SIMULATION-BASED WORKFORCE ASSIGNMENT CONSIDERING POSITION IN A SOCIAL NETWORK

Nurcin Celik
Hui Xi
Dong Xu
Young-Jun Son

The University of Arizona
Systems and Industrial Engineering
Tucson, AZ, 85721, USA

ABSTRACT

Globally distributed software enhancement necessitates joint efforts of workforces across various organizations, which constitutes a multifaceted social network. Here, we propose a novel modeling framework to optimally assign the workforce to software development projects considering both short and long-term benefits of the organization. The proposed framework is composed of the evaluation module, an agent-based simulation model representing the considered social network; and the assignment module, a multi-objective optimization model. The Decision Evolution Procedure of the evaluation module first calculates the position values between each pair of available workforce. Using these position values, the Extended Regular Equivalence Evaluation algorithm of the evaluation module then computes the regular and structural equivalence values between each pair of workforce. Finally, the assignment module selects the optimal workforce mix maximizing both the short (productivity) and long-term performance (robustness) of the organization. The proposed framework is demonstrated with the software enhancement process in Kuali organizational network.

1   INTRODUCTION

Policy and technological developments of the past few decades have given a marked rise to the globalization of industries by which economies, societies, and cultures have become integrated through a globe-spanning network of communication and trade. Along with this change and the advance of communication ways (i.e., fiber optic communications, satellites, and Internet), have become globally distributed team collaborations common in many sectors including information systems and software development (Oshri, Kotlarsky, and Willcocks 2007). Globally distributed software development presents many benefits such as decreased development costs, access to a larger skill pool, and twenty-four hour development by dispersing software development tasks among several remote centers (Nguyen, Wolf, and Damian 2008). On the other hand, efficient workforce assignment plays a crucial role in determining the success of distributed development process especially when there are many tasks/projects that must be completed across a complex social network that includes a broad range of stakeholders (e.g., upper level managers, and testing teams). The effective assignment of labor force leading to reduced workforce idle time, improved workforce utilization, and improved profit becomes even more significant if the considered organization is facing faster product changes, accelerated process innovation, and intense competition (Nembhard and Norman 2002; Mockus and Herbsleb 2002). The literature serve studies analyzing the effect of individual worker's learning and forgetting, cross-training strategies of workforce, labor flexibility, and different programs for workload balancing of workers on process output. To this end, while the impor-

tance of assigning workers to appropriate tasks have been studied during the recent years, not many of them considered the social aspects associated with the management of these team compositions over globally distributed development projects. The goal of this research is to propose a novel integrated modeling framework that will help key stakeholders (in particular, development managers) devise optimal workforce assignments considering both the short-term (productivity) and long-term (robustness of the organization against potential departure of key personnel) needs of the organization in a multi-organizational distributed software development setting considering the employers' positions in their social network.

The proposed framework is composed of the evaluation and the assignment modules. Each time when a workforce assignment has to be made, the Decision Evolution Procedure (DEP) and the Extended Regular Equivalence Evaluation (EREE) algorithm of the evaluation module calculates the position value and the regular and structural equivalence values between each pair of available workforces based on pairs' trustworthiness, influence, reputation, and proximity, respectively. Then, the optimal workforce assignment module selects an optimal workforce mix that maximizes both the short-term (productivity) and the long-term performance (robustness) of the organization. An agent-based simulation model and a multi-objective optimization model is developed for the evaluation and the assignment modules, respectively.

To demonstrate the validity of our work, the proposed framework involving simulation and optimization is illustrated with a software enhancement request process in Kuali, an open source software development project involving 12 organizations and the 46 technical and administrative personnel working on the project (www.kuali.org). The results obtained from the proposed workforce framework considering position notion are compared to the ones obtained from the workforce assignment technique developed by Celik et al. (2010) which considers the underlying social network from relational and structural embeddedness perspective. The effectiveness of the proposed framework is demonstrated for the enhancement request process of Kuali in particular, and for various other organizations and organizational networks that focus on globally distributed software development in general.

2 KUALI ORGANIZATION AND THE ENHANCEMENT REQUEST PROCESS

Kuali Foundation is a non-profit organization involving major universities and colleges developing software that meet the need of all sizes of institutions. As Kuali's initial product, Kuali Financial Systems (KFS), provides a comprehensive suite of financial software. Its modular architecture allows institutions customize and implement only those functional elements they need. The proposed modeling framework is illustrated via the software enhancement request process in KFS. Here, the models aims at being used by development managers in the enhancement request process to contrive optimal workforce assignment. An enhancement request process begins with a participating institution's Subject Matter Expert (SME) filling out required information on the enhancement request form such as the primary component involved, the priority (high, medium and low), business need, impact if not implemented, and questions to be answered by the Development Manager (DM), and submits it to the DM. Then the DM estimates the effort required, technical impact, and system impact (expected functional impact on any other modules), and submits it to the Functional Council (FC). If the request is approved by the FC, the FC meets with the Project Manager (PM) in order to schedule it, and then, the PM delivers this decision to the DM. The DM initiates the series of actions required for assignment of workforce to projects. This includes the evaluation of employee equivalence and assignment of the specific labor force to the projects, which is shown in the gray box in the middle of the sequence diagram of the enhancement process. However, if the request is rejected, the SME may choose to raise it to the board which may affirm or override the FC's decision. Our work aims at building up an optimal workforce assignment framework for the use of DM considering both short and long-term performance factors.

