Simulation

Simulation: Transactions of the Society for Modeling and Simulation International
88(5) 617–633

DOI: 10.1177/0037549711421942
sim.sagepub.com
SAGE

Simulating inter-organizational collaboration network: a multi-relational and event-based approach

Kang Zhao, John Yen, Louis-Marie Ngamassi, Carleen Maitland and Andrea H Tapia

Abstract

In this research, we study inter-organizational collaboration from the perspective of multi-relational networks. We develop an agent-based model to simulate how a collaboration network among organizations emerges from organizations' interactions through another network: the inter-organizational communication network. Our model adds links (or edges) into the collaboration network on the basis of events, which correspond to organizations' formation of collaborative teams for joint projects. The proposed approach also models the competitive yet non-exclusive dissemination of information among organizations, organizations' dynamic prioritization of candidate projects, and network-based influence. Applying the model to a case study of the humanitarian sector, we configure and validate the agent-based simulation, and use it to analyze how to promote inter-organizational humanitarian collaboration by encouraging communication. The simulation results suggest that encouraging communication between peripheral organizations can better promote collaboration than other strategies.

Keywords

agent-based simulation, humanitarian collaboration, information dissemination, inter-organizational network, multi-relational, network influence

1. Introduction

In recent years, inter-organizational collaboration has increased in both for-profit and non-for-profit domains.[1] Inter-organizational collaboration takes place when organizations share authority and responsibility for planning and implementing an action to solve a problem.[2] According to Guo and Acar,[3] inter-organizational collaboration occurs when different organizations work together to address problems through joint effort, resources, decision-making and share ownership of the final product or service.

Inter-organizational collaboration could benefit individual organizations in a community (e.g. improving the ability to address shared problems more effectively, higher potential for cost savings and organizational learning), clients of organizations in a community (e.g. receiving higher quality services or end products), and the community as a whole.[3,4] Research also identifies potential gains that not-for-profit organizations could reap from collaborating with others, including economic efficiencies, more effective response to collective problems, improvements in the quality of services, the spreading of risks, and increased access to resources.[5] As an important topic in organizational research, collaboration has drawn the attention of many researchers. A better understanding of inter-organizational collaboration may reveal ways to facilitate and improve collaboration activities.

While a social network can represent social relationships among individuals, the collaboration relationship among organizations can be denoted by an inter-organizational collaboration network. In such a collaboration network, a node represents an organization and an edge that connects two nodes means that the two organizations are collaborators. At the same time, collaboration is

College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, USA

Corresponding author:
Kang Zhao, College of Information Sciences and Technology, The Pennsylvania State University, University Park, 324 IST Building, University Park, PA 16802, USA.
Email: kxz134@ist.psu.edu

only one of the many types of relationships that could exist among organizations, such as information sharing, business transactions, etc. Also, the existence of other types of relationships among organizations may affect or influence the formation of a collaboration relationships and hence the inter-organizational collaboration network.

The goal of this research is to model the emergence of collaboration networks among organizations. Specifically, we model how organizations' interactions, such as exchanging information, influencing others' decisions and becoming influenced, through a communication network lead to the event of team formation and the subsequent emergence of collaboration networks. We use inter-organizational collaboration in the humanitarian sector as a case study.

In the past few years, the world has suffered from several major natural disasters. Humanitarian efforts after these tragedies have highlighted the need for greater levels of inter-organizational collaboration among humanitarian agencies. Such collaboration can improve humanitarian responses, so that they meet the needs of the affected population to the maximum extent possible.[6] In addition to better understanding of how inter-organizational a collaboration network emerges, we hope this study can also help to provide recommendations on how to promote humanitarian collaboration, which will eventually benefit disaster victims.

The remainder of the paper is organized as follows. Section 2 introduces the background of this research and how our approach differs from existing research. Then, in Section 3, we introduce our agent-based model for the emergence of inter-organizational collaboration network. In Section 4, we describe our implementation and validation of the model and our study of how to promote humanitarian collaboration. The paper concludes with discussions of future research directions.

2. Background

To promote inter-organizational collaboration, one approach taken by humanitarian agencies has been to organize 'coordination bodies', one of whose goals is to improve disaster relief efforts through collaboration among its member organizations. These coordination bodies may be temporary, special initiatives, or permanent incorporated not-for-profit organizations, and provide a non-hierarchical environment for organizations to interact with each other and form collaborative teams for joint projects. As participation in collaborative teams is undertaken on a purely voluntary basis, mutually beneficial joint projects and corresponding teams 'emerge' from the collective behaviors of individual organizations in this non-hierarchical setting.

In our previous interviews and surveys, many humanitarian organizations acknowledged that communication plays a very important role in the formation of collaboration relationships.[7] First, communication often antecedes collaboration and serves as the basis for establishing the future collaboration relationship. This is because organizations need to communicate with acquaintances to obtain information about different joint project initiatives, so that they can identify interesting projects and collaborate on them. Second, an organization's decision on whether to collaborate with others on a joint project is mainly based on its own evaluation of the project. However, through communication, organizations are often able to exert various levels of influence on others' decisions on collaboration. In other words, the collaboration network emerges from individual organizations' interaction through a communication network.

In addition, the analysis of the inter-organizational communication and collaboration networks also revealed the connection between communication and collaboration.[8] In our analysis of assortative patterns, which describes the tendency of nodes in a network being connected with nodes with similar degrees, inter-organizational communication network and collaboration networks have similar dis-assortative patterns. In other words, in both networks, high-degree nodes tend to connect to low-degree nodes. In fact, topologies of the two networks are positively correlated.

Therefore, to model the collaboration network among organizations, we have to adopt a multi-relational perspective and incorporate the impact of the communication network. From the perspective of network modeling, the research needs to simulate how organizations create links (or edges) in the collaboration network. Meanwhile, the decisions to create links are based on information and influence, which are transmitted through another network: the communication network. Hence, our work is related to the existing literature in two areas: (1) link formation in networks; and (2) the dissemination of information and influence. We briefly review the two areas in this section and discuss how our model is different from previous approaches. First we introduce the multi-relational perspective.

2.1 The multi-relational perspective

Connecting a group of nodes with edges (or links), networks are able to represent relationships among entities, such as people, organizations, locations, and so on. If we look at a network from a macroscopic level, an edge connecting two nodes in a network means that the two nodes are somehow related. However, scrutinizing how connected nodes are related to each other in the network, we find the heterogeneous nature of these edges. For example, in a social network, the most general relationships is 'know', i.e. a person knows another person if they have any relationship. A social network can then be specialized by categorizing the relationship as one of business, family, friend, and so on, and each of these can be further sub-classified. For example, Tom and Jack may be brothers with a link between them representing a family relationship; while Jack and David may be graduates of the same college, thus their link reflects an alumni relationship. Similarly, edges between two organizations in an inter-organizational network may denote different types of relationships, such as collaboration, information sharing, trading, etc.

In other words, for many real-world networks, edges in the same network could mean different types of relationships, each of which spans a network of its own. Thus, a network is multi-relational with multiple relationships and heterogeneous edges. The heterogeneity of edges can be further illustrated by representing such a multi-relational network with multiple uni-relational networks. Each type of relationship can be represented by its own uni-relational network. Given a network $G(V, E)$ with node set V and edge set E (note that as multiple types of relationships could exist between two individuals, we allow more than one edge between two nodes in G), we assume N types of relationships are represented by edges $e_i \in E$ in the network. Then we can divide all edges in set E into N disjoint sub-sets, which satisfy $E_1, E_2, \ldots, E_N \subset E$, $\forall i, j \in [1, N]\ (i \neq j) : E_i \cap E_j = \Phi$, and $E_1 \cup E_2 \cup \cdots \cup E_N = E$. Then we are able to divide $G(V, E)$ into N sub-networks that share the same set of nodes but have different sets of edges: $G_1(V, E_1), G_2(V, E_2), \ldots, G_N(V, E_N)$. Each of the sub-network is a uni-relational network that represents only one type of relationship in the aggregated multi-relational network $G(V, E)$.

Many studies on social and organizational networks took a uni-relational perspective. On the one hand, a lot of research focused on networks based on a specific type of relationship. Examples include the email communication among employees within an organization;[9] the collaboration among organizations;[10] the overlapping of board of directors among large corporations,[11] and so on. On the other hand, many studies did not make distinctions between different types of relationships. Instead, they often took a coarse-grained approach and aggregated multiple types of relationships into one network with homogeneous edges. For instance, in studies of online social networks, a link between two users in a social networking website, such as LinkedIn or Twitter, was often only considered to show that the two are somehow related but it was often disregarded whether the link is one of family, co-worker, classmate, etc.[12,13]

While uni-relational approaches are simple and intuitive, they inevitably lose valuable information. First, one uni-relational network may affect or influence another, as catalysts or constraints. For example, the classmate or roommate network among college students may affect their email communication network; the collaboration network among organizations may depend on the communication network; the transportation network may constrain a retailer's distribution network. Second, an individual may exhibit multi-faceted behaviors and possess different structural positions in different uni-relational networks. For instance, in an online social network such as Facebook, a user may have many online 'friends' (the friendship network) but their status or shared links seldom draw comments from others (a comment network); in an inter-organizational network, one could be a hub in the trading network but at the peripheral of the collaboration network.

