



Practicality evaluation of stochastic model on networks for the real economy simulation

Hiromitsu Goto¹ · Wataru Souma¹ · Yuji Fujita¹

Received: 28 December 2018 / Accepted: 8 January 2019 / Published online: 19 January 2019
© Springer Nature Singapore Pte Ltd. 2019

Abstract

We investigate the practicality of the Bouchaud–Mézarid (BM) model for the real economy simulation using an information of Japanese production network, which contains basic financial information for more than a million firms and several million links of supplier–customer links. We found that the BM model can fit the power-law tail of the sales distribution and partially predict the sales for the individual firm.

Keywords Stochastic model · Real economy simulation · Japanese production network

Introduction

One of the most important but challenging issues is to realize the simulation of the real economy in real time. This is because the real economy results from the aggregation of the dynamics of micro level numerous agents such as consumers, firms and financial institutions; they form the multi-layered network roughly estimated in Japan of hundreds of millions of people, millions of firms and hundreds of banks. Inevitably, considerable computation should be involved to simulate whole system. In addition, it must be time-consuming to take into account interaction with non-economic condition such as social media. Although it might be achieved in coming decades by the supercomputer developments and the application of machine learning, there remains a need for an efficient model that can save the computational cost and provide a real understanding at micro level in terms of network structure.

Initial attempts focused on identifying the cause of macro phenomenon with power law by using stochastic process of micro units. While they can help us obtain

Supported by MEXT as Exploratory Challenges on Post-K computer (Studies of Multi-level Spatiotemporal Simulation of Socioeconomic Phenomena, Macroeconomic Simulations).

✉ Hiromitsu Goto
hiromitsu.goto.phys@gmail.com

¹ College of Science and Technology, Nihon University, Chiba 274-8501, Japan

a deep understanding of the phenomenon for the system, over-simplistic assumption tends to make this approach an impractical option in cases such as the real economy simulation. However, the practicality of the stochastic model using the real economic data at micro level remains unclear.

The purpose of this study is to revisit the stochastic model on networks and examine its practicality for the real economy simulation. The Bouchaud–Mézard (BM) model is one of the simplest stochastic models on networks, which was introduced to explain the wealth distribution with a power-low tail [1]. Although the application for complex networks has been studied in [2, 3], there has been no research to apply the BM model to actual economic networks. Therefore, we investigate the practicality of the BM model for the real economy simulation using an information of firm-level Japanese production network, which contains basic financial information for more than a million firms and several million links of supplier–customer links. The approach we use in this study aims to simulate the firm size such as sales in terms of only network structure.

This paper is organized as follows: first, we introduce the firm-level data of Japanese production network and review the BM model briefly. After fitting the power-law tail of the sales size distribution using the BM model, we check the accuracy of our simulation comparing with the real data for each firm. Finally, we summarize our results and conclude.

Method

Since it is difficult to acquire the data of material flow in general, we discuss the flow of money using the information of production network. Therefore, we simulate the distribution process of sales size with the BM model. The BM model includes the interactions of firms which depends on the weighting. To investigate the significance of the network structure to the simulation, our weighting is based on only the degree of firm.

Data: Japanese production network

The data were collected by Tokyo Shoko Research (TSR) on July 2016, which consists of two datasets of ‘TSR Kigyo Jouhou’ (firm information) and ‘TSR Kigyo Soukan Jouhou’ (firm correlation information). They contain basic financial information for 1,546,583 firms and 5,943,073 links of supplier–customer and ownership links. Although TSR collected the data based on questionnaires to firms, for which the maximum number of suppliers and customers reported is limited to 24, it does not mean that the number of links is limited to 24 because the links of large firms can be complemented by those suppliers or customers. For the sake of coherence, we followed the method to define the Japanese production network from TSR datasets used by the previous study [4].