3   PROPOSED SIMULATION-BASED WORKFORCE ASSIGNMENT CONSIDERING POSITION NOTION IN SOCIAL NETWORK

The workforce assignment framework developed in this study is comprised of the equivalence evaluation module and the workforce assignment module (see Figure 1). In the evaluation module, the behavior of a social network is predicted via an agent based simulation model, where each agent represents an individual in the network with its own characteristics (attributes). The evaluation module entails the Decision Evolution Procedure (DEP) and the Extended Regular Equivalence Evaluation (EREE) algorithm, both of which have been developed as part of this study. Given the requirements of a project, the evaluation module first determines the values of the four parameters that are found to reflect the social relationship similarity between each pair of employees involved in a project (trustworthiness, influence, reputation), and to define how close they are in the whole social network in terms of their positions (proximity). These parameters were found to be significant for the considered case of KFS social network based on a regression model that is developed to run offline before the workforce assignment takes place (Provan and Lemaire 2010). Accepting the values of these significant independent parameters as inputs, the DEP computes the position value between each pair of employees. Then, based on these position values, the EREE algorithm is run to determine the regular and structural equivalences among the members in the social network. Next, the assignment module, provided aforementioned regular and structural equivalences from the evaluation module, selects the optimal workforce mix maximizing the estimated short-term (productivity) as well as the long-term performance (robustness) of the organization via a multi-objective optimization involving a weighed summation of the productivity and potential gain in the innovation diffusion and opinion formation capabilities of the social workforce network.
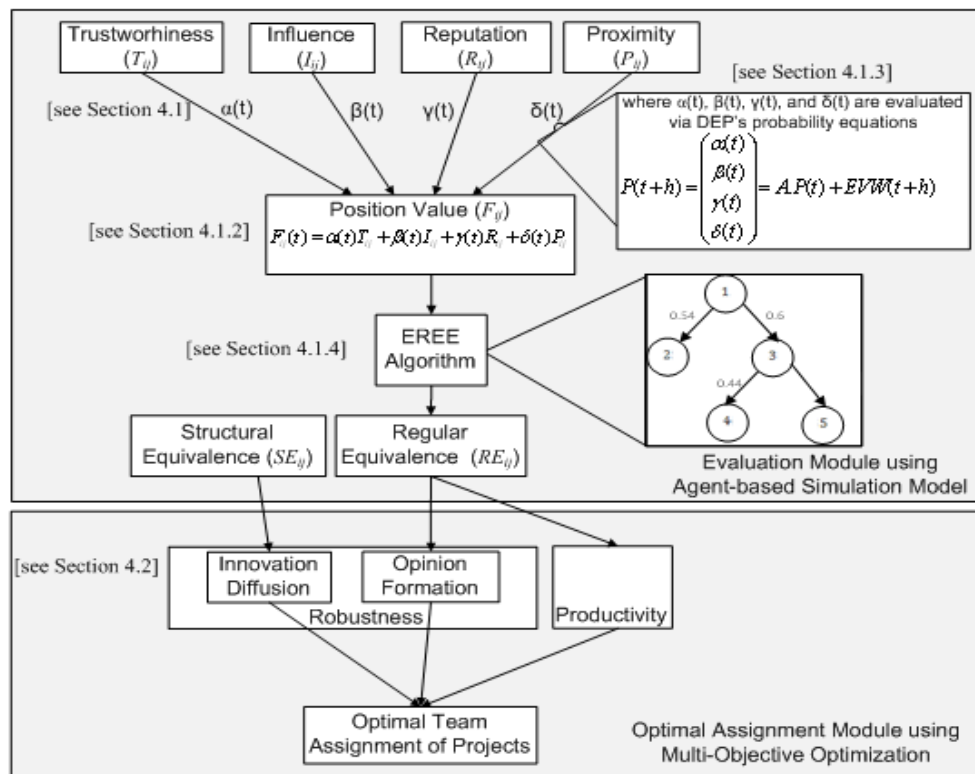

Figure 1: Sequence diagram for the enhancement request process and simulation-based optimal workforce assignment considering position notion in the social network

3.1 Evaluation Module

In network analysis, different forms of equivalence (structural, automorphic, and regular) are defined based on separate theoretical groundings. Two actors are said to be structurally equivalent if they have identical ties to and from all the other actors in the network. Equivalent actors can be grouped into equivalence classes. A major problem with structural equivalence is that it is a mathematical property and actors are unlikely to have exactly the same ties to and from other actors. Therefore, to actually use structural equivalence, measures have been developed to find subsets of actors that are approximately structurally equivalent. Regular equivalence was developed to relax the equivalence assumptions. Regular equivalence is based on identical ties to and from equivalent actors (but not necessarily identical). In fact, any set of actors that are structurally equivalent are also regularly equivalent. The earliest developed measure for regular equivalence and the one most used is the REGE algorithm (Borgatti and Everett 1993). This algorithm uses an iterative procedure, the estimates of the degree of equivalence between a pair of actors is adjusted by the equivalences of alters. Unlike other equivalence measures, REGE counts ties matched between a pair of actors and not the number of alters matched. Regular equivalence can be used to find similarity among actors and allows us to define what type of similarity we are interested in.

In this study, we examine the similarity based on social attributes. In the KFS case, we are first interested in equivalence based on reputation, influence, and trustworthiness. Equivalent actors are those who have similar positions based on the number of out-going and in-coming ties of the given attributes. By defining regular equivalence using these attributes, the goal is to match actors who are trustworthy, reputable, and influential with the ones who are not depending on the need. In a survey of KFS members, we asked respondents to rate the overall qualities of their relationships based on perceived trust, influence, reputation. By basing regular equivalence position on these attributes, we determine which individuals are regarded similarly in the network based on perceptions of these social attributes. Furthermore, because geographical proximity is expected to impact social relationships, a variable for proximity has been included in our regular equivalence model. Next, in order to capture position based on actual ties to others in the network, we compute position based on structural equivalence. Structurally equivalent actors tend to have similar attitudes and behave similarly. The information structurally equivalent actors receive are likely be the same, thus, position based on this type of equivalence influences the speed for innovation diffusion. Data on the ties between KFS members were collected through the survey where respondents were asked to indicate whether they had a working relationship with each individual on the list "over the past six months," and if so, what type of relationship.

