

Incentive Analysis for Cooperative Interactive Multiview Video Streaming

Bo Hu *, H. Vicky Zhao * and Gene Cheung #

* University of Alberta, # National Institute of Informatics

Abstract

In interactive multiview video streaming (IMVS), users can periodically select one out of many captured views available for observation. In single-view video streaming, cooperative strategies where peers share received packets of the same video have proven to be effective in reducing server's upload burden, and incentive mechanisms are designed to stimulate user cooperation. However, exploiting user cooperation in higher dimensional IMVS is difficult, since users watching different views makes it difficult to establish partnership, and users switching views frequently and independently makes it difficult to maintain partnership over time. In this paper, we use a multiview video frame structure for IMVS to support cooperative view-switching, where peers may help each other even if they are observing different views. We then model peers' interaction as an indirect reciprocity game, where each user is assigned a reputation level. To gain a higher reputation level, users help others, which in turn leads to a higher likelihood to receive others' help later. In this work, we focus on how view switching, the key feature of IMVS, affects user cooperation. By modeling users' decision making as a Markov decision process, our analysis shows that users tend to cooperate at some views but not others: given peers can predict their future view navigation paths probabilistically, for a peer who is likely to enter a view-switching path not requiring others' help, he also has less incentive to cooperate. Furthermore, we observe that the game may have multiple Nash Equilibria corresponding to different cooperation levels, e.g., users cooperate at all views in the full cooperation equilibrium, while users only cooperate at certain views in the partial cooperation equilibrium. The particular equilibrium the game will converge to depends on the initial cooperation level of the game. To stimulate user cooperation at all views, we propose a Pay-for-Cooperation (PfC) scheme at the beginning of the game to drive the game to the full cooperation equilibrium to improve system efficiency. Our simulation results show the effectiveness of PfC.

I. INTRODUCTION

Multiview video refers to the simultaneous capturing of multiple videos of the same scene by a large array of closely spaced cameras (e.g. more than 100 cameras in [1]) from different viewpoints. In the emerging interactive multiview view streaming (IMVS) service [2], a client can periodically select one

out of many captured views available for observation as the video is played back in time. In response, a server sends only pre-encoded data for the single requested view (rather than all the captured views) to lower streaming rate. However, in this client-server model based multiview video streaming, the server can be easily overloaded when many users request service at the same time.

In single-view video streaming, to ease server's burden to upload the same video to many users, user cooperation [3]–[5] has been exploited where peers share received packets of the same video, so that a single server can serve a large number of clients. However, exploiting user cooperation in high dimensional IMVS is difficult. First, a small number of peers in a local area are likely to watch different views among a large number of available views, making it difficult for a peer to find partners watching the exact same view to cooperate. In the literature, the work in [2] designed frame structures using distributed source coding (DSC) [6] (based on information theory developed in [7], [8]) for IMVS to enable users watching different views to help each other switch view and to achieve bandwidth-efficient view switching. The work in [9] used DSC for both view switching and cooperative packet loss recovery in a WWAN multiview video multicast system. Following the above works, in this work, we use a DSC based multiview video coding structure to facilitate *cooperative view-switching*, where a user can get help from another peer who has the video data in neighboring views.

Second, even if a peer can find cooperative partners at a time instance, it is difficult to maintain such partnership over time, since peers switch views frequently and independently. This makes it difficult to design incentive mechanisms for cooperation stimulation. In the literature, incentive mechanisms for single view video streaming system have been studied extensively. The works in [10]–[14] proposed direct reciprocity mechanisms, where a user i helps another user j because user j helped i previously. The direct reciprocity schemes are effective when users expect to maintain a long-term partnership, which may not be suitable for IMVS, where users usually have short-term partnerships. The work in [15]–[17] proposed payment-based schemes, where a user earns *points* by uploading data to others and pays *points* to request others' data. These schemes work for the scenario where users change partners frequently. However, users make decisions based on the short-term payoffs only. If the gain from others' payment is smaller than the cost to help, users will not cooperate. The works in [18]–[21] proposed indirect reciprocity games, where each user is assigned a reputation level. A user helps others to gain a higher reputation level, which in turn leads to a higher likelihood to receive others' help.

In this work, similar to [18], [19], we also model user interaction in the IMVS as an indirect reciprocity game, since it effectively stimulates user cooperation when users change partners frequently and when the short-term payoff is negative. Note that in IMVS, users switch views and update reputations frequently,

and thus may change their strategies from time to time. To address this issue, we model their action selection as a Markov decision process (MDP) [22], and study users' optimal action selection at each particular view and reputation to maximize their entire lifetime utilities.

We then summarize the major contributions of this work as follows:

- To the best of our knowledge, this is the first work that provides theoretical analysis on how the multiview video affects users' decision making in cooperative video streaming. In this work, we observe that users may cooperate at some views but not others. This is because peers can predict their future view navigation paths probabilistically, and thus, can estimate the probability that he will need others' help. If a peer is at a view leading to a view-navigation path with lower probability of requiring others' help, he also has less incentive to cooperate.
- We show that a large number of reputation levels provide higher incentive for user cooperation. We first observe that the 2-level reputation system is memoryless, and each user makes decisions based on the short-term utility only. Thus, if a user is at a view where cooperation only results in a negative short-term utility, he will not cooperate. In the R -level reputation system with $R \geq 3$, a user needs to take his future utility into consideration, and may still cooperate even if the short-term payoff is negative. This is because cooperation helps him maintain a high reputation and get others' help in the future. If the future payoffs can compensate his current loss, he will still cooperate.
- We observe that the game may have multiple Nash Equilibria corresponding to different cooperation levels (e.g., users cooperate at all views in the full cooperation equilibrium, while users only cooperate at certain views in the partial cooperation equilibrium). The particular equilibrium the game will converge to depends on the initial state of the game. Thus, we first analyze the sufficient condition for the game to converge to the full cooperation equilibrium. We then propose a pay-for-cooperation (PFC) scheme at the beginning of the game to drive the game to the desired full cooperation equilibrium with the optimal system performance
- We also study the impact of user membership dynamics on user cooperation and system performance. From our theoretical analysis and simulations, we observe that as long as the percentage of users adopting full cooperation strategy exceeds a threshold, full cooperation is a dominant strategy for all users, and they will all cooperate.

The outline of the paper is as follows. We overview the formulation of the IMVS streaming problem, the indirect reciprocity game and users' decision making with MDP in Section II. We analyze users optimal action selection in Section III. We study the Nash Equilibria of the game in Section IV and V. We present simulation results and conclusions in Section VI and VII, respectively.

II. SYSTEM MODEL

In this work, we consider an IMVS system, where a scene is captured by a large one-dimensional array of M evenly spaced cameras. A server compresses video of each view into coding segments of K frames each, and provides IMVS service to a group of N users who are synchronized in playback time. Once a user selects a view, he remains in this view for one segment of K consecutive frames. At the end of this segment, he can switch to another view.

Based on this IMVS system, in this section we first describe an interaction model that captures users' view switching behavior, and a multiview video coding structure that facilitates cooperative view switching among peers. We then propose an indirect reciprocity game to stimulate user cooperation. Finally, we model users' optimal action selection as an MDP. The frequently used notations are listed in Table I.¹

A. View Switching Model

Views are divided into two categories: *anchor views* and *normal views*. Suppose that there are n_a anchor views, which evenly divide normal views into $(n_a + 1)$ view sets of $n_n = (M - n_a)/(n_a + 1)$ views per set. When seeking interested views, a user first browses views coarsely through anchor views. Once he reaches an interested anchor view, he can switch to neighboring normal views to refine view selection. This coarse-to-fine browsing structure is typical in other media navigation applications as well, such as Google Maps². In this work, we assume that users switch interested views frequently. After finding an interested view and remaining for one segment, they will likely seek another interested view for the next segment. Thus, anchor views are more frequently selected (more popular) than normal views.

At each view, we consider the scenario where a user can only switch to his nearby anchor views with probability P_a , or nearby normal views with probability $(1 - P_a)$. Specifically, we model the view transition as a discrete time Markov chain, and construct a $M \times M$ transition matrix \mathbf{T} , where $\mathbf{T}(v, v')$ is the probability of a user selecting view v' in the next segment after viewing v :

$$\mathbf{T}(v, v') = \begin{cases} P_a/|\mathcal{Z}_a| & \text{if } v' \in \mathcal{Z}_a, \\ (1 - P_a)/|\mathcal{Z}_n| & \text{if } v' \in \mathcal{Z}_n, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where \mathcal{Z}_a is the set including v 's nearby anchor views, and \mathcal{Z}_n is the set including v 's nearby normal views. Specifically, if v is an anchor view, \mathcal{Z}_a includes v 's left/right closest anchor views and v itself, and

¹In this table, \mathcal{N}^+ denotes the set of positive integers as in the convention.

²<http://maps.google.com>

TABLE I
FREQUENTLY USED NOTATIONS

\mathcal{A}	$= \{1, 2, \dots, R + 1\}$	The action space
\mathcal{R}	$= \{1, 2, \dots, R\}$	The reputation space
$\bar{\mathcal{R}}$	$= \{t_r - 1, t_r, \dots, R\}$	The reputation set with reputations no less than $t_r - 1$
\mathcal{S}	$= \mathcal{R} \times \mathcal{V}$	The state space in the MDP
\mathcal{V}	$= \{1, 2, \dots, M\}$	The view space
$\underline{\mathcal{V}}, \bar{\mathcal{V}}$	$\subset \mathcal{V}$	The low utility view set, the high utility view set
\mathbf{Q}	2×2 matrix	The reputation updating matrix
\mathbf{T}	$M \times M$ matrix	The view transition matrix
\mathbf{v}	$1 \times M$ vector	The steady state view distribution
\mathbf{x}	$1 \times R$ vector	The reputation distribution
c	$\in [0, +\infty)$	The cost of helping to upload a reconstructed frame
g_v	$\in [0, +\infty)$	The expected short-term gain if the view switching starts from view v and helpers always help
t_r	$\in \mathcal{R}$	The reputation threshold to differentiate beneficial users from non-beneficial users
y	$\in [0, 1]$	The percentage of beneficial users in the network
y_{in}	$\in [0, 1]$	The percentage of selected users in the PfC scheme
$G_{r,v}$	$\in [0, +\infty)$	The expected short-term gain from others' help with reputation r and view v
L	$\in \mathcal{N}^+$	The average interval between two consecutive requests received by a user
M	$\in \mathcal{N}^+$	The number of views
N	$\in \mathcal{N}^+$	The number of users
P_a	$\in [0, 1]$	The probability of switching to anchor views
$P_{r \rightarrow r'}^a$	$\in [0, 1]$	The reputation transition probability from r to r' by action a
R	$\in \mathcal{N}^+$	The highest reputation level
$U_{r,v}^a$	$\in (-\infty, +\infty)$	The expected short-term utility by taking action a at reputation r and view v
$W_{r,v}^\pi$	$\in (-\infty, +\infty)$	The lifetime utility at reputation r and view v by action policy π
$W_{r,v}^{a',\pi}$	$\in (-\infty, +\infty)$	The lifetime utility of the one-shot deviation to a' at reputation r and view v
η	$\in [0, 1]$	The discounting factor
π	$= \{a_s \in \mathcal{A} s \in \mathcal{S}\}$	The action policy
γ	$\in [0, 1]$	The discounting factor after $(t_r - 1)L$ segments

\mathcal{Z}_n includes v 's left/right adjacent normal view sets. If v is a normal view, \mathcal{Z}_a only includes v 's left/right closest anchor views, and \mathcal{Z}_n is the normal view set where v belongs. Given the one-step transition matrix \mathbf{T} , the l -step transition matrix is \mathbf{T}^l (\mathbf{T} raised to the l th power), where $\mathbf{T}^l(v, v')$ is the probability to transition to view v' in l segments after viewing v . The steady state view probability distribution is \mathbf{v} satisfying $\mathbf{v}\mathbf{T} = \mathbf{v}$, where $\mathbf{v}(v)$ is the probability that a user is at view v at the steady state.

For example, for a $M = 3$ views with a single anchor view in the middle, the one-step and two-step

Fig. 1. Example of our multiview video coding structure for $M = 3$ views, segment size $K = 3$. Circles, squares and diamonds denote I-, P- and DSC frames, respectively. Each frame $F_{\tau,v}$ is labeled by its frame index τ and view v .

transition matrices are

$$\mathbf{T} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 1 - P_a & P_a & 0 \\ (1 - P_a)/2 & P_a & (1 - P_a)/2 \\ 0 & P_a & (1 - P_a) \end{pmatrix} \end{matrix} \quad \text{and} \quad \mathbf{T}^2 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} (1 - P_a)(1 - \frac{P_a}{2}) & P_a & P_a \frac{1 - P_a}{2} \\ \frac{1 - P_a}{2} & P_a & \frac{1 - P_a}{2} \\ P_a \frac{1 - P_a}{2} & P_a & (1 - P_a)(1 - \frac{P_a}{2}) \end{pmatrix} \end{matrix} \quad (2)$$

respectively, and the steady state view distribution is $\mathbf{v} = [\frac{1 - P_a}{2}, P_a, \frac{1 - P_a}{2}]$.

