Privacy for Everyone: Learning to Respect Privacy in Multi-party Content Sharing

Hüseyin Aydın[1,2][0000−0002−5746−9702], Alper Demir[3,4][0000−0003−2646−4850], and Pınar Yolum[2][0000−0001−7848−1834]

[1] Middle East Technical University, Ankara, 06800, Turkey
huseyin@ceng.metu.edu.tr
[2] Utrecht University, Utrecht, 3584 CE, The Netherlands
p.yolum@uu.nl
[3] İzmir University of Economics, İzmir, Turkey
[4] The University of Edinburgh, Edinburgh, United Kingdom
alper.demir@ieu.edu.tr

Abstract. Collaborative systems, such as online social networks or co-creating environments, contain large amounts of multi-party content, where the content pertains to multiple individuals. Ensuring the privacy of such content is difficult, since individuals have different privacy expectations. More importantly, these expectations are not simple share or not share decisions as mostly depicted by existing work, but are nuanced based on the various features of the content. Ideally, a privacy assistant should be able to act based on the privacy preferences of its user. Moreover, the interactions between such agents should lead to good privacy outcomes, such that the privacy of all parties are preserved as much as possible. Accordingly, this paper proposes reinforcement learning agents with a factored representation of preferences, equipping them with the ability to reflect user choices in a fine-grained manner and to account for privacy concerns of others in the group. This enables our agents to learn the correlation between similar contents, having akin yet distinct subsets of known features, and the user preferences. Our experimental results in a multiagent simulation of multi-party privacy show that our proposed method can significantly decrease the privacy violations for all parties, not only on previously seen content but also on new content.

Keywords: multi-party privacy, factored representation, reinforcement learning

1 Introduction

An enormous amount of content is produced daily in the collaborative systems such as online social networks or co-creating environments. A large portion of this content pertains to multiple parties, such as group pictures, co-edited documents, or co-created game worlds. Ensuring the privacy of this multi-party content is difficult, because different parties have different and possibly conflicting privacy preferences. What one party finds private, the other party might be willing to

share [31]. Moreover, many times, the privacy preferences of the users could be much more nuanced. A party might not be interested in sharing vacation pictures, but can be happy to share work pictures. Another party might find pictures taken at night private, but might not care about other aspects [15]. As an example, consider the image in Figure 1a, taken at night on a beach. The individuals in the photo may hold conflicting preferences about sharing it, based on different aspects of the scene. One person might be uncomfortable with sharing nighttime photos, while another might be eager to share images taken at the beach.

One way of dealing with conflicts could be veto voting, where even when one person considers the content private, the content is not shared. However, in practice, most users think such a mechanism is too restrictive, and the systems mainly operate with a single user's decision to share, as the uploader of the multi-party content. Although users may try offline solutions to resolve the conflicts by themselves, this is not practical considering the amount of such content. Hence, several agent-based mechanisms [28,29,33,34,17,18], have been proposed to decide over the privacy of a multi-party content, where the agents represent users and act on their behalf.

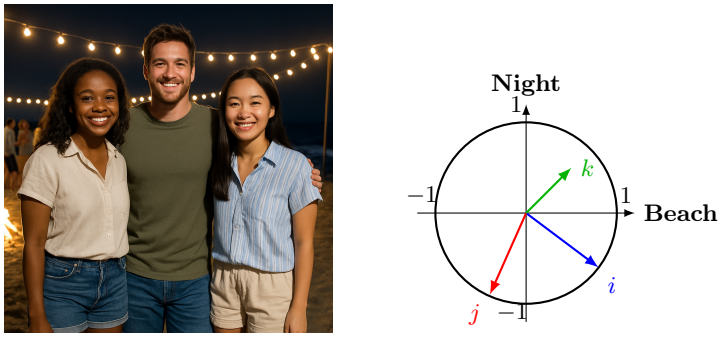
Ideally, such agent-based methods should be able to deal with three crucial aspects of sharing multi-privacy content.

- *Individual effectiveness*: The agent should be able to represent user's privacy preferences in detail and should be able to convert them into privacy sharing actions appropriately.
- *Group awareness*: The agent should be able to observe other users' expectations of privacy and be able to take that into account when generating an action.
- *Content generalizability*: The agent should not require the user to interfere and give input on every new piece of content, but should be able to generalize what the user would expect from previous interactions.

Existing approaches focus on other important aspects of multi-party privacy, such as providing robust mechanisms [25,34], modeling the uncertainty of users about content [3,10], or understanding the usage of such agents by humans [6,26]. However, they have not been addressing how users can express their preferences in detail to account for their own as well as other users' privacy and how this can be realized with minimal intervention of the users.

To address this problem, we propose to represent the users' privacy preferences as feature vectors. Feature vectors enable representation of fine-grained preferences, allowing users to express distinct inclinations (e.g., prefer not to have beach pictures shared but fine with night pictures) rather than expressing an aggregate share or not-share preference for a picture.

This facilitates personalized and precise content sharing decisions. Second, feature vectors support robust generalization across diverse content, as agents can apply established preferences to new, previously unseen items by reasoning over their compositional structure.



(a) (b)

Fig. 1: (a) An example group picture taken on a beach at night, where one individual would like to share beach picture but not if taken at night, while one individual does not share pictures at all and the other is happy sharing such pictures in general. (b) User preferences as vectors over content features: Beach (x-axis) and Night (y-axis).

