Supporting Network Formation through Mining under Privacy Constraints

Florian Skopik, Daniel Schall, Schahram Dustdar
Distributed Systems Group
Vienna University of Technology
Argentinierstraße 8/184-1, A-1040 Vienna, Austria
{skopik|schall|dustdar}@infosys.tuwien.ac.at

***Abstract*—Single professionals and small companies come together and form virtual communities to compete with global players. In these collaboration networks, the actual business partners are discovered and alliances formed on demand. However, it is impossible for single members to keep track of the dynamics in large-scale networks. With the wide adoption of service-oriented architectures (SOA), interactions between partners have become observable. Monitoring collaborations enables the inference of social relations and the identification of successful partner compositions. Measuring the quality of social relations, such as the degree of trust based on the success of past interactions, are a powerful means to support the formation of alliances. However, by applying monitoring, also privacy concerns arise. In this paper we deal with concepts and tools to support group formations. We consider the trade-off between the benefits of sharing personal profiles and accounting for privacy concerns of the individual network members.**

***Keywords*-interaction monitoring, trust inference, group formation, privacy issues, service-centric collaborations**

I. Introduction

Small and medium-sized organizations create alliances to compete with global players, to cope with the dynamics of economy and business, and to harvest business opportunities that a single partner cannot take. In such networks where companies, communities, and individuals form virtual organizations, collaboration support is a major research track. Individuals and companies that are interested in collaborations register at collaboration portals, where they can flexibly discover partners to form temporal alliances [1]. The collaborations in such networks usually span numerous individuals distributed over various organizations and locations. Due to the scale of these networks it is impossible for the individuals to keep track of the dynamics in such networks.

However, the recent adoption of service-oriented concepts and architectures permits the (semi-)automatic inference and management of member profiles and social network structures [2]. In particular, SOA provides the functional means to allow loose coupling of entities and monitoring of interactions for inference of relations through mining logs. Hence, we use SOA to support and guide human interactions; see Human-Provided Services [3]. Thus, negative influences, such as using outdated information, do not exist compared to manually declared relations. Moreover, monitoring of interaction behavior allows timely adaptations in ongoing collaborations, for instance, updates of member profiles based on successes in recent collaborations and collected experiences, without major user intervention. We focus on supporting group formations in virtual environments by accounting for the individuals' social relations, especially *social trust* [2], [4], [5]. Trust reflects the expectation one actor has about another's future behavior to perform given activities dependably, securely, and reliably based on experiences collected from previous interactions.

Use case scenarios for applying *Trustworthy Group Formation* include (i) *Team Formation in Collaboration Environments*, mostly relying on recent collaboration behavior, previous successes, and member recommendations; and (ii) *Social Formation of Campaigns*, mainly focusing on people's interest similarities for targeted social campaigns.

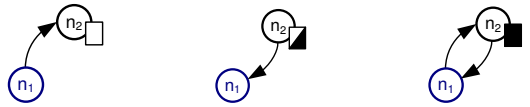
In this work, we introduce concepts and tools to facilitate the formation processes by allowing network members to browse the *Web of Social Trust*. This enables the users to discover trustworthy partners and to study their shared profile information. Hence, members initially providing more information to others are more likely to be able to set up collaborations. However, privacy of members has to be maintained. Thus, it is crucial to account for a balance between sharing and protecting sensible profile information. Finally, two major aspects have to be considered: (i) *which* profile information is shared to facilitate the set up of future collaborations, (ii) with *whom* is this information shared in order to maintain privacy.

II. Privacy-aware Group Formations

We deal with supporting formations, e.g., composing teams and creating communities, in social and collaborative environments, where single members are connected through a *Web of Social Trust* [5]. Our approach supports people who perform formations with the following features: (i) *Dynamic Network Member Profiles.* Dynamically adapting profiles reflect previous collaboration successes, preferences, behavior, and collected experiences. (ii) *Profile Sharing.* Profiles are shared with network members to facilitate collaborations. However, to maintain privacy, information is shared with trusted members only. (iii) *Collaboration Network Visualization.* Relations between network members emerge when performing joint activities. Recently successful compositions are visualized for reuse in future collaborations.

