A Unified Theory of Trust and Collaboration

Guoray Cai and Anna Squicciarini

College of Information Sciences and Technology
Pennsylvania State University, University Park, PA 16802
{cai,asquicciarini}@ist.psu.edu

Abstract. We consider a type of applications where collaboration and trust are tightly coupled with the need to protect sensitive information. Existing trust management technologies have been limited to offering generic mechanisms for enforcing access control policies based on exchanged credentials, and rarely deal with the situated meaning of trust in a specific collaborative context. Towards trust management for highly dynamic and collaborative activities, this paper describes a theory of trust intention and semantics that makes explicit connections between collaborative activities and trust. The model supports inferring trust state based on knowledge about state of collaborative activity. It is the first step towards a unified approach for computer-mediated trust communication in the context of collaborative work.

Keywords: trust communication, access control, information sharing, collaborative work.

1 Introduction

Sharing sensitive information across security domains is almost always a requirement for knowledge intensive collaborative work, but such sharing may impose risks on security and privacy integrity. Research has increasingly recognized the role of trust in minimizing the risk while gaining efficiencies in information sharing. The challenge of trust management is most apparent in digital government applications [14, 25]. Government agencies naturally form multiple security domains (such as DOD, DOE, USDA, USGS) according to different responsibilities of their services and varying sensitivity of information. Some of the most common types of information being shared among government agencies include intelligence, homeland security, law enforcement, and critical infrastructure information. In the literature, the lack of better support for collaboration and the difficulties of information sharing among agencies have been widely recognized in such applications [10, 27]. While each government agency must be responsible for protecting sensitive information they have collected, effective sharing of information among agencies is deemed to be more important when multiple agencies collaborate under high stake missions, such as dealing with large-scale crisis events. As an example, consider the following scenario of bioterrorism investigation.

E. Bertino and J.B.D. Joshi (Eds.): CollaborateCom 2008, LNICST 10, pp. 425–438, 2009.

Scenario A: The biological attacks with powders containing *Bacillus Anthracic* sent through the mail during September and October 2001 led to unprecedented public health and law enforcement investigations, which involved thousands of investigators from federal, state, and local agencies. Following recognition of the first cases of anthrax in Florida in early October 2001, investigators from the Centers for Disease Control and Prevention (CDC) and the Federal Bureau of Investigation (FBI) were mobilized to assist investigators from state and local public health and law enforcement agencies. The response to the anthrax attacks required close collaboration because of the immediate and ongoing threat to public safety. The steps necessary to identify a potential covert bioterrorism attack include a close coordination between those who collect and analyze medical and syndromic surveillance information with the law-enforcement community's intelligence and case-related information.

In this scenario, public health officials and law enforcement agencies must join their knowledge and expertise to identify the sources and consequences of an attack. Collaboration among agencies across CDC and FBI may be difficult for the following reasons: (1) both health and disease information and criminal records are sensitive information to be protected from unintended use; (2) participating parties may not have prior arrangement on trusting relations; therefore trust may need to be developed on demand; (3) the decisions to be made are not 'business as normal'; hence regular security and privacy policies may not apply.

The distinctive feature of this type of applications is that collaboration, trust, and information sharing are tightly coupled and mutually influence each other over extended courses of collaborative activities. These unique characteristics impose new challenges and demand drastically different approaches for trust management. Existing trust management tools for such applications has shown to be cumbersome and are seldom used in real-world applications. Trust management systems, such as PolicyMaker[8], KeyNote[7], and Trust-χ [5], are merely dealing with generic language and mechanisms for specifying and evaluating security policies, credentials, and relationships. They are almost exclusively designed for business transaction applications where stable and uniform policies for security and access control can be enforced. However, managing trust in the contexts of collaboration and information sharing activities is fundamentally different because the meaning of trust always in flux with the situation of the collaboration. Trust is part of the bigger 'picture' of collaborative activities, and can not be understood outside of that context. Methods for supporting this form of trust management must make explicit connection between trust communication and the on-going collaborative activities. For this reason, we need a unified and coherent theory about trust and collaboration, which can serve as the basis for developing computational methods of trust management.

While collaborative activities impose complexities to trust management, we also see collaboration as (potentially) part of the solution to that problem. Trust communication is an inherently complex collaborative activity (to be carried out to support a more primary collaborative activity such as evacuation planning during a crisis). The approach, adopted in this paper, takes advantage of the close coupling among collaborative missions, information sharing, and trust. Our primary objective is to develop an approach for enabling semantic negotiation of trust within the context of collaborative activities. We see trust communication as a sub-process embedded

within a large discourse of the collaborative activity, which allows the knowledge of the ongoing collaboration to be fully utilized. The approach we envisioned uses agent-mediated human-human trust negotiation as a way to encourage human-machine joint problem-solving and to put human in control for trust decisions.

