Privacy and the Market for Private Data: A Negotiation Model to Capitalize on Private Data

Abdulsalam Yassine, Shervin Shirmohammadi
DISCOVER Laboratory
School of Information Technology and Engineering
University of Ottawa, Canada
ayassine@discover.uottawa.ca, shervin@discover.uottawa.ca

Abstract

The market for consumer information is already a lively market, where consumer information and consumer profile data are often among the most valuable assets owned by online retailers. The value of such commodity derives from the ability of firms to identify consumers and charge them personalized prices [1]. We argue that if consumers’ identity and personal information is such a valuable asset, should not consumers benefit from their asset as well? In this paper, we propose a negotiation process between an online consumer agent and an online seller. The online consumer agent acts on behalf of consumers to maximize their social welfare. In our model, the agent derives a quantified privacy risk for each private data and uses it to determine a cost premium value to make the bargaining process manageable. We also provide a computational example to evaluate the model.

1. Introduction

Today, many companies and online retailers, not generally recognized by people, are gathering information and have assembled sophisticated databases that know a great deal about most people [2]. They then sell that information to other firms and organizations that manipulate or add to the data for yet other purposes. Electronic retailers, for example, need not to depend on their own consumer-profile data. The market for consumer information is a lively market, where consumer information is a commodity for trade among web-based marketing firms [1, 2]. These firms collect and sell customers’ data that typically include an individual’s: name, purchasing history, income, size of family, hobbies, lifestyle, etc. In such industry, the tiniest nuggets of personal information have value. The size of such market in the United States alone has reached to $1.93 trillion in revenue in 2006 [8]. As Randleman [2] puts it “Businesses are buying and selling customer data in a dizzying number of ways”.

On the other hand, there is a growing concern from privacy advocates over such market. This concern has reached its peak in recent years. Policy makers, government bodies, and legislators are not sparing any effort to regulate the use and the disclosure of confidential information so that consumers’ rights to privacy are protected. But there is another side to privacy: individuals make choices in which they surrender a certain degree of privacy in exchange for deals (such as price discount, money, better quality, customized offers, specials, etc.) that are perceived to be worth the risk of information disclosure[3,4,5,6,7]. In this respect, private information is an asset with a market price and opportunity cost. The difference between privacy as a right and private data as an asset is that the former draws its authority from laws and humanistic principles while the latter is a property which can be traded for a value [11].

The above distinction between privacy as a right and private data as an asset motivates our research objective: we argue that if consumers’ identity and personal information is a valuable asset, should not consumers benefit from their asset as well? The issue is not whether consumers value online privacy. It is obvious that people value online privacy [9]. The issue, about which prior research is lacking, is the value of private information in the presence of potential benefits from sharing personally identifiable information.

Today there is evidence [7, 10, and 11] to suggest that the market is evolving towards a new industry: agents to negotiate consumer information deals. Companies called “infomediaries” might act as brokers of customer information, marketing it to retailers on consumers’ behalf. The new industry will create
institutions that allow individuals to capture a share of their asset and that impose reputation consequences on violators of their presumed privacy rights.

This paper is investigating a model by which consumers can capitalize on the value of their personal information and get something of value in return. Our main focus is online transactions; consumer concerns are greater online because the collection of private data is a little less obvious. Generally, here is what we think is required: a model that quantifies private data valuation grounded in economic theories; and, of course a negotiation mechanism which allows consumers to be strategic about the revelation of their private information.

The approach followed in this paper is using an online consumer agent. The agent acts on behalf of consumers to maximize their welfare as a community. In our model, the agent derives a quantified privacy risk value for each private data then uses it in a risk-base premium formula to calculate the consumer’s payoff (cost premium). The derived privacy risk premium is context-dependent for each consumer: each consumer is assumed to be stochastically equivalent and has independent distributed valuation for each context. The agent uses a game theoretic negotiation process to negotiate the payoff with the online seller or the Web service provider.