B. Multiview Video Coding Structure and Cooperative View Switching

To address the problem of users having difficulty establishing partnership for cooperation in a high dimensional IMVS, we use a frame structure similar to that in [9]. It supports cooperative view switching, where users may cooperate with each other even if they are observing different views. Fig. 1 shows an example of the frame structure used in this work. Each view is encoded into segments of K frames. We encode the first segment using an intra-coded I-frame with $K - 1$ trailing P-frames. For the next segment, for view switching we encode the first frame $F_{K+1,v}$ into two versions. The first version is an intra-coded I-frame, which can be decoded independently. The second version is a *DSC frame* [6]. To encode the DSC frame, we use the I-frame of the same picture as target, and use at most three decoded P-frames

$F_{K,\max(1,v-1)}, \dots, F_{K,\min(M,v+1)}$ as predictors. As long as one of the predictor frames is available at the decoder buffer, the DSC frame can be correctly decoded, and the decoded frame is bit-by-bit equivalent to the frame decoded from the I-frame. Frame $F_{K+1,v}$ is followed by $K - 1$ trailing P-frames. The following segments have the same structure. In general, an I-frame is much larger than a DSC frame, and a DSC frame is slightly larger than a P-frame [6].

This structure can support cooperative view switching. Using Fig. 1 as an example, suppose that a peer i switches from view 1 to view 3 after the first segment. If another peer watches view 2 in the first segment and is willing to share the reconstructed frame $F_{3,2}$, then peer i only needs to ask the server for the DSC frame of $F_{4,3}$ and the following $(K - 1)$ trailing P-frames to reconstruct the video in view 3. If no one helps peer i (either no user watches view 2 or view 3 in the first segment, or the users who can help are not willing to help), then peer i has to request the I-frame of $F_{4,3}$ from the server.

In this work, we assume that the server's upload bandwidth is limited and expensive. Thus, it charges subscription fees from peers that pull video data from it, and α denotes the price for the transmission of each single bit from the server. As discussed above, when a peer switches to a non-adjacent view, if he can get help from others, he will download the last reconstructed frame in the previous segment from the helper for free, and will only download a DSC frame from the server instead of an I-frame. Thus, he can receive a gain of $\alpha(\text{size}_I - \text{size}_{DSC})$ for paying less to the server, where size_I and size_{DSC} denote the number of bits of one I-frame and one DSC frame, respectively. However, uploading a reconstructed frame will incur a cost to the helper due to the consumed bandwidth, CPU time, etc. In this work, we consider the scenario with homogeneous users who have the same cost to upload a frame. In the following discussion, without loss of generality, we normalize the gain of receiving a reconstructed frame to 1, and let c denote the normalized cost to upload a reconstructed frame to a peer.

C. Indirect Reciprocity Game

To stimulate user cooperation, we model their interaction as an indirect reciprocity game as follows.

1) *Peer Reputation and Interaction:* In this system, each peer i is assigned a discrete reputation $r_i \in \mathcal{R} = \{1, 2, \dots, R\}$, where a larger r_i indicates a higher reputation and peer i is more likely to receive others' help. Users' reputations change as they interact with each other. When a peer needs help, he first needs view information of other peers to find a helper. To implement this, we can either let peers exchange their view information and seek help in a distributed way, or have a central controller that tracks peers' up-to-date view information and assigns helpers to peers who need help. For simplicity, we assume that there is a trustworthy local agent close to the N peers, who tracks peers' view switching,

helps each peer find helpers, observes their interactions, and updates their reputations. Specifically, when peer j needs help, the local agent randomly selects peer i from peers that can help, and sends a request to peer i with peer j 's reputation r_j . Upon receiving a request, peer i will decide whether to help. In this work, we consider that peer i takes a threshold-based action $a_i \in \mathcal{A} = \{1, 2, \dots, R + 1\}$, with \mathcal{A} as the action space, and peer i will only cooperate with others with reputations higher than a_i , i.e., if $r_j \geq a_i$, peer i will help peer j . Otherwise, peer i will not help and peer j has to request the I-frame from the streaming server. In the extreme case where $a_i = R + 1$, peer i will not cooperate with anyone, while $a_i = 1$ means peer i is willing to help all users. With the threshold-based action, users take the same action for requests with different r_j 's, which reduces the decision complexity and simplifies the analysis for the optimal actions.

2) *Social Norm and Reputation Update*: Based on the observed interaction between peer i who receives the request and j who sends the request, the local agent updates i 's reputation following the pre-determined social norm [23] that defines reputation update rules, while j 's reputation remains the same. In this work, we use the social norm similar to one in [19], since it is effective in user cooperation stimulation.

Specifically, we first define a pre-determined threshold $1 < t_r \leq R$. If user i has reputation $r_i \geq t_r$, he has high reputation and is likely to get others' help. Thus, he is called a *beneficial* user. Otherwise, he is not likely to get others' help, and is a *non-beneficial* user. If $r_i \geq t_r - 1$ (i.e., user i is a beneficial user or may become a beneficial user after this interaction), i 's reputation is updated following the matrix,

$$\mathbf{Q} = \begin{matrix} & r_j \geq t_r & r_j < t_r \\ \begin{matrix} a_i \leq r_j, \text{ uploading} \\ a_i > r_j, \text{ not uploading} \end{matrix} & \begin{pmatrix} \min\{r_i + 1, R\} & 1 \\ 1 & \min\{r_i + 1, R\} \end{pmatrix} & \end{matrix}. \quad (3)$$

From (3), if user i cooperates with a beneficial user or denies the request from a non-beneficial user, i complies with the social norm, and he is rewarded by one-step increase of his reputation. Otherwise, he does not comply with the social norm, and he is punished with his reputation being lowered to 1. Thus, with (3), peers are encouraged to help beneficial users, but discouraged to help non-beneficial users.

If user i does not comply with the social norm and his reputation was lowered to 1, similar to [24], the system takes time to forgive his misbehavior. During the forgiveness period, his reputation will be increased by one-step for every time he receives a request, no matter how he responds to such request, until his reputation climbs to $t_r - 1$. Since his reputation is lower than $t_r - 1$ during the forgiveness period, he hardly receives others' help, which results in loss of utility. Here, t_r determines the duration of the forgiveness period. A larger t_r means that it takes a longer time for the system to forgive, and thus gives a harsher punishment. Note that when users make decisions, in addition to the social norm,

they also take other factors into consideration, which will be discussed in details in Section III, IV and V.

In this work, we assume that when users make decisions on the current requests, they also take the future interactions into consideration. Since they do not know with whom they will interact at a later time, the information of peers' reputation distribution \mathbf{x} helps in their decision making, where $\mathbf{x}(r)$ denotes the probability that a user has reputation $r \in \mathcal{R}$. Given that the local agent has the record of all peers' reputations at different time instances, it can estimate \mathbf{x} using

$$\mathbf{x}(r) = \frac{\sum_{t=1}^{T_c} \sum_{i=1}^N I[r_i^t = r]}{NT_c}, \quad \forall r \in \mathcal{R} \quad (4)$$

where T_c is the current segment index, and $I[\cdot]$ is the indicator function. The local agent broadcasts \mathbf{x} to all peers periodically to assist their decision making.

D. Optimal Action Selection with Markov Decision Process (MDP)

In our cooperative IMVS system, users may frequently switch views and their reputations may also change from time to time. Since they may take different actions at different views and reputations, we use MDP to track their strategy dynamics. In our IMVS, the game is played in a sequence of *stages*, and a stage represents an instance when a user receives a request and needs to make a decision. There are $L \geq 1$ segments of video playback between two neighboring stages. Fig. 2 shows an example where a user receives a request at segment t_1 and will receive another request two segments later at $t_2 = t_1 + 2$ with $L = 2$. Following the work in [19], to simplify the analysis, we let L be the average interval between two consecutive requests received by a user in our work.

Given the above defined sequence of stages, an MDP is defined as a four-tuple: the state space \mathcal{S} , the action space \mathcal{A} , the state transition probability \mathcal{P} and the expected short-term utility function U . In our IMVS, a state $s = (r, v)$ represents a user's reputation r and view v when he receives a request. In the following sections, we will interchangeably use s and (r, v) to denote a state. Hence, the state space is denoted as $\mathcal{S} = \mathcal{R} \times \mathcal{V}$, where $\mathcal{V} = \{1, \dots, M\}$ is the view space and $\mathcal{R} = \{1, 2, \dots, R\}$ is the reputation space. At each state (r, v) , a user can select action $a_{r,v}$ from the action space $\mathcal{A} = \{1, \dots, R + 1\}$. In the example in Fig. 2, there are $M = 3$ views and $R = 3$ reputation levels with $t_r = 3$. The state space includes a total of 9 states $\{(r, v)\}_{1 \leq r \leq 3, 1 \leq v \leq 3}$, and the action space includes 4 possible actions $\mathcal{A} = \{1, 2, 3, 4\}$.

A user receives a request at time t and he is at state (r, v) . He takes action $a_{r,v}$ and transitions to another state (r', v') with state transition probability $P_{(r,v) \rightarrow (r',v')}^{a_{r,v}}$ when he receives another request in

the next stage at time $t + L$. By taking action $a_{r,v}$, the user receives an expected short-term utility $U_{r,v}^{a_{r,v}}$, which contains two parts. First, this action may result in a frame upload to another peer, which incurs an expected cost $C^{a_{r,v}}$ immediately at time t . In addition, this action $a_{r,v}$ results in the update of the user's reputation to r' at time $t + 1$, and he keeps reputation r' from time $t + 1$ to $t + L$ until he receives another request. This updated reputation affects the probability of receiving others' help in the following L segments (i.e., from time $t + 1$ to $t + L$), and thus his gain. Given the updated reputation r' and the view v that he is watching at time t , let $\theta(t + l)$ be the expected gain he receives at time $t + l$ for $1 \leq l \leq L$, and define $G_{r',v} = \sum_{l=1}^L \eta^l \theta(t + l)$ as the expected short-term gain, where $\eta \in (0, 1)$ is the discounting factor that quantifies how much a user cares about his future payoff. Then, the expected short-term utility function is $U_{r,v}^{a_{r,v}} = G_{r',v} - C^{a_{r,v}}$. In Fig. 2, when the user receives a request at time t_1 , he is at state $(r = 3, v = 1)$. He selects an action $a_{3,1} \in \{1, 2, 3, 4\}$, receives an expected short-term utility $U_{3,1}^{a_{3,1}}$, and transitions to another state $s' = (r', v')$ with probability $P_{(3,1) \rightarrow (r',v')}$ when he receives another request in the next stage at time $t_2 = t_1 + L$. This process is repeated till the end of the game.

The action policy in MDP is defined as $\pi = \{a_{r,v} \in \mathcal{A} | (r, v) \in \mathcal{S}\}$ that defines the action $a_{r,v}$ at each state (r, v) . The goal of MDP is to find the optimal action policy that maximizes the expected lifetime utility, which is recursively defined as

$$W_{r,v}^\pi = U_{r,v}^{a_{r,v}} + \eta^L \sum_{(r',v') \in \mathcal{S}} P_{(r,v) \rightarrow (r',v')}^{a_{r,v}} W_{r',v'}^\pi \quad \forall (r, v) \in \mathcal{S}, \quad (5)$$

where the second term denotes the user's aggregate lifetime utility after the next stage. In this work, we consider the scenario with homogenous users who have the same cost, and thus take the same optimal action policies at the Nash Equilibrium. To examine whether a policy π gives a Nash Equilibrium, for each user, with the assumption that all others take π , if π also maximizes his lifetime utility, he has no incentive to deviate and π is an equilibrium policy.

III. MDP ANALYSIS AND EQUILIBRIUM ACTION POLICY DISCUSSION

In this section, we first analyze the state transition probability, the expected short-term utility and the lifetime utility of the MDP. We then discuss the properties of equilibrium action policies.

A. MDP Analysis

1) *State Transition Probability*: We first analyze the probability that a user i transitions from state (r, v) to (r', v') after L segments of video playback in the next stage. Note that in our IMVS, the view and reputation transition probabilities are independent. Given the one-step view transition matrix \mathbf{T} in

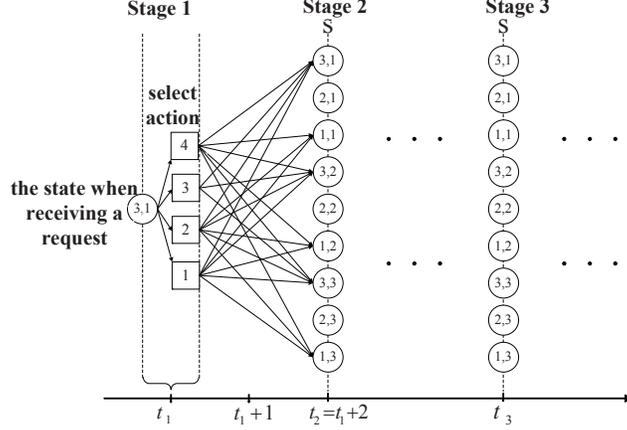


Fig. 2. An example of MDP with $M = 3$ views and $R = 3$ levels in the reputation system. All circles represent states, while all squares represent actions. The average interval between two consecutive requests is $L = 2$.

