Cascade Simulation on Optimized Networks

Takanori Komatsu
Department of computer science,
National Defense Academy,
Yokosuka, Japan
Email: ed10004@nda.ac.jp

Akira Namatame
Department of computer science,
National Defense Academy,
Yokosuka, Japan
Email: nama@nda.ac.jp

KEYWORDS

Cascade Phenomena, Evolutionary Optimization, Genetic Algorithm, Network Topology, Diffusion of innovation

ABSTRACT

Cascade phenomena, which are sequences of adoption by agents, are the important driven forces for the society to make a successful diffusion of innovation. Cascade phenomena on complex networks occur under a range of conditions (or known as cascade windows) defined by both the average degree of the underlying network topology and the threshold of each agents(or nodes). In this paper, we obtain optimal networks for good cascade using generic algorithm (GA). In order to obtain the optimal network for good cascade, the network should have a sufficient number of vulnerable nodes and hub nodes of medium sizes, in other words, the degree distribution of the optimal network should follow a linear combination of Poisson and uniform distribution.

Introduction

Why does new fashion craze every year in the world? Why can one of software company succeed to deploy the operating system markets of the personal computers? Why was Arab Spring able to begin in the Arab world suddenly on Saturday, 18 December 2010 after the dictatorship lasted for many years? These phenomena can be considered as results of coordination, agreements, and information cascade on the underlying network topologies. One of common features of these phenomena is they start from the small fraction of agents on the network and the impact of it spreads into the entire networks with self-reinforcement mechanism. The phenomena with these features are called cascade phenomena and is also called snow ball effects. The cascade consists of good cascade and bad cascade. In the former case, cascade phenomena, which are sequences of adoption by agents, are the important driven forces for the society to make a success of the diffusion of innovation or new products. Many people are also easy to get concert on the wrong decision or choice, and this is an example bad cascade.

The network topology is the most fundamental network structure to consider the performance for cascade phenomena on many networked-systems (for example, Internet, sensor networks, social networks and financial networks). From the high point of view, an agent on each node can denote the human on social network, the router or the terminal on computer network and the company on financial network and each agent are connected by links which denotes there are communications or interactions between agents. The topology, which describe links between them, defines the frequency of interaction and has a certain influence on the dynamics of the application on it. Each agent also has threshold value, which is based on the proportion of the sate of neighbors, to change the own state. If the threshold of the agents is large, it is difficult for external effect to change the state of the agents, but it is easily affected when the threshold is small. In this paper, we study about changing the dynamics of the cascade phenomena using the underlying topology, which needs not to change in other parameters (the number of agents and links and the threshold of agents).

The approach we do in this paper is based on the concept of the dynamic network topology. In some networked systems, it is not easy to change the network topology, in which the network topology is defined by physical layer, but the development of information communication technology allows us to change the topology more easily. For examples, the logical over-lay network is effective to change the topology of the P2P network, the sensor network and the cloud computing network, and the social media (Facebook(Facebook, 2011) or Twitter(Java et al., 2007)) makes the information flow in society faster and more widely. While it is basic and common problem of cascade phenomena to design or find the optimal network topology for good cascade, it becomes difficult when we have large number of nodes and links, because the situation calls on us to check the performance of huge combinations of networks. Then, we need to develop new network design method which has enough scalability compared with human-centric network design. Then, we propose evolutionary optimization by genetic algorithm (GA) to obtain the optimal network for good cascade. The evolutionary optimization is based on the accumulation of heuristic improvements and it also has three key aspects to find optimal networks for good cascade; first, the formulation of the problem requirements; second, understanding and modeling essential properties of optimized network by GA; and finally simulating cascade phenomena on optimized networks to

Proceedings 26th European Conference on Modelling and Simulation ©ECMS Klaus G. Troitzsch, Michael Möhring, Ulf Lotzmann (Editors)
ISBN: 978-0-9564944-4-3 / ISBN: 978-0-9564944-5-0 (CD)

Table 1: Payoff matrix

i \ j	1	0
1	a, a	0, 0
0	0, 0	b, b

confirm those performances. Especially the second is effective to understand the relationship between the topology and the performance. We show the network with a linear combination of Poisson and uniform degree distribution is evolutionary optimal for good cascade, on which cascade phenomena can occur under a wide range of conditions. The overall objective of the research presented in this paper is to change the dynamics of the network application on demand exploiting the underlying topology.

