

Facilities Collaboration in Cloud Manufacturing based on Generalized Collaboration Network

Wenxiang Li¹², Chunsheng Zhu³, Edith C.-H. Ngai⁴, Laurence T. Yang⁵, Lei Shu⁶ and Yuxia Sheng¹²

(1 School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China;

2 Engineering Research Center for Metallurgical Automation and Detecting Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China;

3 Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada;

4 Department of Information Technology, Uppsala University, 751 05 Uppsala, Sweden;

5 Department of Computer Science, St. Francis Xavier University, Antigonish, NS B2G 2W5, Canada;

6 Guangdong Provincial Key Lab. of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, Maoming, Guangdong 525000, China)

Email: liwx2006@hotmail.com; cszhu@ece.ubc.ca; edith.ngai@it.uu.se; ltyang@stfx.ca; lei.shu@lab.gdupt.edu.cn; shengyuxia@wust.edu.cn

Abstract — In cloud manufacturing for regional industrial cluster, there is increasing necessity of collaboration among enterprises or facilities. It is valuable to explore the characteristics of these collaboration behaviors for effectively scheduling dispersed manufacturing facilities and organizing their collaboration. The collaborative relation of manufacturing in regional industrial cluster can be described as a generalized social collaboration network. In this paper, we introduce the relevant entities and relations of facilities collaboration, and propose the method for building Facility Collaboration Network (FCN). We further design the dynamically growing process of FCN for different facility selection strategies, including random selection, balanced selection, random selection with preference and balanced selection with preference. Based on the metrics such as network scale, node degree distribution, act degree distribution, average shortest distance and number of cliques, we present the statistical characteristics of FCN, and analyze relevant characteristics and laws for efficient facilities selection in cloud manufacturing.

Keywords — *Social collaboration network; Complex networks; Cloud manufacturing; Scheduling; Clique; Degree distribution*

I. INTRODUCTION

The rapid development of economy globalization and socialized division of labor, with the increasing of complexity of products, leads to growing necessity of collaborative manufacturing among enterprises [1]. Regional industrial cluster consists of many enterprises of the same industry for a given region (urban agglomeration), and can coordinate resources utilization in these enterprises according to production processes. It can decrease cost and improve efficiency, and becomes a promising direction of global industry chain. On the other hand, Cloud manufacturing [2] is a new paradigm based on cloud computing. By combining advanced information technology and manufacturing technology, it can effectively manage and share resources according to users' demand, and provide networked manufacturing services by automatic configuration and

deployment. Cloud manufacturing provides an effective solution for cooperatively utilizing resources and creating high output value in regional industrial cluster.

There are strong and complex collaborative relations among enterprises and resources in regional industrial cluster, and large amounts of information from users, facilities, raw materials, processes and products is generated as process goes [3]. It is valuable to mine these collaborative relations and data, so we can predict, evaluate, plan and control the collaboration behaviors for convenient and effective cloud manufacturing services. The work will meet the demands for informatization and intellectualization of future industry.

The collaborative relation of production in regional industrial cluster can be described as a generalized social collaboration network [4]. In it the vertices may be enterprises, facilities or various resources, and edges are the collaboration relations of vertices. The network contains rich information in product design, logistics, manufacturing, marketing and service, and is an effective measure for describing and optimizing the industrial process. Now there are limited works on the collaboration relations of facilities in manufacturing and we need the model or tool for explaining and optimizing various phenomenon. In this paper, we take the enterprise cluster for packaging and printing production as an example, and propose the methods for analyzing the collaboration network. We further study the evolution mechanisms and statistical characteristics in Facility Collaboration Network (FCN), and explore the performance of collaboration modes for facilities in cloud manufacturing. The contributions are summarized as follows.

1) This work describes the entities and relations in collaborative production, and proposes the method for building FCN.

2) This work proposes the process of FCN's dynamic growth for different facility selection strategies, including random selection, balanced selection, random selection with preference and balanced selection with preference.

3) Based on the metrics such as network scale, degree distribution, act degree distribution, average shortest distance and cliques, we present the statistical characteristics of FCN.

The rest of this paper is organized as follows. Section II summarizes current work on production collaboration network. Section III describes the entities and relations for building FCN. Section IV introduces an instance of regional industrial cluster. Section V proposes the process of dynamic growth of FCN. Section VI analyzes the statistical characteristics of FCN. And section VII gives the conclusion.

