Agents for Preserving Privacy: Learning and Decision Making Collaboratively

Onuralp Ulusoy(✉) and Pınar Yolum

Utrecht University, Utrecht, The Netherlands
{o.ulusoy,p.yolum}@uu.nl

Abstract. Privacy is a right of individuals to keep personal information to themselves. Often online systems enable their users to select what information they would like to share with others and what information to keep private. When an information pertains only to a single individual, it is possible to preserve privacy by providing the right access options to the user. However, when an information pertains to multiple individuals, such as a picture of a group of friends or a collaboratively edited document, deciding how to share this information and with whom is challenging as individuals might have conflicting privacy constraints. Resolving this problem requires an automated mechanism that takes into account the relevant individuals' concerns to decide on the privacy configuration of information. Accordingly, this paper proposes an auction-based privacy mechanism to manage the privacy of users when information related to multiple individuals are at stake. We propose to have a software agent that acts on behalf of each user to enter privacy auctions, learn the subjective privacy valuations of the individuals over time, and to bid to respect their privacy. We show the workings of our proposed approach over multiagent simulations.

Keywords: Multiagent systems · Online social networks · Privacy

1 Introduction

Collaborative systems enable users to interact online while sharing content that pertains to more than one user. Consider an online social network (OSN), where a user can share pictures that include other users, who are many times able to tag themselves or others, comment on it, and even reshare it with others. Or, an IoT system, in which one security camera would like to share footage of a setting to guarantee security for the people, while one individual would prefer to keep the location of herself secret. In both of these cases, the content being in question relates to multiple entities, who have different privacy concerns or expectations from each other. Even though the content is meant to be shared by a single entity, the content is related to more than the uploader and hence is actually *co-owned* by others [11,21].

When co-owners have different privacy constraints, they should be given the means to make a decision as to either share or not to share the content. However,

N. Bassiliades et al. (Eds.): EUMAS 2020/AT 2020, LNAI 12520, pp. 116–131, 2020.
https://doi.org/10.1007/978-3-030-66412-1_8

current systems enable only the uploader to set privacy settings while publishing contents, but does not allow co-owners to state their constraints. As a result, individuals are left to resolve conflicts via offline methods [14].

Ideally, systems should provide privacy management mechanisms to regulate how content will be shared. Recently, multiagent agreement techniques, such as negotiation [12,21] and argumentation [13] have been used. These approaches have been successful but require heavy computations; that is, they can only be used when the entities can reason on its privacy policies and communicate with others intensively. Moreover, the agents in these systems follow predefined rules but do not learn better ways to preserve their users' privacy over time. An alternative to this is to use auctions [20] where each user bids based on how much she wants to see a content public or private. The decisions are then made based on the winning bids [4,6].

Accordingly, this paper first explains an *agent-based* approach PANO for collaborative privacy management. When a content is about to be shared, agents of co-owners interact over a mechanism to reach a decision. Similar to Squicciarini *et al.* [20], PANO uses Clarke-Tax mechanism, but adapts it to protect users against abuses, and at the same time encourages users to share content online. PANO incorporates a group-wise budget system that ensures that advantages gained by interactions with certain individuals can only be used against the same individuals. Thus, the agents support users in biding automatically for their behalf. Next, we propose an agent architecture called Privacy Auctioning Learning Agent (PANOLA) that uses user's privacy policy as an initial point to bid but then learns to adjust its bidding strategy over time. Learning has been used in context of privacy before, mostly to enable agents to classify whether a user would consider a content private or not [7,18]. However, the learning problem addressed here is different. First, since the content to be shared is co-owned, other agents' actions influence the outcome of a privacy decision. Second, what needs to be learned is not whether a content is private or not, but what the agent would bid to share or not to share the content, given what it has observed and shared before.

Our main contributions in this paper are as follows:

- We provide a fair privacy respecting auctioning method based on Clarke-Tax mechanism, where software agents represent users' privacy requirements and appropriately bid on behalf of the users.
- We develop a privacy-aware bidding strategy for the agents based on reinforcement learning. This gives them the ability to fine-tune their auction bids according to previous experiences and adjust their privacy respecting strategies over time.
- We evaluate the proposed approach over multiagent simulations and show that it achieves superior privacy protection than non-learning cases.

The rest of this paper is organized as follows: Sect. 2 explains PANO in detail, with a focus on how automatic bidding is done for protecting privacy. Section 3 proposes an agent architecture that learns bidding strategies over time. Section 4

describes our multiagent simulation environment and evaluates the effectiveness of learning. Finally, Sect. 5 discusses our work in relation to existing methods in the literature.