By contrast, a multi-relational perspective can shed new light on the study of networks and help us to understand real-world networks in a more systematic way. For example, the structural difference between the sibling network and the farm work assistance network among villagers helped to explain several sociological phenomena.[14] Analyzing various types of networks among online gamers has validated the social balance theory in a large scale.[15] Ahmad et al.[16] inferred the trust network among online gamers from their mentoring network. The incorporation of multiple networks that social media users are involved in also helps to predict users' collective behaviors.[17] Nevertheless, previous multi-relational analysis focused on relationship between different networks' static topologies.

In this research, we adopt the multi-relational network perspective and study the emergence of an inter-organizational collaboration network. Thus, our approach needs to emphasize on how organizations' interaction through the communication network affects the dynamic growth, i.e. formation of links, of the collaboration network.

2.2 Formation of links in networks

Inter-organizational networks have drawn the attention of many scholars in organization and management science.[18,19] Specifically, many studied what drive organizations to form ties and identified important factors such as influence from other organizations, communication, history of relationship, reputation, and level of interdependency.[20,21,22,23] However, these works stayed at the organizational level and relied mainly on organizational characteristics.

At the network level, the link prediction problem has attracted researchers from different disciplines. On the one hand, many researchers studied various statistical attachment rules that guide a node's connections to another node. These models are usually based on connection heuristics and calculate the probability of connections between two nodes using the two nodes' topological attributes. Examples include the preferential attachment model,[24] Jaccard index,[25] Adamic/Adar index,[26] and Katz index.[27] Limited research used this approach to study inter-organizational networks.[10,28] On the other hand, some research used classification or maximum likelihood approaches to predict links. The basic idea is to consider

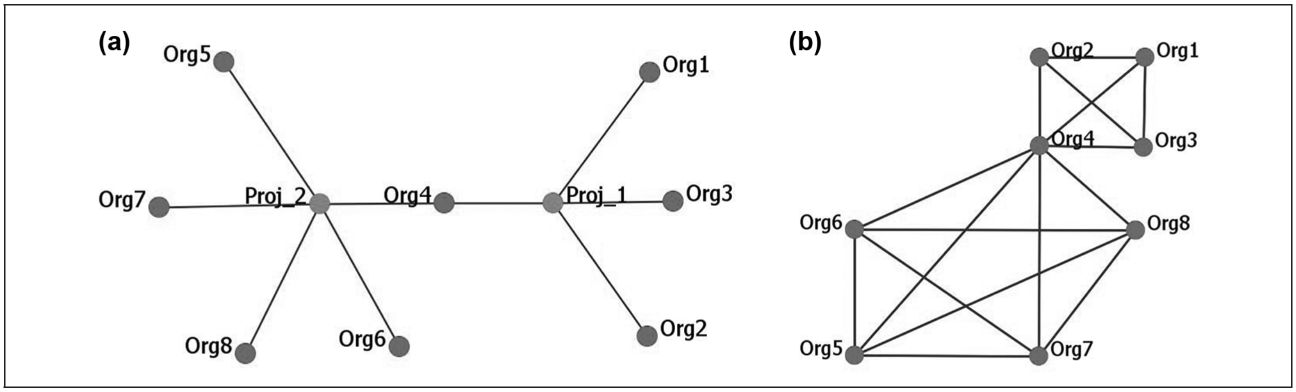

Figure 1. A sample inter-organizational collaboration network: (a) a bipartite graph for collaboration; (b) the collaboration network for the bipartite graph.

connected node pairs as positive samples, and other node pairs as negative samples. Then a classifier is trained[29,30,31] or parameters that maximize the likelihood of data are found[32] using a set of positive and negative samples.

However, most of the aforementioned approaches are rooted in building a dyadic or binary edge between two nodes. As a result, they may not capture two important aspects in humanitarian inter-organizational collaboration.

First, most existing approaches are limited to link formation within one network. In other words, they take a uni-relational approach and do not consider the interaction between different types of relationships. Thus, they are unable to capture how one network affects the formation of links in another, with the recent exception of Ahmad et al.[16] As we mentioned before, inter-organizational collaboration relies heavily on the communication relationship. Thus, our model has to incorporate the multi-relational perspective.

Second, for humanitarian agencies, collaboration is often project-based activities. The joint project they collaborated on could be joint training of staff members, coordinated data collection, shared database, and so forth. A joint project may start by only one or two organizations but could have more than two collaborators. According to our survey, more than 80% of the collaborative projects had more than three collaborating organizations.[7] The Information and Communication Technology (ICT) Skills Building Program of the ReliefTechNet (in this paper, pseudonyms of organizations are used to protect the confidentiality of humanitarian organizations), a major coordination body, is an example of such projects. The goal of this project was to provide training on latest ICTs to staff members of humanitarian agencies, so that their response to emergency and enhance their organizational effectiveness can be improved. This project was initially proposed by one organization in the coordination body ReliefTechNet, but was then developed with inputs and contributions from a team of more than 10 different member organizations.

From a network perspective, the formation of inter-organizational collaboration network is based on collaborative events. An event here is the formation of a collaborative team for a joint project. This type of event-based collaboration relationship can be described as a bipartite graph with two types of nodes: projects and organizations. Organizations are connected to a project if they form a team to collaborate on the project. Figure 1(a) shows a sample bipartite graph with two joint projects and eight organizations: four organizations collaborate on Project 1; Project 2 has five collaborators. The inter-organizational collaboration network can be constructed by converting the bipartite collaboration graph into a network that directly represents the collaboration relationship among organizations. In other words, organizations in the team for the same joint projects will be connected to each other (as shown in Figure 1(b)).

The important implication of the event-based collaboration is that the collaboration relationship is *n-ary*, instead of binary. In other words, several edges in the corresponding inter-organizational network may co-occur at the same time and are not independent of each other. For instance, if organizations in Figure 1(a) fail to form a team for Project 1, there might be no collaboration relationship among the participation organizations of Project 1. In other words, organizations 1, 2, 3, and 4 may not be connected in the inter-organizational collaboration network. Current statistical approaches focus on binary or dyadic relationship, but cannot capture this type of concurrent formation of multiple links and the dependency between links.

2.3 Dissemination in networks

A network often plays a key role for epidemic or social contagions to disseminate across multiple individuals.

Epidemic contagion concerns the spread of infectious diseases, such as flu. Examples include the susceptible–infected–recovered (SIR) model[33] and its variants. Social contagion refers to the diffusion of information, products, fashion, behaviors, and so on in the human society.

Most network dissemination models consider dissemination as an influence-based process. The basic idea is that dissemination happens between network neighbors as one node is able to influence its neighboring node to change status or adopt new behaviors. In terms of how to model such influence among neighboring nodes, there are two types of approaches. One approach is based on independent cascade between neighboring nodes.[34] When a node changes its state, it has one chance to 'infect' its neighboring nodes in other states with a probability. Another type of models is based on the idea of a threshold.[35] In such models, a node will change its state or adopt a new behavior when a certain fraction or number of other nodes have changed or adopted. In this case, the dissemination does not have to happen between neighboring nodes. The threshold of each individual may also vary.

However, existing approaches cannot be directly used in our model for two reasons. First, few dissemination models consider individuals' endogenous factors, which are especially important in social contagion. While exogenous influence from one's peers in a network can often affect whether they are 'infected' by a behavior, information or product, such a decision is also based on one's individual characteristics or their independent evaluation of the 'infectant'. In fact, recent research claimed that the impact of peer influence is overestimated.[36] Instead, homophil[37] plays an important role in the dissemination of the service. Homophily, which represents the similarity among individuals, reflects the endogenous factors that may affect dissemination. Similarly, in the context of inter-organizational collaboration, our survey revealed that when organizations evaluates a candidate joint project, they consider both endogenous and exogenous factors, such as whether the goal of the project aligns with the organizational mission, whether similar projects have been on the organizational agenda, which and how many organizations have decided to collaborate on the project, etc. Therefore, we need a model that incorporate both endogenous and exogenous factors in the dissemination process.

Second, most studies only considered the dissemination of a single 'infectant' but paid little attention to the competitive yet non-exclusive dissemination of multiple 'infectants'. When organizations collaborate, multiple projects may be proposed as candidate joint projects. However, an organization cannot work on all candidate projects, because it has limited resources. Put another way, various candidate joint projects, whose information is disseminated through the communication network, are competing for organizations' resources. The interplay between these candidate projects will eventually affect the outcome of collaboration. A recent study[38] proposed a model for competitive diffusion of political standings in networks. Also, political standings are mutually exclusive of each other: one cannot support both Democrats and Republicans at the same time. However, this type of exclusiveness does not hold in the dissemination of projects information, as an organization can work on multiple joint projects if resources permit. Therefore, our model needs to capture both the competition and the non-exclusiveness of multiple candidate projects in the dissemination process.