The rest of this paper is organized as follows. Section 2 defines a cascade model on networks. Section 3 describes the proposed evolutionary network optimization by genetic algorithms and presents some results. Section 4 simulates cascade phenomena on optimized networks, which are compared with scale-free networks and random networks. Section 5 shows the topology obtained by the evolutionary optimization. Section 6 presents a summary.

A cascade model on networks

Let the given networked system have N agents (nodes) and L links and the state of all agents is 0 as the initial state. We change the state of a few agents from 0 to 1 as a trigger of cascade and observe the spread of the state 1 into the entire network. At each time step, all agents select own state from the space $S = \{0, 1\}$ based on the proportion of the state of neighboring agents at previous time step, to be more exact, agents play a 2×2 coordination game with each neighbor and revise the state using a deterministic myopic-best response to maximize his current payoff given the proportion of neighbors choosing each action in the population. In this framework, selecting 0 or 1 means, for example, selecting product A or product B, adopting old regime or new regime, and accept or reject. Then, this cascade model has the locality and the decision rule based on the proportion, and thanks to this simplicity, the model can be applied on many situations, on which each agent make a binary decision. The payoff of each state for agents is summarized in symmetric matrix (see Table 1).

Let s_i represent the state of agent i and $j \in N_i$ is the set of neighboring agents of agent i. Networks can be directed or undirected. Here, to simplify matters, we consider undirected networks. This implies that neighboring agents mutually affect each other. In addition, let $U(s_i, s_{j \in N_i})$ be the total payoff after a 2×2 coordination game with neighboring agents, which is denoted using the summation of the payoff of each coordination game $u(s_i, s_j)$.

$$U(s_i, s_{j \in N_i}) = \sum_{j \in N_i} u(s_i, s_j) \tag{1}$$

The best response of each agents depends on the proportion of neighbors choosing 1. If the proportion p is larger than the threshold ϕ, then $i's$ best response is to choose 1. Otherwise i choose 0. The dynamics of the choosing state by agents is summarized in Eq. (2).

$$s_i = \begin{cases} 1 & p > \phi \quad (\phi = b/(a+b)) \\ 0 & p < \phi \end{cases} \tag{2}$$

Watts showed there exist some cascade area in term of the threshold ϕ and the average degree(the average number of neighboring agents) Z, where cascade may occur. They define this area as cascade window (Watts, 2002).

Fig.1 shows the example of the cascade window on networks using the cascade condition by Watts(Watts, 2002) (see Eq. (3)). In the cascade window, an initial trigger can make the cascade on the entire network, but on the outside the trigger have only a limited effect and it cannot make the cascade.

$$\sum_{k=0}^{\lfloor 1/\phi \rfloor} k(k-1)P(k) = z \tag{3}$$

where $P(k)$ represents the degree distribution of the network.

It is very interesting that even if the network with same average degree Z, the network with different degree distribution has different the size of cascade window. This implies us that we can change the dynamics of the cascade phenomena exploiting the underlying topology.

Young modeled the diffusion of innovation using cascade model and studied the cascade condition on the lattice network(Young, 2010, 2009, 1993) and Watts showed the importance of the agent with small degree to maximize cascade window(Watts and Dodds, 2007). From their results, it is important there exist a large cluster which consists of agents with a small number of links (connections). López used the technique (mean-filed approximation) of modeling the cascade phenomena and said that a network with an intermediate variance in term of degree (the number of connection) of agents maximizes the size of cascade window because the limitation make a network have a cluster as Yong and Watts introduce. López also showed the following relationship between the topology and the size of cascade window that the exponential network has larger cascade window than the scale free network(López-Pintado, 2006), which is

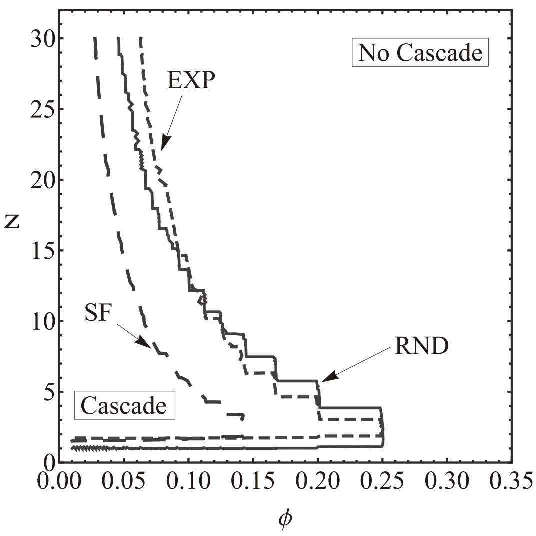


Figure 1: The cascade window as a function of the threshold value and the networks in different network topologies: Scale-Free network(SF), Random network(RND), and Exponential network(EXP).