II. RELATED WORK

Early work for production collaboration network focuses on cooperation of enterprises. By assigning enterprises or functions as vertices, [5] proposed a basic model of collaboration network. [6] pointed out that the demander-supplier network in US economy does not obey scale-free characteristics, and proposed the enhanced growth model to embody the real characteristics of demander-supplier relation. [7] considered such factors as the community characteristics of vertices, bi-directional preferred selection mechanism among vertices and the elimination of collaboration relations among nodes, and proposed weighted evolution model for collaboration network of service-oriented multiple industrial chains. Based on the relations among enterprises and services, [8] analyzed the statistical characteristics (node strength and clustering coefficient) of related weighted collaboration network, for building and managing service-oriented manufacturing network. [9] focused on the clustering characteristics of collaboration network for service-oriented enterprises, and proposed the relevant service architecture. All above work didn't cover the details of production implementation, and provided abstract description of connections among vertices, so they gave little insight on optimal configurations of resources.

Some works have considered the detailed implementation of production process. Liu Li-lan [10] allocated resources with Particle Swarm Optimization, and gave analysis on the Collaborative Manufacturing Resource network of scale-free characteristics. In [11] she also explored the characteristics of relevant statistical indexes and evolution formula based on the collaboration network of product-facility relation, and verified the scale-free distribution of this model by simulation. Further in [12], by assigning individual cloud service for a single subtask as vertices, she built complex network model for cloud manufacturing services. She also explored the statistical indexes such as degree, degree distribution, path length, betweenness of nodes and clustering coefficient. Further, she gave instructions on how to choose cloud services with an example of LED spotlight production. [13] proposed a network model of product development, and divided the elements in production into two categories, i.e., design nodes and resource nodes. By exploring the corresponding evolution model, statistical characteristics (including degree, clustering coefficient, betweenness and robustness) and physical characteristics of collaboration network, the authors also designed community discovering algorithm for load balance. They also verified the effectiveness of relevant methods based on the instance of steam turbine rotor production. Aiming at weighted collaboration network that is composed of resource

nodes in manufacturing grid, [14] explored the community structure and community segmentation method, and proposed an improved algorithm based on Newman fast algorithm. Currently, there is little work on the effect of facility collaboration modes on production, and there are few valuable advices on efficient collaboration methods among facilities. This paper explores the relevant problems based on generalized collaboration network.

III. FACILITY COLLABORATION NETWORK

The relevant entities for production in regional industrial cluster include instances of tasks, types of processes, types of subtasks, types of facilities and instances of enterprises. Each user proposes production requests to the interfaces of cloud manufacturing systems, and all requests are collected as task set $t=\{t_1, t_2, t_3, \dots\}$. Some tasks belong to the same process, so the relationship between task instance and process type is $M:1$. The cloud manufacturing system holds the knowledge of all process types and gets process type set $P=\{P_1, P_2, P_3, \dots\}$. Given subtask type set $S=\{S_1, S_2, S_3, \dots\}$, each process type is composed of several subtasks, and each subtask can take part in many processes, so the relationship between process type and subtask type is $M:N$. Given the facility type set $F=\{F_1, F_2, F_3, \dots\}$, each facility can handle several types of subtasks, and each subtask can be handled in many facility types, so the relationship between subtask type and facility type is $M:N$. All enterprises form the enterprise set $e=\{e_1, e_2, e_3, \dots\}$. Each enterprise consists of several types of facilities, and each facility type can be deployed in many enterprises, so the relationship between facility type and enterprise instance is $M:N$.