In image-related privacy contexts, tags as the features of the image have been already shown to effectively expose sensitive content for individuals [32,2]. However, these approaches focus on individual privacy and do not provide solutions for multi-party privacy. Hence, the benefits of a feature-based factorization for the contents and the preferences in a multi-party setting remain mainly unexplored.

We provide a factored representation of privacy and establish how it can be used by agents to make privacy decisions using reinforcement learning (RL). We design two privacy agents with different capabilities and characteristics that utilize this preference factorization for multi-party content sharing. We evaluate the performances of these agents in challenging scenarios where other group members have dissimilar preferences with them. Through our experimentation with publicly available data sets, we show that the factored preferences enable our agents to behave consistently in different groups for rare and even unseen contents as well as help them decrease privacy violations for themselves and for the groups that they are part of.

The rest of the paper is organized as follows. Section 2 gives the background information for better understanding the scope of the study. Section 3 elaborates on the benefits of factored preferences in multi-party content sharing and the proposed formalization of this problem. While Section 4 presents two different agents that are designed to assess the utility of factorization, Section 5 presents the evaluation of these agents in different configurations. Section 6 provides our results of the posed research questions. Finally, Section 7 discusses the work in the context of related work.

2 Technical Background

The following two bodies of work form the backbone of our study.

2.1 Factored Representations in RL and Planning

In sequential decision-making tasks, factored representations decompose information into a set of features or attributes, providing a more structured alternative to *monolithic* representations, where all information is treated as a single indivisible entity. This approach has been widely adopted in both RL and planning paradigms for Markov Decision Processes (MDPs) [12,14,13,7,27,23,9] and partially observable MDPs [22]. Factored representations effectively characterize both observations (complete or partial) and actions [20] through their constituent features. While the majority of research has focused on single-agent settings, these representational frameworks have also been successfully extended to multi-agent scenarios [11,19,4]. By offering more granular information about observations and actions, factored representations enhance sample efficiency through illuminating the underlying relationships between features. Furthermore, compared to monolithic representations, factored approaches significantly expand representational capacity and facilitate improved generalization to previously unseen data.

2.2 Auction Based Multi-Party Privacy

Handling multi-party privacy decisions require a group-decision mechanism to be in place. Auctions have been used for this purpose, where each agent *bids* on behalf of its user to what extent it would like the content in question to be shared or not shared [25]. The bids for each choice are then aggregated in an auction, and a decision to share or not share the content is reached based on which choice gets the highest total bid. In this way, the outcome of the auction becomes *favorable* to the users who prefer that particular choice regarding the content.

To ensure that agents bid in a realistic range and that agents cannot bid a high amount for every content, a *budget* is assigned to each agent [33]. As is usual, an agent that wins in an auction pays for the bid, reducing its budget; whereas for auctions where the agent loses, there is no cost. In addition, if the agent's bid plays a decisive role in the auction—namely, if the group decision would have been different without it, then an extra amount is charged from its budget as *tax*, following the Clarke-Tax approach [5]. Finally, for each new content to be decided, each agent in the system receives the same amount of additional budget, which they can use on top of their existing one [34].

3 Factored Contents and Preferences

We propose a feature-based factorization within an RL framework, so that agents can learn the privacy preferences of the users in a generalizable manner. We

study this representation in an auction setting. Hence, in this setting, an RL agent should learn to bid effectively on the behalf of its user to advocate their preferences in an auction. While doing so, they should also use their limited budget efficiently in order to maintain this advocation in future auctions as much as possible.

In our approach, user preferences are given for each feature, instead of a monolithic, decision-level preference for the whole content. Equipped with such fine-grained and generalizable information, an agent can better adjust its bid to reflect the strength of the user's opinion for the privacy of the content. Hence, through the factored representation of privacy preferences, the agent can accurately meet the user's expectations without requiring further intervention.

3.1 Factored Representation for Privacy

In such a setting, each content $\mathbf{c} \in \{0,1\}^m$ comprises m binary features, where $\mathbf{c}^x = 1$ indicates that the feature x is present in content $\mathbf{c}$. Each agent i possesses a preference vector $\mathbf{p}^i \in [-1,1]^m$, where positive values ($\mathbf{p}^{i,x} > 0$) indicate that feature x is acceptable for sharing, while negative values ($\mathbf{p}^{i,x} < 0$) signify that feature x is considered sensitive or private and thus not acceptable for sharing. The magnitude of each value in the preference vector reflects the strength of the user's opinion regarding that feature.

With this factored representation, the similarity between the privacy preferences of two agents i and j can be calculated by the cosine similarity

$$\cos(\theta_{\mathbf{p}^{i,j}}) = \frac{\mathbf{p}^i \cdot \mathbf{p}^j}{\|\mathbf{p}^i\| \, \|\mathbf{p}^j\|}, \tag{1}$$

where the values range from -1 to 1, from direct opposite to exactly the same.

