

Towards a Redefinition of Time in Information Networks?

BACK TO THE FUTURE OF THE WEB

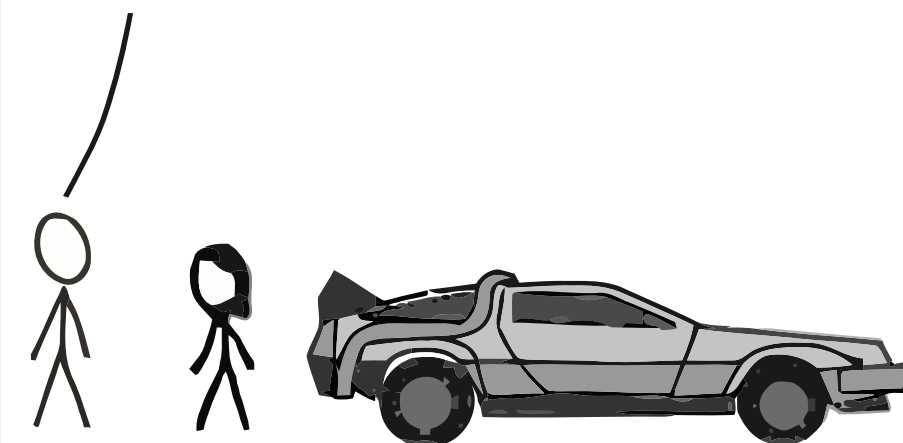
Sébastien Heymann
sebastien.heyman@lip6.fr
Université Pierre et Marie Curie – CNRS LIP6, France

Bénédicte Le Grand
benedicte.le-grand@univ-paris1.fr
Université Paris 1 Panthéon-Sorbonne – CRI, France

Complex networks evolve over time: addition & removal of nodes & links.

Dynamic properties often measured in seconds (*extrinsic time*): endogenous phenomena are difficult to observe [1-2].

Measures based on each change in the network (*intrinsic time*): endogenous phenomena are easier to observe [3].



TIME-VARYING NETWORKS



Fig 1. Local view of internet topology, where nodes represent computers and links represent routes of IP packets. Grey nodes are inactive in the current view [4].

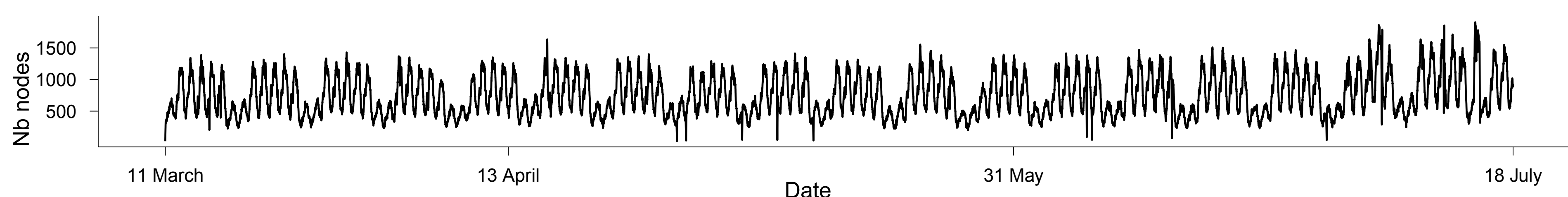


Fig 2. Number of nodes in a sliding window of 1 hour, for every 5 minutes. The temporal evolution of this statistical property for an *extrinsic time* unit shows a daily and weekly pattern.