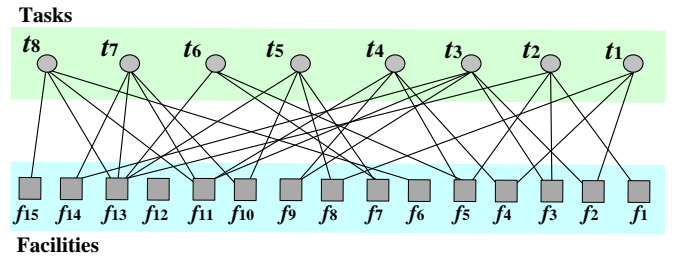


Fig. 1. Example of bipartite graph for the mapping of tasks and facilities

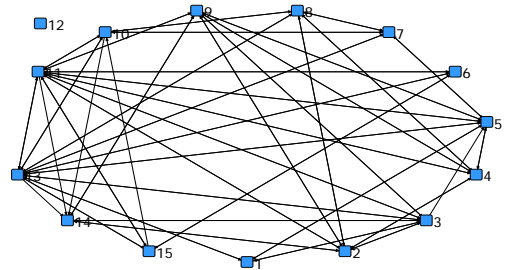


Fig. 2. Projection of bipartite graph for facility collaboration

Considering that a task can be accomplished by the collaboration of several facility types, we can construct a bipartite graph concerning the specific tasks and facilities labeled as $\{f_1, f_2, f_3, \dots\}$, and an example is shown in Fig.1. Based on this, we project the bipartite graph by facilities and

get a one-mode network. If two facilities involve in the same task, there is an edge connecting them. In this way we get an example of FCN in Fig.2.

IV. INSTANCE OF FACILITY COLLABORATION

We choose the case of packaging and printing manufacturing in an enterprises cluster of Hubei province of China from 2011 to 2013 as an example. There are 25 enterprises and 2007 facilities of 51 types in this cluster. With different ways of combination, 22 types of subtasks form 60 process types. There are relatively fixed order of subtasks handling for all processes, and there are also relatively fixed matching between facilities' functions and subtasks. So each enterprise with given facility configuration tends to locate in a relatively fixed position of production process, i.e., upstream, midstream, or downstream. There are exist the introduction of new facilities and the quitting of obsolete facilities. Taking one day as the time unit, the lifetime that each facility (represented by facility ID) is in operation is shown as the vertical strips in Fig.3.

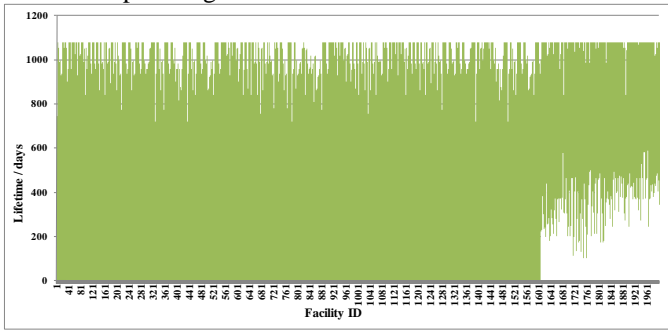


Fig. 3. The operation time span of all facilities

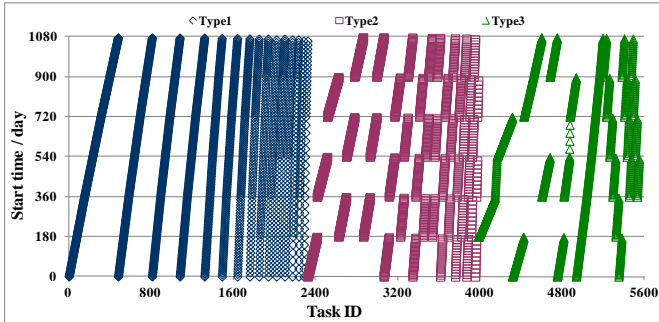


Fig. 4. The duration of each task

There are about 5500 task orders (represented by task ID) in these three years, and their durations are shown in Fig. 4. In it, the tasks that belong to the same process type are placed together and ordered by their starting time. From it we can see that all task demands can be divided into 3 types. Type1 is the constant daily demand that covers the span of three years with different number of tasks, and the range of task ID is about [1, 2300] in Fig.4. Type2 is the periodic demand that covers several months with fixed starting time of each year, and the range of task ID is about [2300, 4000]. Type3 is irregular temporary demand that covers uncertain time span and starts

randomly. The numbers of different types of tasks vary distinctly.

V. SELF-ORGANIZING OF FCN

The time and pattern for selecting facilities are determined by the following issues.

1) The characteristics of process, including the expected volume of output and the technical requirement for handling each subtask.

2) User's preference for some enterprises or facilities.

3) Current processing load of enterprises and facilities.

In this paper, we make analysis and comparison on four typical facility selection strategies:

1) **Random Selection (RAN)**: Select a facility randomly for a subtask among all available facilities with matching functions.

2) **Random Selection with Preference (RAN-P)**: Select a facility randomly for a subtask among the matching facilities in preferred enterprises. If there is no such facility in these enterprises, select a matching facility randomly from other enterprises. This strategy intentionally concentrates subtasks to some facilities.

