# Impact of Small-World Effect on the IP-level Routing Dynamics

Frédéric Tounwendyam Ouédraogo[1(✉)], Tegawendé Bissyandé[2], Sawadogo Daouda[3], Didier Bassolé[2], Abdoulaye Séré[4], and Oumarou Sié[2]

[1] Université de Koudougou, BP 371, av. M. Yameogo, Kdg, Koudougou, Burkina Faso
ouedraogo.tounwendyam@yahoo.fr

[2] Université de Ouagadougou, BP 7021, av. C.D.Gaulle, Ouaga, Burkina Faso
tegawende.bissyande@fasolabs.org, dbassole@gmail.com, oumarou.sie@gmail.com

[3] Univeristé de La Rochelle, 23 av. A. Einstein, 17000 La Rochelle, France
daoudi5@gmail.com

[4] Université de Polytechnique de Bobo-Dioulasso, BP 1091, Bobo, Burkina Faso
abdoulayesere@gmail.com

**Abstract.** Running periodically TRACEROUTE-like measurements at suite frequency from a given monitor towards a fixed set of destinations allows observing a dynamics of routing topology around the monitor. This observed dynamics has revealed two main characteristics: the topology evolves at a pace much higher than expected and the occurrence of observed IP addresses provides a pattern of the IP-level routing dynamics. In this paper, we aim to provide some explanation of these characteristics through the small-world effect, observed on most complex networks. We are able to reproduce the observed dynamics by modeling the measurement on small-world graph. Thus, we show by simulation the influence of the coefficient clustering and the average path lengths on the dynamics.

**Keywords:** Internet · Dynamics · Modeling · Topology · Characterization

## 1 Introduction

Internet is world scale system that evolves over time. Some nodes and links appear and disappear constantly on the measurement. This dynamics is due to the routing, the load-balancing, physical dynamics, or some events like network failure. Understanding this dynamics is important for many applications. It remains a challenge efficient tool to map the Internet. An efficient mapping tool passes by take account the Internet dynamics features. Network protocols development and validation also require a good knowledge of underlying topology. Some applications need to be tested on a model before real deploying.


R. Glitho et al. (Eds.): AFRICOMM 2015, LNICST 171, pp. 26–35, 2016.
DOI: 10.1007/978-3-319-43696-8_3

Many works have been done in order to provide the most efficient and fast tool to map the Internet topology [4,6,8,10]. In this stream of studies, the TRACETREE tool has been proposed to measure the Internet dynamics. This tool performs an ego-centered view measurement periodically from a single monitor towards a set of destinations and provides a series of routing trees where the leaves are the destinations and the root is the monitor. The contribution of this work has brought a little more knowledge on the Internet dynamics measurement and characterization. The analysis of this dynamics has shown two properties that characterize the observed dynamics at IP-level topology [9,11].

Understanding these dynamics behaviors requires sufficient knowledge of the topology. We turn to the simulation to investigate the network properties that cause these dynamics behaviors. In the same goal, previous work has used power-law graph to modeling the IP-level topology of the Internet [11]. Our contribution goes further and studies the observed dynamics on small-world graph. The small-world effect is the fact that most pairs of vertices of the graph are connected by a short path. It may have implications for the network dynamics. For instance, the number of "hops" a packet must take to get from on computer to another on the Internet. The small-world graph has high coefficient clustering and short average path lengths. We address the issue of how to reproduce the observed dynamics on small-world graph. We find the appropriate setting by varying different parameter values.

We show that the small-world effect has a correlation with the observed dynamics of the Internet. This result represents an important step toward Internet dynamics characterization that lead to many applications, including realistic model designing, network routing protocols improvement regarding to some failure, especially for developing countries where the selective power cut makes often the Internet unreachable.

The rest of the paper is organized as follows. Section 2 presents the two properties observed on the dynamics of the IP-level routing topology. Section 3 presents the small-world graph model and shows how we simulate the dynamics on the model, the TRACETREE measurement and the topology evolution. We discuss the simulation results in Sect. 4. Section 5 surveys the related work. Section 6 ends the paper by the conclusion and future work.