Before the above vision of trust management in collaborative applications can be attempted, we must first address the fundamental question of how exactly collaboration, trust, and data sharing are related. As such, this paper presents a theoretical system that provides a unified characterization of collaboration and trust. The theory imposes a mental-state view on both collaborative work and trust communication. The outcome of this work is a plan-based representation of collaboration and trust state, together with a set of modal operators and reasoning scheme for advancing the state of trust and collaboration.

2 A Theory of Collaborative Activities

For the purpose of this paper, we consider trust in the contexts of collaborative work. The ultimate goal of building a trust is to enable the success of a collaborative activity. For this reason, our framework includes a model of collaborative activity as a way to capture the context around a trust object.

Collaborative activities have been the subject of cognitive and computational studies for many years. Cognitive and social theories, such as activity theory [24], distributed cognition [16], and situated actions [32], provide language and conceptual structure for describing the settings and systems of collaborative work, but they do not deal with the design of systems that support collaboration. A commonly accepted philosophical view of collaborative activity is Bratman's notion of ***shared cooperative activity*** (SCA) [9]. According to Bratman, three properties must be met for agents to participate SCA: *mutual responsiveness*, *commitment to the joint activity*, and *commitment to mutual support*. These three properties allow agents to have additional mental attitudes as cognitive resources for communication.

An *activity* consists of a series of actions that are executed by one or more agents to achieve a goal. There actions are related through an underlying intentional structure. For a shared cooperative activity, the intentional structure corresponds to a SharedPlan of the collaboration, according to Grosz and Sidner [15]. A SharedPlan is a formal model of collaborative plans. A SharedPlan for an action includes a mutual belief concerning a way to perform this action / subactions, individual intentions that the action and subactions be performed, and its structure is a hierarchical structure of individual plans and SharedPlans. *Actions* may be basic or complex. Basic actions are performable at will, under certain conditions, while complex actions

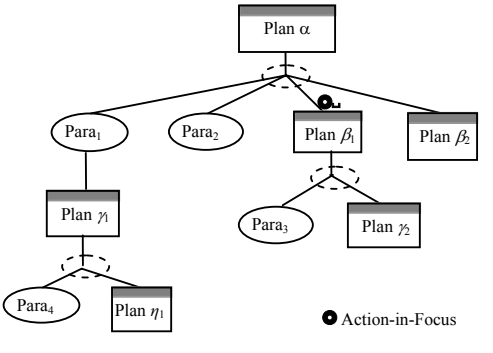

Fig. 1. Structure of an activity represented as a SharedPlan (after[11])

have associated recipes, their performance requiring the performance of each action in their recipe, under certain conditions. *Recipes* represent what agents know when they know a way of doing something. They represent information about the abstract performance of an action, and are composed of a set of constituent acts and associated constraints.

As an example, Figure 1 shows a plan for α, which was elaborated with a recipe that has two parameters ($Para_1$ and $Para_2$) and two subactions β_1 and β_2. Parameters in a recipe are knowledge-preconditions which must be identified before any subactions can be worked on. Identifying a parameter can be a complex process and may need a plan (such as Plan γ_1). A SharedPlan of an activity explains why agents do what they did and is a model of the intentional structure of a collaborative activity. According to Grosz and Kraus [15], a group of agents GR is considered to have a *Full Shared Plan* (FSP) on α when all the following conditions have been established:

0. GR intends to perform action α
1. GR has a recipe *R* for action α
2. For each single-agent constituent action β_i of the recipe, there is an agent $G_{\beta_i} \in GR$, such that
   a. G_{β_i} intends to perform β_i;
      G_{β_i} believes that it can perform β_i;
      G_{β_i} has an individual plan for β_i
   b. The group GR mutually believes 2.a.
   c. The group GR is committed to G_{β_i}'s success
3. For each multi-agent constituent action β_i of the recipe, there is a subgroup of agents $GR_{\beta_i} \in GR$ such that
   a. GR_{β_i} mutually believe that they can perform β_i
   b. GR_{β_i} has a SharedPlan for β_i
   c. The group GR mutually believes 3.a.
   d. The group GR is commited to GR_R's success

Otherwise, the agents' plan will only be a *Partial Shared Plan* (PSP).

SharedPlan theory has been a common framework for modeling collaborative activities (see [1, 20, 21, 26]). However, this theory does not consider the issue of trust among agents. In other words, the theory only works in a fully trusted environment. In fact, agents must place great trust on each other when they elaborate and execute a SharedPlan. We can expect two cases where trust can be an issue:

[**Case 1**] *Mutual dependence on sharing sensitive information*. Agent G_1 may require knowledge about a parameter (e.g., $para_1$) to move forward on its part of the duty, but this piece of knowledge may have to be retrieved from another agent G_2. Such knowledge exchange may be a problem if G_2 considers the piece of knowledge as sensitive and does not have trust on G_1 to make fair use of it.

