Agent-managed Interaction Recommendation for Effective Large-Scale Collaboration

Xu Zhe, Rajiv Ramnath, Jay Ramanathan[1]

OSU-CISRC-5/12-TR07

ABSTRACT

Extreme-Scale Organizational (ExSO) collaboration is defined here to be across the boundaries of existing organizational structures. It is increasingly encouraged both to achieve dynamic responses and to pursue innovative ideas. At the same time such collaboration is challenging mainly because the necessary knowledge maybe highly distributed *across* organizations. Consequently the traditional organizational structures for the successful matching of resources for collaboration are no longer available. While considerable research exists for querying and locating experts in an existing social network, the needed process of enhancing the network to better connect the experts/collaborators needs further research. The research contributions of this paper are three-fold: 1) a conceptual framework that defines ExSO collaboration and its distinct requirements that need to be researched; 2) the design of a recommendation system that identifies human matchers to help connect individuals to facilitate ExSO collaboration, and 3) analysis of real-world public project collaboration data from a highly successful interdisciplinary research center to identify, characterize and illustrate the role of 'matcher' nodes that enhance a network.

Keywords

Computer Supported Cooperative Work, Agents and Intelligent Systems, Interaction Design.

INTRODUCTION

A new type of collaboration is evidenced in situations that address new requirements that are often dynamically identified and require collaborators and knowledge distributed *across* the organizations. For example: "as a motivation technique, Google uses a policy often called 'Innovation Time Off' where Google engineers are encouraged to spend 20% of their work time on projects that interest them [14]." Other examples are community-related volunteering, first response to emergencies, and interdisciplinary academic projects. Such collaborations have goals and tasks that need to be matched with resources that have the right knowledge. However, an organization structure does not always exist to facilitate this matching effectively. A typical organization provides business resources that support collaborations. This type of provisioning might be at reduced levels (or even absent) when going across organizations. This raises the intriguing possibility of crowd sourcing type techniques for certain tasks completed across organizations, and at the same time, the urgency of making such collaboration efficient. We identify and characterize using examples this important emerging trend that we call: *Extreme-Scale Organizational (ExSO) collaboration* (Table 1).

Organization structure	**Collaboration Tasks, Resources and Project, Motivation**
Highly structured organization	Task related to a business need, completion is based on contracted goals. (Examples: Product development, Of shore development). Investment in resources is maximized with structured management organization and workflow.
Combination of crowd sourcing across organizations and structured organization	ExSO: Task related to a customer need, completion is based on individuals across organizations, individual's motivation and business need. (Examples: pursuit of new ideas, ad-hoc collaboration, and community service.). Cost lower, wide access to knowledge, innovation.
Crowd sourced	Task definition flexible, completion is based on individual's motivation. (Example: Wiki). Cost reduced, wide access to knowledge and talent.

Table 1: Characterization of ExSO collaboration.

Note as illustrated in the first row of Table 1 that an organization traditionally provided resources for project. Here we explore technology–mediated ExSO collaboration with potential crowd-sourcing benefits [15]. We begin by noting one main goal of any organization is to provide the management nodes that are the 'matchers' of collaborators that work on tasks. This structure is missing when potential participants are drawn from across organizations. In these situations, existing methods for ExSO matching are limiting and time consuming because these matcher intermediaries are limited (by whom they know or by which social event they happen to know about and attend hoping that something will come out of it).

Our overall research hypothesis here is that we can collaborate more effectively across organizations if we can successfully identify individuals that are good 'matchers' and possess 'matching credentials'. These matchers can help establish new connections that *augment* an existing social network. The connections serve as a virtual organization that can better facilitate ExSO collaborations. Thus the underlying essential challenge of matching collaborators needs us to consider organizational aspects beyond what is addressed in other research based on network structure query approaches [4, 5, 6, 7]. We explore from the following perspectives:

[1] Corresponding author: Ramanathan.2@osu.edu

- Organizational roles are critical in traditional management but absent in enabling ExSO collaboration across organizations.
- Organizational roles provide a deeper understanding of node association patterns in a social network. When we analyze the network structure we also get insights as, we shall show here, from the role sequences reflecting individuals that interact. Each organizational role represents agents that act in some consistent pattern.
- It is easy to get accurate organizational role information. Most organization haves fairly universally accepted organizational labels for each person. As for other attributes like 'motivation', it is usually hard to get such datasets. For example, even getting a complete dataset of friendlists is not as easy as it seems. Some friendships are not recorded on documents due to privacy issues. It is also difficult to get a complete and accurate dataset of profile keywords since people do not include keywords of their expertise. Also there are accuracy challenges in applying natural language processing techniques to extract keywords may fail in accuracy.

Here we will motivate and propose an algorithm to identify matchers by exploring collaboration association patterns based on organization roles. We demonstrate using publicly available project data from the arguably successful CMU Robotics Institute site [16], that the matcher role has *distinct* characteristics and such individuals can be detected and recommended as the human-in-the-loop connection-making nodes essential to enable ExSO collaboration.

The outline of this paper is as follows: we first characterize ExSO collaboration as a distinct type and discuss related work; subsequently we propose the ExSO meta model which also introduces the important role we call 'matcher'; we then discuss the design of the ExSO recommendation system that identifies matchers in the data set and helps connect potential collaborators in order to initiate their collaboration; and, finally, we present the roles in a real-world data set and observations about the organization based on its analysis.

EXSO CHARACTERIZATION AND RELATED WORK

In this section we argue that a *core challenge* of ExSO collaboration is to find the human matchers that have connection-making skills. While other knowledge management skills are also important, we focus on the key one of connection making without this the potential collaborators (abbreviated as 'observers') will not even be connected. We also identify other related work next.

Exso Examples and Problem Characterization

Specific kinds of networking organizations - communities, associations, special interest groups etc. - often exist to facilitate ExSO collaborations. But even participating in these to locate collaborators are found to be extremely time consuming and connection–making chancy. However, with the widespread use of Internet for collaboration and social media, the beneficial potential ExSO collaboration in public, private and academic organizations now needs to be explored. More generally, based on several case studies we have found that the discovery of connections can take many months in a very time consuming way or never occur at all.

To illustrate ExSO we identify just a few specific manifestations of this type of collaboration need. First, in typical vendor enterprises multiple teams often touch the same customer organization at multiple points without knowing of each other's engagement and existence. Another extreme challenge also presents itself among first responders. For example, in Central Ohio hurricane Ike's aftermath required coordination across agencies, private-sector organizations, and community leaders based on local resident-provided knowledge (such as areas of greatest damage). In this case, requests came into different triage centers (like the city's 311 one-stop call center. 311 is a call ceter providing non-emergency city services and city informatin) and many of the requests were actually non-routine. Consequently, the actual matching of need to collaborators was done by humans through communication in a command-central room and outside of the computation systems.

We use the command room analogy above to introduce key project and agent role concepts characteristic of ExSO. This room is in effect a select group of *matchers* that facilitate connections between those that need to interact and who call in from the field. Without such networking across silo organizations, it is difficult to access individuals with knowledge to help complete value-adding interactions effectively. The initially unconnected individuals that are not yet working together are called *observers*. Those that have worked on a project previously are *collaborators*. ExSO *recommendation system* presented here introduces matchers into an existing network. The result of a successful match evidenced by resource sharing and project completion is called a *collaborato*r.

Related Work

The need for large scale collaboration has been very recently identified [15] and architectural frameworks have been proposed [4, 5, 6, 7]. In addition to agent-based architectures and deployed systems, relevant research falls into the following general areas – knowledge management and project environments for collaboration, network query and traversal, and so on. Searching for expertise is currently considered very important. Paper by Zhang et al. [1] compares algorithms such as breadth first search in social networks by simulating on an email dataset. A review of expertise locator systems that locate experts from the web is in [2]. A study on how external factors such as a job role could shape expertise search is conducted in [3].