3. Proposed approach

Based on existing literature and our needs, we chose to use agent-based models (and simulations) for this study. Computational models and simulations, especially agent-based ones, have been widely used to study a variety of social, organizational, and natural phenomenon.[39,40,41] Agent-based models are capable of simulating macro-level structures or patterns resulting from micro-level interactions and decisions of heterogeneous agents within complex systems.[42] An agent-based simulation is especially helpful for decision-makers and policy-makers in organizations, because it is often very difficult to manipulate organizations to evaluate the impact of a policy or a decision. A computational simulation for an inter-organizational collaboration network not only enables us to study the outcome of different policies, but also helps us to gain insights into the patterns and characteristics of inter-organizational collaboration at both micro and macro levels.

In this section, we propose an agent-based model for the formation of inter-organizational collaboration network. The model adopts the multi-relational perspective and simulates how agents' interactions through one uni-relational network (the communication network) lead to the emergence of another uni-relational network (the collaboration network). The model incorporates *the dissemination of project information* through the communication network, *organizations' decision-making* on whether to work on a candidate project in the context of network influence, and how the inter-organizational collaboration network emerges through *the event of forming multi-agent teams for joint projects*.

In the agent-based model, organizations are represented as agents and are embedded in a communication network $G_c(V, E_c)$. $V = \{a_1, a_2, \ldots, a_N\}$ is a set of N agents. The edge $e_{i,j} \in E_c$ denotes that agents a_i and a_j communicate with each other through the network. The weight of an edge $w_{i,j}$ denotes the strength of ties between the two agents. Here $P = \{P_1, P_2, \ldots, P_M\}$ are a set of M

candidate projects. Each agent a_i maintains a prioritized to-do list of projects: $L_i \subseteq P$. A project in agent a_i's to-do list $P_k \in L_i$ is the project that agent a_i would like to work on or collaborate with others. Agent a_i is called a supporter of project P_k if $P_k \in L_i$. Agent a_i also assigns project $P_k \in L_i$ a priority score $C_{i,k}$ that reflects how important this project is to the agent. Projects with higher priority scores are ranked higher in the list. The maximum number of projects S_i in a to-do list may vary from agent to agent. The limited size of to-do lists reflects the resource constraints, which are important in modeling the competitiveness and non-exclusiveness of different candidate projects.

After each agent initializes its to-do list, agent interaction starts as an iterative process with three steps in each iteration. Figure 2 shows the pseudo-code of the model.

The first step is the dissemination of information about candidate projects through the communication network. To find collaborators, an agent first needs to disseminate information about projects in its to-do list, so that other organizations are aware of these candidate projects. An agent a_i spreads project information by proposing the top-ranked project P_t in its list to its neighbors in the communication network. Here a_i's neighbors in $G_c(V, E_c)$ are defined as $B_i \subseteq V$, so that $\forall a_j \in B_i, \exists e_{i,j} \in E_c$.

The second step is the evaluation of candidate projects. Upon receiving a new candidate project $P_k \notin L_i$ proposed by its neighbors, agent a_i will evaluate the project using various criteria of its own and assign the project an initial priority score. As shown in Equation (1), the initial and independent evaluation score $C'_{i,k}$ is determined by the function *Score*, which is based on the characteristics of the project and agent a_i's evaluation criteria R_i. The function *Score* can be configured by the modeler to cater different scenarios:

$$C'_{i,k} = Score(P_k, R_i). \tag{1}$$

Moreover, as we mentioned before, other agents also exert various levels of influence on an agent through the communication network. The priority score that an agent assigns to a project is influenced by other agents' evaluation of the same project. We used a network influence model, which extends the social influence model of Friedkin and Johnsen[43] to handle the exogenous influence from the network. Similar to most influence models, this model also assumes that an agent knows priority scores that other agents assign to candidate projects. Although this assumption may not hold in all situations for collaboration, it is a reasonable one in a collaborative environment, such as a coordination body. As coordination bodies would like to foster communication and collaboration among their member organizations, they often host regular meetings and group discussions, which provide an open forum for member organizations to exchange information and learn about projects others would like to work on. Equation (2) describes how an agent's project evaluation, which is reflected by the priority score assigned to the

```
While (the simulation tick t < T) {
    /**Inter-agent communication and dissemination of project information**/
    For each agent a_i ∈ V{
        Initialize a to-do list L_i, where |L_i| = S_i;
        Calculate priority scores for P_k ∈ L_i;
        Identify P_t ∈ L_i, so that ∀P_k ∈ L_i, C_{i,t} ≥ C_{i,k};
        Send information of top project P_t to all a_i's neighbors B_i in G_c(V, E_c);
    }
    /**Inter-agent influence and projects prioritization**/
    For each agent a_i ∈ V{
        For every P_k ∈ L_i, update the priority score using Equation (2);
        assign a priority score using; Equation (1);
        Adjust L_i and only keep S_i projects with top priority scores;
    }
    /**Possible occurrence of team formation events**/
    Check the supporters SP_k ⊆ V for P_k ∈ P;
    If |SP_k| ≥ TH_k{ /**The team-formation event occurs**/
        All agents Ai ∈ SP_k form a team for project P_k ;
        ∀a_i, a_j ∈ SP_k, add edges e'_{i,j} to G_b(V, E_b);
    }
}
```

Figure 2. Pseudo code for the agent-based model.

project, is iteratively influenced by other agents' evaluations of the same project. The right-hand side of the equation consists of two parts:

$$C_{i,k}(t) = E_i \times F_i \times SC_k(t-1) + (1 - E_i) \times C'_{i,k} \quad (2)$$

where $C_{i,k}(t)$ is the priority score of candidate project P_k assigned by agent a_i at time t $(t > 0)$. Please note that before agent a_i receives project P_k, the agent does not assign any score to the project. In other words, $C_{i,k}(t) = 0$ for $t < t^r_{i,j}$, where $t^r_{i,j}$ is the time stamp when a_i receives the P_k for the first time.

The first part describes the exogenous influence. Here F_i is a $1 \times N$ vector that represents influences on agent a_i from all of the N agents in the communication network, including agent a_i itself. Elements in F_i are called influence indexes, with $F_i[j]$ representing the influence index of agent a_j over agent a_i. The sum of all influence indexes in F_i is 1, as shown in Equation (3)

$$\sum_{j=1}^{N} F_i[j] = 1. \quad (3)$$

Here $SC_k(t)$ is an $N \times 1$ vector that stores project P_k's priority scores assigned by all of the N agents at time t. Namely, $SC_k(t) = [C_{1,k}(t), C_{2,k}(t), \ldots, C_{i,k}(t), \ldots, C_{N,k}(t)]^{\mathrm{T}}$. Thus, the product of F_i and $SC_k(t-1)$ is a score that reflects agent a_i's combined consideration both its independent evaluation and exogenous influence from all other agents' evaluations of the same project P_k at time $t-1$.

The second part is agent a_i's initial and independent evaluation of project P_k. The initial evaluation score is kept because it is made independently by the agent without exogenous influence. This will generally serve as the basis for possible deviations of priority scores even though network influence exists.

The two parts are connected and balanced with the influence coefficient $E_i (0 \le E_i \le 1)$, which denotes how likely agent a_i's project evaluation is influenced by others. A higher influence coefficient means an organization is more subject to exogenous influence, while agents with a lower influence coefficient are more independent when evaluating projects and making decisions on collaboration.

The last step in each iteration is the adjustment of to-do lists. With priority scores for candidate projects from the previous step, an agent may add new projects with higher priority scores to its to-do list, re-evaluate and re-rank existing projects, or remove projects with lower priority scores from the list, as the list is limited in size. In the next round of the iterative process, an agent will again advocate for its top-ranked project and disseminate information about this project to its neighbors, even though this agent receives information about the project from another agent.

As we mentioned in Section 2.3, we need to model the competitive and non-exclusive dissemination of project information. Our design of to-do lists and the adjustment of project prioritization models such as dissemination, as projects compete for positions and rankings in a to-do list that can accommodates multiple projects.

As you can see, agents' interactions are mainly through the communication network. Then how will such interactions through the communication network affect the collaboration network? After the three steps in each iteration, the model will check whether the event of team formation occurs. With the help of inter-agent interactions through the communication network $G_c(V, E_c)$, a candidate project may disseminate to many agents and appear on the to-do lists of some agents. As each agent's to-do list is openly available to all of the other agents, when the number of a candidate project P_k's support reaches a project-specific threshold TH_k, supporters of this project will form a team to work on the project together. As we discussed in Section 2.2, The team-formation event achieved through the communication network also leads to the establishment of collaboration relationships among agents in the same team. Consequently, edges are added to connect all of the team members to each other in the collaboration network $G_b(V, E_b)$, where $e'_{i,j} \in E_b$ denotes the collaboration relationship between agents a_i and a_j. Take one of the teams in Figure 1(a) for example. Organizations (agents) 1, 2, 3, and 4 form a team for Project 1. Thus, six edges connect the four team members in the collaboration network in Figure 1(b).

4. A case study of the humanitarian sector

In this section, we use the proposed agent-based model to simulate the emergence of collaboration networks among organizations in the humanitarian sector. We implement the simulation using the Repast toolkit,[44] a Java library for agent-based simulations. Then we configure and validate the simulation with empirical data. We also conduct an experiment to evaluate the effectiveness of different strategies that aim at facilitating collaborations.

