State of the art
Study of specific cases

DynGraph project
ANR ANR-10-JCJC-0202

March 9, 2011

The three main datasets we will consider are: IP exchanges, local views of the Internet's topology, and P2P exchanges. Since the writing of this project, we have gained the opportunity to study other datasets: contact networks between individuals, acquired with sensors, and brain activity. We describe below these new datasets, and present the existing work concerning the dynamics of all these networks.

1 IP exchanges

Internet traffic at the IP level has been widely studied with motivations ranging from quality of service and security issues, to monitoring of applications and of usage of the Internet, to anomaly detection in order to detect attacks [45, 35, 41, 27, 4, 26, 56, 44, 24, 31, 14, 6, 17]. Well-known aspects of IP traffic are its variability and the multiplicity of traffic, be it legitimate or not, and its non-trivial statistical properties (lack of unique relevant time-scales, long range dependance, non-gaussianity) [37, 23]. These characteristics have been often studied through methods from time-series analysis [1, 8, 43].

Our approach is different, since we consider IP traffic as a graph, in which two IP addresses are linked if they exchange packets. Some work have used this approach, but have considered that the graph is static, i.e. does not evolve with time [29, 22].

2 Internet Topology

The Internet's *physical* topology consists of routers and computers present on the Internet, joined by wires. A good knowledge of this topology is fundamental for understanding the workings of the Internet, designing relevant and efficient protocols on this topology, and know where intervention is needed in case of problems (congestions caused by an increase of traffic with time for instance).

However, the Internet has evolved in a decentralised way, and acquiring data on its topology is a difficult task. In the last ten years, a large body of research has been devoted to measure and analyse this topology [53, 16, 15, 46], and very important advances have been made. Since obtaining a map of the Internet is a heavy and very time-consuming task, most efforts in this domain have therefore been devoted to collect and analyse static views of this map.

Some works have begun to study the dynamics of this topology [28, 33, 34, 39, 36, 25, 38], to which participants of this project have contributed, but these works remain preliminary. Their main outcome is that the dynamics under concern is intense and much more complex than previously expected.

3 P2P exchanges

File exchanges on Peer-to-Peer (P2P) systems have received much attention in the last few years. Because of their distributed nature, collecting information on these systems is a challenging task. Some works have however managed to collect and study datasets about P2P exchanges, in particular in the *eDonkey* system, with the main goal of designing efficient protocols for these systems, see for instance [21, 18, 20, 51, 32, 55, 50].

The papers which studied the dynamics of P2P networks have mainly been concerned with the study of *churn*, i.e. the rate at which users join or leave the system. See for instance [50, 47, 32, 5] and references within. These works have studied different properties related to churn, in particular session lengths, inter-session time, and correlations between consecutive session lengths.

[30] also studies churn, in order to identify nodes which are connected to the system with the same temporal patterns, the aim being to design protocols that take this into account when nodes look for other nodes to connect to in the network.

[49] addresses the question of obtaining a representative sample of a P2P networks, through random walk techniques, in the case where the nodes join or leave the system, possibly at a high frequency.

4 Contact Networks

Proximity between individuals can be detected with wireless sensors carried by people, that are able to detect when they are close to each other. The result is a dynamic network contact, with, at each time step, the information of who is close to whom.

This object has received much attention in the recent years. The main property that has been studied is the characterization of the times of contact and inter-contact [7, 9, 12, 42, 52, 13].

[19] studied the bias on the observed contact duration caused by the fact that some sensors may fail to detect each other at some times.

5 Biological networks

Some authors have studied the dynamical nature of different types of biological networks [2, 3, 40, 54, 48].

We have the opportunity to study a dataset recording the brain activity measured via electroencephalography techniques (EEG). The object under study is an evolving graph in which nodes stand for different brain areas and links for the evidence of an electrical activity between two distinct areas [11, 10]. One of the key question in this context is to identify pertinent statistical properties in order to anticipate specific events in the brain activity (such as seizures for epileptic patients for instance).

