Document downloaded from:

http://hdl.handle.net/10251/38085

This paper must be cited as:

Such Aparicio, JM.; García-Fornes, A.; Botti Navarro, VJ. (2013). Automated buyer profiling control based on human privacy attitudes. Electronic Commerce Research and Applications. 12(6):386-396. doi:10.1016/j.elerap.2013.05.001.

The final publication is available at

http://dx.doi.org/10.1016/j.elerap.2013.05.001

Copyright Elsevier

Automated Buyer Profiling Control based on Human Privacy Attitudes

Jose M. Such[a,b,*], Ana Garcia-Fornes[b], Vicent Botti[b]

[a] *School of Computing and Communications*
Lancaster University, UK
[b] *Departament de Sistemes Informàtics i Computació*
Universitat Politècnica de València, Spain

Abstract

In e-commerce applications, vendors can construct detailed profiles about customers' preferences, which is known as buyer profiling. These profiles can then be used by vendors in order to perform practices such as price discrimination, poor judgment, etc. The use of pseudonyms and, specially, changing pseudonyms from time to time are known to minimise profiling, minimising the capacity of vendors to perform such practices in turn. Although there are some frameworks and tools that support pseudonym change, there are few proposals that suggest or directly change the pseudonym in an automated fashion. Instead, users are usually provided with the mechanisms to change pseudonyms but without any advise on when they should actually use these mechanisms. In this paper, we present an approach to control buyer profiling by means of automated pseudonym changes performed according to human privacy attitudes. We also present an application scenario and an evaluation of our proposal.

Keywords: e-commerce, privacy, buyer profiling, pseudonymity, automated pseudonym change, customer profiling, agent-based e-commerce

[*] School of Computing and Communications, Infolab21, Lancaster University, South Drive, Lancaster LA1 4WA, UK.

Email address: `j.such@lancaster.ac.uk` (Jose M. Such)

1. Introduction

The explosive growth of the Internet in the last decades has caused that more than 2 billion users as of 2012[1]. In this environment, on-line privacy is of great concern. Users are constantly exposed to personal information collection and processing without even being aware of it (Fischer-Hübner and Hedbom, 2008). Information collection refers to the process of gathering and storing data about an individual whereas information processing refers to the use or transformation of data that have been already collected (Solove, 2006) — even possibly inferring new data from the data already collected. There are some directives that try to regulate this massive collection and processing of information (e.g., EU Directives 95/46/EC, 45/2001/EC, and 2002/58/EC). However, due to the very nature of the Internet itself, there is no global governing body that could effectively enforce these regulations in daily digital activity. Therefore, these practices are still possible with the potential to jeopardise privacy.

We focus on a type of information processing in e-commerce environments broadly known as buyer — or customer — profiling (Shaw et al., 2001; Such, 2011), in which vendors obtain detailed profiles of their customers based on previous transactions, and subsequently tailor their offers regarding customers' tastes. These profiles can represent a serious threat to privacy. For instance, these profiles can be used to perform *price discrimination* (Odlyzko, 2003). Vendors could charge customers different prices for the same good according to the customers' profiles, i.e., if a vendor knows that some good is of great interest to one customer, the vendor could charge this customer more money for this good than other customers for the same good. For instance, in 2000, Amazon started to charge customers different prices for the same DVD titles (Spiekermann, 2006). When the story became public, Amazon claimed that this was part of a simple price test and discontinued this practice. Another example of privacy threat due to the use of these profiles is what is known as *poor judgment* (Smith and Milberg, 1996). This is when individuals are judged and subsequently treated according to decisions made automatically based on incorrect or partial personal data. For instance, companies usually divide their potential customers into similar groups based on customers' characteristics — known as customer segmentation. This practice can lead to exclusion of people from services based on potentially distorted

[1] `http://www.internetworldstats.com/stats.htm`

judgments (Spiekermann and Cranor, 2009).

Hansen et al. (2004) point out that pseudonyms[2] should be used and, most importantly, changed from time to time to avoid profiling. Indeed, the most privacy-preserving option is to use transaction pseudonyms (Chaum, 1985), i.e., to use a different pseudonym for each different transaction. The problem is that users are often provided with the mechanisms and infrastructures that allow them to change their pseudonyms but without any mechanism that aids them to decide when they should actually change their pseudonyms. Even the very few approaches that automate pseudonym change (such as Warnier and Brazier (2010); Fritsch (2008); Fonseca et al. (2007)) do not consider the fact that there are many cases in which the user can be interested in reusing the same pseudonym, e.g., users may accept a potential privacy loss when some benefit is expected if they reuse the same pseudonym, such as price discounts, the building of a reputation, etc. (Such et al., 2013a). Indeed, several studies have demonstrated that humans have different general attitudes towards privacy (Ackerman et al., 1999; Westin, 1967; Taylor, 2003; The Direct Marketing Association DMA (UK) Ltd, 2012): privacy fundamentalists are extremely concerned about privacy and reluctant to lose privacy; privacy pragmatists are concerned about privacy but they are willing to lose some privacy when some benefit is expected; and privacy unconcerned do not consider privacy loss.

In this paper, we present an agent-based approach to control buyer profiling automatically based on human attitudes towards privacy. In particular, we present an approach in which an agent automatically decides whether or not to change its pseudonym in its next interaction considering an estimation of the privacy loss and the utility of reusing a pseudonym. The crucial point is that this approach does not require human intervention for each pseudonym change decision but agents will comply with its user's attitude towards privacy. For instance, if a user's attitude towards privacy is unconcerned, her/his agent will only consider the utility of reusing a pseudonym to decide whether or not it changes its pseudonym in its next interaction.

[2] A pseudonym is an identifier of a subject other than one of the subject's real names (Pfitzmann and Hansen, 2010). Human beings have been using pseudonyms in the real world for a long time. For instance, in the 19th century when writing was a male-dominated profession, some female writers used male names for their writings. Nowadays, in the digital world, there are a great number of pseudonyms such as usernames, nicknames, e-mail addresses, sequence numbers, public keys, etc.

The remainder of this article is organised as follows, Section 2 motivates the main contribution of this paper. Section 3 presents our proposal for pseudonym change. Section 4 describes a privacy loss function for agent-based e-commerce domains. Section 5 describes an application scenario for our proposal. Section 6 presents the experiments we performed and the results we obtained when applying our proposal to this application scenario. Section 7 presents some related work. Finally, Section 8 presents our concluding remarks.

2. Motivation

There have been some important efforts in the last decades to minimize privacy threats. Clearly, one of the most important has been the rise of Privacy-Enhancing Technologies (PETs). According to van Blarkom et al. (2003) PETs are "*system[s] of ICT measures protecting informational privacy by eliminating or minimising personal data thereby preventing unnecessary or unwanted processing of personal data, without the loss of the functionality of the information system.*". One of the main PETs that have arisen in the last years to prevent profiling is privacy-enhancing identity management (Clauβ et al., 2005). The building block of privacy-enhancing identity management is Pseudonymity (Hansen et al., 2004), which is the use of pseudonyms as identifiers (Chaum, 1985). However, only using pseudonyms for hiding the real world identity of the users is not enough to prevent profiling. According to Hansen et al. (2008), one of the main questions that is relevant for pseudonyms to avoid profiling is the amount of information that can be gathered by linking the data that have been disclosed under the same pseudonym. Social security numbers in the USA are a clear example of a pseudonym that is usually used for a long time and in different contexts. This allows different pieces of personal information disclosed (even in different contexts) to be linked to each other. Moreover, it also allows the inference of other personal information emerging from the combination of data and the application of learning and inference techniques to obtain detailed profiles. These profiles can then be used as explained in Section 1.

Privacy-Enhancing Identity Management Systems (PE-IMS) (Hansen et al., 2004) (Clauβ et al., 2005) are PETs that support the management of pseudonyms to control the nature and amount of personal information disclosed. These systems provide users with facilities that help them to create and select pseudonyms to be used in different online systems — so that these online systems act as relying parties of PE-IMS for validating the pseudonyms of the users. The problem is that these systems do not usually warn or suggest users about when they should change their pseudonyms. Indeed, only very few proposals suggest or directly change the pseudonym of a user in an automated fashion, such as Warnier and Brazier (2010); Fritsch (2008); Fonseca et al. (2007). Moreover, these approaches that automate pseudonym change usually base on generating a new pseudonym for each new transaction, what is known as transaction pseudonyms (Chaum, 1985). However, they do not consider the fact that there are many cases in which the user can be interested in reusing the same pseudonym across different transactions if some benefit is expected (e.g. price discounts, the building of a reputation, etc.), even though this could cause a potential privacy loss.