2 Agent-Based Auctioning for Privacy: PANO

To enable decisions on co-owned content, we propose co-owners to be represented with software agents. Agents keep track of the privacy preferences of entities and act on behalf of them to reach a decision. We propose PANO, an agent-based privacy decision system, where agents employ auctioning mechanisms to reach decisions on privacy conflicts [24]. PANO uses an extended version of Clarke-Tax Mechanism as an underlying mechanism.

2.1 Background: Clarke-Tax Mechanism

Clarke-Tax mechanism [4] provides an auction mechanism, where participants bid for different, possible actions in the environment. The action that receives the highest total bids from the participants wins and is executed. Different from an English auction, participants who aid in the winning action to be chosen, i.e., that bid towards it, are taxed according to the value they put on it. This is achieved by subtracting the bid values of every single user from the overall values. If the subtraction of a single user's bid changes the overall decision, it shows that the user's bid on this action had a *decisive* value. Thus, the user is taxed with the difference of the actual action's score and the score of action to be taken if that user were not present in the auction [4]. In the context of collaborative privacy, Clarke-Tax mechanism is used to decide on how a content is going to be shared. Squicciarini *et al.* [20] consider three types of sharing actions: *no share*, *limited share*, and *public share*. We follow the same scheme here. When an image is about to be shared, all the relevant participants bid on these three possible actions.

2.2 PANO Auctions

The Clarke-Tax auctions are beneficial for decision making for multiple participants with different opinions, as they support truthfulness [20]. If Clarke-Tax auctions are applied in commerce, then each participant would have their own budget (e.g., money) to bid with. However, since we are emulating the auction idea, the participants are given budgets at the beginning of each auction, which they can use to bid in the current auction or save to bid later. As usual, a participant cannot bid more than her current budget.

When Clarke-Tax auctions are applied in privacy as opposed to commerce, there are two points that need attention: First, users can enter into arbitrary auctions in arbitrary groups to increase their budgets. If budgets earned with one group of users is used to set the privacy in a second group by overbidding, then the system is abused. Second, it is not clear to assign a bid value for privacy.

In commerce, the valuation for an item can be identified more easily, however, for privacy, the difference between values is not easily interpreted. Without clear boundaries to specify the range for bids, agents are left with an uncertainty to express their preferences accurately. We address these two points by offering only group-wise budgets and ensuring boundaries for bid ranges [24].

Group-wise Spending: To prevent abuse of using budgets for trivial auctions with different users, earned budgets can only be used in new contents with the same co-owners. With this, we improve robustness of the system, where malicious users cannot collaborate for increasing their budget and forcing the other users about their own choices. For example, without group-wise Spending, two agents might share arbitrary content over a social network without spending budget for privacy actions, thus increasing their total budget. When they co-own a content with others, they will have extra budget from these previous efforts, and can bid high amounts to force sharing a content over on OSN, while in fact it is a sensitive content for another user that can't outbid the malicious users. With group-wise spending, each agent would have a separate budget for each co-owner group, hence cannot use previously earned budget against a co-owner group if the earned previously budget was with another co-owner group.

Boundaries: Boundaries enable all the agents to bid inside a predefined range. This is beneficial for preventing users that are richer in the budget from dominating the decisions. This also helps agents that participate in the auctions to have better evaluation functions, because they can have a better opinion about the other participants' bids. When the agents know what would be the maximum bid from the others, they can set their bidding strategy accordingly. For example, without the boundaries in place, when an agent considers a content for a privacy action, she would try to bid as much as possible since she would considers others doing the same for the opposite action. But with boundaries, the agent would have a clearer idea about how much to bid, since she will know the amount to outbid in the worse case scenario, where all the agents bid the amount of the maximum boundary for the opposite action.

Definition 1 *PANO:* PANO *auction is defined as a 6-tuple:* $AUC = \{c,AC,A,m,M,BD\}$*, which consists of the auction's related content c, a set of privacy actions (AC), the set of agents (A) that participate in the auction, minimum possible bid (m), maximum possible bid (M) and the set of placed bids (BD), where each bid* $b_{t,a}$ *(*$b_{t,a} \in BD$*) is related to one single action t (*$t \in AC$*) and one single agent a (*$a \in A$*).*

Given a PANO auction defined as in Definition 1, a system can compute the outcome for the agents, and update their budgets accordingly. At the end of each auction, each participant is given an amount that is equal to the half of the maximum possible bid. This prohibits the agent to bid for the maximum possible bid for each auction. That is, the agent would need to save its acquired budget for the next auction to be able to bid higher than average possible bid. Our reason to employ this half of the maximum boundary is that if an agent

acquires more budget than she should use, she would be able to bid the maximum allowed amount for every auction. In this case, it would not make sense for an agent to deliberate the bid amount, since a higher bid would increase her chances to force the action she wants, regardless of the significance of the action. On the extreme opposite case, if the agents would earn very little amount for every auction, they would not be able to bid for many decisions when they consider the content sensitive. In this situation, many privacy violations might occur, and agents would be forced to save their budget for many cases to be able to have a decision in one. Our decision to give half the amount of the maximum possible bid aims to find a balance between these two extreme cases, where agents should deliberate about placing their bids to be able to enforce their decisions only when necessary, but they would still be able to enforce their decisions in the sensitive cases, if they bid reasonably.