Allowing users to browse the *Web of Social Trust*, e.g., to study relations and previously successful compositions, raises several privacy concerns. Usually, sensible data is

shared with close partners only, while general information may be shared with a larger group. In this work, we distinguish three levels of information sharing: (i) *basic profiles* describe some fundamental properties such as name and contact details; (ii) *success stories* comprise information about previous successful activities that have been jointly performed with other members; (iii) *relations to partners*, i.e., referees, are gathered through mining of interactions in prior success stories.

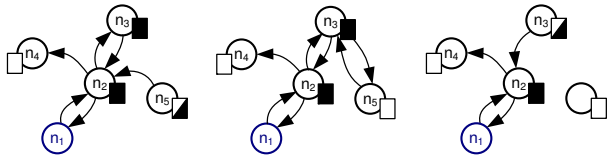


(a) Basic profile. (b) Success stories. (c) Relations.

Figure 1. Fundamental patterns for sharing profile data.

Figure 1 depicts the fundamental patterns for privacy-aware browsing of the *Web of Social Trust*. They depict mandatory relations in the *Web of Social Trust* for sharing profile information. The empty, half-filled, and full filled document symbols reflect the amount of shared profile information with the requester n_1: basic profiles, success stories, relations. Let us assume n_1 browses the *Web of Social Trust* and wants to retrieve profile information from collaboration partners. The first pattern (Figure 1(a)) allows him to reveal the basic profile of trusted network partners. However, n_2 only shares success stories with n_1 if n_2 trusts n_1 to some extent (Figure 1(b)). Relying on mutual trust in collaborations, n_1 and n_2 both share information about relations to their collaboration partners (Figure 1(c)).

With the fundamental patterns, only profiles, success stories, and relations from direct neighbors can be retrieved. Since this would not allow to sufficiently browse the *Web of Social Trust*, there are more advanced patterns to expand the *Circle of Trust*. Within the circle of requester n_1, members share personal data – even if they are not directly connected to n_1 – but still considering their privacy (Figure 2).



(a) Propagation. (b) Blocking Nodes. (c) Success Filtering.

Figure 2. Advanced *Web of Social Trust* browsing patterns.

Propagation of Profiles (Figure 2(a)) allows member n_1 to browse the profile of n_3. However, there is no direct connection between them, both have a transitive relation through n_2. In detail, because n_3 trusts n_2 and thus shares his profile, and n_2 trusts n_1, n_2 shares his perspective on n_3 with n_1. This propagation mechanism can be interpreted as n_2 *recommending* n_3 to n_1 (e.g., realized with FOAF[1]) and extends n_1's *Circle of Trust*. Propagation is enabled by concatenating fundamental sharing patterns along paths with predefined lengths.

[1]Friend-Of-A-Friend Specification: http://xmlns.com/foaf/spec/

Blocking Nodes (Figure 2(b)) terminate the propagation of information in the *Web of Social Trust*. Profile sharing is restricted to members within a certain distance (i.e., the propagation path length). For instance, if the propagation path has a length of two hops, n_5 does not reveal success stories and relations to n_1, even though a path of mutual trust exists between them. Furthermore, it is not possible to propagate success stories or relations over a node that shares only basic profile information itself (here: n_4).

Success Filtering. (Figure 2(c)) means that only distinguished positive collaboration experiences are explicitly highlighted. Spreading information about unsuccessful collaborations, and low trust relations – a form of defamation – is thereby avoided. For instance, let us assume prior collaborations between n_2 and n_5 were not successful, so n_5 and its relations are hidden from n_1.

III. Privacy in Collaborative SOA

A. Flexible Collaborations in SOA

During collaborations, network members interact, for instance, by exchanging documents. Collaborations in SOA means that all interactions are performed through Web services. Even the capabilities of humans are described by WSDL and communication takes place with SOAP messages (see Human-Provided Services [3] and BPEL4People[2]). In the scenario depicted in Figure 3, the two members n_1 and n_2 perform activity a_1, n_2 and n_3 perform activities a_2 and a_3, and so on. Activities [6] represent interaction contexts, reflected by the dashed areas, that hold information about involved actors, goals, temporal constraints, and assigned resources. Hence, an activity holistically describes the context of an interaction in our environment model [2].

