

Towards Merging Models of Information Spreading and Dynamic Phenomena in Social Networks

Erick Stattner, Martine Collard, and Nicolas Vidot

LAMIA Laboratory
University of the French West Indies and Guiana, France
{estattne, mcollard, nvidot}@univ-ag.fr

Abstract. While the impact of network properties on information spreading is now widely studied, influence of network dynamics is very little known. In this paper, we study how evolution mechanisms traditionally observed within social networks can affect information diffusion. We present an approach that merges two models: model of information diffusion through social networks and model of network evolution. Since epidemics provide a reference in application domains of information spreading, we measure the impact of basic network structure changes on epidemic peak value and timing. Then we investigate observed trends in terms of changes appearing in the network structure. Our results provide promising results on how and why network dynamics is a strong parameter to integrate in requirements for information spreading modelling.

1 Introduction

Network analysis has been the subject of an active domain, so-called “Science of Networks” [2,4], an emerging scientific discipline that encompasses the whole diversity of researches on interconnected entities. Intensive effort has been done to study the structure of networks, especially with the emergence of Social Network Analysis (*SNA*) and its applications in various fields such as sociology [17], biology [14], ethology [9] or computer science [3].

In this paper, we address the issue of information dissemination through social networks, a field that has recently been explored with a focus on network modeling [16,21]. *Information* is here considered with a wide meaning and may represent either knowledge, rumor or viruses for instance. It has been argued in several works that the nature of the information does not make much difference for the modeling principles of diffusion [4,10,12]. Whatever can be the kind of information, it is now well admitted that the main concern for modeling the diffusion is the impact of social contacts of individuals [15,6].

However although the effect of social network properties on spreading is currently studied [20,6], the impact of network changes is an emerging field [19,7]. Some solutions have been proposed to model evolution processes leading to specific structural features observed on real world networks [18,11]. It is interesting to note that the issues of *dynamics of networks* (network evolution) and

dynamics through networks (dissemination) are still independent fields. In the latter, the mathematical approach of Gross et al. [13] models the impact of links deletion on spreading. Read et al. [19] show how changes in the frequency of encounters between individuals may impact the dissemination. More recently, Christensen et al. [7] measured the effect of changes in demographic attributes within population on the disease transmission.

Nevertheless, to the best of our knowledge, it seems that no empirical or even comparative study was proposed to explain the impact of network dynamics on the information dissemination process. In this work we focus on the impact of network dynamics by comparing effect of several evolution mechanisms on incidence curves. We show that the dynamics of links plays an important role in spreading through the network, and therefore is a strong requirement to consider for modeling the behavior of information diffusion in real world networks. A concrete example that motivates this study is provided by intervention strategies that are currently proposed in epidemiology and are generally focused on node-based measures. For example, the intervention strategy that gives best results is to vaccinate individuals (nodes) with a highest degree. However, it is obvious that individuals with highest degree at time t will probably not be in the same state at time $(t + 1)$, due to changes that occur in the network. Therefore, dynamic appears to have a strong and real impact on spreading, and have to be taken into account.

As said above, models of information diffusion are very much similar whatever can be the nature of the information. Mathematical approach of compartment models for epidemics modeling like the standard *SIR* (*Susceptible – Infected – Recovered*) model perfectly fits other cases such as knowledge or rumor spreading. These models consider that individuals moves from a state to another with a given probability. The transitions “Susceptible to Infected” and “Infected to Recovered” defined for epidemics are obviously analogous to “Innovator to Incubator” and “Incubator to Adopter” transitions in knowledge diffusion as underlined by Borner et al. [4]. Network models have extended compartment models more appropriately to understand diffusion process since they involve individuals (nodes) and contact links (vertices) among them [8,5,21]. They are able to match various kinds of information too.

In this work, we experiment the approach on the epidemics field since it is well studied and it provides references and resources such as training networks. We propose an approach that merges two models: model of information diffusion through the social network and network dynamics model. The paper is organized in four sections. Section 2 details our approach and Section 3 presents experiments and results. We conclude in Section 4.

