Evolution-generated Communications in Digital Business Ecosystem

Juan Wang
College of Information Engineering
Shenzhen University, China
Email: juanwang@szu.edu.cn

Philippe De Wilde
School of Mathematical and Computer Sciences
Heriot-Watt University, UK
Email: pdw@macs.hw.ac.uk

***Abstract*—This paper discusses a new approach to model a complex evolving system, a Digital Business Ecosystem (DBE). We present a geography-based two coupled networks model. The upper layer is the business network layer regarded as an information ecosystem with multiple evolving agents. The lower layer is the underlying P2P communication layer to support communications. The two networks in the system are modelled as evolving over time and interacting with each other. In the business ecosystem, agents exchange services with each other. Through evolution processes, fit services reproduce and weak services die out. We study how these evolution-generated communications affect the topologies of both the business network and the underlying communication network and also system performance. We provide the details of the discrete event simulation and discusses the experimental results. The approach in this paper can be applied to model and analyze other complex real-world applications as our previous work [1].**

***Index Terms*—Evolutionary communications; Agent-based modeling; Business ecosystem**

I. Introduction

Business ecosystem is a relatively new concept introduced by Moore[2], who describes the increasingly interrelated nature of enterprises as competing and evolving organisms in a business environment. The significance of regarding business networks as ecosystems is that it highlights the importance of understanding the implication of a company's decision on its surrounding environment. The Digital Business Ecosystem (DBE) emphasizes the 'digital' aspect of the business cooperation[3]. It intends to provide an open source distributed environment, where software components, services, applications and also business models are regarded as 'digital species' that can interact with each other, reproduce and evolve according to laws of market selection.

In this paper, our primary concern is with the support for open communities of Small and medium enterprises (SMEs) so to aid them avoiding the ever increasing gap of technology between SMEs and large size enterprises. The DBE provides a platform for SMEs to publish their services as service provider clients of the system and also obtain their required services as service consumer clients of the system. In the ecosystem, SMEs manage and communicate with other SMEs. Services and products provided by SMEs are stored in their habitats. Through selection process, genetic algorithms and exchanges of services between the habitats, individuals of the SME network can benefit by combining their services with those of others, building new products and services so that they can expand their own business easily.

For large-scale complex evolving systems, agent-based modeling and simulation provides a way to model and better understand such systems[4]. The construction of a model is not an easy task. It needs to reveal the most important properties of the system and reflect real-world scenarios but still keeping the simulation possible. In this paper, we present a simplified model of the SME business network as an ecosystem, where numerous SMEs communicate with others to produce service chains and some services of high fittness values are exchanged between SMEs. Communications between different SMEs are supported and executed by a distributed lookup servers system called FADA, which will be described in detail in the following sections.

A lot of research work have presented a number of approaches of information and communication infrastructures to support business processes between different business enterprises. Some work provide basic information infrastructures [5] and some approaches are proposed to improve these basic infrastructures [6], and some work design middleware platform that supports the runtime creation and management of services [7]. Nevertheless, these infrastructures emphasize primarily on how to support and improve the integration of cooperative business processes among multiple enterprises, but do not take the effects of infrastructure itself on the business process into account. Despite business process and information infrastructure interact in practice, efforts on business process engineering (BPR) and information infrastructure are rarely conducted together. Some practitioners have started to pay attention to the correlation between business processes and information infrastructures[8]. However, the study is inadequate and further analysis on this relationship in more detail is needed.

In this paper, we construct a model of a real-world DBE system, consisting of a business network with evolution dynamics generated by the evolutionary environment and a P2P communication network to support the business network. Business processes induce communications between SMEs, which might affect the topology of the business network and be affected by the connectivity of the SMEs. On the other hand, communications between SMEs may in influence the topology of the underlying P2P communication infrastructure. This is a challenging problem.

The rest of the paper is organized as follows. The next section introduces a geography based two coupled networks model. Section III presents evolution-generated communications in the SME business network and section IV shows our experiment design and results. Finally, section V finalizes with the conclusions of the work and the direction of the future work.