the best topology for spreading a disease(Pastor-Satorras and Vespignani, 2002).

$$\phi^*_{\mathrm{SF}} < \phi^*_{\mathrm{EXP}} \tag{4}$$

As a result, we have simple questions "Is the exponential network the optimal network to maximize the cascade window?" or "How do we get the network which meets our requirements on the size of cascade window?" Answering these questions is the first step to change the cascade dynamics using the network topology. Then, in this paper, we consider the maximizing the cascade window for good cascade as an example problem.

Evolutionary optimization

We propose evolutionary network optimization method by genetic algorithm (GA) to obtain the optimal network, which maximize the cascade window, for good cascade. We assume all agents have same threshold ϕ and only consider connected networks, in other words, the one giant cluster should include all nodes.

Chromosome

The network topology, which consists of N nodes, are completely described by the $N \times N$ adjacency matrix $\boldsymbol{A}$. For the simplicity, we use undirected networks, then the element $a_{ij} = a_{ji} = 1$ denotes there is the interaction between agent i and agent j. The sequence of elements of adjacency matrix is the chromosome of genetic algorithm (see Fig.2). In order to generate new networks from two parents networks, we apply uniform crossover into chromosome of them.

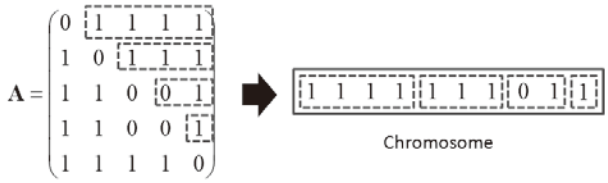


Figure 2: The chromosome of the network topology in genetic algorithms

Fitness function

According to the mean-field analysis of cascade phenomena by Lopez(López-Pintado, 2006), we can estimate cascade window size ϕ^*

$$\phi^* = \arg\min_{\phi\in[0,1]} \sum_{k\geq 1}^{\lfloor 1/\phi \rfloor} k^2 P(k) \tag{5}$$

$$\text{s.t.} \quad \sum_{k\geq 1}^{\lfloor 1/\phi \rfloor} k^2 P(k) > 1$$

In order to maximize cascade window for good cascade, we need to find the network which has the largest ϕ^* and then we use ϕ^* as a fitness function for genetic algorithm.

Procedures

The approach we propose in this section consists of three steps (see Fig.3). First steps is generating initial population by random network and scale free network. Second step is picking up two network as parents from population randomly and generating new networks as children using uniform crossover. Third step is selecting two elite networks based on fitness value to insert them into population and the next step will start from the first step again. The approach we propose do this optimization cycle until the fitness value attain required value or there is no improvement for long time. In addition, we apply programming tips not to change the number of links and isolate no node from the giant cluster after uniform crossover. This implies uniform crossover just change connected-network topology.

Experiments and outputs

Table 2 shows the detail of settings of evolutionary network optimization by GA. The accumulation of improvement by optimization cycle outputs the network which has lager fitness value along the evaluation number is growing (see Fig. 4). We summarize the results of each experiment on Fig.5. As described previously, the exponential network is most suitable to maximize cascade window, but the approach we proposed can find more optimal network which has larger cascade window.

