TIUPAM: A Framework for Trustworthiness-Centric Information Sharing

Shouhuai Xu[1], Ravi Sandhu[2], and Elisa Bertino[3]

[1] Department of Computer Science, Univ. of Texas at San Antonio
shxu@cs.utsa.edu
[2] Institute for Cyber Security, Univ. of Texas at San Antonio
ravi.sandu@utsa.edu
[3] Department of Computer Science, Purdue University
bertino@cs.purdue.edu

Abstract. Information is essential to decision making. Nowadays, decision makers are often overwhelmed with large volumes of information, some of which may be inaccurate, incorrect, inappropriate, misleading, or maliciously introduced. With the advocated shift of information sharing paradigm from "need to know" to "need to share" this problem will be further compounded. This poses the challenge of achieving assured information sharing so that decision makers can always get and utilize the up-to-date information for making the right decisions, despite the existence of malicious attacks and without breaching privacy of honest participants. As a first step towards answering this challenge this paper proposes a systematic framework we call TIUPAM, which stands for "Trustworthiness-centric Identity, Usage, Provenance, and Attack Management." The framework is centered at the need of trustworthiness and risk management for decision makers, and supported by four key components: identity management, usage management, provenance management and attack management. We explore the characterization of both the core functions and the supporting components in the TIUPAM framework, which may guide the design and realization of concrete schemes in the future.

1 Introduction

Information sharing is an important process in human society because it helps make better decisions. However, information should not be arbitrarily disseminated for various reasons including sensitivity and privacy, and truthfulness of information should not be taken for granted. The latter is especially important in adversarial environments, such as business, economics, and military. Traditionally, the research communities and the industrial vendors have focused on enforcing "need to know" via various mechanisms. Recently, a new paradigm known as "need to share" has emerged, primarily to more effectively deal with threats such as terrorist attacks and demonstrated failure of "need to know" in this regard. This brings new challenges because (1) decision makers are potentially even more overwhelmed with information, which should by no means be treated as trustworthy, and (2) the "access control"-centric solution paradigm is not sufficient anymore because the notions of *authorization* and *authentication* are less

E. Ferrari et al. (Eds.): TM 2009, IFIP AICT 300, pp. 164–175, 2009.

explicit. In this paper we propose a solution framework to help the decision makers deal with this new challenge.

Our contributions. We propose a systematic framework called TIUPAM, which stands for "Trustworthiness-centric Identity, Usage, Provenance, and Attack Management". The framework is centered at serving decisionmakers' needs for effectively managing the trustworthiness of information as well as the risk that may be caused by utilizing or not utilizing available information. This core of trustworthiness and risk management is supported by four components.

- *Identity management*. Identity management serves trustworthiness and risk management, provenance management, as well as usage management while receiving services from provenance management and usage management. In particular, it allows the participants to evaluate the trustworthiness of the digital identities and digital credentials for people, organizations, and devices.
- *Usage management*. Usage management serves trustworthiness and risk management while receiving services from identity management. It mainly deals with authorized activities. It extends current generation of usage control by considering, for example, the trustworthiness of both requests and information.
- *Provenance management*. Provenance management serves trustworthiness and risk management by essentially enabling the evaluation of trustworthiness of data, software, and requests, while receiving service from identity management.
- *Attack management*. Attack management serves all of the aforementioned components by dealing with attacks and unauthorized activities. In particular, the evaluation of trustworthiness of information, identity, request, usage, and provenance must be with respect to some specific attack model.

The focus of this paper is on exploration of the characteristics of the core functions and components, rather than specifying any concrete realizations. This characterization would help the design of concrete schemes for realizing the framework in the future.

Paper organization. Section 2 presents the TIUPAM framework as well as its core functions. Section 3 discusses the identity management component, Section 4 presents the usage management component, Section 5 discusses the provenance management component, and Section 6 presents the attack management component. Section 7 summarizes the paper.

2 The TIUPAM Framework

Within this framework and throughout the present paper, we use the term "information" in a broad sense, meaning that it accommodates information items, data items, message items, and knowledge items.

2.1 Framework Components and Their Logical Relationships

As illustrated in Figure 1, the TIUPAM framework is centered at trustworthiness and risk management, which serves the need of decision makers. The supporting components are identity management, usage management, provenance management, and attack management, whose logical relationships are depicted in Figure 2.