## 2 Routing Dynamics Characteristics

Previous work has presented the TRACETREE tool [8] that collects the ego-centered view from a single monitor to a given set of destinations (chosen randomly in the Internet) by measuring the routes from this monitor to each destination. This view of the topology provides a routing tree, in which nodes are IP addresses, and a link exists between two nodes if they are connected.

Performing periodically TRACETREE measurement allows capturing the dynamics of ego-centered views of the routing topology. Many datasets were collected in this way from more than hundred monitors located at different places around the world: Burkina Faso, France, Japon, United States of America and mainly host provided by PlanetLab [5]. Each monitor performs TRACETREE measurement towards a set of 3 000 destinations during one month with a frequency of around 15 min at every round. These datasets are publicly available [5]. The analysis of these datasets revealed two main characteristics of the observed dynamics around a given monitor.

### 2.1 Sustained Discovery of IP Addresses

The number of IP addresses observed at each round measurement is roughly the same, as shown in Fig. 1. Note that this number may be different with other monitors. There are some downward peaks which indicate rounds with *less* IP addresses than usual. These peaks could indicate an event such as a major routing change or failure.[1]

Figure 2 shows the number of IP addresses observed since the beginning of one month measurement.

The unexpected behavior is the pace of the appearance of new IP addresses during the measurement. For measurements that lasted several month, new IP addresses still appears sustainably until the end. This characteristic of the Internet dynamics observed at IP-level topology has been presented in this work [9].

### 2.2 Parabolic Shape of the Dynamics Pattern

The pattern of occurrence of IP addresses follows a parabolic shape. The occurrence of IP addresses around a monitor may be defined by two quantities the number of occurrences and the number of block. The number of occurrences of an IP address represents the total of distinct rounds in which it appears. The number of blocks of an IP address is the number of groups of consecutive rounds in which it is observed. As an example, an IP address which was observed on rounds $1, 2, 3, 5, 6, 8, 9$ has 7 occurrences and 3 blocks.

Figure 3 presents the correlation between these two quantities for a monitor. The plot exposes a clear parabolic shape, with a large number of points close to

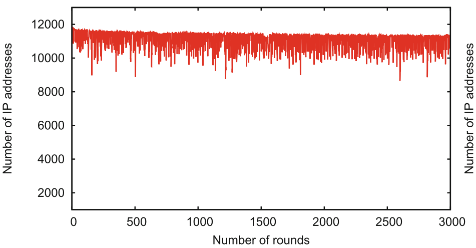


**Fig. 1.** Evolution of the number of IP addresses observed at each round measurement.

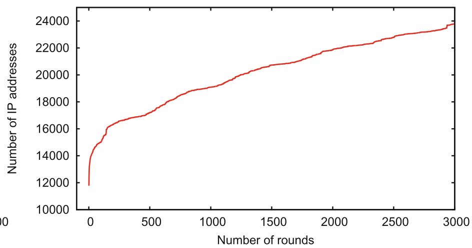


**Fig. 2.** Evolution of the number of IP addresses observed in the union of rounds.

[1] Studying these events is however out of the scope of this paper.

the $x$-axis and to the line $y = x/2$. The presence a large number of IP addresses close to the arc is due to the load-balancing routers. If a load-balancing router randomly spreads traffic among $e$ paths, IP addresses belonging to any of these paths has a probability $p = 1/e$ of being observed at each round, leading to number of occurrences equal to $rp$ approximately.

A given round is then the first of a consecutive blocks of occurrences for one of these routers with the probability $p$ that this IP address was observed in this round, multiplied by the probability $1-p$ that it was not observed in the previous round. Multiplying this probability by $r$ gives the expected number of blocks, which is then equal to $rp(1-p)$ and is the equation of the parabola. In real case where an IP address may belong to paths used by several router performing load balancing. Therefore, an IP address belonging to paths with several load-balancing routers can have any probability $p$ of being observed. This behavior of the dynamics has been presented in previous work by the authors of [11].