Using this similarity metric, the agents that represent the users can be grouped according to the preferences of users. Consider contents and preferences in two dimensions for a picture: being at a beach and taken at night. In this two-dimensional space, contents can be classified as:

- A picture taken at a beach but not at night: $\mathbf{c_1} = [1,0]$
- A picture taken at night but not at a beach: $\mathbf{c_2} = [0,1]$
- A picture taken at night and at a beach: $\mathbf{c_3} = [1,1]$
- A picture taken not at night and not at a beach: $\mathbf{c_4} = [0,0]$

Based on this classification, the image in Figure 1a falls under category $\mathbf{c_3}$. Figure 1b shows the given preferences for three agents (or individuals in Figure 1a) in these two dimensions where

- Agent i with preferences $\mathbf{p}^i = [0.8, -0.6]$ likes sharing beach pictures, avoids night photos,
- Agent j with preferences $\mathbf{p}^j = [-0.4, -0.9]$ is very privacy-conscious, dislikes both,
- Agent k with preferences $\mathbf{p}^k = [0.5, 0.5]$ is comfortable sharing in general.

Considering the similarity between the given preferences of agents i and k, they are likely to make similar bidding decisions when presented with a series of beach pictures. Hence, in time, they may learn to adjust their bidding to an amount that is sufficient to win the auction, yet without being obliged to pay an extra amount as tax. In contrast, agent j, recognizing this alignment between i and k from the previous auctions, may learn to overbid on beach pictures in order to steer the outcome toward its own preference—namely, to preserve the privacy of its user.

The feature based similarity between two contents can be also defined with a metric, for example cosine similarity, like in the case of preferences. While the similarity metric for the preferences may be useful to express different groupings for the agents regarding their bidding strategies, the metric for the content similarity may ease the analysis of the decisions made for a cluster of contents. It is very likely that a decision needs to be made over a new content which the agents have not seen before. Since that particular content can be related to the specific combination of already known features, the decision over this novel content can be easily interpreted based on the previous decisions made for similar contents.

3.2 Multi-party Sharing in RL Framework

Regarding the aforementioned advantages of the factored representation, the multi-party content sharing problem now can be formalized within the RL framework as a partially observable stochastic game [1], which is a tuple $\langle \mathcal{N}, S, \{A^i\}_{i\in\mathcal{N}}, T, \{R^i\}_{i\in\mathcal{N}}, \Omega, \{O^i\}_{i\in\mathcal{N}} \rangle$ where:

- $\mathcal{N} = \{1, 2, ..., N\}$ is the set of agents whose users have a say over the given contents,
- S represents the finite set of states,
- A is the joint set of actions such that $A = A^1 \times A^2 \times ... \times A^N$ with A^i is the finite set of actions for agent i,
- $T : S \times A \to \mathscr{P}(S)$ is the transition function with $\mathscr{P}(S)$ as the set of discrete probability distributions over S,
- $R^i : S \times A \times S \to \mathbb{R}$ defines the reward function for agent i,
- $\Omega : S \times A \to \mathscr{P}(O)$ specifies the observation function
- $O = O^1 \times O^2 \times ... \times O^N$ is the joint set of observations so that the agent i receives a local observation $o^i \in O^i$.

Based on this model, the state $s \in S$ can be defined as:

$$s = \langle c, P, \mathcal{B}, H \rangle, \tag{2}$$

including c as the current content, P denotes the preferences of N users, $\mathcal{B}$ as the current budgets of each agent, and H denotes the satisfactions of all users for the decisions made in the previous auctions.

The satisfaction of the user i, namely h^i, could be defined in various ways. Two aspects are important: i) The user's privacy should be violated as little as possible, and ii) When the user would like to share a piece of content, it should be

shared as much as possible. Thus, we propose to define it based on the weighted average of the privacy violations and the sharing desires that are rejected by the group as follows:

$$h^i = 1 - \frac{w_{viol} \cdot |\mathcal{C}_{viol}| + w_{rej} \cdot |\mathcal{C}_{rej}|}{|\mathcal{C}|}, \tag{3}$$

where $|\mathcal{C}_{viol}|$ denotes the number of contents for which the group decisions were to share, while the user i wanted them to be kept private. $|\mathcal{C}_{rej}|$ is the number of contents that were kept as private by the results of the auctions while the user i wanted to share. w_{viol} and w_{rej} are the weights of these two terms. $|\mathcal{C}|$ represents the number of contents that were subject to auctions so far.

A^i, as the action of agent i, is a pair of $\langle d, b \rangle$, namely the individual choice of agent i for the given content and the bid to support it. Hence, $d \in \{$ *SHARE, NOT SHARE* $\}$ and $b \in [b_{min}, b_{max}]$ as the allowed range to bid at the auctions.

Ideally, the agent should take the right decision by giving as low of a bid as possible. Thus, we define the reward of an agent for its action $\langle d, b \rangle$ as:

$$R_i = w_{\mathcal{U}} \cdot \mathcal{U} + w_{\mathcal{E}} \cdot \mathcal{E}, \tag{4}$$

where $\mathcal{U}$ is the utility term that is the evaluation for the decision d at the auction. $\mathcal{E}$ is the efficiency term to measure how efficient the bid b is. $w_{\mathcal{U}}$ and $w_{\mathcal{E}}$ are the weights to adjust the importance of the utility $\mathcal{U}$ and the efficiency $\mathcal{E}$ respectively.

The efficiency $\mathcal{E}$ of a bid is defined as follows:

$$\mathcal{E} = \mathbb{1}_{b_{paid}>0} \cdot (1 - \frac{b_{paid} \cdot b}{b_{max} \cdot b_{afd}}), \tag{5}$$

where b_{paid} is the bid that the agent pays if it wins the auction. The indicator function checks this term since when the agent loses the auction, then b_{paid} becomes 0, and there is no reason to consider efficiency in this case. b is the bid that the agent initially chose, b_{max} is the maximum bid in the range, b_{afd} is the affordable bid by the agent based on its budget. Namely, if the agent does not have enough budget, then b_{afd} becomes the current budget of the agent as the upper bound for b_{paid}. In summary, this efficiency term is introduced to lead the agent not to bid higher than it needs or higher than it is allowed regarding its current budget.