(1), the probability that user i transitions from view v to v' in L segments is $\mathbf{T}^L(v, v')$. In the example in Fig. 2 with $L = 2$, user i is at view 1 at time t_1 . From (2), after $L = 2$ segments, he will transition to view 1, 2 and 3 with probabilities $\mathbf{T}^2(1, 1) = (1 - P_a)(1 - \frac{P_a}{2})$, $\mathbf{T}^2(1, 2) = P_a$ and $\mathbf{T}^2(1, 3) = P_a \frac{1 - P_a}{2}$, respectively.

To find the reputation transition probability, suppose that user i at state (r, v) takes action $a_{r,v}$ when responding to user j 's request, and user i 's reputation is updated to r' . From Section II-C2, when user i 's reputation is $r < t_r - 1$, his reputation is always increased by 1, and we have $P_{r \rightarrow r+1}^{a_{r,v}} = 1$ and $P_{r \rightarrow r'}^{a_{r,v}} = 0$ for $r' \neq r + 1$. When $r \geq t_r - 1$, user i 's reputation is updated using \mathbf{Q} in (3) and the updated reputation is either $\min\{r + 1, R\}$ or 1. The updated reputation is 1 when i denies the request of a beneficial user (i.e., $a_{r,v} > r_j \geq t_r$) or i cooperates with a non-beneficial user (i.e., $a_{r,v} \leq r_j < t_r$). In addition, $P_{r \rightarrow \min\{r+1, R\}}^{a_{r,v}} = 1 - P_{r \rightarrow 1}^{a_{r,v}}$ and $P_{r \rightarrow r'} = 0$ for all other values of r . In this work, we focus on user i 's optimal actions to maximize his average utility for any requester he may encounter. The requester's reputation r_j is assumed to follow the distribution \mathbf{x} . Thus, the reputation transition probability is:

$$P_{r \rightarrow 1}^{a_{r,v}} = \begin{cases} \sum_{r_j=a_{r,v}}^{t_r-1} \mathbf{x}(r_j) & r \geq t_r - 1, a_{r,v} < t_r \text{ (helping a non-beneficial user),} \\ \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) & r \geq t_r - 1, a_{r,v} > t_r \text{ (not helping a beneficial user),} \\ 0 & \text{otherwise,} \end{cases}$$

$$P_{r \rightarrow \min\{r+1, R\}}^{a_{r,v}} = 1 - P_{r \rightarrow 1}^{a_{r,v}}, \text{ and } P_{r \rightarrow r'}^{a_{r,v}} = 0 \forall r' \neq 1, r' \neq \min\{r + 1, R\}. \quad (6)$$

In the example in Fig. 2 (with $t_r = 3$), if the user takes action 1 at state $(r = 3, v = 1)$ (with $r > t_r - 1$) and cooperates with all users, he will help a non-beneficial user and his reputation will be lowered to 1

with probability $(\mathbf{x}(1) + \mathbf{x}(2))$. The probability to transition to other reputation levels can be calculated similarly.

In summary, the state transition probability is $P_{(r,v) \rightarrow (r',v')}^{a_{r,v}} = \mathbf{T}^L(v, v') \cdot P_{r \rightarrow r'}^{a_{r,v}}$. In the example in Fig. 2, by taking action 1, he will transition to state $(r' = 1, v' = 3)$ in the next stage with probability $P_a \frac{1-P_a}{2} (\mathbf{x}(1) + \mathbf{x}(2))$, and the probability to transition to other states can be calculated similarly.

2) *Expected Short-term Utility and Lifetime Utility*: We now analyze the expected short-term and the lifetime utility functions in (5), and start with the expected short-term utility. From the discussion in Section II-D, the expected short-term utility $U_{r,v}^{a_{r,v}}$ contains two parts: the expected immediate cost $C^{a_{r,v}}$ and the expected short-term gain $G_{r',v}$. With the assumption that the requester j 's reputation follows the distribution \mathbf{x} , the probability that he uploads the frame with action $a_{r,v}$ is $\sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j)$ (with $\sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) = 0$ if $a_{r,v} = R + 1$). Therefore, his expected cost is $C^{a_{r,v}} = c \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j)$.

To analyze $G_{r',v}$, note that taking action $a_{r,v}$ makes user i 's reputation updated to r' at time $t + 1$ and he keeps r' for the following L segments (i.e., from time $t + 1$ to time $t + L$). We then derive the gain he receives at each time $t + l$ (with $1 \leq l \leq L$) given that he watches view v at time t . At time $t + l$, user i receives a positive normalized gain 1 if and only if he switches to a non-adjacent view (i.e., he needs help) and there is a user who can and is willing to help him. Otherwise, his gain is 0. Let $P_{r',v}(t + l)$ denote the probability that user i switches to a non-adjacent view at time $t + l$ and there is a user who can and is willing to help him. Thus, we have $G_{r',v} = \sum_{l=1}^L \eta^l P_{r',v}(t + l)$.

We then derive $P_{r',v}(t + l)$ step by step. Let $v_i(t + l)$ denote the view that user i watches at time $t + l$. For a given view v' , let $\mathcal{V}_{v'} \triangleq \{\max(v' - 1, 1), v', \min(M, v' + 1)\}$ be the set including all adjacent views of v' . In the example in Fig. 2, $\mathcal{V}_1 = \{1, 2\}$, $\mathcal{V}_2 = \{1, 2, 3\}$ and $\mathcal{V}_3 = \{2, 3\}$. Then, we have

$$P_{r',v}(t + l) = \sum_{v'=1}^M P[\mathbb{H}_h | \mathbb{H}_1(v')] P[\mathbb{H}_1(v')] P[(v_i(t + l) = v', v_i(t + l - 1) \notin \mathcal{V}_{v'} | v_i(t) = v)], \quad (7)$$

where $P[v_i(t + l) = v', v_i(t + l - 1) \notin \mathcal{V}_{v'} | v_i(t) = v]$ is the probability that given that user i is at view v at time t , he switches to view v' at time $t + l$ from a non-adjacent view and needs help, $\mathbb{H}_1(v')$ is the event that there is at least one helper who can help user i switch to view v' at time $t + l$ (i.e., there is at least one user who is watching a neighboring view of v' at time $t + l - 1$), and \mathbb{H}_h is the event that the selected helper is willing to help. Note that

$$\begin{aligned} & P[v_i(t + l) = v', v_i(t + l - 1) \notin \mathcal{V}_{v'} | v_i(t) = v] \\ &= \sum_{v'' \notin \mathcal{V}_{v'}} P[v_i(t + l) = v' | v_i(t + l - 1) = v''] P[v_i(t + l - 1) = v'' | v_i(t) = v] = \sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}(v'', v') \mathbf{T}^{l-1}(v, v''). \end{aligned} \quad (8)$$

To find $P[\mathbb{H}_1(v')]$, we actually derive the probability that at least one of the rest $N - 1$ users watch views in $\mathcal{V}_{v'}$ at time $t + l - 1$. Given the stationary view distribution \mathbf{v} , we have

$$P[\mathbb{H}_1(v')] = 1 - \left(1 - \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{v}(v'')\right)^{N-1}. \quad (9)$$

To find the probability that the selected helper k is willing to help, helper k will help user i if user i 's current reputation r' is larger than or equal to helper k 's decision $a_{r_k, v_k(t+l)}$, which depends on helper k 's reputation r_k and view $v_k(t+l)$ at time $t+l$ when k receives the request. Therefore, we have

$$\begin{aligned} P[\mathbb{H}_h | \mathbb{H}_1(v')] &= \sum_{r_k=1}^R \sum_{v_k(t+l)=1}^M \mathbf{x}(r_k) \mathbf{p}_{v'}(v_k(t+l)) I[a_{r_k, v_k(t+l)} \leq r'], \\ \text{where } \mathbf{p}_{v'}(v_k(t+l)) &= P[v_k(t+l) | v_k(t+l-1) \in \mathcal{V}_{v'}] \\ &= \sum_{v'' \in \mathcal{V}_{v'}} P[v_k(t+l) | v_k(t+l-1) = v''] P[v_k(t+l-1) = v'' | v_k(t+l-1) \in \mathcal{V}_{v'}] \\ &= \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{T}(v'', v_k(t+l)) \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})}. \end{aligned} \quad (10)$$

Here, the helper k 's reputation is assumed to follow reputation distribution \mathbf{x} , \mathbf{v} is the steady state view distribution, and $\mathbf{p}_{v'}(v_k(t+l))$ is the probability that given helper k watches a neighboring view of v' at time $t+l-1$, he switches to view $v_k(t+l)$ at time $t+l$. Therefore, we have

$$G_{r', v} = \sum_{l=1}^L \eta^l \sum_{v'=1}^M P[\mathbb{H}_h | \mathbb{H}_1(v')] P[\mathbb{H}_1(v')] P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v]. \quad (11)$$

Based on the above discussion, $G_{r', v}$ is affected by user i 's view navigation path in the next L segments starting from view v . If he has a low probability to switch to non-adjacent views during the next L segments, he will also tend to have a small $G_{r', v}$. It is also easy to observe that $G_{r', v}$ is an increasing function of r' , since a higher r' gives a higher probability to get others' help. Note that in (11), if $P[\mathbb{H}_h | \mathbb{H}_1(v')]$ is always 1, that is, helpers are always willing to help user i , then $G_{r', v}$ becomes

$$G_{r', v} = g_v \triangleq \sum_{l=1}^L \eta^l \sum_{v'=1}^M P[\mathbb{H}_1(v')] P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v]. \quad (12)$$

Thus, g_v is a user's maximum expected short-term gain if he starts view switching from v , and always receives help when needed. Based on the above analysis, together with the fact that the updated r' can only be $\min\{r+1, R\}$ or 1, we can derive user i 's expected short-term utility after taking action $a_{r, v}$ as

$$U_{r, v}^{a_{r, v}} = -c \sum_{r_j = a_{r, v}}^R \mathbf{x}(r_j) + (1 - P_{r \rightarrow 1}^{a_{r, v}}) G_{\min(r+1, R), v} + P_{r \rightarrow 1}^{a_{r, v}} G_{1, v}. \quad (13)$$

Thus, given the action policy π , following (5), we have the lifetime utility $W_{r, v}^\pi$ as

$$W_{r, v}^\pi = U_{r, v}^{a_{r, v}} + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') \left[(1 - P_{r \rightarrow 1}^{a_{r, v}}) W_{\min(r+1, R), v'}^\pi + P_{r \rightarrow 1}^{a_{r, v}} W_{1, v'}^\pi \right]. \quad (14)$$

B. Discussion on the Equilibrium Policies

1) *Elimination of Non-Equilibrium Policies:* We now derive the Nash Equilibrium action policies. In our MDP, the size of the state space is $|\mathcal{R}||\mathcal{V}| = RM$, and we have $(R+1)^{RM}$ possible action policies. To avoid examining all these policies, we need to first eliminate non-equilibrium ones using Theorem 1, which we will discuss one by one in the following.

Theorem 1: In an equilibrium action policy π ,

- a) For all $r < t_r - 1$ and all $v \in \mathcal{V}$, $a_{r,v} = R + 1$.
- b) For all $r \geq t_r - 1$ and all $v \in \mathcal{V}$, $a_{r,v} \in \{t_r, R + 1\}$.
- c) For any view v , a user will take the same action for all reputations $r \geq t_r - 1$, i.e., $a_{t_r-1,v} = \dots = a_{R,v}$.

• *Proof of Theorem 1a):* Theorem 1a) says that if a user is not a beneficial user and cannot become a beneficial user after this decision, then he will not cooperate no matter which view he is watching. This is because when $r < t_r - 1$, for any action he takes, his reputation will be always increased by one, while $a_{r,v} = R + 1$ gives zero cost since he will not help anyone. Thus, $a_{r,v} = R + 1$ dominates the other actions when $r < t_r - 1$.

• *Proof of Theorem 1b):* Theorem 1b) says that if a user is a beneficial user or may become a beneficial user after this decision, then he will either cooperate with beneficial users (i.e., $a = t_r$) or do not cooperate with anyone (i.e., $a = R + 1$). It takes two steps to prove this. We will first show that the action $a_{r,v} = t_r$ dominates all actions $a_{r,v} < t_r$. We then show that any action policy with action $t_r + 1 \leq a_{r,v} \leq R$ cannot be an equilibrium action policy. From these two results, $a_{r,v}$ can only be t_r or $R + 1$.

We first compare the action $a_{r,v} = t_r$ with $a_{r,v} < t_r$ in terms of the incurred cost $C^{a_{r,v}}$ and the updated reputation r' . First, with action $a_{r,v} = t_r$, the expected cost to upload is $C^{t_r} = c \sum_{r_j=t_r}^R \mathbf{x}(r_j)$, and with $a_{r,v} < t_r$, the cost to upload is $C^{a_{r,v} < t_r} = c \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) \geq C^{t_r}$. Second, with action $a_{r,v} = t_r$, from (6), the user's reputation is rewarded with one-step increase with probability 1. However, with $a_{r,v} < t_r$, his reputation is rewarded with one-step increase with probability $1 - (\sum_{r_j=a_{r,v}}^{t_r-1} \mathbf{x}(r_j)) \leq 1$. Thus, $a_{r,v} = t_r$ gives a lower cost but a higher probability to be rewarded with one-step increase of his reputation. Thus, $a_{r,v} = t_r$ dominates all $a_{r,v} < t_r$, and we should have $a_{r,v} \geq t_r$ in the equilibrium.