In this study, a position value is defined for every arc of a directed network, which represents the comprehensive evaluation of agent j given by agent i based on the four attributes denoted as trustworthiness (T_{ij}), influence (I_{ij}), reputation (R_{ij}), and proximity (P_{ij}) (see Figure 1). Each one of these attributes can take values ranging from zero to one, depending on the individual relationships in the network. Mathematically, the position value (F_{ij}) is computed via the function defined in (1) where non-negative parameters $\alpha(t)$, $\beta(t)$, $\gamma(t)$, and $\delta(t)$ denote how significant each corresponding attribute is at a given time for a particular organization's assignment problem. These values are determined via the usage of DEP of the evaluation module as the social network evolves.

$$F_{ij}(t) = \alpha(t)\,T_{ij} + \beta(t)\,I_{ij} + \gamma(t)\,R_{ij} + \delta(t)\,P_{ij} \qquad (1)$$

3.1.1 Decision Evolution Procedure (DEP) in Defining Weight Values

The weight values, $\alpha(t)$, $\beta(t)$, $\gamma(t)$, and $\delta(t)$, denote the organization's preference to each attribute which evolves as the social network of employees changes over time. In reality, different objective-oriented organizations may weigh these four attributes in different ways and may even spend major portion of their attention to a specific attribute among the ones that are given. For instance, diversified organizations handling large cooperative projects may weigh "trustworthiness" attribute (i.e., $\alpha(t)$) more than other types of organizations (i.e., centralized organizations) as employees are required to trust each other in order to

achieve high efficiency (e.g., large-scale automotive manufacturing companies). On the other hand, in centralized national organizations where the network of first tier suppliers are believed to perform more coherently when that are in close distance to each other, more credit may be given to "proximity" attribute (i.e., $\beta(t)$). In academic organizations, seniors (e.g., professors, senior researchers etc.,) and junior fellows (e.g., students, junior researchers etc.,) are usually involved in a project at the same time where they are motivated to take advantage of each other's expertise and excitement to work more collaboratively. Here, peer influence are considered via the greater amount of weight assigned to the "influence" attribute (i.e., $\gamma(t)$). Lastly, service oriented organizations (e.g., consulting companies, restaurants, etc.,), are concerned to increase the reputation of employees since its components (i.e., intangibility, and inseparability characteristics) are essential to the customers.
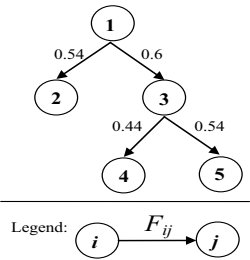
While most organizations apply consistent long-term strategies, it is necessary to adapt their concentration to the changing environment as well as temporary need. For instance, all types of organizations may want to increase the weight of influence during the training period of new employees. In another case, if a centralized company has to accomplish an urgent project with high quality using labors from different departments, it pays more attention to trustworthiness during that particular project time. The preference vector for the aforementioned attributes could not be kept static forever. To address this problem, we employ decision evolution procedure (DEP) to describe the dynamic evolution of preferences among four options and identify the current values of the weights as the network evolves. DEP is developed based on decision field theory (DFT), which is originally built by Busemeyer and Townsend (1993) as a mathematical framework leading to the understanding the cognitive mechanism of the human deliberation process in making decision under uncertainty. However, there are two main differences between the proposed DEP and DFT. First, in DFT, the preference state changes as a mechanism of only cognitive deliberations, while in DEP, the preference state changes as a combined mechanism of social environmental factors, personal deliberations and organizational structural changes. Secondly, the preference matrix in DFT represents the state of the system via a series (vector) of the same parameter, whereas in DEP, the preference matrix represents the state of the system via a combination of four different parameters where each one is defines as a dedicated weight to a separate attribute. In DEP, the preference state is calculated during the deliberation time using a linear system formulation as given in (2) where the preference state matrix $P(t)^T=[\alpha(t), \beta(t), \gamma(t), \delta(t)]$ represents the importance of each attribute at time $t<T_D$, T_D is the decision time and the preference state is updated at every time step h. A demonstrates the stability of preference to each option by its structure. The diagonal elements of A represent the memory from the previous preference state while off-diagonal elements give the inhibitory interactions among competing options. There are two assumptions to give same amount of memory as well as interaction effects to each option: 1) Matrix A is assumed to be symmetric, and 2) diagonal elements of A which represent the memory from the previous preference state are assumed to have the same value. Moreover, all eigenvalues λ_i of A are less than one in magnitude to make the linear system stable ($|\lambda_i|<1$). Given m attributes to n options respectively, V is the value matrix (n x m matrix) representing the subjective perceptions of a decision-maker by $V(i, j)$. The weight vector $W(t)$ allocates the portion of attention to each column j (characteristic) of V through $W(j, 1)$ which is the only dynamically changing element in DEP. Finally, the contrast matrix E compares the weighted evaluations of each option, $VW(t)$. In our case the preference increase of one option lowers the preference of alternative options requiring C to be defined as $C_{ii}=1$ and $C_{ij}=-1/3$ for $i \neq j$. We use DEP to compute the weights of the attributes that are used to compute the position value between each pair of employees. Then, the obtained position values are fed into the EREE algorithm to obtain equivalence values among pairs of employees to conduct the workforce assignment.

$$P(t+h) = AP(t) + EVW(t+h) \tag{2}$$

3.1.2 Extended Regular Equivalence Evaluation (EREE) Algorithm

Given the position values based on the attributes and their weight values which are determined using DEP, the goal of the EREE algorithm developed in this work is to calculate the regular and structural equivalences between each pair of employees. This algorithm adopts a point system that was first introduced by Borgatti and Everett (1993) in the regular graph equivalence algorithm (REGE) for computing role distances. The original point system and its modified version that is adopted in this work are shown in Table 1. In both cases, the point system allows bi-direction arcs to be matched with both one-direction and bi-direction arcs. If an arc $a \leftrightarrow i$, matches another arc $e \leftrightarrow j$, then 2 points are scored. Otherwise, if an arc $a \leftrightarrow i$ matches another arc $e \rightarrow j$ or arc $e \leftarrow j$, only 1 point is scored. In each pair of nodes, algorithm starts counting the match points from one point of view, and then repeats the process from the other point of view. Then, points from both nodes of view are summed and divided by the maximum possible number of matches. For instance, in our earlier simple exemplary network that is depicted in the middle of the Table 1, if we consider nodes 1 and 3, the algorithm would compute as $(1+1+1+1)/(1+1+1+1)=4/4=1$, which means nodes 1 and 3 are totally equivalent in this network in terms of their regular equivalence.

