# INDIVIDUAL VS. NETWORK PREFERENCES

*June Zhang and José M.F. Moura*

Carnegie Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, PA 15213 USA
Email: {junez@andrew, moura@ece}.cmu.edu

**ABSTRACT**

To study signals on networks, to detect epidemics, or to predict blackouts, we need to understand network topology and its impact on the behavior of network processes. The high dimensionality of large networks presents significant analytical and computational challenges; only specific network structures have been studied without approximation. We consider the impact of network topology on the limiting behavior of a dynamical process obeying the stochastic rules of SIS (susceptible-infected-susceptible) epidemics using the scaled SIS process. We introduce the network effect ratio, which captures the preference of individual agents versus the preference of society (i.e., network) and investigate its effects.

***Index Terms—*** network, scaled SIS process, network process, individual vs group, epidemics

## 1. INTRODUCTION

Network processes model dynamical interactions between a large but finite number of agents. They are difficult to study using experiments due to their scale. Computer simulations lack explanatory ability and the ability to investigate what-if scenarios. Investigating what-if scenarios requires re-running the simulation, which may be time-consuming and expensive depending on the complexities of the simulations.

Previously in [1, 2], we studied a network process model called the scaled SIS (susceptible-infected-susceptible) process. The stochastic interaction rules between individuals are based on the general framework of the SIS epidemics model. The generality of the model makes it an appropriate approximation for cascading failures, where infection corresponds to failure and healing to recovery or the diffusion of rumors in social networks [3]. The advantage of the scaled SIS process over other network process models is that the equilibrium distribution can be described in closed-form without resorting to mean-field approximation [4] or assuming specific network topologies [5, 6]. Furthermore, we proved that inference problems such as finding the most-probable configuration, finding the marginal probabilities, finding the expected fraction of infected agents can be solved exactly or approximated efficiently using a polynomial-time algorithm [7].

In this paper, we introduce the network effect ratio, $\sigma$, which captures the preference of individual agents vs. the preference of the overall network. This is interesting when the preferences of individual agents may oppose that of the network effect dynamics. For example, in epidemics, individual agents prefer to be healthy but the network effect facilitates the spread of infection rather than keeping nodes healthy; when the individual preferences dominate, most agents in the network will most likely be healthy whereas when the network preference dominates, most agents in the network will be infected.

The network effect ratio of the scaled SIS process gives insights of the nonlinear dependence of network processes on dynamics parameters. We also show that solution of the Most-Probable Configuration Problem, which solves for the configuration with the maximum equilibrium probability, depends only on the network effect ratio rather than on the individual or network preferences. Section 2 briefly reviews the dynamics of the scaled SIS process and the Most-Probable Configuration Problem; more detail descriptions can be found in [1, 2]. Section 3 defines the network effect ratio, $\sigma$. Section 3.2 shows that the most-probable configuration remains invariant for the same network effect ratio. Section 4 concludes the paper.

## 2. SCALED SIS PROCESS

Consider a population with $N$ agents whose relationships are described by an unweighted, undirected network, $G(V, E)$. Each node in $G$ denotes an individual agent. An edge exists between two agents if there are interactions (i.e., contacts) between them. When $G(V, E)$ is a complete graph, this reflects *homogeneous mixing*; an agent is in contact with every other agent in the network. Network $G(V, E)$ with a less symmetric graph structure reflects *heterogeneous mixing* of the population.

The scaled SIS (susceptible-infected-susceptible) process, $\{X(t), t \geq 0\}$, considers microscopic interactions between individual agents in the population. The state of the scaled SIS process at time $t$ is the $N$-length vector of the state of all the individual agents,

$$X(t) = \mathbf{x} = [x_1, x_2, \ldots, x_N]^T,$$

where $x_i$ denotes the state of the $i$th agent. If the $i$th agent is heathy but susceptible, then $x_i = 0$. If the $i$th agent is infected, then $x_i = 1$. We call $x_i$ the *agent state* and $\mathbf{x}$ the network *configuration*. By retaining the state of individual agents, the scaled SIS process allows much finer degree of detail than traditional compartmental models; it allows us to determine which agents in the population are more vulnerable to infection.