Then, to show that action $t_r + 1 \leq a_{r,v} \leq R$ cannot be in an equilibrium policy, we use the *One-shot Deviation Principle* [25], which says that an action policy is an equilibrium if and only if no one can gain by one-shot deviation when others keep this policy unchanged. Here, one-shot deviation means taking a different action rather than the one defined in the action policy only for the current response to a request, but still following the policy in future responses. We have the following proposition with proof in Appendix A.

Proposition 1: For a policy π with $t_r + 1 \leq a_{r,v} \leq R$ for reputation $r \geq t_r - 1$, one-shot deviation to either $a'_{r,v}=t_r$ or $a'_{r,v}=R+1$ gives a higher lifetime utility. Thus, this π cannot be an equilibrium policy.

From Proposition 1 and the fact that $a_{r,v} = t_r$ dominates all $a_{r,v} < t_r$, $a_{r,v}$ can only be t_r or $R + 1$ in an equilibrium policy, i.e., either cooperate with all beneficial users or do not cooperate with anyone.

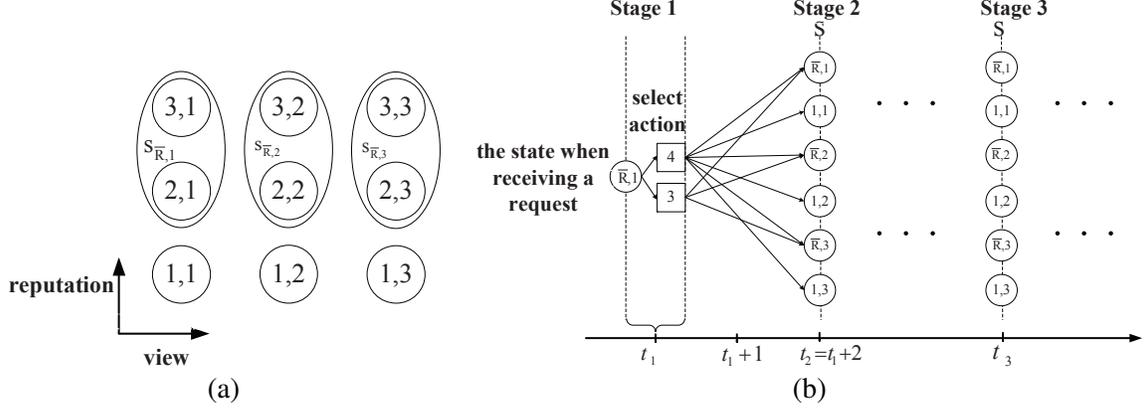


Fig. 3. Example of the state partition and aggregation with 3 views and 3-level reputation system (with $t_r = 3$). (a) The state partition. For each v , the states $(3, v)$ and $(2, v)$ forms a subspace $S_{\bar{R},v}$ and state $(1, v)$ forms a subspace $\{(1, v)\}$ with a single element. (b) The MDP after aggregating the state space $S_{\bar{R},v}$ as one state (\bar{R}, v) , and the corresponding action can only be selected from $\{t_r, R + 1\} = \{3, 4\}$.

• *Proof of Theorem 1c*): We first define a reputation subspace $\bar{\mathcal{R}} = \{r | t_r - 1 \leq r \leq R\}$ including all reputations no less than $t_r - 1$. Then, we define a state subspace $S_{\bar{\mathcal{R}},v} = \{(r, v) | r \in \bar{\mathcal{R}}, v \in \mathcal{V}\}$ that includes all states with view v and reputations no less than $t_r - 1$. Theorem 1c) says that for all states in $S_{\bar{\mathcal{R}},v}$, a user should take the same action, i.e., for any view v , we have $a_{t_r-1,v} = a_{t_r,v} = \dots = a_{R,v}$ in an equilibrium policy. To prove this, we use the concept of *Bisimilarity* [26], which is defined below.

Definition 1: (Bisimilarity) In an MDP, suppose that the state space is divided into m non-overlapping subspaces: $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots \cup \mathcal{S}_m$. For any \mathcal{S}_i ($1 \leq i \leq m$), and any two states $s, s' \in \mathcal{S}_i$ ($s \neq s'$), if for any action a , we have i) $\sum_{s'' \in \mathcal{S}_j} P_{s \rightarrow s''}^a = \sum_{s'' \in \mathcal{S}_j} P_{s' \rightarrow s''}^a$ for all $1 \leq j \leq m$ (i.e., with the same action a , the two states s and s' have the same probability to transition to another state subspace \mathcal{S}_j); and ii) $U_s^a = U_{s'}^a$, (i.e., with the same action a , the two states s and s' have the same expected short-term utility), then all states in the same state subspace \mathcal{S}_i have the bisimilarity relationship, i.e., they are equivalent and can be aggregated as one state ξ_i .

To study states with bisimilarity relationship in our MDP, we first divide the state space \mathcal{S} into subspaces. For any view $v \in \mathcal{V}$, we have $S_{\bar{\mathcal{R}},v}$ defined earlier, which includes states with view v and reputations no less than $t_r - 1$. For any view $v \in \mathcal{V}$ and reputation $1 \leq r \leq t_r - 2$, the state (r, v) forms a state subspace $\{(r, v)\}$ with a single element. All these subspaces are non-overlapping and we have $\bigcup_{v \in \mathcal{V}} (\{(1, v)\} \cup \dots \cup \{(t_r - 2, v)\} \cup S_{\bar{\mathcal{R}},v}) = \mathcal{S}$. For the example in Fig. 2 with 3 views and 3-level reputation system (where $t_r = 3$), Fig. 3a shows the corresponding states partition with 6 subspaces $\{(1, 1)\}$, $S_{\bar{\mathcal{R}},1} = \{(2, 1), (3, 1)\}$, $\{(1, 2)\}$, $S_{\bar{\mathcal{R}},2} = \{(2, 2), (3, 2)\}$, $\{(1, 3)\}$, and $S_{\bar{\mathcal{R}},3} = \{(2, 3), (3, 3)\}$.

We then have the following proposition, and the proof is in Appendix B.

Proposition 2: Following the above state partition, all states in $\mathcal{S}_{\bar{\mathcal{R}},v}$ have bisimilarity relationship, and can be aggregated as one state.

In the example in Fig. 3a, after state aggregation, there are 6 aggregated states: $(1, 1)$, $(\bar{\mathcal{R}}, 1)$, $(1, 2)$, $(\bar{\mathcal{R}}, 2)$, $(1, 3)$ and $(\bar{\mathcal{R}}, 3)$, and the MDP in Fig. 2 becomes the one shown in Fig. 3b.

The next step is to find the transition probability and the expected short-term utility function for the updated MDP. Following Definition 1, given the aggregated states $\{\xi_1, \dots, \xi_m\}$, by taking action a , a user transitions from state ξ_i to state ξ_j with probability $P_{\xi_i \rightarrow \xi_j}^a = \sum_{s' \in \mathcal{S}_j} P_{s \rightarrow s'}^a$ for any $s \in \mathcal{S}_i$, and the expected short-term utility at the aggregated state ξ_i is $U_{\xi_i}^a = U_s^a$ for any $s \in \mathcal{S}_i$. In our MDP, the view and the reputation transitions are independent, and the state aggregation here affects the reputation transition probability only. Therefore, we need to first find the updated reputation transition probability $P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^a$, $P_{\bar{\mathcal{R}} \rightarrow r}^a$ and $P_{r \rightarrow \bar{\mathcal{R}}}^a$ for $r < t_r - 1$. We first study $P_{r \rightarrow \bar{\mathcal{R}}}^a$ for $r < t_r - 1$. When $r < t_r - 1$, the reputation will always be increased by one step to $r + 1$ regardless of the action a . Thus, if $r = t_r - 2$, the reputation will be updated to $t_r - 1 \in \bar{\mathcal{R}}$, and $P_{(t_r-2) \rightarrow \bar{\mathcal{R}}}^a = 1$. For $r \leq t_r - 3$, we have $P_{r \rightarrow \bar{\mathcal{R}}}^a = 0$. We then study $P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^a$ and $P_{\bar{\mathcal{R}} \rightarrow r}^a$. From Section II-C2, with the reputation in $\bar{\mathcal{R}}$, the updated reputation r can only be 1 or $\min\{r + 1, R\} \in \bar{\mathcal{R}}$. Therefore, we have $P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^a = 1 - P_{\bar{\mathcal{R}} \rightarrow 1}^a$ and $P_{\bar{\mathcal{R}} \rightarrow r}^a = 0$ for $2 \leq r \leq t_r - 2$. From the proof of Proposition 2, $P_{r' \rightarrow 1}^a$ is the same for any $r' \in \bar{\mathcal{R}}$. Thus, we have $P_{\bar{\mathcal{R}} \rightarrow 1}^a = P_{r' \rightarrow 1}^a$ for all $r' \in \bar{\mathcal{R}}$. Based on the above analysis, we can find the updated state transition probability $P_{(r,v) \rightarrow (r',v')}^a = P_{r \rightarrow r'}^a \mathbf{T}^L(v, v')$ for any $r, r' \in \{1, \dots, t_r - 1, \bar{\mathcal{R}}\}$.

In the MDP in Fig. 3b, when a user is at the aggregated state $(\bar{\mathcal{R}}, v)$, from Theorem 1b), he will take either action $a = t_r = 3$ or action $a = R + 1 = 4$. With $a = 3$, he complies with the social norm in (3) and his reputation will be updated to 3 with probability 1, that is, $P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^{a=3} = P_{2 \rightarrow 3}^{a=3} = P_{3 \rightarrow 3}^{a=3} = 1$, and $P_{\bar{\mathcal{R}} \rightarrow 1}^{a=3} = P_{2 \rightarrow 1}^{a=3} = P_{3 \rightarrow 1}^{a=3} = 0$. Thus, we can calculate the updated state transition probabilities $P_{(\bar{\mathcal{R}},v) \rightarrow (\bar{\mathcal{R}},v')}^{a=3} = P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^{a=3} \mathbf{T}^L(v, v') = \mathbf{T}^L(v, v')$ and $P_{(\bar{\mathcal{R}},v) \rightarrow (1,v')}^{a=3} = P_{\bar{\mathcal{R}} \rightarrow 1}^{a=3} \mathbf{T}^L(v, v') = 0$. Similarly, we can find the state transition probabilities with action $a = 4$ for the MDP in Fig. 3b.

The last step is to update the expected short-term utility function $U_{\bar{\mathcal{R}},v}^a$. Proposition 2 shows that with the same action a , $U_{r,v}^a$ is the same for all $r \in \bar{\mathcal{R}}$ and thus, $U_{\bar{\mathcal{R}},v}^a = U_{r,v}^a$ for any $r \in \bar{\mathcal{R}}$.

2) *Lifetime Utility Functions:* In the following, we will study how the state aggregation affects the lifetime utility functions. From Theorem 1, when a user has reputation $r < t_r - 1$, he will take $a = R + 1$ in the equilibrium. For $r \in \bar{\mathcal{R}} = \{r | r \geq t_r - 1\}$, from Theorem 1c, for the same v , the actions are the same for all $r \in \bar{\mathcal{R}}$. Thus, to simplify the notation of a action policy π , we first omit its actions with reputations lower than $t_r - 1$. We then omit the reputation index r in the action for $r \in \bar{\mathcal{R}}$. Thus, we has

the simplified $\pi = \{a_1, a_2, \dots, a_M\}$, where $a_v \in \{t_r, R + 1\}$ is the action at view $v \in \mathcal{V}$ with $r \in \bar{\mathcal{R}}$.

