# Knowledge-empowered Agent Information System for Privacy Payoff in eCommerce

Abdulsalam Yassine, Ali Asghar Nazari Shirehjini, Shervin Shirmohammadi, Thomas T. Tran

*Distributed and Collaborative Virtual Environments Research Laboratory*

*School of Information Technology and Engineering SITE, University of Ottawa*

ayassine@discover.uottawa.ca, anazari@discover.uottawa.ca, shervin@discover.uottawa.ca, ttran@site.uottawa.ca

**Abstract.** Today, many online companies are gathering information and assembling sophisticated databases that know a great deal of information about many people, generally without the knowledge of those people. Such endeavor has resulted in the unprecedented attrition of individual's right to informational self-determination. Consumers, on one hand, are powerless to prevent the unauthorized dissemination of their personal information, and on the other, they are excluded from its profitable commercial exchange. This paper focuses on developing knowledge-empowered agent information system for privacy payoff as a means of rewarding consumers for sharing their personal information with online businesses. The design of this system is driven by the following argument: if consumers' personal information is a valuable asset, should they not be entitled to benefit from their asset as well? The proposed information system is a multi-agent system where several agents employ various knowledge and requirements for personal information valuation and interaction capabilities that most users cannot do on their own. The agents in the information system bear the responsibility of working on behalf of consumers to categorize their personal data objects, report to consumers on online businesses' trust and reputation, determine the value of their compensation using risk-based financial models, and finally negotiate for a payoff value in return for the dissemination of users' information. The details of the system as well as a proof-of-concept implementation using JADE (Java Agent Development Environment) are presented here.

***Keywords****: Knowledge, Information System, Agents, Privacy, E-commerce*

## 1 Introduction

Today, many online companies are gathering information and assembling sophisticated databases that know a great deal of information about many people, generally without the knowledge of those people (Cavoukian 2009; Prins 2006). Such information then changes hands or ownership as part of normal e-commerce transactions or during a firm's strategic decisions which often include selling consumers' personal information lists to other firms (Prins 2006). With the increasing economic importance of services based on the processing of personal data, it is clear that firms, especially online service providers[1], have high incentives to acquire consumers' personal information (Taylor 2004). A look at present-day practices reveals that consumers'

[1] Throughout the paper online service provider and online business are used interchangeably

profile data is now considered one of the most valuable assets owned by online businesses (Cavoukian 2009; Taylor 2004; and Schwartz 2004). As a result, a flourishing market in the sale of personal information has been created (Tang et al. 2008; Taylor 2004). Although private data is not traded separately, but when aggregated together, the tiniest nuggets of personal information have value (Prins 2006; Taylor 2004; and Dighton 2003). This paper presents a knowledge empowered agent information system architecture that helps tip the balance of power in favor of consumers.

For many people there is not much difference between the collection and use of their personal information and the shooting of their portrait to be published in a magazine. Both are representatives of the individual. If someone takes a picture of another person and wants to reproduce it in a book, the normal practice is for the photographer to obtain consent from the person to whom the picture belongs. This person is entitled to collect royalties from the use of his picture. But if someone gathers information about an individual - where he lives, what he does, what his habits and lifestyle are - and then reproduces it electronically in a list that is sold over and over again to a variety of online marketing organizations, chances are this individual would not even know it is happening. That is, not until this individual starts to receive such annoyances as junk mail, telemarketing calls, spam e-mail and pop-up web ads in websites he or she visits. Unlike the photographer who may negotiate with the individual a royalty fee against using the picture, the "electronic intruders" collect all the gains for themselves and leave the consumers with nothing except moral and financial hardship (Bergelson 2003). Consumers need mechanisms that help them valuate their personal data and give them the ability to negotiate a payoff or compensation at least in part from the damages that might occur to their privacy.

In this paper, we argue the following: if consumers' personal information is a valuable market asset, should they not be entitled to benefit from their asset as well? Given such argument, a natural question comes to mind is: will online businesses buy into it? From a pure economical perspective, online businesses are expected to benefit from such models as well. Since attention in the information market is one of the main reasons, if not the most important reason, behind the collection of personal information, it is essential for online businesses to find ways to economize on attention (Taylor 2004). In other words, online businesses essentially incur excess costs (junk mail, junk phone calls, junk email etc.) to attract consumers' attention to their products and services. However, if they know precisely what the consumer wants, they could make a much better decision about whether or not to provide the consumer with information about their services.

The scope of this paper is on developing knowledge-empowered agent information system architecture for privacy payoff as a means of rewarding consumers for sharing their personal information. Software agents are becoming a choice technology for carrying out complex tasks that many users cannot do on their own (Ribaric and Hrkac 2008). For example, in order for a user to valuate his or her personal information in an open environment such as e-commerce, the user needs to fulfill certain requirements of personal data categorization, ontological formulation and data composition, risk assessment and quantification, automatic negotiation etc. These tasks are very tedious and sometimes difficult and require certain expertise. Software agents, on the other hand, are capable of performing complex tasks and meet their design objectives in the environment where they reside (Wooldridge and Jennings 1995). As such, we have chosen to design and implement an agent-based information system. Even though the system itself was complex and challenging to build, it makes our proposed eCommerce application easy to use and understand by a user.

The proposed system, described in section 3, consists of several agents that incorporate necessary knowledge in order to maximize the consumers' benefit. The agents in the knowledge empowered information system bear the responsibility of working on behalf of consumers to categorize their personal data objects, report to consumers on online businesses' trust and reputation, determine the value of their compensation using known financial models, and finally negotiate for a payoff value in return for the dissemination of users' information. The viability of the system is demonstrated through detailed analysis and a prototype implementation which is based on the JADE multi-agent framework.

To the best of our knowledge, no other work has considered employing software agents in information system architecture for privacy payoff as a means of rewarding consumers for sharing their personal information in e-commerce similar to our work. Contributions made in this work are as follows:

- We propose a knowledge-empowered agent information system for privacy payoff. This is a rather significant and somewhat complex multi-agent information system that, in the end, makes the user's life easy when dealing with privacy payoff negotiation;
- We propose a new technique to valuate private data based on privacy risks, and use fuzzy logic in the agents for determining the trustworthiness of online businesses;
- We extend existing trust models by introducing a new layer of assessment based on privacy credentials, thus advancing the state of the art in this aspect as well;

- We extend an agent-based negotiation protocol presented by Su el al. (2000) for the goal of maximizing the consumers' benefit;

The remainder of this paper is organized as follows: in the next section, we discuss the related work. In section 3, the proposed system architecture is presented followed by proof of concept implementation in section 4. In section 5 we discuss our work. Finally, in section 6 we conclude the paper and present plans for future work.

# 2 Related Works

This section presents work related on agent-based eCommerce information systems and privacy systems in eCommerce. In particular, we examine a sample of such studies that we believe to be representative and specifically related to the work in our topic. Although far from being exhaustive, this section gives a rather complete idea of the current state of the art.

## 2.1 Agent-based eCommerce Information Systems

The rich content and the dynamic nature of eCommerce have made shopping activity a process that requires large effort (Sierra 2004). A human buyer, for example, requires collecting and interpreting information on merchants, products and services, making optimal decisions and finally entering appropriate purchase and payment information. Developing actionable trading agents help automate a variety of activities, mostly time consuming ones, and thus lower the transaction costs. An example of such work is presented by Cao and He (2009). The study uses agents that can produce optimal trading strategies to be directly used by users to take decision-making actions in eCommerce situations. Other examples of agent-based systems that are extensively studied in eCommerce are negotiation systems. Zhaung et al. (2008) propose a knowledge-based system for automated negotiation in eCommerce. In such a system the authors emphasize the utilization of knowledge originated from historical negotiation data in estimating and fine-tuning the negotiation parameters, for improving the performance of automated negotiation.

Albert el. al. (2004) propose an agent-based electronic market architecture, called GEMS, where the market incorporates ontologies for the different user perspectives. In their system, agent models were used at a high-level where knowledge is included to relate information from different perspectives; for example evaluation knowledge that can be used to derive product evaluations in terms of user ontology from product information based on producer ontology. GEMS focuses on the use of generic agent models, knowledge representation and ontology

design. Besides these ontologies, also domain-specific evaluation and matching knowledge relating the different ontologies has been included in the system. The system has a transparent compositional structure based on a generic broker agent model and is flexible for maintenance in changing circumstances.

Bagnal and Toft (2006) describe an autonomous adaptive agent for single seller sealed bid auctions. The aim of their work is "to determine the best agent learning algorithm for bidding in a variety of single seller auction structures in both static environments where a known optimal strategy exists and in complex environments where the optimal strategy may be constantly changing". Similarly, Aknine et al (2004) propose a solution that enables an agent to manage several simultaneous negotiation processes with the ability to detect failures.

Other forms of agent-based information systems apply the concept of consumer driven eCommerce societies, such as the work of Sensoy and Yolum (2009) which proposes a multi-agent system of consumers that represent their service needs semantically using ontologies. The aim of the system is to allow consumer agents to create new service descriptions and share them with other consumer agents thus creating an eCommerce society of consumers. According to the authors, such a system leads to better offerings from service providers as they compete among each other to provide attractive promotions and target service consumers more effectively.

While our system employs agents to automate activities similar to the above mentioned work, it tackles a rather unique problem in eCommerce (privacy payoff) which is novel and has not been addressed by any previous work. Each agent in our system autonomously assumes a specific activity such as consumers' personal data categorization, trust and reputation, payoff computation, and negotiation. Furthermore, our study, like Sensoy and Yolum (2009) study, uses agents which utilize attribute ontologies to aggregate consumers' information records according to a specific interest thus forming a community of eCommerce consumers. The formation of such a community is beneficial for an individual consumer, because the agent negotiating on behalf of a consumers' community would be in a better position to bargain over the revelation of their personal information and get something of value in return

To give a better understanding of our work, we need to also describe some of the existing work in privacy systems for eCommerce, although they are not agent-based. This is shown next.

## 2.2 Privacy Systems in eCommerce

Since the release of the Platform for Privacy Preferences (P3P) by the World Wide Web Consortium (W3C), several studies were conducted to integrate legal requirements into the mechanism of data revelation such as the work in Desmarais et al. (2007) and Lioudakis et al.

(2007). Both studies focus on the definition of privacy policies and their enforcement on user's private data in eCommerce.

In terms of personal information valuation as a topic, Krause and Horvitz (2008) explored the economics of privacy in eCommerce personalization systems, where people can opt to share personal information in return for enhancements in the quality of an online service. Before Krause and Horvitz (2008), many other studies in the literature discuss similar systems (Acquisti and Grossklags 2003, Syverson 2003, Acquisti 2004; Rafaeli and Raban 2005): from incomplete information about privacy threats and defenses, to bounded ability to deal with their complex trade-offs; from low (and decreasing) privacy sensitivities, to behavioral phenomena, such as immediate gratification.

The work of Lauden (1996) and Preibush (2005) is somehow close to ours. They present mechanisms that allow users to potentially benefit from sharing private data. Preibush (2005) presents a system based on negotiating privacy versus rewards. During the negotiation process the online business starts with a basic offer, consisting of a small discount and few personal data to be asked. The system does not, however, consider the value of private data in question and therefore the online business has the upper hand in the negotiation process. Lauden (1996) proposes the National Information Accounts, a market-based negotiation system in which information about individuals is traded at a market-clearing price, to the level where supply satisfies demand. The use of information would be limited to a specific period of time, and maybe, for specific purposes. This system entails a strict control of the information transfer in the society, possibly imposed and maintained by the government. This model, although interesting and ingenious, does not suit the multiple possibilities to collect, store and process information in a networked society, centered around the Internet, as a global, unregulated communication channel.

Our work is different from these previous studies in many aspects. We propose how to deal with raw personal information records and convert them into a meaningful information asset that satisfies the greatest number of consumers. To do so, we put forward an architecture that employs several autonomous agents leveraging computational and artificial intelligent methods, such as fuzzy logic, ontology formulation, strategic negotiation, and privacy risk quantification for the goal of maximizing consumers' benefit. Thus, the presented architecture would be of scientific value and a contribution to the field .The details of the system are described in the next section.

# 3 System Architecture

In this section, the proposed information system architecture is presented. A high level view is described in the next subsection 3.1, followed by detailed descriptions of each agent in the system; specifically, Information agent in 3.2, Trust and reputation agent in 3.3, Payoff agent in 3.4, and Negotiation agent in 3.5.

In our previous study (Yassine and Shirmohammadi 2008, 2009) we introduced analysis related to the information agent (3.2) and the payoff agent (3.4), however in order to present a self-contained article we are re-presenting them here as well with extensions that are related to the information system design as we are going to see in the next few subsections.