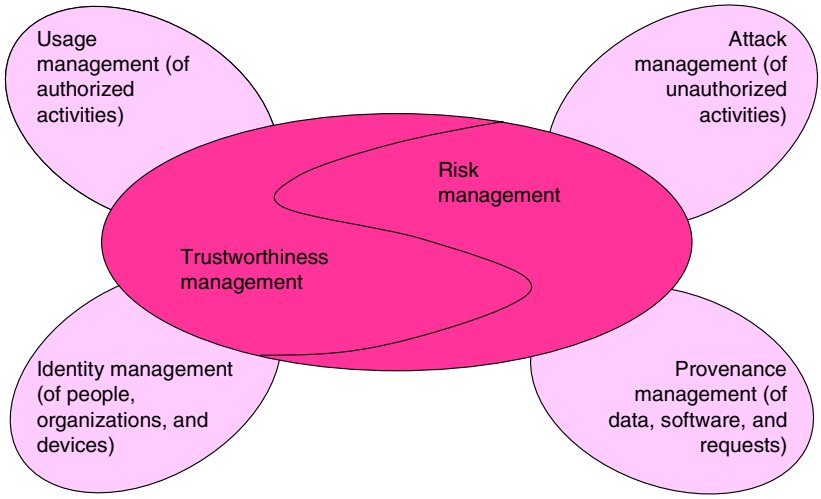

Fig. 1. Key components of the TIUPAM framework

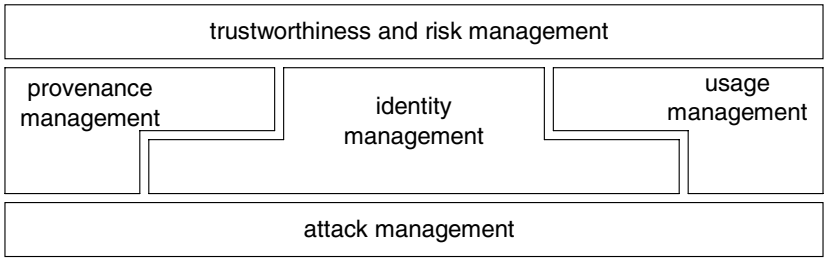

Fig. 2. Logical relationship between the components

In what follows we elaborate on the functionalities of the core as well as the supporting components.

- *Trustworthiness and risk management*: For decision makers, the most important issue is the trustworthiness of the information at hand, which reflects the decision maker's current "snapshot" of the world and may be (in)accurate, (in)correct, misleading, or even maliciously introduced. The term "snapshot" is emphasized because, in the context of the present paper, trustworthiness is meant to capture the *dynamical* evaluation of the degree of information being trustable or trustworthy. The term "dynamical" indicates that one's evaluation of trustworthiness of some information may change with respect to time, as more information is gathered. (In contrast, trust can be invariant regardless of the information currently available; for example, we may still trust an individual even if there is information or rumors against that person.) Corresponding to the non-perfect trustworthy information, any decision based on the "snapshot" bears some risk because its execution

may lead to negative consequences. We stress that trustworthiness is a measure against the snapshot of one's up-to-date observation about the information in question; whereas, risk is a measure against the potential consequences caused by the execution of decisions based on not-necessarily-trustworthy information, based on the state-of-the-art understanding of the world. Therefore, trustworthiness and risk are not necessarily complementary to each other.

- *Identity management*: Identity, including digital credentials, provides a base for trustworthiness, risk, provenance, and usage management. Specifically, in order to support trustworthiness and risk management, we need to measure the trustworthiness of identities of people, organizations, and devices. This is because the aforementioned snapshots are derived, in one way or another, from the statements asserted by the relevant people, organizations, and devices. For example, a software program digitally signed by a software vendor may certify that the output corresponding to a given input to the program is indeed the desired result (e.g., some knowledge extracted from data with respect to the algorithm the program executes); a message digitally signed by an organization would make one tend to believe its trustworthiness; a successful attestation of a remote device may lead us to accept that the remote peering computer is not compromised.
- *Usage management*: Usage management seeks to manage authorized activities by extending traditional access control. It was inspired by the following observations on the limitation of traditional access control: (1) a subject is always trustworthy as long as it passes certain pre-determined authentication, and (2) an object is always trustworthy as long as it is in the filesystem or database. The former preassumption is faulty if the authentication credential of the subject has been compromised, and the later preassumption is faulty if the object itself was malicious or incorrectly provided. Therefore, it is important for usage management to take into account, among other things, the trustworthiness of both data and requests, which in turn requires to take into account the trustworthiness of both provenance and identity.
- *Provenance management*: Provenance management directly serves the higher-layer trustworthiness and risk management by managing the provenance of information, software, and requests etc, while being served by identity management and usage management. Provenance of data allows us to measure the trustworthiness of information; provenance of software helps to evaluate the trustworthiness of software programs; provenance of requests enhances the assurance of the requests' source in that they are invoked by the individual or process in question, rather than by malware.
- *Attack management*: Attack management deals with unauthorized activities, especially malicious attacks that may intentionally introduce wrong or misleading information into the system. In particular, it helps manage the trustworthiness of infrastructure-level services provided to the other components as well as their services in the framework (e.g., authentication services). This is an important problem and more subtle than first glance because traditionally people tend to accept that infrastructures (e.g., public key infrastructures or PKI) as trustworthy simply because of their absolute trust in them.