Aardvark [8] (acquired by Google), exhibits the human-in-the-loop aspect. Here the Aardvark using the social network context of humans likely to know the answer answers questions. This is a recommendation system that search for the right experts to answer people's questions. Researchers have also explored how agent-based model could assist collaboration [9, 10].

Complementing this work, our key point is that some important ExSO project needs have to be researched. For example, a typical social network (with individuals as nodes) has a friends' or node-adjacency list for each node. However, as we have shown in Table 1, the additional organizational structure is important to achieve effective collaboration, considering different cost aspects such as time, social, resource and knowledge management. Applicable research includes algorithms that query social networks considering the communication cost. Work-to-date has been in different network search strategies ranging from purely algorithmic ones that visit every node, to those that consider social cost and visit more selectively (e.g. based on higher ranking). Preferred query strategies keep in mind the negative social impact of explicit communication. That is, if communication is implemented by emails, then we wish to minimize emails. Different traversal strategies result in different numbers of emails.

Other related work [11] is in motivational aspects of the individual, security [12]. Some research shows boundary objects are important in constructing project teams [13]. These are related to current work.

Knowledge management within an organization has been well researched. However, there is less work related to the process managing knowledge spread across organizations. In particular individuals in social networks fail to locate the correct experts because of their local knowledge - not everyone knows what everybody else 'knows, wants or does'. Thus a search path in the real world to locate expertise is actually a series of local knowledge exchanges and interactions among individuals. Virtually, agent-based modeling (ABM) approach is suited to ExSO. Here each node in a social network is considered an agent with local instead of global knowledge about projects. The recommendation of matchers is based on global knowledge that complements the more local knowledge and promotes efficient interactions among agents based on certain principles and rules provided herein.

Thus, here we ask here key questions towards successful ExSO collaboration: What matcher-related features are present in successful ExSO collaboration networks? What kinds of network characteristics must be present for knowledge management to be more successful? How can we design a recommendation on these matcher nodes to make the ExSO collaborations more successful?

ExSO CONCEPTUAL MODEL

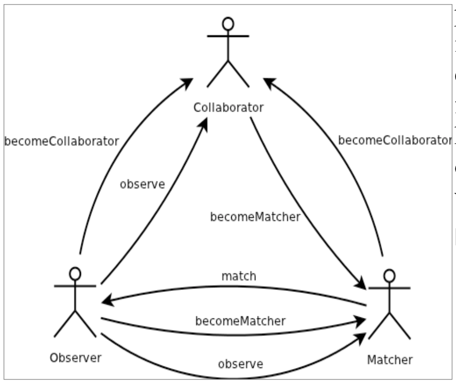
As mentioned above we use three project agents roles to characterize ExSO collaboration: collaborator, observer and matcher. Each collaboration is conducted in the form of project. We use the term "project" to describe the association of a problem with all the required and discovered resources along the way needed to implement. An agent could assume multiple roles and work on multiple projects simultaneously. However, when we talk about the role of an agent, we refer to it in the context of a project.

A collaborator is an agent that is collaborating with other agents associated with a project. A collaborator is assigned with a project. She provides the resources and implements the project. The benefit to a collaborator is that through project completion and value production, there is a strengthening of capability and reputation etc. Collaborators are highly motivated when their work products have been successfully delivered. They often become active observers and matchers in future projects.

An observer is an agent that observes without getting committed to a given project. Before becoming a collaborator or matcher, the agent is an observer. Observers may be interested in certain projects. They observe to decide whether to become collaborators or matchers later. The cause for hesitation towards commitment comes from the lack of motivation, skill, etc. Observers having previous successful project experiences tend to be more highly motivated to get committed. In other words, they would switch to the role of collaborator or matcher more quickly. On the other hand, having failure in previous collaborations makes observers cautious in choosing projects. They would make better decisions due to lessons they have learned.

A matcher matches two observers and assists them to become collaborators based on her knowledge about the observers. A collaborator could also be a matcher for a given project. Matchers know individuals that have an interest or resources that contribute towards solving the problem (skills, time, assets etc.). The benefit to matcher is the ability to create value (e.g. reputation) indirectly without direct provisioning of resources (e.g. time, facilities). Benefit increases if the project succeeds without investment of own resources. If the project fails, matchers' reputation may get jeopardized. But the failure also provides useful experiences. Figure 1 shows the relationships between three agent roles.