The observation function O, and the utilities $\mathcal{U}$ that are used in their training depend on different observation capabilities and characteristics of users based on their perceptions of others' privacy. Thus, we define them per agent type in Section 4.

4 Agent Design

We design two main types of learning agents: self-oriented and privacy-aware. These agents are meant to capture differences in how users behave in terms of privacy.

Table 1: Utilities for *Self-Oriented* Agent during auctions

Condition	**Utility** ($\mathcal{U}$)
Agent was inline with the user's valuation, and the outcome is favorable for the user.	1
Agent was inline with the user's valuation, but the outcome is not favorable for the user.	0
Agent failed to reflect user's valuation, but the outcome is favorable for the user.	0
Agent failed to reflect user's valuation, and the outcome is not favorable for the user.	-1

Despite the differences in their capabilities and characteristics, both types of agents utilize the factorization of user preferences to better strategize their bidding.

4.1 Self-Oriented Agent

When the agent does not get any information from the environment about other members in the group, its observation is based only on the user whom it represents. Constrained by this limited capability, the available observation for the self-oriented agent becomes like a projection of a state described in Equation 2, where all possible information about others is completely filtered out. Hence, an observation o^i for the self-oriented agent i can be formally defined as:

$$o^i = \langle (\mathbf{c} \odot \mathbf{p}^i), b^i_{afd}, h^i \rangle, \tag{6}$$

where the first term, feature-wise product of the current content and preferences gives the preferences of the user that are related to this content. b^i_{afd} is the affordable amount of credits that agent can bid within its current budget. $h^i \in H$ is the satisfaction of its user.

In the same manner, the utilities for this agent merely focus on its user. Hence, it depends on whether the decision made by the agent matches the overall preference of the user, and whether the outcome of the auction is favorable for the the user. Possible cases and corresponding utilities are defined in Table 1. With such a reward function, the agent is trained to become self-oriented, considering only the preferences of its user, and ignores others' privacy.

4.2 Privacy Aware Agent

The second type of agent can observe others in the group to a certain extent. Hence, the choices of others in the previous auction (D^{-i}) are included in the observation of this agent along with the average satisfaction of users as well as those of its own user (h^{-i}_{avg}). With these additional terms, observation of agent i becomes as follows:

$$o^i = \langle (\mathbf{c} \odot \mathbf{p}^i), b^i_{afd}, h^i, D^{-i}, h^{-i}_{avg} \rangle. \tag{7}$$

With this extension in the observation capability of the agent, it can be trained in a more sophisticated manner, with utilities regarding the effect of

Table 2: Utilities for *Privacy Aware Agent* during auctions

Condition	**Utility** ($\mathcal{U}$)
"Agent was inline with the user's valuation for privacy, and preserved it despite the whole group."	1
"Agent was inline with the user's valuation to share, and the content was shared without any violation."	1
"Agent was inline with the user's valuation for privacy, preserved it with the support of other member(s)."	0.75
"Agent was not inline with the user's valuation to share, but preserved the privacy of other member(s)."	0.75
"Agent was not inline with the user's valuation to share, but tried to preserve the privacy of other member(s)."	0.5
"Agent was inline with the user's valuation for privacy, but failed to preserve it."	-0.25
"Agent was not inline with the user's valuation for privacy, but did not cause any violation."	-0.25
"Agent was not inline with the user's valuation to share, but the content was shared without any violation."	-0.25
"Agent was inline with the user's valuation to share, but risked at least one other member's privacy."	-0.25
"Agent was inline with the user's valuation for privacy, but was not able to preserve it as well as at least one other member's privacy.	-0.5
"Agent was inline with the user's valuation to share, but caused a violation for at least one other member."	-0.5
"Agent was not inline with the user's valuation for privacy, and caused a violation for its user."	-0.75
"Agent was not inline with the user's valuation for privacy, and caused a violation for its user and as well as at least one other member."	-1
"Agent was not inline with the user's valuation to share, caused an unnecessary rejection for the whole group."	-1

outcomes for others. Utilities in this function are based on not only the preferences of the agent's user but also conditions like whether the outcome of the auction leads to any privacy violation for others in the group. In this sense, the agent not only fights for the preferences of its user but also tries to avoid any violations of the privacy for the entire group.

As this agent needs to build its policy to preserve privacy for all members in the group, the conditions for its reward function naturally depend on more factors, including privacy violations considering both its user and the group. Furthermore, the risk of any violation should also be taken into account to make the agent behavior consistent for any resulting decision of an auction. Consequently, Table 2 presents the utilities that are constructed for the problem environment including the Privacy Aware Agent.

5 Experiments

Our experiments aim to answer the following research questions:

RQ1 Does factored preferences enable an effective advocacy of the user choices in multi-party content sharing?

RQ2 Can we better preserve the privacy of all parties by using factored preferences compared to decision-level preferences?

RQ3 How robust is the generalization of factored preferences in multi-party content sharing regarding groups including users with dissimilar preferences?