3) **Balanced Selection (BAL)**: Select the facility with the least accumulative load among all matching facilities. This strategy intentionally builds dispersed connections for FCN.

4) **Balanced Selection with Preference (BAL-P)**: Select the facility with the least accumulative load among matching facilities in preferred enterprises. If there is no such facility in these enterprises, select a matching facility with the least accumulative load from other enterprises. This strategy intentionally concentrates the subtasks to some facilities, and builds dispersed connections for FCN.

Based on [15], we propose the task-driven self-organizing process for FCN as follows.

1) **Decomposition**: We decompose the task in processing (labeled as T_i) into $N_{S_i}(i)$ subtasks (labeled as S_{ij} , $0 \leq j \leq N_{S_i}(i)$).

2) **Selection**: Label N_F as the number of all facilities. Each subtask S_{ij} is allocated with a facility, and the probability of allocating facility F_k ($0 \leq k \leq N_F$) to S_{ij} varies according to different selection strategies.

RAN: Label $F(i,j)$ as the facility set that can handle S_{ij} , and the number of facilities in $F(i,j)$ is $N_F(i,j)$, then the probability of S_{ij} selecting F_k is

$$P_k(i,j) = \begin{cases} \frac{1}{N_F(i,j)}, & F_k \in F(i,j) \\ 0, & F_k \notin F(i,j) \end{cases} \quad (1)$$

BAL: Label $FMin(i,j)$ as the facility set with the least load in $F(i,j)$. If the number of facilities in $FMin(i,j)$ is $N_{FMin}(i,j)$, then

$$P_k(i,j) = \begin{cases} \frac{1}{N_{FMin}(i,j)}, & F_k \in FMin(i,j) \\ 0, & F_k \notin FMin(i,j) \end{cases} \quad (2)$$

RAN-P: The facility set that can handle S_{ij} and reside in preferred enterprises is labeled as $FP(i,j)$, and its number is $N_{FP}(i,j)$. So the facility set that can handle S_{ij} and reside in non-preferred enterprises is $F(i,j)-FP(i,j)$, and its number is $N_F(i,j)-N_{FP}(i,j)$. Given the probability $P_E(i,j)$ that S_{ij} is accepted by preferred enterprises, we have

$$P_k(i,j) = \begin{cases} \frac{P_E(i,j)}{N_{FP}(i,j)}, & F_k \in FP(i,j) \\ \frac{1-P_E(i,j)}{N_F(i,j)-N_{FP}(i,j)}, & F_k \in F(i,j)-FP(i,j) \\ 0, & F_k \notin F(i,j) \end{cases} \quad (3)$$

BAL-P: In $FP(i,j)$ the facility set with the least load is labeled as $F_{PMin}(i,j)$, and its number is $N_{F_{PMin}}(i,j)$. In $F(i,j)-FP(i,j)$, the facility set with the least load is labeled as $F_{NPMin}(i,j)$, and its number is $N_{F_{NPMin}}(i,j)$. So we have

$$P_k(i,j) = \begin{cases} \frac{P_E(i,j)}{N_{F_{PMin}}(i,j)}, & F_k \in F_{PMin}(i,j) \\ \frac{1-P_E(i,j)}{N_{F_{NPMin}}(i,j)}, & F_k \in F_{NPMin}(i,j) \\ 0, & F_k \in other \end{cases} \quad (4)$$

Iterate the above process for each subtask in T_i .

3) **Growth:** The facilities selected for T_i are labeled as $\{F_{k1}, F_{k2}, \dots, F_{ks(i)}\}$. Because many subtasks can be handled in one facility, we have $s(i) \leq N_{S(i)}$. Then we add new nodes to FCN.

4) **Connection:** Connect every node pair in $\{F_{k1}, F_{k2}, F_{k3}, \dots, F_{ks(i)}\}$, and get a complete graph.