3. Automated Pseudonym Change

Many empirical studies concluded that humans have different general attitudes towards privacy (Ackerman et al., 1999; Westin, 1967; Taylor, 2003; The Direct Marketing Association DMA (UK) Ltd, 2012). Privacy fundamentalists are extremely concerned about privacy and very reluctant to disclose personal information, they feel that they have already lost too much privacy and are reluctant to lose privacy any more. Privacy pragmatists are concerned about privacy (i.e. they are not willing to lose privacy a priori), but if they expect some utility (e.g. a monetary benefit) they may accept a privacy loss in exchange of this utility. Finally, privacy unconcerned do not consider privacy at all. For instance, a survey made in 2003 among 1010 US adult citizens (Taylor, 2003) shows that 26% of that citizens were considered privacy fundamentalists, 64% privacy pragmatists, and 10% privacy unconcerned. A more recent survey made in 2012 among 1020 UK adult citizens (The Direct Marketing Association DMA (UK) Ltd, 2012) shows that 31% of that citizens were considered fundamentalists, 53% pragmatists, and 16% unconcerned.

We model these attitudes towards privacy to control buyer profiling by means of automated pseudonym changes. In particular, we consider that the decision of whether or not to change a pseudonym is based on a tradeoff between the privacy that will be lost if the pseudonym is not changed and the utility that will be lost if the pseudonym is changed. For instance, in the case of privacy pragmatists, the agent can decide not to change its pseudonym in the next transaction if the privacy to be lost is worth the utility to be gained. We model this problem as a multi-objective optimization problem (Deb, 2005), in which an agent tries to minimize privacy loss while maximizing its utilitarian benefit.

3.1. Option Quality

One of the most used approaches to solve multi-objective optimization problems consists of transforming it into a single-objective problem[3] (Freitas, 2004). This is typically done by assigning a numerical weight to each objective (evaluation criterion) and then combining the values of the weighted criteria into a single value by adding all the weighted criteria.

In our case, agents consider two criteria: privacy loss and utility. Considering these two criteria, agents have two options: to change or not to change its pseudonym in their next transaction. Thus, we are interested in measuring the quality in terms of the privacy loss and the utility of each of these options. An agent will choose the option with the highest quality. We formally define the option set as $\Theta = \{change, nochange\}$. Moreover, we define the quality of an option as:

Definition 1 (Option Quality). *Given a criterion function $c_p(\cdot)$ that evaluates privacy loss, a criterion function $c_u(\cdot)$ that evaluates utility, and weights $w_p, w_u \in [0,1]$ so that $w_p + w_u = 1$, the quality Q_δ of an option $\delta \in \Theta$ is:*

$$Q_\delta = w_p \cdot c_p(\delta) + w_u \cdot c_u(\delta) \tag{1}$$

The specific criterion functions $c_p(\cdot)$ and $c_u(\cdot)$ are domain-dependent. We provide a general privacy criterion function that can be used in e-commerce domains in Section 4. We also provide an example of utility criterion function for a specific scenario of e-commerce in Section 5. Moreover, as privacy loss units may be different from utility units, both criterion functions are expected to return a value in the interval $[0, 1]$ so that they can be comparable. Depending on the final domain, this could require a normalisation process.

[3] We used the single-objective approach to solve the multi-objective problem because it is the most simple and used one, but our proposal is agnostic regarding the approach to solve the multi-objective problem, i.e., our measures of privacy and utility could be used within any other of the approaches to solve this kind of problems in the existing literature on multi-objective optimization. To learn more approaches to solve multi-objective problems refer to, for instance, Deb (2005) and Freitas (2004).

This also implies that the quality of an option $\delta \in \Theta$ will be in that very same interval, i.e., $Q_\delta \in [0, 1]$.

With the option quality formula, agents are able to obtain the quality of each of the options. Thus, they are able to choose whether or not to change their pseudonym in the next transaction. Agents will choose the option with the maximum quality. Formally, an agent will choose an option $\delta^* \in \Theta$ so that:

$$\delta^* = \arg\max_{\delta \in \Theta} Q_\delta \tag{2}$$

3.2. Eliciting weights

We model privacy attitudes by appropriately setting the values for the weights in the option quality formula (Equation 1), i.e., by setting w_p and w_u. An approach to obtain appropriate values for these weights can be based on the existing polls to obtain the privacy attitude of humans, such as The Direct Marketing Association DMA (UK) Ltd (2012), Taylor (2003) or any of the other surveys that Alan Westin conducted between 1978 and 2004 (Kumaraguru and Cranor, 2005). These polls are able to obtain a *degree* of privacy attitude, which can be directly matched to a w_p value (and then we obviously obtain $w_u = 1 - w_p$).

For privacy fundamentalists, w_p will be set to $w_p = 1$ (so $w_u = 0$). This is because privacy pragmatists will only try to minimize privacy loss and will not consider utility at all. For privacy unconcerned, w_p will be set to $w_p = 0$ (so $w_u = 1$), because privacy unconcerned do not care about privacy loss. If $w_p \neq 1 \wedge w_p \neq 0$, we are modelling privacy pragmatists. Moreover, the specific value for w_p and w_u will vary according to the *degree* of privacy attitude of the particular user, i.e., to what extent a user values privacy in front of utility. For instance, a person that has a degree of privacy attitude that is considered pragmatist but it is very close to unconcerned may be modelled with a w_p that will be close to 0.

In the experimental results (Section 6) we use three values that we consider representative for three main pragmatic attitudes w_p: 0.25, 0.5, and 0.75. We set $w_p = 0.25$ to model pragmatic users that value privacy as less important than utility, $w_p = 0.5$ to model pragmatic users that value privacy and utility as equally important, and $w_p = 0.75$ to model pragmatic users that value privacy as more important than utility. However, many other values for w_p and w_u are also possible to model a pragmatic attitude towards privacy as long as $w_p \neq 1$, $w_p \neq 0$, and $w_p + w_u = 1$ are satisfied.

3.3. Pseudonym Management

As our approach is aimed to be completely automated, users do not need to generate pseudonyms by themselves. Instead, we assume that agents are running on top of agent platforms (which are the software infrastructures that facilitate the development and execution of agent-based applications) that provide facilities for generating new unique and random pseudonyms automatically when they decide to change their pseudonym. For instance, Magentix2 (Such et al., 2013b) provides all the needed mechanisms for agents to manage their pseudonyms. Therefore, if an agent that is running on top of Magentix2 decides (using the mechanism presented in this paper) to change its pseudonym, the agent will call to the specific methods of the API that Magentix2 provides to generate a new pseudonym for this agent without requiring direct human intervention. Furthermore, the link between this pseudonym and the identity of its human holder and the link between this pseudonym and the other pseudonyms of the same human holder will not be publicly known. That is, a priori, agents will not know the real world identities of the users they are interacting on behalf of, neither will they know if two different pseudonyms belong to the same agent.

We also assume the use of other privacy-enhancing technologies so that only the pseudonym that an agent is using can be re-identified from interaction to interaction. In particular, we assume that payments are carried out using some kind of anonymous payment mechanism and deliveries are carried out using some anonymous delivery system. Hence, credit card numbers and delivery addresses do not need to be disclosed when an agent acquires a good. For instance, any payment system similar to the untraceable electronic cash presented by Chaum et al. (1990) can be used for anonymous payments. For anonymous deliveries, the privacy-preserving physical delivery system presented by Aïmeur et al. (2005) can be used. Finally, we also assume the use of anonymous communication (e.g. TOR (Dingledine et al., 2004)) so that the IP address and other whereabouts are hidden.

4. Privacy Loss

In this section, we present a general privacy loss function for pseudonymous e-commerce scenarios. Privacy loss is defined in previous works (Lebanon et al., 2006; Li and Li, 2007) as the identifiability of an individual and the personal information that can be linked together and to the individual in case he/she is successfully identified. Identifiability is the ability from an attacker's point of view of sufficiently identify an individual from a set of individuals (the identifiability set) (Pfitzmann and Hansen, 2010; Pfitzmann and Kohntopp, 2001). In our setting, a seller can sufficiently identify a buyer if the buyer re-uses the same pseudonym even if other identifiable information (such as credit card numbers) is hidden by using the technologies described in Section 3.3. This is because, as also described in Section 3.3, pseudonyms are generated randomly and uniquely, so those buyers that do not change their pseudonyms can easily be distinguished from the rest by the sellers. As sellers are able to identify buyers if they re-use their pseudonyms, they are also able to establish the linkability of all the transactions performed by a buyer re-using the same pseudonym. Linkability means that from an attacker's point of view, two different items of interest (in this case two transactions) can be related to each other (Pfitzmann and Hansen, 2010; Pfitzmann and Kohntopp, 2001). The seller can process (as explained bellow) all the linked transactions performed under the same pseudonym and obtain a profile of buyer's tastes, which are consistently considered personal information in the related literature (Rannenberg et al., 2009; Hildebrandt and Gutwirth, 2008). This disclosure of personal information will be seen as acceptable/unacceptable depending on the privacy attitude of the buyer, e.g., a privacy fundamentalist will see any disclosure of personal information unacceptable. Note that the seller does not need to know the real identity of the buyer (or other personal information) to use the constructed profile in future transactions against the buyer (e.g., performing price discrimination, poor judgement, etc.). That is, once the seller has constructed a particular profile about a buyer, the seller only needs that this buyer uses the same pseudonym in the next transaction to be able to identify the buyer as the individual with that particular profile and, thus, to use this profile against the buyer.