## 3.1 High level view

Figure 1 depicts the high level architecture of the proposed system. Consumers open their accounts in the system and record their taste and preferences while online service providers negotiate, through their engaged agents, with the system to acquire the data.

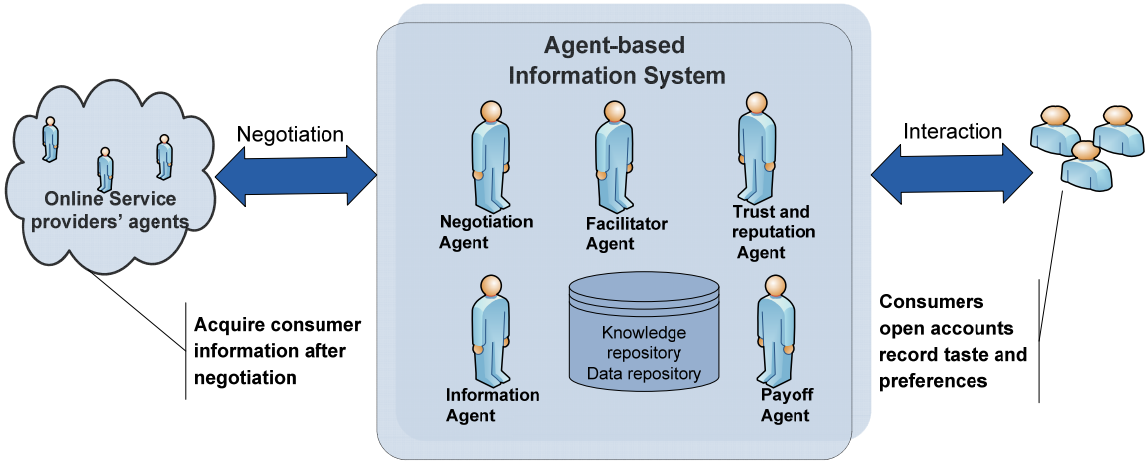


Figure 1. High level architecture of the proposed system

The agents in the system takes on the responsibility of helping consumers valuate their personal data objects that are perceived to be valuable, in order to capitalize on them for the goal of maximizing their market value once the consumer decides to reveal them. Each agent is an independent entity with its own role and information. However, the information under control by an individual agent is not sufficient to satisfy its goal, and so the agent must collaborate with other agents. For example, the payoff agent's goal is to determine the value that the consumer should receive against the revelation of her personal information. However, the payoff valuation is based on the privacy risk that is determined by the information agent. Thus, in order for the

payoff agent to satisfy its goal it must receive this information from the information agent. Next, we will explain the agents' roles and provide details about each agent in the coming subsections.

To explain how the system works, consider the following scenario: Alice is a privacy pragmatic person. According to Spiekermann (2001), privacy pragmatic person is defined as a person who is sensitive to data profiling but generally is willing to provide data. Alice is willing to share her personal data preferences with certain online service providers for a discount value or a reward fee. But Alice has certain requirements before she consents to complete the transaction:

(1) She wants complete information about the service provider's privacy practices and its trustworthiness
(2) She wants to determine the level of risk involved in the transaction based on information from (1)
(3) She wants the system to valuate her data combination risk based on her perceived risk of each private data object
(4) She wants her reward or discount value to be valuated based on the involved risk from (3)
(5) She wants the party that negotiates on her behalf to be strategic during the negotiation process so she can get the maximum benefit
(6) She wants to have the privilege of accepting or denying the final offer

To satisfy Alice's requirements the system employs several agents where each one performs a specific task. The order of task execution is captured in the interaction diagram shown in Figure 2.

(1) Alice opens her account that records her taste, preferences, and personal data (essentially a detailed subscription form that needs to be filled once, but can be updated as desired, an example of such a form is provided in Section 4). Her information is stored in the data repository. The information agent automatically classifies the data into different categories, such as contact data, personal ID data, and hobby data (explained later in subsection 3.2). The decision about data categorization is performed based on ontological formulation. The agent uses ontology attributes to reason about the combination of private data.
(2) When the service provider submits a request to obtain personal data (for example, data about consumers who like sports), the trust and reputation agent assess the trustworthiness of the service provider. The trustworthiness assessment is mainly based on the competency of the service provider with respect to privacy and private data handling. The trust and reputation agent employs fuzzy logic techniques (details

of this technique are discussed in subsection 3.3) to rate the privacy credentials of the online business. Privacy credentials (such as privacy seal, membership to privacy auditing services, security seals, authentication mechanisms, contents of privacy statement, etc.) are attributes which can be thought of as the means by which one can judge the competency of the online business with respect to privacy and private data protection (thus satisfying Alice's requirement 1)

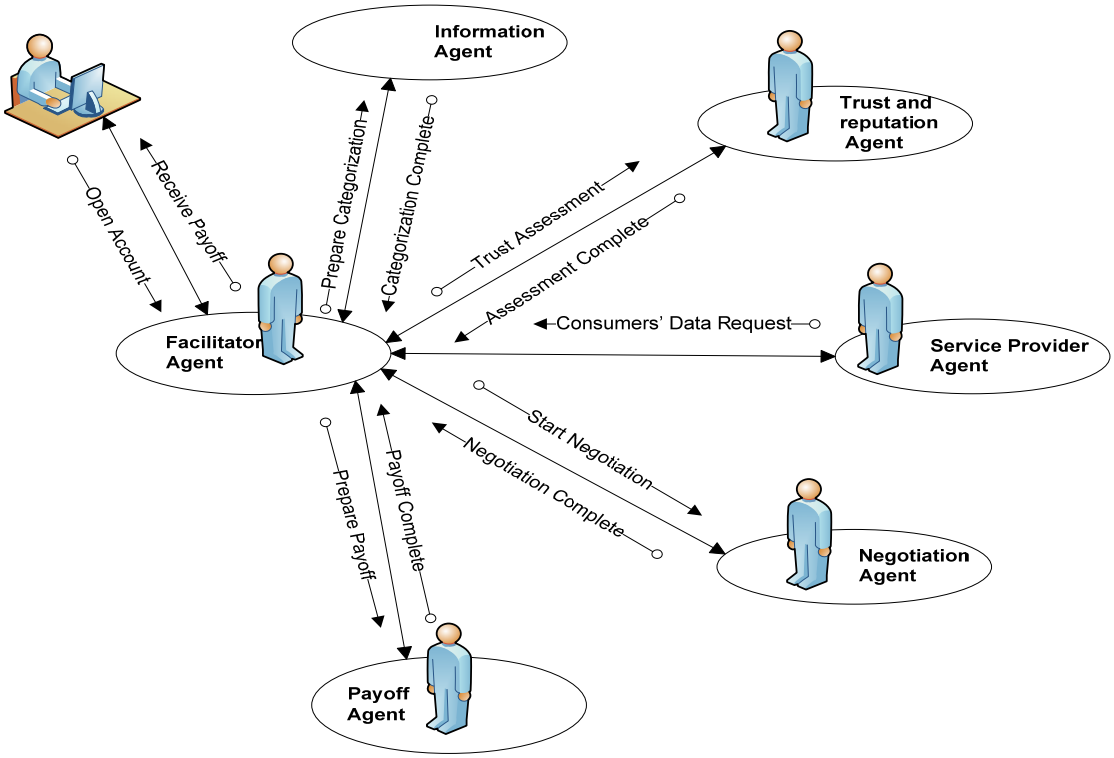


Figure2: Interaction diagram of the proposed system

(3) The trustworthiness rating is communicated back to Alice so she can make an informed decision when assigning the level of privacy risk to the categories of her private data objects (thus satisfying Alice's requirement 2).

(4) After Alice assigns her privacy risk weights to her private data objects, the information agent computes the total privacy risk value for each consumer as well as the combination risk (thus satisfying Alice's requirement 3).

(5) The payoff agent uses the quantified privacy risk value to compute the payoff which Alice should receive if she decides to share her personal information with the online business. The payoff or the compensation is seen as a risk premium valuated in

response to the amount of potential damages that might occur (cost of risk) with respect to the risk exposure (thus satisfying Alice's requirement 4). The payoff agent uses a computational model similar to the models used by the financial and insurance institutions (explained in subsection 3.4).

(6) Once the payoff is determined, the negotiation agent negotiates with the online business on Alice's behalf. The intelligence of the negotiation agent is realized by the negotiation strategy; i.e., the strategy of making concessions and evaluating the incoming offers and making decisions to interact with the negotiation opponent (thus satisfying Alice's requirements 5). The outcome of the negotiation is communicated with Alice. Alice either accepts the offer if it is equal or greater than the expected payoff and in this case shares her data with the online business, or denies the offer and in this case her personal data will not be shared with the online business (thus satisfying Alice's requirements 6).

The order of task execution described above is coordinated by the facilitator agent as seen in Figure 2. The facilitator agent is equipped with knowledge management capabilities. The workspace of the facilitator agent consists of a knowledge repository that keeps the general knowledge of interaction among the agents. All incoming and outgoing tasks are managed by the facilitator agent. Once it dispatches the work order, the execution of the task is completely handled by the responsible agent. While agents can interact with each other in a peer-to-peer fashion if required, we chose to use a facilitator agent for coordination and scheduling of tasks as well as a single point of contact for users as well as for service providers. By so doing, we have simplified the communication among agents and reduced resources for scheduling in each agent to one single point i.e. the facilitator agent.

In the proposed system, we assume that there is some form of security assurances (such as public-key infrastructure) to the consumer that her data is well protected in the system. The agent that manages the consumer data has a complete specification of the private data set without necessarily knowing the actual data values. For example, all the data could be accompanied by digital signatures from the consumers using their private keys. Analysis of mechanisms that ensure the security is out of the scope of this paper. Interested readers may refer to Piolle et al. (2007) for information for security management in user-centered multi-agent systems.

## 3.2 Information agent

This subsection describes the components of the information agent and their details. In particular, we discuss the attribute ontology used in the process of personal data categorization and the computational process of privacy risk quantification.

### *3.2.1 Attribute ontology*

Any operation that involves private data may be viewed as a process that may potentially result in privacy risk depending on the sensitivity level of the private attributes involved in the operation. The information agent categorizes the private data in a way that captures two important characteristics, namely, semantic equivalency and substitution rate given in the following definitions:

*Semantic equivalency*: Consider two transactions that expect as their input a person's home phone number and the same person's family name. One transaction describes the parameters: PhoneNumber and FamilyName while the second transaction describes the parameters: PhoneNumber and LastName. From a privacy perspective both transactions are equivalent. This is due to the semantic equivalence of FamilyName and LastName. To capture this equivalency among attributes, the agent uses ontology sets of semantically defined attributes. The following are examples of such sets:

Set 1 = {FamilyName, LastName, Surname}

Set 2 = {Address, HomeAddress, Location}

*Substitution rate*: The substitution rate of private data captures the level of risk in relation to private data revelation. Private data attributes that are considered substitutable have a constant substitution rate; i.e., the level of exposure risk stays the same. On the other hand, private data attributes that are not substitutable may result in an increase in the privacy risk. Substitution rate is explained next when we describe data classification.

### *3.2.2. Data classification*

The agent receives data specification from consumers which reflects their true personal information and then classifies them into $M$ different categories $(C_1, C_2,..,C_M)$, such as personal identification, contact information, address, hobbies, tastes, etc. In every category $C_i$, private data are further divided into subsets, that is $C_i=\{S_{ik}$ for $k=1, 2,\ldots\}$ based on the attribute ontology that applies to the data set, an example is shown in Figure4.

In figure 4, we consider category Contact, which is divided into two subsets, Telephone and E-mail. The information agent uses the attribute ontology to reason about the classification of the data as follows: first, data in different categories may have different context-dependent weights. The value of the private data may differ from one context to another (personal data value in context sports and personal data value in context health); therefore, its composition may have different implications on the level of revelation.

Second, the substitution rate of private data in the same subset is constant and independent from the current level of revealed data, i.e. assuming that one of the private data has been revealed, revealing the rest of the data in the same subset will not increase the privacy disclosure risk. For instance, a consumer's contact information can be expressed by work phone number or home phone number. Knowing both of them at the same time allows only marginal improvements. This will allow considering each private data subset as one unit.

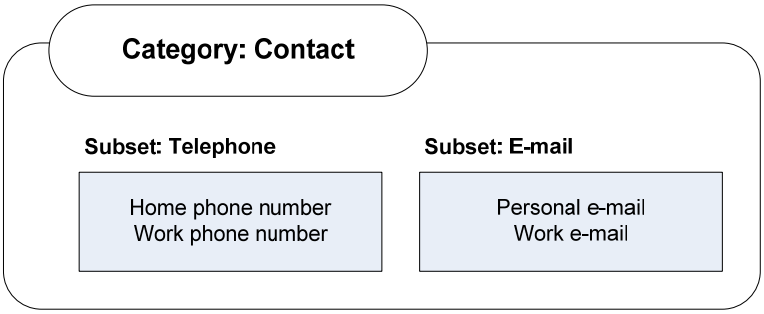


Figure 4. Example of private data categorization

Third, data in different subsets are not substitutable; revealing any one of them will increase the privacy risk. For example, the consumer's telephone number and her email address constitute two possible ways to contact the consumer but they are not completely interchangeable.