An ExSO recommendation goal is to obtain all project resources quickly. If this takes too long the project will starve or fail, as will the population of participants. Thus one of the goals of the recommendation is to find matchers in the shortest time



possible without disturbing agents needlessly. For example, finding interdisciplinary faculty to collaborate even within the same university is time consuming and involves talking to multiple matchers (for example Robotics faculty and adjunct faculty). Within traditional organizations, management hierarchies attempt to help with this type of matching. However, within large organizations even with management nodes, the "right hand often does not know what the left hand is doing".

Figure 1: ExSO Agents

ExSO Matcher Agent

An agent role could be either a human being or a machine. But in this paper, we focus on the matchers, especially human matchers. These are essential to providing the global knowledge within the highly distributed ExSO. Our rationale is given next.

Matcher Human-in-the-Loop: Specifically, a human matcher interacts with observers (those that are uncommitted to a collaboration) in ways irreplaceable by a computerized interaction. This is because machines might tell you "who" to collaborate with. But without additional explanations - they eventually fail to explain "why" as well as the human matcher. This is a well-known limitation of applied artificial intelligence. This explanation is particularly important in the early stage of communication between potential collaborators. A solid explanation of "why" could actually precipitate a collaboration to happen, where none existed before. A machine agent does not do better than a human-in-the-loop even in the case they both have the same static knowledge, let alone in ExSO situations that are typically dynamic and non-routine. Problems include understanding questions, searching for knowledge, decision making based on the questions and presenting answers. Techniques to solve these problems are mostly from AI such as natural language processing, machine learning and knowledge management [17].

Credit history: Good matchers have knowledge attributable to an observer based on past performance, tasks requirements, and so on. Credit (the term we use) is the result of a complex function based on many factors. For instance, a matcher may judge an observer's credits according to projects complete, skills, possession of resources., and so on. The matcher may also reach a conclusion by hearing opinions from other matchers. Humans are better than machines for assigning credits based on objective and subjective information. This knowledge is difficult to acquire as it evolves. However, successful human matchers possess this knowledge. Thus humans-in-the-loop are key to connecting observers with the right skills with those that are required by another observer seeking to collaborate. While the matching might be difficult for machines, keeping track of some of the history in profiles is not.

Cognitive perspective: Two people may seem to be fine working together according to the matching of resources. However, they may fail to collaborate due to some other human factors. For instance, they may have different tempers that do not fit each other. And it could badly jeopardize the collaboration. This is something machine fails to capture.

Here we design the recommendation system that manages the interactions among the human-in-the-loop matchers, observers, and collaborators.

ExSO ABM RECOMMENDER DESIGN

The recommender designed here algorithmically identifies human matchers so that they can be injected into a social network. This provides a virtual organizational element that locally facilitates the connection making through the knowledge advantages that they possess.

Problem definition: We define the problem of identification of human matcher as follows. For any two nodes p and q in a given network that do not collaborate with each other, we define the function *match* (m, p, q) as m matches nodes p with q. Finding the matchers of p and q is to find a set of nodes $M = \{m_1, m_2, ..., m_n \mid match\ (m_i, p, m_{i+1}), 1 <= i <= n\text{-}1, match\ (m_n, p, q)\}$ in the network such that M converts p and q to collaborators for the specific project P.

That is multiple matchers might be needed and the process of matching is iterative: the first matcher matches one potential with a 'second matcher'. And he second matcher matches the potential collaborator with a 'third matcher', and so on. This process terminates under two situations: (1) The second potential collaborator (observer) is reached (successful matching); (2) No more matchers could be added to M and the other second potential collaborator is still not reached (failed matching). It is also important to know that in the above definition p and q could be both collaborators for distinct projects when *match* is applied. The set M could itself contain p and/or q (e. g. p knows q already). A matcher can be a machine or a human agent. As we discussed in previously, we are interested in the human matchers.