References

[1] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch. Multiscale network traffic analysis, modeling, and inference using wavelets, multifractals, and cascades. *IEEE Signal Processing Magazine*, 3(19):28–46, May 2002.

[2] M.M. Babu, L. Aravind, and S.A. Teichmann. Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol.*, 358(2):614–633, 2006.

[3] M.M. Babu, N.M. Luscombe, L. Aravind, M. Gerstein, and S.A. Teichmann. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol.*, 14(3):283–291, 2004.

[4] P. Barford, J. Kline, D. Plonka, and A. Ron. A signal analysis of network traffic anomalies. In *ACM Internet Measurement Workshop*, 2002.

[5] R. Bolla, R. Gaeta, A. Magnetto, M. Sciuto, and M. Sereno. A measurement study supporting p2p file-sharing community models. *Computer Networks*, 53:485–500, 2009.

[6] P. Borgnat, G. Dewaele, K. Fukuda, P. Abry, and K. Cho. Seven Years and One Day: Sketching the Evolution of Internet Traffic. In *Proceedings of the 28th IEEE INFOCOM 2009*. IEEE, 2009.

[7] R. Calegari, M. Musolesi, F. Raimondi, and C. Mascolo. CTG: A connectivity trace generator for testing the performance of opportunistic mobile systems. In *European Software Engineering Conference and the International ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE07)*, Dubrovnik, Croatia, 2007.

[8] O. Cappe, E. Moulines, J. Pesquet, A. Petropulu., and X. Yang. Long-range dependence and heavy tails models for statistical analysis and representation of teletraffic data. *IEEE Signal Processing Magazine*, 19(3), May 2002.

[9] A. Chaintreau, J. Crowcroft, C. Diot, R. Gass, P. Hui, and J. Scott. Pocket switched networks and the consequences of human mobility in conference environments. In *WDTN*, pages 244–251, 2005.

[10] M. Chavez, M. Valencia, V. Latora, and J. Martinerie. Complex networks: New trends for the analysis of brain connectivity. *International Journal of Bifurcation and Chaos*, 20(6):1677–1686, 2010.

[11] M. Chavez, M. Valencia, V. Navarro, V. Latora, and J. Martinerie. Functional modularity of background activities in normal and epileptic brain networks. *Phys. Rev. Lett.*, 104(11):118701, Mar 2010.

[12] A. Clauset and N. Eagle. Persistence and periodicity in a dynamic proximity network. In *DIMACS Workshop*, 2007.

[13] Vania Conan, Jérémie Leguay, and Timur Friedman. Characterizing pairwise inter-contact patterns in delay tolerant networks. In *Proceedings of the 1st international conference on Autonomic computing and communication systems*, page 19, October 2007.

[14] G. Dewaele, K. Fukuda, P. Borgnat, P. Abry, and K. Cho. Extracting hidden anomalies using sketch and non gaussian multiresolution statistical detection procedures. In *SIGCOMM 2007 Workshop LSAD - ACM SIGCOMM 2007 Workshop on Large-Scale Attack Defense (LSAD)*, kyoto, Japan, 2007.

[15] Dimes website: http://www.netdimes.org/new/.

[16] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *ACM SIGCOMM*, 1999.

[17] Romain Fontugne, Pierre Borgnat, Patrice Abry, and Kensuke Fukuda. MAWILab: Diverse Anomaly Detectors for Automated Anomaly Labeling and Performance Benchmarking. In *ACM CoNEXT*, 2010.

[18] Pierre Fraigniaud, Philippe Gauron, and Matthieu Latapy. Combining the use of clustering and scale-free nature of user exchanges into a simple and efficient p2p system. In *LNCS, proceedings of the 11-th international conference Euro-Par*, 2005.

[19] Adrien Friggeri, Guillaume Chelius, Eric Fleury, Antoine Fraboulet, France Mentr, and Jean-Christophe Lucet. Reconstructing social interactions using an unreliable wireless sensor network. *Computer Communications*, 34 (5), 2011.