The proposed privacy risk premium is a quantification model to determine the value of the consumers’ payoff. Risk-base premium is the universal practice of the financial institutions and the insurance industry, and is implied in the stock market and in many other industries. The basic idea is that those who are considered more risky should pay more interest. This means that risk-base premium is a quantified value applied to compensate for unwanted events that might lead to revenue loss or increased cost. The same analogy can be applied to privacy. Privacy risk is the business risk resulting from the collection, use, retention and disclosure of personal information. Like all business risk, privacy risk could result in a loss of revenue.

The rest of the paper is organized as follows: in the next Section we state the intended contribution. In Section 3, we present the related work. In Section 4, we provide a motivation example. In Section 5, the model of privacy risk quantification is discussed with an example, and we show the consumer’s payoff value. We then, in Section 6 present the negotiation process as a game theoretic formulation between the agent and the online seller. In Section 7, we analyze a valuation experiment. Finally, in Section 8, we provide conclusions of our paper, and discuss plans for future work.

2. Contribution of this paper

This paper touches on a relatively untapped subject. It illustrates a novel approach to capitalize on consumers’ private data asset. The intended contribution of this paper is as follow:

- Propose a model that allows consumers to capture a share of their private data asset based on privacy risk valuation
- Employ risk-base premium, a widely accepted concept, to determine consumers’ payoff
- Employ a game theoretic negotiation process that allows the agent to be strategic during the negotiation process

Rigorous quantification evaluation of the proposed approach is performed and the preliminary experimental results are reported in this paper.

To our best knowledge, this approach has not been considered in the open scientific literature by anybody yet.

3. Related work

The previous work mostly focuses on models to protect consumers’ privacy and give them control over the use of personal information on Web sites they visit. The Platform of Privacy Preference P3P [15] developed by the World Wide Web Consortium is emerging as the standard protocol which gives consumers control over their personal information. It is also considered a fundamental model for tracking privacy concerns. Other work like [16] has developed a privacy control protocol, PLUTO, which extend the P3P specification and allows companies to analyze consumer data in a privacy compliance way by considering privacy legislation, privacy specification and data re-identification problem.

Parallel to the development of privacy related technologies, privacy negotiation has been studied in various aspects. In [18] context base negotiation model for the handling of private information is presented. The presented work depended on an OWL ontological representation of the P3P XML-based data schema for the privacy domain to provide relevant substitute data for counteroffers within a negotiation session.

All projects mentioned above focus on the definition of privacy policies and their enforcement to prevent any unauthorized access on user’s private data.

The first work that studied the market for consumer information was presented by Taylor [1]. Taylor assumed that there are two market regimes for consumer information, the anonymity regime in which the trade of consumer information does not exist, and the recognition regime in which consumer information...
is traded among firms. Taylor studied the effect of consumer knowledge of the trade on the firm engagement decision in dynamic prices.

In [17] a framework for negotiating privacy versus personalization is presented. During the negotiation process the service provider starts with a basic offer, consisting of a small discount and few personal data to be asked. The framework proposed two extensions to the P3P protocols to facilitate the negotiation process. The framework does consider the value of private data in question and therefore the service provider has the upper hand in the negotiation process.

In [19] Privacy risk modeling and measuring is presented based on private data categorization. The privacy risk values are then used in QoS selection algorithm for Web services.

In this paper, we adopt the data categorization presented in [19] to derive the privacy risk value. We then use it in a risk-base premium formula to calculate the consumer’s payoff. We consider a game theoretic negotiation process that allows the agent to be strategic during the negotiation process. The simulation experiment presented in this paper showed an interesting result about the power of the market in controlling privacy risk. It showed that consumers and online sellers (Web service providers) are both better off if privacy risk is low.

In the next section, we start by a motivation example inspired by [1].

4. Motivation example

Consider that we have two online sellers 1 and 2, and \( N \) number of consumers. Each consumer is identified by a set of distinct attributes \( a_i \) which can be thought of as the means by which the consumer is recognized. Consider also that we have two purchasing periods \( t = 1, 2 \) in which consumers buy good \( b \). Let \( V_{it} \) be the consumer \( i \)'s willingness-to-pay for good \( t \). Seller 1 offers price \( P_{1t} \) for all consumers and technically \( P_{1t} \) can be seen by seller 2.

