

Topología de Internet: análisis y modelado

(Propuesta para la ECI 2017)

Dr. Ing. José Ignacio Alvarez-Hamelin (UBA-CONICET)

11 de octubre de 2016

Descripción y objetivos

El análisis y modelado de la topología de Internet permite el estudio de nuevos protocolos de ruteo como del dimensionamiento de Internet. Para ejemplificar, las tablas de ruteo BGP poseen hoy en día más de 40 mil entradas (aproximadamente el número de AS activos), lo cual complica el ruteo en redes de alta velocidad. Por lo tanto poder simular nuevos protocolos a nivel global resulta esencial.

El presente curso tiene como objetivo comprender el proceso en su totalidad. El mismo se inicia con la exploración y adquisición de datos de Internet, que es el proceso de tomografía propiamente dicho. Luego se analiza el sesgo de las exploraciones, los mismos pueden ser generados por el método de exploración utilizado. Una vez que se conoce la calidad de los mapas obtenidos, se buscan las propiedades que lo caracterizan. Finalmente se construyen modelos que reproduzcan dichas propiedades, buscando comprender las razones de la existencia de dichas características.

El orden propuesto del curso es el histórico, avanzando luego según la complejidad creciente de los temas.

Resumen de contenidos

1. Motivación del estudio de la topología de Internet (distribución de grados y sus propiedades, contraste con los primeros modelos)
2. Propiedades medidas en las redes (coeficiente de clustering, grado promedio de los vecinos, asortatividad, betweenness, k-núcleos, parámetros en grafos con pesos y multigrafos)
3. Algunos modelos de redes sin escala (grafos aleatorios, acoplamiento preferencial, optimización heurística, competencia y adaptación, modelos anotados)
4. Métodos de exploración de la topología de Internet (basados en: ICMP, UDP, BGP, métodos altamente distribuidos; proyectos actuales)
5. Sesgo de las exploraciones (análisis de la validez de los parámetros medidos, minimización del sesgo, estimación de parámetros).

Requisitos

Conocimiento de redes de datos (Internet), nociones de estadística y programación.

Evaluación (optar por una):

- Implementación de una función en *complexts++* (<https://github.com/ihameli/complexnets>)
- Estudio y análisis de distintos mapas de Internet (obtenidos a partir de CAIDA, RouteViews, etc.).

Programa extendido

- 1. Motivación** Introducción: necesidad de la utilización de los modelos de redes. Primeros modelos de redes: Waxman[30], modelos jerárquicos [34, 15]. Primeras mediciones de Internet: descubrimiento de las distribuciones de colas largas en el número de vecinos de un nodo [18]. Análisis posteriores de las propiedades de la topología de Internet y el origen de sus características [23].
 - 2. Propiedades de la topología:** Necesidad de medir propiedades. Definición y aplicación de coeficiente de clustering, grado promedio de los vecinos, asortatividad, betweenness, k-núcleos; tanto para grafos bidirigidos, dirigidos y multigrafos, en todos los casos con y sin pesos en sus aristas [24, 8, 24, 7, 22, 25].
 - 3. Modelos de topología:** Presentación y análisis de los siguientes modelos de topología de Internet: grafos aleatorios [16], acoplamiento preferencial [6, 33], optimización heurística [17, 5, 21], competencia y adaptación [29], y modelos anotados [14]. En cada caso se analizará la pertinencia, los parámetros que modelan, la posibilidad de análisis matemático y la complejidad de generación de muestras.
 - 4. Exploración de Internet:** Distintos métodos de exploración utilizando paquetes ICMP y UDP (basados en `traceroute`) y en BGP. Proyectos de exploración de Internet: CAIDA [9], DIMES [13], Oregon Router Views [26]. Análisis de sus problemáticas y desafíos.
 - 5. Sesgo en las exploraciones:** Introducción del problema. Presentación de los primeros resultados: validez de la distribución de cola larga en los grados [28, 11]. Estudio teórico de la capacidad de las exploraciones para reproducir las características reales de las redes [12]. Estudio del sesgo en redes reales [4].
- Prácticas:** Se realizarán prácticas de análisis utilizando el *Complex Systems Scilab Toolbox* [3], con mediciones de redes reales obtenidos de los proyectos CAIDA [9], DIMES [13] y Oregon Router Views [26].