### *3.2.3 Risk quantification*

After the agent classifies the personal data into different categories, as explained in the previous subsection, the agent now computes the disclosure privacy risk of private data. But first, it needs to capture the consumers' preferences about revealing different attributes of personal data for each category *i* under context *j*. The context here refers to the situation and the nature of the private data involved in the situation. Individual's privacy cost valuation (cost resulting from privacy risk) differs from one type of information to another. For example, privacy violation of an individual's medical information can be extremely costly for its owner, but the violation of the person's past shopping history would hopefully not be as detrimental. The context under which the private data is revealed could affect the level of privacy cost. For example, providing medical

information to an online drug store is not the same as providing medical information to a government health agency even if it is the same information. In the former, the risk of leaking information about the health condition of the consumer to an insurance company could result in service denial from the insurance company; however, in the latter, the risk of leaking health information to a third party is less likely to happen.

We use the parameter $\beta_{ij} \in [0, 1]$ to capture the privacy risk of each private data in category *i* under context *j*. $\beta_{ij}$ represents the consumer's valuation of her private data under each context (we will explain how the consumer will specify this value in section 4). More specifically, it represents the consumer's type (for example, Alice is privacy pragmatic), where consumers with a higher $\beta$ incur higher privacy costs as a result of privacy violation. Each consumer is assumed to be stochastically equivalent and has independent distributed valuation for each context. According to Preibush (2005) and Yu et al. (2006), individuals have various global valuation levels for each transaction that involves private data.

Let us consider a consumer with identity I, and several private data attributes $\Lambda = \{\Lambda_1, \Lambda_2, ..., \Lambda_n\}$. We let the privacy risk to vary with $\Lambda$ and $\beta$ according to the functional form $\Psi(\Lambda, \beta)$, where $\Psi(\Lambda, \beta)$ characterizes the magnitude of privacy risk resulting from the composition of different private data attributes $\Lambda$ and the cost $\beta$ (privacy risk cost) of revealing them. The calculation of $\Psi(\Lambda, \beta)$ is as follows (from now on, and for the sake of clarity, we will drop the symbols $\Lambda$ and $\beta$ ):

Let Xi be the cardinality of Ci (i.e., Xi = |Ci|). We perform normalization over the whole private data set. The normalized data size NXi reflects the ratio of each category in a consumer's privacy.

$$NX_i = \frac{X_i}{\sum_{k=1}^{M} X_k} \tag{1}$$

where *M* is the number of the private data categories.

After the consumer assigns the value $\beta_{ij}$ for each category *i* under context *j*, the agent computes the weighted privacy risk, as in (2) and then normalizes it, as in (3). The value $N\Psi_{ij}$ in (3) reflects the risk of revealing all data in category i under context j. The purpose of this normalization is to put the privacy risk of each category into the range of [0,1].

$$\Psi_{ij} = NX_i * \beta_{ij} \tag{2}$$

$$N\Psi_{ij} = \frac{\Psi_{ij}}{\sum_{n=1}^{i} \Psi_{nj}} \quad (3)$$

The privacy risk weight of revealing $\alpha_i$ subsets from category *i* is calculated as in (4).

$$\Psi_i = \frac{\alpha_i}{X_i} * N\Psi_{ij} \quad (4)$$

Example: Consider that we have two consumers, Alice and Bob. Alice is privacy pragmatic and Bob is privacy fundamentalist (i.e. Bob is very concerned about the use of his personal data, as described in Spiekermann [2001]). For the sake of simplicity, consider that both of them have similar categories as explained below:

Category C1 (Contact) has two subsets S11 (telephone) and S12 (e-mail). Subset (telephone) has private data d1 (home phone number), d2 (work phone number), while subset (e-mail) has private data d3 (personal e-mail) and d4 (work e-mail).

Category C2 (Hobbies) has two subsets, S21 (online games) and S22 (sports), subset (online games) has two private data, d5 (Ogame) and d6 (Hattrick), while subset (sports) has private data d7 (soccer).

Category C3 (Income) has one subset, S31 (salary), with private data d8 ($100K).

Consider that an online business called sportsworld.com is interested in acquiring information about individuals who like sports, and the provider is interested in getting the following information: home phone number, personal email, soccer, and the salary value. Assume that Alice and Bob have obtained the trustworthiness assessment of this provider (i.e., its reputation and its privacy credential ratings) from the trust and reputation agent. Both of them assign the privacy risk weights to their private data categories as shown in Table 1.

Table 1: Context-dependent weights for Alice and Bob

| Weight | Category (contact) | Category (hobbies) | Category (income) |
|---|---|---|---|
| Context j = sports | Bob W1j = 0.4 | Bob W2j = 0.1 | Bob W3j = 0.5 |
| | Alice W1j = 0.2 | Alice W2j = 0.1 | Alice W3j = 0.2 |

Applying equations (2), (3), and (4), the overall privacy risk value $\Psi$ for Alice and Bob is 0.74 and 0.92 respectively. The privacy risk values 0.74 and 0.92 represent the reluctance level of each individual with respect to the revelation of his or her personal information. In reality, some people value their personal information more than others; say a celebrity's cell phone number versus a student's cell phone number. Therefore, the computation performed by the information agent captures the different valuations that people might assign to their private data.

This value will be used by the payoff agent to place a reward amount that reflects the level of risk perceived by each individual.

Next we present the details of the trust and reputation agent.

### 3.3 Trust and reputation agent

As mentioned earlier in section 3, in order for Alice to assign privacy risk weights to her private data objects, she wants to know the reputation and the competency of the online business with respect to privacy and private data handling.

In recent years, trust and reputation systems have emerged as a way to reduce the risk entailed in interactions among total strangers in electronic marketplaces. While there are many definitions of trust, we follow Gambetta (1988), where trust is defined to be "*a particular level of subjective probability with which an agent assesses that another agent will perform a particular action, both before the assessing agent can monitor such an action and in a context in which it affects the assessing agent's own action*." According to Gambetta (1988), trust is the perception of confidence. It helps reduce the complexity of decisions that have to be taken in the presence of many risks. In the context of eCommerce (or any kind of online transaction), perceived privacy is often thought to be another important antecedent of trust (Chen and Barnes 2007) and is defined as the subjective possibility of accordance between consumers' anticipation and their cognition of how their private information is being used (Yang 2005). It is the perception that the online business will adhere to an acceptable set of practices and principles (Lee and Turban 2001).

On the other hand, reputation is based on past behavior where derived reputation scores are assumed to predict likely future behavior. In the context of eCommerce, reputation acts as a signal of the service provider's skill and integrity. Consumers can use reputation information to distinguish dishonest sellers from good ones, thereby overcoming adverse selection. However, it should be mentioned here that highly reputable websites do not always mean that the service provider has high privacy credential ratings. Therefore, consumers need an additional mechanism to make them aware of the privacy practices of online service providers.

To understand how the service provider is going to handle the personal information, it is essential to assess its trustworthiness based on privacy credential attributes which the online provider has obtained (e.g., trustmark seals, privacy certificates, contents of the privacy statement, encryption mechanisms, etc.). Empirical research on the impact of such attributes on individuals' perceived trust, e.g., that of Ratnasingam and Pavlou (2003), Sha (2009), and Teo and Liu (2003), to name few, report that the impact varies depending on the type of the attribute. Examples of such attributes are shown in Table 2.

In this paper, the rating of each attribute as shown in Table 2 under column "Significance" is based on studies on consumer research analysis such as those of Chouk et al. 2007, Gideon 2006, Hu et al. 2003, Larose 2004, Sha 2009, Kimery and McCord 2002 and industry studies such as those of Cline 2003 and Hussain et al. 2007. These studies analyze the impact of different trusting attributes on consumers' online purchasing behavior and their perceived trust For example, in Sha (2009) and Gideon et al. (2006) the "use of encryption mechanism" and "periodic reporting on privacy matters" have low impact on the perceived trust according to consumer research analysis.

The attributes and their classifications are stored in a database dedicated for the privacy credential ratings. The trust and reputation agent uses a fuzzy logic mechanism (discussed in the next subsection) to determine the privacy credential score and the privacy report for each service provider (implementation of such report is shown in section 4).

Table 2: Examples of trusting attributes and their significance

| **Credential** | **Type** | **Significance** |
|---|---|---|
| Trustmark seals | e.g., Reliability Seals, Security Seals, Vulnerability Seals, Privacy Seals | Varies depending on the type of the seal and the issuing party (Sha 2009, Zhang 2005) |
| Dispute resolution | Independent resolution mechanisms, e.g., American Arbitration Association | Medium (Chouk et al. 2007) |
| Membership to privacy compliance auditing service | e.g., Pricewaterhousecooper, PrivaTech | High (Cline 2003, Pennington et al. 2003) |
| Privacy statement | Privacy policy generator, e.g., Organization for Economics Cooperation and Development OECD | Varies depending on the contents (Meinert et al. 2006) |
| Authentication-based disclosure of information | e.g., Kerberos-based vs. username/password authentication | Low (Sha 2009, Hu et al. 2003) |
| Periodic reporting on privacy matters | e.g., reports on privacy training, reports on privacy compliances | Low (Sha 2009) |
| Allows consumers to opt out | Opt out of data sharing and marketing solicitations | High (Gideon et al. 2006) |
| Use of Encryption Mechanisms | e.g., 128-bit is better than 64-bit encryption scheme | Low (Sha 2009, Gideon et al. 2006) |

Before discussing the fuzzy logic mechanism used by the agent, it should be mentioned here that in the proposed information system, we assume two methods of collecting the privacy credential attributes and store them in the database: automated, and manual. The automated process relies on machine-readable mechanisms used by the service provider to publish its

privacy attributes. For example, a service provider publishes an XML document in a policy, such as Enterprise Privacy Authorization Language, P3P (The Platform for Privacy Preferences) protocol, or SOAP (Simple Object Access Protocol) messages detailing how personal information is handled, the security level that is supported, dispute resolution etc. Other possible way of populating the database is the use of software agent where an agent travels into a website and report on the existence or otherwise of the trust attributes. In the manual process, administrators use an interface to populate the privacy attributes of the online businesses. This process is used only when automation is not possible, such as the case when the service provider does not use P3P or other mechanisms to describe its privacy policies.

### 3.3.1 Fuzzy logic-based rating

Privacy attributes can be thought of as the means by which one can judge if the online business is trustworthy to handle the consumers' private data. As shown in Table 2, attributes differ in their impact according to their significance (low, medium, and high).

Let us assume that $\zeta_k$ is the set of K attributes, written as follows: $\zeta_k = \{\zeta_1, \zeta_2, \zeta_3 ..., \zeta_K\}$; k = 1,..,K. As mentioned earlier, each attribute has its own impact on the perceived trust as shown in Table 2. The set $\zeta_k$ is classified into three subsets $\zeta_f^h, \zeta_g^m$, and $\zeta_w^l$ of high, medium, and low credentials respectively, such that $\zeta_f^h \cup \zeta_g^h \cup \zeta_w^l = \zeta_k$. Subsets $\zeta_f^h$, $\zeta_g^m$ and $\zeta_w^l$ will be used as reference sets against which the online business will be measured. In reality, the service provider may have obtained different types of attributes, each of which has a different impact on the perceived trust. We introduce here a factor, called Online Provider Reliability $\Phi$, which helps determine the competency of the online business with respect to privacy and private data handling. The competency is an indication to what extent the service provider can be trustworthy to respect consumer's private data and adhere to the best practices of respecting consumers' privacy. Let $\varphi$ be the set of credential attributes that the online business has. In reality, $\varphi$ would include attributes of different types i.e., high, medium, and low (e.g. Sha 2009, Chouk et al. 2007, Zhang 2005). According to Sha 2009, Kim et. al. 2008, Holger et. al. 2007, and Larose 2004 the higher the number of credential attributes, the more reliable is the online provider to protect consumers' personal data. However, as mentioned earlier, the service provider may have obtained different types of attributes. Therefore, its reliability $\Phi$ is based on such combination. We denote by $\Theta$ to

represent the combination of X, Y, and Z then we determine the online service provider's reliability $\Phi$ as follows:

$\Phi = X \Theta Y \Theta Z$, such that $X \leq \left|\zeta_f^h\right|, Y \leq \left|\zeta_g^m\right|$, and $Z \leq \left|\zeta_w^l\right|$, where $\left|\zeta_f^h\right|, \left|\zeta_g^m\right|$, and $\left|\zeta_w^l\right|$ are the cardinalities of sets $\zeta_f^h, \zeta_g^m$ and $\zeta_w^l$ respectively.

