-frequency characterization. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, pages 165–171, New York, NY, USA, 2010. ACM.

[4] Mingming Chen, Meng Xu, and Ke Xu. A delay-guiding source selection method in network topology discovery. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1–6, June 2011.

[5] I. Cunha, R. Teixeira, and C. Diot. Measuring and characterizing end-to-end route dynamics in the presence of load balancing. In *PAM'11 Proceedings of the 12th international conference on Passive and active measurement*, pages 235–244, 2011.

[6] I. Cunha, R. Teixeira, D. Veitch, and C. Diot. Dtrack: A system to predict and track internet path changes. *Networking, IEEE/ACM Transactions on*, 22(4):1025–1038, Aug 2014.

[7] Radar data. http://data.complexnetworks.fr/Radar/.

[8] B. Donnet, P. Raoult, T. Friedman, and M. Crovella. Efficient algorithms for large-scale topology discovery. In Derek L. Eager, Carey L. Williamson, Sem C. Borst, and John C. S. Lui, editors, *Proceedings of the International Conference on Measurements and Modeling of Computer Systems, SIGMETRICS 2005, June 6-10, 2005, Banff, Alberta, Canada*, pages 327–338. ACM, 2005.

[9] B. Eriksson, G. Dasarathy, P. Barford, and R. Nowak. Efficient network tomography for internet topology discovery. *Networking, IEEE/ACM Transactions on*, 20(3):931–943, June 2012.

[10] Xun Fan and John Heidemann. Selecting representative ip addresses for internet topology studies. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, pages 411–423, New York, NY, USA, 2010. ACM.

[11] Global and local data. http://www.univ-koudougou.bf/Radarbf/.

[12] Jean-Loup Guillaume, Matthieu Latapy, and Damien Magoni. Relevance of massively distributed explorations of the internet topology: Qualitative results. *Computer Networks*, 50.

[13] Hamed Haddadi, Steve Uhlig, Andrew W. Moore, Richard Mortier, and Miguel Rio. Modeling internet topology dynamics. *Computer Communication Review*, 38(2):65–68, 2008.

[14] B. Holbert, S. Tati, S. Silvestri, T. La Porta, and A. Swami. Network topology inference with partial path information. In *Computing, Networking and Communications (ICNC), 2015 International Conference on*, pages 796–802, Feb 2015.

[15] M. Latapy and C. Magnien. Complex network measurements: Estimating the relevance of observed properties. In *INFOCOM 2008. 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 13-18 April 2008, Phoenix, AZ, USA*, pages 1660–1668. IEEE, 2008.

[16] M. Latapy, C. Magnien, and F. Ouédraogo. A radar for the internet. In *Workshops Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), December 15-19, 2008, Pisa, Italy*, pages 901–908. IEEE Computer Society, 2008.

[17] C. Magnien, F. Ouedraogo, G. Valadon, and M. Latapy. Fast dynamics in internet topology: Observations and first explanations. In *Proceedings of the 2009 Fourth International Conference on Internet Monitoring and Protection*, ICIMP '09, pages 137–142, Washington, DC, USA, 2009. IEEE Computer Society.

[18] P. Marchetta and A. Pescape. Drago: Detecting, quantifying and locating hidden routers in traceroute ip paths. In *Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on*, pages 109–114, April 2013.

[19] A. Medem, C. Magnien, and F. Tarissan. Impact of power-law topology on ip-level routing dynamics: Simulation results. In *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*, pages 220–225, March 2012.

[20] Jian Ni, H. Xie, S. Tatikonda, and Y.R. Yang. Efficient and dynamic routing topology inference from end-to-end measurements. *Networking, IEEE/ACM Transactions on*, 18(1):123–135, Feb 2010.

[21] F. Ouédraogo and C. Magnien. Impact of sources and destinations on the observed properties of the internet topology. *Computer Communications*, 34(5):670–679, 2011.

[22] J.-J. Pansiot. Local and dynamic analysis of internet multicast router topology. *Annales Des Tlcommunications*, 62(3-4):408–425, 2007.

[23] Y. Shavitt and E. Shir. Dimes: Let the internet measure itself. *SIGCOMM Comput. Commun. Rev.*, 35(5):71–74, October 2005.

[24] Y. Shavitt and U. Weinsberg. Quantifying the importance of vantage points distribution in internet topology measurements. In *INFOCOM 2009, IEEE*, pages 792–800, April 2009.

[25] Feng Ying and Zhao Hai. Analysis of the hierarchical characteristics of ip-level topology dynamic node. In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2014 Sixth International Conference on*, volume 2, pages 338–341, Aug 2014.