We assume that seller agents follow an approach to build buyer agents' profiles similar to Serrano et al. (2013). Based on this, a seller agent marks goods that are not accepted by a buyer agent over the course of a purchase (or transaction) as a negative instance (class "-"), while a seller agent marks goods that buyer agent accepts to buy as a positive instance (class "+"). The seller agent uses all the collected instances about a buyer agent to train a statistical classifier. Thus, the resulting trained classifier models the buyer agent's tastes with a given accuracy.

To avoid profiling, buyer agents need to prevent seller agents from obtaining an accurate classifier for the tastes of the buyer agents' users. Thus, a privacy loss metric for this domain can be defined regarding the estimation that a buyer agent can make about the possible accuracy of the classifier that a seller agent might have constructed based on their previous transactions. To this aim, what we propose is that buyer agents train themselves classifiers with the very same instances that seller agents will have available to construct their own classifier, i.e., the transactions that each buyer agent completed using the same pseudonym with the same seller. Then, buyer agents can test the accuracy of their classifiers with their users' real preferences to know the number of correctly classified instances and have an estimation of the accuracy of the classifiers that sellers might have constructed about them.

We denote by $G = \{g_1, \ldots, g_l\}$ the nonempty and finite set of on-sale goods. $A = \{a_1, \ldots, a_k\}$ is a nonempty and finite set of attributes so that $a : G \rightarrow V_{a_i}$ is a partial function for any $a \in A$, where V_a is the domain of a. We also denote $I_n^{p_i,p_j} = \{\{g_1, c_1\}, \ldots, \{g_m, c_m\}\}$ as the set of m instances (positive and negative) resulting from n transactions[4] of a buyer agent under pseudonym p_i with a seller agent under pseudonym p_j, where each $g_i \in G$ is a good and $c_i \in \{+, -\}$ is the class that states whether a good complies with the preferences of the corresponding user or not. Moreover, we define a function $h_{I_n^{p_i,p_j}} : G \rightarrow \{+, -\}$, which is the statistical classifier that is trained with $I_n^{p_i,p_j}$. Although determining the best classifier to be used is out of the scope of the paper, we used three different classifiers for our experiments that have been proved to offer accurate results in this domain (as detailed in Section 6). We also define a function $f : \{+, -\} \times \{+, -\} \rightarrow \{0, 1\}$, which returns 1 if the two classes passed are equal or 0 otherwise. Finally, we define a set $T = \{(g_i, c_i) \mid g_i \in G \wedge c_i \in \{+, -\}\}$ of random generated goods with its correct class according to the users' preferences. This set is used to test the classifier to obtain its accuracy. Based on this, we define the privacy loss function as:

Definition 2 (Privacy Loss). *Given a buyer agent with pseudonym p_i and a seller agent with pseudonym p_j, the set of instances $I_n^{p_i,p_j}$ resulting from n transactions between p_i and p_j, the classifier $h_{I_n^{p_i,p_j}}$ (trained with $I_n^{p_i,p_j}$), and a set of test goods T,the privacy loss $L(p_i, p_j, n)$ is defined as:*

[4] Note that as shown later on in Section 6, one transaction could result in more than one instance.

$$L(p_i, p_j, n) = \frac{1}{|T|} \cdot \sum_{(g,c) \in T} f(h_{I_n^{p_i,p_j}}(g), c) \tag{3}$$

With this function, the buyer agent can estimate the accuracy of the profile that the seller agent could have about the buyer agent's preferences after a number of transactions.

We now define the privacy criterion function based on this privacy loss function:

Definition 3 (Privacy Criterion Function). *Given a buyer agent with pseudonym p_i and a seller agent with pseudonym p_j, and the number n of completed transactions between them under these pseudonyms, the privacy criterion function $c_p(\delta)$ of an option $\delta \in \Theta$ for the next transaction $n+1$ is:*

$$c_p(\delta) = \begin{cases} 1 - L(p_i, p_j, n+1) & \text{if } \delta = \mathit{nochange} \\ 1 & \text{if } \delta = \mathit{change} \end{cases}$$

For the *nochange* option we estimate the privacy loss that the next transaction $n+1$ will cause, considering $I_{n+1}^{p_i,p_j} = I_n^{p_i,p_j} \bigcup \{(g_{n+1}, +)\}$, so that g_{n+1} is the good that the buyer agent is willing to purchase in the next transaction. For the *change* option we consider that it has the highest quality ($c_p(\mathit{change}) = 1$). This is because we assume that when a buyer agent changes its pseudonym in its next transaction with the seller agent, the seller agent is not able to re-identify the pseudonym[5].

5. Application Scenario

We consider an electronic market where seller agents and buyer agents trade wines on behalf of their users. Seller agents act on behalf of wine merchants. Buyer agents act on behalf of the users that are interested in acquiring wines. Agents in the e-marketplace follow the negotiation protocol depicted in Figure 1 (Serrano et al., 2013). A buyer agent makes a request to purchase a bottle of wine with a *request* message. This message can be replied by the seller agent with either a *model* message (which means that

[5] This assumption should be relaxed in other domains such as inter-vehicle communication systems in which pseudonym changes have been proved as not being enough to prevent profiling (Wiedersheim et al., 2010).

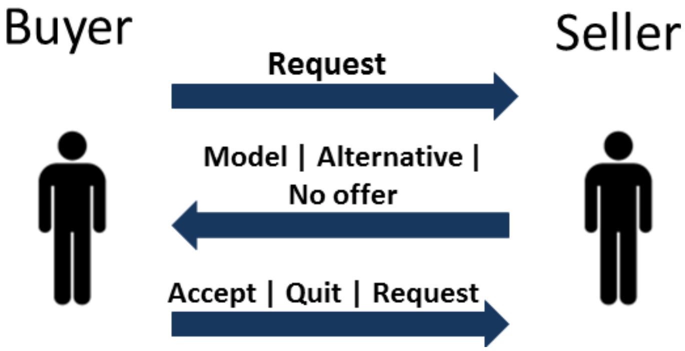

Figure 1: Negotiation Protocol for the Wine e-marketplace scenario.

the requested wine is available) or an *alternative* message (which means that the requested wine is not available but there is another one that is very similar). Then, the buyer agent can reply to both messages with: an *accept* message (which means that the buyer agent accepts the wine offered), a *quit* message (which means that the negotiation was broken by the buyer agent), or a *request* message (which means that the agent requests a new different bottle of wine).

We based on the wine attributes considered in the preference modelling approach described in Aydoan and Yolum (2010). Thus, we consider the following attributes to describe wines: colour, body, flavour, sugar, and country. The possible values for each of these attributes are shown in Table 1.

Attribute	**Values**
Colour	red, rose, white
Body	light, medium, full
Flavour	delicate, moderate, strong
Sugar	dry, offDry, sweet
Country	France, Portugal, Spain, Italy, USA Germany, Australia, New Zealand

Table 1: Considered Wine Attributes.

We define a utility function for this specific scenario based on customer loyalty programs[6]. Retaining customer loyalty is crucial in electronic commerce because the value of an on-line store is largely determined by the number of its loyal customers (Lee et al., 2000). Many vendors use different approaches to achieve customer loyalty such as price discounts, allotment of points that can be used for future purchases, etc.

Loyalty programs usually reward buyers with benefits as they buy more goods. The incentive is for buyers to spend enough to gain access to different levels of rewards (Dowling and Uncles, 1997). We specifically assume a loyalty program with two levels: emerald level, and gold level. When a buyer agent interacts for the first time with a seller agent, the seller agent does not consider the buyer agent as member of any of the levels in the loyalty program. However, if the buyer agent interacts for the second time with a seller agent, the seller agent considers the buyer agent as emerald level member. This means that the seller agent provides a 5% discount to the overall buyer agent purchase in this transaction. From this moment on, each time the buyer agent interacts with the seller agent, it will receive a 5% discount as well as one point. When the agent reaches 50 points, the seller agent will upgrade the buyer agent to gold level member. From this moment on, the seller agent provides a 10% discount to the overall buyer agent purchases.