VI. ANALYSIS ON CHARACTERISTICS OF FCN

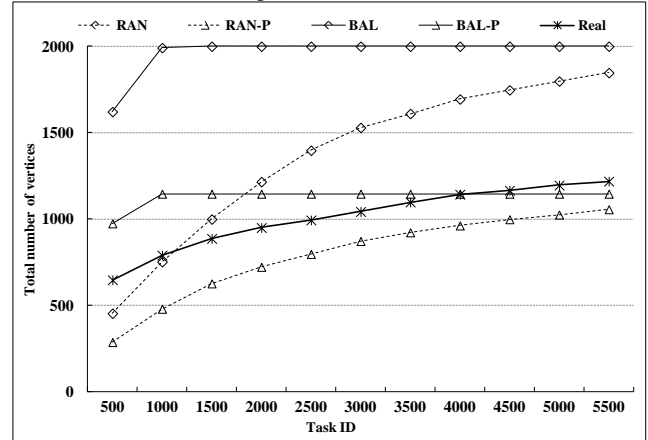
Among the about 5500 task orders, we designate a task scenario at the interval of 500 tasks, so we get 11 task scenarios when the number of tasks is 500, 1000, 1500, ... 5500. We conduct simulations of FCN growth with four facility selection strategies respectively. With the real statistics from production, we can get 5 FCNs (4 from BAL, BAL-P, RAN and RAN-P, and 1 from real data) for each task scenario. We explore the statistical characteristics of the following metrics for these scenarios.

A. Network scale

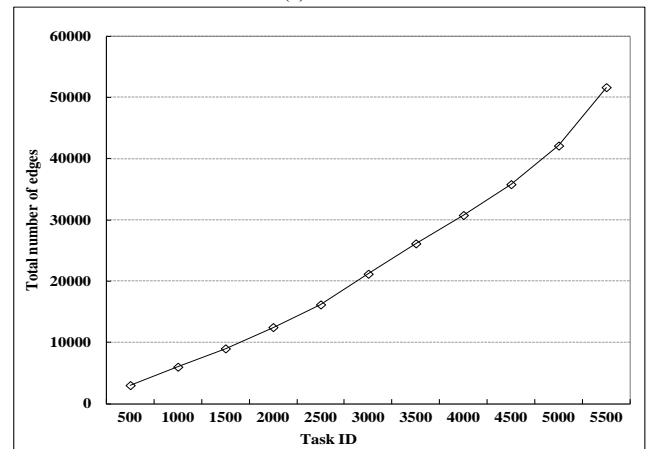
The numbers of vertices and edges as network grows are shown in Fig. 5. In Fig. 5(a), the numbers of vertices of load balanced strategies (BAL and BAL-P) reach maximum soon and then keep stable, while those of other strategies increase continuously with decreased acceleration.

Under the same condition, the numbers of vertices for BAL and BAL-P are much larger than those for RAN and RAN-P respectively. The reason is that load balanced strategies tend to utilize unused facilities, so all candidate facilities are utilized early. Because the rate of facility renewal is much lower than task occurrence, at late task scenarios there are few new vertices for BAL-P and BAL. However, random selection strategies utilize some facilities for many times, so more

unused facilities can be candidates for new vertices at late task scenarios for these strategies.



(a) vertices



(b) edges

Fig. 5. Growth of vertices and edges

The numbers of vertices for preferred selection strategy (BAL-P and RAN-P) are much smaller than those for non-preferred selection (BAL and RAN). The reason is that the selected facilities for preferred selection strategy amass in some enterprises, and they are utilized more frequently. The growth characteristics for real case (labeled as Real in Fig. 5) are similar to RAN-P with higher number of vertices. This illustrates that real case is the hybrid of RAN and RAN-P.

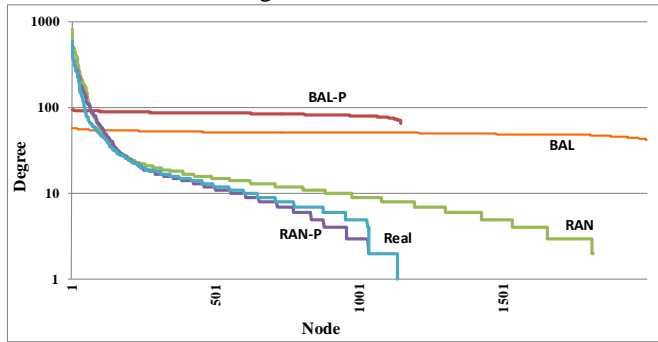
In Fig. 5(b), the numbers of edges for different strategies are identical at the same task scenario. Because the number of task orders increases steadily every year, the number of edges increases with slight acceleration as network grows.