In order to study the three research questions, we conducted a series of reinforcement learning experiments. We compare our factored approach to a setting that uses monolithic (one-hot) encodings for both content and agent preferences. In other words, this setting produces the primitive versions of our agents where their observation spaces are not factorized based on the content and user preferences. We will refer these agents as *Self-Oriented Monolithic Agent* and *Privacy Aware Monolithic Agent*.

Self-Oriented agents are used to compare the effectiveness in advocacy for the user preferences in factored and monolothic representations. Hence, the experiments that used these agents are held to answer **RQ1**. Meanwhile, to answer **RQ2**, the benefits of factorization in preserving the group privacy are evaluated in the experiments including Privacy Aware agents. Finally to answer **RQ3** and test the robustness of the factored approach, all agents are evaluated in two different data sets including different types of contents along with *dissimilar* preferences for the agents.

To ensure controlled comparisons, we evaluate a single learning agent that interacts with other baseline agents following fixed policies. This setup allows us to capture the impact of representation on learning dynamics. For the sake of simplicity, we assume that the preference vector can be reduced to a scalar by calculating the mean of all non-zero preferences for the given features. This allows us to compare the agent's decision with the user's designated choice for the content. While this calculation gives the plain information about the content to the monolithic agents, it also constructs the decision mechanism of the non-learning agents.

5.1 Baseline Agents

This section presents the fixed decision and bidding strategy for the non-learning agent we used as a baseline in our experiments. For both factored and monolithic type contents, this agent uses the scalar preference for the overall content, which is calculated by the mean of non-zero features of the content. Depending on how strong the preference is for the content, the agent bids in a corresponding range. Table 3 presents the complete setting for the baseline agents. When the agent's preference value p_c is between -0.2 and 0.2, its decision is randomly constructed with a lower bid, reflecting its indecisiveness. Even though baseline agents do not learn over time, predicting their behavior by other agents is still challenging, because the baseline agents choose bids from a range, rather than always providing the same scalar bid value.

Table 3: *Baseline* agent's decision selection and bidding ranges based on given scalar preferences calculated for the content.

Preference	**Decision**	**Bid**
$-1 \leq p_c < -0.5$	*NOT SHARE*	[15,20]
$-0.5 \leq p_c < -0.2$	*NOT SHARE*	[10,15)
$-0.2 \leq p_c \leq 0.2$	*NOT SHARE* or *SHARE*	[1,5]
$0.2 < p_c \leq 0.5$	*SHARE*	[10, 15)
$0.5 < p_c \leq 1$	*SHARE*	[15, 20]

5.2 Data Sets

We use two different data sets for the experimentation. For both of the data sets, we created a small subset of features that are mainly based on YourAlert data set [24], where the users were asked to annotate photos from their own photo collections. In that study, feature extraction from these photos was done by an automated software, so that the annotation can be associated with the features without violating the users' privacy.

We included some additional features that are not present in the original data set in order to cover further aspects that can be important for the multi-party content sharing. For example, certain events like a graduation or party where the photo is taken can be a criterion for the users' privacy preferences [30]. As a result, we had 26 features in total for the contents and preferences that are going to be used in the experimentation (see Appendix A for the complete list). While the data sets are created with the features that focus on the images, the agents are agnostic to the content and feature types and only operate on a given vector in an abstract manner.

Based on these features we first created a set of contents where we can control the dependencies among the features and preferences to make the possible cases reasonable and easy to interpret. 1000 contents including maximum 5 features at the same time were created for the experimentation with this *controlled* data set. We used randomly generated discrete preferences $\{-1, -0.5, 0, 0.5, 1\}$ for the said 26 features matching with a Likert-scale for shareability of the feature from the users' point of view.

The second data set, called *random*, is arbitrarily generated as 1000 contents including maximum 15 features at the same time. The purpose of evaluating the agents with this data set is to test the generalization capability of the methodology with dense contents, independently from any assumption that we made for the first data set. Hence, we randomly generated continuous preferences in the range of [-1, 1] for the same features in the second data set. The data sets and the source code used in the experiments are available online along with the supplementary material[5].

[5] `https://github.com/huseyinaydinmetu/mpc_sharing_w_rl`

5.3 Experimental Setup

For the evaluation of our agents with these data sets, we mainly followed the auction configuration in the previous studies that are based on auction-based mechanisms [33,34]. Hence, the initial budget for each party is set as 50. After each auction, agents are given 10 more credits. Bids in an auction are limited to values from 1 to 20. In the case where an agent changes the outcome of the auction, 5 credits are deducted as tax from the agent's budget. This amount is fixed for the sake of simplicity, but can easily be extended to depend on the bids as in the Clarke-Tax mechanism [5]. In order to avoid over-prioritization of the budgets during training, $w_{\mathcal{E}}$ is set as 0.25, while $w_{\mathcal{U}}$ is set as 1 in Equation 4, for the reward functions of all agents. In line with the control theory of privacy [21], we assumed that the users are more sensitive about privacy violations than the rejection of their sharing desires. Hence, w_{viol} is set as 1.5, while w_{rej} is set as 0.5 in Equation 3.

We trained the learning agents using the DQN [16] algorithm with the following parameters: a learning rate of 0.001, batch size of 128, discount factor (γ) of 0.9, and soft update coefficient (τ) of 0.05. The experience replay buffer stored 10,000 samples, and we updated the target network every 100 steps. The exploration rate (ϵ) decreased linearly from 0.8 to 0 during the first 80,000 training steps.