Table 1: Equivalence evaluation between actors a and e by using the point system of REGE algorithm (Borgatti and Everett 1993) and the proposed EREE algorithm

(a) REGE Algorithm			
Alter j / Alter i	$e \rightarrow j$	$e \leftarrow j$	$e \leftrightarrow j$
$a \leftarrow i$	1	0	1
$a \leftarrow i$	0	1	1
$a \leftrightarrow i$	1	1	2

(b) Proposed EREE Algorithm			
Alter j / Alter i	$e \rightarrow j$	$e \leftarrow j$	$e \leftrightarrow j$
$a \rightarrow i$	$1*F_{ae}$	0	$1*F_{ae}$
$a \leftarrow i$	0	$1*F_{ae}$	$1*F_{ae}$
$a \leftrightarrow i$	$1*F_{ae}$	$1*F_{ae}$	$2*F_{ae}$

The EREE algorithm developed in this study differs from the original REGE algorithm in two main aspects. First, the original REGE algorithm is build to find the regular equivalences in the network whereas our EREE algorithm is aimed to find both regular and structural equivalences between pair individuals. The partial pseudo code of EREE algorithm for computation of structural equivalence among the members of social network is provided in Figure 2. Second, while the original REGE algorithm is built to determine the network equivalences based on single attribute; the EREE algorithm determines the network equivalences based on the four attributes, trustworthiness, influence, reputation, and proximity, which have been shown to be significant in determining the network equivalences. Based on the same scoring principles as those of REGE point system, a revised point system considering attributes effect has been implemented in our EREE algorithm. For example, if we consider the same nodes 1 and 3, EREE would compute as $(0.54+0.60+0.44+0.54)/(1+1+1+1)=0.53$ as the regular equivalence.

```
LOAD the initial network data
FOR  i=1: Total Number of Nodes //i and j are nodes (agents) in the social network
   FOR  j=i+1: Total Number of Nodes // i ≠ j
   Check if pair (i, j) is connected, and determine the connection type/weights
      FOR k=0: Total Number of Nodes
          Increment count by one if i and j both have same relationships
      End FOR
   Check the count
   IF i and j have the same relation to all other nodes, structural equivalence is 1
   Else their structural equivalence is 0
   End FOR
End FOR
```

Figure 2: Partial pseudo code of EREE algorithm for computation of structural equivalence

3.2 Assignment Module

The workforce assignment problem can be solved for multiple projects or a single project at a time depending on the need. (3) denotes a multiple projects assignment problem. When n is set to 1 and k is deleted from the formulation, it denotes a single project assignment problem. The objective function of the assignment problem denotes organization's short-term and long-term interests in terms of productivity and robustness factors, respectively. In (3), m is the number of employees, n is the number of projects, A_i is the maximum total number of projects that can be assigned to employee i, B_k is the number of employees that need to be assigned to project k, $x_{ik}(t)$ is the binary decision variable, meaning that employee i is assigned to project k at time t when $x_{ik}(t)$ is equal to 1, and $P_k(t)$ is the profit function where $0 \leq \alpha \leq 1$, $RE_{aver}(t+\Delta t)$, $RE_k(t)$ and $SE_k(t)$ are predicted via simulation in a Δt time window.

In the proposed assignment formulation, α can reflect how significant it is to obtain short-term benefits for a particular organization when the assignment is made for a set of projects. Increased α values represent that obtaining immediate results are more significant whereas decreased α values represent that steady state results are more critical for the organization. To this end, α value is expected to be high for high-tech manufacturers (e.g., Intel, Samsung), where timely introduction of new products against changing markets are important and is expected to be low for government institutions, public services and construction business, where the existence of the organization is dependent on their long-term success.

$$Max \sum_{i=1}^{m}\sum_{k=1}^{n} P_k(t)x_{ik}(t)$$

where
$$x_{ik}(t) \in \{0,1\}$$

$$\sum_{k=1}^{n} x_{ik}(t) \leq A_i \qquad \forall i \in \{1, 2, \ldots, m\} \quad P_k(t) = \alpha.f(RE_{aver}(t)) + (1-\alpha).g(RE_k(t+\Delta t), SE_k(t+\Delta t)) \qquad (3)$$

$$\sum_{i=1}^{m} x_{ik}(t) = B_k \qquad \forall k \in \{1, 2, \ldots, n\} \qquad g(RE_k(t+\Delta t), SE_k(t+\Delta t)) = \frac{3.RE_k(t+\Delta t) + SE_k(t+\Delta t)}{4}$$

3.2.1 Productivity (*f* in the Objective Function)

In this section, we explain how the short-term goals (productivity) of different types of organizations are formalized based on the regular equivalence. In general, at every certain time period t, we have the regular equivalence matrix $RE(t)$ for the entire workforce in the company. In this matrix, every two employee, i and j, have a value a_{ij} (which is between 0 and 1) to express their regular equivalence (which depends on the common trustworthiness, influence, reputation as well as proximity that these two people have). The matrix value of $RE(t)$ is a measure of regular equivalence between each two people at a certain time period, which changes dynamically through the development stages of the company when the time proceeds. Values in this measure represent the strength of ties between each two people. Then, when there is a project assigned to a certain workforce combination including k persons in it, they form a k-persons team to do the project. Here, we use the arithmetic average to formulize the average regular equivalence in this k-persons team as shown in (4).

$$RE(t) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & & \ddots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{bmatrix}$$

$$f_1(RE_{aver}(t)) = 1 - \sqrt{e^{-9RE_{aver}(t)}}$$

$$f_2(RE_{aver}(t)) = 0.95RE_{aver}(t)$$

where $a_{mm} = 1, 1 \leq m \leq n, 0 \leq a_{ij} = a_{ji} \leq 1, 1 \leq i, j \leq n, i \neq j$

$$f_3(RE_{aver}(t)) = \frac{e^{5RE_{aver}(t)-1}}{16} \qquad (4)$$

$$f_4(RE_{aver}(t)) = 2.8RE_{aver}^2(t) - 1.12RE_{aver}(t) + 0.212$$

$$RE_{aver}(t) = \frac{\sum_{1}^{N_k} a_{ij}(t)}{N_k}, \quad \text{where } N_k = \binom{k}{2}$$