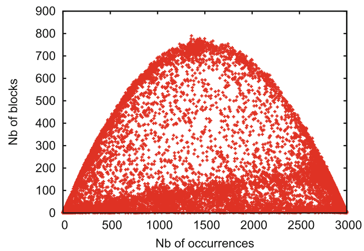


**Fig. 3.** Occurrences of IP addresses. Each point represent an IP address obtained by its number of occurrences on the $x$-axis and its number of blocks on the $y$-axis.

## 3 Modeling

The simulation model consists to reproduce the routing topology and dynamics (include routing change and load-balancing) and TRACETREE measurement on a given monitor to a set of destinations. Let us note that the model goal is explain the characteristics observed in previous work through the small-world networks properties.

### 3.1 IP-level Topology

We represent the Internet topology at IP-level by the undirected and connected graph $G = (V, E)$ where the set of vertices $V$ represents IP addresses and each edge in the set $E$ represents the links between IP addresses. The edges are not weighted.

### 3.2 TRACETREE Measurement

In real measurement the destinations are chosen randomly, and similarly the monitor location in the network. We made the same with the simulation. We randomly chose 3 000 destinations and the monitor among the set of vertices $V$. From the monitor, we perform a breadth-first search (BFS) on the graph $G$ and obtain a tree. Afterwards, we remove recursively all the leaves of the tree which are not destinations. At the end, the leaves of the remaining tree are destinations, and the root represents the monitor.

### 3.3 Dynamics Modeling

We distinguish two dynamics in the model. The load-balancing and the routing change. We simulate the load-balancing when performing the BFS. Each vertex chooses at random the next vertex on a shortest path to the destination, therefore two BFS on the same graph $G$ lead to different trees. The routing change corresponds to a modification of the topology new edges between vertices. We suppose that the number of news IP addresses is insignificant. We only consider the dynamics of edges between vertices in the model by swapping edges. Let $(a, b)$ and $(u, v)$ two edges of $G$ chosen randomly. The swap consists to replace $(a, b)$ and $(u, v)$ by $(a, u)$ and $(b, v)$ in the graph $G$. We simulate two consecutive rounds TRACETREE measurement by performing a fixed number of swaps between two consecutive BFS.

### 3.4 Small-World Properties

The small-world network is mainly characterized by two quantities: a short average path lengths[2] and a high clustering coefficient[3] [1].

We used the Watts-Strogatz model to generate a small-world graph. Given the number of vertices $n$ and the mean degree $D$ (assumed $> \ln(\mathrm{n})$), the model constructs the graph in two steps:

- construct a ring network in which each node is connected to the same number $D$ nearest neighbors, $D/2$ on each side.
- perform a rewiring on every edge with a probability $p$, $0 \leq p \leq 1$.

These two quantities decrease when the rewiring probability $p$ increases but the path lengths decrease more quickly that the clustering coefficient. These properties of the ring network allow having an small-world network for small value of $p$, until some threshold. When $p$ reaches the maximum value 1, all edges are rewired and the ring network becomes a random graph having a small clustering coefficient and a short average path lengths.

[2] If we denote $d(x, y)$, the distance (or shortest path) between the vertices $x$ and $y$, the mean of distances of the vertex $x$ to the other vertices of $G$ is its average path lengths. The average path lengths of $G$ is the mean of average path lengths of all vertices.

[3] Given a vertex $x$, the clustering coefficient is a measure of the probability to which two vertices connected to $x$ tend to be connected.

## 4 Simulation

In this section we aim to find if it is possible, the appropriate settings of the small-world graph that will make possible to reproduce the dynamics with the characteristics presented in Sect. 2.

We are going further to investigate on the relation between the clustering and the average path with respect to the observed dynamics.

### 4.1 Appropriate Settings

In order to address the question of how to reproduce on small-world graph the observed dynamics. We perform simulations varying the values of the parameters. In this ways we found it is possible with suitable values to reproduce the dynamics with the characteristics describe in Sect. 2. The most meaningful parameters are the probability $p$ of rewiring, the number $s$ of swap and the number $n$ of vertices of the graph. As we cannot have the sample as huge as the Internet, we make sure that the size of the sample is enough to leave invariant the parameters of the simulation. We vary the number of vertices $n$ of the graph $G$. We simulate the measurement on a small-world graph with a size varying from 20 000 to 200 000 and the other parameters fixed $p = 0.15$, $d = 3000, s = 50$.