2.3 Privacy Policy

Each agent should have an evaluation mechanism on the importance of a content, and how much it is willing to bid for its preferred actions. Since the action set can differ significantly in terms of size, the evaluation mechanism of the agents should rely on some generic, but still comprehensive representation of the represented individuals' privacy preferences. Thus, we propose a 5-tuple privacy policy structure to represent the privacy related choices of the individuals.

Definition 2 *PANO Policy: A* PANO *policy P is a 5-tuple* $P = \{a,n,p,q,i\}$*, where a is the agent that the policy belongs to, n is the audience of the policy who are the users affected by the outcome, p is the contextual properties for the content that the policy will be applied, q is the privacy related action and i is the importance of the policy, which is a rational value between* 0 *and* 1*.*

An example policy of an agent that represents Alice, who wants to share its blood pressure information received from an IoT device with her doctor and nurse can be defined as:

$P = \{Alice,\{doctor[Alice],nurse[Alice]\},info[BloodPressure],share,0.9\}$.

3 Learning to Bid

Existing work in PANO assumes that the agents are homogeneous and bid in a predefined manner. However, this is rarely the case in real life. First, different users have different privacy understandings that can affect their bidding strategies. Second, users do not know their valuations accurately. Third, some users' privacy expectations can change over time, requiring them to bid differently for the same content at two different time points.

In general, users (and thus agents) are not experts of privacy domains. Even though users claim that they care about privacy and can express their privacy concerns, they tend to act differently and their actions can possibly contradict with their privacy requirements [1]. Hence, presenting privacy related actions in

a way that users can understand and fit their privacy requirements with ease becomes essential. For a privacy auctioning mechanism, agents would find it difficult to place an exact bid on a privacy action, but presenting a range from which they can provide their bids, rather than a single value could be easier. Depending on the context, the extent of the range can vary and providing bids on one end of the range versus the other can significantly change the outcome of the bid. For this reason, it is best if an agent can learn over time a range from which it can generate its bids.

In order to facilitate this, we use reinforcement learning [22]. With reinforcement learning, agents can learn how to improve their privacy actions over time by making use of the only few possible visible outcomes in the system and with simple computations. In our adoption of reinforcement learning to PANO; over time, agents' desired actions are rewarded or their bad choices are penalized. According to these, agents explore their set of actions, in order to adapt and act in the best possible way for the current state of the environment. The convergence to learn the best possible action depends on the exploration/exploitation balance of the agents. An adventurous agent can explore from a wider range of actions while risking being penalized, while a conservative agent can avoid taking risk and adapt slowly, but might get stuck in local minima since the best possible action has a bigger probability of never being explored.

In light of the aspects mentioned above that can affect the privacy decisions, we introduce our learning agent, called *Privacy Auctioning Learning Agent (PANOLA)*. PANOLA employs reinforcement learning to learn the bidding ranges, build strategies using defined coefficients and adapt its bidding according to the outcome of previous decisions. In addition, we ensure that PANOLA can act coherently with agents' privacy policies even when previous decisions are not available.

3.1 Bidding Ranges

With the given minimum and maximum boundaries for PANO, we introduce bidding ranges, where the agents can pick from the possible ranges within the boundaries and bid integers between the picked ranges. All the possible bidding ranges within boundaries are stored by the agents themselves; each of them accompanied by a rational *utility* value, in the range of [0–1] that denotes how suitable a range is for bidding for a privacy action; 0 meaning the least suitable and 1 the most suitable. Since the agents cannot have any previous experience when first introduced to a domain, the initial utilities are computed according to the distance of the ranges' mean values to the agents' initial bid evaluations extracted from their privacy policies.

Example 1 Figure 1 depicts two bidding range examples ($r_1 = [4, 12]$ and $r_2 = [14, 18]$) for action t between minimum and maximum boundaries (m and M respectively), assigned as 0 and 20. The set of ranges contains more than these two, since we include all possible integer ranges between m and M. $b_{t,a}$ shows the initial bidding evaluation for action t, which is given as 6 and means that the agent would initially bid 6 for t for the incoming content.