[**Case 2**] *Mutual dependence on each other's capability to perform actions*. Since the success of a larger activity is dependent on the success of component acts which can be executed by different agents, agents place great trust on each other in their capabilities.

Such need for trust was not made explicit as a research topic in previous work, but will be the focus of the next section.

3 A Theory of Trust and Trust Communication

Here we articulate a theory that explains the way in which trust and collaborative activities are coupled and how they become problems and solutions to each other. This theory applies a mental state view to both trust and collaboration, and allows semantic connections through the intentional structures of collaboration and trust communication.

3.1 Definition of Trust

The concept of trust has been intensively investigated in the fields of philosophy, sociology, psychology, management, marketing, ergonomics, human–computer interaction (HCI), industrial psychology and electronic commerce (e-commerce). These disciplines study domain-specific forms of trust to address different types of problems, and they hardly agree on what trust is. When we come to the task of modeling trust in computing systems, we have a 'pudding of trust' [6] to deal with, each emphasizing different sets of issues. Despite such diversity of definitions, many [18, 19] believe that there exists a conceptual core that provides a general construct to model a variety of senses of trust in different application domains. We share the same view in this study. The search for a conceptual core of the trust concept is still an on-going process, but the current literature seems to suggest that trust can be analyzed as a trust relationship (in terms of "A trust B on doing z") and a set of beliefs (e.g. trustor's belief on trustworthiness of the trustee, etc) associated with such a relationship. All the challenges related to managing trust can be eventually translated into the task of representing and reasoning on the trust relations and associated beliefs.

In this study, we limit discussion to the basic form of trust which involves two agents. We use T (A, B, z) to denote the trust object that captures all the information about "A trusts B on achieving the effect of z."

Definition 1. A trust object T (A, B, z) refers to a trust relation together with a set of mental states (see Figure 2). It has three major components.

T = {TR(A, B, z), MST(A, TR), MST(B, TR)}

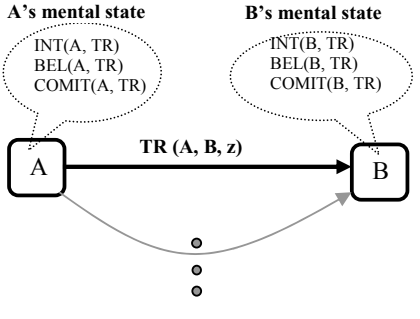

Fig. 2. Triadic relation of trust

(1) **Trust Relation.** A trust object defines a triadic (three place) relation, in the form of 'A trusts B to do z', or TR(A, B, *z*). Here A stands for the trustor (the peer that trusts someone). B stands for the trustee (the peer being trusted). *z* is some state of affair that A hopes to achieve. A places some *dependency* on B to achieve z and A is also risked by the possibility that B not do *z* as expected. A is potentially rewarded if *z* is achieved.

(2) **Mental States of A towards TR.** Agent **A** holds a set of individual mental attitudes towards trust relation TR. Such attitudes are constantly in flux, and they change as more evidences are exchanged and taken into account. The state of mental attitude at any given time is called *mental state*. We use MST(A, TR) to refer to the mental state of A towards trust relation TR. For the purpose of this study, we limit the contents of A's mental state to three kinds of mental attitudes: intention, belief, and commitment, represented as INT(A, TR), BEL(A, TR), COMMIT(A, TR), respectively. One important component of A's belief is BEL(A, TW(B, TR)) which refers to "*A's belief on the trustworthiness of B in relation to TR*." TW(B, TR) is a measure of how much B is able and willing to act in A's best interest. BEL(A, TW(B, TR)) serves as the bases for A to place trust on B. In order to promote A's trust on B, B can manipulate A's belief by communicating the trustworthiness of B, TW(B, TR), to A (through sharing trust-promoting credentials or demonstrating trust-provoking behaviors).

(3) **Mental States of B towards TR.** MST(B, TR) also has three components: INT(B, TR), BEL(B, TR), COMMIT(B, TR), respectively. One important piece of BEL(B, TR is the *trustfulness* of the trustor as perceived by the trustee. This is represented as BEL(B, TF(A, TR)). Trustfulness, TF(A, TR), refers to the capacity of the trustor to take a risk that the trustee will not behave according to a special agreement even if promising to do so [34]. Trustfulness is likely to be influenced by the intentionality, past behavior, social relationship, the risk that is taken by the trustor, and the reward if the trust is realized. The trustee's knowledge on the trustfulness of the trustor may become motivations for the trustee to conform to behavioral rules in place or even form helpful attitude towards the trust relation.