Logged interactions are periodically analyzed and aggregated. Various metrics describe the collaboration behavior and attitude of network members, including responsiveness, availability, or support reciprocity. Finally, interaction metrics are interpreted according to pre-defined domain rules, and the degree of trust between each pair of previously interacting members is determined. The exact mechanism has been studied in [2]. Using interaction logging and mining in collaborative service-oriented environments enables the automatic emergence of social relations, and the dynamic adaptation of individual profiles.

B. Adaptive Profile Sharing Model

As previously mentioned, we enable members to share basic profiles, success stories, and personal relations. For

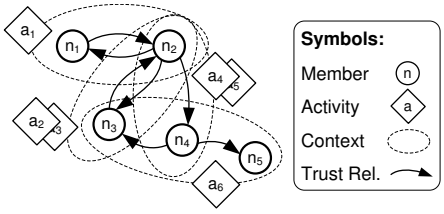


Figure 3. Activity-centric collaboration model.

[2]WS-BPEL Extension for People, Version 1.0, 2007.

that purpose, we utilize three different models (i) the *Basic Profile Model*, (ii) the *Activity Involvement Model*, and (iii) the *Trust Graph Model*. Whenever one member requests profile information about a neighbor, s/he receives parts from respective models with regard to trust paths in the *Web of Social Trust*. Therefore, we do not only allow members to share their own isolated profiles, but also enable the propagation of profiles along transitive relations and sharing of joint activity information.

Trust Graph Model. Let $G_\tau = (N_n, E_\tau)$ be a graph reflecting the *Web of Social Trust* with N_n denoting the set of network members and E_τ the set of directed edges connecting pairs of nodes. Each edge is associated with further metrics that are used to determine trustworthy behavior.

Activity Involvement Model. The involvement of members in activities is modeled as bipartite graph $G_a = (N_n, N_a, E_a)$. It comprises collaboration success stories; in our model successfully performed activities $a_i \in N_a$ and their participating members $n_j \in N_n$. An edge $e_a(n_j, a_i) \in E_a$ reflects that member n_j has been involved in activity a_i. A list of further properties may be assigned to an edge, for instance, the degree of participation or involvement role. Members of finished activities can decide themselves which ones shall be explicitly included in their profiles; thus, providing a personalized view on success stories.

Basic Profile Model. The basic profile, attached to each node in N_n, comprises fundamental data about a member, such as name, organizational affiliations, and contact details. This basic profile is mainly static.

Shared Network Data. Finally, network members share subsets of (i) basic profiles bound to nodes in N_n, (ii) success stories reflected by G_a, and (iii) trust relations in G_τ. Figure 4 shows an example for data shared with n_1 when n_1 incrementally extends its view and requests data about its neighbors. On the left side shared success stories are depicted, while on the right side shared personal relations are shown. Members share different amounts of information with n_1 through propagation according to dynamic trust G_τ for the given scenario in Figure 3. For instance, n_1 has no view on the trust relation from n_4 to n_5, since there is no mutual trust path from n_1 to either n_4 or n_5.

Algorithm 1 deals with sharing of basic profiles in N_n, success stories in G_a and personal relations in G_τ. The shared network segment S contains subsets of data managed by these models, and is incrementally extended. This enables a requester, i.e., the origin node n_o to browse through the *Web of Social Trust* by extending its view, i.e., S, through one of the connected nodes n_e step by step. The functions `predec()`, `succ()`, `neighbors()` provide the predecessors, successors, and neighbors of a node, and `isShared()` determines if a user shares a given activity as success story.Finally `addToS()` extends the shared network information segment S with provided nodes, edges, and activities. For the sake of clarity, we neglect blocking behavior when exceeding threshold distances.

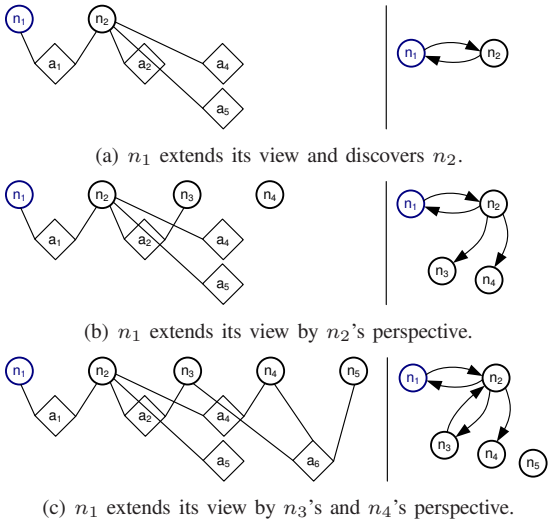


Figure 4. Example applying browsing patterns from n_1's perspective.