2 Proposed Approach

Networks are alive and animated objects, in which nodes can appear and disappear, links can be created, removed, or can even evolve. Complex networks such as human sexual contact often do not have an engineered architecture but

instead-of are self-organized by the actions of a large number of individuals. In the network paradigm, these actions are often modeled as a set of rules, that create or delete links leading to particular topological features.

Thus, we compare effects of four well known basic evolution models, accepted in the literature as reproducing changes observed in real world networks [11,18,4]. In this preliminary work, we have deliberately chosen simple evolution mechanisms that are restricted to link creation only:

- **Random (R)**: random creation of a link between two nodes.
- **Triadic Closure (TC)**: a node creates links with neighbors of its neighbors i.e. “friends of my friends become my friends”.
- **Global Connection (GC)**: a node creates links with other nodes outside of its local neighborhood i.e. beyond friends of its friends.
- **Preferential Attachment (PA)**: a node is more likely to connect to one with high degree i.e. “rich get richer”.

To address the problem of diffusion in evolving networks, we integrate a dynamic layer into the diffusion (epidemic) model. Thus, we measure the impact of evolution strategies on the diffusion process by introducing the information (a disease) in an evolving network. We assume, as is often the case in real life, that individuals behavior does not change with the occurrence of the disease, i.e. the network still follows to the same evolution strategy.

The two models of evolution are concerning: **(1)** the network that is evolving and **(2)** the disease that is spreading into this evolving network. In the commonly used *SIR* model of disease spreading, parameters are the β probability of transmission per contact and the γ probability of recover. We assume that a susceptible individual i has a probability $1 - (1 - \beta)^{k_t^i}$ to become infected, where k_t^i is the number of infected neighbors of the node i at time t . Incidence curves of *SIR* epidemics spread are illustrated in Figure 1. According to this model, the epidemic is dependent on the initial infected population size and on probabilities β and γ . For instance it is demonstrated that for high values of β and low values of γ , the diffusion is improved as depicted on Figure 1.

Our approach to merge both spreading and dynamics models is illustrated in Figure 3. Let N be the number of individuals within the network. We denote G_t the state of the network at time t , L the infection list that stores the percentage of infected nodes at each time t , W the network evolution model (R , TC , GC or

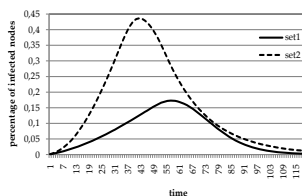


Fig. 1. Example of incidence curves obtained by *SIR* with two sets of parameters: set1 ($\beta = 0.55$ and $\gamma = 0.1$) and set2 ($\beta = 0.775$ and $\gamma = 0.05$)

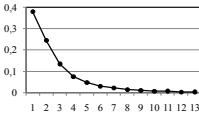
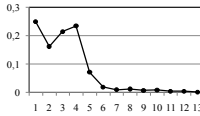
	N1	N2
Origine	Generated	Portland
#nodes	3233	4829
#links	5154	7455
Density	0.000986	0.0006395
Avg Degree	3.188	3.087
Max Degree	118	17
Distribution Degree		
Clust. Coeff.	0.00427	0.60880

Fig. 2. Main features of seed networks

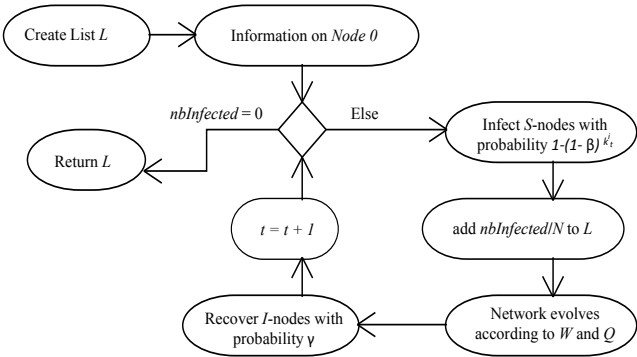


Fig. 3. Proposed approach to merge dynamics and diffusion

PA) and *Q* the evolution speed (number of links created at each iteration). We use two seed networks (see Fig. 2) and the classic *SIR* model and we introduce evolution mechanisms and measure their effects on each network. *N1* is a kind of networks most commonly observed in real world, such as the Internet, telephone calls networks, sexual networks or friendship networks known as scale-free networks. In was generated with the BarabasiAlbert model [1]. *N2* is a synthetic network extracted from *EpiSims*, an epidemiological simulation system prior to *EpiSimdemics* [3].