$$f_5(RE_{aver}(t)) = -8.33[RE_{aver}(t) - 0.82]^2 + 0.92$$

In this study, we consider three main development stage of the company: the early development, the growth and the maturity stages. During early growth stage, the company concerns mainly on the expansion of its sales to the market and the recruitment of employees. Therefore, during this period of development, most of the organizations concentrate on communication, coordination as well as cooperation in order to develop products, retain the business growth, and penetrate the market. Thus, the regular equivalence based on trustworthiness, influence, reputation as well as proximity play a significant role during this period of organization development. For companies in the early and growth stages, we propose different productivity functions based on their total regular equivalence.   Figure 3(a) depicts the organizations whose productivity increases faster when regular equivalence intensity grows in early (initial) stage, rather than the growth stage (see f_1 as well). Here, it would not be very beneficial to intensify the regular equivalence beyond the initial stage since it would have only a modest effect on productivity. This class encompasses organizations that might operate with minimal support staff and financial assurance, and struggle to survive (e.g., AT&T). Figure 3(b) depicts a positive linear relationship between regular equivalence intensity and productivity (see f_2). Thus, the more intensity of the relationship is strengthened, the more productivity is achieved. This class of organizations includes those that usually focus on multidisciplinary projects with the involvement of various specialized groups where high trustworthiness, better influence as well as good reputation become the key to success. Organizations, which are formed by multiple universities to reach their common objectives, belong to this class as well (i.e., Kuali organization). Figure 3(c) represents that a company's productivity increases faster when the regular equivalence increases during the growth stage rather than the early stage, or from early to growth stage. In these organizations, it is very beneficial to intensify the regular equivalence relationship during the growth stage since it has a tremendous effect on productivity growth. A company that relies on a new technology, or an expanded product line for its survival (i.e., Lenovo Computer) belongs to this category (see f_3). Figure 3(d) depicts that during the early stage of company development, the increase of regular equivalence decreases the productivity, because it needs time to make the workforce become familiar to each other, to cooperate as well as create a healthy atmosphere to conduct the project. Afterwards, the productivity function becomes similar to the one depicted in Figure 3(c), in which the productivity grows much faster when the regular equivalence increases. The companies operating in large scale and having many distributed stores but relatively low level of technology (i.e., Food Store) belong to this category (see f_4). After the early development and growth stage of development, the company evolves to a different development stage, which is usually called as the maturity stage. The relationship between the productivity function and the regular equivalence in this stage is more complex than the relationships given for earlier stages of a company. Figure 3(e) presents the case in which after evolving into the maturity stage, a company has expanded enough products in the market, and the business sales tends to be steady. The business has reached the point where the fruits of labor can be realized and the company has attacked its market well. However, after this point, the productivity cannot be increased due to the excess strength of regular equivalence. This happens when the transfer of business is hard to realize, the conflict of interest exists among some high regular equivalence teams or the hardship occurs for the new employee to emerge to the high regular equivalence team (i.e., CircuitCity when bought by Systemax Inc.) (see f_5).
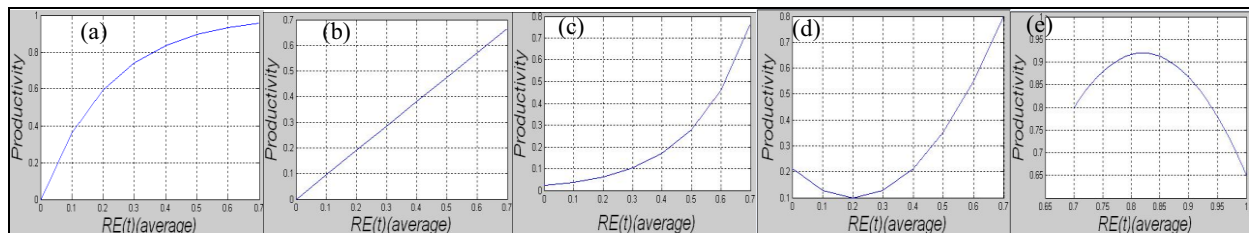

Figure 3: Relationship between RE(t) and productivity for an organization of types f_1, f_2, f_3, f_4, and f_5

3.2.2 Robustness (*g* in the Objective Function): Opinion Formation and Innovation Diffusion

In this study, two metrics which are widely studied in the literature in theoretical grounding, opinion formation and innovation diffusion, are adopted to estimate the long-term benefits of the assignment. Because individuals depend on others to shape their view of the world and other people's views do influence acquisition of a new technology (Wu and Huberman 2008; Kuandykov and Sokolov 2010), an organization's procedure of opinion formation and innovation diffusion indicate the robustness of its social network. In an organization, employees can exchange and spread opinions while collaborating or socially interacting with each other. The process of cumulating daily ideas and forming a concrete opinion is defined as opinion formation. Latane (1981) used the concept of social distance (distance between two individual's beliefs and ideals) to determine their social impact on each other when personal opinion updates and formulated that each individual *i* updates its opinion state, σ_i with a probability, $p(A_i)$ based on (5). In (5), η is the social noise, I_j is the influence of an individual *j* over the social network, RE_{ij} is the regular equivalence between individual *i* and *j* , σ_i is the opinion of each individual *i*, and *h* is the external information. When RE_{ij} has larger values (*i* is regularly more equivalent to *j*), $p(A_i)$, the probability that two individuals have similar opinions, tends to be larger. The limitation of this method is that every individual's opinion has to be taken into account which may not be realistic in some organizations.

$$\sigma_i = \begin{cases} +1 & \text{with probability } p(A_i) \\ -1 & \text{with probability } 1 - p(A_i) \end{cases} \quad \text{where} \quad p(A_i) = \frac{e^{A_i/\eta}}{e^{A_i/\eta} + e^{-A_i/\eta}} \quad \text{and} \quad A_i = \sum_{j=1, j \neq i}^{m} \frac{I_j \sigma_j}{RE_{ij}} + h \quad (5)$$

