Studying propagation dynamics in networks through rule-based modeling

Jason Vallet* Bruno Pinaud† Guy Melançon‡
University of Bordeaux, CNRS UMR 5800 LaBRI, Talence, France

Abstract

Modeling propagation dynamics on networks is an amazingly fertile and active area of research. Roughly speaking, network models aim at gaining a better understanding of how actors influence the overall network behaviour through their individual actions. However, considering the extended literature surrounding the subject, one is entitled to think that moving beyond the state-of-the-art in network modeling requires the ability to compare models, or consider slight variations of a model. This requires having a common language describing all considered models, allowing to objectively compare them and unfold their inherent properties and complexity. This also assumes users can easily run models, steer them and interactively evaluate their performance and behaviour.

The approach we describe aims at providing a framework turning network propagation modeling into rule-based modeling (aka graph rewriting). That is, models are described as a set of algorithmic transformation rules acting locally. Our approach has partially been validated by providing such a description of a well-known model relying on probabilistic rules, where nodes trigger actions depending on their neighbor's influences. The results so obtained confirm rule-based modeling as a promising avenue. The use of a visual analytics framework to conduct such tasks is vital and motivated us to further develop and adapt a general purpose visual analytics system for graph rewriting to the particular case of network propagation.

Index Terms: I.2.4 [Artificial Intelligence]: Knowledge Representation—Formalisms and Methods; F.1.2 [Computation by Abstract Devices]: Modes of Computation—

1 Introduction

Network propagation is a fundamental paradigm to study a large spectrum of phenomenon such as social communication and influence [2, 3, 4], information spreading and topic dissemination [7] or epidemics [5], for instance. Models found in the literature are diverse, not only because they focus on different aspects of network propagation, but also because they are expressed using different formalisms. Models are sometimes based on equations (differential or probabilistic), and some other times rely on algorithmic descriptions. Their validation moreover relies on *ad hoc* implementations rarely revealed or discussed by their authors.

This diversity makes it hard to objectively compare models. Kempe *et al.* [6] have proposed a unified framework making it possible to express linear threshold models and cascading models using a same formalism. Theoretically speaking, this shows the possibility to express all models using a common language, although the framework does not offer specifications on how models should be run. Comparing two models in order to identify similarities or nuances is just one of many tasks one may want to conduct. Being able to locally or slightly change the model and observe the impact at a global level is yet another task of interest. These tasks however may be hard to achieve on an equation-based description of a model.

We have recently worked on a solution offering users the necessary control on the model description and use, allowing them to easily compare models or observe how a model is impacted when modified. Our work suggests that rule-based modeling offers the necessary expressiveness to describe any propagation model on a network. Moreover, because rules can be embedded in a visualization system [9] and because the model can be simulated and steered from its visualization, our solution appears useful and powerful enough to satisfactorily support common comparison tasks when studying network propagation models. Incidentally, rule-based modeling has proven to be quite relevant and useful to model bio-molecular interactions [10].

2 Rule-based modeling and network propagation

Roughly speaking, propagation consists in sequences of incremental state modifications of network nodes. In most cases, the propagation simply turns a node into *active* mode after some conditions are met. By nature, these conditions are local and depend on the states of neighbour nodes. In two words, a node becomes active when enough of its neighbours themselves are active. Of course, models introduce subtleties as to when and how a node is influenced by its neighbours.

As it turns out, these local modifications can all be realized by applying *rewriting rules* encoding local state transformations. The applicability of a rule depends on whether a node and its neighbour satisfy a given pattern. At a global level, the evolution of the network follows from the repeated applications of rules. This process is inherently non-deterministic, since rules may apply in different order, but also because the patterns allowing the applicability of rules may overlap.

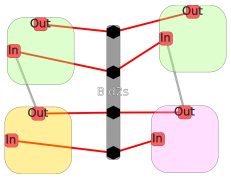


Figure 1: Rewriting rule example

Part of our work aims at proving, either theoretically or empirically, that *any* propagation model can be described as a rule-based model. For now on, we have successfully shown how rules capture the probabilistic model proposed by Goyal *et al.* [3], focusing our efforts on one of the most cited model in this area. Obviously, our claim requires that we indeed cover a much larger number of models. Our project is ongoing, and we plan to address this issue in the coming year. Our proof that Goyal *et al.*'s model can indeed be realized as a rule-based model has requested much of our efforts implementing it into PORGY [8, 1, 9]. However, because PORGY was not originally designed to support propagation models, it has to be significantly extended.

*e-mail: Jason.Vallet@labri.fr
†e-mail:Bruno.Pinaud@labri.fr
‡e-mail:Guy.Melancon@labri.fr

3 Rewriting rules

Rules are specified using a graphical equation $L \rightarrow R$ meaning that any pattern matching the left hand side of the rule L can be transformed into the right hand side of the rule R. In most papers, rule systems (also known as graph rewriting systems) are traditionally specified graphically where possible ambiguities are easily discarded. The implementation of rules and use of graphical editors to build and modify them forces the use of artifacts capable of dissolving any ambiguities, as shown in Fig. 1. Potential ambiguities are solved by placing a central rectangle holding red threads (lines) indicating how left and right items match – although simpler cases do not require to make this correspondence explicit. These red threads *are not* part of the network being rewritten.