Let us denote a money flow of supplier–customer link as $i \rightarrow j$, where firm j is a supplier to another firm i , and consider Japanese production network as directed

Table 1 Japanese production network structure: sizes of different components

Component	# of firms (with sales data)	Ratio (%)
GSCC	530,174 (521,419)	49.7
IN	278,880 (274,070)	26.2
OUT	219,927 (216,809)	20.6
TE	37,056 (36,312)	3.5
GWCC	1,066,037 (1,048,610)	100.0

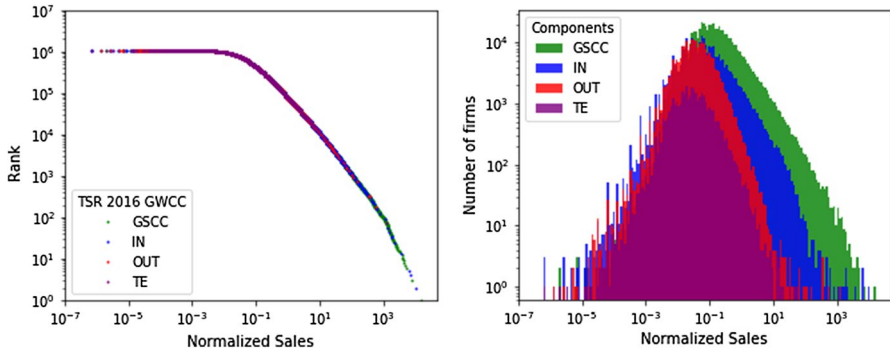


Fig. 1 Rank plot of the normalized sales and the distributions of each component

graph using TSR datasets. To exclude inactive and failed firms, we used an indicator flag for them in the firm information. Note that we eliminated self-loop and parallel links which are duplicated links recorded in the data. The network has the giant weakly connected component (GWCC), which corresponds to the largest connected component when viewed as an undirected graph. The GWCC can be decomposed into the parts defined as follows:

$$GWCC = IN + GSCC + OUT + TE, \tag{1}$$

where the giant strongly connected component (GSCC) is the largest connected one when viewed as a directed graph and the firms which can reach to and from the GSCC via a directed path are named IN and OUT components, respectively. The rest of the GWCC are named as tendrils (TE). We show the number of the firms each components in Table 1. It has been known that Japanese production network has a walnut structure [4].

As shown in Fig. 1, the distribution of sales normalized by its mean has power-law tail mainly composed of the GSCC, whose tail exponent can be computed as $\mu = 1.0223 \pm 0.0150$, where $p(x) \propto x^{-(1+\mu)}$ in large x . In general, the simulation on a directed network requires some extension to deal with a dead end as can be seen in PageRank [5]. Note that from the aspect of the money flow, the whole production network is also incomplete for the lack of the workers and consumers. Hence, we only focus on the GSCC hereafter. We show the cumulative distribution function for the in- and out-degree of the GSCC in Fig. 2 (left). As shown in Fig. 2 (right), it is

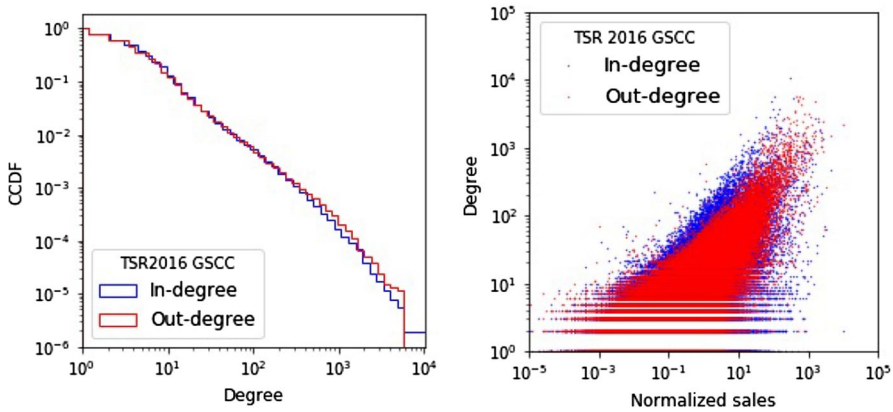


Fig. 2 Left: complementary cumulative distribution function (CCDF) for the in- and out-degree of the GSCC. Right: correlation diagram between the normalized sales and degree for the GSCC. The blue and red colors correspond to in- and out-degree, respectively

apparent that there is correlation between the degree and the sales size of firm. Since the simulation does not always reflect these features, we simulate using the weighting based on network structure for confirmation.