With the above simplification of notations, we then study the lifetime utility with the aggregated states. Since we only need to focus on the action selection with reputation in $\bar{\mathcal{R}}$, (14) can be rewritten as

$$W_{\bar{\mathcal{R}},v}^\pi = U_{\bar{\mathcal{R}},v}^{a_v} + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') \left[(1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v}) W_{\bar{\mathcal{R}},v'}^\pi + P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v} W_{1,v'}^\pi \right]. \quad (15)$$

Note that (15) has a recursive term $W_{1,v}^\pi$. First, a user with reputation $r < t_r - 1$, his reputation always increases by 1 for every time he receives a request, until his reputation climbs to $t_r - 1$, i.e., to $\bar{\mathcal{R}}$. Before he climbs to $\bar{\mathcal{R}}$, following Theorem 1a), he always uses action $R + 1$ and does not help anyone. Thus, his cost is always zero. In addition, Theorem 1b) indicates no one helps users with reputation smaller than t_r . Therefore, he does not receive any gain from others' help with $G_{r',v} = 0$. Thus, his expected short-term utility is always zero. Based on this analysis, $W_{1,v}^\pi$ can be expanded following (14) as

$$\begin{aligned} W_{1,v}^\pi &= 0 + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{2,v'}^\pi = \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') \left[0 + \eta^L \sum_{v''=1}^M \mathbf{T}^L(v', v'') W_{3,v''}^\pi \right] \\ &= \eta^{2L} \sum_{v''=1}^M \mathbf{T}^{2L}(v, v'') W_{3,v''}^\pi = \dots = \eta^{(t_r-2)L} \sum_{v'=1}^M \mathbf{T}^{(t_r-2)L}(v, v') W_{t_r-1,v'}^\pi \\ &= \eta^{(t_r-2)L} \sum_{v'=1}^M \mathbf{T}^{(t_r-2)L}(v, v') W_{\bar{\mathcal{R}},v'}^\pi. \end{aligned} \quad (16)$$

Note that in (15) and (16), $\bar{\mathcal{R}}$ is a common reputation index in the subscripts of $W_{\bar{\mathcal{R}},v}^\pi$ and $U_{\bar{\mathcal{R}},v}^{a_v}$, which can be omitted for notation simplicity. We then substitute (16) into (15), and rewrite (15) as

$$W_v^\pi = U_v^{a_v} + \eta^L (1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v}) \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{v'}^\pi + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v} \sum_{v'=1}^M \mathbf{T}^{(t_r-1)L}(v, v') W_{v'}^\pi. \quad (17)$$

where $\gamma = \eta^{L(t_r-1)}$ is the discounting factor after receiving $t_r - 1$ requests.

To determine a policy π to be a Nash Equilibrium, we need to show that it can resist any one-shot deviation, where the user takes action a'_v other than the action a_v defined in π only for the current response to a request, and he will follow π in all later responses. The lifetime utility with the one-shot deviation to action a'_v is

$$W_v^{a'_v, \pi} = U_v^{a'_v} + \eta^L (1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v}) \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{v'}^\pi + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v} \sum_{v'=1}^M \mathbf{T}^{(t_r-1)L}(v, v') W_{v'}^\pi. \quad (18)$$

Comparing (17) and (18), one-shot deviation to a'_v gives a different expected short-term utility $U_v^{a'_v}$ and a different reputation transition probability $P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v}$. In later discussion, we use the one-shot deviation principle to examine whether a policy π is an equilibrium policy, that is, from a user's perspective, given that other users all take π unchanged, π is an equilibrium policy if and only if $W_v^\pi \geq W_v^{a'_v, \pi}$ for any v and a'_v .

C. Stationary Reputation Distribution

From the previous analysis, users' reputation distribution \mathbf{x} , where $\mathbf{x}(r)$ being the probability that a user has reputation r , affects the state transition probability and users' expected short-term utilities.

Thus, it affects users' decision making. If an equilibrium exists in the game, in the simple scenario with homogeneous users, all users should use the same strategy in the equilibrium, and it is expected that the reputation distribution should also converge to a stationary state. In the following, given a policy π adopted by all users, we determine whether there exists a stationary reputation distribution \mathbf{x} .

We first let $y = \sum_{r=t_r}^R \mathbf{x}(r)$ be the probability that a user is a beneficial user. Assume that user i receives a request from user j , and both of their reputation follow distribution \mathbf{x} . Then, following the social norm in Section II-C2, if user i 's current reputation is $r \leq t_r - 2$ (with probability $\mathbf{x}(r)$), his reputation will be increased to $r + 1$ for any action he takes. Thus, in the updated reputation distribution \mathbf{x}' , we should have $\mathbf{x}'(r) = \mathbf{x}(r - 1)$ for $2 \leq r \leq t_r - 1$. If the stationary state exists, the reputation distribution should remain the same and $\mathbf{x}'(r) = \mathbf{x}(r)$. Therefore, we have $\mathbf{x}(1) = \mathbf{x}(2) = \dots = \mathbf{x}(t_r - 1)$. In addition, given $y + \sum_{r=1}^{t_r-1} \mathbf{x}(r) = 1$, we have $\mathbf{x}(1) = \dots = \mathbf{x}(t_r - 1) = (1 - y)/(t_r - 1)$.

If user i 's reputation is $r \in \bar{\mathcal{R}}$ (which happens with probability $y + \mathbf{x}(t_r - 1)$), his action $a_v = \{t_r, R + 1\}$ only depends on his view v when receiving a request, and at the stationary state he is at view v with probability $\mathbf{v}(v)$. i 's reputation will then be updated to either 1 or $\min\{r + 1, R\}$. Given his possible action $a_v = \{t_r, R + 1\}$ and from (6), his reputation is updated to 1 if and only if he takes action $R + 1$ and user j who sends the request is a beneficial user with reputation $r_j \geq t_r$. Therefore, user i 's reputation is reduced to 1 with probability $P_{\bar{\mathcal{R}} \rightarrow 1}^a = I[a_v = R + 1]y$, and he is a beneficial user with probability $1 - P_{\bar{\mathcal{R}} \rightarrow 1}^a$. Therefore, after the reputation update, user i is a beneficial user with probability

$$y' = [y + \mathbf{x}(t_r - 1)] \sum_{v \in \mathcal{V}} \mathbf{v}(v)(1 - P_{\bar{\mathcal{R}} \rightarrow 1}^a) = [y + \mathbf{x}(t_r - 1)] \sum_{v \in \mathcal{V}} \mathbf{v}(v)(1 - yI[a_v = R + 1]). \quad (19)$$

For a given policy $\pi = \{a_v\}$, if the stationary state exists, we should have $y' = y$, and y should satisfy

$$y' - y = \left(y + \frac{1 - y}{t_r - 1} \right) \sum_{v \in \mathcal{V}} \mathbf{v}(v) \{1 - yI[a_v = R + 1]\} - y = 0. \quad (20)$$

We first observe that given π , the left hand side (LHS) of (20) is a quadratic function of y . When $y = 1$, $LHS = \sum_{v \in \mathcal{V}} \mathbf{v}(v) \{1 - I[a_v = R + 1]\} - 1 \leq \sum_{v \in \mathcal{V}} \mathbf{v}(v) - 1 = 0$, and when $y = 0$, $LHS = \frac{1}{t_r - 1} \sum_{v \in \mathcal{V}} \mathbf{v}(v) > 0$. Thus, (20) has a single root in the range $[0, 1]$. Therefore, given π , there exists a unique stationary reputation distribution \mathbf{x} . To find \mathbf{x} for a given policy π , we first solve (20) and find y , and then calculate $\mathbf{x}(1) = \dots = \mathbf{x}(t_r - 1) = \frac{1 - y}{t_r - 1}$.

IV. EQUILIBRIUM ACTION POLICY DERIVATION

In this section, we analytically derive the equilibrium action policies of the game. We first consider a simple scenario with a single anchor view and derive the equilibrium policies in Section IV-A. We then extend our analysis to the general case with multiple anchor views in Section IV-B.

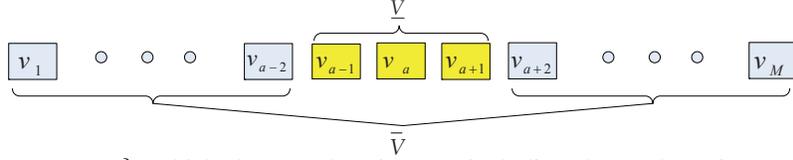


Fig. 4. $\underline{\mathcal{V}} = \{\sigma - 1, \sigma, \sigma + 1\}$, which denotes the view set including the anchor view σ and its left and right adjacent views. $\bar{\mathcal{V}} = \mathcal{V} \setminus \underline{\mathcal{V}}$, which denotes the view set including the rest views.

A. Game Analysis with A Single Anchor View

1) *View Switching Model with A Single Anchor View*: Following the view switching model described in Section II-A, with a single anchor view as shown in Fig. 4, this anchor view is in the middle, and partitions the rest $M - 1$ normal views into two normal view sets with $(M - 1)/2$ views per set. (Here, we assume M is an odd number.) Let $\sigma = (M + 1)/2$ denote the anchor view index. Following Section II-A, the one-step view transition matrix is

$$\mathbf{T} = \begin{matrix} & 1 & \cdots & \sigma - 1 & \sigma & \sigma + 1 & \cdots & M \\ \begin{matrix} 1 \\ \vdots \\ \sigma - 1 \\ \sigma \\ \sigma + 1 \\ \vdots \\ 1 \end{matrix} & \begin{pmatrix} \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} & P_a & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} & P_a & 0 & \cdots & 0 \\ \frac{(1-P_a)}{M-1} & \cdots & \frac{(1-P_a)}{M-1} & P_a & \frac{(1-P_a)}{M-1} & \cdots & \frac{(1-P_a)}{M-1} \\ 0 & \cdots & 0 & P_a & \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & P_a & \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} \end{pmatrix} & \end{matrix}. \quad (21)$$

Given (21), it is easy to show that the steady-state view distribution is $\mathbf{v}(\sigma) = P_a$ and $\mathbf{v}(v) = (1 - P_a)/(M - 1)$ for all other views.

2) *Expected Short-term Gain with A Single Anchor View*: For IMVS with a single anchor view, we first study the expected short-term gain $G_{r',v}$ for different views. In Appendix C, we show that with a single anchor view, $G_{r',v}$ in (11) can be rewritten as

$$G_{r',v} = \left(\sum_{r_k = \ell_{r-1}}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k) I[a_{v_k} \leq r'] \right) g_v, \quad (22)$$

where g_v defined in (12) is the maximum expected short-term gain when a user switches views starting from v and always receives help whenever needed. g_v is the only term in (22) that is affected by view v . In the following, we compare g_v with different v 's.

We first divide the view space \mathcal{V} into two sets $\underline{\mathcal{V}} = \{\sigma - 1, \sigma, \sigma + 1\}$ and $\bar{\mathcal{V}} = \mathcal{V} \setminus \underline{\mathcal{V}}$ as shown in Fig. 4. $\underline{\mathcal{V}}$ includes the anchor view and its left and right adjacent views, and $\bar{\mathcal{V}}$ includes the rest views. Then, we have the following Proposition 3, and the proof is in Appendix D.

Proposition 3: In a high dimensional IMVS (where the total number of views, M , is large, e.g., $M \geq 30$) with a single anchor view, all views in $\underline{\mathcal{V}}$ have approximately the same g_v 's, and for views in $\bar{\mathcal{V}}$, their g_v 's are also approximately the same.

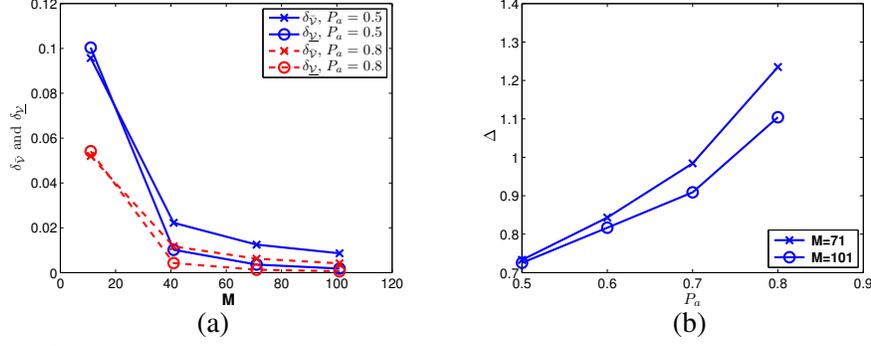


Fig. 5. (a) $\delta_{\underline{y}}$ and $\delta_{\bar{v}}$ with different M . (b) Δ with different P_a .

To numerically show this, we first let $g_{\underline{y}} \triangleq \sum_{v \in \underline{y}} g_v / |\underline{y}|$ denote the average g_v of all $v \in \underline{y}$, and similarly, define $g_{\bar{v}} \triangleq \sum_{v \in \bar{v}} g_v / |\bar{v}|$. We then let $\delta_{\bar{v}} \triangleq \frac{\max_{v \in \bar{v}} g_v - \min_{v \in \bar{v}} g_v}{g_{\bar{v}}}$ denote the maximum difference of g_v in \bar{v} normalized by the average $g_{\bar{v}}$. Similarly, we define $\delta_{\underline{y}} \triangleq \frac{\max_{v \in \underline{y}} g_v - \min_{v \in \underline{y}} g_v}{g_{\underline{y}}}$ for the set \underline{y} .

Fig. 5a plots $\delta_{\underline{y}}$ and $\delta_{\bar{v}}$ with M in the range [11, 101]. In this example, we have $N = 10$ users, and the forgetting factor is $\eta = 0.95$. We test the probability to switch to the anchor view $P_a = 0.5$ and 0.8. We observe the same trend for other values of the system parameters. From Fig. 5a, we observe that $\delta_{\underline{y}}$ and $\delta_{\bar{v}}$ are decreasing functions of M and P_a , and the difference between g_v in the same set becomes smaller as M and P_a increases. In the following analysis, we consider the scenario where M and P_a are large (e.g., $M \geq 30$ and $P_a \geq 0.5$), and the difference between g_v in the same set is very small and can be ignored. In this scenario, $g_{\bar{v}}$ ($g_{\underline{y}}$) can be used to denote the g_v for all views in the set \underline{y} (\bar{v}).