Firstly, we observe that it is possible to reproduce on the small-world graph the sustained discovery of IP addresses as observed with real data. Secondly, the number of vertices discovery increases with the size of sample until some threshold (around 150 000). Beyond this threshold the number of vertices discovery becomes invariants with respect to the size. Next, we choose to fix the size of sample beyond the threshold at 300 000 vertices.

Now, we focus on the evolution of the number of discovery vertices over time when varying the number $s$ of swaps. The swaps simulate the dynamics of route changes on the topology. The number of swaps varies from 0 to 1 000 with the fixed parameters $n = 300\,000$, $p = 0.15$ and $d = 3\,000$.

Figure 4 presents the number of vertices observed at each and round. As observed with real measurement data, the number of vertices observed at each

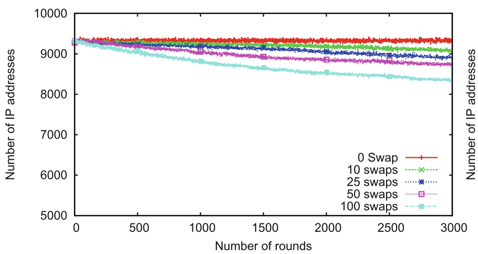


**Fig. 4.** Impact of the number of swaps on the number of vertices observed at each round.

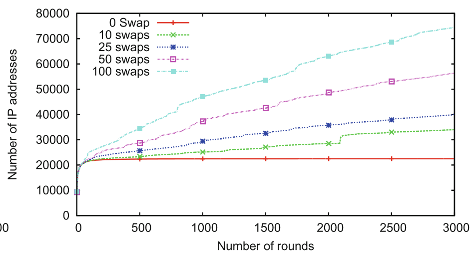


**Fig. 5.** Impact of the number of swaps on the sustained discovery of vertices.

round is roughly the same with slight decrease at the end. When increasing the number of swaps, the slope of the curve becomes large.

Figure 5 shows the evolution of the cumulative number of vertices observed since the beginning of the measurement. We find a relation between the number of swaps and the speed of discovery new vertices. The curve of the high number of swaps is above and has a greater slope. This means that more swaps induce a faster discovery of new vertices. When there is no swap the number of discovered vertices remains stable. This means that only load-balancing cannot reproduce the sustained discovered of vertices as observed in Internet. Increasing beyond hundred the number of swaps lead to faster discovered of vertices than what observed on real measurement. We assume that the number of 50 swaps is relevant to fit simulation on the real data.

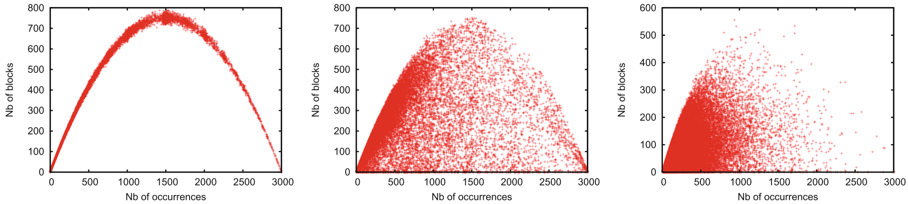


**Fig. 6.** Occurrences of IP addresses for different values of the number of swaps. Left: 0 swap. Middle: 50 swaps. Right: 500 swaps

We have the same result with the dynamics pattern. When there is no swap the load-balancing presents arc shape. The number of swaps spread the vertices under the arc. We obtain similar pattern of the Internet when the number of swaps is approximately between 25 and 100. The high number of swaps leads to faster discovery of the vertices and nearly all vertices that can be observed are discovered in short time. We suppose that the number of 50 swaps is appropriate to reproduce the observed dynamics of Internet topology.