An important facet of the above combination is the rules that determine how the value of $\Phi$ will be determined. This is because the weights of the attributes in the combination are often set by using linguistic variables (Sha 2009, Schalgar and Pernul 2008, Chouk et. al. 2007,) such as "I believe that if the service provider provides a readable, easy-to-understand privacy statement, then my trust in his intentions is high." Setting a label of HIGH may be interpreted to represent more than one numerical value. This is because human perception is vague and uncertain. To avoid assigning a specific numeric value to this rather subjective concept, we follow the work of (Xin et al. 2006, Schalgar and Pernul 2008) and use fuzzy rules where we can subdivide a range [0,1] into a number of linguistic labels such as VeryLow, Low, Moderate, High, and VeryHigh. Not to mention that the valuation of the acceptance of the service provider's reliability value (from human perspective) is too complicated by using conventional mathematical methodologies (Schalgar and Pernul 2008). Fuzzy logic helps us determine the final linguistic label which is translated into a single crisp value, which could be a numeric scale if needed. In our case, the final linguistic label is translated into a scale composed of five "Stars" such that one star is VeryLow, two stars is Low, and so on and so forth. Using fuzzy logic, the reliability of the online business is determined as follows:

IF $\Phi$ is high

THEN privacy risk is low

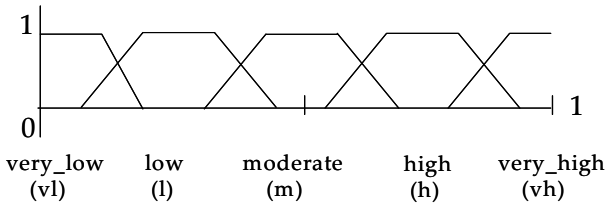


Figure 5: Fuzzy sets for variable $\Phi$

Figure 5 shows the fuzzy sets for the variable $\Phi$. The heuristic behind this rule is that online businesses which show a high degree of competency to protect consumers' data are assumed to honor their promises and can be trusted. Table 3 shows a possible set of fuzzy rules

which the mechanism uses to determine the reliability $\Phi$. In Table 3, *vh, h, m, l,* and *vl* stands for *very high*, *high*, *medium*, *low* and *very low* respectively. Note that the rules presented in Table 3 tend to ensure the minimum credentials to determine the value of the variable $\Phi$. For example, the first rule imposes a condition that all online businesses which obtain at least half of the required attributes in each category are considered very reliable. These rules are defined based on common sense. That is, it is quite normal for a consumer to trust a service provider with high privacy credential attributes. There are empirical studies such as those of Sha (2009) and Zhang (2005) suggests that as the number of credential increases the confidence in the service provider increases. In this paper, we propose a minimum threshold at which we can construct a meaningful result, but yet it is up to the user to set the threshold as desired. It should be clear, however, that if the threshold is set too high, no service provider will be able to meet the condition. And if the threshold is set too low, then all service providers will be indifferent and personal data will be shared with all of them equally which is not desirable.

Table 3: Example of reliability rules

| **Rules** |
|---|
| $\mathbf{IF}\ (X \geq \frac{\lvert\zeta_f^h\rvert}{2}, Y \geq \frac{\lvert\zeta_g^m\rvert}{2}, Z \geq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ vh$ |
| $\mathbf{IF}\ (X \geq \frac{\lvert\zeta_f^h\rvert}{2}, Y \geq \frac{\lvert\zeta_g^m\rvert}{2}, Z \leq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ vh$ |
| $\mathbf{IF}\ (X \geq \frac{\lvert\zeta_f^h\rvert}{2}, Y \leq \frac{\lvert\zeta_g^m\rvert}{2}, Z \geq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ h$ |
| $\mathbf{IF}\ (X \geq \frac{\lvert\zeta_f^h\rvert}{2}, Y \leq \frac{\lvert\zeta_g^m\rvert}{2}, Z \leq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ h$ |
| $\mathbf{IF}\ (X < \frac{\lvert\zeta_f^h\rvert}{2}, Y \geq \frac{\lvert\zeta_g^m\rvert}{2}, Z \geq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ m$ |
| $\mathbf{IF}\ (X < \frac{\lvert\zeta_f^h\rvert}{2}, Y \geq \frac{\lvert\zeta_g^m\rvert}{2}, Z \leq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ m$ |
| $\mathbf{IF}\ (X < \frac{\lvert\zeta_f^h\rvert}{2}, Y \leq \frac{\lvert\zeta_g^m\rvert}{2}, Z \geq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ m$ |
| $\mathbf{IF}\ (X < \frac{\lvert\zeta_f^h\rvert}{2}, Y \leq \frac{\lvert\zeta_g^m\rvert}{2}, Z \leq \frac{\lvert\zeta_w^l\rvert}{2})\ \mathbf{THEN}\ \Phi\ is\ l$ |

| |
|---|
| **IF** $(X < \frac{\left\|\zeta_f^h\right\|}{2}, Y < \frac{\left\|\zeta_g^m\right\|}{2}, Z \geq \frac{\left\|\zeta_w^l\right\|}{2})$ **THEN** $\Phi$ *is l* |
| **IF** $(X < \frac{\left\|\zeta_f^h\right\|}{2}, Y < \frac{\left\|\zeta_g^m\right\|}{2}, Z \leq \frac{\left\|\zeta_w^l\right\|}{2})$ **THEN** $\Phi$ *is vl* |
| **IF** $(X < \frac{\left\|\zeta_f^h\right\|}{2}, Y < \frac{\left\|\zeta_g^m\right\|}{2}, Z < \frac{\left\|\zeta_w^l\right\|}{2})$ **THEN** $\Phi$ *is vl* |

### 3.4 Payoff agent

In this section, we present the details of the payoff agent. The payoff agent computes the reward value which the consumer should receive as a result of revealing their personal information. The payoff is computed in response to the amount of potential damages that might occur (cost of risk) with respect to the risk exposure. The payoff, which can be seen as a risk-base premium or compensation, is the worldwide practice of the financial institutions and the insurance industry, and is implied in many other industries (Bakshi et al. 2006). The basic idea is that those who are considered more risky should pay more interest. This means that risk-based premium is a quantified value applied to compensate for unwanted events that might lead to revenue loss or increased cost. Privacy risk is the business risk resulting from the use, mishandling of personal information. Like all business risk, privacy risk could result in financial hardship (Cavoukian 2009, Lauden 1996).

Previously, in subsection 3.2.3, we showed how the information agent determines the privacy risk value $\Psi(\Lambda, \beta)$ that the consumer expects when revealing his or her personal information. The payoff agent's responsibility is to determine the payoff value that the consumer should receive given the privacy risks that are involved. Intuitively, we want to associate high benefit/compensation with $\Psi(\Lambda, \beta)$ that allow high identification of $I$ (the identity of the consumer) given $\Lambda$ and high risk $\beta$. One approach for determining this value is to construct a risk premium that is valuated in response to the amount of potential damages that might occur (cost of risk) with respect to the risk exposure. In this manner, the compensation paid to the consumer is justified, at least in part, from the damages that might occur. This approach will help make service providers more conservative when handling users' personal information. This is because privacy risk penalties and reputation consequences on the violators of users' presumed

privacy rights are more likely to be costly. This approach is widely used in the financial and auto insurance industry (Danthine and Donalaon, 2002).

Theoretically, the expected payoff of a risky asset in conjunction with expectations of the risk-free return should be used to construct the risk premium (Danthine and Donalaon, 2002). The basic idea is that those who are considered more risky should pay more interest. The standard model (Danthine and Donalaon, 2002) is:

$$Expected\ return = RiskfreeRate + AssetRisk \cdot Market\ Risk\ \ premium \tag{5}$$

where:

Expected return is the payoff given to the consumer against revealing their private data (i.e., their asset);

RiskfreeRate is the expected return value at zero risk;

Market Risk premium is the expected market risk and it is positive; and

AssetRisk is the calculated risk of the private data (i.e. $\Psi$ calculated from subsection 4.3).

Assuming that the RiskfreeRate is zero (this assumption is valid since consumers do not expect to be compensated for private data if the level of risk is zero), this model requires two inputs. The first is the risk value of the private data asset being analyzed, and the second is the appropriate market risk premium(s). Equation (5) is now formally written as follows:

$$E(U_q) = \Psi_q \bullet E(\Re) \tag{6}$$

where:

E (Uq) is the expected return U of revealing private data q;

$E(\Re)$ is the expected market risk premium; and

$\Psi_q$ is the calculated privacy risk of revealing private data q.

Equation (6) is simply saying that consumers' demand for revealing their personal information is measured based on the perceived risk and the market risk premium. The market risk premium is a monetary value estimated by looking at historical premiums over long time periods (Danthine and Donalaon, 2002). According to Clinebell et al. (1994) and Bakshi et al. (2008), the general behavior of market risk premiums follows a random variable, random walk, or autoregressive process over time. Here, we assume that the risk premium follows a random process and fluctuates over time following a geometric Brownian motion, where the expected risk premium value will be the underlying mean. The geometric Brownian motion assumption has the advantage that the process depends only on two variables, the drift and the standard deviation.

The distribution function of a geometric Brownian motion is lognormal, which has the favorable property that negative future values have zero probability. A stochastic (random) process $\Re$ is said to follow a geometric Brownian motion if it satisfies the following stochastic differential equation (Ross 1995):

$$d\Re_t = \mu\Re_t dt + \sigma\Re_t dW_t \tag{7}$$

Where $\{W_t\}$ is a Wiener process or Brownian motion, $\mu$ is the drift, and $\sigma$ is the standard deviation. Using Ito's lemma (Ross, 1995), we can derive the analytical solution of equation (7) as follows:

$$d\ \ln\ \Re(t) = \frac{1}{\Re(t)} d\Re(t) - \frac{1}{2}\frac{1}{\Re(t)^2} d\Re(t)^2 \tag{8}$$

$$= \frac{1}{\Re(t)}\Re(t)[\mu dt + \sigma dW(t)] - \frac{1}{2}\frac{1}{\Re(t)^2}\Re(t)^2[\sigma^2 dW(t)^2] \tag{9}$$

$$= \mu dt + \sigma dW\ (t) - \frac{1}{2}\sigma^2 dt \tag{10}$$

The integral of (10) from 0 to t is:

$$\int_0^t d\ \ln\Re(t) = \int_0^t (\mu dt + \sigma dW\ (t) - \frac{1}{2}\sigma^2 dt) \tag{11}$$

$$\ln\Re(t) - \ln\Re(0) = (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t) \tag{12}$$

Basic calculus gives us:

$$\Re_t = \Re_o e^{((\mu-\sigma^2/2)t+\sigma W_t)} \tag{13}$$

where $\Re_o$ is the risk premium value at time t = 0; i.e., it is the initial risk premium value.

We now take the expectation $E[\Re_t]$ for (13):

$$E[\Re_t] = E[\Re_o e^{((\mu-\sigma^2/2)t+\sigma W_t)}] \tag{14}$$

Applying the law of normal variables with mean $\mu$ and variance $\sigma^2$ knowing that the Brownian motion $\sim N(0,t)$ we get:

$$E[\Re_t] = \Re_o e^{\mu t} \quad (15)$$

The risk premium is assumed to be known at time 0. Some studies such as (Hann 2003) estimate the market privacy risk premium to be between $45 and $57 per personal data record. Equation (15) estimates the value of the payoff as it varies over time.

**Example:** Consider the example we described in section 3.1. Alice's and Bob's privacy risk value was 0.74 and 0.92 respectively. Assume, for the sake of simplicity, that the estimated risk premium value is $100 for the current period (We note that historical record of risk premiums on risky assets is usually provided by industry agencies or specialized government bodies. In our study we assume that these values are somehow known). Then Alice and Bob are entitled to receive $74 and $92 for their data objects respectively if they decide to share their information.

Next, we provide the details of the negotiation agent.

## 3.5 Negotiation agent

It is a daunting prospect for an individual consumer to bargain with a huge service provider about a desired payoff against revealing personal information. Therefore, having an agent working on behalf of a group of consumers would be in a better position to bargain over the revelation of their personal information and get something of value in return. To set the stage for specifying the negotiation model, some rules and assumptions are given as follows:

- Since a bargaining negotiation is fundamentally time- dependent (Bo et al. 2008), we assume that both the agent representing consumers and the service provider agent utilize a time dependent strategy while making a concession. This assumption is an appealing one since human traders easily understand deadlines, and it is trivially simple to specify a deadline to a software agent. Thus the consumers' agent has a deadline $T^{CA}_{deadline}$ and the service provider agent has a deadline $T^{SP}_{deadline}$.
- The negotiation is non-cooperative with incomplete information and competition among consumers is not considered
- The agents are negotiating over a single issue i.e. the payoff that the consumers should receive against the revelation of their personal information

Having studied the rules of the negotiation, we can now present the negotiation strategy.