We also observe that views in \bar{v} gives higher short-term gain than \underline{y} , i.e., $g_{\bar{v}} \geq g_{\underline{y}}$. To show this, we define $\Delta \triangleq \frac{g_{\bar{v}} - g_{\underline{y}}}{g_{\underline{y}}}$ as the difference between $g_{\bar{v}}$ and $g_{\underline{y}}$ normalized by $g_{\underline{y}}$. Fig. 5b shows Δ with different M and P_a . For different M 's, we observe the same trend. From Fig. 5b, when $M \geq 71$ and $P_a \geq 0.5$, the difference between $g_{\bar{v}}$ and $g_{\underline{y}}$ is at least 70%, and is an increasing function of P_a . This is because when P_a is larger, all users have a higher probability to switch to the anchor view. Since users in the set \underline{y} do not need others' help in this switch while those in \bar{v} need help from others, a larger P_a results in a lower $g_{\underline{y}}$ and a higher $g_{\bar{v}}$, and thus increases the difference.

3) *State Aggregation*: With the above observation, we can aggregate more states in the MDP to further simplify the analysis. We first classify the state space. For any $r \leq t_r - 2$, we define $\mathcal{S}_{r, \underline{y}} = \{(r, v) | v \in \underline{y}\}$ and $\mathcal{S}_{r, \bar{v}} = \{(r, v) | v \in \bar{v}\}$. We then define $\mathcal{S}_{\bar{r}, \underline{y}} = \{(\bar{r}, v) | v \in \underline{y}\}$ and $\mathcal{S}_{\bar{r}, \bar{v}} = \{(\bar{r}, v) | v \in \bar{v}\}$. Those state subspaces are non-overlapping, and $\mathcal{S}_{\bar{r}, \underline{y}} \cup \mathcal{S}_{\bar{r}, \bar{v}} \cup_{r \leq t_r - 2} (\mathcal{S}_{r, \underline{y}} \cup \mathcal{S}_{r, \bar{v}}) = \mathcal{S}$. Fig. 6a shows an example of the state classification with M views and 3-level reputation system, where $t_r = 3$ and $\bar{r} = \{2, 3\}$. In Fig. 6a, there are four non-overlapping state subspaces $\mathcal{S}_{1, \underline{y}}$, $\mathcal{S}_{1, \bar{v}}$, $\mathcal{S}_{\bar{r}, \underline{y}}$ and $\mathcal{S}_{\bar{r}, \bar{v}}$.

We then have the following proposition with the proof in Appendix E.

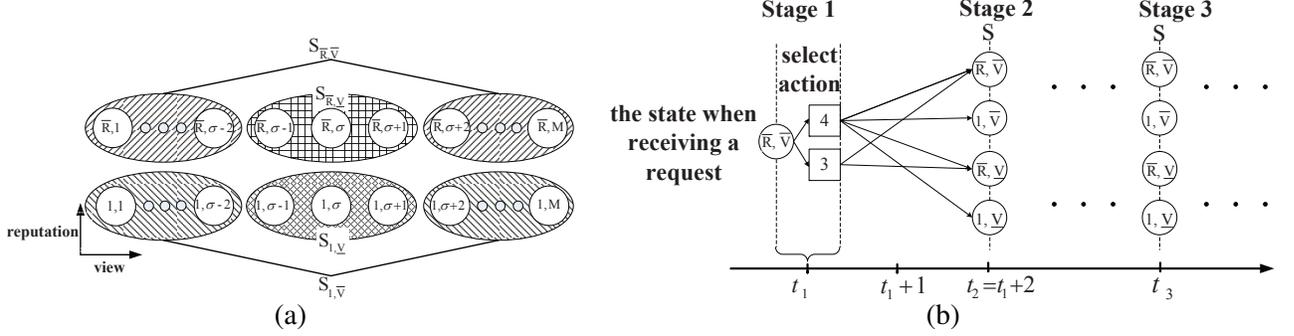


Fig. 6. Example of the state classification and aggregation with M views and 3-level reputation system ($t_r = 3$). (a) The state classification, where we have 4 state subspace $\mathcal{S}_{1, \underline{\mathcal{V}}}$, $\mathcal{S}_{1, \bar{\mathcal{V}}}$, $\mathcal{S}_{\bar{\mathcal{R}}, \underline{\mathcal{V}}}$ and $\mathcal{S}_{\bar{\mathcal{R}}, \bar{\mathcal{V}}}$. (b) The MDP after state aggregation, where we have only 4 states in the state space, $(1, \underline{\mathcal{V}})$, $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$, $(1, \bar{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$.

Proposition 4: With the above state classification, states in each subspace have bisimilarity relationship and can be aggregated as one state.

From Proposition 4, all states in the same subspace can be aggregated into one state. Thus, for the example in Fig. 6a, there are four aggregated states denoted as $(1, \underline{\mathcal{V}})$, $(1, \bar{\mathcal{V}})$, $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$, and Fig. 6b shows the updated MDP after state aggregation. From Theorem 1 and the discussion in Section III-B, for the aggregated state with reputation $r < t_r - 1$, users will always take action $a = R + 1$ and do not cooperate with anyone. Therefore, we only need to consider the aggregated states, $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$, and let $a_{\underline{\mathcal{V}}}$ and $a_{\bar{\mathcal{V}}}$ denote actions taken at these two aggregated states, respectively.

The next step is to study the state transition probability for the aggregated states. Note that the reputation and view transition probabilities are independent, and the reputation transition probabilities are the same as in Section III-B1. Therefore, we only need to analyze the updated view transition probabilities. Note that given the one-step view transition matrix in (21), starting from any view $v \in \underline{\mathcal{V}}$, after one segment, it will transition to views in $\underline{\mathcal{V}}$ with the same probability $\sum_{v' \in \underline{\mathcal{V}}} \mathbf{T}(v, v') = P_a + \frac{2(1-P_a)}{M-1}$, and to views in $\bar{\mathcal{V}}$ with the same probability $\sum_{v' \in \bar{\mathcal{V}}} \mathbf{T}(v, v') = 1 - P_a - \frac{2(1-P_a)}{M-1}$. Therefore, with the aggregated states, the one-step view transition probability is denoted as $\mathbf{T}(\underline{\mathcal{V}}, \underline{\mathcal{V}}) = P_a + \frac{2(1-P_a)}{M-1}$, which is the probability that a user transitions from $\underline{\mathcal{V}}$ (i.e., from any view in $\underline{\mathcal{V}}$) to views in $\underline{\mathcal{V}}$. Similarly, we also have $\mathbf{T}(\bar{\mathcal{V}}, \underline{\mathcal{V}}) = P_a + \frac{2(1-P_a)}{M-1}$ and $\mathbf{T}(\underline{\mathcal{V}}, \bar{\mathcal{V}}) = \mathbf{T}(\bar{\mathcal{V}}, \bar{\mathcal{V}}) = 1 - P_a - \frac{2(1-P_a)}{M-1}$. Thus, at the steady-state view distribution, a user will be at views in $\underline{\mathcal{V}}$ with probability $\mathbf{v}(\underline{\mathcal{V}}) = P_a + \frac{2(1-P_a)}{M-1}$ and $\mathbf{v}(\bar{\mathcal{V}}) = 1 - P_a - \frac{2(1-P_a)}{M-1}$, where we still use \mathbf{v} to denote the steady state view distribution over those two view sets for notation simplicity. Therefore, with the aggregated states, the state transition probability is $P_{(r,v) \rightarrow (r',v')}^a = P_{r \rightarrow r'}^a \mathbf{v}(v')$ for all $r, r' \in \{1, \dots, t_r - 2, \bar{\mathcal{R}}\}$ and $v, v' \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$.

In the example in Fig. 6b, given the current state $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$ at stage 1, from Theorem 1, the possible equilibrium actions are $a = t_r = 3$ and $a = t_r + 1 = 4$. When taking action $a = 3$, the user follows the

social norm, and his reputation stays at $\bar{\mathcal{R}}$ with probability $P_{\bar{\mathcal{R}} \rightarrow \bar{\mathcal{R}}}^{a=3} = 1$, and the state transition probabilities are $P_{(\bar{\mathcal{R}}, \underline{\mathcal{V}}) \rightarrow (\bar{\mathcal{R}}, \underline{\mathcal{V}})}^{a=3} = \mathbf{v}(\underline{\mathcal{V}})$, $P_{(\bar{\mathcal{R}}, \bar{\mathcal{V}}) \rightarrow (\bar{\mathcal{R}}, \bar{\mathcal{V}})}^{a=3} = \mathbf{v}(\bar{\mathcal{V}})$, and $P_{(\bar{\mathcal{R}}, \bar{\mathcal{V}}) \rightarrow (1, v')}^{a=3} = 0$ for any $v' \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. Similarly, we can derive the state transition probabilities when the action $a = 4$ is taken.

We then derive the expected short-term utility of the updated MDP after state aggregation. From the proof of Proposition 4, all views in the same view set $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$) give the same expected short-term utility. Thus, for the aggregated state, we have $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}} = U_v^{a_{\underline{\mathcal{V}}}}$ for any $v \in \underline{\mathcal{V}}$, and $U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}} = U_v^{a_{\bar{\mathcal{V}}}}$ for any $v \in \bar{\mathcal{V}}$.

After this state aggregation, the lifetime utility in (17) can be written as

$$\begin{cases} W_{\underline{\mathcal{V}}}^{\pi} = U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}} + [\eta^L(1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}}) + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}}] [\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}], \\ W_{\bar{\mathcal{V}}}^{\pi} = U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}} + [\eta^L(1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}}) + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}}] [\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}], \end{cases} \quad (23)$$

and the action policy can be simplified as $\pi = \{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\}$ for $r \in \bar{\mathcal{R}}$.

4) *Equilibrium Analysis with 2-Level Reputation System:* In this section, we consider a simple scenario with a 2-level reputation system (i.e., $R = 2$), and derive the equilibrium action policies of the game. Since the threshold $1 < t_r \leq R = 2$, t_r can only be 2. Note that the 2-level reputation system is memoryless. This is because if a user's behavior complies with the social norm, his reputation is updated to 2. Otherwise, his reputation is updated to 1 regardless of his past reputation.

Note that with $t_r = R = 2$, $\bar{\mathcal{R}} = \{r \geq t_r - 1\} = \{1, 2\} = \mathcal{R}$. Therefore, after state aggregation in Section IV-A3, there are only two aggregated state $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$. As discussed in IV-A3, we study the action policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\}$ for $r \in \bar{\mathcal{R}}$. Since $a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}} \in \{t_r, R + 1\} = \{2, 3\}$, where $a = 2$ means cooperation with beneficial users, and $a = 3$ means no cooperation with anyone. Thus, we have 4 possible action policies $\{a_{\underline{\mathcal{V}}} = 2, a_{\bar{\mathcal{V}}} = 2\}$, $\{a_{\underline{\mathcal{V}}} = 2, a_{\bar{\mathcal{V}}} = 3\}$, $\{a_{\underline{\mathcal{V}}} = 3, a_{\bar{\mathcal{V}}} = 2\}$ and $\{a_{\underline{\mathcal{V}}} = 3, a_{\bar{\mathcal{V}}} = 3\}$. By examining each of them, we have the following Proposition 5, and the proof is in Appendix F.

Proposition 5: For an IMVS with a single anchor view and 2-level reputation system,

- a) If $g_{\underline{\mathcal{V}}} \geq c$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 2\}$ is an equilibrium policy, where users cooperate at all views (*full cooperation*).
- b) If $g_{\bar{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}}) \geq c \geq g_{\underline{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}})$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 2\}$ is an equilibrium policy, where users only cooperate at views in $\bar{\mathcal{V}}$ with high expected short-term gains, but not at views in $\underline{\mathcal{V}}$ with low expected short-term gains (*partial cooperation*).
- c) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 3\}$ is always an equilibrium policy, where users do not cooperate at all (*no cooperation*).
- d) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 3\}$ is not an equilibrium policy.

From Proposition 5, there are multiple Nash Equilibriums coexisting. In addition, from Proposition 5.a, users cooperate at views in $\underline{\mathcal{V}}$ only when $g_{\underline{\mathcal{V}}} \geq c$. This is because the 2-level reputation system is

memoryless, and users decide their actions only based on the expected short-term utility. If cooperation at views in $\underline{\mathcal{V}}$ gives a negative expected short-term utility ($g_{\underline{\mathcal{V}}} < c$), users will not cooperate.

5) *Equilibrium Analysis with R-level ($R \geq 3$) Reputation System:* The R -level ($R \geq 3$) reputation system is non-memoryless, and users need to take their future utilities into consideration. We also study the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\}$ for reputation $r \geq t_r - 1$, where $a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}} \in \{t_r, R + 1\}$. Thus, we also have 4 possible policies $\{a_{\underline{\mathcal{V}}} = t_r, a_{\bar{\mathcal{V}}} = t_r\}$, $\{a_{\underline{\mathcal{V}}} = t_r, a_{\bar{\mathcal{V}}} = R + 1\}$, $\{a_{\underline{\mathcal{V}}} = R + 1, a_{\bar{\mathcal{V}}} = t_r\}$ and $\{a_{\underline{\mathcal{V}}} = R + 1, a_{\bar{\mathcal{V}}} = R + 1\}$. By examining each of them, we have the following proposition.