2.2 Core Functions of Trustworthiness and Risk Management

Figure 2 highlighted the logical relationships between the components. In what follows we discuss the core functions that should be realized by the TIUPAM framework to the applications. The relationships between the functions are highlighted in Figure 3 and elaborated below.

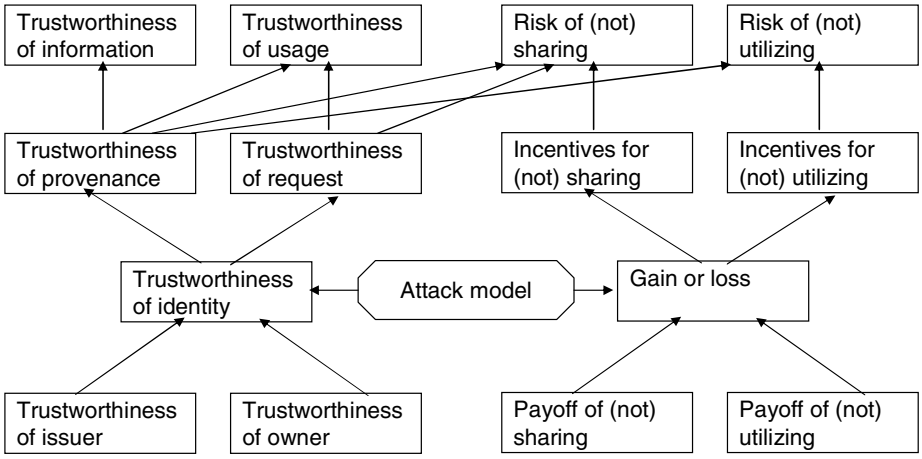

Fig. 3. Structures of the trustworthiness and risk functions (attack models are not elaborated for a better visual effect)

Decisionmakers would often need to resolve the following questions: How trustworthy is a given information? How trustworthy is the adherence of an information consumer to the usage policy? What is the risk incurred by sharing or not sharing a certain information? What is the risk because of utilizing or not utilizing a given information? Corresponding to these questions, we may define the following function families.

- **Trustworthiness of information:** It should be a function of the trustworthiness of the provenance of the information in question. Therefore, there are families of functions $\{f_1\}$, $\{f_{11}\}$, and $\{f_{111}\}$ such that

  $trustworthiness\_of\_data = f_1(trustworthiness\_of\_provenance)$, where
  $trustworthiness\_of\_provenance = f_{11}(trustworthiness\_of\_identity)$, and
  $trustworthiness\_of\_identity = f_{111}(trustworthiness\_of\_issuer,$
  $trustworthiness\_of\_owner, attack\_model)$.

  Note that the *attack_model* is always an input argument to some "low level" functions, meaning that it is implicit in the "high level" functions such as $trustworthiness\_of\_data$. The motivation is that in order to evaluate the functions in a consistent fashion, the same *attack_model* should be used in the bottom-up evaluation of the functions.

- **Trustworthiness of usage:** It should be a function of both trustworthiness of provenance and trustworthiness of request. Therefore, there are families of functions $\{f_2\}$ and $\{f_{21}\}$ such that

  $trustworthiness\_of\_usage = f_2(trustworthiness\_of\_provenance, trustworthiness\_of\_request)$, where
  $trustworthiness\_of\_request = f_{21}(trustworthiness\_of\_identity)$, and
  $trustworthiness\_of\_identity = f_{111}(trustworthiness\_of\_issuer, trustworthiness\_of\_owner, attack\_model)$.