Algorithm for Identifying Human Matchers using organizational roles: So how do we identify the actual set of matchers M? The nodes in a social network have many attributes, for instance, profile keywords or friends lists. In a comprehensive algorithm, these attributes would be weighted and used. However, we begin this research exploration with the essential roles.

Next we introduce the definitions to allow us to detect role combination patterns. A *role pattern* is composed of a sequence of organizational roles played by participants in that project. Let a project collaboration instance be C. We have participants c_1, c_2, …, c_k in C. Each participant c_i has a corresponding organizational role r_i for $1 <= i <= k$. And we say C presents a collaboration pattern $r_1 + r_2 +\ldots + r_k$ according to organizational roles. There is no order within the same pattern. There will be many instances of this collaboration pattern that we aggregate as a role pattern with a overall fractional contribution of this pattern. A role pattern thus shows the particular combination of organizational roles made up of participants, and its frequency.

Our algorithm is based on the following observations. For *match* (m, p, q), the connection between m and 'p' is stronger between that between 'p' and 'q'. And the connection between m and 'q' should also be stronger than that between 'p' and 'q'. Otherwise, introducing m to 'p' and 'q' will make no sense since weaker connections are of no use to encourage collaboration. In other words, m knows or interacts with 'p' or 'q' in a way better than the direct communication between 'p and q'. And this is how matchers bridge the gap.

Another observation is about the proportional frequency of a role pattern in the total set of patterns. If a pattern is more successful in achieving project completion than others, a role-based recommendation system should encourage observers using that successful pattern more often than other patterns. Many different measures can be used to define how successful a pattern is. We call these measures fitness measures. We use the frequency of the role pattern to indicate its fitness in this paper. This is based on the observation that successful collaborations tend to have large number of occurrences.

We define the fitness of a pattern more formally as follows. For a corpus of T collaborations, let the success of the role pattern R be F(R). Let the number of collaborations that follow R be T_R. $F(R) = T_R / T$.

The algorithm to identify matchers uses the above role patterns starting from an empty set M. The input to the algorithm is p and q. The output is M matcher individuals to match p and q. In each step that the algorithm proceeds, one node is added to M. The size of M keeps increasing until one of the conditions is satisfied: p is successfully connected with q, or there are no more nodes, which can be added into M. The criteria to choose a node for adding to M is based on the current set of M and fitness values of role patterns. Let the current M be $\{m_1', m_2', ..., m_s'\}$ and their corresponding organizational roles be R $\{r_1', r_2', ..., r_s'\}$. Let the organizational roles of p and q are r_p and r_q separately. From the corpus of patterns we use all the patterns that follows the form: $r_p + r_q + r_1' + r_2' + \ldots + r_s' + rx$. Recall, these patterns were sorted according the fitness value. Therefore we choose the one having the largest fitness value. And we randomly pick up a node playing the corresponding role r_x and add that node into M.

The random strategy will obviously not always work well for individual nodes. However, for now, this paper focuses on the overall populations of agents. And, we are going to show in the case study that our algorithm provides useful insights from the perspective of the population under consideration. In future work, we are going to combine organizational roles with other features to improve the recommendation for individuals.

In the next section we will give an example application of the algorithm.

EXSO CASE STUDY

The goal here is to use real-world data to 1) identify that matcher-related features exist in successful ExSO collaborations, and 2) rationalize the conceptual model for thinking about the problem of ExSO projects.

Selection of the Data Set

It would be ideal to have a dataset that shows organizational roles of project-based collaborations. However, we did not find such a one existing. We selected Carnegie Mellon's Robotics Institute projects data set which has important characteristics: 1) Publically available project collaborations records from an arguable successful inter-disciplinary center [16], and 2) about 5000 projects records within the same environment. The original data does not contain information about the organizational roles. It only listed the names of collaborators in each project. However, form this we were able to extract the actual organizational structure (i.e. the roles: professor, research scientist, graduate student). Extracting the organizational roles of all collaborators in 5000 projects requires huge manual and work and is unnecessary. Thus, we took the last 150 records and identified the organizational roles by searching their profiles in the CMU site manually. A project is on-going or complete but does not affect fitness in our example.