The network configuration, $X(t)$, evolves stochastically in time. Based on the standard SIS framework, infected agents heal and healthy agents become infected. The scaled SIS process is a binary-state, continuous-time Markov process. See [1] for details. The dynamics are determined by 3 parameters: $\lambda, \gamma, \mu$ and the underlying network topology, $G(V, E)$. The scaled SIS process

This work was partially supported by NSF grants CCF1011903 and CCF1513936

978-1-4673-8576-3/15/$31.00 ©2015 IEEE

assumes that the infection rate of a healthy agent is exponentially dependent on the number of infected neighbors. The process assumes the following state transitions:

1. $X(t)$ transitions to the configuration where the $i$th agent ($i = 1, 2, \ldots, N$) is infected with transition rate

$$q(\mathbf{x}, H_i\mathbf{x}) = \lambda\gamma^{m_i}, \tag{1}$$

where $m_i = \sum_{j=1}^{N} \mathbb{1}(x_j = 1)A_{ij}$ is the number of infected neighbors of node $i$. The symbol $\mathbb{1}(\cdot)$ is the indicator function, and $A = [A_{ij}]$ is the adjacency matrix of $G(V, E)$.

2. $X(t)$ transitions to the configuration where the $j$th agent ($j = 1, \ldots, N$) is healed with transition rate:

$$q(\mathbf{x}, H_{j\bullet}\mathbf{x}) = \mu. \tag{2}$$

We call $\mu > 0$ the healing rate. The parameter $\lambda > 0$ is the exogenous (i.e., spontaneous) infection rate since, if $m_i = 0$, the infection rate is $\lambda$. The parameter $\gamma > 0$ is the contagion factor since it is scaled by the number of infected neighbors. The scaled SIS epidemics model does not have an absorbing state because of exogenous infection and healing. The state space of the scaled SIS process is $\mathcal{X}$ and the size of the state space is $2^N$.

In [1], we proved that the equilibrium distribution of the scaled SIS process is

$$\pi(\mathbf{x}) = \frac{1}{Z} e^{H(\mathbf{x})}, \tag{3}$$

where

$$H(\mathbf{x}) = 1^T\mathbf{x} \log\left(\frac{\lambda}{\mu}\right) + \frac{\mathbf{x}^T A\mathbf{x}}{2} \log(\gamma). \tag{4}$$

The partition function, $Z$, ensures that the equilibrium distribution is a true probability distribution. The adjacency matrix, $A$, of the underlying network explicitly appears in the equilibrium distribution. The term $1^T\mathbf{x}$ is the number of infected agents in configuration $\mathbf{x}$ and $\frac{\mathbf{x}^T A\mathbf{x}}{2}$ is the number of infected edges (i.e., edges whose end nodes are infected).

### 2.1. Regime Parameters

The equilibrium distribution depends on $\frac{\lambda}{\mu}$ and $\gamma$, which are both unitless scaling factors. The preference of the individual agents is represented by $\frac{\lambda}{\mu}$, which we call the effective healing rate. When $0 < \frac{\lambda}{\mu} < 1$, individual agents prefer the healthy state more as the healing rate, $\mu$, is larger than the spontaneous infection rate, $\lambda$. Individual agents prefer the infected state more when $\frac{\lambda}{\mu} > 1$. Individual agents have no preference for either the infected or healthy state when $\frac{\lambda}{\mu} = 1$.

When $\gamma > 1$, an additional number of infected neighbors increase the infection rate of a healthy agent; contagion rate is biased toward using the network structure for spreading infection. When $0 < \gamma < 1$, an additional number of infected neighbors decrease the infection rate of a healthy agent. When $\gamma = 1$, the infection rate of a healthy agent does not depend on its number of infected neighbors.

We can divide the parameter space into 4 different regimes:

I) **Healing Dominant:** $0 < \frac{\lambda}{\mu} \leq 1, 0 < \gamma \leq 1$;

II) **Endogenous Infection Dominant:** $0 < \frac{\lambda}{\mu} \leq 1, \gamma > 1$;

III) **Exogenous Infection Dominant:** $\frac{\lambda}{\mu} > 1, 0 < \gamma \leq 1$;

IV) **Infection Dominant:** $\frac{\lambda}{\mu} > 1, \gamma > 1$.

In Regime II) and Regime III), individual preference opposes the network preference. Regime II) **Exogenous Infection Dominant:** $\frac{\lambda}{\mu} > 1, 0 < \gamma \leq 1$ models the standard epidemics/cascading failures assumptions: individual agents prefer to remain in the healthy state but network effect spreads infection.