Our definition of trust explicitly recognizes that the mental states of agents are integral part of a trust. Trust communication can be understood as manipulating the mental states of others by communicative actions. A few notes:

- Trust relation TR(A, B, *z*) implies that A depends on B in some sense in order to accomplish *z*. This dependency may take one of many possible senses. Based on the literature [13, 23], the most prevalent senses of trust dependencies are *competence, responsiveness, credibility, security, cooperativeness*, etc. Each trust relation can involve one or more of these basic trust dependencies.
- There can be many trust relations between two agents, each characterized as a trust object. We define **T**(A, B) = {T(A, B, *)} as the set of all the trust objects defined from A to B. Also, we define T(A) = { T(A, *, *) } as the set of all the trust objects that A play the role of trustor.
- Our idea of using BEL(A, TW(B, TR)) as the basis of trust is motivated by the work of Bhargava and his colleagues [6] who define trust as 'the trusting peer's belief in the trusted peer's willingness and capability to behave as expected by the trusting peer in a given context at a given time'. Jones [18] identified two kinds of

belief: (1) *rule belief* and (2) *conformity belief*. The rule-belief refers to the belief that there exist some rules (pertaining to behavioral regularities, norms, obligations of the trustee) leading to the expectation that the trustee will do z. The conformity-belief states that the trustor holds beliefs on the trustee to actually behave in the way not violating those rules. These beliefs on rules and conformity are necessary basis for establishing belief on trustworthiness based on from observable properties and behavior of the trustee.

The above definition can be re-stated using modal logic notations.

Notations	Meaning
T (A, B, z) <==>	"A trust B on achieving z"
a) TR(A, B, z) &	"A has a trust dependency on B to achieve z "
b) MST(A, TR) &	"A's mental state relevant to TR"
c) MST(B, TR) &	"B's mental state relevant to TR"

The recognition of a conceptual core in our framework directly points to the importance of mental beliefs in the theories of trust. This is consistent with prior conception of trust (for recent surveys, see [6, 12, 23, 35]). The variations across these trust definitions can be explained by applying the above core conceptual structure (relation + beliefs) to a specific application context. The generality of this conceptual view lies in the fact that it does not require prior and direct binding of trust with any characteristics of the trustees. Instead, the belief components of a trust serve as the mediator for such binding, and can be done with late-binding. In this way, the concept of trust is allowed to exist in an abstract sense, and will take a concrete meaning only after it is elaborated in a specific situation.

3.2 Trust Communication

Trust communication can be understood as the process of exchanging evidences of trustworthiness and trustfulness. Evidences of trustworthiness are properties of the trustee, but such evidences must be communicated to the trustor and become part of the trustor's belief in order to have an effect on the trust.

Agents begin with a partially developed trust and extend the trust towards a fully developed trust. A *fully developed trust* (FD-Trust) has the following properties:

(i) The purpose of the trust is known. (This refers to the z component of T (A, B, z));
(ii) The trustor A and trustee B have been identified;
(iii) The nature of the trust dependency is determined and mutually believed by both the trustor and the trustee;
(iv) The trustor has established the highest possible belief on the trustworthiness of the trustee based on knowledge about trustee's credentials, observed behavior, reputation, etc;
(v) The trustee has established the highest possible belief on the trustfulness of the trustor based on knowledge about trustor's risk-taking capacity, goals, risks and rewards, etc.

When the above conditions are not fully met, we say that the trust object is a *partially developed trust* (PDT). In real world situations, agents may have to act based on a partially developed trust, due to lack of knowledge, extra cost of trust communication, or making decisions under urgent conditions. When a trust is perceived as inadequate by the participating agents, they will form a shared intention to further develop the trust. Developing a trust involves elaborating on components of a trust objects. Such elaboration process on a trust will continue until one of the following two conditions is met: (a) the trust becomes adequate for the purpose; or (b) the agents do not have any other ways to advance the trust.

4 Meshing of Collaboration and Trust

In the last two sections, we have presented separate theories for trust and collaboration. Both theories were formulated using mental state operators. Now we are ready to connect these into a larger theory of trust-mediated collaboration. We will focus on two semantic relations between trust and collaboration. At one case, trust is the prerequisite for advancing collaboration on a domain activity. On another case, the state of a collaborative activity serves as trust-requiring situation that guides trust communication. We will discuss these two points in more detail next.

4.1 Trust as Prerequisite for Advancing the State of a Collaborative Activity

As agents develop and execute their collaborative plan, they heavily depend on their collective ability to bring their collaboration to certain desirable state. Such effort often requires two kinds of preconditions to be met, as discussed by Lochbaum [21]. One is called "knowledge preconditions", which are denoted as follows:

- *has.recipe*(G, α, R): a group of agents G has a recipe R for action α
- *Id.Param*(G, α(p_1, …, p_n)): G can identify parameters needed for action α.