4.1 Configuration of the simulation

In order to implement a trustworthy simulation, we need to configure our simulation properly. We first configure the simulation using empirical data of humanitarian organizations' demographics and project preference. Then we validate the configured simulation by comparing the simulated inter-organizational collaboration network with the real-world network. If the two networks are not similar, we tweak the configuration and run the simulation again,

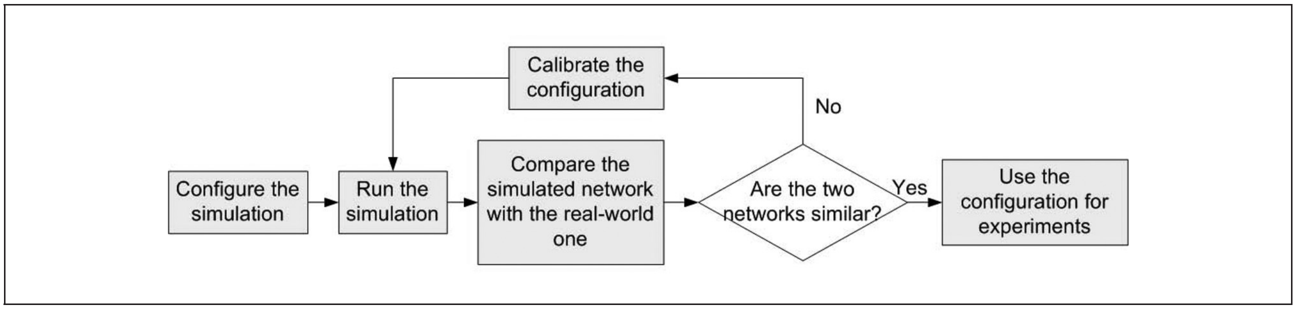

Figure 3. The process of simulation configuration and calibration.

until we obtain a configuration that leads to satisfactory results. Figure 3 illustrates the configuration process.

The configuration has to be backed by empirical data. Thus, we conducted two surveys and numerous interviews among member organizations of GlobalSympNet, a major coordination body with 119 member organizations. We collected these organizations' demographic data, including missions, focus regions, numbers of full-time employees. Table 1 lists the nine major missions and seven focus regions of member organizations in GlobalSympNet. Note that an organization may have multiple missions and more than one focus region. In addition, we also gathered data about 30 collaborative projects that humanitarian organizations worked on. The data includes where the project was implemented, the goal of the project, and how many organizations got involved, etc. We augmented data of the 30 projects and generated 300 synthesized candidate projects.

The surveys and interviews also helped us to understand how humanitarian organizations evaluate candidate collaborative projects and how their decisions are influenced by others. It is worth noting that, although we expected collaborative projects that were implemented immediately after disasters, most of the projects we found were pre-disaster projects whose goals are to improve humanitarian organizations' capabilities in disaster response and relief. Therefore, time pressure, which is very important in forming teams for post-disaster projects, does not seem to be a key issue when organizations evaluate pre-disaster projects.

Further, as our focus is on the multi-relational perspective, we need to find different types of relationships among organizations. We identified two uni-relational networks among organizations: communication and collaboration. Note that, the communication network is based on the inter-organizational relationship of advice seeking and giving. Research has shown that such advice exchange behavior often plays a major role in the exchange of information among humanitarian organizations.[45] More importantly, the advice exchanged among humanitarian organizations through the network is mostly about humanitarian projects, thus we can consider this network as a more focused communication network with stronger ties.

Table 1. List of missions and focus regions for organizations in GlobalSympNet.

Mission	Focus region
1. Provide food	1. Sub-Saharan Africa
2. Provide shelter	2. Middle East & North Africa
3. Provide water	3. Europe & Central Asia
4. Provide sanitation	4. South Asia
5. Provide medical care	5. South East Asia
6. Provide funding	6. North America
7. Provide information services	7. Latin America & Caribbean
8. Provide training and advice	
9. Provide IT infrastructure and/or applications	

In the collaboration network, two organizations are connected by an edge if they used to collaborate on humanitarian information management projects.

To find a proper configuration for the simulation, we simulate the collaboration network among 30 member organizations of the GlobalSympNet. The simulation takes as inputs the 30 organizations' demographic data, synthesized data of candidate projects, and the 30 organizations' communication network, which was identified in May 2008 (see Figure 4(a)). We run the simulation for 40 ticks, allowing 8 rounds of inter-agent interactions. After the simulation stops, we obtain a collaboration network among the 30 organizations.

Then we compare the simulated collaboration network with the actual collaboration network, which we gathered in a follow-up survey in October 2009 (see Figure 4(b)). To compare the two networks, we evaluate how close the simulated one is to the actual one using several metrics, including the number of total edges, the clustering coefficients, the average path length, and the accuracy of link prediction. On the basis of evaluation outcomes, we adjust and calibrate the configuration of the simulation and re-run the simulation until we find a satisfactory configuration.

Here we illustrate one such configuration. Our survey results suggest that when evaluating a candidate project, an

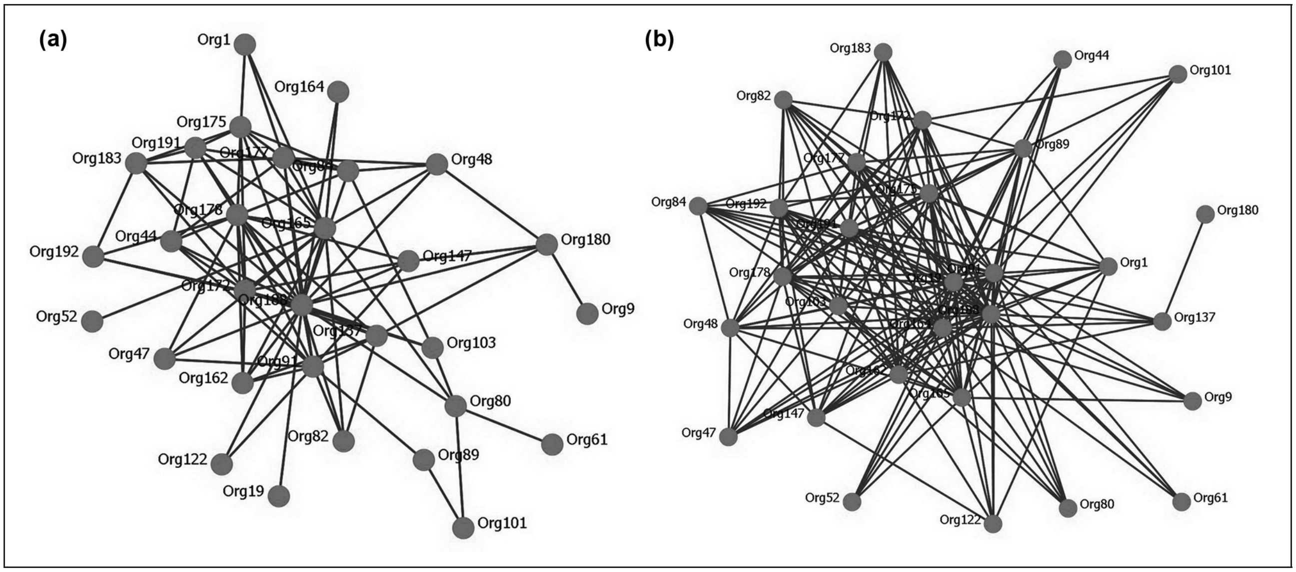

Figure 4. Inter-organizational networks among 30 organizations. (a) The communication network among 30 organizations of GlobalSympNet as of May 2008. (b) The collaboration networks among 30 organizations in GlobalSympNet as of October 2009.

organization considers two key factors: whether the purpose of the project matches the mission of the organization, and whether the beneficial (geographical) area of the project is within the organization's focus regions. Thus, agent a_i calculates the initial priority score $C^{'}_{i,k}$ for project P_k using

$$C^{'}_{i,k} = \alpha_m \sum_{m=1}^{9} \mathcal{H}_{\text{match}}(\mathcal{M}_i(m), \mathcal{M}_k(m)) + \alpha_r \sum_{r=1}^{7} \mathcal{H}_{\text{match}}(\mathcal{R}_i(r), \mathcal{R}_k(r)), \quad (4)$$

where

$$\mathcal{H}_{\text{match}}(x, y) = \begin{cases} 1 & \text{if } x = 1 \text{ and } y = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Here $\mathcal{M}_i(m)$ specifies whether organization a_i has mission m. If organization a_i has mission m, $\mathcal{M}_i(m) = 1$; otherwise, $\mathcal{M}_i(m) = 0$. Similarly, $\mathcal{M}_k(m)$ denotes whether m matches one of project P_k's purposes. Likewise, $\mathcal{R}_i(r)$ and $\mathcal{R}_k(r)$ refer to whether region r is within organization a_i's focus region and whether project P_k will be implemented in or bring benefit to region r, respectively. Equation (4) calculates the number of overlapping missions/purposes and the number of overlapping regions between organization a_i and project P_k. Coefficients α_m and α_r denote the relative importance of the two numbers. On the basis of organizations' average rankings of the two factors, we choose $\alpha_m = 5$ and $\alpha_r = 10$. Also, for the purpose of simplicity, we apply the two coefficients to all of the agents. The initial priority score $C^{'}_{i,k}$ that agent a_i assigns to project P_k is the weighted sum of the numbers of mission and area matches. Intuitively, the better the project's purposes and beneficiary regions match the organization's missions and focus regions, the higher the score is. This initial score also reflects homophily-based project selection, because organizations with similar missions and regions may prefer similar candidate projects.