### 2.2. Most-Probable Configuration Problem

The Most-Probable Configuration Problem solves for the configuration with the highest equilibrium probability. The *most-probable configuration*, $\mathbf{x}^*$, is the network configuration that we would most likely observe at equilibrium:

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} \pi(\mathbf{x}) = \arg\max_{\mathbf{x}\in\mathcal{X}} H(\mathbf{x}). \tag{5}$$

The equilibrium distribution of the scaled SIS process (3) is a Gibbs distribution. Realize that the most-probable configuration is also

$$\mathbf{x}^* = \begin{cases} \arg\max_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2} \frac{\log(\gamma)}{\log\left(\frac{\lambda}{\mu}\right)} & \text{if } \frac{\lambda}{\mu} > 1 \\ \arg\min_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2} \frac{\log(\gamma)}{\log\left(\frac{\lambda}{\mu}\right)} & \text{if } 0 < \frac{\lambda}{\mu} < 1. \end{cases} \tag{6}$$

The solution of the Most-Probable Configuration Problem (5) depends on the adjacency matrix of the underlying network, $A$, and the dynamics of the process through $\frac{\lambda}{\mu}$ and $\gamma$. The most-probable configuration is particularly interesting in Regime II) **Exogenous Infection Dominant**, where $\frac{\lambda}{\mu} > 1, 0 < \gamma \leq 1$. In this regime, the topology-dependent process (controlled by $\gamma$) opposes the topology-independent process (controlled by $\lambda, \mu$). We showed in [1, 2] that many of the most-probable configurations are non-degenerate; in these non-degenerate most-probable configurations, subsets of agents are infected while others are healthy. In Regime II), the solution of the Most-Probable Configuration Problem corresponds to the minimum of a submodular function [2]. We can use Max-Flow/Min-Cut algorithm to solve efficiently this combinatorial optimization problem [10].

## 3. NETWORK EFFECT RATIO

We define the network effect ratio, $\sigma$, as

$$\sigma = \left| \frac{\log(\gamma)}{\log\left(\frac{\lambda}{\mu}\right)} \right|, \tag{7}$$

where $|\cdot|$ is the absolute value function. Taking the absolute value is a convenience since $\sigma$ may be positive or negative depending on the value of the dynamics parameters $\lambda, \gamma, \mu$.

The network effect ratio is the ratio of the network preference, as represented by the endogenous infection parameter, $\gamma$, and the individual preference, as represented by the parameter $\frac{\lambda}{\mu}$. When the numerator term dominates, the equilibrium behavior of the scaled SIS process is determined by the contagion dynamics; in the case of epidemics, this means that the scaled SIS process is driven toward increasing the number of infected agents. When the denominator term dominates, the equilibrium behavior is determined by the dynamics of individual agents; in the case of epidemics, this means that the scaled SIS process is driven toward increasing the number of healthy agents.

### 3.1. Nonlinear Behavior

According to (7), the network ratio, $\sigma$, does *not* scale linearly with changes of parameters, $\frac{\lambda}{\mu}$ and $\gamma$. For example,

$$\alpha\sigma \neq \left|\frac{\log(\alpha\gamma)}{\log\left(\frac{\lambda}{\mu}\right)}\right|.$$

Instead the network ratio scales linearly with exponential changes of parameters,

$$\alpha\sigma = \left|\frac{\log(\gamma^{\alpha})}{\log\left(\frac{\lambda}{\mu}\right)}\right|.$$

Further, the ratio of network preference to individual preference is highly nonlinear due to the natural log function. Figure 1 plots the natural log, $\log(x)$, and the absolute natural log function, $\left|\log(x)\right|$. The function $\left|\log(x)\right|$ is not smooth. The behavior of the function depends on if $0 < x < 1$ or if $x > 1$; there is a bend at $x = 1$. This shows that the network effect ratio is sensitive to the parameter regime of the scaled SIS process, especially when the parameters are either $>> 1$ or $<< 1$.

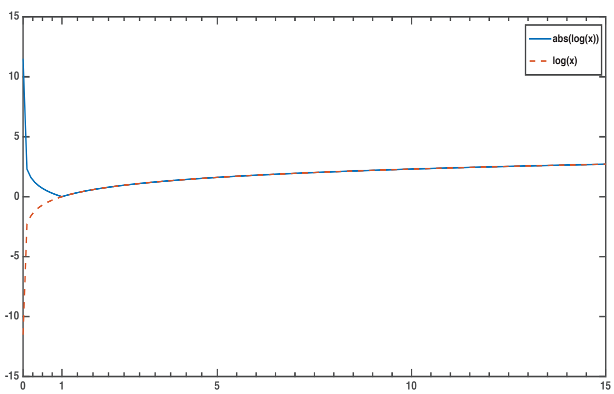



**Fig. 1**: Plot of $\log(x)$ and $\left|\log(x)\right|$

In regime II), the network prefers to spread infection $\gamma > 1$; as $\gamma$ increases, corresponding to stronger contagion effect and a stronger preference for the network spread infection, the effect on the network effect ratio, $\sigma$, gradually reaches near saturation.