The consumer information list is formed based on the first purchasing period and it merges the consumer identification attributes with his purchasing decision \( D_{it} \approx \{0, 1\} \) for all \( i \in N \). If seller 2 buys the list from seller 1, then seller 2 can use this information to engage in personalized pricing. Seller 2 offers:

\[
P_{12} = \begin{cases} 
  p_{1}^2 & \text{for consumers who bought good} \\
  p_{0}^2 & \text{for consumers who did not buy good} 
\end{cases}
\]

From [1] we have:

\[
\text{Payoff}_{\text{consumer}} = (V_{11} - P_{1}) \cdot D_{11} + (V_{12} - P_{12}) \cdot D_{12}
\]

If seller 2 bought the list from seller 1 for a price \( \phi \), then the payoff to seller 1 is

\[
\text{Payoff}_{\text{seller1}} = N_1 \cdot P_{1} + \phi
\]

Where \( N_1 \cdot P_{1} \) is the payoff from selling good 1 to consumer mass \( N_1 \subseteq N \) for price \( P_{1} \)

The payoff for seller 2 is

\[
\text{Payoff}_{\text{seller2}} = N_1 \cdot P_{2}^1 + N_2 \cdot P_{2}^0 - \phi
\]

Where \( N_1 \cdot P_{2}^1 \) is the payoff from selling good 2 to consumer mass \( N_1 \) and \( N_2 \cdot P_{2}^0 \) is the payoff from selling good 2 to consumers mass \( N_2 \) such that \( N_1, N_2 \subseteq N \). Prices \( P_{2}^1 \) and \( P_{2}^0 \) are based on prior knowledge of consumer preferences found in the list.

Observations

- Seller 1 made an extra profit \( \phi \) from selling consumers information list to seller 2.
- Seller 2 incurred an extra cost \( \phi \) when purchased the consumer information list, but compensated the cost by adjusting the prices based on prior knowledge of consumers preferences.
- Consumers in the first period provided their private data for zero return and suffered price discrimination in the second period. In the first period, seller 1 built an asset at zero cost. In the second period, seller 2 invested in this asset for a return paid by the original owner of the asset, which is in our case the consumer.

From this example, we can see that a mechanism is needed to allow consumers to be fully strategic about the revelation of their information and to be far-sighted as sellers. The challenge is to determine private data objects that are perceived to be valuable to capitalize on them for the goal of maximizing their market value once the consumer decides to reveal them. In this paper, we propose the following methods to address the problem of valuating private data:

1. Define data sets taking into account the different overall sensitivity levels towards each private data and different importance one may assign to a specific data;
2. Define context-dependent weights to fit different situation as one consumer rule about any single private data item may not fit all situations;
3. Determine the expected return value of private data asset based on the quantified privacy risk of revealing the data. The higher the privacy risk, the higher the expected cost of risk;
5. The model

5.1. Assumptions

In part to set the model of the negotiation process between the seller and the agent, we briefly state the basic assumptions that apply to our work:
1. The agent knows the consumer’s private data specification set, but may not know the actual data;
2. The consumer assigns the release strategy of private data under different context;

5.2. Data classifications

The agent receives data specification from consumers and classifies them into \((M)\) different categories \((C_1...C_m)\), such as personal identification, contact information, address, hobbies, and salary. In every category, private data are further divided into subsets \((S_{ij}, i=1...M)\), each subset may contain one or more private data as shown in the example below.