Data Characteristics: We found there are about 300 people involved in the 150 projects. 21 of these roles were identified as listed in Table 2. Some role profiles could not be found by searching, so we gave the corresponding individuals the role "other". We believe it is reasonable to consider the "other" as people outside CMU, since the CMU site provides a search engine that has high coverage for searching people. Another role needing explanation is "external collaborator". It is a label provided by CMU to indicate they come from outside. We classified 21 organizational roles into 3 categories depending where they originated from: Robotics Institute in CMU, other institutes in CMU and other organizations outside CMU. We also calculated the frequency of each role (Table 2). The frequency is generated by dividing the number of projects containing that role by the number of all projects.

Some other role patterns seem confusing at the first glance since organization structure does not help much in explaining them. These patterns are happening across organizations. Our algorithm can effectively help understand how these patterns happen. One example is “other + Robotics visiting scholar”. This one has pretty high fitness value and the cause of this pattern is not self-evident. “Other” is somebody who is supposed to be outside the Robotics Institute. And a Robotics visiting scholar is expected to work with people within the Institute. Collaboration could happen between these two roles. But from an organization perspective, this pattern should not be so frequent. Yet it ranks 7th in the list.

Category	Organizational Role	Frequency
Robotics Institute in CMU	Robotics Faculty	0.76
	Robotics Student	0.41
	Robotics Adjunct Faculty	0.33
	Robotics Visiting Scholar	0.22
	Robotics Research Programmer	0.13
	Robotics Post doctor	0.11
	Robotics Project Scientist	0.10
	Robotics System Scientist	0.08
	Robotics NREC Commercialization Specialist	0.08
	Robotics Guest Lecturer	0.01
Other Institutes in CMU	Computer Science Faculty	0.15
	Electrical Computer Engineering Student	0.02
	Computer Science Student	0.02
	Mechanical Engineering Student	0.01
	Electrical Computer Engineering Faculty	0.01
	Biomedical Faculty	0.01
	Biomedical Student	0.01
	Human-Computer Interaction Institute Faculty	0.01
	Mechanical Post doctor	0.06
Other Organizations outside CMU	Other	0.44
	External Collaborator	0.10

Table 2: Robotics Institutes (RI) Organizational Roles and Frequencies.

How could this happen? To answer this, if we associated “other + Robotics visiting scholar” with other patterns “Robotics faculty + Other”, “Robotics faculty + Robotics visiting scholar” and “other + Robotics faculty + Robotics visiting scholar”, things get clearer. As we mentioned earlier, the matcher should have a stronger connection with each targeted observer compared to the connection between the observers. Fitness value is an indicator of the strength of this connection. High fitness value implies strong connection. Either “Robotics faculty + Other” or “Robotics faculty + Robotics visiting scholar” has a larger fitness than that of “other + Robotics visiting scholar”. Thus, “Robotics faculty” is a matcher for “other” and “Robotics visiting scholar”. In other words, due to the assistance of “Robotics faculty”, the “other” population could collaborate frequently and successfully with “Robotics visiting scholar”. This explanation provides a thorough understanding about how a cross-organization collaboration originates from within-organization collaborations.

Observations

We constructed our example query as follows: find the matchers that connect an agent playing the organizational role “other” with another agent having the role “Robotics visiting scholar”.

In order to apply the recommender algorithm above, we first got the role pattern list and the corresponding fitness value from the 150 records. As the complete pattern list contains hundreds of pattern due to combinations, thus we only list the top 30 patterns having the largest fitness values in Table 3.

Observe that the top patterns have no more than 3 distinct organizational roles. According to the algorithm design applied to our query, we are looking for the patterns containing 3 roles of which two of them are “other” and “Robotics visiting scholar”.