We define a utility function that models this loyalty program as follows:

Definition 4 (Utility). *Given a buyer agent with pseudonym p_i and a seller agent with pseudonym p_j, and the number n of transactions between them under these pseudonyms, the utility is:*

$$U(p_i, p_j, n) = \begin{cases} 0 & \text{if } n = 1 \\ 0.5 + \frac{n-2}{100} & \text{if } n \geq 2 \wedge n \leq 51 \\ 1 & \text{if } n > 51 \end{cases}$$

The rationale for this utility function is described as follows. In the first transaction, the buyer agent receives no utility. In the second transaction it receives utility 0.5 because the buyer agent is receiving half the maximum discount that the seller agent provides, that is, 5% discount in front of 10% maximum discount. Moreover, from $n > 2$ to $n \leq 51$ the buyer agent

[6] Choosing the most suitable utility function for either this specific scenario or other agent-based e-commerce scenarios is out of the scope of this paper.

receives $0.5 + \frac{n-2}{100}$ because this models that the buyer agent is obtaining the 5% discount but it is also cumulating points in order to reach 50 points and be gold level member. Finally, from $n > 51$ the buyer agent receives utility 1 because it is considered as gold level loyalty program member so that it receives the maximum discount possible, that is, 10% discount.

We define the utility criterion function for this scenario as:

Definition 5 (Utility Criterion Function). *Given a buyer agent with pseudonym p_i and a seller agent with pseudonym p_j, and the number n of completed transactions between them under these pseudonyms, the utility criterion $c_u(\delta)$ of an option $\delta \in \Theta$ for the next transaction $(n+1)$ is:*

$$c_u(\delta) = \begin{cases} U(p_i, p_j, n+1) & \text{if } \delta = \textit{nochange} \\ 0 & \text{if } \delta = \textit{change} \end{cases}$$

We consider that the utility criterion function returns 0 when changing the pseudonym in the next transaction $(n+1)$, this is because the seller agent will not recognize the buyer agent with the new pseudonym. The seller agent will consider that it is the first time that it is interacting with that buyer agent. Therefore, the seller agent will not consider the buyer agent as member of any of the loyalty program levels.

6. Experimental Results

We conducted several simulations considering buyer and seller agents. In each of these simulations, each buyer agent performs 100 different transactions, i.e., each buyer agent performs 100 different purchases of a bottle of wine. Moreover, to support the findings with statistic significance, each transaction was repeated 1000 times. Each transaction involves a negotiation with a seller agent to get the desired wine. We assume that negotiations are always successful, i.e., buyer agents always purchase a bottle of wine. However, we consider that negotiations can randomly involve from 1 up to 10 rounds of the protocol depicted in Figure 1. That is, we simulate negotiations in which a buyer agent and a seller agent perform a maximum number of 10 rounds of the protocol. Based on this, a seller agent marks wines that are not accepted by a buyer agent as a negative instance (class "-"), while a seller agent marks the first requested wine and the wine of the last step in the protocol (i.e., the wine that the buyer agent accepts to buy) as a positive instance (class "+").

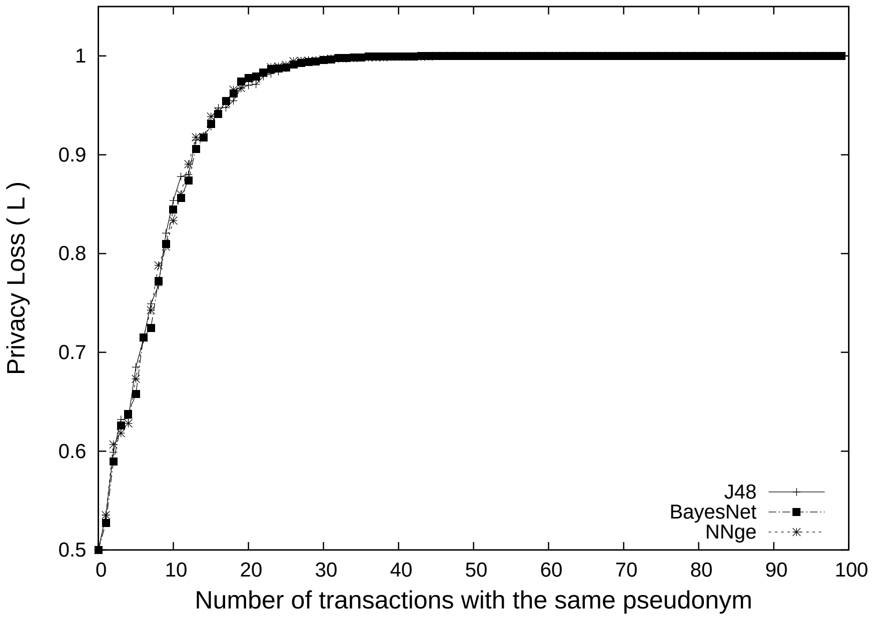

Figure 2: Privacy lost as the number of transactions with the same pseudonym increases.

For each new transaction, the buyer agent uses all of the wines that it has available — those generated in all of its previous transactions with the seller agent as well as the wine that is willing to purchase in the next transaction — as instances to train a classifier. The buyer agent then uses the privacy loss function (Equation 3) to calculate the number of correctly classified instances out of the total number of instances from an extra set of test instances (positive and negative). The buyer agent randomly generates these test instances according to its user's preferences. The result of the privacy loss function is used as an estimation of the accuracy of the profile that the seller agent may have constructed on the buyer agent's user.

6.1. Pseudonym Change Effectiveness

In this section, we describe the experiment we performed to ascertain to what extent changing pseudonyms can reduce the accuracy of the profiles that seller agents can construct about buyer agents. To this aim, we considered a buyer agent that represents a user with the following preferences:

$(body = light \wedge flavour = delicate) \vee (sugar = dry \wedge country = Portugal)$

We repeat the overall simulation (i.e., 100 purchases) three times so that each time the buyer agent is using a different classifier. Specifically, we consider the same classifiers as in Serrano et al. (2011), which are proved to be effective and accurate to learn preferences in this kind of domains. These classifiers are: the J48 decision tree algorithm (an implementation of the C.45 algorithm), the NNge classification rules algorithm (Nearest neighbor like algorithm using non-nested generalized exemplars) and the BayesNet classifier that is a classifier based on Bayesian networks. We use the implementation of these classifiers that is freely available in the Weka[7] open source data mining software.

Figure 2 shows the results we obtained. As it can be observed, the more transactions the buyer agent carries out with the seller agent, the more accurate are all the models that the seller agent constructs with all of the classifiers. Moreover, this relationship among transactions and accuracy is not linear. Instead, the accuracy rapidly achieves the maximum of 100% of accuracy. After 10 transactions, the accuracy is almost 80%. Moreover, after 30 transactions the seller agent is able to obtain a 100% accurate model of the preferences of the buyer agent's user. This means that the seller agent has a complete and certain profile on the preferences of the buyer agent's user, which clearly represents a threat for the privacy of the buyer agent's user.

With only one transaction, the seller is only able to construct classifiers that obtain an accuracy of 50%, which means that the corresponding classifier is not able to distinguish between what the buyer agent's user prefers and what he/she does not prefer. This confirms the hypothesis made in most of the privacy-enhancing technologies literature, i.e., the most privacy-preserving option is to use a different pseudonym for each transaction (known as transaction pseudonyms in the privacy-enhancing technologies literature (Hansen et al., 2008)).

6.2. Automated Pseudonym Change

In this section, we assume the same preference profile for the buyer agent as the one described in the previous section. However, we repeated the simulation changing the privacy attitude of the buyer agent, i.e., we considered different values for w_p and w_u for the buyer agent. In particular, we re-

[7] http://www.cs.waikato.ac.nz/ml/weka/

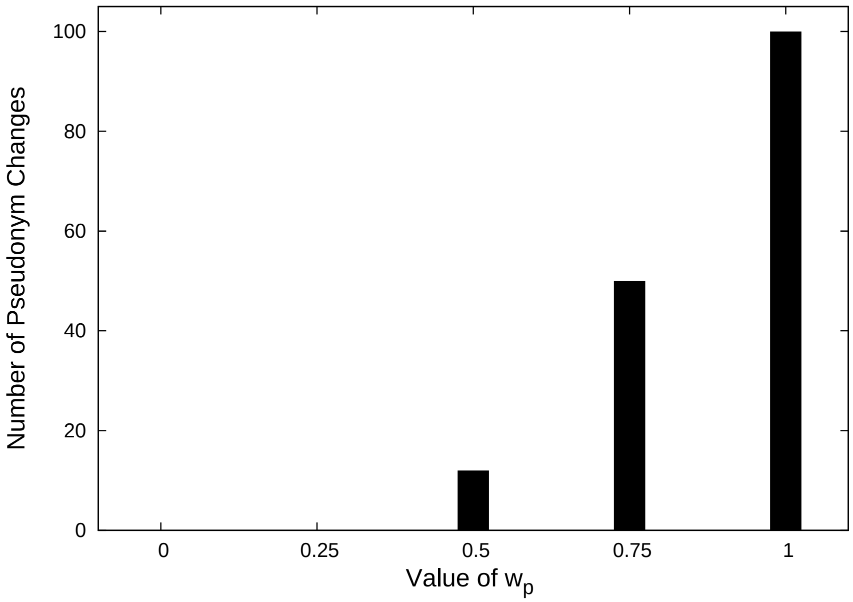

Figure 3: Number of pseudonym changes per each w_p value.

peated the simulation 5 times with the following values for w_p (recall that $w_u = 1 - w_p$): $0, 0.25, 0.5, 0.75, 1$.