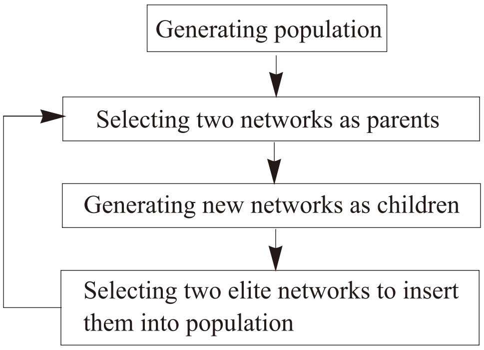
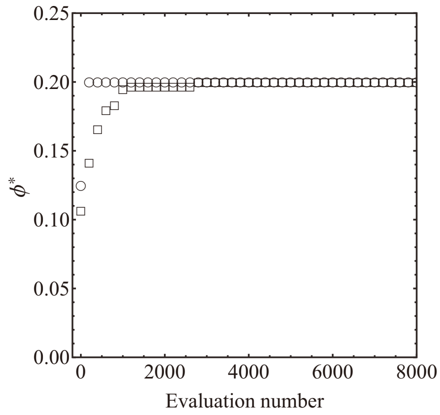


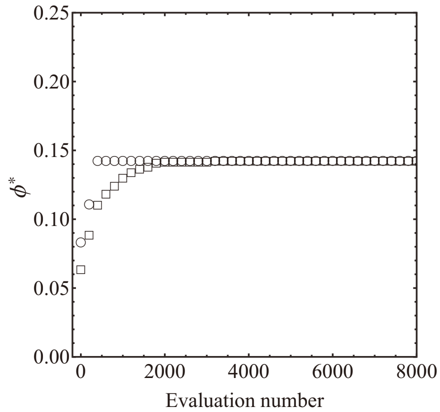
Figure 3: The procedures of evolutionary optimization by genetic algorithms

Table 2: Parameters for genetic algorithm

Genetic algorithm model	Minimum Generation Gap model(H.Sato et al., 1997)
Initial population size	20
” Child ” population size	20
Fitness function	see Eq.(6)
Crossover	Uniform crossover
Mutation	Not used
Selection	An elite selection strategy
The number of evaluations	Over 20000



(a) $\langle k \rangle = 10$



(b) $\langle k \rangle = 20$

Figure 4: The fitness value of the network as a function of the evaluation number: ◯:largest fitness value, □:average fitness value

Cascade Simulation on Evolutionary Optimized Networks

In this section, we present a set of simulation results. The simulations are conducted to confirm the optimal networks by GA have the largest cascade window, compared with scale-free network, random network and exponential network, especially, the exponential network is considered to have the largest cascade window(López-Pintado, 2006).

Simulation scenario

The fitness function (Eq.(6)) is directly derived from mean-filed theory, which can technically apply to the network which has no loop and degree correlation. They are not tiny assumptions, and then we need to confirm the properties of the optimal network by numerical simulations. We use three type (SF:scale-free network, RND:random network, EXP:exponential network) networks, which has 500 nodes, to compare with the optimal network by GA. Initially, all nodes have the state 0 and we change the state of only one node from 0 to 1. We run same simulation 1000 times and observe the average final proportion of agents choosing 1.

Simulation results

We plot border point, on which we observe the cascade which spreads into the entire network at least one time out of 1,000 trials (see Fig.6). It is very clear that the optimal network by GA has the largest cascade window compared with other network topologies. The results on all networks reflect the trend of the theoretical cascade window. We summarize these results with same manner of López as

$$\phi^*_{\mathrm{SF}} < \phi^*_{\mathrm{EXP}} < \phi^*_{\mathrm{GA}} \qquad (6)$$

Topology of evolutionary optimized networks

The degree distribution of evolutionary optimized networks(GA networks) show that the network has a wide range of nodes in terms of degree(see Fig.7(a), Fig.8(a)). Note that, in the degree distributions, not merely a large proportion of nodes having small degree is to be found, but also hub nodes with a certain size are to be found.

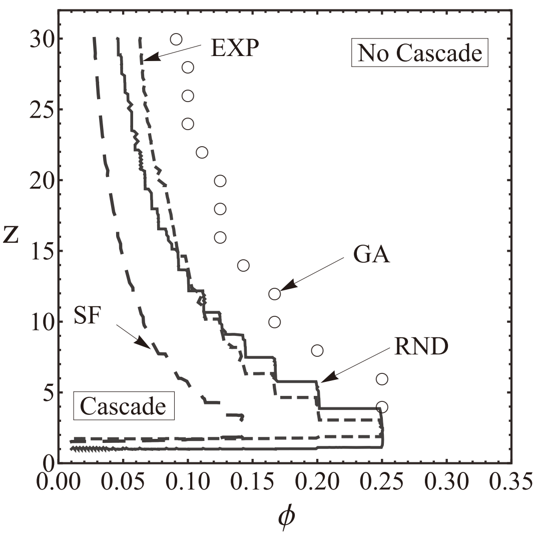


Figure 5: The cascade window of evolutionary optimal network for good cascade as a function of average degree Z and the threshold value ϕ of the agents