- Risk of sharing information: It should be a function of the trustworthiness of the provenance of the information in question, the trustworthiness of request (which may be explicit in pull-based information sharing and implicit in push-based information sharing), and incentives for sharing. Therefore, there are a family of functions $\{f_3\}$, $\{f_{31}\}$, $\{f_{311}\}$ such that

  $risk\_of\_sharing\_information = f_3(trustworthiness\_of\_provenance, trustworthiness\_of\_request, incentive\_for\_sharing\_information)$, where
  $incentive\_for\_sharing\_information = f_{31}(gain\_because\_of\_sharing)$, and
  $gain\_because\_of\_sharing = f_{311}(payoff\_of\_sharing, attack\_model)$.

  Note that *gain_because_of_sharing* must take *attack_model* into consideration because in different attack models the outcome can be completely opposite (e.g., the information receiver is truly the claimed authorized user vs. the information receiver is actually the attacker who can perfectly impersonate the user because the attacker has compromised the user's identity). Similarly, we can define the function families for specifying the risk of not sharing.
- **Risk of utilizing received information:** It should be a function of the trustworthiness of information provenance and the incentives for utilizing the information in question. Therefore, there are a family of functions $\{f_4\}$, $\{f_{41}\}$, $\{f_{411}\}$ such that

  $risk\_of\_utlizing\_information = f_4(trustworthiness\_of\_provenance, incentive\_for\_utilizing\_information)$, where
  $incentive\_for\_utilizing\_information = f_{41}(gain\_because\_of\_utilizing)$,
  $gain\_because\_of\_utilizing = f_{411}(payoff\_of\_utilizing, attack\_model)$.

  Similarly, we can define the function families for specifying the risk of not utilizing a given information (e.g., because of its low or uncertain degree of trustworthiness).

It should be noted that risk comes from two aspects: (1) the consequences that may be caused by utilizing incorrect or malicious information (e.g., because it is accompanied with a high degree of trustworthiness); (2) the consequences that may be caused by not utilizing the not-known-to-be, but indeed trustworthy, information (e.g., because it is accompanied with a low or uncertain degree of trustworthiness). The risk will be evaluated whenever a relevant decision is being made; for example, whether to allow the use or exchange of information.

We reiterate that the components aim to evaluate and maintain the trustworthiness of information in a dynamic fashion because the trustworthiness should always be updated. For example, we may treat a data item as fully trustworthy today even though it was only partially trustworthy yesterday because of new insights obtained since then.

We emphasize that it is not our aim in this paper to define the function families mentioned above, which should be specific to the applications. Rather, we want to clearly state the framework by detailing the relationship between the components.

3 Identity Management

In this section we discuss the key properties of desired identity management systems.

Extensibility. Any good identity management should be easily extended to accommodate or integrate emerging new identity systems. This is important because as the computing environments evolve, so do the individual identity management systems. Moreover, the diversity of applications often implies diversity in digital identity or credential systems. This is important because while we are used to digital identities such as public keys or attribute certificates, facilitated by a public key infrastructure (PKI), other types of identities may emerge. For example, in the case of mobile computing, two users with no common trusted third party could establish a mutual trust on their own. Moreover, this individual trust may further bootstrap future trust establishment between their friends because social networks are becoming an indispensable part of future computing paradigms. In turn, this means that future identity management systems should be easily extensible.

Automated trustworthiness. A key support of identity management systems to trustworthiness and risk management, usage management, and provenance management is the trustworthiness a verifier can put on a digital identity in question. This requires the identity management component to provide automated trustworthiness service by ensuring the following.

- *Compromise containment*. This states that the consequences due to the compromise of some computers or identities are contained and, ideally, minimized. There are several typical scenarios.
  - *Compromise of servers that authenticate users through their identities or credentials is contained*. This is relevant when the authentication is based on symmetric cryptography, including symmetric key cryptosystems and passwords. In this case, compromising a server could cause the compromise of the users' authenticators directly (e.g., when symmetric keys are used) or indirectly (e.g., after launching off-line dictionary attack when passwords are used). This is also relevant when the authentication is based on asymmetric cryptography, such as when servers store the public keys of users. For example, the attacker could tamper with the access history of the users, erase the access events incurred by the attacker, insert bogus user entries, or modify the public keys of the users. In all of these cases, the damage should be contained and, ideally, minimized.
  - *Compromise of some users' identities or credentials is contained to those users and, ideally, to those compromised (e.g., stolen) identities and credentials*. It it not unusual that every user has multiple digital identities and credentials, which may or may not be independent of each other at all (e.g., one user reuses a password for multiple accounts). There is a possibility that compromising a user's

computer could cause the compromise of all the user's identities or credentials, which corresponds to the worst-case scenario. Therefore, it is important to ensure containment in the following sense: Compromising of digital identity or credential for accessing one server does not cause the compromise of digital identity or credential for accessing another server.