Figure 3 shows the number of pseudonym changes performed by the buyer agent during 100 transactions with the seller agent for each w_p value. For the sake of simplicity and clarity, we only comment the results obtained with the J48 decision tree algorithm because the results obtained with the other two classifiers were very similar (as one could expect from the results obtained in the experiment detailed in the previous section in which all classifiers performed very similar). As it can be observed, for low values of w_p (0 and 0.25) the buyer agent did not perform any pseudonym change. This is because with this w_p values the buyer agent acts either as a privacy unconcerned ($w_p = 0$) or as a privacy pragmatist that values more utility than privacy ($w_p = 0.25$). For the rest of w_p values the buyer agent performs pseudonym changes in some of its 100 transactions with the seller. Moreover, the higher the value of w_p, the more concerned is the buyer agent about its privacy. Thus, the buyer agent performs the needed pseudonym changes so that it loses the less possible privacy. The highest number of pseudonyms changes occurs when $w_p = 1$, because the buyer agent models a fundamentalist approach to pri-

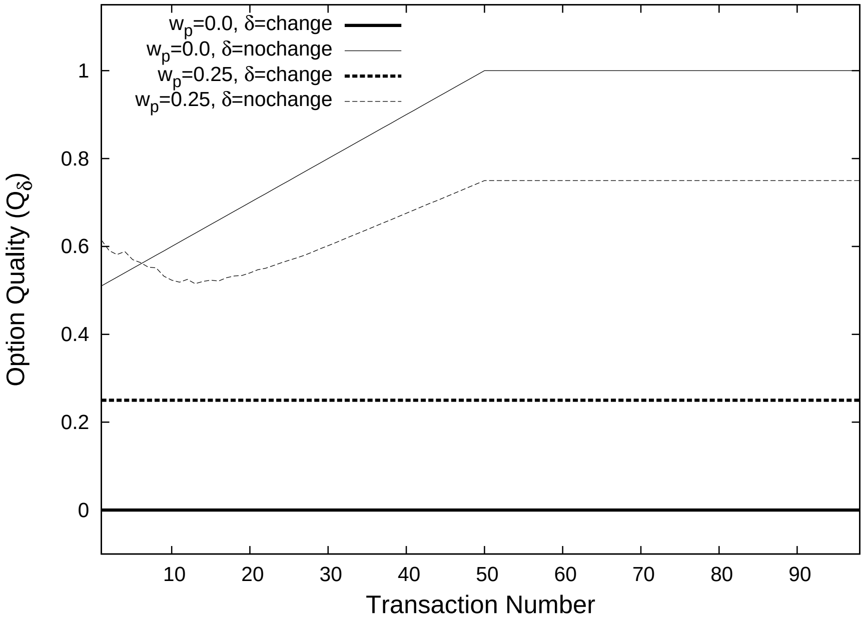

Figure 4: Q value for changing (or not) the current pseudonym in the next transaction. Values for w_p: $0.0, 0.25$.

vacy so that it tries to prevent privacy loss to the greater possible extent. Thus, it changes its pseudonym for each new transaction.

In order to provide the reader with more details about the performance of the buyer agent according to the different values for w_p, Figures 4 and 5 show the quality value (Q) of changing and not changing the pseudonym per each transaction performed with the seller agent. Moreover, to improve visibility, we split the results into two figures: Figure 4 shows the results obtained for w_p: $0.0, 0.25$; while Figure 5 shows the results obtained for w_p: $0.5, 0.75, 1.0$.

Figure 4 shows that for w_p: $0, 0.25$ the quality value of not changing the pseudonym in the next transaction is always higher than the quality value of changing the pseudonym in the next transaction. Thus, the agent does not perform any pseudonym change in any of both cases. Moreover, the results obtained for both cases are very similar. The quality of changing is always the same. This is because the utility criterion function of changing returns 0 and the privacy criterion function of changing returns 1. Therefore, the quality of changing the pseudonym is always 0 for $w_p = 0$ (i.e., the privacy quality is 1 but after being weighted it becomes 0) and 0.25 for

$w_p = 0.25$. Regarding the quality of not changing, when $w_p = 0$, it follows exclusively the pattern described by the utility function defined in Section 5. Thus, when the transaction number is 1, the quality value of not changing the pseudonym in the next transaction is 0.5 (the value returned by the utility function for transaction 2). This corresponds to the emerald level loyalty program membership. Thus, during the following 50 transactions, the quality of not changing increases according to the utility function until reaching the maximum possible value. Then, it obtains the maximum utility of 1, which corresponds to the gold level loyalty program membership. From this moment on, the quality of not changing is always 1.

The quality of not changing when $w_p = 0.25$ behaves a little different. In the first transactions (from 1 to 20), the quality of not changing exhibits a decreasing pattern. This is because the privacy that is being lost devalues the utility that is achieved. However, from 20 transactions on, the privacy loss reaches its maximum (see Figure 2) because the seller agent achieves a nearly perfect model of the buyer agent's preferences. Thus, the privacy quality will be 0. This implies that the quality of not changing starts behaving like the utility function but weighted with 0.75. This is because when $w_p = 0.25$, $w_u = 0.75$. Therefore, when the utility function reaches its maximum (from 52 transactions), the maximum of the quality of not changing is 0.75.

As we can see in Figure 5, for $w_p = 1$ the quality value of changing the pseudonym in the next transaction is always higher than the quality value of not changing the pseudonym in the next transaction. Therefore, when $w_p = 1$, the buyer agent always changes its pseudonym in the next transaction. As a consequence, the seller agent can infer very little from the buyer agent, so the privacy loss is very low and constant. This is the most privacy-preserving option and models a fundamentalist attitude towards privacy, i.e., it minimizes privacy loss without considering the utility that reusing the same pseudonym could cause ($w_u = 0$).

Figure 5 also shows the results for pragmatic attitudes with w_p: $0.5, 0.75$. In both cases, we can observe that the quality value of not changing the pseudonym is higher than the quality value of changing the pseudonym most of the times. However, there are some times in which the quality value of not changing the pseudonym falls under the quality value of changing the pseudonym. This is when the buyer agent decides to change its pseudonym for the next transaction.

We can also see in Figure 5 that the quality of not changing the pseudonym in the next transaction exhibits a very similar behaviour in both cases (w_p:

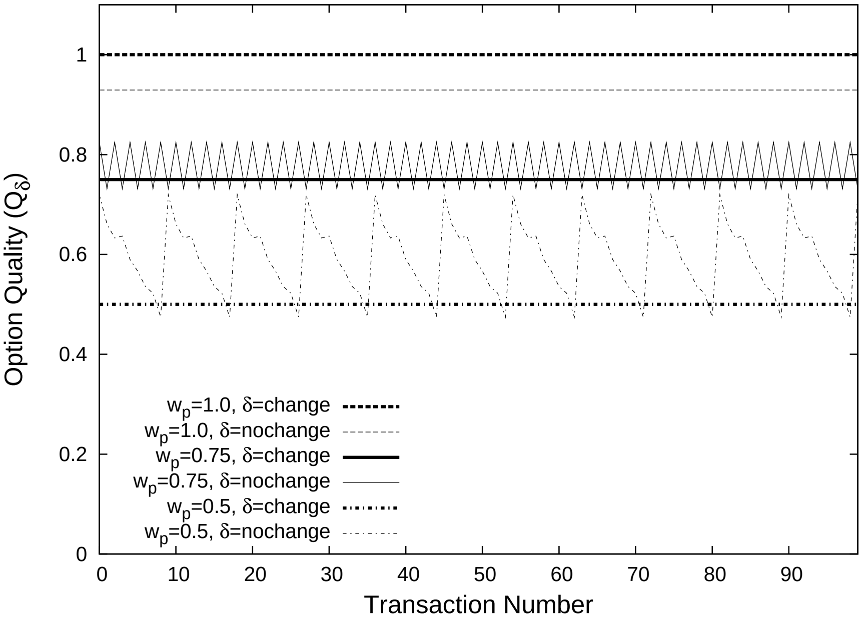

Figure 5: Q value for changing (or not) the current pseudonym in the next transaction. Values for w_p: $0.5, 0.75, 1.0$.