Both networks *N1* and *N2* are appropriate candidates for applying the dynamic models *R*, *TC*, *GC* and *PA* that add new links. And indeed, in real world networks such as scale free networks or random networks like *N1* and *N2*, link updates are mostly creations and link removing is rare.

Thus the approach was to experiment the resulting spreading-dynamic model and to observe how the network evolution may have an impact on spreading.

3 Experiment and Results

In this section, the approach is first experimented to analyze effects of evolution models (R , TC , GC , PA) on the epidemic strength in terms of value and in occurrence time. Afterwards, we investigate explanations on network properties.

3.1 Effects on Spreading

Disease behavior depends on many parameters such as the number of initial infected individuals, the probability of transmission or the probability of recover. However, changes in these parameters often influence only the virulence of the epidemic and are quite well known. In the issue we address, the most relevant parameter seems to be the evolution speed of the network. Thus to study the impact of this parameter on the disease behavior, we trained the two networks, evolution models and different evolution speeds.

The data collection was done over a period of 120 iterations. The probability of transmission was set to 0.1 and the probability of recover was set to 0.2, i.e. $\beta = 0.1$ and $\gamma = 0.2$. Each test was performed upon 100 runs and the average was computed.

On Figures 4 and 5, we compare the results obtained for each model. As a first analysis, with Figure 4, one can compare the **incidence curves** obtained with two arbitrary speeds 50 and 150. This first test allows us to make several observations on the evolution of the epidemic peak. For a more complete analysis, Figure 5 presents **summary results** on the evolution of the epidemic peak by focusing on its value and its occurrence in time.

On Figure 4, the percentage of infected nodes at each iteration is plotted according to the kind of network and evolution speed (x -axis). The incidence curve obtained without any evolution mechanism is plotted as a reference. We can observe that although all evolution strategies generate an epidemic peak

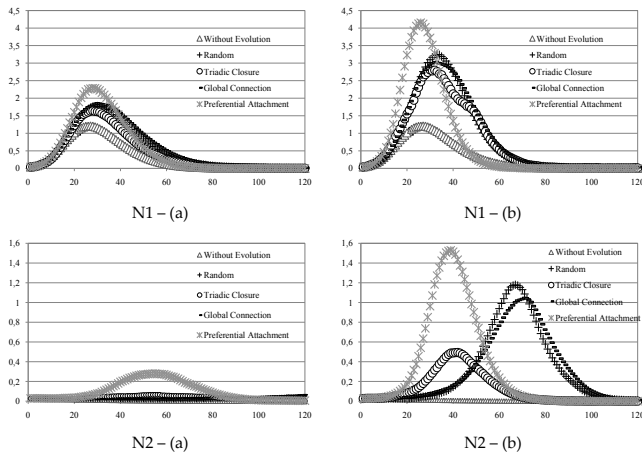


Fig. 4. Incidence curves according to networks and evolution speeds (x -axis)

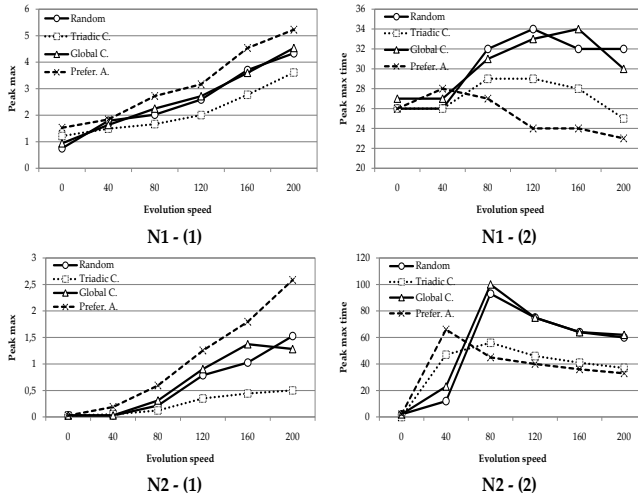


Fig. 5. Synthetic results for $N1$, $N2$ according to speed: (1) values of peak maximum, (2) peak maximum occurrence time