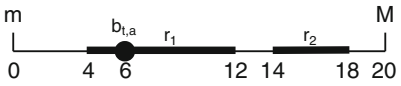

Fig. 1. A depiction of two ranges between minimum (m) and maximum (M) bidding boundaries and the initial bidding evaluation of agent a for action t

In time, utility values of bidding ranges change according to success or failure of the picked bids. Agents do not share the utility values with the environment or other agents. Each agent updates its utilities independently according to the outcome of the auctions. Reinforcement learning is used to make agents learn to pick the most suitable range for a given content type, using information that results from PANO auctions, such as the amount they paid from their budget according to their bids, the deducted tax amount if any tax was paid and the action chosen by the auction, which can be considered as the most important factor for the learning process. We employ all these factors in our computations for learning the suitability of the ranges. The agents pick the range with the highest utility for a given content and bid an integer value inside this range according to their bidding strategy for their preferred action.

3.2 Effective Auctions

An important aspect in facilitating reinforcement learning is to balance exploration of new bid ranges with exploitation of already found ones. The exploration/exploitation balance is not binary in most of the real life domains, since the uncertainty and non-determinism is usually present. Therefore, we make use of continuous utility ranges with several coefficients that represent properties of the auction outcomes to compute the balance.

Like most of the approaches in reinforcement learning [3,5,23], the unsuccessful range pickings are penalized with a decrease in the utility value, while the successful ones have an increase in the utility. In our approach, the utilities are based on the *effectiveness* of the previous auctions. Intuitively, an auction has been effective for an agent if a preferred action has been decided, while the agent did not bid too high and was not taxed too much. We formalize this intuition below using three coefficient values. Table 1 summarizes the important parameters for the proposed approach.

– Bid Coefficient (BC) captures the preference of winning an auction with lower bids. Having a higher BC means that spending less is more important while winning. This is essential when an agent has a limited budget, since winning with a lower bid would enable the agent to have spare budget for the future auctions. In contrast, a rich agent would prefer a lower BC value since bidding more than it should would still leave budget for the future auctions, without the need to search of another winning bid with a lower value.

Table 1. Coefficients and values for utility calculations

Name *Abbreviation*	Short description	Equation/Function	Range
Bid Coef. BC	Used for distinguishing between winning with lower and higher bids	$BC \rightarrow 0$: decrease effect of BC $BC \rightarrow 0.5$: increase effect of BC	[0–0.5
Tax Coef. TC	Changes the importance of taxes in utility calculation	$TC \rightarrow 0$: decrease effect of TC $TC \rightarrow 0.5$: increase effect of TC	[0–0.5]
Action Coef. AC	Assigned by the agents according to their action choice preferences	$AC \rightarrow BC + TC$: decrease effect of AC $AC \rightarrow 1$: increase effect of AC	$[(BC + TC)-1]$
Distance D	Used in the initial utility value calculations	$D = (M - \lvert Mean(rng) - b_{t,a}\rvert)/M$	[0–1]
Effectiveness E	Calculates agent's effectiveness in an auction	$E = AC - (BC * b_{t,a}/M + TC * Tax/M)$	[0–1]

– Tax Coefficient TC has a similar purpose to BC, but it focuses on the amount of taxed budget on winning bids instead of the bids themselves. Similar to BC, a higher TC increases the importance of taxes in utility computation.
– AC enables each agent to decide the importance order of the privacy actions. Agents assign coefficient values between $BC + TC$ and 1 to each action according to their action ordering preferences, the highest coefficient value being the AC of the most important action.

These three aforementioned coefficients are used in computing the final effectiveness. For the Effectiveness (E) value, a higher amount means that the agent's preferred action has been chosen with lower bidding and lower taxing. The ratio of $b_{t,a}$ to the maximum possible bid M gives the magnitude of the bid. The higher this value, the less effective the auction will be. This magnitude is adjusted with BC to account for the fact that different agents would care about this differently. The ratio of Tax to maximum possible bid M gives the magnitude of the budget loss for the agent. Again, the higher this amount, the less effective the auction would be. Adjusting it with TC enables the agent to account for different contexts, e.g., when the agent has high budget and would not be affected by being taxed. The effectiveness of the auction is then the difference between the value gained by the decided action AC and the cost of bidding and taxing as shown in Table 1. The sum of Tax Coefficient TC and the Bid Coefficient BC should be lower than the Action Coefficient AC, so that when an auction is successful,

E will have a positive value and can increase the utility of the picked range for the auction.