$0.5, 0.75$). It has a phase in which the quality value of not changing the pseudonym increases, followed by a phase in which the quality value decreases. We explain this behaviour as follows. The quality of not changing increases in each transaction because the increase in utility is higher than in privacy loss. However, there is a point in which the privacy that is expected to be lost is high enough so that the utility that is expected to be gained is not worth it. Thus, the quality of not changing starts decreasing. This tendency stops when the quality of not changing decreases below the quality of changing so that the buyer agent performs a pseudonym change. Due to the pseudonym change, the privacy loss is restarted because the buyer agent assumes that the seller agent has no instance corresponding to the new pseudonym to train its classifier. Moreover, the utility is also restarted because when a buyer agent changes its pseudonym, the seller agent will treat the buyer agents as not pertaining to any loyalty program level. From this moment on, with a privacy loss criterion function that returns a good quality in terms of privacy loss, utility starts becoming again the driving force. Therefore, the quality of not changing starts increasing again. This pattern is repeated throughout all of the transactions. Moreover, we can observe that the length of this pattern (in terms of number of transactions) depends on how the buyer agent values privacy in front of utility. Thus, for $w_p = 0.75$ (the buyer agent values privacy more than utility) this length is shorter than for $w_p = 0.5$ (the buyer agent values privacy and utility equally).

6.3. Multiple Buyers

Another important factor that should be considered is the complexity of users' tastes (or preferences) regarding the specific product to be acquired (i.e., in our case the wines that the user likes). In this way, we claim that more complex tastes are more difficult to learn, and thus, for the same number of transactions they involve less privacy loss. That is, a buyer agent with simple tastes will need equal or greater pseudonym changes than a buyer agent with complex product preferences in order to comply with the privacy attitude of its human user.

ID	**Wine Preferences**
1	$(body = light)$
2	$(body = light \wedge flavour = delicate)$
3	$(body = light \wedge flavour = delicate)$ $\vee$ $(sugar = dry \wedge country = Portugal)$
4	$(sugar = dry \wedge colour = red \wedge body = medium \wedge$ $flavour = moderate \wedge country = France)$ $\vee$ $(sugar = sweet \wedge colour = white \wedge body = light \wedge$ $flavour = delicate \wedge country = USA)$
5	$(sugar = dry \wedge colour = red \wedge body = medium \wedge$ $flavour = moderate \wedge country = France)$ $\vee$ $(sugar = sweet \wedge colour = white \wedge body = light \wedge$ $flavour = delicate \wedge country = USA)$ $\vee$ $(sugar = sweet \wedge colour = rose \wedge body = medium \wedge$ $country = Portugal)$ $\vee$ $(sugar = offDry \wedge colour = red \wedge body = full \wedge$ $flavour = strong \wedge country = NewZealand)$

Table 2: Buyer agent's wine preferences (or tastes), ordered by complexity — i.e., ordered in terms of how difficult they are to be learned from individual transactions.

To prove this claim, in this section we consider multiple buyers with different wine preferences (or wine tastes) and with different privacy attitudes (specified by giving different values to w_p). In particular, we consider five different wine preferences shown in Table 2 and the very same five privacy attitudes considered in the previous section ($w_p = \{0, 0.25, 0.5, 0.75, 1\}$). Each buyer is a combination of one wine preference and one privacy attitude, so we consider 25 different buyers.

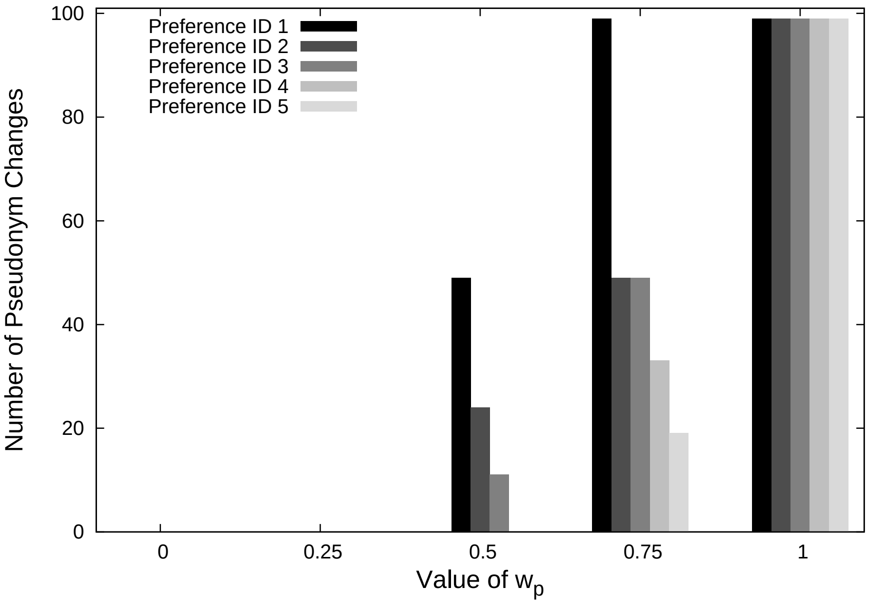

Figure 6: Number of pseudonym changes per each preference ID and per each w_p value.

We conducted a simulation in which each buyer performs 100 different transactions with the same seller (see Section 6.4 for experiments considering multiple sellers). Figure 6 shows the results we obtained for this experiment. As we can see, buyers that have a privacy attitude that is either unconcerned ($w_p = 0, w_u = 1$) or pragmatic but that value utility as being much more important than privacy loss ($w_p = 0.25, w_u = 0.75$) never change their pseudonym regardless the complexity of their wine preferences (or tastes).

This is because buyer agents with that privacy attitudes and for this particular environment will always consider that the utility they are receiving in the form of price discounts is worth the privacy loss. That is, on the one hand buyers that are unconcerned do not consider privacy loss at all. On the other hand, pragmatic buyers that value utility as more important than privacy loss consider that the discounts the seller will apply to their purchases (represented in the form of a utility function in Section 5) are worth the privacy loss due to not changing their pseudonyms.

We can also see in Figure 6 that the specific wine preferences do have a clear effect on the number of changes that buyers perform if they follow a pragmatic attitude that values privacy as equally or more important than utility ($w_p = 0.5, w_u = 0.5$ and $w_p = 0.75, w_u = 0.25$ respectively). In particular, we can clearly see that the more complex are the wine preferences the less pseudonym changes buyers need to carry out. This is because, for the same number of transactions, the more complex are the wine preferences the less accurate the classifier constructed (that models buyers' tastes) is. Thus, for the same number of transactions buyers with more complex wine preferences will experience less privacy loss. As a consequence, they will need less pseudonym changes for the same number of transactions than other buyers that have wine preferences that are easier to learn and that will imply more privacy loss. Indeed, we can see that, regarding buyers with privacy attitude $w_p = 0.5$, they do not need to perform any pseudonym change to comply with that privacy attitude if their wine preferences are either 4 or 5, which are the most difficult to learn.

Another interesting result is that buyers that have a privacy attitude $w_p = 0.75$ and wine preferences 2 and 3 perform the same number of pseudonym changes. This is because of the very nature of the privacy loss function, which is not completely lineal with the number of transactions (though it will always be monotonically increasing with the number of transactions performed). Note however, that the exact privacy loss will be different for both cases but it will not be different enough to imply a different amount of pseudonym changes. Thus, we can state that given two buyers with the same privacy attitude but different wine preferences, for the same number of transactions the buyer with the most easy-to-learn preferences (from the two different wine preferences) will need equal or greater pseudonym changes than the buyer with the least easy-to-learn preferences to comply with the privacy attitude of its user.

Finally, we can see in Figure 6 that for buyers with a fundamentalist attitude ($w_p = 1, w_u = 0$), the results are equal for each all of the preferences. This is because fundamentalists will only consider privacy loss, i.e., they will try to prevent privacy loss to the greater possible extent. The most privacy preserving option is always to change their pseudonym (transaction pseudonyms), because not changing them will always imply a loss of privacy regardless the complexity of its preferences. That is, the more complex preferences are, the less privacy loss, but there will be always a privacy loss by re-using the same pseudonym, so a fundamentalist will always try to avoid it.

To sum up, we can conclude that although the most important factor that will drive pseudonym change to comply with a privacy attitude is precisely the privacy attitude itself, it is clear that how difficult it is to learn the tastes of a buyer will always influence the final number of pseudonym changes needed to comply with the privacy attitude, because this will determine the amount of privacy that is lost by re-using a pseudonym. Thus, for the same privacy attitude, different complexity of the buyer's tastes will imply a different number of pseudonym changes required to comply with that privacy attitude.