B. Distribution of node degree and act degree

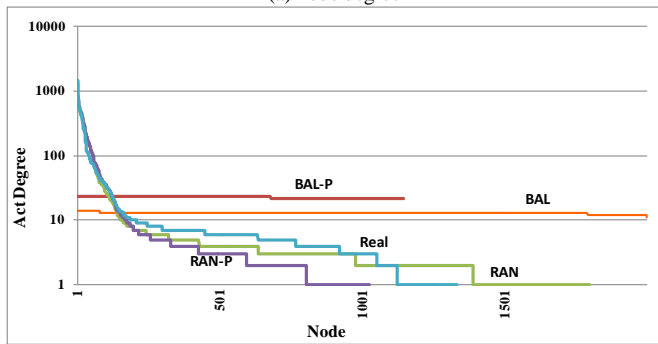
The node degree of generalized collaboration network denotes the number of neighbors that one node connects to. Act degree stands for the number of tasks that a facility joins in. The results the last task scenario are shown in Fig.6, in which node degree and act degree decrease as number of nodes increase.

The vertices have similar node degree and act degree for each load balanced strategy, so BAL and BAL-P achieve balance of accumulative load among relevant facilities, in spite

of the unbalanced numbers of different types of tasks. The facilities utilized in BAL-P are fewer than that in BAL, so the average node degree and act degree in BAL-P are higher than those of BAL. The distributions of node degree and act degree show the scale-free characteristic for RAN, RAN-P and real case, the reason is the distinct difference of numbers of different types of tasks in Fig. 4, i.e., some process types own large number of task instances and lead to more frequent utilization of some types of facility. The real case has the hybrid characteristics of RAN and RAN-P, so we don't discuss it in the following.



(a) node degree



(b) act degree

Fig. 6. Degree distribution

C. Average shortest distance

Given the vertices set V and the shortest distance d_{ij} from vertex i to j , the average shortest distance l of network shows the average distance from a vertex to another.

$$l = \frac{1}{N(N-1)} \sum_{i,j \in V (i \neq j)} d_{ij} \quad (5)$$

The l at 11 scenarios of FCN are shown in Fig.7. It can be observed that the l for RAN and RAN-P is always slightly higher than 2 with obvious small-world effect. However, the l drops to slightly higher than 2 from a large value for BAL and BAL-P. This shows the severe isolation of facilities for BAL and BAL-P at the early stage of network growth.

D. Clique

n -Clique is the maximal complete graph with at least n vertices, and embodies the highest level of connectivity for relevant vertices. We get the numbers of 4-clique, 5-clique and 6-clique in Fig.8.

The numbers of 4-cliques of RAN and RAN-P are similar. And in early task scenarios, the dispersed connection from BAL and BAL-P leads to low network density (the ratio of the actual number of edges to the number of all possible edges) and small number of Cliques, so there are much fewer 4-cliques of BAL and BAL-P. However, in late task scenarios, 4-cliques of BAL-P increases dramatically, the reason is the cliques' emergence when the network density increases above a threshold in BAL-P's FCN with much fewer vertices.

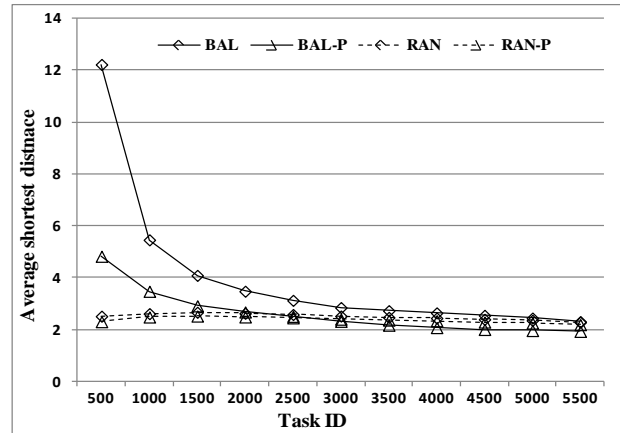


Fig. 7. Variation of average shortest distance

The numbers of 5-cliques of RAN and RAN-P are similar to each other and much larger than those of BAL and BAL-P. In early task scenarios, the numbers of 5-cliques of BAL and BAL-P are similar. However, in late task scenarios, 5-cliques of BAL-P increase dramatically for the same reason of cliques' emergence.