Agents participating in the auction may exhibit different preferences regarding content shareability. To investigate these scenarios, especially to evaluate the robustness of the generalization with the factorization, we analyzed a challenging configuration with a learning agent interacting with three agents holding dissimilar preferences. Figure 3 in Appendix C illustrates the preference similarity matrices for these agent configurations across both *controlled* and *random* data sets. In order to provide novel contents in the auctions, we split the data sets into training and testing sets with 80% - 20% split. The agents are trained with the training set and the learned policies are tested for generalization on the test set.

The agents and the simulation environment are implemented with Python. The simulations are carried out using a compute server that has two AMD EPYC 7H12 processors and 1 TB RAM with 64-bit Ubuntu 22.04 operating system.

5.4 Metrics

We evaluated our agents based on the following metrics:

- **Reward** is used to determine how the factorization affects the learning process for both agent types.
- **Alignments** measures the number of auctions where the user gets favorable outcomes. Note that the agent's decision might differ from the user's valuation. Hence, the number of alignments is not equal to the number of auctions that the agent wins.
- **Own Privacy Violations** measures the number of auctions that resulted in sharing of the contents where the user would want to keep them private.

– **Others' Privacy Violations** measures the number of auctions that resulted in sharing of the contents and at least one other member would want to keep them private.
– **Average Payment** measures how much the agent spends on average for an auction that it wins.

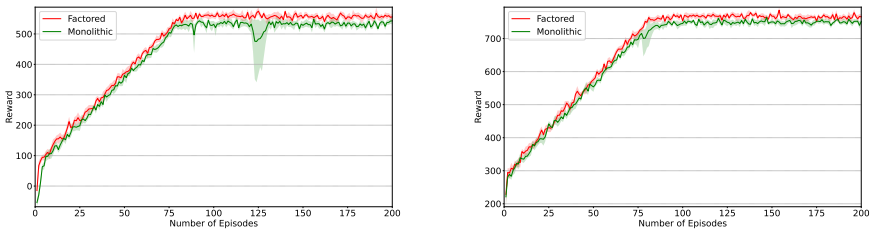
A good policy should show high alignment with the user it represents, low privacy violations and low payment signaling better use of bidding.

6 Results

6.1 Alignment with User Preferences

Figure 2a illustrates the performances of the Self-Oriented Agent utilizing factored versus monolithic representations in *controlled* data set when interacting with dissimilar agents in terms of the reward collected per episode during the training phase. In these episodes, learning allows the agent to increase the reward (defined according to the utilities in Table 1 and the function in Equation 4). The rapid convergence to high reward values indicates effective learning performance and success in the advocation of user's preferences.

A more explicit comparison can be done in terms of favorable outcomes that the Self-Oriented agents get for their users. While the factorization leads to 506.16 (with std of 51.90) alignments during the auctions against the agents with dissimilar preferences, the monolithic agent with the same 493.26 (with std of 52.00) alignments. All of the results are averaged over 10 experiments, each of which includes 200 episodes. As evidenced by both higher reward and number of alignments, factored representation converges to a stable and effective policy more rapidly.



(a) Self-Oriented agents (b) Privacy Aware agents

Fig. 2: Reward of agents during training in *controlled* data set.

6.2 Privacy Preservation

The performance of the Privacy Aware agents in *controlled* data set when interacting with dissimilar agents is shown in Figure 2b, with respect to the reward (determined by Equation 4 as well, but based on the utilities in Table 2) collected per episode during the training. The factorization leads the learning agent to reach higher rewards in less time.

These agents can be directly evaluated with privacy violations, namely the number of contents that are decided to be shared in the auctions while at least one user wants to keep them private. Privacy Aware Monolithic Agent fails to preserve the privacy of 227.42 (with std of 63.66) contents on average during the auctions against the dissimilar agents. The average number of such contents reduces to 220.96 (with std of 65.39) for the Privacy Aware Agent tested in the same setting. All of the results are again averaged over 10 experiments, each of which includes 200 episodes. As suggested by higher rewards and reduced number of violations for each member, factorization leads to better privacy preservation for everyone in the group.

6.3 Generalization to Unseen Contents

Table 4 presents the performance of the learning agent using factored and monolithic representations when tested on unseen content against agents with dissimilar preferences in the *controlled* dataset. The results clearly demonstrate that factored representations enable the learning agent to better capture user preferences, resulting in improved or comparable alignment, fewer privacy violations for both the learning agent and other participants on content not included during training. Additionally, the results reveal that agents using monolithic representations exhibit greater standard deviations during testing. Notably, factored representations facilitate superior user modeling, mainly better optimization in average payment.

Table 5 presents the performance metrics of the learning agents with factored versus monolithic representations in *random* data set when tested on unseen content. In this data set, characterized by more dense features and continuous preferences, factored representations consistently yield better alignment with user preferences while reducing privacy violations for all agents. The agent utilizing monolithic representations demonstrates poorer performance on unseen content

Table 4: Performance on unseen contents with *controlled* data set. The results are averaged over 10 episodes and 10 experiments with standard deviations in parentheses where each episode takes 1000 steps.