### *3.5.1. Negotiation Strategy*

In our setting of incomplete information, the intelligence of the negotiation agent is realized by the negotiation strategy. In particular, the strategy of making concessions and evaluating the incoming offers and of making decisions to interact with the negotiation opponent. This paper adopts the work of Bo et al (2008) in defining a concession strategy. Each agent is assumed to have a different time preference, i.e. its time deadline. In round $0 < t < \min(T^{CA}_{deadline}, T^{SP}_{deadline})$ if the proposal $P^{SP \to CA}_{t-1}$ at round t – 1 is not acceptable to the consumers' agent, the agent may make a concession to service provider's agent at round *t* as reaching an agreement is always better than failing to reach an agreement.

In general, the proposal of agent *A* to its trading partner at round t $(0 \le t \le T^{A}_{deadline})$ is modeled as a time dependent function as follows (Bo et al. 2008):

$$P^{A}_{t} = IP^{A} - \phi^{A}(t) \times (IP^{A} - RP^{A}) \tag{16}$$

where *t* is the current trading time, *IP* is the initial price, *RP* is the reservation price, and $\phi^{A}(t)$ is a time-dependent function based on the time deadline preference.

The time dependent concession strategy is used to decide the amount of concession in each round of the negotiation process. The time-dependent function $\phi^{A}(t)$ is determined with respect to time preference $\eta^{A}$ and deadline $T^{A}_{deadline}$ (where $\eta^{A} \ge 0$ and $T^{A}_{deadline} > 0$ is finite) and is given as:

$$\phi^{A}(t) = \left( \frac{t}{T^{A}_{deadline}} \right)^{\eta^{A}} \tag{17}$$

The open literature has described a large number of negotiation strategies with respect to the remaining trading time (one for each value of $\eta^{A}$). In this paper, we adopt the "sit-and-wait" strategy proposed by Bo et al (2008). The "sit-and-wait" strategy is used in the case when the negotiation issue does not devaluate over time and has been proven by Bo et al (2008) that this strategy is the dominant strategy for an agent using time-dependent negotiation process, regardless of the strategy that its trading partner adopts. Unlike commodities or other services that might be devaluated over time, personal information as a trading object does not have such problem. Therefore, this strategy is suitable to the type of problem we are facing. For the consumers, even if the offer happens only at some future stage of negotiation, as long as it is

anticipated, the consumers' agent who makes this offer has bargaining power. Another important aspect of bargaining power arises from impatience. In order to achieve certain competitive advantages, online businesses have particular service delivery threshold that makes them more impatient. Therefore, they are more likely to make bigger concessions in order to seal the deal as soon as possible.

The "sit-and-wait" strategy for the consumer agent is as follows: At the time when $t < T_{deadline}^{CA}$, it follows that $(t / T_{deadline}^{CA})^{\infty} = 0$, and $P_t^{CA} = IP^{CA}$. When $t = T_{deadline}^{CA}$, it follows that $(T_{deadline}^{CA} / T_{deadline}^{CA})^{\infty} = 1$, and $P_t^{CA} = RP^{CA}$. Let $Q_t^{CA}$ and $Q_{T_{deadline}^{CA}}^{CA}$ be the amounts of concession at $t < T_{deadline}^{CA}$ and $T_{deadline}^{CA}$, respectively. Before the deadline, the agent does not make any concession but "waits" for the service provider agent to concede, since $Q_t^{CA} = P_{t-1}^{CA} - P_t^{CA} = 0$ $(0 \le t < T_{deadline}^{CA})$. It only concedes at its deadline $Q_{T_{deadline}^{CA}}^{CA} = P_{T_{deadline}^{CA}-1}^{CA} - P_{T_{deadline}^{CA}}^{CA} = IP^{CA} - RP^{CA}$. The service provider agent, on the other hand, concedes to its reservation price at $T_{deadline}^{SP}$. In reality, the time deadline for the consumer agent is set to be a large value so the consumer agent can benefit from the impatient of the online business.

### *3.5.2. Offer Construction*

Consider that the negotiation is to acquire consumers' information records which are aggregated together according to a specific interest (e.g. a list of consumers who like soccer or like to travel to England). As mentioned earlier in section 3.5 it is a daunting prospect for an individual consumer to bargain with a huge service provider about a desired payoff against revealing personal information. Therefore consumers are grouped together based on their interest and have one agent bargain on their behalf. The consumers' agent and the service provider's agent therefore will have the following objectives in their offers.

*Consumers' agent offer:* The consumer agent's objective is to maximize the consumers' payoff as a community that is, maximizing their social welfare SW. Consider that N is the number of records in the list then (for simplicity we consider that each consumer is represented by one record), at each round of the negotiation the consumers' agent construct an offer as follows:

$$N' = \arg\max_{N' \subseteq N} \sum_{i \in N'} SW_i \Big| U_q \quad (18)$$

Such that $N' \subseteq N$ and $U_q$ is the payoff per personal data record q (personal data record refers to the vector of information fields that will be revealed). When the consumers' agent receives an offer from the service provider's agent, it calculates the number of records $N'$ that satisfy this offer. The decision about the number of consumers is taken based on an offer/demand curve according to the consumer's valuation of his or her personal information record.

*Service provider's agent offer*: Assume that the service provider's agent has a utility value $V_r$, where $r$ refers to the consumer's record in the list. Let $C_r$ be the cost of acquiring record $r$ (that is, the payoff offer given to the consumer agent per record) then the agent's goal is to maximize the function ($V_r$ - $C_r$). The service provider's agent constructs its offer so that the value $C_r$ is minimized. The optimal value that maximize the utility is when $C_r=0$, but such offer will end the negotiation process in the first round and the service provider will walk away with nothing. The strategy of the service provider's agent is to start with a small value $C_r$ (for example a predefined percentage of the utility $V_r$) and then in each round of the negotiation gradually increases the offer at a step rate equal to $\theta$ (an arbitrary bid increment value determined by the provider and can be adapted during the negotiation if required) until it reaches its *RP* or the deadline expires and in this case concedes to *RP*.

Theoretically, the maximum reservation price which the provider can offer is equal to $V_r$ which yields to zero profit. But practically, this is not acceptable for the service provider. Hence, in practice the upper limit of *RP* is the value that yields minimum acceptable utility. The minimum acceptable utility is at which the difference between the utility and the reservation price is relatively small, that is, ($V_r$ - $C_r$) = $\tau$ where $\tau$ is a positive small value slightly greater than zero.

### *4.5.3. Negotiation Protocol*

A basic condition for the automation of the negotiation process among agents is the existence of a negotiation protocol, which encodes the allowed sequences of actions. Although FIPA ( Foundation for Intelligent Physical Agents) provides a plethora of protocols, such as FIPA brokering, FIPA English auctions, FIPA Contract net protocol etc., we found that there is no agreed upon standard interaction protocol for 1-1 automated negotiation. As a result, we adopt the negotiation protocol proposed by Su et. al. (2000) and implemented by Skylogiannis et. al. (2007). This protocol is a finite state machine that must be hard-coded in all agents participating into negotiation. While we understand that hard-coded protocols in agents may lead to

inflexibility, the focus of this paper is not on protocol design but rather a negotiation strategy that maximizes the consumers' benefit. Following (Skylogiannis et al. 2007), our protocol is a finite state machine with discrete states and transition, see Figure 6.

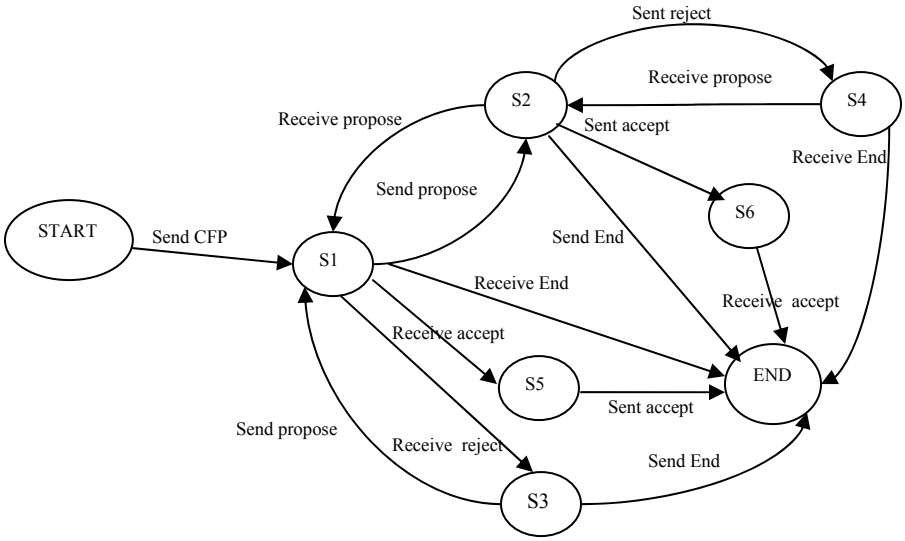


Figure 6: State machine of 1:1 negotiation protocol

In Figure 6, the notations Start, S1, S2, S3, S4, S5, and S6 represents the different state of the negotiation and END is the final state in which there is an agreement, or a failure of agreement between the participants. *Send* and *Receive* primitives specify the interactions between the two agents and cause state transitions. For example, the sequence of transition START->S1->S2->S6->END can be interpreted as follows: the consumer agent initially sends a call for proposal message (CFP) to the provider agent (START->S1), then it receives a propose message (S1->S2) and after the evaluation it decides to send an accept message (S2->S6). Lastly it receives an accept message and the negotiation terminates successfully (S6->END). In the next section, we present a proof of concept implementation.

# 4 Proof of concept implementation

A prototype of the proposed system has been implemented using JADE (Java Agent Development Environment) platform. JADE provides a multi-agent environment which is composed of the FIPA standard agents and of a set of application dependent agents realized by the application developer. The communication component is implemented as a set of classes that inherit the jade.Core.Agent and jade.lang.acl.ACL message of existing classes of the JADE platform. These classes provide a means to construct, send, and receive messages via several

FIPA communication performatives. Figure 7 shows a snapshot of the agents developed under container-1.

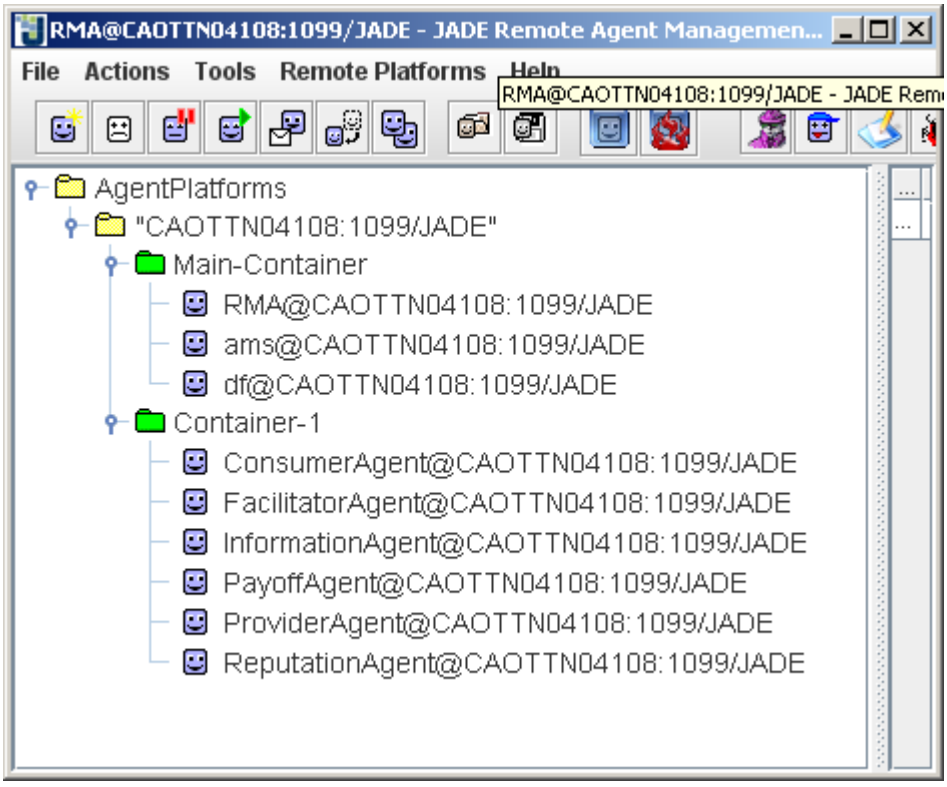


Figure 7: Screen capture of the Remote Management Agent GUI for JADE set-up

Agents developed under container-1 may (but do not have to) exist on a remote machine. For convenience, the agents in our prototype reside on the same host. The users open their accounts and record their preferences through a GUI interface as shown in Figure 8.