Meeting knowledge conditions often requires that agents share their knowledge with each other. When knowledge to be shared is considered sensitive and the environment is not fully trusted, a trust negotiation process must be introduced.

Another set of preconditions for collaboration concerns the ability of agents in executing the collaborative plan. A complex plan is often executed by different agents, each executing some subset of actions in the plan. A requirement for having a SharedPlan is that agents must trust each other in their ability of doing individual share of the task. This kind of trust may not be automatic, but need explicit effort to negotiate.

The SharedPlan theory of collaboration (as described in section 2) does not consider the issue of trust in collaborative work. It works only under the assumptions of a fully trusted environment. We extend such theory of collaboration with the following modal operators:

- *has.recipe.Sensitive*(G, α, R)
- *Id.Param,sensitive*(G, α(p_1, …, p_n))
- *Can,Execute.uncertain*(G, α, R)

When one of these operators is invoked, a trust negotiation process is initiated and inserted as a sub process of the overall collaboration.

4.2 Collaborative Activity as Trust-Requiring Situation

An important property of trust communication within a collaborative activity is that the process of trust negotiation itself is a collaborative activity. As described in Section 3, trust is defined by a set of mental states, which can be concretized only in a real situation. In the context of collaborative work, the goal of a trust communication process is to enable further advances to the larger activity. The state of collaboration on the main activity serves as the motivations behind a trust communication session and determines when and how a trust gets started, developed, and ended. The goal-oriented nature of trust has been widely recognized in the philosophical literature [18], but the articulation of how goals relate to trust has been vague and informal.

The mental-state view of trust serves as a schema with which a trust-requiring situation is recognized, interpreted, acted upon by the agents. *A trust-requiring situation* is a situation that a trust is needed in order to advance the ongoing activity. When a situation is known, an agent will actively interpret the situation in order to decide or update the values/contents of the components in a trust object. At the end of Section 2, we have identified two general classes of situations where trust communication needs to be introduced. Here we will refine that discussion, using the following as an example of a situation:

> **Scenario B:** After a major earthquake disaster, many wounded at the events are moved into The Good Health Hospital (GHH), waiting to be treated and cared. However, the hospital runs short of qualified nurses, and has announced a few temporary positions to be filled quickly. Alice has recently completed her training as a nurse from The Care Hospitals in India, and she is motivated to help local residence in fighting this crisis. She applies on-line for the open position at The Good Health Hospital. The hospital needs to verify her capabilities as a nurse before hiring her while the nurse wants to ensure that the Hospital is of a reputable standing. Further to prove her capabilities, the Hospital requests that she shows her Degree Certificate and her Training Certificate, while the nurse requests that the hospital prove it respects HIPAA (Health Insurance Portability and Accountability Act) rules.

Give the above situation, the following can be inferred.

(1) *A situation determines if a trust object is to be created and when*. In a collaborative activity, there are certain moments when collaboration can not proceed without first establishing trust between entities. In other words, a collaborative activity creates *trust-requiring situations* (following the work of Riegelsberger [28]). In scenario B, there is no need to create (or activate) a trust between GHH and Alice until the moment GHH evaluate Alice's application in order to make a hiring decision.

(2) *A situation determines who are involved in a trust, and what roles they play* (either as trustor or trustee). For example, in Scenario B, Alice needs to gain trust from GHH in order to get the job. Hence, GHH serves as the trustor and Alice serves as the trustee in this relation.

(3) *A situation determines the nature of the trust relation to be established*. In the above, GHH has a concern on the Alice capability in performing a nursing job, while Alice wants to make sure that GHH is a reputable place to work. Depending on what the actual concern was raised, the strategy for elaborating the trust can be quite different.

5 A Running Example

In order to demonstrate the applicability of our theory, we will present a brief walk-through on the analysis of scenario B where a nurse Alice seeks to join medical team at The Good Health Hospital in USA. Using the theory presented in this paper, we can analyze this scenario in two levels.