The numbers of 6-cliques of RAN-P are generally a little larger than those of RAN, for fewer vertices in RAN-P's FCN lead to higher network density. The numbers of 6-cliques of balanced strategies are very small, this illustrates that these strategies leads to much lower level of high order concentration.

VII. CONCLUSIONS

In this paper, we propose the method for building FCN with relevant entities in regional industrial cluster, and propose the dynamic growing process of FCN for different facility selection strategies. From analysis on simulation results and real data, we observe that the network scale for balanced selection is larger than that for random selection. And the network scale for preferred selection is much smaller than that for non-preferred selection. The vertices have similar node degree and act degree for each balanced strategy. While, the distributions of node degree and act degree show the scale-free characteristic for RAN and RAN-P. From the average shortest distances, RAN and RAN-P show obvious small-world effect. The numbers of n -cliques of RAN and RAN-P are generally larger than those of BAL and BAL-P with much higher level of concentration of nodes. In late task scenarios, 4-cliques and 5-cliques of BAL-P increase dramatically with emergency.

These conclusions provide useful insights for the design and management of cloud manufacturing services. FCN can

provide us with the typical facilities cooperation modes, and warn us to adjust the load of facilities with high node degree or act degree, or adjust the distribution of facilities in enterprises. Further, it is valuable to analyze the effects of various tasks occurrence patterns on facilities collaboration, and explore the effect of multiple connections for two facilities with the view of weighted network.

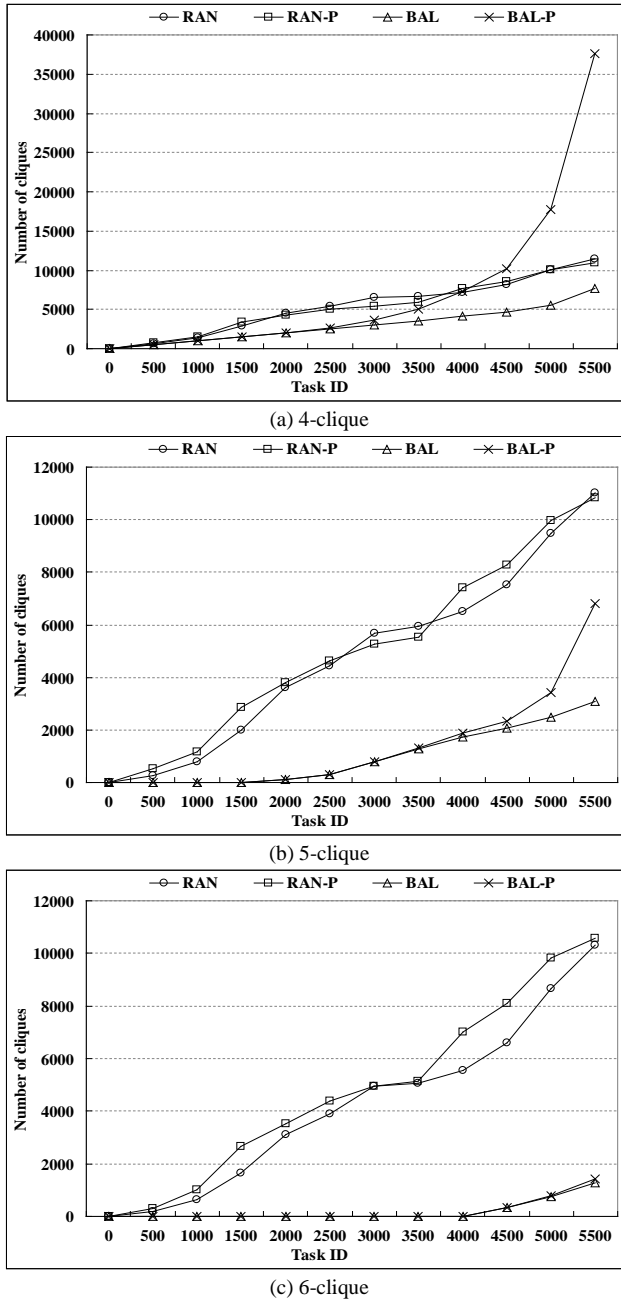


Fig. 8. Number of cliques as network grows

ACKNOWLEDGMENT

This work was supported by the Scientific Research Foundation for the Returned Overseas Chinese Scholars from State Education Ministry of China, the Science and Technology Research Project from Education Department of