Proposition 6: For an IMVS with a single anchor view and R -level reputation system where $R \geq 3$,

a) If $\bar{c}_1 \triangleq \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}})}{1 - \gamma} + g_{\underline{\mathcal{V}}} \geq c$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$ is an equilibrium policy, where users cooperate at all views (*full cooperation*).

b) If

$$\begin{aligned} & \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}}) [(1 - \gamma)g_{\bar{\mathcal{V}}} - (\eta^L - \gamma)\mathbf{v}(\underline{\mathcal{V}})(1 - y)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})]}{1 - \eta^L + \eta^L y - y\gamma} \geq c \\ & \geq \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}}) [\mathbf{v}(\bar{\mathcal{V}})(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) + (1 - \gamma)g_{\underline{\mathcal{V}}}]}{1 - \eta^L + \eta^L y - y\gamma}, \end{aligned} \quad (24)$$

$\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R + 1, t_r\}$ is an equilibrium policy, where users cooperate at views in $\bar{\mathcal{V}}$ with high expected short-term gains but not at views in $\underline{\mathcal{V}}$ with low expected short-term gains (*partial cooperation*).

c) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R + 1, R + 1\}$ is always an equilibrium policy, where users do not cooperate at all (*no cooperation*).

d) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R + 1\}$ is not an equilibrium policy.

Proof: In the following, we will prove Proposition 6.a, and the rest of the proof is in Appendix G.

For the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$, we determine when it is an equilibrium. To do this, we first assume that all users use this policy and study the corresponding stationary reputation distribution \mathbf{x} following the discussion in Section III-C. Then, using the one-shot deviation principle, we examine whether a user has incentives to unilaterally deviate to any one-shot deviation at any view.

As discussed in Section III-C, by solving (20), we have $y = 1$ and $\mathbf{x}(1) = \dots = \mathbf{x}(t_r - 1) = 0$. This is because all users' actions comply with the social norm in (3) and thus have the highest reputation R .

We then exam the one-shot deviation principle. First, the given policy is $a_{\underline{\mathcal{V}}} = a_{\bar{\mathcal{V}}} = t_r$. When a user with reputation $r \in \bar{\mathcal{R}}$ receives a request at $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, by taking action $a_v = t_r$, he will upload the requested frame with probability 1, since all other users have reputation R . Thus, the expected immediate cost is $C^{a=t_r} = c$. In addition, $a_v = t_r$ complies with the social norm in (3), and the his reputation will be lowered to 1 with probability $P_{\bar{\mathcal{R}} \rightarrow 1}^{t_r} = 0$, and he is a beneficial user with probability 1. Since others also take policy $a_{\underline{\mathcal{V}}} = a_{\bar{\mathcal{V}}} = t_r$, he will always receive others' help and have the maximum expected

short-term gain g_v for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. Therefore, following the policy $a_v = t_r$, his expected short-term utility is $U_v^{a=t_r} = -c + g_v$ for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, and the lifetime utility in (23) becomes

$$\begin{cases} W_{\underline{\mathcal{V}}}^{\pi} = -c + g_{\underline{\mathcal{V}}} + \eta^L \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right], \\ W_{\bar{\mathcal{V}}}^{\pi} = -c + g_{\bar{\mathcal{V}}} + \eta^L \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right]. \end{cases} \quad (25)$$

Note that (25) is a linear system with two unknowns $W_{\underline{\mathcal{V}}}^{\pi}$ and $W_{\bar{\mathcal{V}}}^{\pi}$, and we have

$$W_{\underline{\mathcal{V}}}^{\pi} = \frac{g_{\underline{\mathcal{V}}} - c + \eta^L \mathbf{v}(\bar{\mathcal{V}})(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})}{1 - \eta^L}, \quad \text{and} \quad W_{\bar{\mathcal{V}}}^{\pi} = \frac{g_{\bar{\mathcal{V}}} - c - \eta^L \mathbf{v}(\underline{\mathcal{V}})(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})}{1 - \eta^L}. \quad (26)$$

Now we study the user's lifetime utility if he takes one-shot deviation. As discussed in Section III-B, action t_r and $R+1$ dominate other strategies, and thus, we only need to exam the one-shot deviation to $R+1$. First, with $a'_v = R+1$ for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, this user does not help anyone and the immediate cost is 0. Since all other users have reputation R , the action $a'_v = R+1$ makes his reputation lowered to 1 with probability $P_{\bar{r} \rightarrow 1}^{a'_v=R+1} = 1$. Thus, he cannot receive others' help in the following L segments, and the expected short-term gain is $G_{1,v} = 0$. Therefore, the expected short-term utility by one-shot deviation to $R+1$ is $U_v^{a'_v=R+1} = 0$. Thus, the lifetime utility of one-shot deviation in (18) can be rewritten as

$$\begin{cases} W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} = \gamma \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right], \\ W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} = \gamma \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right] \end{cases} \quad (27)$$

Substitute (26) into (27) and compare W_v^{π} with $W_v^{a'_v=R+1,\pi}$ for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. We have

$$\begin{aligned} W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} &= \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c - g_{\underline{\mathcal{V}}})(1 - \gamma)}{1 - \eta^L} + (g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}), \\ W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} &= \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c - g_{\underline{\mathcal{V}}})(1 - \gamma)}{1 - \eta^L}. \end{aligned} \quad (28)$$

It is easy to observe that $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} > W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi}$. Thus, as long as $W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} = \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c - g_{\underline{\mathcal{V}}})(1 - \gamma)}{1 - \eta^L} \geq 0$, i.e., $\bar{c}_1 \triangleq \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}})}{1 - \gamma} + g_{\underline{\mathcal{V}}} \geq c$, we have $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} > W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} \geq 0$, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$ is an equilibrium policy. This completes the proof of Proposition 6.a. ■

From Proposition 6, we also observe multiple Nash Equilibria coexisting. Fig. 7 shows an example for a single-anchor view IMVS with $M = 101$ views, $N = 10$ users and $R = 10$ levels of reputations with $t_r = 10$. In the view switching model, users switch to the single anchor view with probability $P_a = 0.5$. The discounting factor is $\eta = 0.95$, and the expected short-term gains for the view set $\bar{\mathcal{V}}$ and $\underline{\mathcal{V}}$ are $g_{\bar{\mathcal{V}}} = 0.82$ and $g_{\underline{\mathcal{V}}} = 0.33$, respectively. From Fig. 7, when $c \in [0, 0.56]$, the full cooperation policy is an equilibrium. When $c \in [0.30, 0.36]$, the partial cooperation policy is an equilibrium. The non-cooperation policy is always an equilibrium for all $c \geq 0$. Thus, we have three equilibrium policies (full, partial and no cooperation) when $c \in [0.30, 0.36]$, and we have two equilibrium policies (partial and no cooperation) when $c \in [0, 0.30) \cup (0.36, 0.56]$.

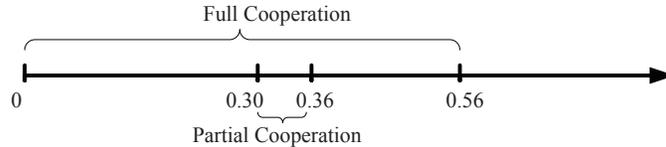


Fig. 7. An example of the coexistence of multiple equilibrium policies for a single anchor view IMVS.

When comparing Proposition 6a and Proposition 5a for the conditions of full cooperation, we observe that when $g_{\underline{\mathcal{V}}} < c \leq \bar{c}_1$ and when cooperation at views in $\underline{\mathcal{V}}$ gives him a negative expected short-term utility, a user will not cooperate at views in $\underline{\mathcal{V}}$ in the 2-level reputation system; while he may still fully cooperate at all views in the R -level reputation system with $R \geq 3$. This is because different from the memoryless 2-level reputation system, the R -level reputation system is non-memoryless, and a user needs to consider his future utilities when making a decision. Although cooperation at views in $\underline{\mathcal{V}}$ gives a negative expected short-term utility, this also helps him maintain a high reputation and keep receiving others' help in future view switching. As long as the expected future gain can compensate his current loss, he will still cooperate.

B. Game Analysis with Multiple Anchor Views

For the general IMVS with multiple anchor view and R -level ($R \geq 3$) reputation system, similar to the analysis in Section IV-A5, non-cooperation at all views is always an equilibrium, and partial cooperation and full cooperation may be equilibrium policies in certain scenarios. With a large view space \mathcal{V} , we have many different partial cooperation policies, and the analysis for each partial cooperation policy is much more complicated than that in the single anchor view system. Note that from the system designer's perspective, the full cooperation equilibrium makes all users cooperate whenever possible, minimizes the consumed upload bandwidth at the server's side, and thus is the desired equilibrium policy. Therefore, in this work, for IMVS with multiple anchor views, we focus on the analysis of full cooperation and derive the conditions for full cooperation to be an equilibrium policy.

Similar to the proof in Proposition 6.a, we first assume that all users take the full cooperation policy $\pi = \{a_1, a_2, \dots, a_M\} = \{t_r, t_r, \dots, t_r\}$ and derive the corresponding reputation distribution \mathbf{x} . We then examine whether π can resist the one-shot deviation to $a'_v = R + 1$ for any $v \in \mathcal{V}$.

If all users cooperate with the policy π , they will keep the highest reputation R , and the reputation distribution is $y = 1$ and $\mathbf{x}(1) = \dots \mathbf{x}(t_r - 1) = 0$. For a user receiving a request at view $v \in \mathcal{V}$, he will help upload with probability 1 by following π . Thus, the expected immediate cost is c . In addition, with $P_{\mathcal{R} \rightarrow 1}^{t_r} = 0$, his reputation remains to be R . Therefore, he always receives others' help, and receives the maximum expected short-term gain g_v . Thus, his expected short-term utility is $U_v^{a_v=t_r} = -c + g_v$, and his lifetime utility is

$$W_v^\pi = -c + g_v + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{v'}^\pi, \forall v \in \mathcal{V}. \quad (29)$$

To solve (25), we expand the recursive term W_v^π at the right side of (29) and have

$$\begin{aligned}
W_v^\pi &= -c + g_v + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v')(-c + g_{v'}) + \eta^{2L} \sum_{v'=1}^M \sum_{v''=1}^M \mathbf{T}^L(v, v') \mathbf{T}^L(v', v'') W_{v''}^\pi \\
&= -c + g_v + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v')(-c + g_{v'}) + \eta^{2L} \sum_{v''=1}^M \mathbf{T}^{2L}(v, v'') W_{v''}^\pi \\
&= \dots = -c + g_v + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^M \mathbf{T}^{nL}(v, v')(-c + g_{v'}) \\
&= \underbrace{g_v + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^M \mathbf{T}^{nL}(v, v') g_{v'}}_{\triangleq \mathbb{G}_v} - c - c \sum_{n=1}^{\infty} \eta^{nL} \left(\sum_{v'=1}^M \mathbf{T}^{nL}(v, v') \right) = \mathbb{G}_v - \frac{c}{1 - \eta^L}. \quad (30)
\end{aligned}$$

In (30), \mathbb{G}_v is the maximum lifetime gain a user can receive (when helpers always help) if he starts view switching from view v , and $\frac{c}{1 - \eta^L}$ is his lifetime cost to help others whenever asked.³ From (30), a necessary condition for the full cooperation policy $\pi = (t_r, \dots, t_r)$ to be an equilibrium is to enable a non-negative lifetime utility with $W_v^\pi = \mathbb{G}_v - \frac{c}{1 - \eta^L} \geq 0$ for all views, that is, $c \leq (1 - \eta^L) \min_v \mathbb{G}_v$. Otherwise, users have no incentive to cooperate.

We then derive the lifetime utility with one-shot deviation to $a'_v = R + 1$. Similar to the analysis in Section IV-A5 for the single anchor view IMVS, we have

$$W_v^{a'_v=R+1, \pi} = \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') W_{v'}^\pi = \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \left(\mathbb{G}_{v'} - \frac{c}{1 - \eta^L} \right). \quad (31)$$

Comparing W_v^π in (30) and $W_v^{a'_v=R+1, \pi}$ in (31), we have

$$\begin{aligned}
W_v^\pi - W_v^{a'_v=R+1, \pi} &= \mathbb{G}_v - \frac{c}{1 - \eta^L} - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \left(\mathbb{G}_{v'} - \frac{c}{1 - \eta^L} \right) \\
&= \mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \mathbb{G}_{v'} - \frac{c}{1 - \eta^L} + \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \frac{c}{1 - \eta^L} \\
&= \mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \mathbb{G}_{v'} - \frac{1 - \gamma}{1 - \eta^L} c. \quad (32)
\end{aligned}$$

Define $\bar{c}_2 \triangleq \frac{1 - \eta^L}{1 - \gamma} \min_v \left\{ \mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \mathbb{G}_{v'} \right\}$. From (32), $c \leq \bar{c}_2$ (i.e., $W_v^\pi - W_v^{a'_v=R+1, \pi} \geq 0$ for all v 's) is also a necessary condition for the full cooperation policy $\pi = \{t_r, t_r, \dots, t_r\}$ to resist any one-shot deviation.