EXTRINSIC TIME

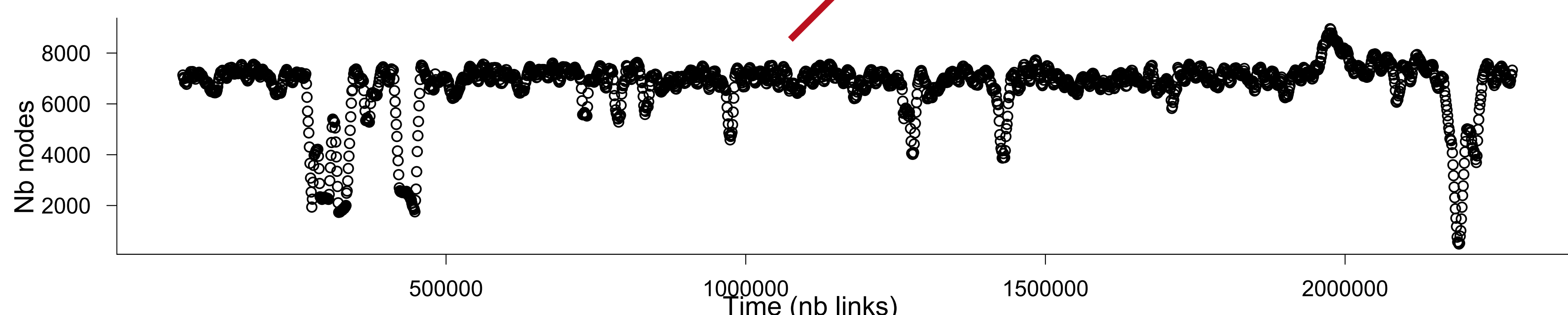


Fig 3. Number of nodes in a sliding window of 10,000 links, for each set of 1000 links. The temporal evolution of this statistical property for an *intrinsic time* unit is very stable. The events corresponding to peaks may clearly be extracted from the overall trend: the graph has normal dynamics in the statistical sense and statistical anomalies may be identified.

INTRINSIC TIME

GITHUB SOCIAL NETWORK

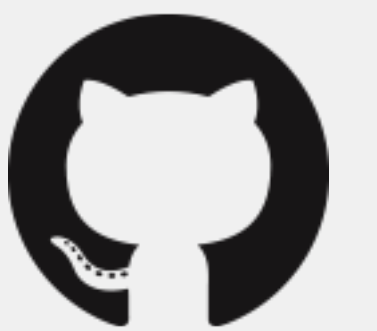
Github.com is an online platform created in 2008 to help developers share open source code and collaborate. Built on the Git decentralized versioning system, it facilitates contributions and discussions by providing a Web interface for source code repositories.

Considered activities: commit and push source code, open and close issues for bug reports, comment on issues, commits or pull request, create or delete branches and tags, and edit of the repository wiki.

Dataset: "who contributes to which repository"

Activity between users and repositories from March 11, 2012 to July 18, 2012
336 000 nodes
2.2 million links with timestamps.

Source: <http://githubarchive.org>



PERSPECTIVES

- The *intrinsic time* is an innovative concept of time in networks, based on changes observed in the network space. Using a novel time unit, we observed a **new kind of dynamics** which is not captured by traditional time units.
- Suitable for **network streams** such as IP traffic, wireless sensors logs, user-system interactions data [5].
- Produces stable time series which are simpler to analyze than periodic time series.
- Calls for the design of new on-line methodologies of measurement and **real-time analysis** of evolving networks, and for the analysis of diffusion processes.



CUMULATIVE GRAPH

Observe the growth of the graph over time, captured in a single time window.

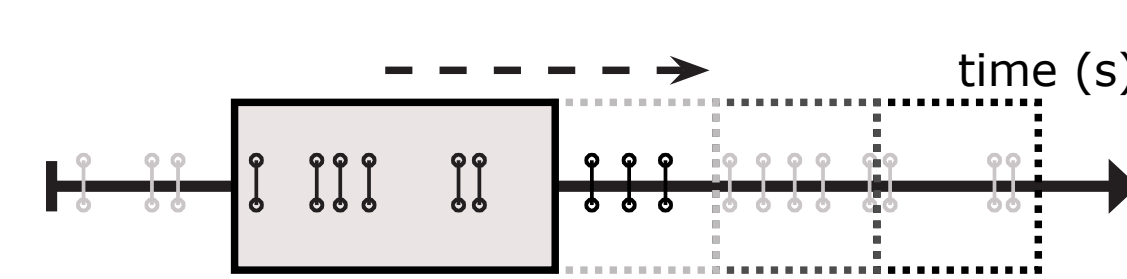


Fig 4. Stream of appearing links cumulated in a single -growing- time window.