II. A Geography Based Two Coupled Networks Model

In this section, we present a model for an example of a real-world distributed agent system - DBE. It is based on the following real-world scenario: SMEs may be providers or consumers of services, who are all clients of the system. SME provider clients use the system to submit their services descriptions. The system server aggregate and recombine these services into service chains and form complex services. Based on the feedback from users, successful service chains cross-over and mutate while useless service chains die. Thus, when SME consumer clients request services to the system, they can obtain the most suitable service according to their requirements and invoke the services in the SME providers end. The model of the system is a two coupled networks model, as illustrated in the Fig.1 below.

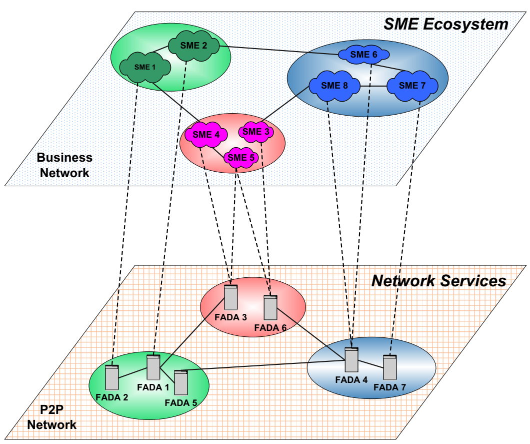

Fig. 1. Architecture of the two coupled network of the DBE

The upper layer is the business network layer, in which SMEs interact with each other for service exchange and migration. SMEs in different physical location have been clustered according to their region. We can regard the business network layer as an undirected graph, in which vertices represent individual SMEs and edges correspond to open connections between the SMEs.

The lower layer is the P2P agents network layer. In the system, SME provider clients who offered services have different physical location and built on disparate platforms, so it is difficult to develop a shared architecture for them. The proposed solution to this problem is SUN Microsystems' protocol of Federated Advanced Directory Architecture (FADA)[9]. It is ideal for communications in the business network environment. Furthermore, FADA is designed as a peer-to-peer dynamic distributed system in which different lookup servers (FADA nodes) work together to provide lookup server function from any entry point of the system and access services through the network. In our model, the P2P communication network is based on the protocol of FADA. It can be viewed as an undirected graph where nodes represent different lookup servers which are called FADA nodes in the following sections, and edges correspond to virtual connections between the nodes.

In the real world, capacities and bandwidths of links between nodes located in different geographical regions are various. Business interactions happen more frequently between SMEs within the same geographical region, therefore SMEs are modelled to be virtually grouped according to their geographical locations. As FADA nodes in the communication infrastructure support the business processes of SMEs, they are also virtually grouped depending on the locations of the SMEs that they are serving. SMEs in a region can only communicate with FADA nodes which are in the same region as them.

Service provider agents need to use the system to publish their services. On the other hand, service consumer agents need to use the system to find the suitable services. Accordingly, the communication process and interactions between agents in the two networks in the model are shown in Fig.2 and the algorithm follows the steps below.

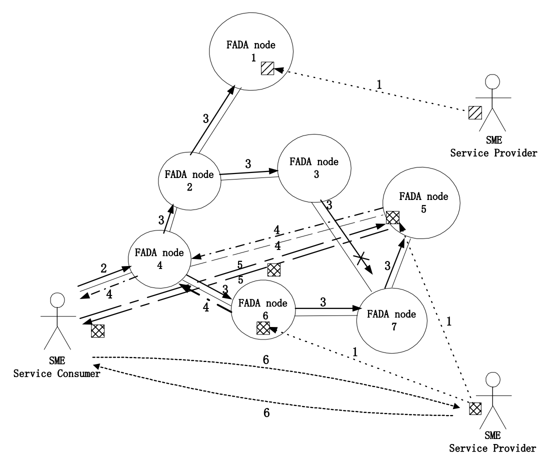

Fig. 2. Communications in the two coupled networks model.