In regime II), the individual prefers to remain in the healthy state with $0 < \frac{\lambda}{\mu} \geq 1$; as $\frac{\lambda}{\mu}$ decreases, corresponding to faster healing rate and stronger individual preference to remain healthy, the effect on the network effect ratio, $\sigma$, increases rapidly. This means that a small strengthening of the individual preference for *all* agents to remain healthy (i.e., increasing healing rate) can *significantly* impact the equilibrium behavior of the scaled SIS process provided that $\frac{\lambda}{\mu} << 1$.

On the other hand, when the network preference is already strong (i.e., $\gamma >> 1$), a small increase of the contagion rate will have *little* impact on the equilibrium behavior of the scaled SIS process. However, if the network preference is weak (i.e., $\gamma \approx 1$), then a small increase in the contagion rate will have more impact on the equilibrium behavior.

### 3.2. Invariance of the Most-Probable Configuration Problem

The most-probable configuration, $\mathbf{x}^*$, is invariant if the network effect ratio, $\sigma$, remains the same.

**Theorem 3.1.** *Consider two scaled SIS processes on the same network, $\{X_1(t), t \geq 0\}$ and $\{X_2(t), t \geq 0\}$, with parameters $\left(\frac{\lambda_1}{\mu_1}, \gamma_1\right)$ and $\left(\frac{\lambda_2}{\mu_2}, \gamma_2\right)$ in the same parameter regime. If $\frac{\lambda_1}{\mu_1}, \frac{\lambda_2}{\mu_2} \neq 1$, and*

$$\sigma_1 = \left|\frac{\log\gamma_1}{\log\frac{\lambda_1}{\mu_1}}\right| = \sigma_2 = \left|\frac{\log\gamma_2}{\log\frac{\lambda_2}{\mu_2}}\right|,$$

*then the most-probable configuration of both processes, $\mathbf{x}_1^*$ and $\mathbf{x}_2^*$, is the same.*

*Proof.* By definition of the most-probable configuration and depending on the value of $\frac{\lambda_1}{\mu_1}$,

$$\mathbf{x}_1^* = \begin{cases} \arg\max_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2}\frac{\log(\gamma_1)}{\log\left(\frac{\lambda_1}{\mu_1}\right)} & \text{if } \frac{\lambda_1}{\mu_1} > 1 \\ \arg\min_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2}\frac{\log(\gamma_1)}{\log\left(\frac{\lambda_1}{\mu_1}\right)} & \text{if } 0 < \frac{\lambda_1}{\mu_1} < 1, \end{cases} \tag{8}$$

and

$$\mathbf{x}_2^* = \begin{cases} \arg\max_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2}\frac{\log(\gamma_2)}{\log\left(\frac{\lambda_2}{\mu_2}\right)} & \text{if } \frac{\lambda_2}{\mu_2} > 1 \\ \arg\min_{\mathbf{x}\in\mathcal{X}} 1^T\mathbf{x} + \frac{\mathbf{x}^T A\mathbf{x}}{2}\frac{\log(\gamma_2)}{\log\left(\frac{\lambda_2}{\mu_2}\right)} & \text{if } 0 < \frac{\lambda_2}{\mu_2} < 1. \end{cases} \tag{9}$$

Since the parameters are in the same regime and $\sigma_1 = \sigma_2$, then (8) is equal to (9).