Hubei Province of China (Q20141110, Q20151101), Engineering Research Center of Metallurgical Automation and Detecting Technology of Ministry of Education (Wuhan University of Science and Technology, China) (MARC201304, MARC201307). This work was also supported by funding from the Natural Sciences and Engineering Research Council of Canada. E. C.-H. Ngai's research was partly funded by the SSF ProFuN project and Vinnova GreenIoT project in Sweden. The work of L. Shu was supported by 2013 Special Fund of Guangdong Higher School Talent Recruitment, Educational Commission of Guangdong Province, China Project (2013KJCX0131), Guangdong High-Tech Development Fund No. 2013B010401035, 2013 top Level Talents Project in "Sailing Plan" of Guangdong Province, National Natural Science Foundation of China (61401107), and 2014 Guangdong Province Outstanding Young Professor Project. C. Zhu is the corresponding author.

REFERENCES

- [1] L. Gan, J. Zhang and J. Hong, "Research on partner selection of industrial cluster enterprises based on knowledge complementary", *WIT Transactions on Information and Communication Technologies*, 2014, vol. 56, pp. 255-262.
- [2] N. Liu, X.-P. Li and W.-M. Shen, "Multi-granularity resource virtualization and sharing strategies in cloud manufacturing", *Journal of Network and Computer Applications*, 2014, vol. 46, pp. 72-82.
- [3] X.-M. Yi, F.-M. Liu, J.-C. Liu, et al, "Building a network highway for big data: architecture and challenges", *IEEE Network*, 2014, vol.28, no.4, pp. 5-13.
- [4] L. Wang and Y.-H. Ma, "Competition and fitness in one-mode collaboration network", *Communications in Nonlinear Science and Numerical Simulation*, 2015, vol. 25, no. 1-3, pp. 136-144.
- [5] G.Schuh, L. Monostori, B.C. Csaji, et al, "Complexity-based modeling of reconfigurable collaborations in production industry", *CIRP Annals - Manufacturing Technology*, 2008, vol. 57, no. 1, pp. 445-450.
- [6] A. Enghin, H. Ali, R. James, et al, "Network structure of production", in *Proc. of the National Academy of Sciences of the United States of America*, 2011, vol. 108, no. 13, pp. 5199-5202.
- [7] J. Chen and L.-F. Sun, "Model of business-related multi-industrial chain collaborative network", *Computer Integrated Manufacturing System*, 2010, vol. 16, no. 5, pp. 1088-1095. (in Chinese)
- [8] J. Li, R. Mo and L.-L. Liu, "Analysis of service-oriented manufacturing network based on complex network", *Applied Mechanics and Materials*, 2013, vol. 271, no. PART 1, pp. 401-405.
- [9] F.-G. Zhang, P.-Y. Jiang, Q.-Q. Zhu, et al, "Modeling and analyzing of an enterprise collaboration network supported by service-oriented manufacturing", *Journal of Engineering Manufacture*, 2012, vol. 226, no. 9, pp. 1579-1593.
- [10] L.-L. Liu, Z.-S. Shu, X.-H. Sun, et al, "Optimum distribution of resources based on particle swarm optimization and complex network theory", *Lecture Notes in Computer Science*, 2010, vol. 6329, no. 2, pp. 101-109.
- [11] L.-L. Liu, X.-H. Sun, Z.-S. Shu, et al, "A production-collaboration model for manufacturing grid", *Lecture Notes in Computer Science*, 2010, vol. 6328, no. PART 1, pp. 166-175.
- [12] L.-L. Liu, C. Chen and T. Yu, "Characteristics analysis of cloud services based on complex network", *Communications in Computer and Information Science*, 2014, vol. 462, pp. 318-330.
- [13] L.-J. Fu, P.-Y. Jiang and C. Wei, "Modeling and performance analysis of product development process network", *Journal of Network and Computer Applications*, 2013, vol. 36, no. 6, p 1489-1502.
- [14] Y. Yin, Y. Li and C.-Y. Zhang, "Study on community structures in manufacturing grid and the algorithm for community partition of its resource nodes", *Applied Mechanics and Materials*, 2014, vol. 526, pp. 222-229.
- [15] J. J. Ramasco, S. N. Dorogovtsev and P.-S. Romualdo, "Self-organization of collaboration networks", *Physical Review E* 70, 03610, 2004.