The effectiveness of an auction will determine the likelihood of a bidding range to be picked again. However, at the beginning, the agent does not have any effectiveness values, as it has not participated in any previous auctions. Yet, they still need a mechanism to assign bids. Distance (D) formula is used for this purpose of initial utility value calculations. This formula favors bidding ranges that are closer to the agent's initial privacy policy. That is, the distance formula assigns higher utility values to the ranges that have a close mean value to the agents' initial bid evaluations, and lower values to the distant ranges. According to D (in Table 1), if the mean of all the integer values within a range is equal to the initial bid evaluation of the agent, D will be equal to 1, which will be a top pick for the first auction for a related content. The normalization according to the maximum auction boundary ensures that the furthest difference between the range mean and initial bid evaluation would be the difference between the maximum boundary and the minimum boundary (zero for our simulation), since the furthest distance could be the initial bid evaluation to be at one end of the boundary and the mean of the range on the other end. In such case, $|Mean(range) - b_{t,a}|$ part of the D calculation will always be equal to the maximum boundary M, thus the D value will be computed as 0. In addition to enabling first time utilities with D, we also ensure that initial bids are as close as possible to the agents' intended privacy requirements. A utility value closer to 1 would mean that the agent is indeed willing to bid around the mean of the picked range, and the privacy action outcome of the first auction would be similar with when the agent does not employ a learning strategy and bids a value according to its own privacy policies.

Example 2 Referring back to the examples of two ranges in Fig. 1, the mean of r_1 and r_2 are 8 and 16 respectively. If we assume that there are no previous auctions for agent a, the initial bid $b_{t,a}$ is given as 6, which is the amount a is willing to bid for action t, if the learning process with ranges are not available. According to the equation of D, r_1 has the initial utility of 0.9 and r_2 has 0.5. As the mean of r_1 is closer to $b_{t,a}$, it has a higher D value than r_2 and can be considered a better candidate for a bidding range of t for an incoming auction.

3.3 Utility Update

After the initialization with the Distance value, utility computation depends on the Effectiveness value and the total number of auctions entered. Utility for a range called r_x is simply computed with the formula below:

$$Utility\{r_x\} = \frac{\sum_{i=1}^{n} E_i + D_{r_x}}{n+1} \tag{1}$$

According to Formula 1, utility value of r_x after n auctions is the sum of all previous E values and the initial D value divided by the number of entered auctions plus one, considering D.

Example 3 According to the example in Fig. 1, the initial utilities of the ranges according to D value would be $[0-1] : 0.725, [0-2] : 0.75, ..., [4-12] : 0.9, ..., [14-18] : 0.5, ..., [18-20] : 0.35, [19-20] : 0.325$.

If we ignore the ranges that are not shown in the examples above, r_1 ([4–12]) is the one to be picked for the next bid, since it has the highest utility. Assume that the agent picked r_1, won the auction with a bid within the range, and got an E value of 0.8 out of it. The utility of r_1 will become $(0.9 + 0.8)/2$, equaling to 0.85. Since this value is still higher than other ranges above, it will have the highest probability to be picked for the next auction.

4 Evaluation of Learning for Preserving Privacy

The above setup shows us how reinforcement learning can be used by the agents to generate bids. Some important questions that follow are: does this approach enable agents to learn accurately, do the agents that learn bidding ranges perform better in PANO auctions, do other personal values affect preserving privacy and if so, how.

In order to answer these questions, we design and implement a multiagent simulation environment, where PANO and PANOLA agents with different privacy policies enter PANO auctions. The environment consists of a set of agents, and different types of contents, where the agents have predetermined evaluations to rely on. According to these content evaluations, agents have an initial opinion about which privacy actions to support in an auction, and how much they are willing to bid for it. The environment also keeps track of the budget balances of the agents, and their success rate (i.e., the percentage of won auctions in all entered auctions) for further performance evaluations. The content types and the number of actions may vary in the environment, and the required information is fully observable to the agents so they can evaluate on how to bid for a given content type and the set of privacy actions. As in the original Clarke-Tax algorithm, the agents cannot see the bids of the other agents before or after an auction, but they are informed of the winning action as well as the amount of tax to pay, in case they are taxed.

4.1 Simulation System

We have developed a simulation system to evaluate the performance of PANOLA agents in different setups. The environment supports both PANO agents, which do not employ any learning for bidding and PANOLA agents, which learn how to bid over time. The simulation includes multiple action choices and all the agents have predetermined evaluations about how important they consider different action types and how much their initial bid should be accordingly. After the agents are loaded into environment, the simulation cycles for all the contents, and agents enter PANO auctions to collaboratively decide which action to take for the given auctions.

To understand whether an agent is successful, we use a success metric, which calculates the percentage of auctions for which an agent's preferred privacy action is chosen. Recall that the auctions are set up in a such a way that the privacy expectations of the agents conflict. As a result, if an agent's most preferred action is the result of the auction, then this agent has won and the remaining agents have lost. That said, it is possible to have two privacy actions that end with the same highest bid. In those cases, we disregard the auction from calculations of success. Thus, the total wins of all the agents equals the total count of the auctions. This simple metric enables us to see which agents have been the most successful in selecting privacy actions as measured by the percentage of total auctions.