### 4.2 Clustering vs Average Path Lengths

The coefficient clustering and the average path lengths are two important parameters of the complex networks like Internet. In the model of Watts-Strogatz, the clustering and the average path decrease until their low value when the probability of the rewiring $p$ increases until reach 1. The average path lengths decreases faster than the coefficient clustering when $p$ increases. Therefore, the small-world effect is obtained with small values of $p$ which is not enough to decrease strongly the clustering. We have chosen the rewiring probability $p = 0.15$ as appropriate to perform the simulation. Now we study the influence the rewiring on the capacity to reproduce observed dynamics. We keep the other parameters fixed at their appropriate setting and we vary $p$ from 0 to 1.

Figure 7 shows simulation result for three values of $p$. When $p = 0$, the graph is ring lattice with high coefficient clustering and the average path lengths is

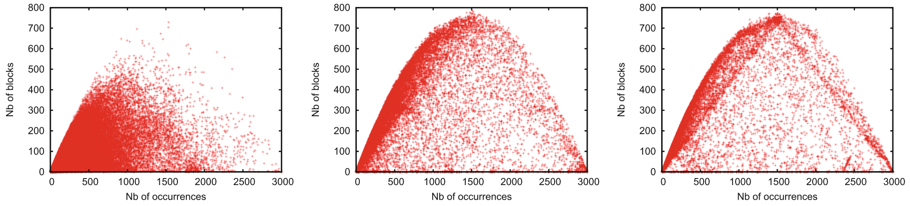


**Fig. 7.** Occurrences of IP addresses for different probabilities of rewiring $p$. Left: $p = 0$. Middle: $p = 0.25$. Right: $p = 1$.

at its maximal value. We do not see the arc shape representing the effect of the load-balancing. The points are concentrated on the left. Nearly all vertices in the graph are discovered before applying the first swaps. It is impossible to reproduce the sustained discovered of vertices and the dynamics pattern on the ring lattice.

When $p$ increases a little bit more, for instance $p = 0.05$, the situation becomes different. Then, it is possible to reproduce the observed dynamics, until the probability $p$ reaches 1. We notice that we are able to reproduce the observed dynamics only when the average path lengths becomes small. We are able to reproduce the observed dynamics only when the average path lengths becomes small. While the coefficient clustering seems to have weak influence on the observed dynamics.

## 5 Related Work

Many works on the Internet dynamics concern the measurement and characterisation [3,8,12–14,16,18]. The load-balancing has been identified as responsible of some observed dynamics on the internet and induces artefacts on the measurement [17]. The authors of [3] characterize the end-to-end paths dynamics with the presence of the load-balancing [3]. Recent work on the same topic provides a tool to predict and track Internet path changes [4].

The contributions on the Internet dynamics modeling concerns mostly the AS-level topology [2,7,15,19]. The work of these authors of [11] concerns the dynamics modeling at IP-level topology. Their main goal is not to obtain a realistic model but to analyze the impact of the power-law on the dynamics characteristics observed at IP-level topology. Our contribution is in the same stream of studies. We address the role of the small-world effect on the Internet dynamics observed at IP-level.

## 6 Conclusion and Perspectives

The Internet dynamics analysis at IP-level topology is at its beginning. Previous studies focused more the measurement. In this paper we provided a simulation results to explain some dynamics behaviors observed at IP-level of the topology.

Particularly our goal was to highlight the role of the Small-world effect on this observed dynamics. Using the model of Watts-Strogatz graph with appropriate setting we have been able to reproduce the observed dynamics characteristics. We used greedy approach in order to find the suitable values of the different parameters. It consisted to vary values of the target parameter until reach the suitable ones while the other parameters are fixed.

Our contribution is a step in the comprehension of the Internet dynamics and the results are useful for many applications, including new routing protocols development and modeling. In the continuation of this work, studying the correlation between other complex network properties and the observed dynamics of Internet is necessary for more comprehension of the network properties at the origin of the observed dynamics. Another future work should be the dynamics characterization. For instance, investigate whether the dynamics behaviors observed at IP-level topology is also the same at the AS-level.