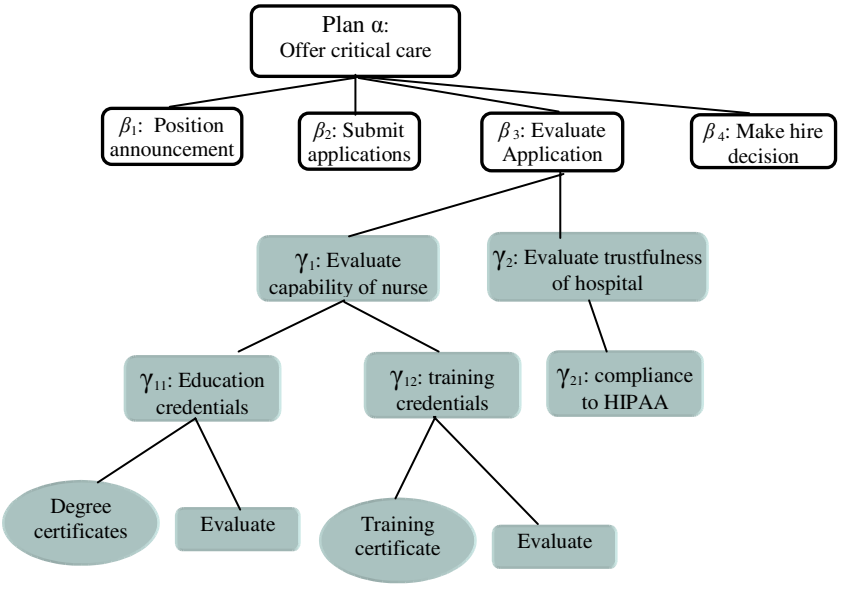

Fig. 3. Collaborative plans of Scenario B

5.1 Intentional Structure of the Domain Activity

The upper portion of Figure 3 shows the intentional structure of the domain activity. The top-level collaborative activity is about finding people who have expertise in providing critical care at an emergency situation. Both Alice and the Hospital share a common goal of providing medical care services to the wounded. To achieve this goal, a process (recipe) adopted by this scenario is that the hospital issued position announcement, and interested candidates are expected to submit applications through an on-line utility, followed by evaluating applications at the Hospital. The hiring decision can not be made until β_1, β_2, β_3 are done. The plan structure of domain level activity is shown in the top part of Figure 3 (none shaded plan nodes).

5.2 Intentional Structure of Trust Communications

When agents elaborate the plan to the point γ_1, the system recognizes that there is a need to create a trust object T_1(GHH, Alice, "qualified nurse"). Following the principles stated in Section 4.2, the system will use the knowledge about the current state of collaboration (specifically, the plan graph that T_1 is rooted) to determine the nature of the trust dependency and the belief components of T_1.

- Determining the trust relation involves identifying trustor and trustee and how the trustor depends on the trustee. Based on the knowledge about plans α and β_3, we can infer that the trustor is GHH and the trustee is Alice. GHH depends on Alice's *capability* of performing a nurse function.
- Determining the mental state components of trust is also relying on the knowledge about α and β_3. Due to emergency situation of this recruiting, the intention to have established trust T_1 is clearly set. The belief component BEL(GHH, TW(Alice, TR)) (i.e., GHH's belief on the trustworthiness of Alice in relation to TR) is zero. TW(Alice, TR) is reflected by two properties of Alice: her degree certificates, and her training certificates. This leads the agents to adopt a recipe for γ_1 that includes two sub actions: γ_{11} and γ_{12}. To contribute these subgoals, Alice shared her degree certificate and training certificate with GHH, and GHH went through an evaluation process.
- Establishing BEL(Alice, TF(GHH, TR)). As part of this trust, Alice would like to insure that GHH is a reputable institution. Based on Alice's request for evidences, GHH shared the certificate of its membership with HIPAA. This will also allow Alice to trust GHH that that information she provided to GHH will not be abused.

6 Discussion and Conclusion

We have presented a unified theory of trust and collaboration using a mental state perspective. As demonstrated by the analysis of scenario B, we have observed that our theory is capable of explaining the way that trust and collaboration are coupled in real activities. Such coupling effect creates opportunities for dealing with difficult semantic issues in trust management. Our work stands at crossroads of two research areas: trust negotiation and theories of collaboration and communication.

The work presented in this paper is conceptual in nature, and is our first step towards effective trust management in collaborative applications. While much research efforts has been placed into the foundations of trust negotiation –such as languages for expressing resource access control policies [3, 4], protocols and strategies for conducting trust negotiations [30, 39], and logic for reasoning about the outcomes of these negotiations [29, 37], little effort has been posed to understanding how the articulated notion of trust fits within negotiations, beyond the access control aspect. Existing trust negotiation theories have formalized the interaction protocols that parties should follow in order to ensure correct negotiation executions. In particular, Yu *et al.* [39] investigated the notion of negotiation strategies which control the exact content of the messages: which credentials to disclose, when to disclose them, and when to terminate a negotiation. Strategies and their interoperability have been further investigated by the same authors

[38, 39] who have proposed a unified scheme, called *Unipro*, to model resource protection, including policies. Additionally, in order to address the need for managing the dynamics of trust, Ma and Orgun [22] proposed a formal theory of trust evolution for multiagent systems. Their theory uses Typed Modal Logic (TML) expressions to represent beliefs and operators on beliefs.