The evaluating and adopting process when a new idea is formed and spread through an organization, is defined as innovation diffusion. In the literature, one of the key factors affecting an organization's innovation diffusion capability is agreed to be organizational structure. One model aimed to define the innovation diffusion relying on the processing capability of the agents in the organization is introduced by DeCanio and Watkins (1998). In this work, a modified version of their state update equation is adopted while their parameter called connectedness is replaced with our structural equivalence so that it could be applied to more general cases with various types of organization structures studied in this work. To this end, the probability of an individual *i* in the organizational network changes its state as defined in (6), where the binary variable y_j is the state of individual *j*, SE_{ij} is the structural equivalence between individuals *i* and *j*, and *c* is the processing power of agents that change in a monotonic way. Similar to opinion formation, when individuals *i* and *j* have structurally equivalent position in the organization, it is likely that their innovation will be accepted and executed by their subordinates in a faster pace. However, when the said individuals do not possess the same social role (are not structurally equivalent), formation of opinions may be hindered due to the lack of mutual understanding. In our agent based simulation, the values of the parameter, $SE_{ij}(t)$, shows smaller values if the considered agents *i* and *j* are not structurally equivalent which results in decreased probability of change of state of individual *i* in the organizational network. In contrast, the larger values of $SE_{ij}(t)$ shows that the considered agents *i* and *j* are structurally equivalent resulting in increased probability of change of state of individual *i* in the network. Similar to opinion formation, it may not be realistic for some organizations that innovations would diffuse until they are held by all members of the organization.

$$P_i = f(x_i) = \frac{1}{[1 + e^{-[\frac{x_i \cdot c - a}{b}]}]} - \frac{1}{1 + e^{\frac{a}{b}}} \quad \text{where} \quad x_i = \sum_{j=1, j \neq i}^{m} \frac{y_j}{SE_{ij}(t)} \text{ and } y_i = \begin{cases} 1 & \text{with prob. } P_i \\ 0 & \text{with prob. } 1 - P_i \end{cases} \quad (6)$$

4 EXPERIMENTS AND RESULTS

In this section, results obtained from the proposed simulation-based workforce assignment framework considering position notion in a social network are presented with a focus on the enhancement request process in Kuali. In Kuali, the workforce assignment of projects is performed by the development managers, who assign developers to projects mainly by considering the workload and qualifications of the developers but not the underlying social roles of the said developers. In the agent-based simulation model that represents the Kuali network, each developer as well as development manager is modeled as a sepa-

rate agent having its own attributes, social roles and connections. The model contains 46 different agents whose related information is originally derived from an interactive survey conducted on the Kuali employees. The factors affecting the social position are also identified from the same survey. The historical data on enhancement requests from Kuali workflow system is used to determine characteristics of the projects (e.g., range for project duration, and average number of people involved). The model assigns 30 projects, each having different number of developers and lasting for two months on average. The optimization problem of the assignment is solved using the Greedy algorithm (Martello and Toth 1990).

Regarding the parameters that are used to demonstrate the results of this study, the 'innovation diffusion' represents the number of time steps necessary for a recently presented idea to be accepted by all of the stakeholders. Therefore, the lesser the values, the more promising the results are since smaller values show that it takes shorter amount of time for the network to diffuse innovations. In this case, the structure of the organization is being shaped in such a way that the organization members comply with new concepts and capable of quickly adapting to altering settings. Larger values of 'innovation diffusion' represents that the structure of the organization is being shaped in such a way that the organization members are not receptive of unaccustomed, newly introduced ideas and instead the social network's drive is towards preserving already accustomed principles, thoughts, and relations. Analogously, the smaller values are favorable in 'opinion formation', which demonstrates the number of time steps needed for the evolution of a set of opinions considering the interactions of the individuals of the network.

The dynamic change of regular equivalence levels (total for each developer) in the developed agent based simulation model at different times of a single simulation run when nine workforce assignments and 30 workforce assignments are conducted, respectively are represented in Figure 4. The agents depicted with figures in white, dotted white, gray and black represent the regular equivalence levels 1, 2, 3 and 4 respectively where lower levels represent the lower levels of equivalences. Next, in Figure 4, the dynamic change of the preference state matrix $P(t)^T=[\alpha(t), \beta(t), \gamma(t), \delta(t)]$ is shown when DEP is called to update at every specified simulation time advancement (i.e., h). Here, DEP is used to compute the weights of the attributes that are used to compute the position value between each pair of employees. Then, the obtained position values are fed into the EREE algorithm to compute equivalences among each pair of agents to finally conduct the workforce assignment for the incoming projects.
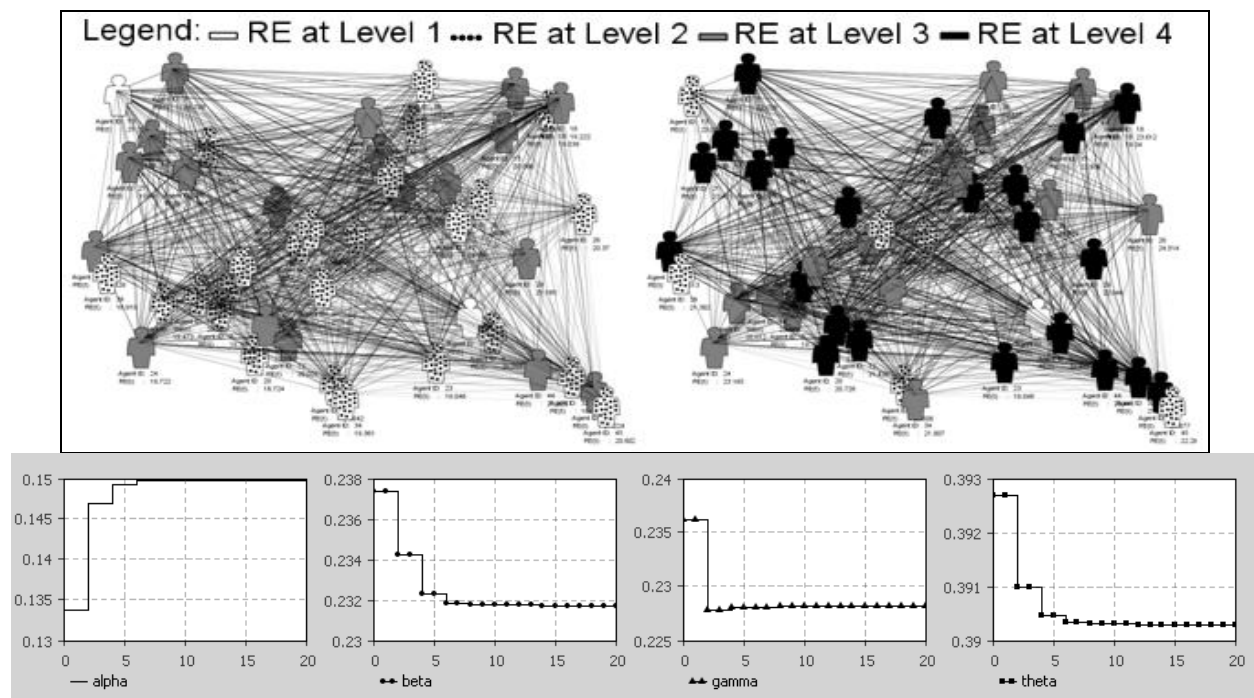