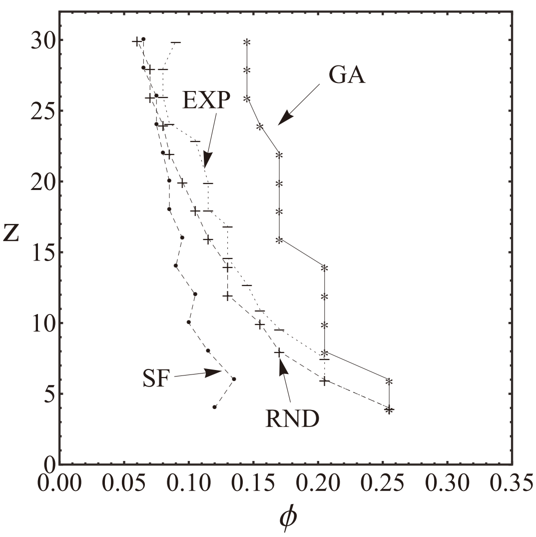


Figure 6: The cascade window by numerical simulations as a function of average degree Z and the threshold value ϕ of the agents

Fig.9 and Fig.10 shows how nodes in exponential and GA networks are connected to others. More preciously, we plot each point on (d_i, d_j) when node i with degree d_i and node j with degree d_j are connected. The diameter of each point is also proportional to the logarithm of the frequency of the cases. After that, the figure is converted to be symmetry for the only visualization. Although in the exponential networks, we found one big cluster and the center of it is near from $(\langle k\rangle, \langle k\rangle)$, in the GA networks, we found two cluster and the center of the cluster in the circle is located on (k_c, k_c) where $k_c < \langle k\rangle$. The nodes of the cluster in the circle have small number of links and the cumulative fraction of those nodes is about $80 - 90\%$ which can percolate the almost entire network(see Fig.7(b) and Fig.8(b)). These topological properties of the GA networks imply that the cascade occurs more easily on GA networks compared with the exponential network.

We also try to model the degree distribution of the GA networks by using a linear combination of two different degree distributions, because each cluster of GA networks may have a different degree distribution. From results of preliminary experiments(they are not included in this paper), we use Poisson and uniform distribution to model the degree distribution of each cluster in GA networks. The model is shown as:

$$p(k) = \frac{1}{A}\frac{\lambda^k e^{-\lambda}}{k!} + B \tag{7}$$

where A and B are controlling parameters for the network design.

We plot the degree distribution and complementary cumulative distribution function of the model on Fig.7 and Fig.8 where each coefficient was adjusted heuristically in this case. These distributions of the model fit the distributions of GA networks very well.

From the results obtained, the GA networks have a degree distribution which follows a linear combination of Poisson and uniform degree distribution and the distribution enables the network has a sufficiently large cluster, of which nodes have relatively small number of links and facilitate cascade phenomena.

Then, we obtain the relationship of the threshold of each network for cascade as:

$$\phi^*_{\text{SF}} < \phi^*_{\text{EXP}} < \phi^*_{\text{Poisson-Uniform}} \tag{8}$$

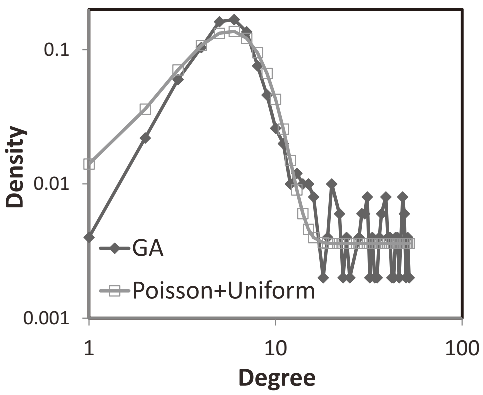
Conclusion

In this paper, we study changing the behavior of the application on networks exploiting the underlying network topology and, as an example problem, maximize cascade in terms of a region of cascade conditions, which is important driven forces for the diffusion of the innovation. Instead of a human-centric method, which is creative but not scalable, we propose evolutionary network optimization by genetic algorithm to obtain a network required. The approach we propose can find that networks with a linear combination of Poisson and uniform distribution have largest cascade window compared with exponential networks, random networks and scale free networks by the accumulation of the improvements.