And the resulting top pattern we got is “other + Robotics faculty + Robotics visiting scholar”. The algorithm now directs that a good matcher to connect an “other” with a “Robotics visiting scholar” is a “Robotics faculty” due to its high probability. We then introduce this matcher into the interaction between “other” and “Robotics visiting scholar”. Now the set of matchers M is contains a “Robotics faculty” agent. If the collaboration does not work out, it means we need to add another matcher.

Similarly, we sort patterns that have 4 roles and 3 of them are “other”, “Robotics faculty” and “Robotics visiting scholar”. And we take the pattern, which has the largest fitness value, pick up an agent that has the identified matcher role in the top pattern and add it to M. If the matching does not change the collaboration status, we repeat the whole process until it is unrepeatable.

How do we know that the “Robotics faculty” identified by the recommendation algorithm is actually the correct matcher for “other” and “Robotics visiting scholar”? We support the algorithms recommendation by looking at the data from an organizational perspective. From Table 3, we can see some patterns can be derived directly from organization structure. One pattern is “Robotics faculty + Robotics visiting scholar”. Another pattern is “Robotics faculty + Other”. These two patterns are within the top 5 patterns, which have high fitness values. It means they are quite successful role patterns. And they fits our expectation of the Robotics Institute as an organization: a Robotics visiting scholar should work with a Robotics faculty member and Robotics faculty members are encouraged to collaborate others outside the institute. These patterns are representative examples of collaborations clearly fitting with the organization structure.

Role Pattern	**Fitness**
Robotics Faculty + Robotics Student	0.35
Other + Robotics Faculty	0.29
Robotics Adjunct Faculty + Robotics Faculty	0.19
Robotics Faculty + Robotics Visiting Scholar	0.19
Other + Robotics Adjunct Faculty	0.17
Computer Science Faculty + Robotics Faculty	0.15
Other + Robotics Visiting Scholar	0.14
Other + Robotics Faculty +Robotics Visiting Scholar	0.12
Robotics Faculty + Robotics Research Programmer	0.10
Robotics Adjunct Faculty + Robotics Student	0.10
Other + Robotics Student	0.09
Robotics Faculty + Robotics Project Scientist	0.08
Robotics NREC Commercialization Specialist + Robotics Faculty	0.08
Robotics Adjunct Faculty + Robotics Postdoctor	0.08
External Collaborator + Robotics Faculty	0.08
Robotics Faculty + Robotics System Scientist	0.07
Robotics Faculty + Robotics Postdoctor	0.07
Computer Science Faculty + Robotics Student	0.07
External Collaborator + Other	0.07
Computer Science Faculty + Robotics Faculty + Robotics Student	0.07
Robotics Adjunct Faculty + Robotics Faculty + Robotics Student	0.07
Other + Robotics Research Programmer	0.07
Other + Robotics Faculty + Robotics Student	0.07
Other + Robotics Postdoctor	0.07
Robotics Adjunct Faculty + Robotics Faculty + Robotics Postdoctor	0.06
Computer Science Faculty + Robotics Adjunct Faculty	0.06
Computer Science Faculty + Other	0.06
Computer Science Faculty + Robotics Adjunct Faculty + Robotics Faculty	0.06
Other + Robotics Adjunct Faculty + Robotics Postdoctor	0.06
Computer Science Faculty + Other + Robotics Faculty	0.06
External Collaborator + Robotics Student	0.06

Table 3: Collaboration Patterns and Corresponding Fitness.

Using the same method, we discovered a few more matchers in Table 4. The matcher column indicates the matcher identified from the corresponding pattern. You can see that Robotics faculty plays an important role as a matcher, and this follows our intuition.

More generally we have shown that when agents exhibit the ability to work between organizations, their original organizational roles can be algorithmically analyzed and the individuals can be identified and recommended as matchers.