- *Accountability*. Digital identities or credentials live and operate in a hostile environment wherein many computers, including well-protected servers, can be compromised. This puts in question accountability, enforcement of which deters many attacks. For example, if an access is launched through the use of stolen identity or credential, who should be held accountable for the consequences? It is arguable that the user, whose identity or credential was stolen, is a victim as well. Things could become much more complicated when, for example, a malicious user intentionally abuses this fact to hide its own unlawful activities. This calls for good forensics mechanisms to deter, if not absolutely hold the malicious users or attackers accountable for attacks and abuses.

4 Usage Management

The concept of usage control has recently emerged as a paradigm for next generation access control transcending the traditional access matrix model [2]. To accommodate modern applications, concepts such as trust management, digital rights management, obligations and attribute-based access control were proposed in the past decade. Usage control provides a unified framework for modeling these and other access-control extensions. Usage control maintains the classic access control abstraction of a right as a privilege that a subject must hold to access an object in different modes. Unlike the traditional access matrix, in usage control the existence of a right is determined when an access is attempted by a subject and may continue to be determined as the right is used. This usage decision is made based on subject attributes, object attributes, authorizations, obligations, and conditions. Specifically, authorizations are predicates that determine whether the subject (requester) holds the requested rights on the object, obligations are predicates that verify the subject has performed required actions prior or during the usage, and conditions are predicates on environmental or system state. Usage control explicitly recognizes a pre, ongoing and post phase for each usage of a resource. Another feature of usage control not present in conventional access control models is mutable attributes attributes of subjects and attributes that are modified before, during, or after a usage session. Collectively these features allow for consumable rights and instant and preemptive revocation.

Considering the requirements of trustworthiness-centric information sharing discussed above, we identify two limitations of current usage control models, viz., future (or post) obligations and system obligations. For example, a physician accessing a patients electronic health record in an emergency may have a pre-obligation to acknowledge that this is an emergency situation so that access is opened up. After the usage is completed, she may incur a post-obligation to file a statement confirming that the emergency access was justified. Even though the post-obligation occurs after access, it validates the circumstance of the completed access. Completion of the post-obligation

may be deferred into the future, allowing it to be done when the physician has some downtime. This then truly becomes a future obligation. The concept of a system obligation comes into play when an object is moved from one security domain to another. For example a document D from domain A could be made available in domain B but with policy requirements specified by domain A, such as make D accessible to no more than five users in domain B, store D in encrypted form and delete D after one month. These policy requirements are stated by domain A but enforced by domain B. We say domain B has an obligation to enforce the policy specified by domain A.

The dynamic charactersitics of usage control are well suited to the problem of trustworthiness-centric information sharing. Usage control decisions are made at access time and continue to be revisited during access. As such the basic elements for dynamic decisions with respect to access are fundamental to usage control. However, trustworthiness and risk should be more explicitly incorporated into future versions and manifestations of usage control. While the current models for usage control provide the necessary foundational framework much research needs to be done to incorporate attributes and rules that effectively capture trustworthiness and risk as attributes.

5 Provenance Management

5.1 Functional Requirements

Without loss of generality, we assume that information may move within distributed/decentralized systems in the format of messages. Moreover, new messages may be produced by algorithms that may take other messages as inputs. That is, we are primarily dealing with information provenance management in distributed or decentralized systems, which might often be large-scale.

From a functional perspective, we believe that a secure provenance management system should cover the entire lifecycle of information as well as their associated provenance. In this context, we classify information lifecycle into the following procedures of generation and processing. We note that this lifecycle is somewhat tailored to secure provenance management systems, and thus may not be appropriate for other systems.

- *Generation*: An information item originally enters into a provenance management system through some participant; such participant is the party responsible for the initial generation and insertion of the information item into the system.
- *Processing*: Each participant, source or intermediate node, can produce new information items based on the items it received from other participants. Various (e.g., datamining or knowledge extraction) algorithms and functions are possible for producing information items. For example, such a function can simply consist of endorsing an information item another participant is disseminating.

5.2 Security Requirements

A secure provenance management system should provide information trustworthiness management service to higher layer applications.In general, information trustworthiness depends on the trustworthiness of the source, the trustworthiness of the intermediate nodes as well as their processing algorithms. However, things quickly become

complex when some participants (i.e., sources and intermediate nodes) may be malicious. In what follows we discuss some representative issues that relevant to how information trustworthiness should be managed.