In terms of exogenous influence, the survey results indicate that organization a_i is more likely to be influenced by organization a_j in the following scenarios: (1) the two organizations directly communicate with each other; (2) a_j is widely considered a leader in the community; and (3) a_j is a larger organization (and, thus, generally has more resources). Therefore, in the following equation, the configuration incorporates these factors into the calculation of $F_i[j]$, organization a_j's influence index on organization a_i:

$$F_i[j] = \frac{f_i[j]}{\sum_{k=1}^{N} f_i[k]}, \quad \text{where } f_i[j] = \mathcal{D}(i,j)\mathcal{L}(j)\mathcal{S}(i,j). \quad (6)$$

According to Equation (6), $F_i[j]$ is essentially normalized to $f_i[j]$, which is a product of multiple parts. Function $\mathcal{D}(i,j)$ is based on the geodesic distance $dist(i,j)$, i.e. the number of hops, between organizations a_i and a_j in the communication network. Intuitively, the longer the distance between a_i and a_j, the smaller the value of $\mathcal{D}(i,j)$. This configuration represents $\mathcal{D}(i,j)$ with an exponential function,

$$\mathcal{D}(i,j) = e^{-[dist(i,j)-1]} \quad (7)$$

so that the influence degrades very fast as the distance increases.

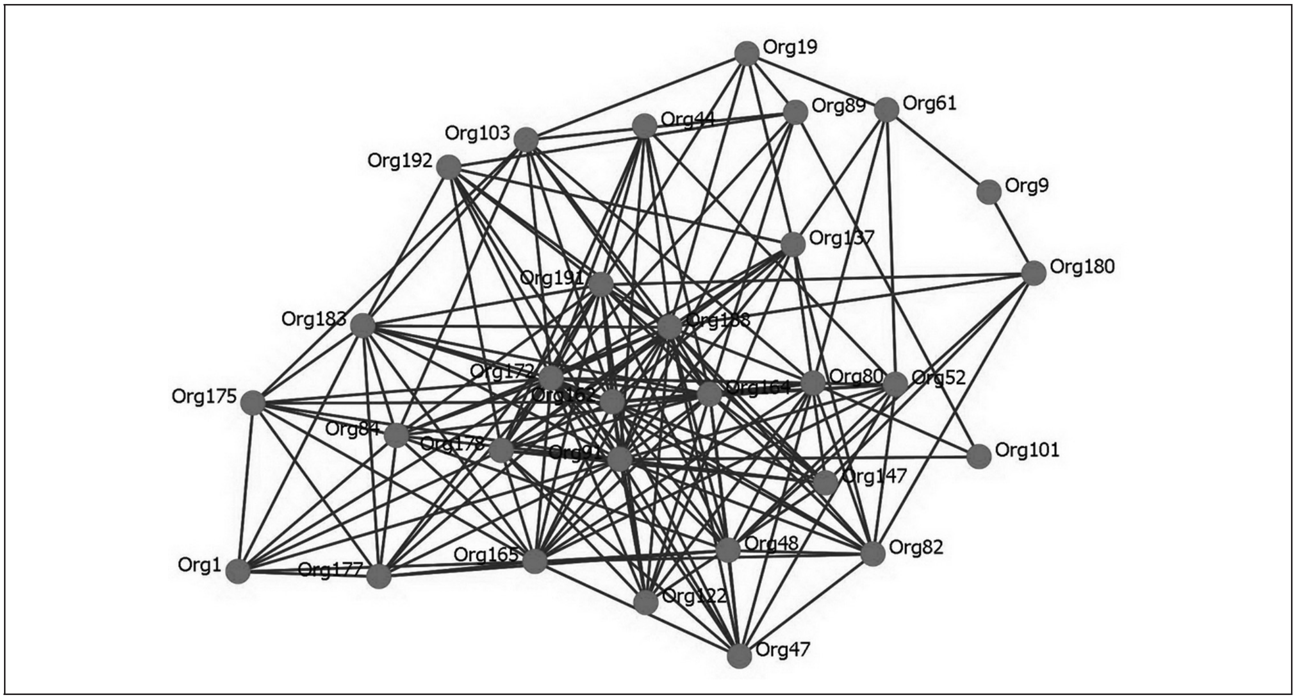

Figure 5. The predicted collaboration networks among 30 organizations in GlobalSympNet.

Function $\mathcal{L}(j)$ concerns whether organization a_j is considered a leader in this community. A leader organization a_j will have $\mathcal{L}(j) = 1.2$, while $\mathcal{L}(j) = 1$ for non-leaders. Function $\mathcal{S}(i, j)$ reflects how sizes of the two organizations affect the influence power. On the basis of the number of full-time employees in each organization, we categorize organizations into micro, small, medium, large and very large organizations. Here $\mathcal{S}(i, j)$, defined as

$$\mathcal{S}(i, j) = \left[\frac{size(a_j)}{size(a_i)}\right]^{\phi} \tag{8}$$

should yield a higher value if a_j is larger than a_i, and a lower value if a_i is larger than a_j. In this configuration, we pick $\phi = 0.3$.

The influence coefficient E_i, which connects the independent evaluation and the external influence in Equation (2), is based on each organization's response in a survey question about how likely their decisions will be influenced by those of other organizations.

With the above configuration, we run the simulation 30 times and get 30 simulated inter-organizational collaboration networks. Figure 5 shows one of them. Table 2 lists the basic statistics of simulated collaboration networks, along with those of the actual one. Statistics of simulated networks are the average of results from 30 different runs. The simulated results are very close to the statistics of the actual collaboration network. In terms of link predication accuracy, simulations with this configuration obtain an average accuracy rate of 70%, with an average sensitivity of 64%, and an average specificity of 75%. This means that the simulated network can predict whether two specific nodes are connected or not with a success rate of 70%.

Table 2. Statistics of the simulated and actual collaboration network.

	Simulated network	Actual network
Number of edges	183.6 (178.9–188.3)	186
Clustering coefficient	0.69 (0.68–0.71)	0.73
Average path length	1.64 (1.63–1.65)	1.63

Note: 95% confidence intervals are reported in parentheses.

Overall, taking the communication network as one of the inputs, the properly configured simulation is able to generate collaboration networks that are very similar with the actual collaboration network in the number of edges, average path length, and clustering coefficient. Although the simulation does not excel at link prediction accuracy, the main goal of this simulation is not to predict whether two specific organizations are connected or not in the collaboration network either. In fact, the validity of the configured simulation in basic statistics of the simulated collaboration network paves the way for our experiment in the following section, because the number of edges in the collaboration network is used as a key metric to evaluate the effectiveness of different strategies.

4.2 An experiment on how to facilitate inter-organizational collaboration

The GlobalSympNet is very interested in finding strategies that could facilitate or promote inter-organizational collaboration among its member organizations. However, as the GlobalSympNet is a coordination body without formal hierarchy and does not participate in any humanitarian project, it may not become involved directly in the process of identifying collaborative projects and forming teams. From a network perspective, it cannot directly help to build connections in the inter-organizational collaboration network. In contrast, as many organizations indicate that communication is important for and often serves as the prerequisite for collaboration, the GlobalSympNet could focus its effort on another uni-relational network, the communication network among organizations. In other words, it could try to facilitate collaboration by promoting communication among its member organizations.

The analysis of the inter-organizational communication network (Figure 6(a)) inside the GlobalSympNet reveals that organizations in the community are polarized in their network positions. The network has 95 nodes and 574 edges. As the degree distribution of the communication network in Figure 6(b) shows, there are some highly active core organizations with high degrees. In other words, some organizations communicate with a lot of other organizations and are in the core of the community. Meanwhile, many organizations only talk to a few other organizations and are at the periphery of the community.[8] Then the question for the GlobalSympNet is, among many organizations that have not communicated with each other before, which ones should the GlobalSympNet pick so that its staff members can try to introduce them to each other and encourage them to communicate. This provides a good scenario to use our simulation, because trying different strategies on its member organizations in the real world is often difficult, risky, or expensive.

Therefore, in this experiment, we use our simulation to explore how three strategies that enhance the communication network can facilitate collaboration: *Strategy 1* encourages core members to communicate more with other core members; *Strategy 2* encourages core members to communicate with peripheral members; *Strategy 3* encourages the communication between peripheral members.