Step 1.Once a service provider wants to publish a service, it writes a service proxy that can communicate with the real service. It then registers the proxy in at least one FADA node.

Step 2.Then a SME, as a service consumer joins the system. It sends a lookup request containing information such as price, functionality of a specific service to a FADA node which is called an initiator node of the request.

Step 3.The initiator node lookups in itself and uses flooding algorithm to propagate the request to its neighbors, and its neighbors' neighbors, limited by the TTL (Time To Live) and distance (number of hops away from the initiator).

Step 4.FADA nodes who find matches return these proxies to the initiator via a direct call instead of passing through previous nodes, which allows a fast response and results in a new connection in the network.

Step 5. Service consumer selects the best solution among the matches and contacts the FADA node which contains the selected proxy, and downloads the proxy from it.

Step 6. By using this proxy, the service consumer can communicate with the service provider for the real service.

III. Evolution-generated Communications in the SME Network

When a request is sent to, for example SME i, an evolving population in the service habitat of the SME i is activated. Service habitats of SMEs are the places where copies of pointers to services provided by their own SMEs and also other SMEs are stored. According to the requirements of the request, fitness values of different services are computed by using a fitness function. Through a set of simple genetic algorithm, an optimal solution is created. For the purpose of exchanging and sharing good services between SMEs, copies of pointers to services in the optimal solution generated from the service habitat of SME i are migrated to the service habitats of the SMEs which are neighbors of the SME i. For those habitats of SMEs that the copies of pointers have been migrated to, the migrated copies result in an evolution of populations of services in the habitats. The strategy of the evolution is that popular services will reproduce and less-used services will die out. Accordingly, the registrations of proxies of these services in the FADA network will be changed. The optimal solution, which contains the IDs of the selected services, is then sent to a randomly selected FADA node available to the SME i. The lookup procedure is then performed in the FADA network to look for appropriate proxies for the selected services. With the proxies, SME i can invoke actual services from different SMEs. The details of the algorithms above are described in the following subsections.

A. Evolution of Populations of Services in SMEs

In the SMEs business network of the two coupled networks model, SMEs exchange services with each other in order to achieve better profit. Inspired by evolutions of species in natural ecosystems, where fit species reproduce and weak species die out, populations of different services in SMEs can also evolve according to the fitness values of different services in this way. In order to define the rules for the evolution of populations of services, replicator dynamics in evolutionary game theory has been applied to the model. The replicator dynamics describes the evolution of frequencies of strategies in a population as in [10].

In the SMEs business network layer, each SME owns a service habitat, in which services of different business types exist. For instance, for a SME which is a travel agency, its habitat may maintain services such as airline, car rental, hotel, restaurant, cleaning services and so on. The providers of these services can be the SME itself or by other SMEs. When a request is sent into a SME, some services in the SME are selected to construct a service chain. The selection of services is based on their fitness values, which are depending on the differences between services and the request and also their usage histories. The usage history of a services means the number of times that has been used in the past. Therefore, if a SME can be guided by some rules so that it knows how to keep and reproduce services of high fitness values and get rid of services of low fitness values, solutions that the SME can offer may be better, benefiting not only the individual SME but also the whole DBE system.

Let p_i^j be the number of services of type i in population j (SME j) and x_i^j be the proportion of services of type i in population j. Therefore,

$$x_i^j(t) = \frac{p_i^j(t)}{\sum_{i=1}^{n} p_i^j(t)}, \tag{1}$$

where n is the number of types of different services in all populations and

$$\overrightarrow{x}^j = (x_1^j, x_2^j, \ldots, x_n^j). \tag{2}$$

At each time step, each service in the population of services in a SME is decided to be reproduced, deleted or maintained depending on an associate value v_e, which is changing over time. For instance, $v_{e_i}^j$ is associated with service type i in SME j. It indicates the probability of reproducing service i in j (when $v_{e_i}^j \geq 0$) and the probability of deleting the service i in j (when $v_{e_i}^j < 0$).