Example of private data categorizations

<table>
<thead>
<tr>
<th>Category (contact)</th>
<th>Subset (telephone number), Internet address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset (telephone number)</td>
<td>Private Data (work phone number), Private Data (home phone number)</td>
</tr>
<tr>
<td>Subset (email)</td>
<td>Private Data (work email), Private Data (personal email)</td>
</tr>
</tbody>
</table>

The privacy sensitivity of each private data depends on parameters \((W_{ij}, \psi_{ij})\) given in the following definitions:

- Context privacy risk weight \(W_{ij}\): is the release strategy of private information category \(i\) under context \(j\). It is a value that quantifies the consumer sensitivity to private data protection under a specific context.

- Privacy risk \(\psi_{ij}\): is the overall weighted privacy risk value of revealing private data of category \(i\) under context \(j\).

The presented data classification has three interesting characteristics in the context of profile data subsets that are related to each other: First, data in different categories may have different context-dependent risk weights. Since the value of the private data may differ from one context to another also 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 is one of the private data has been revealed, revealing the rest of the data in the same subset will not increase the disclosure risk of privacy. For instance, user’s age information can be expressed by age, year of birth or high school graduation year. Knowing all of them at the same time allows only marginal improvements. This will allow us to consider each private data subset as a one unit. 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 his email address constitute two possible ways to contact the consumer but they are not completely interchangeable.

5.3. Risk quantification

Let \(T_i\) be the private data size: i.e. the number of subsets in category \(i\). If \(f_j\) subsets have been revealed, then the weighted privacy risk of revealing private data \(f_j\) normalized over the number of subsets in category \(i\) is calculated as follows:

\[
\psi_{ij} = \frac{f_j}{T_i} \cdot W_{ij}
\]

(5.3.1)

Where the following properties apply:
- \(W_{ij}\) is linear in the interval \([0,1]\);
- The sum of context privacy risk weights for all categories under each context is 1.

To put the privacy risk of each category in the range of \([0, 1]\), the privacy risk in equation (5.3.1) is further normalized over the privacy risk value of each category under context \(j\); the example below illustrates the privacy risk calculation.

Example: Consider that a consumer has the following private data categorization depicted in figure 1. Category \((C_1)\) has two subsets \((S_{11})\) and \((S_{12})\). Subset \((S_{11})\) has private data \((d_1)\) and \((d_2)\), while subset \((S_{12})\) has private data \((d_3)\) and \((d_4)\). Category \((C_2)\) has two subsets \((S_{21})\) and \((S_{22})\), subset \((S_{21})\) has two private data \((d_5)\) and \((d_6)\), while subset \((S_{22})\) has private data \((d_7)\). Category \((C_3)\) has one subset \((S_{31})\) with private data \((d_8)\).

Under context \(j\), the context-dependent weight \(W_{ij}\), which reflects the release strategy of private information in category \(i\), is given in table 1.

| Table 1: Context-dependent weight for categories 1, 2, and 3 |
|-----------------|-------|-------|-------|
| Context \(j\)   | \(W_{1j} = 0.4\) | \(W_{2j} = 0.3\) | \(W_{3j} = 0.3\) |
Assume that the consumer has revealed the following private information \((d_1), (d_3), (d_6),\) and \((d_8);\) his overall privacy risk value is calculated in table 2 below.

In real world, the value of \(\psi\) is an indication of how much the consumer values the requested private data.

\[
\psi = \frac{\sum_{i=1}^{n} \psi_{ij} \cdot W_{ij}}{\sum_{j=1}^{m} \psi_{ij}}
\]

\[
\psi_{norm} = \frac{\psi_{ij}}{\sum_{j=1}^{m} \psi_{nj}}
\]

\[
\psi_{1j} = 0.4 \quad \psi_{2j} = 0.15 \quad \psi_{3j} = 0.3
\]

\[
\psi_{norm} = 0.47 \quad \psi_{norm} = 0.17 \quad \psi_{norm} = 0.35
\]

\[
\psi = 0.99
\]

\[\text{Expected Return} = \text{RiskfreeRate} + \sum_{z=1}^{k} \text{Risk}_z \cdot \text{Risk Premium}_z\]

Where \(z\) is the risk factor

5.4. Consumers’ payoff

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. In the case of private data assets, risk premium is evaluated in response to the amount of potential damages that might occur (cost of risk) with respect to the risk exposure. In this manner, the premium paid to the consumer is justified, at least in part, from the damages that might occur. Risk-base premium is a widely used concept in the insurance and financial industry. The standard model [12] is:

\[\text{Expected Return} = \text{RiskfreeRate} + \sum_{z=1}^{k} \text{Risk}_z \cdot \text{Risk Premium}_z\]

Assuming that the RiskfreeRate is zero (this assumption is valid since consumers do not expect to capitalize on 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 risk premium(s) for the factor or factors in the model.