Referencias

- [1] J. I. Alvarez-Hamelin and J. R. Busch. Edge connectivity in graphs: an expansion theorem. [*math.GM*], arXiv:0803.3057v1, 2008.
- [2] J. I. Alvarez-Hamelin, L. Dall’Asta, A. Barrat, A. Vespignani. Large scale networks fingerprinting and visualization using the k-core decomposition. In Y. Weiss, B. Schölkopf, J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 41–50, Cambridge, MA, 2006. MIT Press.
- [3] J. I. Alvarez-Hamelin, V. Estrada, J. I. Orlicki. Complex Systems Scilab Toolbox: <http://sourceforge.net/projects/complex-sys-sci>, 2006-2009.
- [4] J. I. Alvarez-Hamelin, Luca Dall’Asta, Alain Barrat and Alessandro Vespignani. K-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases. *Networks and Heterogeneous Media*, (3):2, 371-393, 2008.
- [5] J. I. Alvarez-Hamelin N. Schabanel. An Internet Graph Model Based on Trade-Off Optimization. *Eur. Phys. J. B, special issue on “Applications of networks”*, 38(2):231–237, march II 2004.
- [6] A.-L. Barabási R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [7] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. (USA)*, 101:3747, 2004.
- [8] V. Batagelj M. Zaversnik. Generalized Cores. *CoRR*, cs.DS/0202039, 2002.
- [9] CAIDA. Cooperative Association for Internet Data Analysis, 2003. URL <http://www.caida.org/tools/measurement/skitter/>.

- [10] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, E. Shir. MEDUSA - New Model of Internet Topology Using k-shell Decomposition. *cond-mat/0601240*, 2006.
- [11] A. Clauset C. Moore. Accuracy and scaling phenomena in the Internet. *Phys. Rev. Lett.*, 94:018701, 2005.
- [12] L. Dall'Asta, J.I. Alvarez-Hamelin, A. Barrat, A. Vazquez, A. Vespignani. Exploring networks with traceroute-like probes: theory and simulations. *Theor. Comput. Sci.*, 351(1):6–24, 2006.
- [13] DIMES. Distributed Internet MEasurements & Simulations, 2005. URL <http://www.netdimes.org/science.html>.
- [14] X. Dimitropoulos, D. Krioukov, A. Vahdat, and G. Riley. Graph Annotations in Modeling Complex Network Topologies, *ACM Transactions on Modeling and Computer Simulation*, v.19, n.4, p.17, 2009
- [15] M. Doar. A better model for generating test networks. In *IEEE Globecom*, November 1996.
- [16] P. Erdős A. Rényi. On random graphs I. *Publ. Math. (Debrecen)*, 6:290–297, 1959.
- [17] A. Fabrikant, E. Koutsoupias, C. H. Papadimitriou. Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet. *LNCS*, 2380:110–, Jun 2002.
- [18] M. Faloutsos, P. Faloutsos, C. Faloutsos. On power-law relationship of the Internet topology. *Comput. Commun. Rev.*, 29:251–263, 1999.
- [19] B. Huffaker, M. Fomenkov, D. Moore, and k. claffy. Macroscopic analyses of the infrastructure measurement and visualization of Internet connectivity and performance. In *Passive and Active Measurement (PAM)*, April 2001.
- [20] LaNet-vi. LARge NETwork VIualization tool, 2005. URL <http://xavier.informatics.indiana.edu/lanet-vi/>.
- [21] L. Li, D. Alderson, R. Tanaka, J. C. Doyle, W. Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications (extended version). *Internet Math.*, 2(4):431–523, 2005.
- [22] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitripoulos, Kc. Claffy, and A. Vahdat. Lessons from three views of the internet topology. *Preprint cs.NI/0508033*, 2005.
- [23] A. Medina, I. Matta, J. Byers. On the origin of power laws in internet topologies. *Computer Communications Review*, 30(2):18–28, apr 2000.
- [24] M. E. J. Newman. Assortative Mixing in Networks. *Phys. Rev. Lett.*, 89:208701–, 2002.
- [25] R. Pastor-Satorras A. Vespignani. *Evolution and structure of the Internet: A statistical physics approach*. Cambridge University Press, 2004.
- [26] Router-Views. University of Oregon, Router Views Project, 2005. URL <http://www.routeviews.org/>.
- [27] J. I. Orlicky, V. Estrada, and J. I. Alvarez-Hamelin. Complex Systems SCILAB Toolbox. <http://sourceforge.net/projects/complex-sys-sci/> .
- [28] T. Petermann P. De Los Rios. Exploration of scale free networks. *Eur. Phys.J. B*, 38:201, 2004.
- [29] M. A. Serrano, M. Boguñá, A. Diaz-Guilera. Competition and adaptation in an Internet evolution model. *Phys. Rev. Lett.*, 94:038701, 2005.
- [30] B. M. Waxman. Routing of Multipoint Connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, December 1988.

- [31] K. Yamasaki, K. Matia, D. Fu, S. V. Buldyrev, F. Pammolli, M. Riccaboni, H. E. Stanley. A Generalized Preferential Attachment Model for Complex Systems. *CoRR*, physics/0502082, 2005.
- [32] Y Shavitt and E Shir. Dimes: Let the internet measure itself. *Preprint cs.NI/0506099*, 2005.
- [33] S. H. Yook, H. Jeong, A.-L. Barabási. Modeling the internet's large-scale topology. *Proc Natl Acad Sci USA*, 99(21):13382–6, Oct 2002.
- [34] E. W. Zegura, K. L. Calvert, M. J. Donahoo. A quantitative comparison of graph-based models for Internet topology. *IEEE/ACM Transactions on Networking*, 5(6):770–783, 1997.