It can be noted that the value $v_{e_i}^j$ can be regarded as the payoff function, and it plays an important role in the population dynamics. As is in other large complex systems, interactive behaviours between individuals in the model in this research are difficult to observe and quantify. Therefore instead of defining a clear form for the payoff function, a function to compute the value of the $v_{e_i}^j$ for each service is defined as its evolution guidance.

As is in other large complex systems, interactive behaviours between individuals in the model in this research are difficult to observe and quantify. Therefore instead of defining a clear form for the payoff function, a function to compute the value of the $v_{e_i}^j$ for each service is defined as its evolution guidance. In view of both system aspects and the functionality of the SME network of providing services, $v_{e_i}^j$ is defined as a function of N, Δu and r. Here, N is the total number of services in a SME habitat, $\Delta u = u_i - \overline{u}$ is the difference between the usage history of service i and the average usage of all services in a SME habitat including copies of services migrated from other SME habitats, and $r = \frac{n_i}{n_{max}}$, where n_i is the number of service i and n_{max} is a fixed number to limit the amount of all service types in a habitat.

The considerations regarding to the function of $v_{e_i}^j$ are as follows: as the memory of each SME habitat is limited, evolutions of services in SME habitats have to take N into account so that the size of the population will not exceed the memory limit. Moreover, the amount of services that have been used for more times should be larger than other services in the population, which is reflected by the variable Δu. However, in order to maintain the diversity of types of services and also avoid the case that a few strong services occupy the whole population, the number of services of a certain type in a population should be limited, as represented by r. In a word, the function for $v_{e_i}^{j}$ is defined through a set of comprehensive considerations. It enables services in SMEs habitats to evolve 'rationally'.

B. Optimizations of Solutions to Requests

1) Description of services: Each service in a SME habitat is primarily described by two entries as following:

- busiType: describes the business type of a service, such as representing the airline name, flight destination, hotel accommodation and so on. It is presented by a randomly chosen number in
- SM: representing the description of a service in Business Modelling Language (BML).
- history: the number of times that a service has been used.
- smID: ID of a service manifest.

2) Description of requests: It is modelled that at each time step, one request is created and sent to the system looking for services to be provided by SMEs. In a request, required services is composed of the following entries:

- requestID: the ID of a request.
- smeID: the ID of a randomly chosen SME which a request is sent to.
- chainLength: the length of a request, which counts the number of services required.
- busiTypes: the business types of services required in a request.
- requiredSM: the description of services required in a request
- requestTime: the record of the time step when a request is sent to the system.

3) Optimizing solutions to requests: When a request is sent to a SME, the SME creates an evolving population, in which corresponding services are aggregated to construct a number of service chains to the request. A simple optimization process by using crossover genetic algorithm is applied to the service chains generated in the evolving population in order to provide the optimal solution to the request. The procedure of answering a request in the DBE is illustrated in Fig. 3.

In the system, for instance, SME 1 sends a request 1 to the DBE system. The request is in the form of a service chain, and it contains the information of the types and the service manifest descriptions of the required services. The DBE system sends the request to a suitable SME, e.g. SME 2. In Fig. 3, each shape represents the business type of a specific service, and each color represents the service manifest descriptions of a specific service. SME 2 receives the request 1 from the DBE system, and it generates an evolving population corresponding to the request 1. In the evolving population, it first collects all the services whose business types are the same as the required services in the request, but these services are of different service manifest description. Then the SME 2