The illustrated rule shows how a pair of nodes, one green (active), one yellow (inactive), is *rewritten*. Assuming they are connected (pale gray link) through dedicated ports (smaller red dots) each in a given state, additional properties (probabilities not shown here that may otherwise be color-mapped) allows the yellow node to become active with its color changed to pink. An additional round completes activation and ultimately sets the color of the node to green.

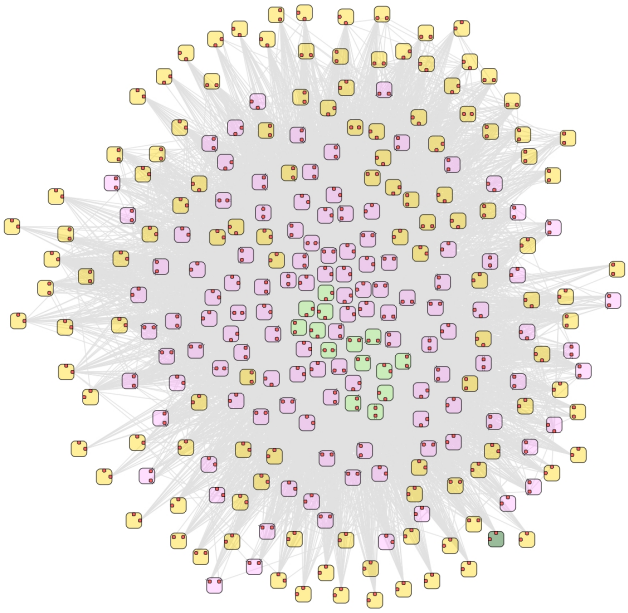


Figure 2: Propagation simulation example

As mentioned earlier, rules forms the heart of the model. Our system supports rule adjustments by allowing users to edit rules (graphs). Sequences of rule applications give rise to a tree of all possible rule applications (the *derivation tree*). As a consequence, studying the propagation model is turned into the visual examination of the tree, as well as the individual graphs. Tracking nodes to see when and how they are modified is made easier and can be further improved by using visualization techniques adapted to dynamic graphs, offering, for instance, mental map preservation.

Figure 2 shows a graph with nodes colored according to their internal states. Further applications of propagation rules show regions where nodes become active. Inspecting local regions by zooming in, filtering out edges and nodes allows to better understand how the model operates. Other derivations can be obtained by triggering alternate rule applications from an existing graph in the tree.

Validation of our approach implementing propagation models through rewriting rules was done using Goyal *et al.*'s [3] dataset. The data describes a social network where a person is linked to acquaintances, and where probabilities (states) of a node influencing a neighbour evolves in time. The active state of a node correspond to a person signing in into a service, or enrolling in a specific group (a Google+ circle, or a LinkedIn group, for instance). Running the model, we are able to compute global measures on the network, estimating how many of its nodes are active at any given time.

4 Conclusion and future Work

We believe rule-based modeling to be a promising approach allowing easier comparison of network propagation models. On the one hand, rule-based formalism allows to compare models on a theoretical level. On the other hand, the use of a visual analytics framework to steer their simulation provides full control over the models, and allows users to introduce slight changes and evaluate their impact in the overall behaviour of the model.

Our work is ongoing and much work is needed to fully validate our claims. We plan to translate other existing propagation models, able to handle both static and dynamic data, and express them with rewriting rules, so we could compare Goyal *et al.*'s to competitor models and exhibit the pros and cons for each of them. We also plan to further develop PORGY into a complete visual analytics environment to study network propagation models.

Acknowledgements

The authors wish to thank Hélène Kirchner for her invaluable help. This work was supported in part by a grant from the French research agency (ANR EVIDEN ANR-10-JCJC-0201).

References

[1] M. Fernandez, H. Kirchner, and B. Pinaud. Strategic Port Graph Rewriting: An Interactive Modelling and Analysis Framework. 2014. Presented at the 3rd Workshop On GRAPH Inspection and Traversal Engineering (GRAPHITE 2014). See http://sysma.imtlucca.it/graphite/. To appear in EPTCS (Electronic Proceedings in Theoretical CS).

[2] M. Gomez Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. In *16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, pages 1019–1028. ACM, 2010.

[3] A. Goyal, F. Bonchi, and L. V. Lakshmanan. Learning influence probabilities in social networks. In *3rd ACM International Conference on Web Search and Data Mining*, pages 241–250. ACM, 2010.

[4] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan. A data-based approach to social influence maximization. *Proceedings of the VLDB Endowment*, 5(1):73–84, 2011.

[5] T. House, G. Davies, L. Danon, and M. Keeling. A motif-based approach to network epidemics. *Bulletin of Mathematical Biology*, 71(7):1693–1706, 2009.

[6] D. Kempe, J. M. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.

[7] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang. Mining topic-level influence in heterogeneous networks. In *19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 199–208. ACM, 2010.

[8] B. Pinaud, J. Dubois, and G. Melançon. Porgy: Interactive and visual reasoning with graph rewriting systems. In *IEEE Conference on Visual Analytics Science and Technology (VAST) (poster)*, pages 293–294, 2011.

[9] B. Pinaud, G. Melançon, and J. Dubois. Porgy: A visual graph rewriting environment for complex systems. *Computer Graphics Forum*, 31(3pt4):1265–1274, 2012.

[10] M. A. Smith, W. Xu, Y. Sun, R. F. J., and E. G. Marai. Rulebender: Integrated visualization for biochemical rule- based modeling. In *1st IEEE Symposium on Biologial Data Visualization*, 2011.