Category	**Pattern**	**Matcher**
Robotics Institute in CMU	Robotics Adjunct Faculty + Robotics Faculty + Robotics Postdoctor	Robotics Faculty
	Robotics Adjunct Faculty + Robotics Faculty + Robotics Student	Robotics Faculty
Other Institutes in CMU	Computer Science Faculty + Robotics Adjunct Faculty + Robotics Faculty	Robotics Faculty
	Computer Science Faculty + Robotics Faculty + Robotics Student	Robotics Faculty
Other Organizations outside CMU	Other + Robotics Faculty + Robotics Student	Robotics Faculty
	Other + Robotics Adjunct Faculty + Robotics Postdoctor	Robotics Adjunct Faculty
	Computer Science Faculty + Other + Robotics Faculty	Robotics Faculty

Table 4: Patterns with Matcher Identified

CONCLUSIONS

There are many examples of the type of collaboration we call ‘ExSO’ or Extreme Scale Organizational collaboration. For example, finding interdisciplinary faculty to collaborate even across the departments of the same university is time consuming and involves talking to multiple 'matchers'. Within traditional organizations, management hierarchies attempt to help with this type of matching. In this paper we characterize and study the problem of ExSO collaboration. An ExSO collaboration occurs across organizations.

Thus across organizations an ExSO recommender is needed to facilitate projects by inserting matchers into the social network.We propose an agent-based model to characterize ExSO recommender that identifies matchers that convert observers into collaborators that complete project tasks. A human matcher is an important concept that augments networks to promote collaborations. We also introduce a matcher identification algorithm to identify role patterns and use them in to suggest matchers. A case study based on the project dataset demonstrates our method and illustrates how the recommender identifies matchers.

More generally we show that an environment to facilitate ExSO collaboration not only requires an understanding social network structure, but also 'knowledge management' processes related to the matcher, collaborator, observer roles along with credibility established due to previous project successes, other interests and motivations of individuals. The ABM model for ExSO provides a foundation towards a deeper understanding of how population of matchers, observers and collaborators could evolve in a network. Leveraging this in real-world applications is the focus of our future research.

REFERENCES

[1] Zhang, J., and Ackerman, M. S. Searching for expertise in social networks: a simulation of potential strategies, GROUP '05.

[2] Becerra-Fernandez, I. Searching for experts on the Web: A review of contemporary expertise locator systems", ACM Transactions on Internet Technology (TOIT) 2006.

[3] Ehrlich, K., and Shami, N. S. Searching for expertise, CHI'08.

[4] Schilling, M. A., and Phelps, C. C. Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation, Management Science, Volume: 53, Issue: 7, 2007.

[5] Hawryszkiewycz, I.T. System requirements for large scale community collaboration, Collaboration Technologies and Systems (CTS), 2011

[6] Wiggberg, M., and Daniels, M. Reflecting on running large scale student collaboration projects, Frontiers in Education Conference, 2008. FIE 2008.

[7] Yang, S., Jiang, J., and Shi, M. A Scalable Framework for Large-Scale Distributed Collaboration, Computer Supported Cooperative Work in Design, 2006.

[8] Horowitz, D., and Kamvar, S. D. "The anatomy of a large-scale social search engine", WWW '10.

[9] Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems, National Acad Sciences 2002.

[10] Cabri, G., Ferrari, L., and Leonardi, L. Agent role-based collaboration and coordination: a survey about existing approaches.Systems, Man and Cybernetics, 2004.

[11] Huberman, B. A., and Loch, C. H. Collaboration, motivation, and the size of organizations, Journal of Organizational Computing and Electronic Commerce, Volume 6, Issue 2, 1996

[12] Van Veenstra, A. F., and Ramilli, M. Exploring information security issues in public sector inter-organizational collaboration, EGOV'11.

[13] Phelps, A. F., and Reddy, M. The influence of boundary objects on group collaboration in construction project teams, GROUP '09

[14] http://en.wikipedia.org/wiki/Google#Employees

[15] Ramanathan, J., Ramnath, R., Xu, Z., Bolinger, J., Herold, M., Khuc, V., Lynch, T., Raje, S., and Shivade, C., Extreme-Scale Social (ExSo) Collaboration: Ontology, Challenges and Opportunities, CETI Tech Report 2011

[16] http://www.ri.cmu.edu/

[17] O'shaughnessy, D., Speech communications: human and machine, Universities press, 2000.