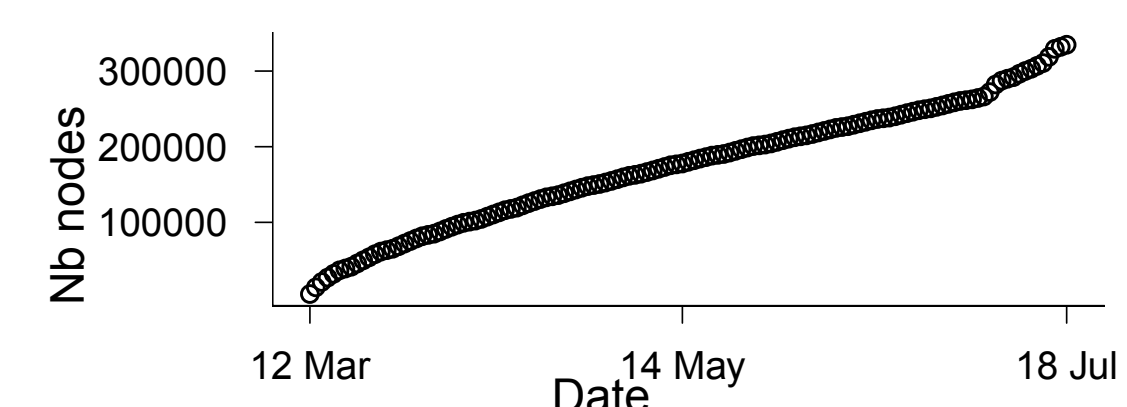


Fig 5. Cumulative number of distinct nodes as a function of the total number of observed links.

SPLIT WINDOW

Split streams of links with contiguous time windows to build a series of sub-graphs.

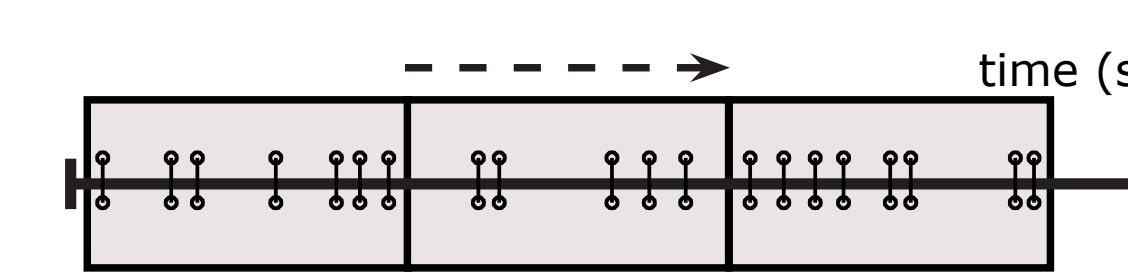


Fig 6. Stream of appearing links split in contiguous time window.

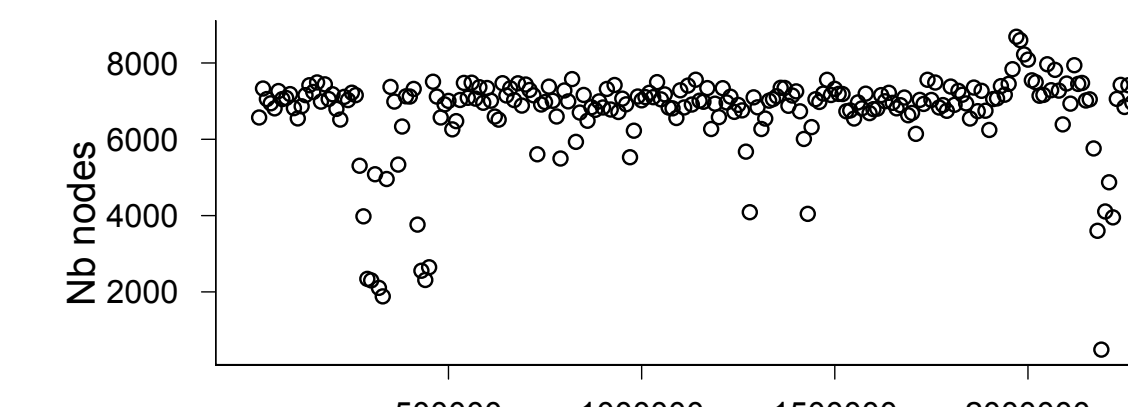


Fig 7. Number of distinct nodes in the union of 10,000 consecutive links for every 1,000 links, as a function of the number of observed links.

SLIDING WINDOW

Extract consecutive sub-graphs from a sliding time window.

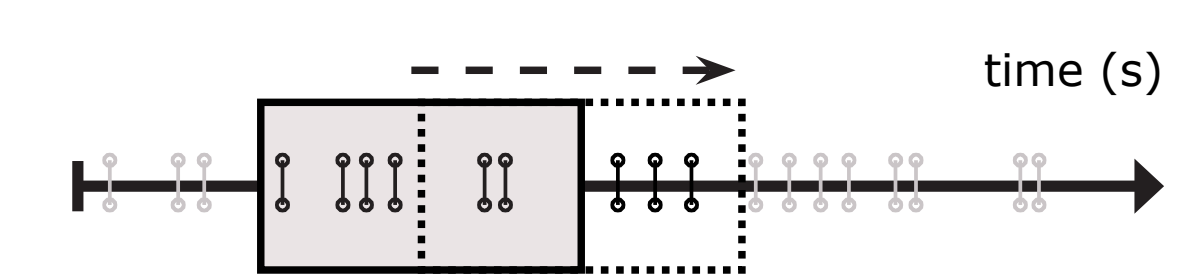


Fig 8. Example of a sliding time window over a stream of appearing links.