4.2 PANOLA vs. PANO Agents

In our multiagent simulations, there are PANOLA agents that learn how to bid over time and the remaining agents are opposing PANO agents that have different action choices than PANOLA agents. These opposing PANO agents have a static strategy, meaning that they always bid the same pre-evaluated amount for the same type of content.

We perform ten simulation runs of 100 contents for each to evaluate preservation of total budget, amount spent for each content and success for entered auctions (e.g. successful if the first action choice of the agent is the outcome of an auction and unsuccessful if not). The experiments where we include both PANO agents and PANOLA are executed with a single PANOLA against a PANO agent setup, since we aim to measure PANOLA's success with different characteristics against PANO agent opponents that do not learn how to bid over time. The experiments for comparing PANOLA agents with different values against each other are conducted with one-against-one auctions, since our purpose for this comparison is to measure a learning characteristic against another one.

In our first experiment, we evaluate the success of PANOLA against PANO agents in terms of privacy decisions. For all 100 content, our scenario sets the privacy actions of PANOLA and PANO agents always in conflict, thus in each auction the agents oppose each other to ensure their own privacy action becoming the final privacy decision. One of the goals of PANO auctions is to enable every agent to participate for making privacy decisions in the long run, by taxing the winners of the auctions to give a higher chance for the losing agents for the future auctions. Referring back to Sect. 2.2, since we allow agents to earn limited budget (i.e., half of the maximum possible bid) after each auction, even when the agent learns the right bidding range, they might not be able to bid due to lack of budget. Hence, we evaluate whether PANOLA agents learn the right bidding range, we perform auctions with and without budget restrictions. Figure 2 shows the privacy success percentages of PANOLA against PANO agent in both conditions.

As expected, PANOLA learns to outbid the PANO agent after a few auctions, and wins every auction afterwards for the unlimited budget condition.

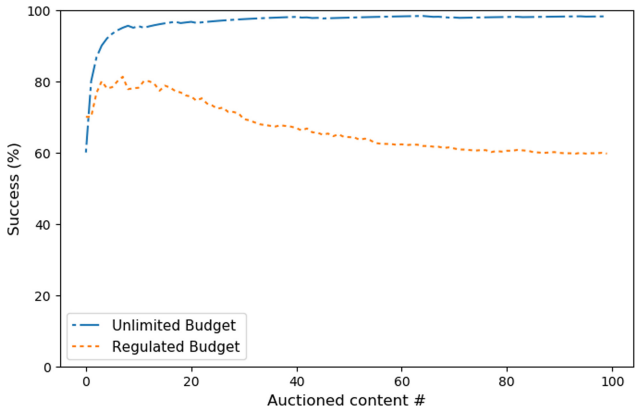

Fig. 2. Privacy success of PANOLA against PANO agents in unlimited and regulated budget scenarios.

This shows that PANOLA indeed learns the correct range to bid from and if PANOLA owns enough budget, it will always choose the correct amount to bid for its privacy actions. When the budget regulation is in place, it is expected for both agents to decide on some privacy actions in the long run, as this is a desired outcome in our mechanism. For the evaluation with the regulated budget, PANOLA still performs better in the long run than PANO agent (˜60% privacy success after 100 auctions); but this time PANO agent is able to give the decisive privacy action for some auctions. The main reason for this is that even though PANOLA learns how to outbid the opponent, it will run out of budget after winning some auctions, and in that case the opponent can win the auction. However, we can also conclude that learning how to bid is beneficial for agents, since adapting the bids for their desired privacy actions enables them to obtain significantly more desired collaborative privacy decisions in their favor than the agents that do not adapt over time.

4.3 Exploration Within Bid Ranges

While learning which range a bid will be given from is the first step, deciding on the actual bid is an important second step. Intuitively, the agent can pick a bid from the range based on a given distribution. Currently, we implement two types of agents, namely *adventurous* and *conservative*. *Adventurous* agents bid randomly within the picked bidding range, while *conservative* agents bid according to normal distribution in Gaussian.

We compare the performance of the adventurous and conservative PANOLA agents against each other. We investigate the success rate and total owned budget of the agents over 100 auctions. Figure 3 shows the success rates and total owned budget over 100 auctions for both agents.