Although relevant, this body of work fails to elaborate on the notion of trust, and to consider how parties' collaboration may substantially alter the negotiation flow. Rather, most of the current trust negotiation approaches have been designed under the assumption that the negotiating parties are inherently not collaborative, although they adhere to the trust negotiation protocol. This motivated a large amount of work focusing on privacy and on cryptographic based negotiations [31, 36]. We believe that this assumption is very restrictive, and that it hindered the deployment of trust negotiation protocols in many real-world domains where collaboration is essential. The only work on negotiation considers the effect of cooperative work on trust negotiation are work by Baselice et al. [2], Jin et al. [17], and Svirskas [33]. However, these works are mostly about interoperability of trust management across domains, and they do not deal with meaning of trust in collaborative applications. Also, a simple collaborative approach for trust negotiations is proposed in [30], where a cooperative strategy is proposed.

Our work fills a gap in the literature by making the connections between theories of collaboration and concepts of trust. Existing theories about collaboration and trust have been isolated. On one hand, theories of collaboration works only with perfectly trusted environment. On the other hand, methods of trust communication assume collaboration-neutral environment, and do not deal with the semantics of trust. Our current work provides a unified theory of trust and collaboration based on the SharedPlan theory of collaboration.

We plan to further validate our theory by investigating more practical scenarios and by conducting extensive case analysis. At the same time, we have been using this theory to guide the design of a new experimental system CollTrust-X. CollTrust-X builds on top of the TRUST-X architecture [5, 30] and adds a semantic layer for managing trust objects and collaborative plans.

Acknowledgement

This work is partially supported by a grant from the National Science Foundation under Grants No EIA-0306845 and by the National Visualization and Analytics Center (NVAC), a U.S. Department of Homeland Security Program, under the auspices of the Northeast Regional Visualization and Analytics Center (NEVAC).

References

1. Balkanski, C., Hurault-Plantet, M.: Cooperative requests and replies in a collaborative dialogue model. International Journal of Human-Computer Studies 53, 915–968 (2000)
2. Baselice, S., Bonatti, P.A., Faella, M.: On interoperable trust negotiation strategies. In: Eighth IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2007), Bologna, Italy, June 13-15, pp. 39–50 (2007)

3. Bertino, E., Ferrari, E., Squicciarini, A.: Privacy preserving trust negotiations. In: 4th International Workshop on Privacy Enhancing Technologies, Toronto, Canada (2004)
4. Bertino, E., Ferrari, E., Squicciarini, A.: Trust negotiations: concepts, systems, and languages. Computing in Science & Engineering 6(4), 27–34 (2004)
5. Bertino, E., Ferrari, E., and Squicciarini, A.: Trust-X桚 Peer to Peer Framework for Trust Establishment. IEEE Trans. Knowledge and Data Eng. 16(7), 827–842 (2004)
6. Bhargava, B., Lilien, L., Rosenthal, A., Winslett, M., Sloman, M., Dillon, T.S., Chang, E., Hussain, F.K., Nejdl, W., Olmedilla, D., Kashyap, V.: The pudding of trust. IEEE Intelligent Systems 19(5), 74–88 (2004)
7. Blaze, M., Feigenbaum, J., Ioannidis, J., Keromytis, A.D.: The KeyNote Trust-Management System Version 2 (1999)
8. Blaze, M., Feigenbaum, J., Lacy, J.: Decentralized trust management. In: Proceedings, 1996 IEEE Symposium on Security and Privacy, pp. 164–173 (1996)
9. Bratman, M.E.: Shared cooperative activity. Philosophical Review 101, 327–341 (1992)
10. Butler, J., Mitchell, L.C., Friedman, C.R., Scripp, R.M., Watz, C.G.: Collaboration between Public Health and Law Enforcement: New Paradigms and Partnerships for Bioterrorism Planning and Response. Emerging Infectious Diseases 8, 1152–1156 (2002)
11. Cai, G., Wang, H., MacEachren, A.M., Fuhrmann, S.: Natural Conversational Interfaces to Geospatial Databases. Transactions in GIS 9(2), 199–221 (2005)
12. Corritore, C.L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. International Journal of Human-Computer Studies 58(6), 737–758 (2003)
13. Grabner-Krauter, S., Kaluscha, E.A.: Empirical research in on-line trust: a review and critical assessment. International Journal of Human-Computer Studies 58(6), 783–812 (2003)
14. Grimes, J.G.: Department of Defense Information Sharing Strategy, Department of Defense (2007)
15. Grosz, B.J., Kraus, S.: Collaborative plans for complex group action. Artificial Intelligence 86, 269–357 (1996)
16. Hollan, J., Hutchins, E., Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research. ACM Transactions on Computer-Human Interaction 7(2), 174–196 (2000)
17. Jin, J., Ahn, G.-J., Shehab, M., Hu, H.: Towards trust-aware access management for ad-hoc collaborations. In: CollaborateCom 2007. International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2007. CollaborateCom 2007, pp. 41–48 (2007)
18. Jones, A.J.I.: On the concept of trust. Decision Support Systems 33(3), 225–232 (2002)
19. Jones, K.: Trust: Philosophical Aspects. In: Smelser, N.J., Baltes, P.B. (eds.) International Encyclopedia of the Social & Behavioral Sciences, pp. 15917–15922. Pergamon, Oxford (2001)
20. Lesh, N., Rich, C., Sidner, C.L.: Using Plan Recognition in Human-Computer Collaboration. In: Proceedings of the seventh international conference on user modelling, Banff, Canada, pp. 23–32 (1999)
21. Lochbaum, K.E.: A collaborative planning model of intentional structure. Computational Linguistics 24(4), 525–572 (1998)
22. Ma, J., Orgun, M.A.: Trust management and trust theory revision. IEEE Transactions on Systems, Man and Cybernetics, Part A 36(3), 451–460 (2006)
23. McKnight, D.H., Chervany, N.L.: The meanings of trust. In: Falcone, R., Singh, M., Tan, Y.-H. (eds.) AA-WS 2000. LNCS (LNAI), vol. 2246, pp. 27–54. Springer, Heidelberg (2001)