**Acknowledgment.** This work is supported by the National Scientific Research Fund, MESS/BF/2014.

## References

1. Barrat, A., Weight, M.: On the properties of small-word network models. Eur. Phys. J. B **13**(3), 547–560 (2000)
2. Chang, H., Jamin, S., Willinger, W.: Internet connectivity at the as-level: an optimization-driven modeling approach. In: ACM SIGCOMM MoMeTools Workshop (2003)
3. Cunha, Í., Teixeira, R., Diot, C.: Measuring and characterizing end-to-end route dynamics in the presence of load balancing. In: Spring, N., Riley, G.F. (eds.) PAM 2011. LNCS, vol. 6579, pp. 235–244. Springer, Heidelberg (2011). doi:10.1007/978-3-642-19260-9_24
4. Cunha, I., Teixeira, R., Veitch, D., Diot, C.: Dtrack: a system to predict and track internet path changes. IEEE/ACM Trans. Networking **22**(4), 1025–1038 (2014)
5. Radar data. http://data.complexnetworks.fr/Radar/
6. Donnet, B., Raoult, P., Friedman, T., Crovella, M.: Efficient algorithms for large-scale topology discovery. In: Eager, D.L., Williamson, C.L., Borst, S.C., Lui, J.C.S. (eds.) Proceedings of the International Conference on Measurements and Modeling of Computer Systems, SIGMETRICS, 6–10 June 2005, Banff, Alberta, Canada, pp. 327–338. ACM (2005)
7. Haddadi, H., Uhlig, S., Moore, A.W., Mortier, R., Rio, M.: Modeling internet topology dynamics. Comput. Commun. Rev. **38**(2), 65–68 (2008)
8. Latapy, M., Magnien, C., Ouédraogo, F.: A radar for the internet. In: Workshops Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008), 15–19 December 2008, Pisa, Italy, pp. 901–908. IEEE Computer Society (2008)
9. Magnien, C., Ouedraogo, F., Valadon, G., Latapy, M.: Fast dynamics in internet topology: observations and first explanations. In: Proceedings of the 2009 Fourth International Conference on Internet Monitoring and Protection, ICIMP 2009, pp. 137–142. IEEE Computer Society, Washington, DC (2009)

10. Marchetta, P., Pescape, A.: Drago: detecting, quantifying and locating hidden routers in traceroute ip paths. In: 2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 109–114, April 2013
11. Medem, A., Magnien, C., Tarissan, F.: Impact of power-law topology on ip-level routing dynamics: simulation results. In: 2012 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 220–225, March 2012
12. Ni, J., Xie, H., Tatikonda, S., Yang, Y.R.: Efficient and dynamic routing topology inference from end-to-end measurements. IEEE/ACM Trans. Networking **18**(1), 123–135 (2010)
13. Oliveira, R.V., Zhang, B., Zhang, L.: Observing the evolution of internet as topology. SIGCOMM Comput. Commun. Rev. **37**(4), 313–324 (2007)
14. Pansiot, J.-J.: Local and dynamic analysis of internet multicast router topology. Ann. Telecommun. **62**(3–4), 408–425 (2007)
15. Park, S.-T., Pennock, D.M., Giles, C.L.: Comparing static and dynamic measurements and models of the internet's as topology. In: IEEE Infocom (2004)
16. Paxson, V.: End-to-end internet packet dynamics. IEEETON **7**(3), 277–292 (1999)
17. Viger, F., Augustin, B., Cuvellier, X., Orgogozo, B., Friedman, T., Latapy, M., Magnien, C., Teixeira, R.: Detection and prevention in internet graphs. Comput. Netw. **52**, 998–1018 (2008)
18. Wang, F., Mao, Z.M., Wang, J., Gao, L., Bush, R.: A measurement study on the impact of routing events on end-to-end internet path performance. SIGCOMM Comput. Commun. Rev. **36**(4), 375–386 (2006)
19. Wang, X., Loguinov, D.: Wealth-based evolution model for the internet as-level topology. In: IEEE INFOCOM (2006)