According to Figure 3, it can be seen that conservative bidding achieves slightly more successful results after the agent learns the environment through

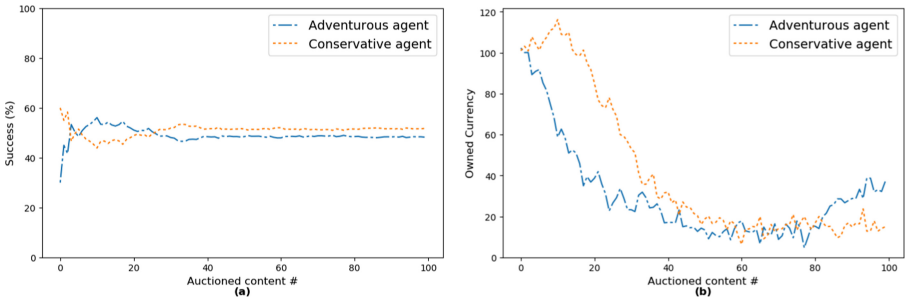

Fig. 3. Success (a) and Owned Budget (b) of adventurous and conservative PANOLA against each other

some auctions. It is also more successful at the first few auctions, while spending more reasonably than the adventurous bidding with random distribution. Around the tenth auction, adventurous agent's success passes conservative, since the adventurous agent tries to increase its bids to beat conservative, while conservative does not increase its bids since it already wins auctions. But after the next few auctions, conservative agent also adjusts its bids accordingly, and stays steadily around 4% more successful than the adventurous agent. The main reason for this difference relies on the Clarke-Tax mechanism; when a conservative agent outbids the adventurous, the tax amount payed tends to be a small amount, since the conservative agent sticks closer to its winning range and not reaching the maximum boundaries. In the opposite position, an adventurous agent can win by trying bids closer to the maximum boundary, but get taxed with a bigger amount which decreases its budget significantly for the next auction. According to this evaluation, it can be said that when two learning agents have the same importance evaluation for an incoming content, using a conservative approach leads to more successful bids in the long run.

With these results, we can conclude that employing conservative strategy in biddings is more beneficial than the adventurous strategy in most cases. However, the learning curve of an adventurous agent while losing is steeper than the conservative one. Thus, when the agent loses most of the bids, trying an adventurous strategy while trying to pick from higher ranges could be useful to find out the winning privacy bids over opponents.

5 Discussion

Privacy in ubiquitous systems started to receive attention around early2000 s, with the Internet becoming accessible to most of the people in the world and enabling easy sharing and access of private information over the web. Langheinrich [15] is one of the first works that investigate the open issues for privacy-respecting approaches for ubiquitous computing. Spiekermann and Cranor [17] and Gürses *et al.* [10] study the grounds of engineering privacy, explaining how

information related domains can be designed to employ privacy-preserving methods. Paci *et al.* [16] provide an extensive survey for literature about access control over community centric collaborative systems; laying down the key issues and giving a roadmap for future challenges. Bahri *et al.* [2] show the challenges of preserving privacy over decentralized OSNs, and provides a review of previous work done for overcoming these challenges. These studies all show that privacy is an important aspect of collaborative information systems and address the need for effective mechanisms.

Even though the systems the main goal is intended to satisfy the general good for the collaborative privacy decisions, the agents that represent entities naturally have the goal to force their privacy requirements to the others.

Collaborative privacy management is investigated in the literature for different domains. Fong [9] introduce Relationship Based Access Control (ReBAC) mechanism, and provides a model to make it applicable to OSNs, where users can define their privacy constraints related to the relations that are available in OSNs, such as friends or colleagues. Multi-party Access Control Model by Hu *et al.* [11] is another work which focuses on determining a single final policy according to privacy requirements of the users. PANO offers [24] a fair mechanism to decide on which action to take, which uses Clarke-Tax auctions at its core with some economic modifications such as group-wise spending, bidding boundaries and income-expenditure balance levels. For the competitiveness of the agents, we introduce a learning mechanism that is based on reinforcement learning, where agents can adapt according to the visible information resulting from the outcome of previous auctions. We also use an evaluation distance coefficient to overcome the cold start problem for the agents that have no prior information about auctions or their opponents.

The use of machine learning for privacy is gaining momentum and the research area is still open for further improvement. Fogues *et al.* [8] provide an agent-based approach which requires user input when required to learn incrementally about user policies, and recommends privacy policies for sharing content for multiuser scenarios. Vanetti *et al.* [25] propose a machine learning approach for filtering unwanted textual contents in OSNs. Squicciarini *et al.* [19] infer privacy policies of OSN users for photographic contents. Zhong *et al.* [26] employ contextual image properties in a different way: they extract and learn from the image features in a way to detect possible privacy conflicts to take further action.

Our work on this paper opens up interesting research directions. The first direction is to use the findings of this paper to build an agent that can change its behavior as needed as well as build models of other agents' in the auctions to make better decisions. The second direction is to capture the dynamics between agents, especially that of trust. When agents trust each other more, they could reflect that differently when bidding, leading to better overall decisions. The third direction is understanding and derivation of social norms into PANO, which could be beneficial to create learning agents according to their normative behavior.