Simulation with Bouchaud–Mézard model

The BM model consists of the multiplicative stochastic process and the network effect from interactions of firms as follows:

$$X_i(t + 1) = a_i(t)X_i(t) + \sum_{j(\neq i)}^N (J_{ij}X_j(t) - J_{ji}X_i(t)) , \tag{2}$$

where $X_i(t)$ is the sales size of the i th firm and a_i is a Gaussian random variable $N(m, \sigma^2)$. The interaction coefficients J_{ij} constitute a dealing matrix from the j th firm to the i th firm. In this work, we assume a dealing matrix as

$$J_{ij} = JA_{ij}/k_j^{\text{out}} , \tag{3}$$

where k_j^{out} and A_{ij} is the out-degree of the j th firm and the adjacency matrix, respectively. In our weighting, Eq. (2) can be rewritten as

$$X_i(t + 1) = a_i(t)X_i(t) + J(\bar{X}_i(t) - X_i(t)) , \tag{4}$$

where \bar{X}_i is a local mean field of i th firm, which corresponds to the additive noise from network effects, defined as

$$\bar{X}_i(t) = \sum_{j(\neq i)}^N A_{ij}X_j(t)/k_j^{\text{out}} . \tag{5}$$

Hence, all the firms distribute $JX_i(t)$ equally to the supplier. Moreover, it has been known that the multiplicative stochastic process with additive noise obeys the power-law distribution [6]. Therefore, the additive network effect can be seen as additive noise in the previous literature to explains the power-law tail. Note that the additive term in our weighting is proportional to its in-degree; $\bar{X}_i(t) \propto k_i^{\text{in}}$.

To have a stationary distribution, we use the normalized sales by the mean $x_i(t) = X_i(t)/\bar{X}(t)$ where $\bar{X}(t) = N^{-1} \sum_k X_k(t)$. Under the mean approximation in large N limit, which is the case ($k_i^{\text{out}} = N$), the probability density function is known as

$$p(x) = \frac{(\mu - 1)^\mu}{\Gamma[\mu]} \exp\left\{-\frac{\mu - 1}{x}\right\} x^{-(1+\mu)}, \quad \mu = 1 + \frac{J}{\sigma^2}, \tag{6}$$

where the exponent μ is the Pareto index in economics. In general, therefore, the tail exponent of the normalized sales distribution can be fitted by only the parameter σ^2 and J in the BM model. After fitting the power-law tail, we compute the ratio between the simulation result of the BM model and the data for each firm, $r_i = x_i/x_i^{\text{data}}$ to estimate the simulation accuracy.

Simulation results and discussion

The main purpose of this work was to test the practicality of the simulation of the real economy using the simplified model. As mentioned previously, we sought to explain the sales size of firms by using only the BM model and the information of a real network. Numerical simulations were carried out for the GSCC of Japanese production network and repeated ten times using Eq. (2) with an initial condition $x_i(0) = 1$, a Gaussian random variable $N(1.01, \sigma^2)$ and various values of J . The distribution at $t = 10^4$ was taken as the stationary distribution.

As shown in Fig. 3, we found that the BM model can fit the power-law tail of the normalized sales distribution with $J = 0.003$, $\sigma^2 = 0.02$, and the ratio between the simulation result of the BM model and the data seemed to follow the log-normal