	Alignment	**Own Privacy Violations**	**Others' Privacy Violations**	**Average Payment**
Self-Oriented Agent	**524.04 (25.79)**	**155.75 (17.91)**	**62.37 (12.90)**	13.30 (0.81)
Self-Oriented Monolithic Agent	347.01 (41.00)	270.11 (58.43)	146.21 (48.02)	**11.02 (4.03)**
Privacy Aware Agent	**395.80 (15.67)**	**156.89 (12.30)**	**59.35 (8.01)**	**8.90 (0.65)**
Privacy Aware Monolithic Agent	359.22 (16.83)	188.30 (19.59)	79.71 (13.87)	10.42 (1.87)

Table 5: Performance on unseen contents with *random* data set. The results are averaged over 10 episodes and 10 experiments with standard deviations in parentheses where each episode takes 1000 steps.

	Alignment	**Own Privacy Violations**	**Others' Privacy Violations**	**Average Payment**
Self-Oriented Agent	**732.71 (14.16)**	**129.95 (13.09)**	**112.94 (12.69)**	12.21 (0.18)
Self-Oriented Monolithic Agent	452.23 (58.75)	373.60 (51.66)	335.96 (49.02)	**8.66 (2.43)**
Privacy Aware Agent	**555.36 (17.98)**	**79.53 (9.53)**	**66.70 (7.67)**	11.51 (0.20)
Privacy Aware Monolithic Agent	521.65 (16.65)	112.29 (14.85)	92.40 (13.41)	**10.43 (2.04)**

and struggles to win auctions. The primitive agents seem to be better in optimizing the average payment, yet the results become comparable with the greater standard deviations.

7 Discussion

In this work, we propose a feature-based factorization of user preferences instead of assigning a decision for each content. In this manner, the agent that represents the user is provided a fine-grained information, instead of all-or-nothing decisions. Furthermore, this type of factorization leads to the generalization of the contents in terms of the associated features, so the agent can be more robust for rare or unseen contents.

In order to evaluate the effect of this factorization, we design two privacy agents that utilize the feature-based user preferences in an RL setting to manage the multi-party content sharing with an auction based group-decision mechanism. While one of the agents only cares about the preferences of its user, the other one tries to preserve the privacy of the whole group. Despite the differences between their characteristics, they both outperform their primitive versions which are given only a scalar to represent the overall preference of the user for the given content. Our empirical analysis shows that factorization supports the generalization of the contents and enables the learning agent to construct more robust and effective policies, especially in terms of alignments that its user gets for their privacy choice with the outcome of the auctions and the number of privacy violations.

In addition to the generalization capability, the factorization provides more transparency for a multi-party content sharing system. By feature-level explanations, the decision-making process becomes more interpretable for the users. This capability gives the opportunity to analyze the sharing decision for a given set of features, while such an analysis is not possible where each content is independently evaluated in a monolithic structure. Hence, it also paves the way for the sharing systems like online social networks to establish a trustworthy relation with their users.

Advantages of the feature based factorization in the privacy of user images have been addressed by Tonge and Caragea [32] and Spyromitros-Xioufis et al. [24]. However, these studies focus on the classification of the images from individual perspectives without considering multi-party cases. Such et al. [30]

attack the problem for this type of cases where multiple users have a say over the images. Yet, the aim of this work is to investigate the characteristics of conflicting preferences of users and how users normally try to resolve their differences with mostly offline and corrective manners on online social networks with an empirical study rather than addressing the issue with a specific solution.

Such and Criado [28,29] propose an agent-based alternative to this mechanism. Mosca et al. [17] present a utility and value-driven approach which simultaneously addresses explainability, role-agnosticism and adaptability requirements for a multi-party content sharing system. In their follow-up research [18], they present empirical evidence demonstrating that these agents surpass existing methods by effectively balancing individual utility, value alignment, and user satisfaction regarding system explainability. While these studies try to construct generalizable policies for the users in terms of audience for a single multi-party content, they do not address the features of the content which can lead those preferences. Although Erdoğan et. al. [8] propose an approach that can model the user's and the other preferences with a probability for willingness to share, the study focuses on the individual facial expressions rather than all of the features can be present in a content.

While the features and the user preferences are assumed to be given to the agent within this study, the modularity of the framework we built enables the integration of new components that can handle feature extraction and user modeling based on these features in the future.

Such a component can also enable different techniques to determine the user's overall valuation for the content to state a decision. Instead of assuming mean of all non-zero features, more precise measurements can be done with weighted averages of features, based on the user's tendencies. In the same manner, utilities and the reward function might be subject to refinement in another research direction. More importantly, the overall framework can be evaluated by incorporating human interactions in future research.

Acknowledgments. This research is financially supported by the Hybrid Intelligence Center, a 10-year programme funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research (grant number 024.004.022). Hüseyin Aydın and Alper Demir are supported by the Scientific and Technological Research Council of Turkey, through BIDEB 2219 International Postdoctoral Research Scholarship Program.

References

1. Albrecht, S.V., Christianos, F., Schäfer, L.: Multi-agent reinforcement learning: Foundations and modern approaches. MIT Press (2024)
2. Aycı, G., Özgür, A., Şensoy, M., Yolum, P.: Can we explain privacy? IEEE Internet Computing **27**(4), 75–80 (2023)
3. Ayci, G., Sensoy, M., Özgür, A., Yolum, P.: Uncertainty-aware personal assistant for making personalized privacy decisions. ACM Transactions on Internet Technology **23**(1), 1–24 (2023)