References

1. Acquisti, A., Brandimarte, L., Loewenstein, G.: Privacy and human behavior in the age of information. Science **347**(6221), 509–514 (2015)
2. Bahri, L., Carminati, B., Ferrari, E.: Decentralized privacy preserving services for online social networks. Online Soc. Netw. Media **6**, 18–25 (2018)
3. Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning. Discr. Event Dyn. Syst. **13**(4), 341–379 (2003)
4. Clarke, E.: Multipart pricing of public goods. Pub. Choice **11**(1), 17–33 (1971)
5. Diuk, C., Cohen, A., Littman, M.L.: An object-oriented representation for efficient reinforcement learning. In: Proceedings of the 25th International Conference on Machine Learning, pp. 240–247. ICML 2008, ACM, New York, NY, USA (2008)
6. Ephrati, E., Rosenschein, J.S.: The clarke tax as a consensus mechanism among automated agents. In: Proceedings of the Ninth National Conference on Artificial Intelligence, Vol. 1, pp. 173–178. AAAI 1991, AAAI Press (1991)
7. Fang, L., LeFevre, K.: Privacy wizards for social networking sites. In: Proceedings of the 19th International Conference on World Wide Web, pp. 351–360. WWW 2010, ACM, New York, NY, USA (2010)
8. Fogues, R.L., Murukannaiah, P.K., Such, J.M., Singh, M.P.: SoSharP: recommending sharing policies in multiuser privacy scenarios. IEEE Internet Comput. **21**(6), 28–36 (2017)
9. Fong, P.W.: Relationship-based access control: protection model and policy language. In: Proceedings of the First ACM Conference on Data and Application Security and Privacy, pp. 191–202. CODASPY 2011, ACM (2011)
10. Gürses, S., Troncoso, C., Diaz, C.: Engineering privacy by design. Comput. Priv. Data Prot. **14**(3), 25 (2011)
11. Hu, H., Ahn, G.J., Jorgensen, J.: Multiparty access control for online social networks: model and mechanisms. IEEE Trans. Knowl. Data Eng. **25**(7), 1614–1627 (2013)
12. Kekulluoglu, D., Kökciyan, N., Yolum, P.: Preserving privacy as social responsibility in online social networks. ACM Trans. Internet Technol. **18**(4), 42:1–42:22 (2018)
13. Kökciyan, N., Yaglikci, N., Yolum, P.: An argumentation approach for resolving privacy disputes in online social networks. ACM Trans. Internet Technol. **17**(3), 27:1–27:22 (2017)
14. Lampinen, A., Lehtinen, V., Lehmuskallio, A., Tamminen, S.: We're in it together: interpersonal management of disclosure in social network services. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3217–3226. CHI 2011, ACM, New York, NY, USA (2011)
15. Langheinrich, M.: Privacy by design — principles of privacy-aware ubiquitous systems. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 273–291. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45427-6_23
16. Paci, F., Squicciarini, A., Zannone, N.: Survey on access control for community-centered collaborative systems. ACM Comput. Surv. **51**(1), 6:1–6:38 (2018)
17. Spiekermann, S., Cranor, L.F.: Engineering privacy. IEEE Trans. Softw. Eng. **35**(1), 67–82 (2009)
18. Squicciarini, A., Caragea, C., Balakavi, R.: Toward automated online photo privacy. ACM Trans. Web **11**(1), 2:1–2:29 (2017)

19. Squicciarini, A.C., Lin, D., Sundareswaran, S., Wede, J.: Privacy policy inference of user-uploaded images on content sharing sites. IEEE Trans. Knowl. Data Eng. **27**(1), 193–206 (2015)
20. Squicciarini, A.C., Shehab, M., Paci, F.: Collective privacy management in social networks. In: Proceedings of the 18th International Conference on World Wide Web, pp. 521–530. WWW 2009, ACM, New York, NY, USA (2009)
21. Such, J.M., Rovatsos, M.: Privacy policy negotiation in social media. ACM Trans. Auton. Adapt. Syst. **11**(1), 41–429 (2016)
22. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (2018)
23. Tan, M.: Multi-agent reinforcement learning: Independent vs. cooperative agents. In: Proceedings of the Tenth International Conference on Machine Learning, pp. 330–337. Morgan Kaufmann (1993)
24. Ulusoy, O., Yolum, P.: Collaborative privacy management with auctioning mechanisms. In: Ito, T., Zhang, M., Aydoğan, R. (eds.) ACAN 2018. SCI, vol. 905, pp. 45–62. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-5869-6_4
25. Vanetti, M., Binaghi, E., Ferrari, E., Carminati, B., Carullo, M.: A system to filter unwanted messages from OSN user walls. IEEE Trans. Knowl. Data Eng. **25**(2), 285–297 (2013)
26. Zhong, H., Squicciarini, A., Miller, D.: Toward automated multiparty privacy conflict detection. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 1811–1814. CIKM 2018, ACM, New York, NY, USA (2018)