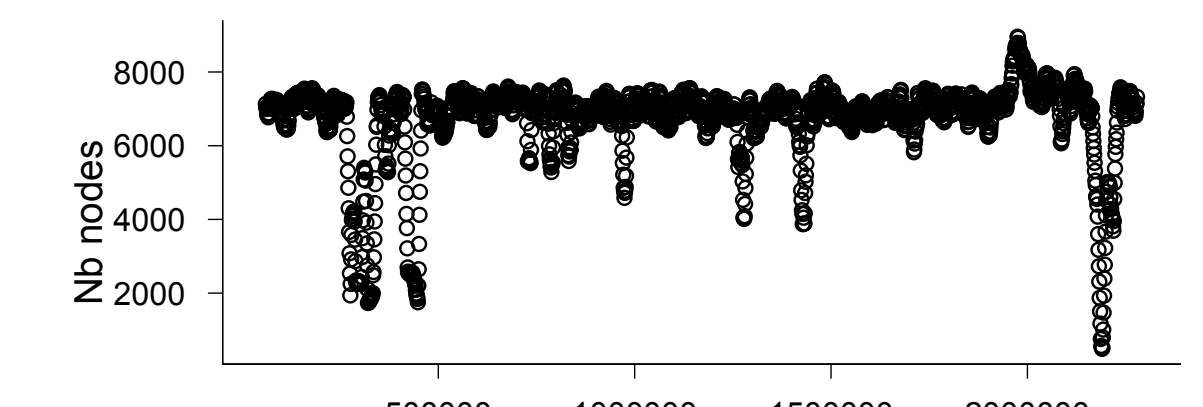
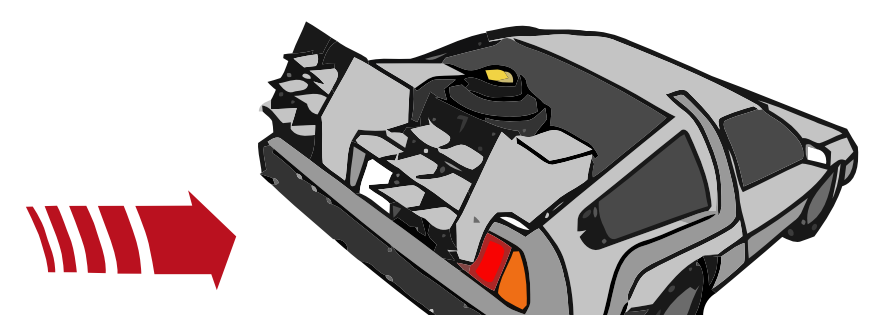


Fig 9. Number of distinct nodes in the union of 10,000 consecutive links for every 1,000 links, as a function of the number of observed links.



References

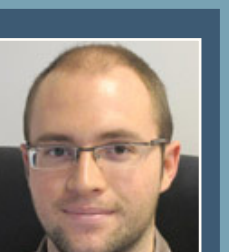
- Thomas Aynaud and Jean-Loup Guillaume. Multi-Step Community Detection and Hierarchical Time Segmentation in Evolving Networks, in Proc. of 5th SNA-KDD Workshop'11.
- Mark R. Meiss, Filippo Menczer, Santo Fortunato, Alessandro Flammini, Alessandro Vespignani. Ranking web sites with real user traffic, in Proc. of ACM WSDM'08.
- Sébastien Heymann and Bénédicte Le Grand. Towards A Redefinition of Time in Information Networks? to appear in Proc. of ACM WebSci'13.
- Matthieu Latapy, Clémence Magnien and Frédéric Ouédraogo. A Radar for the Internet, in Complex Systems, 20 (1), 23-30, 2011.
- Sébastien Heymann, Bénédicte Le Grand. Monitoring User-System Interactions through Graph-Based Intrinsic Dynamics Analysis, to appear in Proceedings of the 7th IEEE RCIS 2013 Conference.

Contact

I'm Sébastien Heymann, PhD candidate. I work under the supervision of Bénédicte Le Grand to study the dynamics of real-world networks (social networks, computer networks, the web...).

Contact me to discuss on the future of these studies!

Phone: +33 671-818-139
Web: <http://sebastien.pro>



This work is licensed under a Creative Commons Attribution-NonCommercial 2.5 License. Dolorian artworks come from <http://iStock.com>