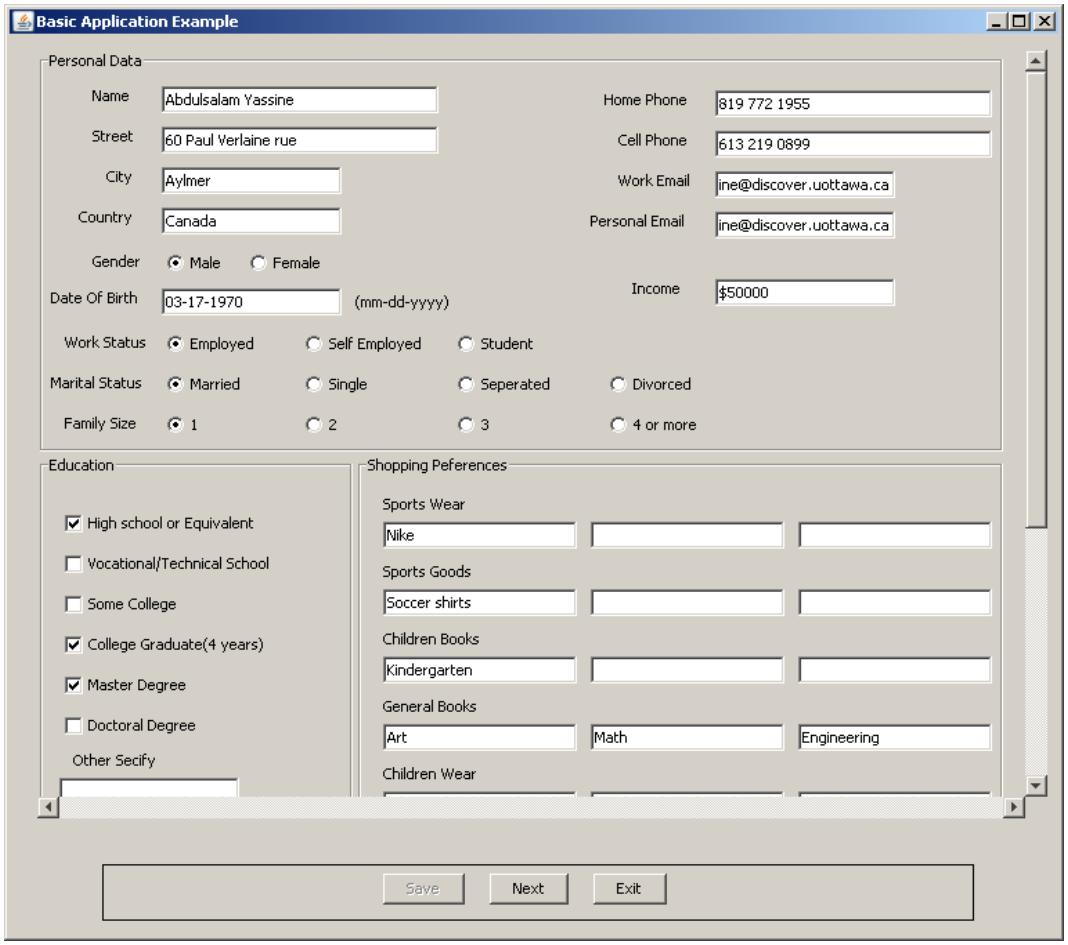


Figure 8: User interface for consumers to open account and record their preferences

Once the user opens the account the facilitator agent sends a message to the information agent to categorize the personal information as explained in subsection 3.2. The INFORM message sent by the facilitator agent contains the information "Prepare Categorization".

*FacilitatorAgent: Send message to Information Agent for data availability*
*DatabaseAgent: received the following message :*
*(INFORM*
*:sender ( agent-identifier :name FacilitatorAgent@CAOTTN04108:1099/JADE :addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))*
*:receiver (set ( agent-identifier :name InformationAgent@CAOTTN04108:1099/JADE ) )*
*:content "Prepare Categorization" )*

Figure 9 shows a snapshot of the user interface which the consumers use to assign weights to their private data categories based on the service provider's privacy credentials and reputation score.

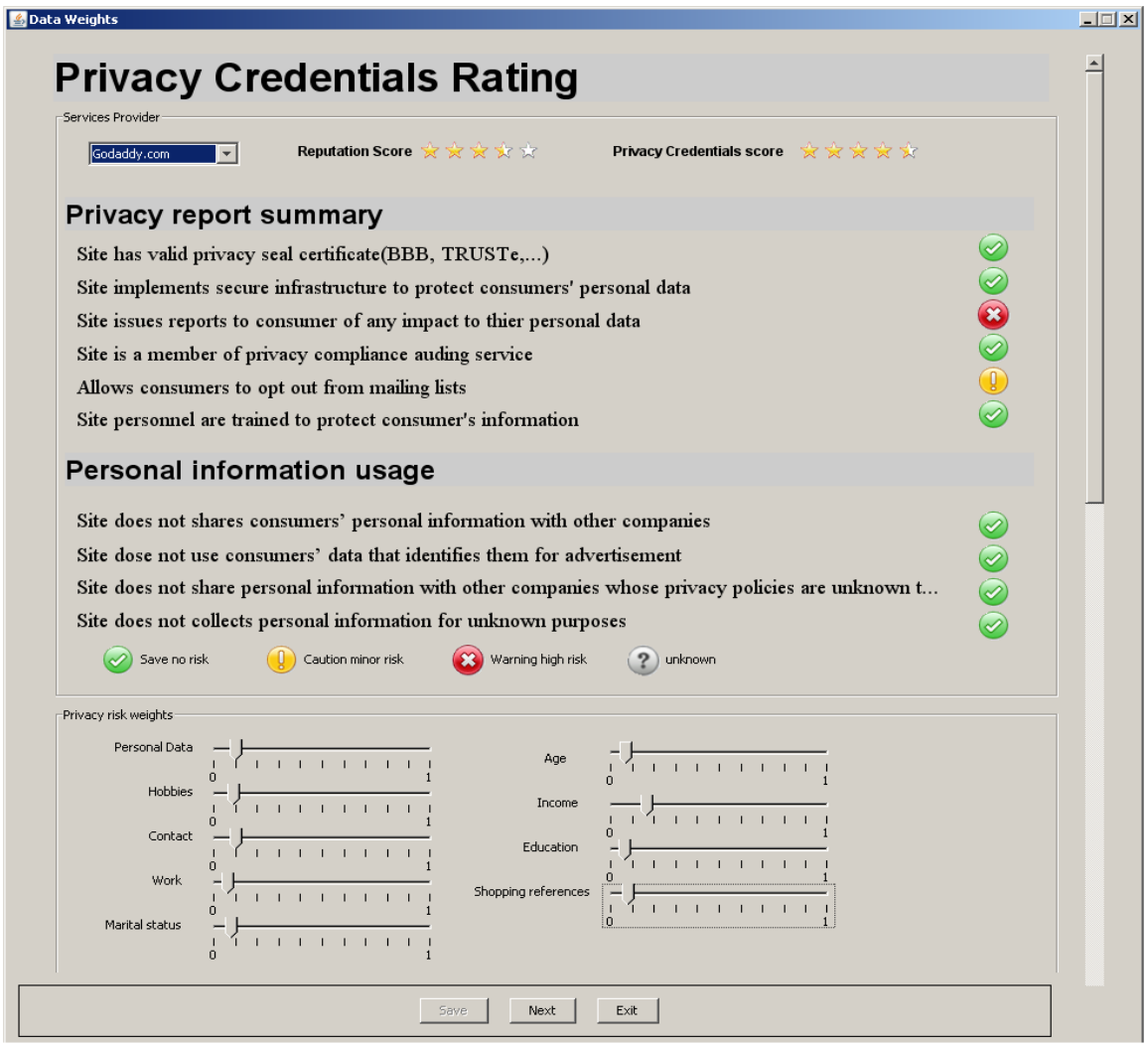


Figure 9: User interface for consumers to assign their privacy risk weights

In this example the trust and reputation agent prepares the privacy credential ratings of a service provider called godaddy.com. The privacy credential report emphasizes key items of privacy concerns that are likely to be most interesting to users; for example, information about the provider' data-sharing practices and information about whether the provider allows opt out of data sharing and marketing solicitations. For the reputation score we collected reputation information from iVouch.com and Bizrate.com. These websites provide reputation services about online businesses based on customers' testimonials. The score provided in the report is the average of the reputation score provided by the commercial websites iVouch.com and Bizrate.com. The user uses the credential report to learn about the service provider's privacy practices and then to assign the level of privacy risk to his personal data categories based on the perceived privacy risk. The system prompts (with an error exception) the user to revise his selection if he/she makes an ill characterization such as exceeding the privcy risk sum of 1 for all

categories. Once the user assigns the privacy risk weights, the information agent quantifies the total privacy risk of the user personal information. The payoff agent uses the quantified information to calculate the payoff value for the consumer. The message received by the payoff agent is shown below.

```
FacilitatorAgent: Send message to Payoff Agent for payoff calculation
PayoffAgent: received the following message:
(INFORM
 :sender  ( agent-identifier :name FacilitatorAgent@CAOTTN04108:1099/JADE
:addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))
 :receiver  (set ( agent-identifier :name PayoffAgent@CAOTTN04108:1099/JADE ) )
 :content  "0.6428572" )
```

In the above message the privacy risk value calculated by the information agent is 0.657143 for the user account shown in Figure 8. The payoff agent uses this value to calculate the payoff that the consumer should receive. For this example, we assumed the risk-premium value to be $50, then the consumer should receive $32.142 as a payoff for revealing his personal information record.

After the payoff is calculated, the consumer agent (the agent that negotiates on behalf of consumers) receives a message to start the negotiation.

```
FacilitatorAgent: Send message to Consumer Agent to start negotiation...
ConsumerAgent: received the following message :
(INFORM
:sender  ( agent-identifier :name FacilitatorAgent@CAOTTN04108:1099/JADE
:addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))
:receiver  (set ( agent-identifier :name ConsumerAgent@CAOTTN04108:1099/JADE
) )
:content  "Start Negotiation")
```

In this setup, we have created two agents namely the consumer agent (working on behalf of consumers) and the provider agent (working on behalf of the online service provider). Table 4 shows the parameters we used to setup the negotiation process. We have simulated 2000 user accounts stored in the database. To capture different types of users (that is, privacy pragmatic, privacy fundamentalists, etc.), we randomly assigned a range of privacy risk levels (between 0 and 1) for each record.

In this example, when the provider agent reaches its deadline and concedes to its reservation price, the agent sends the offer in two separate messages so the consumer agent knows that this is the last offer. This scenario is one of different possible scenarios related to the setup of the reservation price and the time deadline, but we are confident that it is sufficient to show the effectiveness of the negotiation strategy.

Table 4: Negotiation parameters for both the consumer agent and the provider agent

| **Consumer Agent** | **Provider Agent** |
|---|---|
| Reservation price = 45 | Utility = 70 |
| Time Deadline = 10 | Reservation price = 35 |
| Number of consumer record = 2000 | Time deadline = 6 |
| | Bid Increment = 3 |

The trace of the messages exchange between the two agents is captured via the sniffer agent provided by JADE as shown in Figure 10. The consumer agent sends a call for proposal (CFP) message to the provider agent. The provider agent sends its proposal to the consumer agent. Each time the consumer agent receive an offer it calculates the number of records $N'$ (as explained in subsection 3.5.2) and send it to the provider agent in a propose message. In this scenario, the exchange of messages continues until the provider agent reaches its deadline and concedes to the reservation price.