Figure 4: Representation of dynamic change of regular equivalence in the agent based simulation model

In this study, the performance of the proposed position-based workforce assignment methodology is validated against the results obtained from the embeddedness-based and workload-based assignment techniques (Celik et al. 2010). In order to the test the robustness of organizations, 'innovation diffusion' and 'opinion formation' indexes are used. The average values and the 95% confidence intervals of both indexes are computed from 40 independent replications of the developed agent-based simulation. The simulation makes assignments for 30 different projects during each run considering the underlying social network structure. The 'innovation diffusion' and 'opinion formation' results obtained from embeddedness-based workforce allocation and the proposed simulation-based assignment framework are summarized in Table 2, for various types of organizations. Among the functions (the relationship between the total regular equivalence and productivity for individuals of an organization) that are used to test the validity of the proposed approach, functions f1, f2 and f3 are comparable with the same numbered functions tested by Celik et al. (2010). However, the behavior represented in functions f4 and f5 in our study are quite different from the ones tested in the earlier study, hence comparison is not made towards these functions. Values obtained for both 'innovation diffusion' and 'opinion formation' using our proposed assignment framework are substantially lower than the ones obtained from embbeddedness-based allocation for the set of functions f1 through f3. Furthermore, for the same set of functions, the average and maximum improvement in 'innovation diffusion' is observed as 24.28% and 47%, respectively. Similarly, the average and maximum improvement in 'opinion formation' is observed as 15.47% and 31%.

Figure 5 compares the specific 'innovation diffusion' and 'opinion formation' results obtained from the proposed position-based workforce assignment framework with the ones obtained from embeddedness-based and workload-based assignment methods for the case of Kuali against varying α values. As the workload-based assignment does not retain any weights, the same average value attained for 'innovation diffusion' and 'opinion formation' are utilized for varying levels of α. It is observed from Figure 5(a) that while the embeddedness-based workforce assignment outperforms the workload-based assignment in most of the cases (especially for α values greater than 0.8), the performance of proposed position-based assignment framework surpasses the performance of both embeddedness-based and workload-based workforce assignment techniques regardless of the α values. The 'opinion formation' values obtained from embeddness-based assignment gets closest to that of the values obtained from position-based assignment when the α value approaches to 1 where the organization is only focused on its long term goals than its short-term productivity goals. Figure 5(b) demonstrates that the results acquired from the embeddedness-based assignment in terms of 'opinion formation' is similar to the ones obtained from the workload-based assignment and neither one of them shows significant improvement against each other. However, the proposed position-based workforce assignment method shows at least 18% lesser value in 'opinion formation' when compared to the best assessment obtained from any of the said methods.

Table 2: Results obtained from workforce allocation structure and the proposed simulation-based workforce assignment framework in this work considering position notion

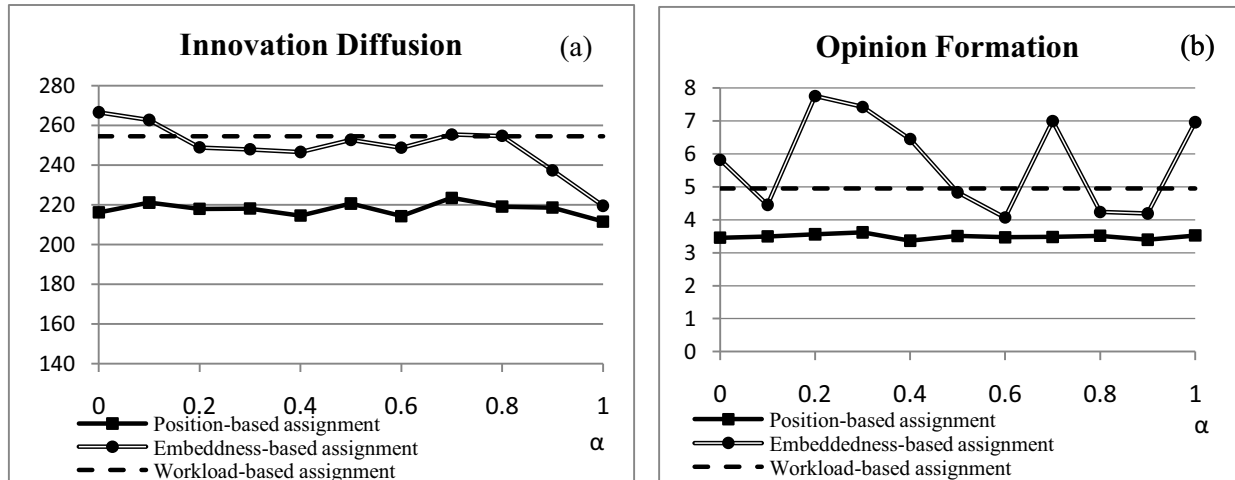| Function	Celik et al. (2010)		Proposed workforce assignment framework	
	Innovation diffusion	Opinion formation	Innovation diffusion	Opinion formation
f1	223.99 ± 3.19	3.94 ± 0.88	218.752±3.434	3.502±0.077
f2	266.03 ± 5.82	3.61 ± 1.02	215.581±3.693	3.522±0.085
f3	314.35 ± 4.91	4.64 ± 1.39	213.746±3.598	3.531±0.047
f4	---	---	218.586±3.356	3.501±0.074
f5	---	---	218.654±3.533	3.532±0.069