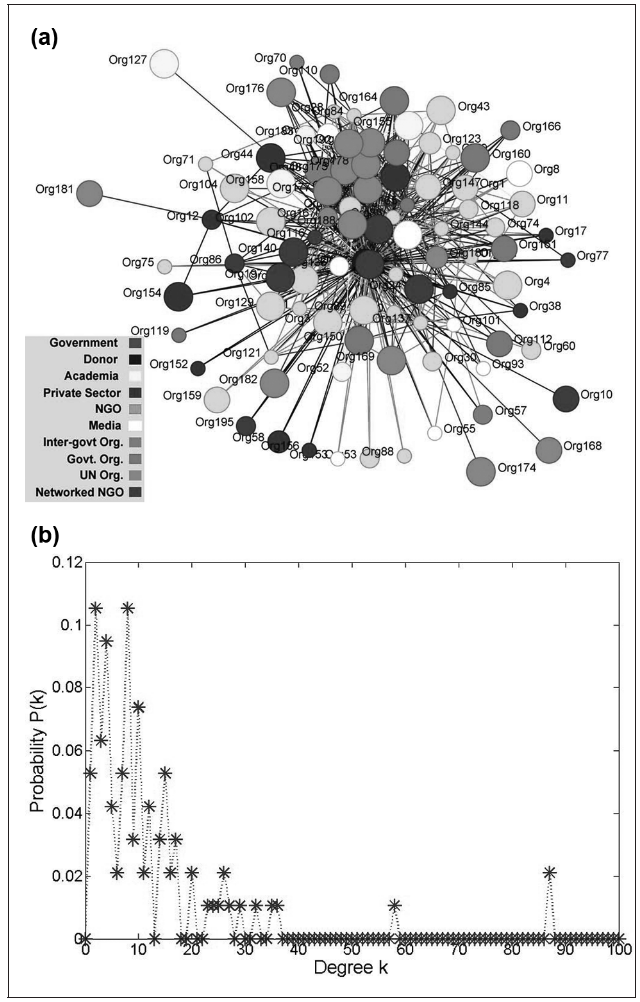

Figure 6. The communication network among the 95 humanitarian organizations as of October 2009. (a) Visualization of the network. Node colors denote organization types. (b) Degree distribution of the network.

4.3 Simulation setup and results

As we would like to evaluate how different strategies to enhance the communication network will affect the simulated collaboration network, we design four scenarios to manipulate the communication network topology: one baseline scenario with no changes to the communication network in Figure 6(a), and three scenarios with enhanced communication networks as the simulation input. To simulate the Strategy 1, 57 new edges (about 10% of the total number of existing edges) are added to the communication network. Each new edge lies randomly between two high-degree nodes, whose degrees are within the top 25% of all nodes. For the Strategy 2, the simulation also adds 57 edges to the original communication network but each edge has to connect a random high-degree node, whose degrees are within the top 25%, and a random low-degree node, whose degrees are within the bottom 25%. Similarly, Strategy 3 is simulated by adding 57 edges between randomly chosen low-degree nodes, whose degrees are within the bottom 25%.

In addition to the topology of the communication network, two other groups of factors or parameters will affect the outcome of our simulation: (1) agents' attributes and criteria of project evaluation; and (2) candidate projects that agents put in their initial to-do lists. In order to balance between internal and external validity, we control parameters in the first group and ensure that all simulations for the four scenarios will have the same agent attributes and project evaluation criteria. This is because agents' attributes and evaluation criteria, such as size, focus regions, whether it is a community leader, etc., are based on real-world data we collected and have been validated in Section 4.1. Meanwhile, as we are using synthesized project data and have no empirical data about which project an agent will pick at the very beginning, we also introduce some level of randomness into the second group of factors. In different runs of the simulation, we allow an agent to pick projects randomly from the same pool of 300 synthesized candidates projects, and put them into their initial to-do lists, as long as these projects match one of the agent's missions or focus regions. Thus, repeating multiple runs of a simulation scenario will help us to improve the validity of the results.

After running simulations 30 times (each run lasts for 40 ticks, as we did in Section 4.1) for each of the four scenarios, we evaluate the effectiveness of the four strategies to see which one promotes more collaboration. An effective strategy is one that can facilitate or promote more collaboration, but how do we decide which strategy is more effective?

Here we first consider the density of the collaboration network as an important and intuitive indicator for how well collaboration is promoted, because more edges in a collaboration network often mean more collaboration among organizations and a more collaborative environment in a coordination body. The increase in collaboration helps organizations to get more collaborators and access more resources that may be unavailable internally.[46] More edges in a network will general decrease the distance (such as the average shortest path length) between nodes and make the community more close-knit. Network density is also among the commonly used metrics to evaluate an inter-organizational network[47] in organizational research. Admittedly, the metric of density emphasizes more on the quantity of collaboration. Although the quality of collaboration is also important, it is often out of the control of coordination bodies and thus is beyond the scope of this research.

Then, the strategy, whose corresponding simulation scenario can generate a collaboration network with more edges in our experiment, is considered more effective at facilitating inter-organizational collaboration. Figure 7 shows the number of edges in simulated collaboration networks after implementing different strategies on the communication network. Each data point is the average of 30 runs. Vertical bars at data points indicate the 95% confidence interval.

The comparison first suggests a surprising result: Strategy 1, adding edges between core members in the communication network, performs worse than adding no

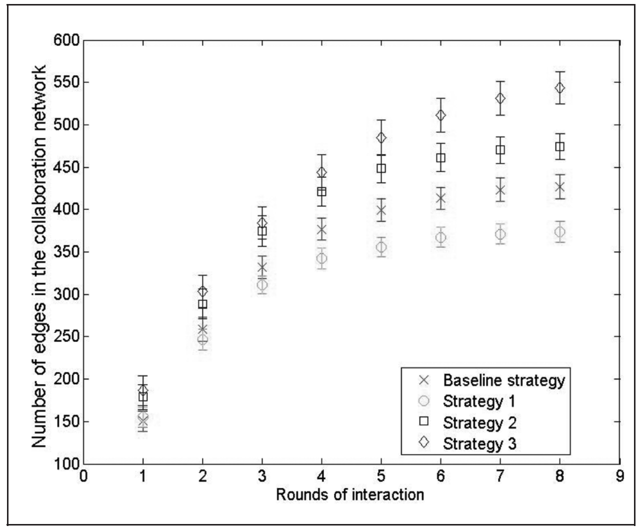

Figure 7. Comparing the effectiveness of different strategies (applied to the communication network) on the collaboration network.

edges to the communication network. In other words, although it is expected that adding edges to the communication network will always have positive impact on the collaboration network, adding edges only between high-degree nodes does not help. In the context of the GlobalSympNet, focusing only on promoting communication among its core members may not facilitate collaboration.

Meanwhile, Strategies 2 and 3, especially Strategy 3, can increase the density of the resulting collaboration network, compared with the baseline strategy. In other words, if GlobalSympNet can encourage peripheral members to get more involved in the community by introducing them to other organizations, especially other peripheral members, collaborations among humanitarian organizations will be facilitated.

4.4 Discussion

Why does Strategy 1 fail to work? Why does Strategy 3 outperform Strategy 2? From a multi-relational network perspective, we hypothesize that the dissemination of candidate project information through the communication network may have contributed to the difference in the densities of simulated collaboration networks. Figure 8 shows the total number of unique candidate projects that are evaluated by all agents. This number serves as a good metric of how well project information disseminates through the communication network. If more projects are evaluated by organizations, there is a higher chance that more collaboration can happen and thus more edges are formed in the collaboration network. As you can see from the figure, the curves of the number of evaluated projects

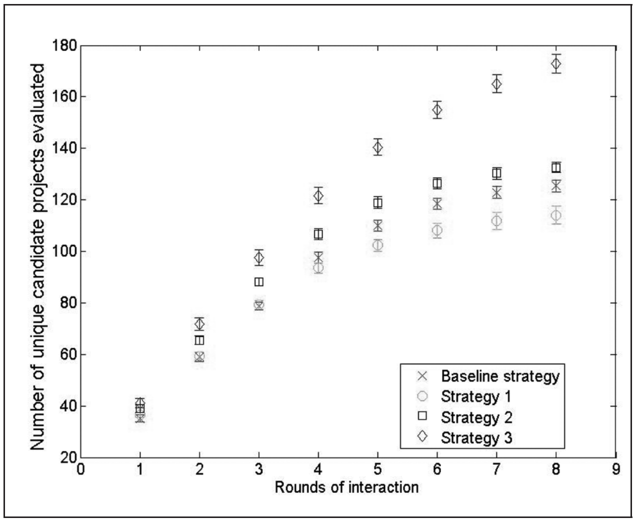

Figure 8. The total number of unique candidate projects that agents evaluate.

are very similar to the curves for the number of collaboration edges in Figure 7. Strategy 1 has the lowest number of evaluated projects, while Strategy 3 generates up to 30% more evaluated projects than Strategy 2 does.

Now we know different levels of dissemination in the communication network lead to different collaboration networks. Then what characteristics of the communication network could reflect such difference in the collaboration network? As all three strategies add the same number of edges to the communication network, the densities of the three enhanced communication networks are the same. That leaves us with average path length and clustering coefficient. Thus, we compare the two measures for inter-organizational communication networks enhanced by the three strategies in Table 3.