One important aspect of private data compared to other assets such as financial assets, is the fact that they are subject to the context in which they are being used. This means that private data are subject to a single-factor generation process [12].

Under the single-factor return process, the above model can be written as follows:

\[\text{Expected Return} = \text{Risk} \cdot \text{Risk Premium}\]

As far as the risk premium is concerned, we would like to know what consumers, on average, demand as return for revealing their private data. In practice, risk premiums are estimated by looking at historical premiums over long time periods [12]. According to
The general behavior of risk premiums follows a random variable, random walk, or autoregressive process over time. In this paper, we assume that the risk premium follows a random process 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. Formally, the expected payoff will be written as follows:

\[ E \left[ \psi \right] = \psi \cdot (E[\mathcal{R}]) \quad (5.4.1) \]

Where \( E[\cdot] \) is the expectation operator, \( \psi \) is the private data asset in question, and \( \psi \) is the overall risk of revealing private data asset \( q \). A stochastic (random) process \( \mathcal{R} \) is said to follow a geometric Brownian motion if it satisfies the following stochastic differential equation:

\[ d\mathcal{R} = \mu \mathcal{R} \, dt + \sigma \mathcal{R} \, dW, \quad (5.4.2) \]

Where \( \{W_t\} \) is a Wiener process or Brownian motion, \( \mu \) is the drift, and \( \sigma \) is the standard deviation. Using Ito’s lemma [22], we can derive the analytical solution of equation (5.4.2) as follows:

\[ d \ln \mathcal{R} = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW(t) \]

The integral of (5.4.3) from 0 to \( t \) is:

\[ \int_0^t d \ln \mathcal{R} = \int_0^t \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma W(t) \]

Basic calculus gives us:

\[ \mathcal{R}_t = \mathcal{R}_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W(t)} \quad (5.4.4) \]

We now take the expectation \( E[\mathcal{R}] \) for (5.4.4):

\[ E[\mathcal{R}_t] = E[\mathcal{R}_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W(t)}] \quad (5.4.5) \]

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

\[ E[\mathcal{R}_t] = \mathcal{R}_0 e^{\mu t} \quad (5.4.6) \]

The risk premium is assumed to be known at time 0.

The derived value of the risk premium gives consumers a fair indication of the payoff they should get once decided to reveal their private information.

In the next section, we show a game theoretic negotiation process between the agent and the online seller. We use the example presented in section 4 “Motivation example” to be our running example.

6. Agent negotiation

The negotiation process in this section is based on a game theoretic model. Game theory is a branch of applied mathematics that is often used in the context of economics [20]. Below we provide some preliminary definitions of game theory elements that apply to our work:

- **Players**: The players are the individuals who make decisions.
- **Strategy**: The strategy of a game is a rule that tells the player which action to choose at each instance of the game, given his information set.
- **Player’s Payoff**: The expected utility the player receives as a function of the strategies chosen by him and the other players.
- **Equilibrium**: Equilibrium is a strategy combination consisting of a best strategy for each player in the game.
- **Non-cooperative game**: A non-cooperative game is a game in which the players cannot make binding commitments.

Consider that the agent is working on behalf of consumers in the example presented in section 4, “Motivation example”. For simplicity and without loss of generality, we assume that consumers’ consumption of good1 is one unit per consumer. This means that the demand for good1 equals to the number of consumers. The agent objective is to negotiate a payoff paid to consumers in return for their private information revelation. The payoff can be equivalent to a price discount offered to the consumer once the negotiation process completed with acceptance.

