State of the art
Event detection

DynGraph project
ANR ANR-10-JCJC-0202

March 9, 2011

1 Generic event detction

The problem of event detection is far from being new. This is a classical problem in many contexts.

Many studies target event detection in the dynamics of various systems [7]. Two main approaches are followed, named anomaly-based and signature-based.

The underlying principle of anomaly-based approaches [4, 3] is that one knows the normal behavior of the system. Then, any observation that differs from this normal behavior is considered as an event. This approach is very appealing as it is able to detect any kind of event, including kinds that were never observed. It however relies on a precise knowledge of the dynamics of the considered object, and evolution of the normal behavior makes the method ineffective. In the case of the internet topology, these two limitations make this approach unapplicable.

Signature-based approaches rely on the knowledge of characteristic features of events to detect, which may be inferred from a set of known events (typically with machine learning techniques) [18, 7]. If the observed dynamics matches these features at some point, then one considers that this is an event. This approach is very effective in cases where the events may be described, like some computer viruses for instance. In our case, though, very limited knowledge of events in the internet, and no description on their impact on observed topology, are available.

Much attention has been paid to event detection in internet traffic, see for instance [1, 15, 23, 19, 13, 16, 9, 5, 10]. The methods used are mainly based on the quantity of traffic; some works study the whole traffic, and some subdivise the multi-dimensional space induced by the origin and destination of the traffic; some works also correlate informations from different sites [8]. These works have produced valuable results, but the strategies based on the volume have reached their limits, mainly due to the scale-free nature of the IP traffic that makes outlier detection difficult if not impossible.

2 Anomaly/event detection in graphs

Some works have addressed the questions of detecting anomalies in *static* graphs, i.e. parts of the graph that are not coherent with its global structure [17, 20, 6]

Some works have also studied the behaviors of nodes in dynamic networks, and isolated some behaviors which are statistically different from the expected behavior [2].

Some works have also studied the temporal evolution of graphs, to try and detect temporal and/or spatial anomalies [21, 22, 11, 14]. In particular, [12] studies the impact of the smoothing window used to compute statistics on the observed properties of a dynamic network, and notices that some window sizes lead to event detection, while others do not. However much remains to be done in this context.

References

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

[2] Lamia Benamara and Clémence Magnien. Estimating properties in dynamic systems: the case of churn in P2P networks. In *Proceedings of the Second International Workshop on Network Science for Communication Networks (Netscicom 2010), In Conjuction with IEEE Infocom 2010*, 2010.

[3] Damiano Bolzoni and Sandro Etalle. Approaches in anomaly-based intrusion detection systems, 2006.

[4] Damiano Bolzoni, Sandro Etalle, and Pieter Hartel. Poseidon: a 2-tier anomaly-based network intrusion detection system. In *UNIVERSITY OF TWENTE*, pages 144–156. IEEE Computer Society, 2006.

[5] 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.

[6] Deepayan Chakrabarti. AutoPart: parameter-free graph partitioning and outlier detection. In *Proceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases PKDD'04*, pages 112–124, September 2004.

[7] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3), 2009.

[8] Parminder Chhabra, Clayton Scott, Eric D. Kolaczyk, and Mark Crovella. Distributed spatial anomaly detection. In *Proceedings of Infocom 2008*, April 2008.

[9] 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.

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

[11] Assia Hamzaoui, Matthieu Latapy, and Clémence Magnien. Detecting Events in the Dynamics of Ego-centered Measurements of the Internet Topology. In *Proceedings of International Workshop on Dynamic Networks (WDN), in conjunction with WiOpt 2010*, 2010.

[12] Gueorgi Kossinets and Duncan J Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, January 2006.

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

[14] Ravi Kumar, Jasmine Novak, Prabhakar Raghavan, and Andrew Tomkins. On the Bursty Evolution of Blogspace. *World Wide Web*, 8(2):159–178, June 2005.

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

[16] 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.

[17] Caleb C. Noble and Diane J. Cook. Graph-based anomaly detection. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03*, page 631, New York, New York, USA, August 2003. ACM Press.

[18] Janak J. Parekh, Ke Wang, and Salvatore J. Stolfo. Privacy-preserving payload-based correlation for accurate malicious traffic detection. In *Proceedings of the 2006 SIGCOMM workshop on Large-scale attack defense*, LSAD '06, pages 99–106, New York, NY, USA, 2006. ACM.

[19] 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.

[20] Jimeng Sun, Huiming Qu, Deepayan Chakrabarti, and Christos Faloutsos. Relevance search and anomaly detection in bipartite graphs. *ACM SIGKDD Explorations Newsletter*, 7(2):48–55, December 2005.

[21] Jimeng Sun, Dacheng Tao, and Christos Faloutsos. Beyond streams and graphs: Dynamic tensor analysis. In *Proceedings of KDD*, 2006.

[22] Jimeng Sun, Yinglian Xie, Hui Zhang, and Christos Faloutsos. Less is more: Compact matrix decomposition for large sparse graphs. In *Proceedings of SDM'07*, 2007.

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