Although previous research argued that dissemination is easier in networks with shorter path lengths,[48] Table 3 reveals that the average path length does not correlate well with network dissemination in our study. Intuitively all three strategies can reduce the average path length because they add new edges to the original network. Despite the shorter average path length, the network enhanced by Strategy 1 still hinders the dissemination of project information compared with the original network. In addition, Strategy 2 is able to generate networks with the lowest average path length by connecting peripheral nodes with hub nodes. However, Strategy 2 is still outperformed by Strategy 3 in terms of facilitating dissemination and collaboration. Thus, the average path length of the communication network does not seem to be the key factor in affecting the outcome.

By contrast, the clustering coefficient sheds light on the problem. As Table 3 shows, Strategy 1 increases the clustering coefficient. Strategies 2 and 3, on the other hand, are able to decrease the clustering coefficient. The order of networks (by descending clustering coefficients) correlates well with the order of strategies (by descending effectiveness in facilitating inter-organizational collaboration). In other words, the clustering coefficient of the inter-organizational network based on communication relationship has an important impact on the density of the inter-organizational collaboration network.

From a multi-relational network perspective, the simulation data confirms the close relationship between communication and collaboration networks, as the dissemination of project information through the communication network plays an important role in the formation of the collaboration network. It also suggests that simply decreasing the average path length in the communication network may not necessarily increase the density of the collaboration network. Instead, adding more edges for low-degree nodes in the communication network can better facilitate the dissemination of project information and foster a better connected inter-organizational collaboration network. In addition, adding more edges between high-degree nodes in the communication network may have negative impacts on the collaboration network. A possible reason is that, in a highly clustered communication network, paths for project information dissemination to diffuse across the network is often controlled by high-degree nodes. Adding edges between high-degree nodes reinforced their key roles in the network. By contrast, new edges for low-degree nodes add shortcuts that may connect some poorly connected nodes or loose clusters a little bit. Consequently, project information does not have to go through the few high-degree nodes and have more alternative paths to be disseminated to another nodes or clusters.

Table 3. Comparison of four communication networks.

Networks	Average path length	Clustering coefficient
The original network	1.9559	0.6600
The network enhanced with Strategy 1	1.9350 (1.9342–1.9359)	0.7031 (0.7008–0.7054)
The network enhanced with Strategy 2	1.9098 (1.9068–1.9129)	0.6345 (0.6309–0.6381)
The network enhanced with Strategy 3	1.9263 (1.9254–1.9272)	0.5620 (0.5590–0.5649)

Note: 95% confidence intervals are reported in parentheses if available.

From an organizational management perspective, if GlobalSympNet would like to facilitate collaboration among its members, it may want to get peripheral members more involved in the community by encouraging them to communicate with others, especially other peripheral organizations. Empirical studies found that core organizations in the communication network are often larger or general-purpose humanitarian organizations, such as the Red Cross or the Food and Agriculture Organization of the United Nations. These organizations are often well funded and have less need from others. Meanwhile, peripheral organizations are often smaller. They may also specialize in a specific humanitarian area, such as land mine detection, or a geographic region, such as North America. They may have with limited information, resources or expertise. Thus, peripheral organizations have greater need for external resources and information and are generally more motivated to collaborate.

If core organizations are more densely connected to each other, project information is often exchanged among this highly connected group. If an organization at the periphery of the communication network would like to send out information of a candidate project, for which it wants to solicit collaborators, or to get information about other candidate projects from peer organizations, it has to rely on its one or two points of contact among core organizations. If a peripheral organization has a candidate project that fails to get support from core organizations, the project will have little chance to be received and evaluated by other organizations, who may be very interested in collaborating on the project.

Conversely, if communication between core and peripheral organizations is encouraged, information about candidate projects can diffuse more easily between information- and resource-rich organizations and organizations who desperately need more information and resources. Further, if a peripheral organization can talk to more peripheral ones, those who are highly motivated to collaborate are connected directly, which will likely to lead to more dissemination and collaboration.

5. Conclusions and future work

In this paper, we simulate the emergence of inter-organizational collaboration networks from individual organizations' interaction and decisions in a non-hierarchical context. The model adopts a multi-relational perspective and investigate how organizations' communication networks affect the growth of the collaboration network. It also uses an event-based approach and generates collaboration networks from the formation of teams for joint projects. When modeling organizations' decision-making, we capture both endogenous and exogenous factors that affect organizations' decisions on whether to collaborate on a project and to disseminate project information. To the best of the authors' knowledge, this study represents the first attempt to use agent-based models for the emergence of inter-organizational networks.

The model is implemented as an agent-based simulation for a case study of the humanitarian sector. After configuring and validating the simulation, we use the simulation to study how to promote collaboration among humanitarian organizations by enhancing the communication network. The simulation helps to compare the effectiveness of three different strategies that enhance the communication network. From the perspective of multi-relational network analysis, the simulation results suggest that adding edges for low-degree nodes, which will lead to a lower clustering coefficient, in the communication network may help to improve the dissemination on the communication network and the connectivity of the collaboration network. The organizational implication for humanitarian coordination bodies is that encouraging peripheral organizations to communicate more with others, especially with other peripheral organizations, can help them to reach out and consequently facilitate more inter-organizational collaboration in the humanitarian sector. The case study helps to provide recommendations on how to promote humanitarian collaboration, which will eventually benefit disaster victims.

In addition, the implication of this research is not limited to this case study only. First, this research illustrates that one network (the communication network) could have an impact on the growth and structure of another network (the collaboration network). Revealing the importance of the multi-relational perspective, this research should have implications to the study of complex networks that involve multiple types of relationships, such as social networks and supply-chain networks.[49] Second, we model the formation of edges in a collaboration network using an event-based approach, which can better reflect the *n*-ary nature of a collaboration relationship in the real world. This approach can also be applied to the modeling of other networks with event-based *n*-ary relationships. For example, in a co-authorship network, an event refers to scholars' collaboration on a paper; in an online social network, an event could be several users' participation in the same threaded discussion. Third, our study proposes a novel way to model the dissemination of competitive yet non-exclusive information through networks. Incorporating individuals' endogenous prioritizations and exogenous influence from the network, the new approach can be used to model the network dissemination of fashion, behaviors, products, and so on. For instance, different digital gadgets, such as iPads and netbooks, are competing in the sense that they all need users' investment, but they can co-exist too, because a user may not be limited to only one gadget. Finally, when properly configured, our model may also be used to simulate emerging collaboration networks in other domains, where individuals or agencies interact and have an influence on each other in an environment that has no formal hierarchy and welcomes collaboration. Example domains include social services, environmental protection, open-source software development, education, academic research, and so on.

For future research, we would like to configure and calibrate the simulation configuration further, and improve the accuracy of edge-by-edge prediction, so that this simulation can be used for more experiments on other inter-organizational collaboration issues. One possible way to improve the simulation is to introduce more heterogeneity among agents. For example, in Equation (4), we can assign different coefficients values (α_m and α_r) to agents, which means they evaluate candidate projects in different ways. This will bring the simulation closer to real-world situations. Another possible way is through the use of more data. In this research, we are limited by the amount of data we can collect through traditional data collection methods (surveys and interviews). Thus, we need to consider more effective methods to collect more data, especially on how networks among humanitarian organizations evolve over time. More data will help us to improve and validate our simulation.

Moreover, this research has studied how the communication network affects the collaboration network and the interaction between the two networks is one-way only. We also hope to explore how the collaboration network in turn affects the influence and information dissemination among nodes in the communication network. Another possible research direction is to incorporate other relationships in this multi-relational network, such as the business transaction network and the funding network, into the agent-based model.

Funding

This research has been partially supported by the U.S. National Science Foundation (grant number CMMI-0624219).