In part of setting the negotiation game between the agent and the seller, we state the following rules:

1- The number of consumers \( N \) is known to the agent, but not to the seller. This means that the agent will use his knowledge of \( N \) as a bargaining power to force the seller for a better offer.

2- The price per record in the consumer information list is known to the seller but not to the agent. This means that the seller has an upper limit to the offer and beyond this limit the profit will be negative. The agent is unaware of this information otherwise the game will end from the first negotiation round;
3- Seller t is the monopoly seller of good t. This means that the seller is free to set his expected payoff. However, the seller’s limitation is: if the offered discount is low, then the demand (equal to the number of consumers) will decrease, and the size of the consumers information list will decrease;

4- The agent and the seller are two non-cooperative negotiators;

6.1. Seller’s strategy

Assume that the list will be sold to seller 2 for a price $P_r$ per consumer record. Then, if the list contains $N$ number of consumers the profit of seller 1 is $Z = N \cdot P_r$. However, if the seller offers a price discount as a payoff for consumers in return for their private information, then this discount will be shown as a cost incurred on the consumers’ list. Hence, seller 1’s profit will be $Z = N \cdot (P_r - C_r)$ where $C_r$ is the cost equal to the discount offered for each consumer. Theoretically, the maximum discount the seller can offer is equal to $P_r$ which yields a zero profit. However, in practice the upper limit of the discount offer is the value that yields minimum acceptable profit. The minimum profit is at which the difference between the price and the cost is relatively small, that is, $(P_r - C_r) = l$ where $l$ is a positive small value slightly greater than zero. Then, the seller strategy is gradually increasing the offered discount at a step rate equal to $\theta$ such that $0 < \theta \leq 1$.

6.2. Agent’s strategy

The agent objective is to maximize consumers’ payoff as a community that is, maximizing their social welfare $SW$. The agent, as mentioned in rule 1 of the game, has a prior knowledge of $N$. Therefore, his strategy at each round of the negotiation process is to reveal $N' \subseteq N$ given the offered payoff $U_q$ such that:

$$N' = \arg \max_{N' \subseteq N} \sum_{i \in N'} SW_i U_q$$  \hspace{1cm} (6.2.1)

The decision about the number of customers is taken based on an offer/demand curve, which is determined by the agent based on a pool of individual customers he represents.

6.3. Negotiation process

There are two phases in the negotiation process: The preparation phase and the bargaining phase. During the preparation phase, the seller and the agent work together to identify the required data set. During the bargaining phase, the seller and the agent negotiate the payoff value.

6.3.1. Preparation phase. Consider that in the preparation phase the seller would like to collect five private data such as (name, contact, hobbies, preference, and family size). Assume that the requested data is in the scope of consumer data specification. This eliminates any potential disagreement between the agent and the seller in this phase. Based on subsection 5.2, the agent calculates the overall privacy risk weights for each consumer. Then the payoff is determined as shown in section 5.4.

6.3.2. Bargaining phase. We assume that the payoff is equivalent to a price discount. The problem the negotiators are going to negotiate is: how many consumers are willing to provide the asked private information for the offered discount.

For this scenario, we have developed a high-level messaging protocol that supports an exchange of messages between the agent and the seller during the negotiation process. Before describing the protocol, we wish to present the steps of negotiation between the agent and the seller:

1. The agent sends a request to the seller asking for a price discount;
2. The seller first replies with a discount that maximizes his profit. The seller is assumed to be monopoly, but striving for the largest demand at the lowest cost. However, if the seller offers zero discount then no agreement is reached and both the seller and the agent walk away with zero payoff;
3. Based on the offered discount the agent replies with demand $D = N'$ number of consumers such that $0 \leq N' \leq N$ and asks the seller to go higher than the current offer.
4. The seller increases the discount offer by step $\theta$, and then step 3 is repeated.
5. The negotiation ends when either the seller reaches $(P_r - C_r) = l$ or when $D = N$ the maximum number of consumers has reached.