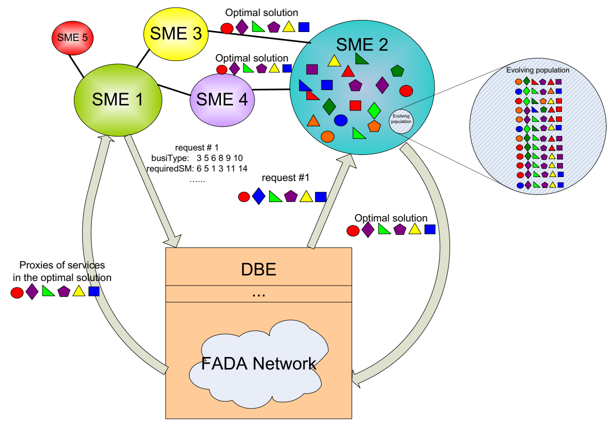

Fig. 3. The procedure of answering a request in the DBE

generates a number of possible solutions by aggregating these services into service chains. Due to the memory limitation of SMEs, the number of solution chains generated in the evolving population in SME 2 is bounded by S = 1000. SME 2 computes the fitness value of each solution chain. The solution with the highest fitness value is selected as the optimal solution to the request 1, and is sent to the DBE. The FADA network in the DBE then provides the proxies of the services in the optimal solution to SME 1, where the request is originally sent from.

The details of how the optimal solution is generated from the evolving population of SME 2 are explained as follows. The solution to the request is in the form of service chains. The fitness value of each solution is evaluated according to Eq. 3:

$$f(i) = \alpha(\exp(-\frac{d(i)}{\sigma})) + \beta h(i), \tag{3}$$

where

$$d(i) = \sum |d(i,j)| = \sum |s(i,j) - r(i,j)|, \tag{4}$$

and

$$h(i) = \sum h(i,j). \tag{5}$$

Here $d(i,j)$ is the measurement of the difference between a service provided in a solution and a service required in a request and $h(i,j)$ is the usage history of a service provided in a solution. The α, β and c are system parameters. Eq. 3 implies that a relatively high fitness value is assigned to a solution which contains services that not only are similar to services required in a request but also have good usage records.

Due to the memory limitation of SMEs, the number of solution chains generated to a request is limited. In order to produce the optimal solution among these solutions in an evolving population of a SME, some genetic algorithms is needed to the system. In genetic algorithms, crossover is a genetic operator which combines two parents to produce a new child, which may be better than the parents by possibly taking the best characters from each of the parents [11]. In this paper, the simplest type, the one point crossover is used in the model. It is operated by randomly selecting a crossover point within a chromosome and then interchanges the two parents chromosomes at this point to produce two new children. The fitness values of two new solutions represented by two new children generated from crossover operation are computed by using Eq. 3 and new solution replace the existing solution of the lowest fitness value if its fitness value is higher.

IV. Simulations and results

The simulation procedure consists of the initialization stage that builds up the initial state of the network topology and evolving stage that models the dynamic topology network.

Recently, topological theory in complex networks has been widely applied to complex agent networks. The random graph is one of the well-known topologies, which enables the analysis of emergent properties of a system. Random graph was introduced by Erdös and Rènyi[12], in which edges are distributed randomly and the presence or absence of any edge between two nodes in the network is dependent on a fixed connection probability p.

Stage I (initialization):

- The FADA network is set up by using the ER random graph model with a fixed connectivity probability p_{FADA}.
- SMEs initially register their services in the FADA nodes which are available to them. The number of FADA nodes in which a service can be registered is denoted by n_{reg}.

Stage II (evolving):

- As time evolves, SMEs join and leave the system frequently. The SMEs are modelled *joining* the system at rate r_{join}. FADA nodes are generated to serve new SMEs accordingly, with $n_{edg} = p_{FADA} \cdot N$ (N is the size of the FADA network) new links generated to connect new FADA nodes to others in order to maintain the ER random graph with fixed p_{FADA}.
- At each time step, each FADA node checks leases of proxies that are registered in it, where the lease of a proxy is the amount of time allocated to a proxy indicating how long a service proxy can be kept in the FADA. If the lease of a proxy is to expire, the node has to make a decision whether to renew it or not, depending on the remaining space of the node.
- The requests are generated at rate r_req per time step. When a request is sent to, for example, SME i, an evolving population in it is activated. According to the requirements of the request, fitness values of different services are computed by using the fitness function (defined in Eq. 3). Through a set of simple genetic algorithm, an optimal solution is created.
- For the purpose of exchanging and sharing good services between SMEs, copies of pointers to services in the optimal solution generated from the service habitat of SME i are migrated to the service habitats of the SMEs which are neighbours of the SME i.
- For those habitats of SMEs that the copies of pointers have been migrated to, the migrated copies result in an evolution of populations of services in the habitats. The strategy of the evolution is that popular services will reproduce and less-used services will die out. Accordingly, the registrations of proxies of these services in the FADA network will be changed.