6.4. Multiple Sellers

Finally, we would also like to ascertain whether the number of sellers in a particular scenario could also have any impact on the number of pseudonym changes required to comply with users' privacy attitudes. In particular, we repeated an experiment considering a varying number of sellers from 1 to 50. In each repetition, we considered 5 buyers with the same product preferences (or tastes) and with different privacy attitudes w_p (i.e., $0, 0.25, 0.5, 0.75, 1$) that perform 100 transactions. For each transaction, each buyer picks randomly a different seller[8].

Figure 7 shows the number of pseudonym changes per number of sellers considered and for each type of agent. We can see two different patterns, one in which the number of pseudonym changes remains the same, and another

[8] We are only interested in ascertaining how interacting with different sellers could affect the number of pseudonym changes needed to preserve the desired level of privacy in exchange of the price discounts. Thus, the decision to select the most appropriate seller in each interaction is beyond the scope of this paper, e.g., this decision could be made based on who the sellers that offer the highest discounts are, etc.

one in which the number of pseudonyms changes drops as the number of sellers increases:

1. The number of pseudonym changes remains the same regardless the number of sellers when $w_p = \{0, 0.25, 1\}$. For w_p 0/1 it is clear that the buyer will always maintain/change its current pseudonym in its next interaction because of its unconcerned/fundamentalist privacy attitude, i.e., it is only interested in utility/privacy respectively. For $w_p = 0.25$ we have that, as shown in Section 6.2 for only one seller, the buyer never changes its pseudonym due to its privacy attitude that values more the expected benefit than the privacy loss, and in this case the discount it receives is worth the privacy it losses. Thus, when the number of sellers increases the privacy loss will be even less than with only one seller, so that the utility received will still be worth the privacy loss and no pseudonym change will be performed.
2. The number of pseudonym changes drops as the number of sellers increases for $w_p = \{0.5, 0.75\}$. This is because for these values of w_p the privacy attitude is one pragmatic that values privacy as being greater or equally important to privacy loss. Thus, when the number of sellers increases, the probability of interacting with the same seller decreases (recall that we are modelling buyers that choose sellers randomly). As the number of interactions with the same seller decreases, the privacy loss will decrease as well. Therefore, these agents need less pseudonym changes to maintain the same amount of privacy loss.

7. Related Work

There have also been proposed agent-based approaches to support the management of pseudonyms in e-commerce, but only to a limited extent. For instance, users connect to the IntelliShopper agent (Menczer et al., 2002) using a pseudonym to avoid the link between the profiles that IntelliShopper has about customers and their real identity. Moreover, users can use different pseudonyms for IntelliShopper to have separate profiles for separate activities. However, the authors of this work leave users with the responsibility of creating their pseudonyms and they do not provide any pseudonym management facility.

Another agent-based approach for providing general support for pseudonymity is providing this support from the agent platform (AP), which is the software

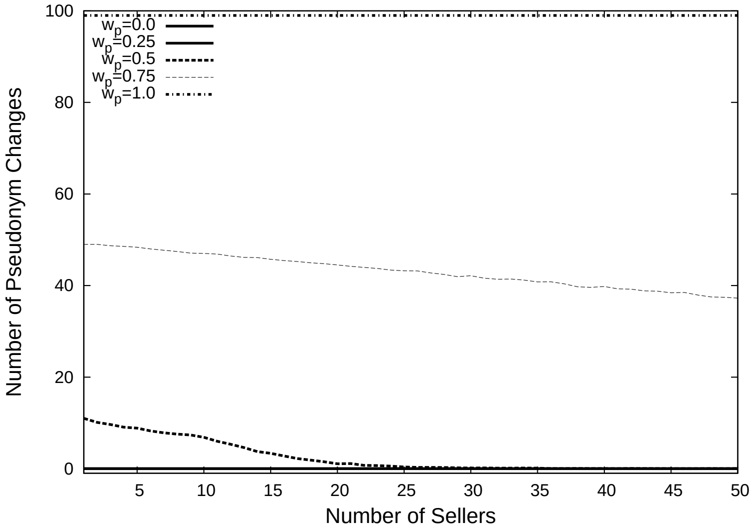

Figure 7: Number of pseudonym changes per number of sellers and for each type of agent.

infrastructure that facilitates the development and execution of agent-based applications. This support aids agent developers to use pseudonymity without having to implement their own solutions. However, only few APs implement some kind of support for pseudonymity. Magentix (Such et al., 2011a), Secmap (Ugurlu and Erdogan, 2005), AgentScape (Quillinan et al., 2008) and Cougaar (Newman, 2004) assign a unique identity for each agent that it can use to authenticate itself to other agents. Using this identity, agents can act pseudonymously, i.e., agents can act on behalf of their principal without using the identity of their principal. However, agents cannot hold more than one pseudonym, i.e., principals should use a different agent each time they want to use a different pseudonym.

Warnier and Brazier (2010) also present a mechanism for the AgentScape AP that offers pseudonymity. At will, agents can ask AgentScape for new pseudonyms. AgentScape also offers an automatic pseudonym change service. For each new transaction, the service generates a new pseudonym. Again, this does not consider that changing the pseudonym for each new transaction may not be always appropriate, i.e., the user can be interested in reusing the same pseudonym across different transactions if some benefit is expected.

Moreover, AgentScape itself must be completely trusted. This is because AgentScape knows the link of pseudonyms to each other and to the principal involved. This usually implies that the organization or company that hosts the specific system (e.g. eBay in the case of an e-marketplace) knows this link as well. Therefore, this organization or company can collect and process information constructing profiles about the users that run their agents on the system.

Other more general agent-based approaches have been proposed to support pseudonymity in (van Blarkom et al., 2003) and (Such et al., 2011b). Both approaches propose the use and integration of Privacy-Enhancing Technologies (PETs) and agent technologies. On the one hand, Van Blarkom et al. (van Blarkom et al., 2003) propose the use of Identity Protectors. Identity Protectors are in charge of converting the identity of the user involved into one or more pseudonyms. They propose that the Identity Protector is placed either between the user and the agent or between the agent and the environment. However, they do not provide any specific design or implementation of an Identity Protector. On the other hand, Such et al. (2011b) present a proposal based on the aforementioned PE-IMS. This proposal has been integrated into the Magentix2 AP (Such et al., 2013b). In this way, Magentix2 relies identity management on external trusted Identity Providers, which are PE-IMS. Therefore, this management is decoupled from the system where the pseudonyms are to be used, and the system (e.g. eBay in the case of an e-marketplace) would encounter more difficulties to perform the association of pseudonyms to each other and to users[9]. Agents running in Magentix2 can obtain new pseudonyms at will and they can select which pseudonym to use in their next interaction automatically. However, nothing is said about when a pseudonym should be changed or not. That is, there are the technical means to change a pseudonym, but they would require direct human intervention to decide whether a pseudonym change is appropriate or not. Thus, as it has been identified in (Such et al., 2013a) and as we have pointed out over the course of this article, automated control of buyer profiling by

[9] Note that this may not completely prevent information processing. There is still the possibility that an agent running on Magentix2 or Magentix2 itself could collude with some of the Identity Providers in order to be able to link a pseudonym to its corresponding real identity. However, agents could (partially) address this by obtaining different pseudonyms from different Identity Providers so as to decrease the probability of being traced back in case of collusion (Such et al., 2013b).

means of pseudonym changes was still an open challenge. In this article, we have presented an approach for automated buyer profiling control based on human attitudes towards privacy.

8. Conclusions

The main contribution of this paper is, to the best of our knowledge, the first automated mechanism for buyer profiling control based on pseudonym changes in e-commerce environments. To this aim, our proposed mechanism considers both the privacy/utility loss of reusing/changing a pseudonym. In particular, agents decide whether to change a pseudonym or not based on the specific attitude towards privacy of their users. This specific attitude is what determines to what extent an agent values the privacy loss and the utility of reusing/changing a pseudonym. We also contribute a general privacy loss function that can be used in e-commerce environments that measures the accuracy of the profile built by sellers based on previous interactions with them.

We also presented an application scenario and the experiments we performed to validate our proposed mechanism. The results we obtained prove that changing pseudonyms can prevent buyer profiling in e-commerce scenarios, though, of course, at the expense of the benefits of re-using a pseudonym — this is actually the reason why a model like the one presented is needed, because different persons with different privacy attitudes will see this expense as acceptable or not to preserve their privacy. The results also validated that the pseudonym changes suggested by our proposed mechanism will follow the privacy attitude of the user involved. Moreover, we obtained results that confirm that there are other factors that will determine the final number of pseudonym changes needed to comply with a particular privacy attitude. In particular, the results we obtained point out that the final number of pseudonym changes to be performed will also depend on the complexity of users' tastes, i.e., users' tastes that are more difficult to learn will imply less privacy loss so that less pseudonym changes will be needed. Finally, we also proved that the number of different sellers with whom the buyers interact will also impact the number of pseudonym changes required to comply with a given privacy attitude.