24. Nardi, B.A. (ed.): Context and Consciousness: Activity Theory and Human-computer Interaction. MIT Press, Cambridge (1996)
25. Relyea, H.C., Seifert, J.W.: Information Sharing for Homeland Security: A Brief Overview. Congressional Research Service Reports on Homeland Security (2005)
26. Rich, C., Sidner, C.L., Lesh, N.: Collagen: Applying collaborative discourse theory to human-computer collaboration. AI Magazine 22(4), 15–25 (2001)
27. Richards, E.P.: Collaboration between Public Health and Law Enforcement: The Constitutional Challenge. Emerging Infectious Diseases 8(10) (2002)
28. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: The mechanics of trust: A framework for research and design. International Journal of Human-Computer Studies 62(3), 381–422 (2005)
29. Seamons, K.E., Winslett, M., Yu, T.: Requirements for Policy Languages for Trust Negotiation. In: Third IEEE International Workshop on Policies for Distributed Systems and Networks, Monterey, CA (2002)
30. Squicciarini, A.C., Bertino, E., Ferrari, E., Paci, F., Thuraisingham, B.M.: PP-Trust-X: A system for Privacy Preserving Trust Negotiations. ACM Transactions on Information Systems Security 10(3), 1–50 (2007)
31. Squicciarini, A.C., Bertino, E., Ferrari, E., Ray, I.: Achieving privacy in trust negotiations with an ontology-based approach. IEEE Transactions on Dependable and Secure Computing 3(1), 13–30 (2006)
32. Suchman, L.: Plan担 and Situated Actions: The Problem of Human崩achine Communication. Cambridge University press, Cambridge (1987)
33. Svirskas, A., Isachenkova, J., Molva, R.: Towards secure and trusted collaboration environment for European public sector. In: CollaborateCom 2007. International Conference on Collaborative Computing: Networking, Applications and Worksharing, pp. 49–56 (2007)
34. Tullberg, J.: Trust–The importance of trustfulness versus trustworthiness. Journal of Socio-Economics (2008) (in press)
35. Wang, Y.D., Emurian, H.H.: An overview of online trust: Concepts, elements, and implications. Computers in Human Behavior 21(1), 105–125 (2005)
36. Winsborough, W.H., Li, N.: Protecting sensitive attributes in automated trust negotiation. In: Proceedings of the ACM Workshop on Privacy in the Electronic Society, pp. 41–51. ACM Press, New York (2002)
37. Winsborough, W.H., Seamons, K.E., Jones, V.E.: Automated trust negotiation. In: DARPA Information Survivability Conference and Exposition, Piscataway, New Jersey, pp. 88–102. IEEE Press, Los Alamitos (2000)
38. Yu, T., Winslett, M.: A Unified Scheme for Resource Protection in Automated Trust Negotiation. In: The 2003 IEEE Symposium on Security and Privacy, May 11-14, p. 110. IEEE Computer Society, Los Alamitos (2003)
39. Yu, T., Winslett, M., Seamons, K.E.: Supporting structured credentials and sensitive policies through interoperable strategies for automated trust negotiation. ACM Transactions on Information and System Security (TISSEC) 6(1), 1–2 (2003)