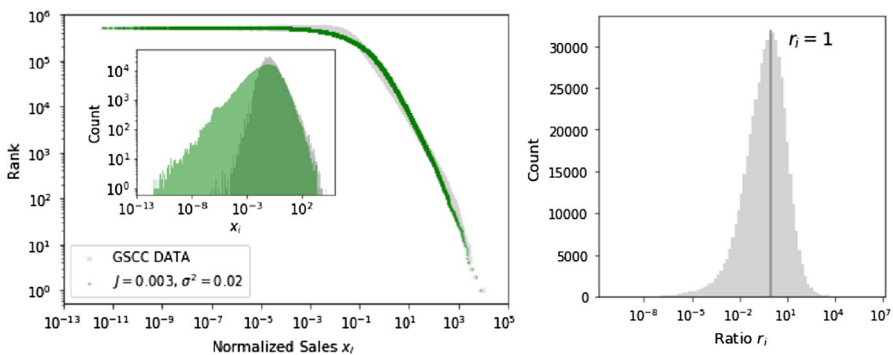


Fig. 3 Left: comparison of BM model (green) and data (gray) with respect to the distribution of sales. Right: histogram of the ratio $r_i = x_i/x_i^{\text{data}}$

distribution, whose central value was located at $r_i = 1$. On comparing the simulated distribution with the TSR data as shown in Fig. 3 (left), it appeared that the BM model underestimated the firms located in the middle range of sales as lower. Moreover, it is clear that the BM model is unfit for the purpose of simulation requiring high precision as can be seen from Fig. 3 (right). Although this is not an ideal result, it is surprising that the firms simulated with $0.5 < r_i < 2$ accuracy account for 19.8% of the total firms, despite that the stochastic model was only used for the information of network.

To improve our results, it is important to understand what kind of firms the BM model can describe accurately in our weighting. In Fig. 4, we visualized the ratio r_i in terms of the network structure. The white dot corresponds to $r_i = 1$. It changes to blue (red) as the ratio r_i becomes larger (smaller) than one. As shown in Fig. 4 (left), the ratio increases in a downward direction. It should be noted that the IN and OUT components of the GWCC are ignored in our simulation, and locate above and below the GSCC. Therefore, these results indicate the possibility of improvement by extension to the whole network. As shown in Fig. 4 (right), we found that the simulation with the BM model was better suited for firms which have larger in-degree. Nevertheless, it does not necessarily correspond to the sales. In future, reasonable results for specific firms might be obtained by choosing an appropriate weight.

Finally, we carried out the same simulation using the randomized network, whose in- and out-degrees were conserved. In addition, to compare the network effects from actual and randomized network, we used the same random variable $a_i(t)$ for

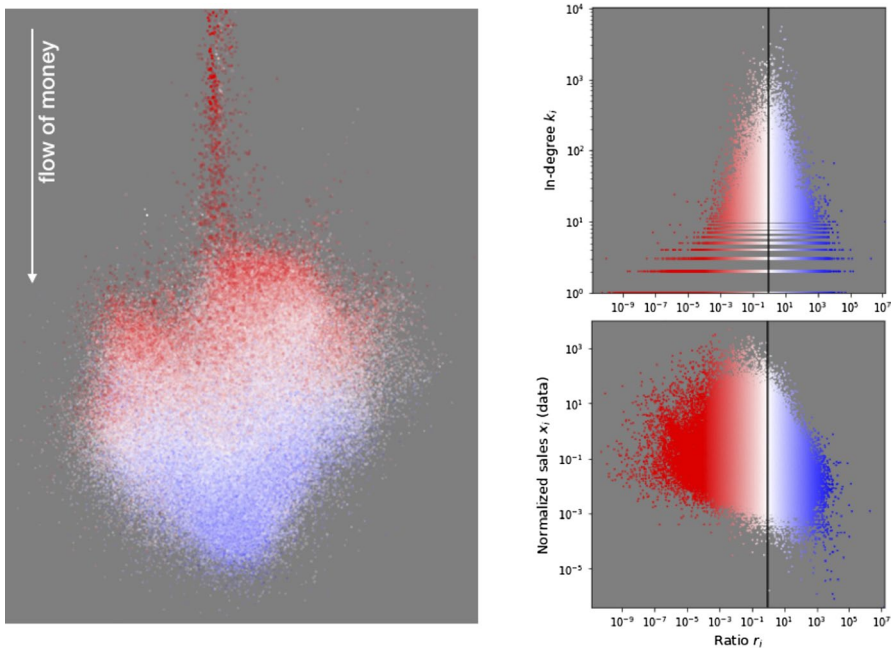


Fig. 4 Visualization of simulation accuracy. The white dots and the black solid lines correspond to $r_i = 1$ and they change to blue (red) as the ratio $r_i = x_i/x_i^{\text{data}}$ becomes larger (smaller) than one