Finally, as future work we plan to develop a complete privacy-enhancing agent-based e-marketplace application that will build on our mechanism presented in this paper implemented in agents running on top of the Magentix2 agent platform. To this aim, there are still other functionalities that this application would need to provide and that are beyond the scope of this paper, e.g., models to search for sellers that sell the goods that buyers would like to acquire, models that allow agents to select the most appropriate seller to interact with from all the sellers that provide these goods, etc.

Acknowledgements

This work has been partially supported by CONSOLIDER-INGENIO 2010 under grant CSD2007-00022.

References

Ackerman, M.S., Cranor, L.F., Reagle, J., 1999. Privacy in e-commerce: examining user scenarios and privacy preferences, in: EC'99: Proceedings of the 1st ACM conference on Electronic commerce, ACM, New York, NY, USA. pp. 1–8.

Aïmeur, E., Brassard, G., Onana, F.S.M., 2005. Privacy-preserving physical delivery in electronic commerce, in: Proceedings of IADIS International Conference on e-Commerce, pp. 25–33.

Aydoan, R., Yolum, P., 2010. Learning opponents preferences for effective negotiation: an approach based on concept learning. Autonomous Agents and Multi-Agent Systems , 1–37.

van Blarkom, G., Borking, J., Olk, J. (Eds.), 2003. Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents. College bescherming persoonsgegevens.

Chaum, D., 1985. Security without identification: transaction systems to make big brother obsolete. Commun. ACM 28, 1030–44.

Chaum, D., Fiat, A., Naor, M., 1990. Untraceable electronic cash, in: CRYPTO '88: Proceedings on Advances in cryptology, Springer-Verlag New York, Inc., New York, NY, USA. pp. 319–27.

Clauβ, S., Kesdogan, D., Kölsch, T., 2005. Privacy enhancing identity management: protection against re-identification and profiling, in: DIM '05: Proceedings of the 2005 workshop on Digital identity management, ACM, New York, NY, USA. pp. 84–93.

Deb, K., 2005. Multi-objective optimization, in: Burke, E.K., Kendall, G. (Eds.), Search Methodologies. Springer US, pp. 273–316.

Dingledine, R., Mathewson, N., Syverson, P., 2004. Tor: The Second-Generation Onion Router, in: 13th USENIX Security Symposium, San Diego, CA, USA. pp. 303–20.

Dowling, G., Uncles, M., 1997. Do customer loyalty programs really work"?". Sloan management review 38, 71–82.

Fischer-Hübner, S., Hedbom, H., 2008. Benefits of privacy-enhancing identity management. Asia-Pacific Business Review 10, 36–52.

Fonseca, E., Festag, A., Baldessari, R., Aguiar, R., 2007. Support of anonymity in vanets-putting pseudonymity into practice, in: Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE, IEEE. pp. 3400–5.

Freitas, A., 2004. A critical review of multi-objective optimization in data mining: a position paper. ACM SIGKDD Explorations Newsletter 6, 77–86.

Fritsch, L., 2008. Profiling and location-based services (lbs), in: Hildebrandt, M., Gutwirth, S. (Eds.), Profiling the European Citizen. Springer Netherlands, pp. 147–68.

Hansen, M., Berlich, P., Camenisch, J., Clau, S., Pfitzmann, A., Waidner, M., 2004. Privacy-enhancing identity management. Information Security Technical Report 9, 35 – 44.

Hansen, M., Schwartz, A., Cooper, A., 2008. Privacy and identity management. IEEE Security & Privacy 6, 38–45.

Hildebrandt, M., Gutwirth, S., 2008. Profiling the European Citizen: Cross-Disciplinary Perspectives. Springer Publishing Company, Inc.

Kumaraguru, P., Cranor, L., 2005. Privacy indexes: A survey of westin's studies. Technical Report CMU-ISRI-5-138. Carnegie Mellon University, School of Computer Science, Institute for Software Research International.

Lebanon, G., Scannapieco, M., Fouad, M.R., Bertino, E., 2006. Beyond k-anonymity: A decision theoretic framework for assessing privacy risk, in: In Privacy in Statistical Databases, pp. 217–32.

Lee, J., Kim, J., Moon, J.Y., 2000. What makes internet users visit cyber stores again? key design factors for customer loyalty, in: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA. pp. 305–12.

Li, N., Li, T., 2007. t-closeness: Privacy beyond k-anonymity and l-diversity, in: In Proceedings of IEEE International Conference on Data Engineering.

Menczer, F., Street, W.N., Vishwakarma, N., Monge, A.E., Jakobsson, M., 2002. Intellishopper: a proactive, personal, private shopping assistant, in: Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3, pp. 1001–8.

Newman, A.E., 2004. Cougaar developers' guide. `http://www.cougaar.org`.

Odlyzko, A., 2003. Privacy, economics, and price discrimination on the internet, in: Proceedings of the 5th international conference on Electronic commerce, ACM, New York, NY, USA. pp. 355–66.

Pfitzmann, A., Hansen, M., 2010. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. http://dud.inf.tu-dresden.de/Anon_Terminology.shtml. V0.34.

Pfitzmann, A., Kohntopp, M., 2001. Anonymity, unobservability, and pseudonymity - a proposal for terminology, in: Designing privacy enhancing technologies, Springer. pp. 1–9.

Quillinan, T.B., Warnier, M., Oey, M., Timmer, R., Brazier, F., 2008. Enforcing security in the agentscape middleware, in: Proceedings of the 2008 workshop on Middleware security, ACM. pp. 25–30.

Rannenberg, K., Royer, D., Deuker, A. (Eds.), 2009. The Future of Identity in the Information Society: Challenges and Opportunities. Springer Publishing Company, Incorporated.

Serrano, E., Rovatsos, M., Botia, J., 2011. Mining qualitative context models from multiagent interactions (extended abstract), in: Proceedings of the tenth international joint conference on Autonomous agents and multiagent systems.

Serrano, E., Such, J.M., Botia, J., García-Fornes, A., 2013. Strategies for avoiding preference profiling in agent-based e-commerce environments. Applied Intelligence .

Shaw, M., Subramaniam, C., Tan, G., Welge, M., 2001. Knowledge management and data mining for marketing. Decision Support Systems 31, 127–37.

Smith, H.J., Milberg, S.J., 1996. Information privacy: measuring individuals' concerns about organizational practices. MIS Quarterly 20, 167–96.

Solove, D., 2006. A taxonomy of privacy. University of Pennsylvania Law Review 154, 477–560.

Spiekermann, S., 2006. Individual price discriminaton - an impossibility?, in: International Conference for Human-Computer Interaction (CHI'2006), Workshop on Privacy and Personalization.

Spiekermann, S., Cranor, L.F., 2009. Engineering privacy. IEEE Transactions on Software Engineering 35, 67–82.

Such, J.M., 2011. Enhancing Privacy in Multi-agent Systems. Ph.D. thesis. Departament de Sistemes Informàtics i Computació, Universitat Politècnica de València.

Such, J.M., Alberola, J.M., Espinosa, A., García-Fornes, A., 2011a. A Group-oriented Secure Multiagent Platform. Software: Practice and Experience 41, 1289–302.

Such, J.M., Espinosa, A., García-Fornes, A., 2013a. A Survey of Privacy in Multi-agent Systems. Knowledge Engineering Review In press.

Such, J.M., Espinosa, A., Garcia-Fornes, A., Botti, V., 2011b. Partial identities as a foundation for trust and reputation. Engineering Applications of Artificial Intelligence 24, 1128–36.

Such, J.M., García-Fornes, A., Espinosa, A., Bellver, J., 2013b. Magentix2: a privacy-enhancing agent platform. Engineering Applications of Artificial Intelligence 26, 96–109.

Taylor, H., 2003. Most People Are "Privacy Pragmatists" Who, While Concerned about Privacy, Will Sometimes Trade It Off for Other Benefits. Harris Interactive. Retrieved February 27, 2011, from http://www.harrisinteractive.com/vault/Harris-Interactive-Poll-Research-Most-People-Are-Privacy-Pragmatists-Who-While-Conc-2003-03.pdf.

The Direct Marketing Association DMA (UK) Ltd, 2012. Data privacy: What the consumer really thinks.

Ugurlu, S., Erdogan, N., 2005. An overview of secmap secure mobile agent platform, in: Proceedings of Second International Workshop on Safety and Security in Multiagent Systems.

Warnier, M., Brazier, F., 2010. Anonymity services for multi-agent systems. Web Intelligence and Agent Systems 8, 219–32.

Westin, A., 1967. Privacy and Freedom. New York Atheneum.

Wiedersheim, B., Ma, Z., Kargl, F., Papadimitratos, P., 2010. Privacy in inter-vehicular networks: Why simple pseudonym change is not enough, in: Wireless On-demand Network Systems and Services (WONS), 2010 Seventh International Conference on, pp. 176 –83.