on $N1$ at speed 50, the difference between strategies remain very low for this speed. In the case of network $N2$ a speed of 50 links per iteration is not sufficient to generate an epidemic expect with PA the only strategy able to generate an epidemic on this network.

Globally, two main observations can be made. Dynamics has an obvious impact on the epidemic virulence, since peaks values increase with the speed with a different range from one evolution model to another. Dynamics has also an impact on the epidemic timing, since when the evolution speed increases, the epidemic peak appears more or less early according to the evolution model.

These results suggest the direct impact of the network evolution speed on: (1) the peak value, and (2) its occurrence time. Figure 5 shows how these two indices behave when the speed is varying. Figure 5(1) (resp. (2)) gives the peak value (resp. the occurrence time).

Common trends appear for the two networks. PA tends to give an epidemic curve with a peak that is systematically higher than in the other three strategies. It induces the earliest occurrence of the epidemic. The peak obtained by TC is always the lowest and is the second to appear in time. R and GC give very similar curves: the peak occurs later than for PA and TC .

3.2 Changes in Networks Properties

To explain the observed trends, we can investigate what happens at the network level. For this, we focus on changes that occur on network structure, by studying the evolution of main network properties. As shown in Figure 4, there is a strong observable difference between strategies with speed 150. Thus we have compared changes occurred on network features according to each strategy R , TC , GC and

		min degree	avg degree	max degree	Clust. C.
N1	R	1.0	5.673	120.27	0.0045
	Tria. C.	1.0	5.367	146.04	0.3036
	Glo. C.	1.0	5.546	119.82	0.0020
	Pre. A.	1.0	5.540	142.65	0.0142
N2	R	1.0	4.54	18.88	0.3882
	Tria. C.	1.0	3.95	23.40	0.6204
	Glo. C.	1.0	4.58	18.77	0.3763
	Pre. A.	1.0	4.25	33.31	0.4588

Fig. 6. Changes in main properties of *N1* and *N2* with each evolution model after epidemics with speed 150

PA, after an epidemic diffusion with speed 150. Results are shown in Figure 6, and were obtained by averaging 100 runs.

To understand effects on these evolution mechanisms on spreading, the analysis should be conducted at two levels. (1) Comparing network properties before and after evolution (see Fig. 2 *VS* Fig. 6). (2) Comparing properties resulting from different evolution models (see Fig. 6).

Preferential Attachment reinforces links of the most connected nodes. This can be observed on the growth of max degree. For example, max degree of *N2* is 17 against 18.88 for *R*, 23.40 for *TC*, 18.77 for *GC* and 33.31 for *PA* after evolution. *PA* enables rapidly the emergence of individuals sufficiently connected to result in a strong and fast transmission of disease within the network. The virulence and earliest occurrence of the epidemic peak are thus explained.

Triadic Closure strengthens links within groups of nodes to result in a significant increase in the overall clustering coefficient: from 0.00427 to 0.3036 for *N1* and from 0.60880 to 0.6204 for *N2*. It is interesting to note that other models may even reduce the clustering coefficient. While the strategy *TC* allows the emergence of highly connected nodes, it also generates network with a high clustering coefficient. So the epidemic appears relatively early and is less virulent, since the transmission occurs mainly within a same community.

Random and **Global Connection** both tend to shorten the range of degree value. However, *GC* allows a node to connect only with any node outside its immediate community. This explains that, except for clustering coefficient, observed properties with *R* and *GC* strategies are very close. *GC* does not allow creating *triangles* while *R* is likely to do it. *R* and *GC* provide very similar results on spreading, since their effects on network properties prove to be very similar.