Based on the description above, discrete-event simulations have been carried out. Primary simulation parameters are summarized in table I.

TABLE I
Parameters settings in the simulations.

Simulation parameters	values
connectivity probability of SME network p_{SME}	0.2, 0.4, 0.6, 0.8
connectivity probability of FADA networkp_{FADA}	0.2
maximum size of a proxy s_{proxy}	200
maximum lease time for a proxy $lease_{max}$	100
the number of FADA nodes one proxy registered inn_{reg}	3
maximum number of hops in lookup procedure $hops_{max}$	5
rate of sending requests, r_{req}	1

Through the simulations, a comprehensive analysis on the system can be implemented. As described in the previous section, communication flows generated due to the migration of copies of pointers to good services and the changes of registrations of services proxies are depending on the topology of the SME business network. Thus, it is interesting to study how the topology of the SME business network affects the system performance, which is our main concern here.

In the simulation, ER random graph model is applied to both the SME network and the FADA network. Due to the objectivity of the study of this section, the connectivity probability of the FADA network is fixed as $p_{FADA} = 0.2$, while the connectivity probability of the SME network p_{SME} is of various values as shown in table I. As the migration and the evolution of populations of services happen at all time, services that are kept in any SME habitat are changing accordingly. How does the value of p_{SME} affect the usage records of services maintained in SMEs' habitats? The result is shown in Fig. 4. It can be seen that as time elapses, the

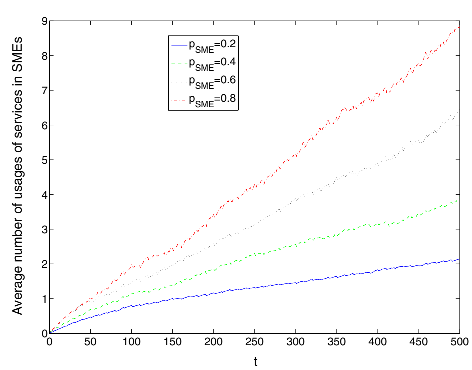

Fig. 4. The average number of usages of services in SME habitats in a SME business network of different connectivity probabilities.

average usage record of services in SMEs' habitats increases. It implies that SMEs are able to keep services which have good usage records and delete less used services. It is also shown that the higher the p_{SME}, the larger the average usage history of services in habitats is. This is because that if the p_{SME} is larger, popular services can be migrated to more SMEs' habitats. Furthermore, it is found the scaling law between the p_{SME} and the average usage of services in habitats is linear.

As illustrated above, the connectivity of SME network shows the impact on the usage records of services in habitats, does it also affect the fitness value of solutions? According to the fitness function defined in Eq. 3, fitness values of solutions are dependent on the differences between services provided by SMEs and services required in requests, and also the usage records of services maintained in the habitats. The potential solution whose fitness value is the highest is selected as the optimal solution. Here the fitness values of optimal solutions offered by the SME business network with different connectivity probability p_{SME} are investigated. The result is presented in Fig. 5. For the purpose of clarity, the results of

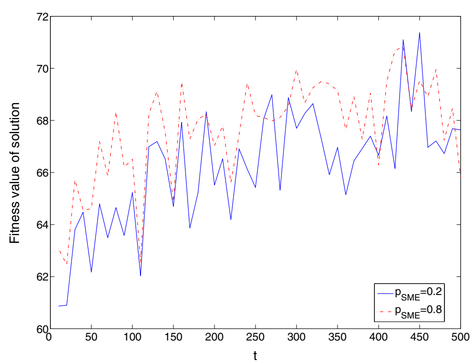

Fig. 5. The fitness value of solution offered by the SME business network of different connectivity probabilities.