Algorithm 1 Dynamically extend shared network data S.

Input: origin node n_o, extension node n_e
Global: $G_\tau = (N_n, E_\tau)$, $G_a = (N_n, N_a, E_a)$, $S = (G^s_\tau, G^s_a)$
function EXTENDVIEW(n_o, n_e)
  /* add basic profiles of all trustors and trustees */
  $N'_n \leftarrow \text{predec}(n_e, G_\tau) \cup \text{succ}(n_e, G_\tau)$
  for each $n \in N'_n$ **do**
    addToS(n, N^s_n)
  /* add success stories */
  $N'_n \leftarrow \text{predec}(n_e, G_\tau)$
  for each $n \in N'_n$ **do**
    $N'_a \leftarrow \text{neighbors}(n, G_a)$
    for each $a \in N'_a$ **do**
      if isShared(a, n) **then**
        addToS(a, N^s_a)
        addToS(edge(n, a, G_a), E^s_a)
  /* add personal relations */
  if (predec(n_e, G^s_τ) $\cap$ succ(n_e, G^s_τ)) $\neq \emptyset \vee n_e = n_o$ **then**
    $N'_n \leftarrow \text{predec}(n_e, G_\tau) \cap \text{succ}(n_e, G_\tau)$
    for each $n \in N'_n$ **do**
      if $\exists$ edge(n_e, n, G_τ) **then**
        addToS(edge(n_e, n, G_τ), E^s_τ)

IV. Evaluation and Discussion

A. Portal Application for Privacy-Aware Formation

We evaluated the efficiency of introduced *Web of Social Trust* browsing patterns with an implementation of a Web-based *Network Browser* shown in Figure 5. This tool depicts the expanded network on the left side, and shared profile information on the right side. Clicking a node reveals the basic profile and some calculated metrics of a member (shown here), while clicking an edge reveals information about joint activities, where this relation emerged. Solid edges represent relations that are described by further metrics used to determine trust; see `Link Property` *Activity Success*. Dashed lines reflect trust relations that exist due to joint success stories, however, associated metrics are not shared with the tool user. The size of nodes and width of edges respectively are proportional to the visualized metrics selected by `Partner Property` and `Link Property`. The formation use

case starts with visualizing the user of the tool as a single (yellow) node. The user is then able to expand the network to discover his collaboration partners. Depending on trust, partner relations and joint success stories with third parties are propagated. So, the user can incrementally discover larger segments of the *Web of Social Trust*. The tool user evaluates the members' profiles, success stories, and community embeddings, and picks single members to join a team or form a group in social campaigns. This step is supported by embedded communication tools; e.g., potential partners can be contacted with an integrated Skype client to negotiate their admission.

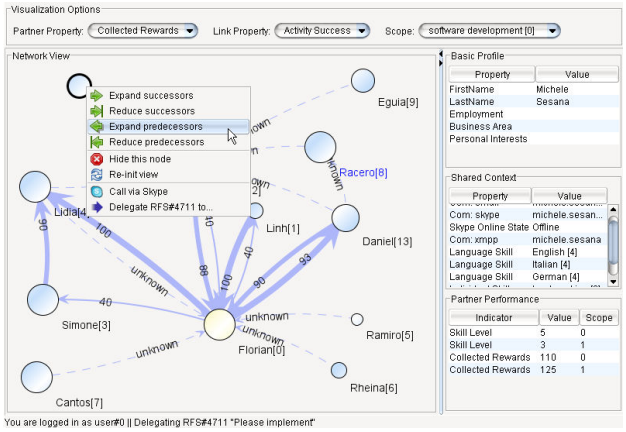


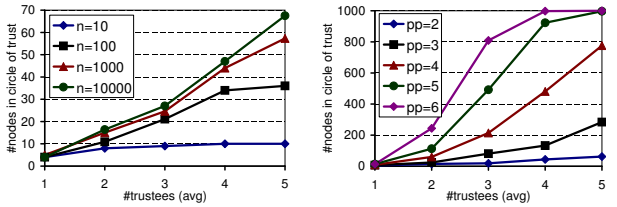
Figure 5. Collaboration network browser.