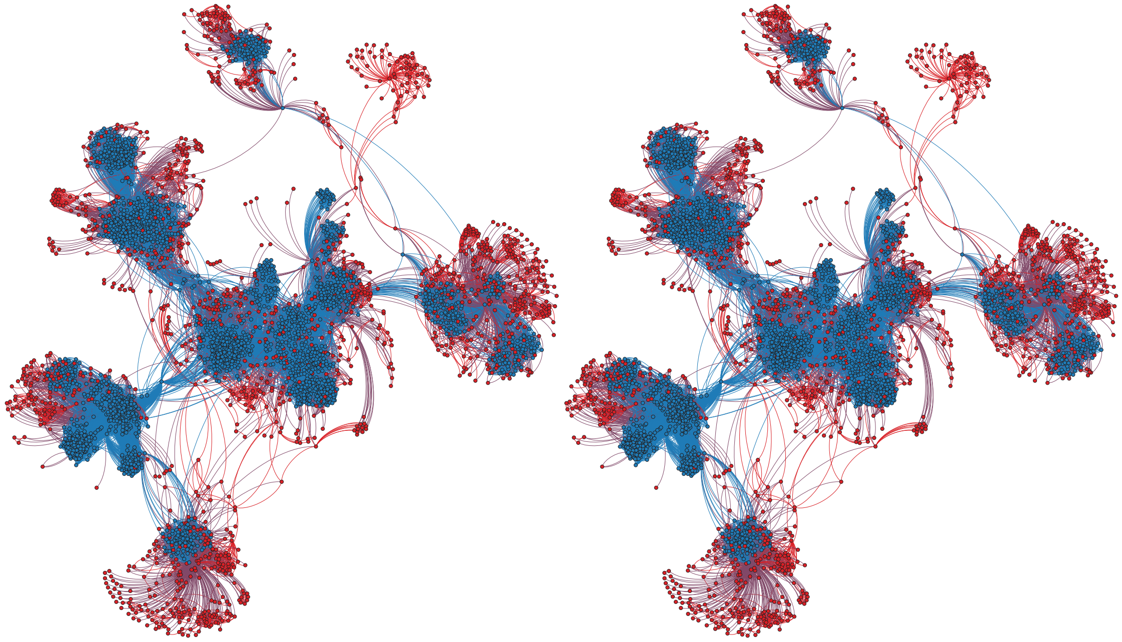
□

The network shown in Figure 2 is a small sample of the Facebook network with $N = 4039$ users and $|E| = 88,234$ connections [11]. This means that the scaled SIS process has $2^{4039}$ possible configurations. As stated in Theorem 3.1, we can see that the most-probable configuration is the same for scaled SIS processes with different parameters $\lambda, \gamma, \mu$ in Regime II) but with the same network effect ratio, $\sigma$.

## 4. CONCLUSION

Phenomena such as epidemics and cascading failures are results of dynamical interactions between multiple agents. We are interested in scenarios where the preference of individual agents opposes the preference of the network dynamic. In epidemics, individuals prefer to remain in the healthy state whereas the network dynamic evolves to spread infection through contagion. We showed, using the scaled SIS process, an analytical network process model, that the equilibrium behavior depends on the ratio of the network preference (i.e., how quickly infection spreads from infected agent to healthy agent) and the individual preference (i.e., how quickly individual agents heal). We called this the network effect ratio, $\sigma$.

For the scaled SIS process, the network effect ratio depends not on the ratio of the contagion factor, $\gamma$, and the effective healing rate, $\frac{\lambda}{\mu}$, but on the ratio of $\log(\gamma)$ and $\log\left(\frac{\lambda}{\mu}\right)$. This means that the equilibrium behavior of the scaled SIS process scales *nonlinearly* with



(a) $\left(\frac{\lambda}{\mu} = 0.01, \gamma = 1.5, \sigma = 0.088\right) : 1^T \mathbf{x}^* = 2653$ (b) $\left(\frac{\lambda}{\mu} = 0.0005, \gamma = 1.953, \sigma = 0.088\right) : 1^T \mathbf{x}^* = 2653$

**Fig. 2**: Most-Probable Configuration, $\mathbf{x}^*$, for Different $\left(\frac{\lambda}{\mu}, \gamma, \sigma\right)$. Blue = Infected, Red = Healthy

$\frac{\lambda}{\mu}$ and $\gamma$. Furthermore, depending on if the parameter value is less than or greater than 1, slight changes in parameter value may have a significant impact on the equilibrium behavior. We also proved that processes with the same $\sigma$ have the same most-probable configuration, $\mathbf{x}^*$ at equilibrium.

## 5. REFERENCES


[1] J. Zhang and J. M. F. Moura, "Diffusion in social networks as SIS epidemics: Beyond full mixing and complete graphs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 537–551, Aug 2014.

[2] ——, "Role of subgraphs in epidemics over finite-size networks under the scaled SIS process," *Journal of Complex Networks*, 2015.

[3] M. O. Jackson, *Social and Economic Networks*. Princeton University Press, 2008.

[4] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008, vol. 1.

[5] P. Van Mieghem and E. Cator, "Epidemics in networks with nodal self-infection and the epidemic threshold," *Phys. Rev. E*, vol. 86, p. 016116, Jul 2012. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevE.86.016116

[6] R. Durrett, "Some features of the spread of epidemics and information on a random graph," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4491–4498, 2010. [Online]. Available: http://www.pnas.org/content/107/10/4491.abstract

[7] J. Zhang, "Network process: How topology impacts the dynamics of epidemics and cascading failures," Ph.D. dissertation, Carnegie Mellon University, 2015.

[8] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT press, 2009.

[9] A. K. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems: Basics, Algorithms and Statistical Mechanics*. John Wiley & Sons, 2006.

[10] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[11] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, Jun. 2014.