[20] J.-L. Guillaume, S. Le-Blond, and M. Latapy. Clustering in p2p exchanges and consequences on performances. In *IPTPS*, 2005.

[21] S. Handurukande, A.-M. Kermarrec, F. Le Fessant, and L. Massoulié. Exploiting semantic clustering in the edonkey P2P network. In *Proceedings of ACM SIGOPS*, 2004.

[22] Marios Iliofotou, Prashanth Pappu, Michalis Faloutsos, Michael Mitzenmacher, Sumeet Singh, and George Varghese. Network monitoring using traffic dispersion graphs. In *Proceedings of Internet Measurement Conference*, pages 315–320. ACM, 2007.

[23] T. Karagiannis, M. Molle, and M. Faloutsos. Long-range dependence - ten years of internet traffic modeling. *IEEE Internet Computing*, September 2004.

[24] B. Krishnamurty, S. Sen, Y. Zhang, and Y. Chen. Sketch-based Change Detection: Methods, Evaluation, and Applications. In *Proceedings of ACM IMC*, Miami, 2003.

[25] M. Lad, D. Massey, and L. Zhang. Visualizing Internet Routing Changes. *IEEE Transactions on Visualization and Computer Graphics, special issue on Visual Analytics*, 2006.

[26] A. Lakhina, M. Crovella, and C. Diot. Diagnosing Network-Wide Traffic Anomalies. In *Proceedings of ACM SIGCOMM '04*, 2004.

[27] Nicolas Larrieu and Philippe Owezarski. Towards a Measurement Based Networking approach for Internet QoS improvement. *Computer Communications*, 28(3):259–273, 2005.

[28] M. Latapy, C. Magnien, and F. Ouédraogo. A radar for the internet. In *Proceedings of ADN'08: 1st International Workshop on Analysis of Dynamic Networks, in conjonction with IEEE ICDM 2008*, 2008.

[29] Matthieu Latapy and Clémence Magnien. Complex Network Measurements: Estimating the Relevance of Observed Properties. In *2008 IEEE INFOCOM - The 27th Conference on Computer Communications*, pages 1660–1668. IEEE, 2008.

[30] Stevens Le Blond, Fabrice Le Fessant, and Erwan Le Merrer. Finding Good Partners in Availability-Aware P2P Networks. In *Stabilization, Safety, and Security of Distributed Systems*, volume 5873 of *Lecture Notes in Computer Science*, pages 472–484, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[31] X. Li, Bian. F., M. Crovella, C. Diot, R. Govindan, A. Lakhina, and G. Iannaccone. Detection and Identification of Network Anomalies Using Sketch Subspaces. In *Proceedings of ACM IMC*, Rio de Janeiro, 2006.

[32] Thomas Locher, David Mysicka, Stefan Schmid, and Roger Wattenhofer. A Peer Activity Study in eDonkey and Kad. In *International Workshop on Dynamic Networks: Algorithms and Security (DYNAS)*, 2009.

[33] Clémence Magnien, Frédéric Ouédraogo, Guillaume Valadon, and Matthieu Latapy. Fast Dynamics in Internet Topology: Observations and First Explanations. In *2009 Fourth International Conference on Internet Monitoring and Protection*, pages 137–142. IEEE, 2009.

[34] Ricardo Oliveira, Beichuan Zhang, and Lixia Zhang. Observing the Evolution of Internet AS Topology. In *Proceedings of ACM SIGCOMM*, 2007.

[35] (P.) OWEZARSKI, (N.) LARRIEU, (L.) BERNAILLE, (W.) SADDI, (F.) GUILLEMIN, (A.) SOULE, and (K.) SALAMATIAN. Distribution of traffic among applications as measured in the french metropolis project. *Annals of Telecommunication, Special issue on Analysis of traffic and usage traces on the Internet - From network engineering to sociology of uses*, 2007.

[36] Jean-Jacques Pansiot. Local and Dynamic Analysis of Internet Multicast Router Topology. *Annales des télécommunications*, 62:408–425, 2007.