4 Conclusion

Dynamics is an intrinsic property of real world networks. This work tackles the emerging and fundamental issue of spreading in evolving networks. We have addressed here this issue by comparing effects of various dynamics models on

incidence curves in epidemic spreading. As currently admitted, the nature of the information spread through social networks (rumor, knowledge, virus, disease. . .) does not make much difference. Our work has highlighted a set of trends related to evolution mechanisms and provides an interesting view about the way a disease spreads through an evolving network.

In this preliminary work, the evolution strategies have been restricted to link creation and we have shown that network properties are differently modified from one strategy to another. These promising results have highlighted the impact of the network dynamics upon diffusion. They should be useful to define requirements that cannot be ignored to control and model spreading phenomena. We are currently investigating more complex dynamics model in a similar approach.

References

1. Albert, R., Barabasi, A.L.: Statistical mechanics of complex networks. *Reviews of Modern Physics* 74, 51 (2002)
2. Barabasi, A.L.: *Linked: The New Science of Networks*. Perseus Books (2002)
3. Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: Episindemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: *ACM/IEEE Conference on Supercomputing* (2008)
4. Borner, K., Sanyal, S., Vespignani, A.: Network science. In: Cronin, B. (ed.) *Annual Review of Information Science & Technology*, vol. 41, pp. 537–607 (2007)
5. Chen, Y., Tseng, C., King, C., Wu, T., Chen, H.: Incorporating geographical contacts into social network analysis for contact tracing in epidemiology: a study on taiwan sars data. In: *Conference on Intelligence and Security Informatics* (2007)
6. Christakis, N.A., Fowler, J.H.: Social network sensors for early detection of contagious outbreaks. *PloS One* 5(9) (2010)
7. Christensen, C., Albert, I., Grenfell, B., Albert, R.: Disease dynamics in a dynamic social network. *Physica A: Statistical Mechanics and its Applications* 389(13), 2663–2674 (2010)
8. Christley, R.M., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R., Turner, J.: Infection in social networks: Using network analysis to identify high-risk individuals. *American Journal of Epidemiology* 162(10), 1024–1031 (2005)
9. Croft, D.P., James, R., Krause, J.: *Exploring Animals Social Networks*. Princeton University Press (2008)
10. De, P., Das, S.K.: Epidemic Models, Algorithms, and Protocols in Wireless Sensor and Ad Hoc Networks, pp. 51–75. John Wiley & Sons, Inc. (2008)
11. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. *Adv. Phys.* (2002)
12. Eubank, S., Anil Kumar, V.S., Marathe, M.: Epidemiology and Wireless Communication: Tight Analogy or Loose Metaphor? In: Liò, P., Yoneki, E., Crowcroft, J., Verma, D.C. (eds.) *BIOWIRE 2007. LNCS*, vol. 5151, pp. 91–104. Springer, Heidelberg (2008)
13. Gross, T., D’Lima, C.J., Blasius, B.: Epidemic dynamics on an adaptive network. *Physical Review Letters* 96(20) (2006)
14. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.-L.: The large-scale organization of metabolic networks. *Nature* 407, 651–654 (2000)
15. Klov Dahl, A.S.: Social networks and the spread of infectious diseases: the aids example. *Soc. Sci. Med.* 21(11), 1203–1216 (1985)

16. Lopezpintado, D.: Diffusion in complex social networks. *Games and Economic Behavior* 62(2), 573–590 (2008)
17. Milgram, S.: The small world problem. *Psychology Today* 1, 61–67 (1967)
18. Newman, M.E.J.: The structure and function of complex networks. *Siam Review* 45, 167–256 (2003)
19. Read, J.M., Eames, K.T.D., Edmunds, W.J.: Dynamic social networks and the implications for the spread of infectious disease. *J. R. Soc. Interface* 5(26) (2008)
20. Salathe, M., Jones, J.H.: Dynamics and control of diseases in networks with community structure. *PLoS Comput Biol.* 6(4) (2010)
21. Tripathy, R.M., Bagchi, A., Mehta, S.: A study of rumor control strategies on social networks. In: 19th ACM International Conference on Information and Knowledge Management, pp. 1817–1820 (2010)