B. Simulations

We created artificial networks with fixed amounts of nodes and power-law distributed edges [7] to evaluate the effects of propagating profile information. The first experiment investigates the average size of the *Circle of Trust*, depending on the number of trustees for different network sizes n and propagation path lengths pp. For that purpose, we discover for a set of random users recursively all partners who share at least their joint success stories profile. Figure 6 shows that for highly cross-linked graphs (i.e., $\#trustees > 2$), only short pps (max. 3 or 4 hops) are feasible. Otherwise, virtually all members are in the *Circle of Trust*. The second experiment highlights the computational complexity of determining the *Circle of Trust*. While the size of the network does not considerably influence the number of required graph operations (at least for small pp), increasing pp in highly cross-linked graphs leads to exponential costs. Graph operations include retrieving referenced nodes and edges, neighbors, predecessors and successors in G_τ and G_a.

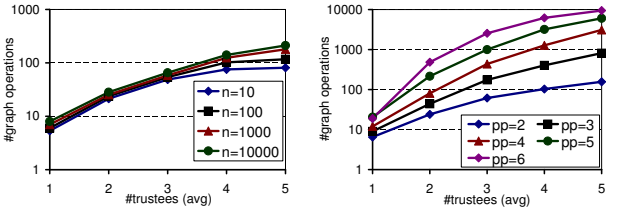
V. Related Work

People interact, including communication, coordination or execution actions, to successfully accomplish their goals in activity-centric environments [6]. Trust [5] is an effective concept to individually rank collaboration partners regarding social criteria. Depending on the environment, it may rely on the outcome of previous interactions [2], [5], and the similarity of skills and interests [4]. Note, trust is not simply a synonym for *quality of service* (QoS). Instead,



(a) Depending on n ($pp = 2$). (b) Depending on pp ($n = 1000$).

Figure 6. Size of the circle of trust with respect to average number of trustees for different network sizes n and propagation path lengths pp.



(a) Depending on n ($pp = 2$). (b) Depending on pp ($n = 1000$).

Figure 7. Required trust graph operations with respect to average number of trustees for different network sizes n and propagation path lengths pp.

metrics expressing social behavior and influences are used in certain contexts. As investigated by [8] trust is strongly related to information disclosure, and thus, privacy.

VI. Conclusion and Outlook

In this paper we discussed the importance of privacy-awareness in social and collaborative networks when sharing profile information. In our case, shared information does not only contain isolated personal profiles, but also joint success stories and personal relations, as common in social networks (see FOAF). Currently, we our concepts in end-user environments of the funding EU project COIN. As a consequence, we will account for user feedback to improve the tool support.

Acknowledgments

This work is funded by the EU project COIN (216256).

References

[1] L. M. Camarinha-Matos and H. Afsarmanesh, "Collaborative networks," in *PROLAMAT*, 2006, pp. 26–40.
[2] F. Skopik, D. Schall, and S. Dustdar, "Trustworthy interaction balancing in mixed service-oriented systems," in *ACM SAC*, 2010, pp. 801–808.
[3] D. Schall, H.-L. Truong, and S. Dustdar, "Unifying human and software services in web-scale collaborations," *IEEE Internet Computing*, vol. 12, no. 3, pp. 62–68, 2008.
[4] J. Golbeck, "Trust and nuanced profile similarity in online social networks," *ACM Transactions on the Web*, vol. 3, no. 4, pp. 1–33, 2009.
[5] T. Grandison and M. Sloman, "A survey of trust in internet applications," *IEEE Communications Surveys and Tutorials*, vol. 3, no. 4, 2000.
[6] P. Moody, D. Gruen, M. J. Muller, J. C. Tang, and T. P. Moran, "Business activity patterns," *IBM Systems Journal*, vol. 45, no. 4, pp. 683–694, 2006.
[7] A. Reka and Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, June 2002.
[8] C. Dwyer, S. R. Hiltz, and K. Passerini, "Trust and privacy concern within social networking sites: A comparison of facebook and myspace," in *AMCIS*, 2007.