[37] K. Park and W. Willinger. Self-similar network traffic: An overview, 1999.

[38] S.-T. Park, D. M. Pennock, and C. L. Giles. Comparing static and dynamic measurements and models of the Internet's AS topology. In *IEEE Infocom*. IEEE, 2004.

[39] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, 2003.

[40] Robert J J. Prill, Pablo A A. Iglesias, and Andre Levchenko. Dynamic properties of network motifs contribute to biological network organization. *PLoS Biol*, 3(11):1881–1892, 2005.

[41] METROSEC project. http://www.laas.fr/METROSEC.

[42] A. Scherrer, P. Borgnat, E. Fleury, J.-L. Guillaume, and C. Robardet. Description and simulation of dynamic mobility networks. *Computer Network*, 52:2842–2858, 2008.

[43] A. Scherrer, N. Larrieu, P. Borgnat, P. Owezarski, and P. Abry. Non gaussian and long memory statistical modeling of internet traffic. In *4th Workshop IPS-MoMe*, pages 176–185. Salzburg Research, February 2006.

[44] A. Scherrer, N. Larrieu, P. Owezarski, P. Borgnat, and P. Abry. Non gaussian and long memory statistical characterisations for internet traffic with anomalies. *IEEE Trans. on Depend. and Secure Comp.*, 4(1):56–70, January 2007.

[45] Colleen Shannon, David Moore, Ken Keys, Marina Fomenkov, Bradley Huffaker, and kc Claffy. The internet measurement data catalog. *ACM SIGCOMM Computer Communication Review*, 35(5):97 – 100, 2005.

[46] Caida – Skitter project. `http://www.caida.org/tools/measurement/skitter/`.

[47] Moritz Steiner, Taoufik En-Najjary, and Ernst W. Biersack. A global view of kad. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 117–122, New York, NY, USA, 2007. ACM.

[48] Ralf Steuer, Adriano Nunes Nesi, Alisdair R. Fernie, Thilo Gross, Bernd Blasius, and Joachim Selbig. From structure to dynamics of metabolic pathways: application to the plant mitochondrial TCA cycle. *Bioinformatics*, 23(11):1378–85, June 2007.

[49] D. Stutzbach, R. Rejaie, N.G. Duffield, S. Sen, and W. Willinger. On unbiased sampling for unstructured peer-to-peer networks. *IEEE/ACM Transactions on Networking*, 17(2), 2009.

[50] Daniel Stutzbach and Reza Rejaie. Understanding churn in peer-to-peer networks. In *Proceedings of the 6th ACM SIGCOMM on Internet measurement - IMC '06*, page 189, New York, New York, USA, 2006. ACM Press.

[51] Daniel Stutzbach, Shanyu Zhao, and Reza Rejaie. Characterizing files in the modern Gnutella network. *Multimedia Systems*, 13(1):35–50, 2007.

[52] Pierre Ugo Tournoux, Jérémie Leguay, Marcelo Dias de Amorim, Farid Benbadis, Vania Conan, and John Whitbeck. The Accordion Phenomenon: Analysis, Characterization, and Impact on DTN Routing. In *Proceedings of the 28rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 1116–1124. IEEE, 2009.

[53] Traceroute@home website: http://www.tracerouteathome.net/.

[54] A. Vázquez, R. Dobrin, D. Sergi, J. P. Eckmann, Z. N. Oltvai, and A. L. Barabási. The topological relationship between the large-scale attributes and local interaction patterns of complex networks. *PNAS*, 101:17940–17945, 2004.

[55] Xiaoming Wang, Zhongmei Yao, and Dmitri Loguinov. Residual-based estimation of peer and link lifetimes in P2P networks. *IEEE/ACM Transactions on Networking (TON)*, 17(3):726–739, 2009.

[56] Y. Zhang, Z. Ge, A. Greenberg, and M. Roughan. Network anomography. In *Proceedings of ACM IMC*. ACM, 2005.