$p_{SME} = 0.2$ and $p_{SME} = 0.8$ are only presented here. It is shown that the solution provided by the SME network of larger p_{SME} is better than of smaller one, benefitting from the higher average usage records of services in SMEs' habitats in the SME network of larger $p_{SME} = 0.2$ and $p_{SME} = 0.8$ are only presented here. It is shown that the solution provided by the SME network of larger p_{SME} is better than of smaller one, benefitting from the higher average usage records of services in SMEs' habitats in the SME network of larger p_{SME}.

V. Summary and Directions

In this paper, a complex evolving system, DBE, is modelled by a two coupled networks model. The SME business network is regarded as an information ecosystem, where behaviours of business processes such as exchanges of services between SMEs and optimizations of service solutions are modelled. The FADA network is the communication infrastructure which supports the business processes of the SME network. These two networks interact with each other and evolve over time.

Although business processes and communication infrastructure that supports the processes interact in practice, very few research has worked on how the changes on any of the two could have impacts on the other. The discrete-event simulation tool was used to show how to study and understand the dynamic relationship between the two. Our study here primarily focused on investigating the effects of graph connectivity of the SME network on the system. It is found that the graph connectivity of the SME network matters. The SME network is constructed and evolved based on the ER random graph model with various connectivity probability p_{SME}. With larger p_{SME}, the fitness values of solutions that the DBE system can offer are higher. Furthermore, it was found that the scaling law between the p_{SME} and the usage records of services in SMEs is linear.

In the presented geography-based two coupled networks model, we are keen to further study some other important aspects, including how the the effects of different types of topologies of the SME network on the system and on the FADA P2P communication network and also the scalability of the FADA communication network when it supports real-time business processes in the SME network. It is worth noticing that through the simulations in the model, the interrelation between applications in one layer and communications infrastructure in the other layer that the applications are running on, was examined. This approach can be extended to many other real-world applications, such as planning of communication network and power network in a city and so on.

Acknowledgment

This work was jointly supported by the EU FP6 funded Project Digital Business Ecosystem, contract number 507953 and by the Natural Science Foundation of China under the project number 60773203 and 60602066.

References

[1] J. Wang and P. D. Wilde, "Properties of evolving e-mail networks," *Physical Review E*, vol. 70, no. 066121, 2004.

[2] J. F. Moore, "Predators and prey: a new ecology of competition," *Harvard business review*, vol. 71, no. 3, pp. 75–86, May-June 1993.

[3] Digital Business Ecosystems (DBE) EU IST Integrated Project No 507953, Available http://www.digital-ecosystem.org [15 Jan 2008].

[4] E.K.Vuori, "Knowledge-intensive service organizations as agents in a business ecosystem," in *proceddings of 2005 International Conference on Services Systems and Services Management*, vol. 2, June 2005, pp. 908–912.

[5] Object Management Group. The Common Object Request Broker: Architecture and Specification, July 2002.

[6] S. Y. W. Su, H. Lam, and M. Lee, "An information infrastructure and e-services for supporting internet-based scalable e-business enterprises," *5th International Enterprise Distributed Object Computing Conference*, September 2001.

[7] P. M. Nirmit Desai and S. Tai, "Service communities: A structuring mechanism for service-oriented business ecosystems," in *2007 Inaugural IEEE International Conference on Digital Ecosystems and Technologies*, February 2007, pp. 122–127.

[8] R. J. Paul, "Collaborative information systems and business process design using simulation," *in Proceddings of the 37th Hawaii international conference on system sciences*, January 2004.

[9] Federated Advanced Directory Architecture (FADA), Sun Microsystems, Available http://fada.sourceforge.net [15 Jan 2008].

[10] J. Weibull, *Evolutionary game theory*. Cambridge, MA: The M.I.T Press, 1995.

[11] J. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: University of Michigan Press, 1975.

[12] P. Erdös and A. Rènyi, "On the evolution of random graphs," *Publ. Math. (Debrecen)*, vol. 6, p. 290, 1959.