4. Bianchi, F., Zorzi, E., Castellini, A., Simao, T., Spaan, M.T., Farinelli, A., et al.: Scalable safe policy improvement for factored multi-agent mdps. PROCEEDINGS OF MACHINE LEARNING RESEARCH **235**, 3952–3973 (2024)
5. Clarke, E.H.: Multipart pricing of public goods. Public Choice **11**, 17–33 (1971), http://www.jstor.org/stable/30022651
6. Colnago, J., Feng, Y., Palanivel, T., Pearman, S., Ung, M., Acquisti, A., Cranor, L.F., Sadeh, N.: Informing the design of a personalized privacy assistant for the internet of things. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2020)
7. Degris, T., Sigaud, O., Wuillemin, P.H.: Learning the structure of factored markov decision processes in reinforcement learning problems. In: Proceedings of the 23rd international conference on Machine learning. pp. 257–264 (2006)
8. Erdogan, E., Aydın, H., Dignum, F., Verbrugge, R., Yolum, P.: Mitigating privacy conflicts with computational theory of mind. In: 24th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2025. pp. 695–703. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS) (2025)
9. Feng, F., Magliacane, S.: Learning dynamic attribute-factored world models for efficient multi-object reinforcement learning. Advances in Neural Information Processing Systems **36**, 19117–19144 (2023)
10. Filipczuk, D., Baarslag, T., Gerding, E.H., Schraefel, M.: Automated privacy negotiations with preference uncertainty. Autonomous Agents and Multi-Agent Systems **36**(2), 49 (2022)
11. Guestrin, C., Koller, D., Parr, R.: Multiagent planning with factored mdps. Advances in neural information processing systems **14** (2001)
12. Guestrin, C., Koller, D., Parr, R.: Solving factored pomdps with linear value functions. In: Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01) workshop on Planning under Uncertainty and Incomplete Information. pp. 67–75. Citeseer (2001)
13. Guestrin, C., Koller, D., Parr, R., Venkataraman, S.: Efficient solution algorithms for factored mdps. Journal of Artificial Intelligence Research **19**, 399–468 (2003)
14. Kearns, M., Koller, D.: Efficient reinforcement learning in factored mdps. In: IJCAI. vol. 16, pp. 740–747. Citeseer (1999)
15. Kurtan, A.C., Yolum, P.: Assisting humans in privacy management: an agent-based approach. Autonomous Agents and Multi-Agent Systems **35**(1), 7 (2021)
16. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. nature **518**(7540), 529–533 (2015)
17. Mosca, F., Such, J., McBurney, P.: Towards a value-driven explainable agent for collective privacy. In: Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems (2020)
18. Mosca, F., Such, J.M.: Elvira: An explainable agent for value and utility-driven multiuser privacy. In: Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems. pp. 916–924 (2021)
19. Peng, B., Rashid, T., Schroeder de Witt, C., Kamienny, P.A., Torr, P., Böhmer, W., Whiteson, S.: Facmac: Factored multi-agent centralised policy gradients. Advances in Neural Information Processing Systems **34**, 12208–12221 (2021)
20. Sallans, B., Hinton, G.E.: Reinforcement learning with factored states and actions. Journal of Machine Learning Research **5**(Aug), 1063–1088 (2004)
21. Schoeman, F.D.: Philosophical dimensions of privacy: An anthology. Cambridge University Press (1984)

22. Shani, G.: Task-based decomposition of factored pomdps. IEEE transactions on cybernetics **44**(2), 208–216 (2013)
23. Simão, T.D., Spaan, M.T.: Safe policy improvement with baseline bootstrapping in factored environments. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33, pp. 4967–4974 (2019)
24. Spyromitros-Xioufis, E., Papadopoulos, S., Popescu, A., Kompatsiaris, Y.: Personalized privacy-aware image classification. In: Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. p. 71–78. ICMR '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2911996.2912018, https://doi.org/10.1145/2911996.2912018
25. 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 (2009)
26. Stöver, A., Hahn, S., Kretschmer, F., Gerber, N.: Investigating how users imagine their personal privacy assistant. Proceedings on Privacy Enhancing Technologies (2023)
27. Strehl, A.L., Diuk, C., Littman, M.L.: Efficient structure learning in factored-state mdps. In: AAAI. vol. 7, pp. 645–650 (2007)
28. Such, J.M., Criado, N.: Resolving multi-party privacy conflicts in social media. IEEE Transactions on Knowledge and Data Engineering **28**(7), 1851–1863 (2016)
29. Such, J.M., Criado, N.: Multiparty privacy in social media. Communications of the ACM **61**(8), 74–81 (2018)
30. Such, J.M., Porter, J., Preibusch, S., Joinson, A.: Photo privacy conflicts in social media: A large-scale empirical study. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. p. 3821–3832. CHI '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3025453.3025668, https://doi.org/10.1145/3025453.3025668
31. Thomas, K., Grier, C., Nicol, D.M.: unfriendly: Multi-party privacy risks in social networks. In: Privacy Enhancing Technologies: 10th International Symposium, PETS 2010, Berlin, Germany, July 21-23, 2010. Proceedings 10. pp. 236–252. Springer (2010)
32. Tonge, A., Caragea, C.: Privacy-aware tag recommendation for accurate image privacy prediction. ACM Transactions on Intelligent Systems and Technology (TIST) **10**(4), 1–28 (2019)
33. Ulusoy, O., Yolum, P.: Pano: Privacy auctioning for online social networks. In: Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems. pp. 2103–2105 (2018)
34. Ulusoy, O., Yolum, P.: PANOLA: A personal assistant for supporting users in preserving privacy. ACM Transactions on Internet Technology **22**(1) (2021)