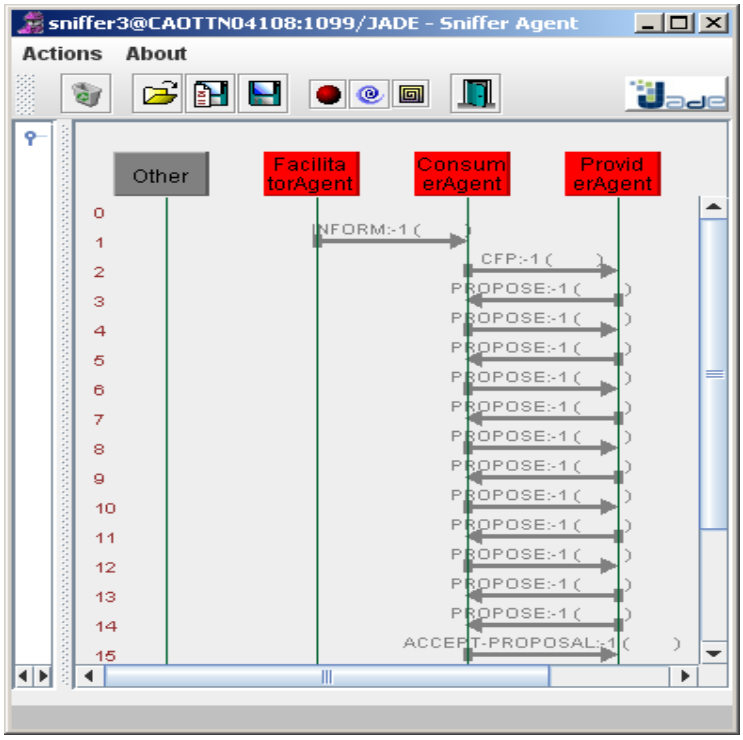


Figure 10: Traces of per record negotiation between the consumer agent and the provider agent

Below we show the exchange of offers in round 4 (a negotiation round corresponds to the process of sending a propose message and receiving a reply) as well as the final accept message.

**Exchange of messages in round 4:**

*(PROPOSE*
*:sender ( agent-identifier :name ProviderAgent@CAOTTN04108:1099/JADE :addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))*
*:receiver (set ( agent-identifier :name ConsumerAgent@CAOTTN04108:1099/JADE ) )*
*:content "inbidPR:16.0" )*

*(PROPOSE*
*:sender ( agent-identifier :name ConsumerAgent@CAOTTN04108:1099/JADE :addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))*
*:receiver (set ( agent-identifier :name ProviderAgent@CAOTTN04108:1099/JADE ) )*
*:content "inbidCA:371")*

***ACCEPT-PROPOSAL* message:**
*Provider Agent: received the following message:*
*(ACCEPT-PROPOSAL*
*:sender ( agent-identifier :name ConsumerAgent@CAOTTN04108:1099/JADE :addresses (sequence http://CAOTTN04108.ad3.ad.alcatel.com:7778/acc ))*
*:receiver (set ( agent-identifier :name ProviderAgent@CAOTTN04108:1099/JADE ) )*
*:content "inbidCA:1538" )*

The propose message sent by the provider agent contains the provider's offer of $16 per record (shown in the content section of the message). The replied propose message sent by the consumer agent contains an offer equal to 371 consumers' records (shown in the content message). The exchange of message continues until the provider agent sends its final offer and the consumer agent accepts it. The content of the accept proposal message include 1538 records of consumers that potentially will accept to share their personal information for the agreed upon payoff.

## 4.1 Results

Figure 11 depicts the values of the exchanged offers between the two agents. Each time the provider agent increases its offer, the number of consumers who are willing to share their personal information increases. This is because individuals valuate their perceived privacy risks differently. While some people are willing to accept lower offers to reveal their personal information (this type of individuals according to Spiekermann (2001) are marginally not concerned), others expect offers that are worth the risk of exposure (e.g. privacy pragmatics and privacy fundamentalists).

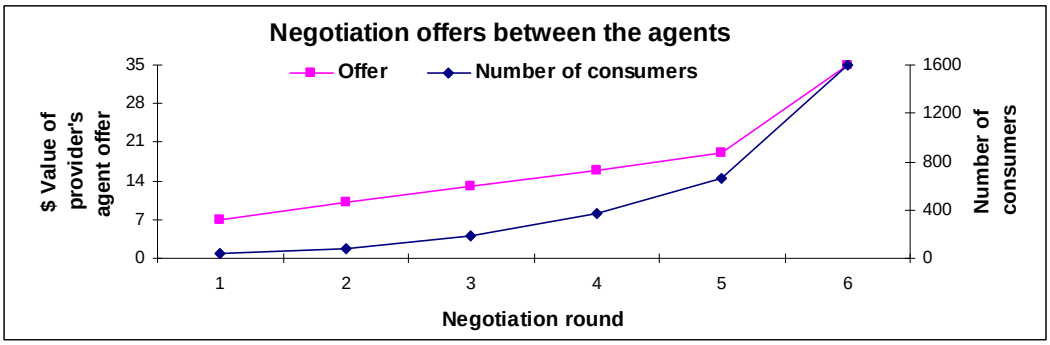


Figure 11: Negotiation offers between the consumer agent and the provider agent

In this scenario, the final payoff value is \$35 for each record which means that the total gain of the consumers' community is 1538*\$35 = \$53830. But, as mentioned above, consumers valuate their personal information differently, therefore the distribution of the surplus resulted from the difference in valuation should reflect the contribution of each consumer in the community (Bonnevay et al. 2005).

Let $G_i = \{g_1,...,g_x\}$ be the gain vector of all consumers according to their initial valuation of their personal information. Let $C$ be the consumers' community in this deal, then the gain surplus $S(C)$ of the community $C$ is:

$$S(C) = v(C) - \sum_{j \in C} g_j \qquad (19)$$

The surplus of $C$ is the difference between the community gain $v$(C) and the sum of the actual gain of each consumer in $C$ according to their initial valuation of their personal information. $S(C)$ is the value that can be divided among consumers of community $C$. The community gain $v$(C) = \$53830. The total initial value gain for each consumer is \$28019.99. Then according to (13) the surplus gain is \$25810.01. Following Bonnevay et al. (2005), we distribute the surplus among consumers according to the contribution of each consumer's valuation. For example if two consumers valuated their personal information to be $g_1$=$g_2$=\$25, then both should receive the same surplus gain. However, if the valuation is $g_1$= \$19 and $g_2$=\$25, then each one of them should obtain a surplus value reflect the percentage of contribution in the community which is in this case \$17.50 and \$23.03 respectively. Figure 12 depicts the surplus distribution for each consumer in the community and Figure 13 depicts the total gain of each consumer.

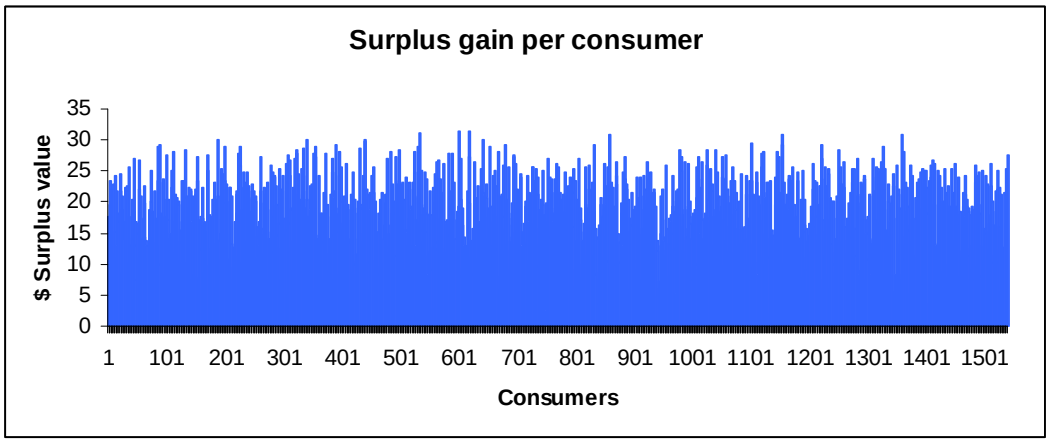


Figure 12: Surplus distribution among consumer according to the percentage of contribution

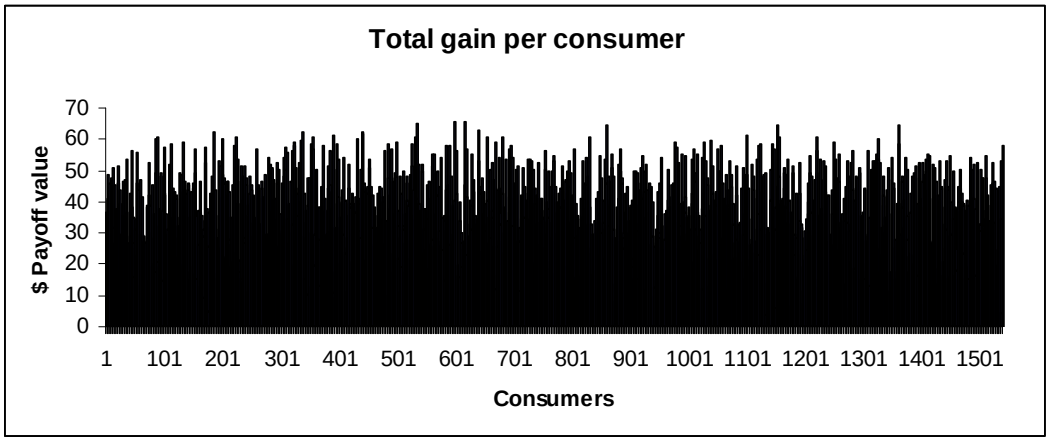


Figure 13: The total payoff distribution among consumer including the surplus gain

# 5 Discussion

The goal of this paper is the development of a knowledge-empowered agent information system architecture that allows users to participate in the information market and benefit from sharing their personal information. This novel architecture is based on a detailed analysis of existing systems and the privacy challenges that consumers are facing in the emerging information market. One of the main contributions is the proposal of a new technique to categorize and valuate privacy risks derived from users' private data set based on attribute ontology—namely, the semantic equivalence and the substitution rate of personal information which was used in the categorization process. This new technique of personal information categorization takes into account the granularity of private data which is achieved by associating the contextual usage risk to the personal data objects. Our system allows users to assign these contextual privacy risks to their private data so that the software agent can quantify the privacy costs for each user. The

assignment of these credit values is based on the users' perceived risk, trust, and the level of protection. As such, this paper contributed to the process of making privacy risk a measuring principle for the quantification of the privacy payoff value. This was one of the main design requirements of the proposed information system; specifically, determining reliable decision criteria for sharing personal information based on trust, risk, and protection level of the private data. Also essential was determining the risk value of a given private data object by means of credit assigning and weighting concept given the risk of potential damages resulting from the misuse of personal information.

One of the challenges that users face in the online world is how service providers use their personal information. The paper advances the state of the art by proposing a new assessment scheme to determine the trustworthiness of service providers based on privacy credentials. This is achieved through the design and development of a new mechanism that allows consumers to see "good" signs and "bad" signs when visiting online stores. The trust and reputation agent, who is responsible of making such knowledge available to consumers, provides a privacy credential score and a report detailing the competency of the service provider with respect to privacy and private data handling. The score is designed based on a fuzzy logic technique that takes into consideration the significance of each trust attribute that the service provider has obtained. The rating report also includes privacy-related items that concern most consumers. The importance of such assessment is that it serves as a source of knowledge that allows consumers to learn how their personal information will be handled by a potential service provider.

Yet another contribution is the development of a model by which we can valuate the payoff for the consumers against the revelation of their personal information. The reward amount which the consumer should receive must be based on the perceived privacy risk of each data object as well as the risk resulting from the composition of each data object. In this paper, we were the first to use a financial model to construct the amount of the payoff based on the risk exposure. This model is a well-known one in the financial and insurance industry. Financial models (such as the one we used) are designed so that those who are considered more risky should pay more interest. Since the value of the private data is associated with the amount of privacy risk, financial models were very helpful to determine the fair value that users should obtain. By so doing, we were able to valuate the payoff value that the consumers should receive against the revelation of personal information.

Concerning the negotiation with online service providers, this process is proven to be a daunting prospect for an individual consumer as explained in subsection 3.5. The extension of the "sit-and-wait" strategy allows us to empower the negotiation agent with a dominant strategy for

an agent using time-dependent negotiation process. This is because personal information does not devaluate over time. That is, even if the offer happens only at some future stage of negotiation, as long as it is anticipated, the consumers' agent who makes this offer has bargaining power. Another important aspect of bargaining power arises from impatience. In order to achieve certain competitive advantages, online businesses have particular service delivery threshold that makes them more impatient. Therefore, they are more likely to make bigger concessions in order to seal the deal as soon as possible.

In the proof of concept implementation that inspired by the scenario presented in subsection 3.1, the validity of the system made clear that the proposed system is feasible. Furthermore, we have shown through an experiment that a strategically placed agent can help consumers attain the maximum gain during the negotiation process.

While some people would argue that such system could lead to an intensified disclosure of personal data, it is unlikely to happen for the following reasons: first, the payoff or the compensation given to the consumers is for personal information that they usually provide for free anyway while online businesses rack up lucrative amounts of money from personal information profiling. At best, consumers sometimes exchange a certain degree of privacy (i.e., providing some of their personal information) for small deals such as price discounts, customized offers, specials, etc. (Hui et al. 2006, Acquisti 2004, Chellappa and Sin 2005, Harn et al. 2002, Grover and Teng 2001, Hann et al. 2003).