The messages between the agent and the seller are shown in table 3 below.

<table>
<thead>
<tr>
<th>Direction</th>
<th>Messages</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent to Seller</td>
<td>Req(U)</td>
<td>Request for price discount</td>
</tr>
</tbody>
</table>
Offer a higher discount than the current discount \( U \)

No demand is granted for the offered discount \( D, U \)

Acknowledged with demand \( D, U \)

The offered discount is \( U \)

The offered discount \( U \) is the end offer

In figure 3, we present the sequence diagram of a general negotiation scenario between the agent and the seller. Next, we will show that the negotiation process terminates at Nash Equilibrium.

### 6.4. Negotiation termination and Nash equilibrium

The negotiation process in the presented model guarantees to terminate after a finite amount of exchange messages. Not only that, but also it guarantees that the negotiation process is always terminating (with acceptance) at Nash equilibrium.

**Nash Equilibrium** is an action profile that belongs to the strategy profile \( \sigma = \{s_1, s_2, ..., s_n\} \) with the property that no player can increase his or her payoff by choosing a different action, given the other players’ actions. Formally: \( u_n(s^n_n, s^*_{-n}) \geq u_n(s_n, s^*_{-n}) \)

Where \( u_n \) is the player’s utility corresponding to his Nash equilibrium strategy \( s^n_n \) and the other player’s Nash equilibrium strategy represented by \( s^*_{-n} \).

**Proposition:**

Let \( U \) be the discount such that the demand \( D(U) = N \), where \( N \) is the number of consumers, then the negotiation process will terminate at discount:

a) \( U^* = U \) if \( (P_r - C_r) = \tau > l \), where \( \tau \) is a positive number and \( C_r = U^* \)

b) \( U^* < U \) if \( (P_r - C_r) = l \), and \( C_r = U^* \)

c) \( U^* = U \) if \( (P_r - C_r) = l \), and \( C_r = U \)

a) **Proof:** Let \( Z = D(U^*), (P_r - U^*) \) represents the payoff of the seller for the offered discount \( U^* \). If \( Z \) is not Nash equilibrium, then there exists a utility \( \hat{Z} > Z \) such that \( \hat{Z} = D(U**)((P_r - \hat{U}*) \) and \( \hat{U}** < U^* \) is best response. Assume the seller decided to terminate at \( U** \), the agent then will reply with \( D(U**) < D(U^*) \), since the seller has no prior knowledge of \( D(U^*) \) value, therefore the risk of loosing utility \( Z > \hat{Z} \) is credible. Hence the best response is \( U^* \), that is, to select Nash equilibria that do not involve credible risks.

b) **Proof:** Let \( Z = D(U^*), (P_r - U^*) \) represents the payoff of the seller for the offered discount \( U^* \). Then, \( Z^* \) is the minimum payoff the seller will get since any \( U** > U^* \) will result in \( \hat{Z} \leq 0 \), which is undesirable. The agent in the other hand will not reply with \( D(U**) > D(U^*) \) if the offered discount is not \( U** > U^* \). Hence the negotiation process will terminate at \( U^* \) with no incentive for any player to deviate from his last action.

c) **Proof:** This proof is straightforward following (b).

### 7. Simulation experiment and analysis

In this section, we evaluate the negotiation process with numeric values. The setting of the experiment is as follows: we consider that a group of 1000 consumers submit requests to the agent to complete the transaction with seller 1. The different values of the overall privacy risk \( \psi \) for each consumer is shown in
The overall privacy risk value $\psi$ is derived from private data categorization and the context privacy risk weight for each consumer.

**Figure 4:** Privacy risk values for each consumer

Seller 1 expected price for each consumer record in the list to be $10 and the price for good1 is $80. The risk premium is modeled using a geometric Brownian motion as shown in subsection 5.4. The initial value of the risk premium $R_0$ is 0.01. The experiment is performed with drift values $\mu (1, 1.1, 1.2, \ldots, 2)$.