the simulations. As can be seen in Fig. 5 (left), distributions were almost identical in both cases. This is because the additive noise from network effect is determined by the nearest-neighbor interaction in our simulation, which means each additive noise was proportional to its own in-degree. It could be inferred, therefore, that the correlation between the sales and in-degree shown in Fig. 2 (right) may have affected our simulation directly. We show the correlation between the results from actual and randomized network in Fig. 5 (right), which is gradationally colored by the difference in in-degree. It is evident that the results obtained from both networks have good agreement at large in-degree regions. This trend is in line with the previously discussed feature of our simulation.

Conclusion

We have investigated the practicality of the simulation of the real economy by using the BM model for the GSCC of firm-level Japanese production network. We found that the BM model can fit the power-law tail of the distribution of sales size, and the accuracy defined by the ratio between the result and the real data seems to follow a log-normal distribution, whose central value is located at one. Moreover, we confirmed that the accuracy of our simulation results reflects the correlation between the degree and the sales size. To the knowledge of the authors, this is the first result to evaluate the practicality of the stochastic model using real data. Although they seem to be unfit for the purpose of simulation requiring high precision, the results have potential to improve by changing our weighting.

The real economy simulation in real time can be achieved in coming decades by the development of high-performance computing. To implement the model on multi-layered network overlapping each others, it is important to partially approximate the system in accordance with the intended use; from the central bank's point of view, the interactions of firms can be replaced by the stochastic model. Therefore,

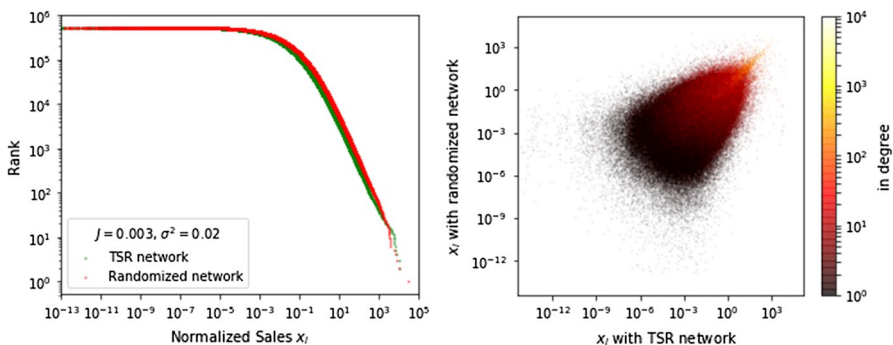


Fig. 5 Comparison of the simulation result between actual and randomized network. Left: rank plot of the normalized sales simulated by using the BM model with the actual (green) and randomized (red) network. Right: correlation between the results from the actual and randomized network gradationally colored by the difference of in-degree

our study can provide support for the modeling of the whole system to simulate the real economy.

References

1. Bouchaud, J.-P., & Mézard, M. (2000). Wealth condensation in a simple model of economy. *Physica A: Statistical Mechanics and its Applications*, 282(3), 536–545.
2. Souma, W., Fujiwara, Y., & Aoyama, H. (2001). Small-world effects in wealth distribution. eprintarXiv:cond-mat/0108482.
3. Ichinomiya, T. (2012). Wealth distribution on complex networks. *Physical Review*, E86, 066115.
4. Chakraborty, A., Kichikawa, Y., Iino, T., Iyetomi, H., Inoue, H., Fujiwara, Y., & Aoyama, H. Hierarchical communities in walnut structure of Japanese Production Network, RIETI Discussion Paper Series 18-E-026
5. Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1–7), 107–117.
6. Takayasu, H., Sato, A., & Takayasu, M. (1997). Stable infinite variance fluctuations in randomly amplified Langevin systems. *Physical Review Letters*, 79, 966.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.