Second, the payoffs are risk-base premiums. Therefore, online businesses are expected to be more conservative when handling consumers' personal information as privacy risk penalties (risk-base premiums) and reputation consequences on violators of consumers presumed privacy rights are more likely to be costly.

Third, since attention in the information market within eCommerce is one of the main reasons, if not the most important reason, behind the collection of personal information, service providers have a financial stake in seeking ways for accurate information. As profiles become more accurate by participation, targeted services can be delivered to the right consumers, thus achieving a particular delivery threshold and minimizing opportunity costs. In other words, service providers essentially incur excess costs (junk e-mails, junk advertisements, etc.) to attract consumers' attention to their products and services. However, if they knew precisely what the consumer wanted, they could make a much better decision about whether or not to provide the consumer with information about their services.

Fourth, from a purely practical perspective, a negative reaction could cause consumers to turn away from the service and the service provider, thus counteracting any marketing

improvements to the service. In that regard, knowledge of consumer and societal perceptions of privacy invasion are as important as knowledge of the consumer's behavior and habits. With that knowledge, any measure taken by service providers to compensate consumers for their information dissemination in an appropriate manner will potentially be regarded by consumers and maximize their loyalty.

# 6 Conclusion and future work

This paper reports on a knowledge-empowered agent information system for consumers' privacy payoff in eCommerce. The key ideas and the overall system architecture are described, and a proof of concept implementation is presented. We plan to extend our work in various ways:

The first direction is to consider the effect of future learning in the collected personal information. The model presented in this paper assumes that the possession of information today does not influence the possession of information in the future, and therefore it assumes a linear view of the information valuation with respect to the risk at the time of revealing the personal information. However, in reality a future learning effect might impose different usage of personal information at different time intervals and therefore a different level of risk at each time. This means that when the agent computes compensation prices should consider a learning premium in the belief that what personal information the consumer reveal today will allow online businesses to learn more about this same consumer in the future. In this case, when we generalized the concept of compensation prices of risk to private data, the linear correlation assumption becomes inadequate as the distribution of risk becomes non-normal.

The second direction is to engage real scenarios with actual users, this is because empirical studies, such as those of Taylor (2003) and Huberman et al. (2005), to name few, of the value consumers assign to privacy risk have highlighted a dichotomy between professed attitudes and actual behavior, raising questions about individuals' awareness of privacy trade-offs and their true valuation of privacy.

# References


Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. ACM EC'04, New York, USA, 17–20 May

Acquisti, A. and Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Security and Privacy, Vol. 3, pp.26–33

Aknine S., Pinson S., Shakun M. (2004). An Extended Multi-Agent Negotiation Protocol. Autonomous Agents and Multi-Agent Systems, 8, 5– 45

Albert M., Jonker C. M., Karami M., Treur J. Agent models and different user ontologies for an electronic market place. Knowledge and Information systems, 2004, Volume 6, Number 1, 1-41

Bagnal A., Toft I. (2006) .Automated adaptive Agents for single seller sealed bid auction. Autonomous Agents and Multi-Agent Systems, 12, 259–292

Bakshi, G., Carr, P. and Wu, L. (2006). Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. Journal of Financial Economics, Elsevier, Vol. 87, No. 1, pp.132–156

Bergelson, V. (2003). It's Personal But is it Mine? Toward Property Rights in Personal Information. UC Davis Law Review, Vol. 37, No. 379, . Available at SSRN: http://ssrn.com/abstract=870070

Blažič, A.J., Dolinar, K., Porekar, J. (2007). Enabling privacy in pervasive computing using fusion of privacy negotiation, identity management and trust management techniques. Proceedings of the First International Conference on the Digital Society (ICDS)

Bo An, Sim K.S., Tang L.G., Mia C.Y., Shen Z.Q., Cheng D.J. (2008). Negotiation agents' decision making using Markov Chains. Studies in Computational Intelligence (SCI) Springer 89, 3-23,

Bonnevay S., Kabachi N., Lamure M. (2005). Agent-based simulation of coalition formulation in cooperative games. Proceeding of the IEEE/WIC/ACM International conference on Intelligent Agent Technology IAT

Cao L., He, T. (2009). Developing actionable trading agents. Knowledge and Information systems 2009, 18:183-198

Cavoukian, A., (2009). Privacy as a negative externality: The solution Privacy by Design. WEIS Workshop on the Economics of Information Security in London, U.K

Chellappa, R. and Sin, R.G. (2005).Personalization versus Privacy: An Empirical examination of the Online Consumer's Dilemma. Springer, Vol. 6, Nov. 2–3

Cranor F.L., Arjula M., and Guduru P. (2002). Use of P3P User Agent by Early Adopters. Proceedings of the ACM Workshop on Privacy in the electronic society

Cvrcek, D., Kumpost, M., Matyas, V. and Danezis, G. (2006). A study on the value of location privacy. WPES Proceedings of the 5th ACM Workshop on Privacy in Electronic Society, ACM Press, pp.109–118

Danthine, J.P. and Donalaon, I.B. (2002). Intermediate Financial Theory. New Jersey: Pearson Education 1st edition.

Desmarais, C., Shen, X., Shirmohammadi, S., Cameron, A., Goerganas, N.D. and Kerr, I. (2007). PLUTO-a privacy control protocol for e-commerce communities. Proceeding IEEE Conference on Enterprise Computing, E-Commerce and E-Services, Tokyo, Japan

Dighton, J.A. (2003). Market solutions to privacy problems?. Chap. 6 in Digital Anonymity and the Law – Tensions and Dimensions, Harvard Business School, The Hague: T.M.C. Asser Press

FIPA specification http://www.fipa.org/

Laudon, K.C. (1996) .Markets and Privacy. Communications of the ACM, Vol.39, Issue 9, pages 92-104, NY, USA

Lioudakis, G.L, Koutsoloukas, E.A, Delaas, N.l., Tselikas, N., Kapellaki, S., Preaerakos, G.N., Kaklamani, D.I., Venieris,I.S. (2007). A middleware architecture for privacy protection. Elsevier computer networks 51 4679-4696

Gopal, R., Garfinkel, R., Nunez, M. and Rice, D. (2006). Electronic markets and private information: electronic and security considerations. IEEE Proceeding of the 39th International Conference on System Sciences

Grover, V. and Teng, J. T.C. (2001). E-commerce and the information market. Communication of the ACM, Vol. 44, No. 4

Hann, IL-Horn, Hui, K.L., Lee, T.S., Png, I.P.L. (2003).The Value of Online Information Privacy: An Empirical Investigation. AEI-Brookings joint center for regulatory studies

Harn, I., Lee, T. S., and Png, I.P.L. (2002) .Online information privacy: Measuring the cost-benefit trade-off. 23rd International Conference on Information Systems

He, Y. and Jutla, D.N., (2006). Contextual e-negotiation for handling of private data in e-commerce semantic. Web Proceeding of the 39th Hawaii International Conference on System Sciences, HICSS'06, Vol. 3, pp.62a–62a.

Hui, K.L., Bernard C.Y.Tan, and Goh, C.Y. (2006). Online Information Disclosure: Motivators and Measurements. ACM Transactions on Internet Technology, Vol. 6, No. 4, November 2006, PP. 415–441

Huberman, A.B., Ader, E., and Fine L.R. (2005). Valuating Privacy. IEEE Security and Privacy, Vol. 3, no.5, Sep-Oct. 2005, pp. 22-25

Krause, A. and Horvitz, E. (2008) . A Utility-Theoretic Approach to Privacy and Personalization. Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence

Lioudakis, G.L, Koutsoloukas, E.A, Delaas, N.l., Tselikas, N., Kapellaki, S., Preaerakos, G.N., Kaklamani, D.I., Venieris,I.S. (2007). A middleware architecture for privacy protection. Elsevier computer networks 51 4679-4696

O'Hara K., Tuffield M.M., Shadbolt N. (2008). Lifelogging: Privacy and empowerment with memories for life. Identity and information society IDIS 1:155-172

Palen, L. and P. Dourish, (2003). Unpacking privacy for a Networked World. CHILetters, 5(1): pp. 129-136.http://guir.berkeley.edu/projects/denim/denim-chi-2000.pdf

Piolle G., Demazeau Y., and Caelen J., (2007). Privacy Management in User-centred Multi-agent Systems. ESAW 2006, Springer LNAI 4457, pp. 354-367.

Preibush, S. (2005). Implementing Privacy Negotiation Techniques in E-Commerce. Proceeding of the seventh IEEE International Conference on E-Commerce Technology

Prins, C. (2006). When Personal Data, Behavior and Virtual Identities Become a Commodity: Would a Property Rights Approach Matter?. SCRIPT-ed, Vol. 3, No. 4

Ribaric S., Hrkac T. (2008). TeMS-a multi-agent system for temporally rich domains. Knowledge and Information systems 2008, 15:1-30

Ross, S. (2004) .Stochastic Processes. John Wiley and Sons, Inc., Chap. 8, 2nd ed.pp.366–381, 1995

Schwartz, P.M. (2004). Property, privacy, and personal data. Harvard Law Review, Vol. 117, pp.2056–2127

Sheizaf R. and Daphne R. R. (2005). Information sharing online: a research challenge. Int. J. Knowledge Learning 1(1/2): 62-79

Skylogiannis T., Anotniou G., Bassiliades N. Governatori G., Bikakis A. (2007). DR-NEGOTIATE – A system for automated agent negotiation with defeasible logic-based strategies. Data and Knowledge Engineering Vo 63, Issue 2, Pages 362-380

Sensoy M., Yulom P., (2009). Evolving service semantics cooperatively: a consumer-driven approach. Autonomous Agents and Multi-Agent Sytems 18:526–555

Spiekermann, S. (2001). Online Information search with electronic agents: drivers, Impediments, and Privacy Issues. http://edoc.hu-berlin.de/dissertationen/spiekermann-sarah-2001-11-22/PDF/Spiekermann.pdf

Sierra C. (2004). Agent-Mediated Electronic Commerce. Autonomous Agents and Multi-Agent Sytems, 9, 285–301

Syverson, P. (2003). The paradoxical value of privacy. 2nd Annual Workshop on Economics and Information Security (WEIS 2003)

Su S.Y.W., Huang C. and Hammer J. (2000). A Replicable Web-based Negotiation Server for E-Commerce. In Proc. 33rd Hawaii International Conference on System Sciences.

Tang, Z., Hu, J. and Smith, M.D. (2008). Protecting online privacy: self-regulation, mandatory standards, or caveat emptor. Journal of Management Information Systems, Vol. 24, No. 4, pp.153–173.

Taylor, C.R. (2004). Consumer privacy and the market for customer information. RAND Journal of Economics, Winter, Vol. 35, No. 4, pp.631–650

Taylor H. (2003). Most people are privacy pragmatists, who, while concerned about privacy, will sometimes trade it off for other benefits. The Harris Poll, 17

Voulodimos, A.S., Patrikakis, C.Z. (2009). Quanitfying privacy in terms of entropy for context aware services. Identity and Information Society, Springer, 0026-2, 2009

Yassine A., Shirmohammadi S. (2009). An Intelligent Agent-Based Framework for Privacy Payoff Negotiation in Virtual Environments. Proc. IEEE Workshop on Computational Intelligence in Virtual Environments, in Proc. IEEE Symposium Series on Computational Intelligence, Nashville, TN, USA

Yassine A., Shirmohammadi S. (2008). A business privacy model for virtual communities. International journal of web based communities, Vol. 5, No. 2

Yassine A., Shirmohammadi S. (2009). Measuring Users' Privacy Payoff Using Intelligent Agents *IEEE computational Intelligence Society, International Conference on Computational Intelligence for Measurements Systems and Applications*, CIMSA09

Yassine A., Shirmohammadi S. (2008). Privacy and the Market for Private Data: A Negotiation Model to Capitalize on Private Data, *Proc. ACS/IEEE International Conference on Computer Systems and Applications*, Doha, Qatar, March 31-April 4 2008, pp. 669–678

Yu, T., Zhang, Y. and Lin, K.J. (2006). Modeling and measuring privacy risks in QoS web services. Proceeding of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, Vol., No., pp.4–4.

Yee, G. and Korba, L. (2004). Privacy Policies and their Negotiation in Distance Education. National Research Council of Canada NRC 46555

W3C (2005). The Platform for Privacy Preferences 1.1 (P3P1.1) Specification. W3C Working Draft 4 http://www.w3.org/TR/2005/WD-P3P11-20050104/

Xin W., Xiaojun S., Nicolas D. G. A Fuzzy Logic Based Intelligent Negotiation Agent (FINA) in E-commerce. IEEE CCECE/CCGEI, Ottawa, May 2006

Zhuang Y., Fong S., Shi M. (2008). Knowledge-empowered automated negotiation system for e-Commerce. Knowledge and Information systems 2008, 17:167-191