**Figure 5:** Number of consumers who completed the transaction

The number of consumers who completed the transaction is shown in figure 5. The expected profit of the seller from the consumer information list is presented in figure 6. Table 4 shows consumers social welfare with respect to the discount offer.

**Table 4:** Consumers' SW corresponds to the discount offer

<table>
<thead>
<tr>
<th>Drift Value</th>
<th>Discount offer</th>
<th>$\text{Consumers' SW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.081</td>
<td>6522</td>
</tr>
<tr>
<td>1.1</td>
<td>0.09</td>
<td>7208</td>
</tr>
<tr>
<td>1.2</td>
<td>0.099</td>
<td>7966</td>
</tr>
<tr>
<td>1.3</td>
<td>0.121</td>
<td>9575</td>
</tr>
<tr>
<td>1.4</td>
<td>0.121</td>
<td>8884</td>
</tr>
<tr>
<td>1.5</td>
<td>0.121</td>
<td>8855</td>
</tr>
<tr>
<td>1.6</td>
<td>0.121</td>
<td>8796</td>
</tr>
<tr>
<td>1.7</td>
<td>0.121</td>
<td>8037</td>
</tr>
<tr>
<td>1.8</td>
<td>0.121</td>
<td>7239</td>
</tr>
<tr>
<td>1.9</td>
<td>0.121</td>
<td>6500</td>
</tr>
<tr>
<td>2</td>
<td>0.121</td>
<td>5945</td>
</tr>
</tbody>
</table>

Based on the presented results we can identify the following two cases:

**Case 1.** Values $\mu = (1, 1.1, \text{and } 1.2)$ represent the situation when the risk premium is considered low. In this situation, we noticed that all consumers completed the purchase. The seller accumulated a profit of $3477, $2791, and $2033, while the social welfare of the consumers was $6522, $7208, and $7966 respectively. The negotiation process stopped at discount offers 8%, 9%, and 9.9%. In each case of the negotiation process the number of consumers has reached maximum (1000). The agent could not press for a better offer, since there is no incentive for the seller to offer a higher discount if the agent is not replying with a higher number of consumers.

**Case 2.** Values $\mu = (1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, \text{and } 2)$ represent the situation when the risk premium is high. The negotiation process stopped at a discount offer 12% for all cases. Although, the discount offer was higher than case1, the number of consumers who completed the transaction was less, see figure 5. The reason behind that is the fact that consumers were expecting a higher discount corresponding to the value of the privacy risk. On the other hand, the seller could not afford to offer a better discount. Any discount offer higher than 12% will result in profit losses, which is undesirable. The sellers' profit has dropped significantly as privacy risk increases, see figure 6. From table 4, we can see that consumers’ SW was gradually decreasing as the value of the risk premium increases (drift value increases). Consumers were expecting a higher discount compared to the risk taken, a discount the seller cannot afford.

The experiment showed an interesting result about the power of the market in controlling privacy risk. It showed consumers and online sellers (Web service providers) are both better off if privacy risk is low. The results showed that private data assets are no exception to the fact that assets are matters safely and usually better left to markets.
8. Conclusion and future work

In this paper we presented a novel approach to capitalize on private data assets. We showed that consumers and service providers are better off if privacy risk is low. In this respect, markets complement regulations in a sense that markets impose reputation consequences on violators of consumers presumed privacy rights.

The initial work can be extended in different directions: The first direction can be to study the effect of competition between service providers, (in this paper we examined only the case of a monopoly) that is, to consider two or more service providers competing to attract consumers while considering privacy risk premiums. The second direction is to consider risk premium in Web service composition. Specifically, how Web service provider will fulfill consumers requests while minimizing the risk premium in Web service composition.

9. References

[18] Yingxin He, Dawn N.Jutla” Contextual e-Negotiation for Handling of private data in e-commerce on a semantic Web” Proceeding of the 39th Hawaii International Conference on System Sciences, 2006