References

1. Powell WW. Learning from collaboration: Knowledge and networks in the biotechnology and pharmaceutical industries. *California Management Review* 1998; 40: 228–240.
2. Wood DJ and Gray B. Toward a comprehensive theory of collaboration. *Journal of Applied Behavioral Science* 1991; 27: 139–162.
3. Guo C and Acar M. Understanding collaboration among non-profit organizations: Combining resource dependency, institutional, and network perspectives. *Nonprofit and Voluntary Sector Quarterly* 2005; 34: 340–361.
4. Jang HS and Feiock RC. Public versus private funding of nonprofit organizations: Implications for collaboration. *Public Performance and Management Review* 2007; 31: 174–190.
5. Gazley B. Beyond the contract: The scope and nature of informal government–nonprofit partnerships. *Public Administration Review* 2008; 68: 141–154.
6. Bennett J. *NGO Coordination at Field Level: A Handbook*. Oxford: INTRAC, 1994.
7. Maitland C, Tapia A, Ngamassi L, Zhao K and Maldonado E. *Global Symposium +5 Information for Humanitarian Action Second Follow-up Survey Report*. Technical report, United Nations Office for the Coordination of Humanitarian Affairs, 2010.
8. Zhao K, Ngamassi L-M, Yen J, Maitland C and Tapia A. Assortativity patterns in multi-dimensional inter-organizational networks: A case study of the humanitarian relief sector. In Chai S-K, Salerno JJ and Mabry PL (eds), *Advances in Social Computing–Proceedings of the 2010 International Conference on Social Computing, Behavioral Modeling, and Prediction (Lecture Notes in Computer Science, vol. 6007*). Berlin: Springer, 2010, pp. 265–272.
9. Diesner J, Frantz T and Carley K. Communication networks from the enron email corpus ‘it’s always about the people. enron is no different’. *Computational and Mathematical Organization Theory* 2005; 11: 201–228.
10. Powell WW, White DR, Koput KW and Owen-Smith J. Network dynamics and field evolution: The growth of inter-organizational collaboration in the life sciences. *The American Journal of Sociology* 2005; 110: 1132–1205.
11. Davis GF, Yoo M and Baker WE. The small world of the american corporate elite, 1982–2001. *Strategic Organization* 2003; 1: 301–326.
12. Leskovec J, Backstrom L, Kumar R and Tomkins A. Microscopic evolution of social networks. In *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM Press, 2008, pp. 462–470.
13. Huberman BA, Romero DM and Wu F. Social networks that matter: Twitter under the microscope. *First Monday* 2009; 14: 9.
14. Entwisle B, Faust K, Rindfuss RR and Kaneda T. Networks and contexts: Variation in the structure of social ties. *American Journal of Sociology* 2007; 112: 1495–1533.
15. Szell M, Lambiotte R and Thurner S. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences of the USA* 2010; 107: 13636–13641.
16. Ahmad MA, Borbora Z, Srivastava J and Contractor N. Link prediction across multiple social networks. In *Domain Driven Data Mining Workshop (DDDM2010) at ICDM 2010*, 2010.
17. Tang L and Liu H. Toward predicting collective behavior via social dimension extraction. *IEEE Intelligent Systems* 2010; 25: 19–25.
18. Oliver AL and Ebers M. Networking network studies: an analysis of conceptual configurations in the study of inter-organizational relationships. *Organization Studies* 1998; 19: 549.
19. Powell WW and Smith-Doerr L. *Networks and Economic Life*. Princeton, NJ: Princeton University Press, 1994, pp. 368–402.
20. Gulati R and Gargiulo M. Where do interorganizational networks come from? *The American Journal of Sociology* 1999; 104: 1439–1493.

21. Oliver C. Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review* 1990; 15: 241–265.
22. Ostrom E. *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press, 1990.
23. Van de Ven AH and Walker G. The dynamics of interorganizational coordination. *Administrative Science Quarterly* 1984; 29: 598–621.
24. Barabasi AL and Albert R. Emergence of scaling in random networks. *Science* 1999; 286: 509–512.
25. Jaccard P. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin del la Société Vaudoise des Sciences Naturelles* 1901; 37: 547–579.
26. Adamic LA and Adar E. Friends and neighbors on the web. *Social Networks* 2003; 25: 211–230.
27. Katz L. A new status index derived from sociometric analysis. *Psychometrika* 1953; 18: 39–43.
28. Rosenkopf L and Padula G. Investigating the microstructure of network evolution: alliance formation in the mobile communications industry. *Organization Science* 2008; 19: 669–687.
29. Benchettara N, Kanawati R and Rouveirol C. A supervised machine learning link prediction approach for academic collaboration recommendation. In *Proceedings of the 4th ACM Conference on Recommender Systems*. New York: ACM Press, 2010, pp. 253–256.
30. Al Hasan M, Chaoji V, Salem S and Zaki M. Link prediction using supervised learning. In *Proceedings of SDM 06 Workshop on Link Analysis, Counterterrorism and Security*, 2006.
31. Wang C, Satuluri V and Parthasarathy S. Local probabilistic models for link prediction. In *7th IEEE International Conference on Data Mining (ICDM)*. Piscataway, NJ: IEEE Press, 2008, pp. 322–331.
32. Clauset A, Moore C and Newman MEJ. Hierarchical structure and the prediction of missing links in networks. *Nature* 2008; 453: 98–101.
33. Bailey NTJ. *The Mathematical Theory of Infectious Diseases and Its Applications*. London: Griffin, 1975.
34. Goldenberg J, Libai B and Muller E. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters* 2001; 12: 211–223.
35. Granovetter M. Threshold models of collective behavior. *The American Journal of Sociology* 1978; 83: 1420–1443.
36. Aral S, Muchnik L and Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the USA* 2009; 106: 51: 21544–21549.
37. McPherson M, Smith-Lovin L and Cook JM. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 2001; 27: 415–444.
38. Broecheler M, Shakarian P and Subrahmanian VS. A scalable framework for modeling competitive diffusion in social networks. In *2010 IEEE Second International Conference on Social Computing (SocialCom)*, 2010, pp. 295–302.
39. Carley K and Gasser L. *Computational Organization Theory*. Cambridge, MA: The MIT Press, 1999, pp. 299–330.
40. Zhang Y and Leezer J. Simulating human-like decisions in a memory-based agent model. *Computational and Mathematical Organization Theory* 2010; 16: 373–399.
41. Cui X, Beaver J, Stiles E, Pullum L, Klump B and Treadwell J. The swarm model in open source software developer communities. In *2010 IEEE Second International Conference on Social Computing (SocialCom)*, 2010, pp. 656–660.
42. Berry BJL, Kiel LD and Elliott E. Adaptive agents, intelligence, and emergent human organization: Capturing complexity through agent-based modeling. *Proceedings of the National Academy of Sciences of the USA* 2002; 99: 7187–7188.
43. Friedkin NE and Johnsen EC. Social positions in influence networks. *Social Networks* 1997; 19: 209–222.
44. North MJ, Howe TR, Collier NT and Vos RJ. The repast simphony development environment. In Macal CM, North MJ and Sallach D (eds), *Agent 2005 Conference on Generative Social Processes, Models, and Mechanisms*, 2005, pp. 159–166.
45. Leonardi PM. Activating the informational capabilities of information technology for organizational change. *Organization Science* 2007; 18: 813–831.
46. Boyd B. Corporate linkages and organizational environment: A test of the resource dependence model. *Strategic Management Journal* 1990; 11: 419–430.
47. Provan KG and Sebastian JG. Networks within networks: Service link overlap, organizational cliques, and network effectiveness. *The Academy of Management Journal* 1998; 41: 453–463.
48. Kleinberg J. *Cascading Behavior in Networks: Algorithmic and Economic Issues*. Cambridge: Cambridge University Press, 2007, pp. 613–632.
49. Zhao K, Kumar A, Harrison TP and Yen J. Analyzing the resilience of complex supply network topologies against random and targeted disruptions. *IEEE Systems Journal* 2011; 5: 28–39.

Author biographies

Kang Zhao is currently a PhD candidate at the College of Information Sciences and Technology, the Pennsylvania State University, USA. He received BE in Electrical Engineering from Beijing Institute of Technology and an MS in Computer Science from Eastern Michigan University. At Penn State University, he was the recipient of the Graham Endowed Graduate Fellowship and the Network Science Exploration Grant. His research interests include the simulation and analysis of complex social, inter-organizational and supply-chain networks, social computing and agent-based systems.

John Yen received his PhD in Computer Science from University of California, Berkeley, in 1986; MS in Computer Science from University of Santa Clara in 1982, and BS in Electrical Engineering from National Taiwan University in 1980. He is currently University Professor and Director of Strategic Research Initiatives at the College of Information Sciences and Technology at the Pennsylvania State University. His research interests include social network analysis, intelligent agents, social media mining, and cloud computing. He has published a book, two edited volumes, and over 170 refereed papers. Sponsors of his research include NSF, ARL, ARO, ONR, AFOSR, DOE, NIH, and MIT Lincoln Laboratory. He received the NSF Young Investigator Award in 1992, and was named a Fellow of IEEE in 2000.

Louis-Marie Ngamassi Tchouakeu received his PhD from the College of Information Sciences and Technology at Penn State University. His research focuses on information and communication technology (ICT) use for inter-organizational co-ordination and social networks among humanitarian organizations.

Carleen Maitland is a Program Manager in the Office of International Science and Engineering (OISE) at the U.S. National Science Foundation. She is on leave from the College of Information Sciences and Technology at Penn State University. She holds a doctorate in Economics from the Technical University of Delft, The Netherlands as well as Bachelor and Master's degrees in engineering from Stanford University and Worcester Polytechnic Institute. As Associate Professor of Information Sciences and Technology her research investigates international and inter-organizational mechanisms influencing the production of, access to and use of information and communication technologies (ICTs). These studies are carried out in the U.S., Europe, Africa and Asia, and in several domains including humanitarian relief and wireless communications.

Andrea H Tapia is an Associate Professor of Information Sciences and Technology at the Pennsylvania State University. She also holds appointments with the Sociology Department, Labor and Industrial Relations Department and Science and Technology Studies Program at the same institution. Dr. Tapia earned her doctorate in the field of Sociology at the University of New Mexico, followed by a Post-Doctoral Fellowship at the University of Arizona. Dr. Tapia's research has contributed to addressing some of the most critical social problems facing the world today, such as humanitarian relief, emergency response and has directly contributed to the policy-making bodies of the United Nations, the U.S. Obama Administration and the largest international relief and development organizations globally.