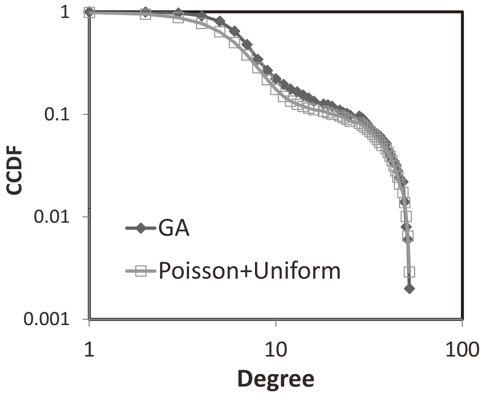
REFERENCES

Facebook (2011). Statistcs. *http://www.facebook.com/press/info.php?statistics*.

H.Sato, O.Isao, and K.Shigenobu (1997). A new generation alternation model of genetic algorithms and its assessment. *Journal of Japanese Society for Artificial Intellligence*, 12(5):734–744.

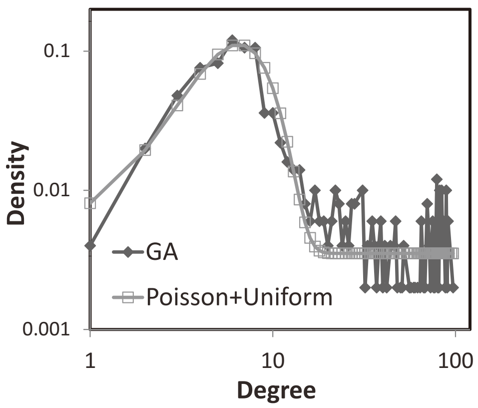


(a) Degree distribution

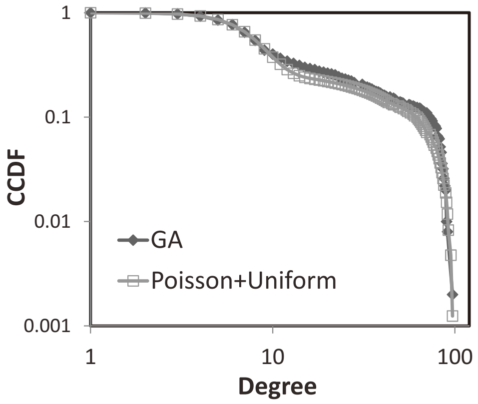


(b) Complementary cumulative distribution function

Figure 7: The density and the complementary cumulative distribution function of the degree of the evolutionary optimal network(GA) and the combination distribution of Poisson and uniform distribution($A = 1.2$, $B = 0.0036$ in Eq.(7)) as a function of degree k on logarithmic scale:Both of these distribution have same average degree $\langle k \rangle = 10$



(a) Degree distribution



(b) Complementary cumulative distribution function

Figure 8: The density and the complementary cumulative distribution function of the degree of the evolutionary optimal network(GA) and the combination distribution of Poisson and uniform distribution($A = 1.4$, $B = 0.0035$ in Eq.(7)) as a function of degree k on logarithmic scale:Both of these distribution have same average degree $\langle k \rangle = 20$

Java, A., Finin, T., Song, X., and Tseng, B. (2007). Why we twitter: Understanding microblogging usage and communities. *The Joint 9th WEBKDD and 1st SNA-KDD Workshop*.

López-Pintado, D. (2006). Contagion and coordination in random networks. *International Journal of Game Theory*, 34(3):371–381.

Pastor-Satorras, R. and Vespignani, A. (2002). Epidemic dynamics in finite size scale-free networks. *Physical Review E*, 65(3):035108+.

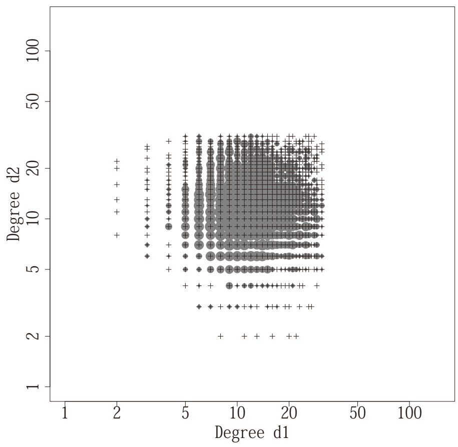
Watts, D. and Dodds, P. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34:441–458.

Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771.

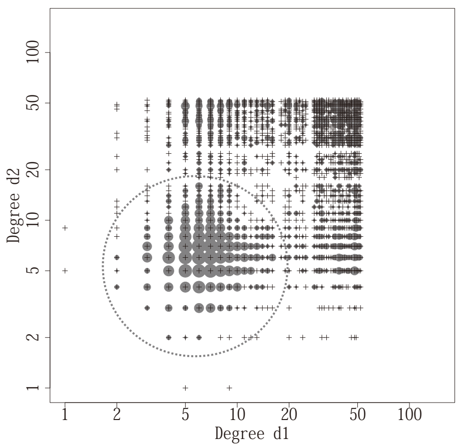
Young, H. (1993). The evolution of conventions. *Econometrica*, 61:57–84.

Young, H. P. (2009). Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American Economic Review*, 99(5):1899–1924.

Young, H. P. (2010). The dynamics of social innovation. Department of Economics, University of Warwick, Economic Theory Workshop.

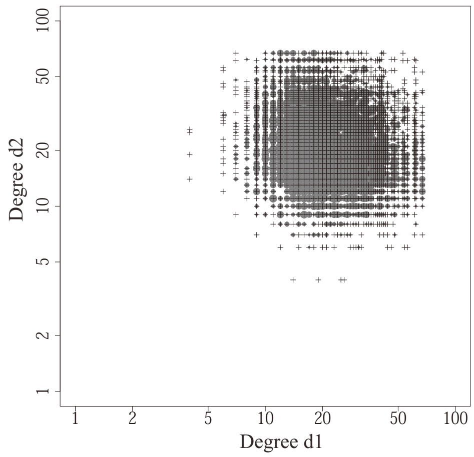


(a) Exponential network

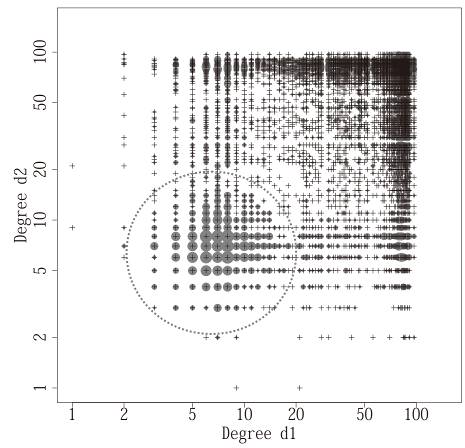


(b) GA network

Figure 9: This map shows the relationship of degrees between nodes, which are connected by links. The average degree of each network is $\langle k \rangle = 10$. The diameter of each point on the map is proportional to the logarithm of the frequency.



(a) Exponential network



(b) GA network

Figure 10: This map shows the relationship of degrees between nodes, which are connected by links. The average degree of each network is $\langle k \rangle = 20$. The diameter of each point on the map is proportional to the logarithm of the frequency.

AUTHOR BIOGRAPHIES

Takanori Komatsu holds the bachelor's degree of Science in Electrical and Electronic Engineering at Tokyo University of Agriculture and Technology in Japan and master's degree of Science in Computer Science and Engineering at National Defense Academy of Japan. Currently he is doing his PhD in Computer Science at National Defense Academy of Japan. His research interests include network design, network security and Multi-agents. His email is `ed10004@nda.ac.jp` and his personal webpage at `http://www.nda.ac.jp/~hsato/taka/`.

Akira Namatame is professor of Dept. of Computer Science National Defense Academy of Japan. He holds the degrees of Engineering in Applied Physics from National Defense Academy, Master of Science in Operations Research and Ph.D. in Engineering-Economic System from Stanford University. His research interests include Multi-agents, Game Theory, Evolution and Learning, Complex Networks, Economic Sciences with Interaction Agents and A Science of Collectives. He is a member of the AAAI, the IEEE, the SICE. His email is `nama@nda.ac.jp` and her personal webpage at `http://www.nda.ac.jp/~nama`.