Figure 5: Result obtained by the proposed position-based assignment methodology against embeddedness-based and workload-based assignment with varying values of α

5   CONCLUSION AND FUTURE WORK

In this research, we have proposed a novel simulation-based workforce assignment framework based on position notion in a social network to aid project managers of globally distributed software development conduct optimal assignments. The proposed framework operates via two separated (evaluation and assignment) modules and considers both short-term (productivity) and long-term (robustness of the organization against potential departure of key personnel) benefits of the organization. For the evaluation module, an agent-based simulation model is developed to represent the social network under consideration; DEP and EREE algorithms are embedded to this module to calculate the changing position values, and the regular and structural equivalences between each pair of workforces, respectively. The assignment module incorporates a multi-objective optimization model. The use of proposed modeling framework is demonstrated with the software enhancement request process in Kuali, an open source software development organization involving 12 organizations and for various other types of organizations.

The performance of the proposed workforce assignment methodology is validated against the results obtained from the embeddedness-based assignment framework developed by Celik et al. (2010) and the ones obtained from workload-based assignment. The average and the 95% confidence intervals of 'innovation diffusion' and 'opinion formation' values are computed from for various types of organization cultures where the relationship between the total regular equivalence and productivity changes. Preliminary results obtained from Kuali organization show 24% average and 47% maximum improvement in 'innovation diffusion'. Similarly, the average and maximum improvement in 'opinion formation' are observed as 15% and 31%. Furthermore, while the embeddedness-based workforce assignment outperforms the workload-based assignment in most of the cases (especially for α>0.8), the performance of proposed position-based assignment framework surpasses the performance of both embeddedness-based and workload-based workforce assignment techniques regardless of the α values, with at least 18% lesser value in both 'opinion formation' and 'innovation diffusion'. Future studies focus on conducting exhaustive experiments to secure the validity of the proposed framework and enhance its capabilities by including various other factors that might affect the best possible assignment such as the qualifications of the workforce, and the quality and speed of the project.

ACKNOWLEDGMENTS

REFERENCES

Borgatti, S.P., and M.G. Everett. 1993. Two algorithms for computing regular equivalence. *Social Networks* 15:361-376.

Busemeyer, J.R., and J.T. Townsend. 1993. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review* 100:432-459.

Celik, N., S. Lee, E. Mazhari, Y. Son, R. Lemaire, and K. Provan. 2010. Simulation-based workforce assignment in multi-organizational social network for alliance-based software development. Accepted with minor revisions to *IEEE Transactions on Software Engineering.*

DeCanio, S.J., and W.E. Watkins. 1998. Information processing and organizational structure. *Journal of Economic Behavior and Organization* 36:275-294.

Kuandykov, L., and M. Sokolov. 2010. Impact of social neighborhood on diffusion of innovation S-curve. *Decision Support Systems* 48: 531-535.

Latane, B. 1981. The psychology of social impact. *American psychology*, 36:343-356.

Martello, S., and P. Toth.1990. *Knapsack problems: Algorithms and computer implementations*, revised ed. Chichester, NY: John Wiley & Sons.

Mockus, A., and J.D. Herbsleb. 2002. Expertise browser: A quantitative approach to identifying expertise. In *Proceedings of the International Conference on Software Engineering 2002*, 503–512.

Nembhard, D.A., and B.A. Norman. 2002. Worker efficiency and responsiveness in cross-trained teams, *Technical Report*, Department of Industrial Engineering, University of Pittsburgh, 15261, 2002.

Nguyen, T., T. Wolf, and D. Damian. 2008. Global software development and delay: Does distance still matter? *In Proceedings of the International Conference on Global Software Engineering 2008*.

Oshri, I., J. Kotlarsky, and L.P. Willcocks. 2007. Global software development: Exploring socialization and face-to-face meetings in distributed strategic projects. *J. of Strategic Info. Systems* 16(1): 25-49.

Provan, K.G., and R.H. Lemaire. 2010. A matter of intensity: Structural embeddedness in a virtual multi-organizational task network, *University of Arizona working paper*.

Wu, F. , and B.A. Huberman. 2008. Social Structure and Opinion Formation. Available via: <http://arXiv.org/pdf/cond-mat/0407252v3> [accessed April 14, 2010].

AUTHOR BIOGRAPHIES

NURCIN CELIK is a Ph.D. candidate in the Department of Systems and Industrial Engineering at U. of Arizona (UA). She received her B.S. in Industrial Engineering from Bogazici University, Turkey and M.S. in Systems and Industrial Engineering from UA. She has received several awards such as the IIE Outstanding Graduate Research Award (2008) and the Best Ph.D. scientific poster award in the Ph.D. Colloquium at IERC2009. She is the treasurer of INFORMS Student Chapter and a student member of IIE, INFORMS and SME. She can be reached at <nurcinkoyuncu@gmail.com>.

DONG XU is a Ph.D. student in the Department of Systems and Industrial Engineering at U. of Arizona. He received his B.S. in Logistics Systems Engineering from Huazhong University of Science & Technology, China. He can be reached at <dongxu@email.arizona.edu>.

HUI XI is a Ph.D. student in the Department of Systems and Industrial Engineering at U. of Arizona. She received her B.S. in Automation Measurement and Control from Harbin Institute of Technology, China. She is a student member of INFORMS. She can be reached at <huix@email.arizona.edu>.

YOUNG-JUN SON is a Professor of Systems and Industrial Engineering at U. of Arizona. He is an associate editor of the *Int. J. of Modeling and Simulation* and the *Int. J. of Simulation and Process Modeling*. He has received the SME 2004 Outstanding Young Manufacturing Engineer Award, IIE 2005 Outstanding Young Industrial Engineer Award, IERC Best Paper Awards (2005 in modeling and simulation; 2008 in homeland security; 2009 in modeling and simulation), and the Best Paper of the Year Award in 2007 from *Int. J. of Industrial Engineering*. He can be reached by email at <son@sie.arizona.edu>.