- For a source, it is necessary to know about the trustworthiness of an information item that has to be entered into the system. It is also necessary that, when a source realizes that it has entered into the system inaccurate or even misleading information (e.g., deceptive information deliberately provided by adversary), the source be able to inform all the relevant participants about this fact (and possibly also to provide updated information).
- For an intermediate node, it is necessary to know about the trustworthiness of both the source and the prior intermediate nodes so that, for example, a decision may be made whether to re-disseminate the processed information. It is also important to allow a node to notify upstream nodes, e.g., that some information items they provided are inaccurate or even misleading (we may call this "backward information correction"), and to notify downstream nodes, e.g., that some information items they received are inaccurate or even misleading (we may call this "forward information correction").
- For an information consumer, it is necessary to be able to evaluate the trustworthiness of an incoming information item. Moreover, the consumer must be cautious in making decisions that rely on such items because the decisions may not be reversible and, once enforced, may cause severe consequences.
- For an administrator, it is important to know who has a large influence or impact on the evolution of information in the networks? Enhancing security of such participants would significantly improve security from a whole-system perspective.

6 Attack Management

In order to enable trustworthiness and risk management, identity management, usage management, and provenance management, attack management seeks to systematically model the attacks against each of the relevant processes and procedures. Corresponding to Figure 3, we articulate the following attack models attempting to manipulate most of the functions. Note that the attacks accommodate those targeting application layer and those targeting infrastructure layer as well.

Attacks attempting to manipulate the trustworthiness of information. Trustworthiness of information can be manipulated by compromising the provenance of the information. There are several "attack points" at which the attacker can tamper with the provenance information. The attack points correspond to the generation of the information (e.g., a malicious user enters false information into the system), processing of the information (e.g., a malicious user claims that the information is the output of some legitimate application program), the dissemination of the information (e.g., a malicious user claims that the originator of the false information is a trustworthy source). One way to successfully launch the above attacks is to manipulate the trustworthiness of identities, which can be done by compromising and abusing the compromised

identity to impersonate the identity owner, or by compromising a victim identity issuer or becoming a malicious identity issuer.

Attacks attempting to manipulate the trustworthiness of usage. Trustworthiness of usage can be attacked by undermining the trustworthiness of information provenance so that, e.g., malicious information thereby spreads to a large population of users, or by manipulating the trustworthiness of request. The latter can be done by compromising the trustworthiness of identity (e.g., compromising the credential in question). It is also possible to manipulate trustworthiness of usage by attacking the management of *who could read/write/modify* as well as *who have read/written/modified* which information.

Attacks attempting to manipulate the risk management. Risk management can be undermined by manipulating the trustworthiness of the information in question (e.g., highly trustworthy information is deemed as low trustworthy, low trustworthy information is deemed as high trustworthy), or by manipulating the trustworthiness of the request (e.g., unauthorized users may becomes highly trustworthy in requesting the information).

Attacks attempting to manipulate the privacy of honest participants. Privacy is important in many applications and is relevant in all the components. Privacy protection may be at odds with trustworthiness because, for example, a malicious user may intentionally introduce misleading information into the system by abusing the anonymity protection shield so as to not be held accountable. Privacy protection can be dealt with by appropriate risk management in deciding whether or not to share some information, or whether or not to utilize some received information. Privacy protection is crucial to usage management because a malicious user may leak certain information without being held accountable.

7 Conclusion and Future Work

We have explored a systematic framework for trustworthiness-centric information sharing we called TIUPAM. Our framework consists of a core component — the trustworthiness and risk management, and four supporting components — identity management, usage management, provenance management, and attack management.

We highlighted the properties a desired solution should possess, and it is beyond the scope of the present paper for designing concrete solutions. As such, this paper introduces a range of challenging research problems for future work. For example, the identity and attack management explored in the paper is even more demanding than the state of the art in managing the trustworthiness of certain cryptographic credentials [3]; the provenance management explored in the paper is even more challenging than the provenance security discussed in [1].

Acknowledgement. This work is supported in part by AFOSR MURI award FA9550-08-1-0265.

References

1. Braun, U., Shinnar, A., Seltzer, M.: Securing provenance. In: HotSec 2008 (2008)
2. Park, J., Sandhu, R.: The UCON ABC usage control model. ACM Transactions on Information and System Security (TISSEC) 7(1), 128–174 (2004)
3. Xu, S., Yung, M.: Expecting the unexpected: Towards robust credential infrastructure. In: FC (2009)