To summarize, when $c \leq \min \left\{ (1 - \eta^L) \min_v \mathbb{G}_v, \bar{c}_2 \right\}$, full cooperation with $\pi = \{t_r, t_r, \dots, t_r\}$ is an equilibrium policy.

³In (30), \mathbb{G}_v includes an infinite series. Since $\eta < 1$, it is easy to show that this series converges, and \mathbb{G}_v is finite, which users can calculate offline.

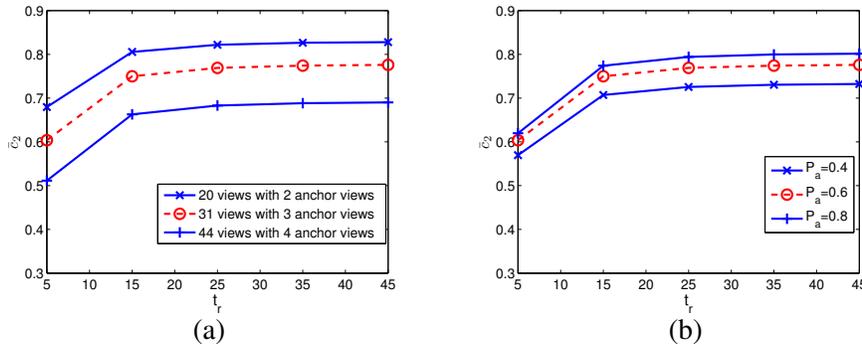


Fig. 8. \bar{c}_2 with t_r under different system setup. (a) $P_a = 0.6$. (b) $M = 31$ with 3 anchor views.

V. REPUTATION SYSTEM OPTIMIZATION AND FULL COOPERATION INITIATION

In this section, we study from system designer perspective and analyze how to stimulate full user cooperation to the highest possible level. We first study the optimal parameter selection for the reputation system. Then, observing that there are more than one equilibrium policies in the game and the MDP initial state determines the final equilibrium to which the game converges, we analyze the sufficient condition on the initial state to drive the game to the desired full cooperation equilibrium, and propose a Pay-for-Cooperation (PfC) scheme to ensure users' full cooperation at all views.

A. Optimal t_r and R of The Reputation System

For single anchor view IMVS, note that from Proposition 6a, a larger \bar{c}_1 gives a larger range of cost for users to have full cooperation as an equilibrium. Thus, from system designer perspective, a larger \bar{c}_1 is preferred to stimulate user cooperation. In \bar{c}_1 , we have the term $\gamma = \eta^{(t_r-1)L}$, where t_r is a system parameter describing the punishment a user receives if his action does not comply with the social norm. It is easy to show that $\partial\bar{c}_1/\partial t_r > 0$. Thus, the reputation system should select the highest $t_r = R$ that gives the harshest punishment to stimulate user cooperation to the highest level.

For the multi-anchor view IMVS, from the analysis of \bar{c}_2 in Section IV-B, t_r affects not only the term γ but also the summation term $\sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v')\mathbb{G}_{v'}$, which makes the analysis difficult. In the following, we consider an example with a $R = 45$ level reputation system, $N = 10$ users and a discounting factor of $\eta = 0.95$. For the above example, Fig. 8 shows the numerical results of \bar{c}_2 with different numbers of views M and different numbers of anchor views. Fig. 8 shows that \bar{c}_2 is an increasing function of t_r with different M and different view switching probability P_a . We observe the same trend for other values of the system parameters. Thus, similar to the single anchor view IMVS, in the multi-anchor view IMVS, a larger t_r also gives a larger \bar{c}_2 and thus a larger range of c for full cooperation being an equilibrium policy, and the largest possible $t_r = R$ should be chosen.

B. Full Cooperation Initiation

From the discussion in the previous sections, for both single- and multi- anchor view IMVS, we observe multiple equilibrium policies. When full cooperation is an potential equilibrium, we observe that the MDP's initial state, especially the percentage of cooperative users y_{in} , who are willing to cooperate at all views with other beneficial users, determines whether the game can converge to full cooperation. For example, in the extreme case with $y_{in} = 0$ where no user cooperates, a user who unilaterally cooperates will receive a negative utility due to the cost of frame uploading, and thus, is unwilling to cooperate. In the other extreme case with $y_{in} = 1$ where all users cooperate, if a user who unilaterally deviates from cooperation will receive penalty with a long term utility loss, and thus, he is willing to cooperate. In this section, we first derive a sufficient condition on y_{in} at the initial state, under which users are all willing to cooperate. We then propose a *Pay-for-Cooperation* (PfC) scheme at the beginning of the game to achieve the condition of y_{in} , and thus, drive the game to full cooperation equilibrium.

1) *Sufficient Condition of y_{in} for Full Cooperation*: To derive the sufficient condition on y_{in} for full cooperation at the initial state, we first assume that the local agent assigns each user the highest reputation R at the beginning of the game. This is because users may cooperate only when they have reputation larger than $t_r - 1$, and assigning each user the highest reputation R makes him have cooperation as an option. We then study how a user makes decisions if he is not one of the y_{in} percentage cooperative users. First, he considers the y_{in} percentage cooperative users will cooperate with the full cooperation policy $\pi_c = \{t_r, t_r, \dots, t_r\}$. Second, for the rest $1 - y_{in}$ percentage users, since he does not know how those users will behave, he considers the worst case scenario, where the $1 - y_{in}$ percentage users do not cooperate at all with the non-cooperation policy $\pi_n = \{R + 1, R + 1, \dots, R + 1\}$. Based on the above assumptions, we then use the one-shot deviation principle to exam whether the cooperation policy π_c is also his optimal policy. Since all users have reputation R at the beginning, by taking action t_r following π_c , he will cooperate and help to upload with probability 1 whenever being requested and his reputation is lowered to 1 with probability $P_{\mathcal{R} \rightarrow 1}^{t_r} = 0$. Furthermore, since he assumes that only the y_{in} percentage users will cooperate, his expected short-term gain is $g_v y_{in}$. Thus, his expected short-term utility is $U_v^{t_r} = -c + g_v y_{in}$, and his lifetime utility using policy π_c is

$$W_v^{\pi_c} = -c + g_v y_{in} + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{v'}^{\pi_c}. \quad (33)$$

Same as in (30), we also expand the recursive term $W_{v'}^{\pi_c}$ at the right side of (33), and have

$$W_v^{\pi_c} = -c + g_v y_{in} + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^M \mathbf{T}^{nL}(v, v') (-c + g_v y_{in}) = y_{in} \mathbb{G}_v - \frac{c}{1 - \eta^L}. \quad (34)$$

where \mathbb{G}_v is defined in (30). Similarly, a necessary condition for full cooperation to be his optimal strategy is $W_v^{\pi_c} \geq 0$, i.e., $y_{in} \geq \frac{c}{(1 - \eta^L) \mathbb{G}_v}$.

Similar to the proof of Proposition 6a, we then study the lifetime utility of the one-shot deviation to $a'_v = R + 1$. By taking action $R + 1$, he will upload with probability 0 and his reputation is lowered to 1 with probability $P_{\mathcal{R} \rightarrow 1}^{t_r} = 1$. Thus, he cannot receive others' help in the next L segments and the expected short-term gain is 0. Therefore, his expected short term utility is also zero. Thus, the lifetime utility with one-shot deviation to $a'_v = R + 1$ is

$$W_v^{a'_v=R+1, \pi_c} = \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') W_{v'}^{\pi_c}. \quad (35)$$

We then substitute (34) into (35), compare (34) and (35), and have

$$\begin{aligned} W_v^{\pi_c} - W_v^{a'_v=R+1, \pi_c} &= y_{in} \mathbb{G}_v - \frac{c}{1-\eta^L} - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \left[y_{in} \mathbb{G}_{v'} - \frac{c}{1-\eta^L} \right] \\ &= y_{in} \mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') y_{in} \mathbb{G}_{v'} - \frac{1-\gamma}{1-\eta^L} c. \end{aligned} \quad (36)$$

It is easy to show that (36) is an increasing function of y_{in} given that full cooperation is an equilibrium with $(\mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \mathbb{G}_{v'}) \geq 0$. By solving $W_v^{\pi_c} - W_v^{a'_v=R+1, \pi_c} \geq 0$ for all views, we have

$$y_{in} \geq \bar{y}_{in} \triangleq \frac{(1-\gamma)c}{(1-\eta^L) \min_{v \in \mathcal{V}} \left[\mathbb{G}_v - \gamma \sum_{v'=1}^M \mathbf{T}^{L(t_r-1)}(v, v') \mathbb{G}_{v'} \right]}. \quad (37)$$

Thus, if $y_{in} \geq y_{in}^{min} \triangleq \max\left(\frac{c}{(1-\eta^L)\mathbb{G}_v}, \bar{y}_{in}\right)$ is satisfied, π_c is the optimal strategy for him, and he will cooperate at the beginning of the game. Therefore, users will converge to the full cooperation equilibrium. From (37), we observe that \bar{y}_{in} is an increasing function of c , i.e., with a higher cost c , we need more cooperative users for cooperation initiation. This is because with a higher cost c , a user requires more cooperative users to cooperate with him to compensate his cost for cooperation.

2) *Pay-for-Cooperation Scheme*: The above analysis provides the condition on y_{in} for full cooperation initiation. To achieve such condition, we propose one possible solution, PfC scheme, where the local agent randomly selects $y_{in} \geq y_{in}^{min}$ percentage of users, and pays them for their cooperation with other beneficial users. Here, the payment is higher than the cost c , and thus, cooperation with beneficial users becomes the dominant strategy for the selected users. The local agent also announces y_{in} to the other unselected users, who are not paid for cooperation, to assist their decision making. Since with $y_{in} \geq y_{in}^{min}$, cooperation is also the optimal strategy for unselected users, they will start to cooperate at the beginning of the game and full cooperation can be initiated. Once the local agent observes that all unselected users have started to cooperate, it will gradually stop paying the selected users one by one. Note that once a selected user is stopped from being paid, he has to estimate other users' actions and makes his own decision on whether to continue cooperation. If the local agent stops paying all selected users at the same time, each selected user may have different estimation of other selected users' actions. Thus, their behavior is unpredictable, which may also affect the unselected users' cooperation. The proposed strategy where the local agent

stops paying the selected users one by one can avoid this problem. This is because if at a time only one selected user is stopped from being paid, he considers that the other users either still get paid for cooperation or have started to cooperate. In such a case, continuing cooperation is a dominant strategy for him.

VI. EXPERIMENTATION

This section evaluates the system performance by simulations. In the simulation setup, we have a $R = 10$ level reputation system and select the optimal $t_r = R$ as discussed in Section V-A. The server provides IMVS with $M = 31$ views to a group of $N = 10$ users, and each user is assigned reputation 10 at the beginning of the game. In the view transition model, users switch to anchor views with probability $P_a = 0.5$. The discounting factor is $\eta = 0.95$. Since an IMVS with a single anchor view is a special case of that with multiple anchor views, in this section, we only show the results with multiple anchor views, and let view 8, 16 and 24 be the three anchor views. In this work, we let the local agent keeps tracking the percentage of users who help upload frames and broadcast this information to assist users' decision making. When a user at view v receives a help request, he will use this information broadcasted by the local agent and follows (36) to calculate $W_v^{\pi_c} - W_v^{a'_v=R+1, \pi_c}$. If $W_v^{\pi_c} - W_v^{a'_v=R+1, \pi_c} \geq 0$, he will cooperate with action $a_v = t_r$. Otherwise, he plays non-cooperatively with $a_v = R + 1$.

A. Cooperation Initiation Verification

Following the discussion in Section IV-B, we first calculate the condition for full cooperation to be an equilibrium to be $c \leq 0.7$. We then select $c = 0.65 \leq 0.7$ to have full cooperation as an equilibrium. Following the discussion in Section V-B, the sufficient condition to initiate user cooperation is $y_{in} \geq y_{in}^{min} = 0.72$. In the following, we test different y_{in} selected by the local agent in the PfC scheme to verify our theoretical analysis. In our experiments, once the local agent observes all unselected users have started to cooperate, it will stop paying selected users one by one, and the local agent will stop paying after the 50th segment in all scenarios.

Fig. 9 shows the simulation results when the local agent selects $y_{in} = 0.8$. Since $y_{in} = 0.8 > y_{in}^{min} = 0.72$, it can initiate user cooperation. Fig. 9b shows the percentage of users who use the cooperative action $a = R$ when their reputations are no less than $t_r - 1$. We observe that all users cooperate, and thus they all have reputation R as shown in Fig. 9a. We then test that the local agent selects $y_{in} = 0.5 < 0.72$, and the results are in Fig. 10, from which we observe that it cannot initiate user cooperation. We observe that From Fig. 10b, in the first 50 segments when the selected users are paid to cooperate, the percentage of cooperative users increases gradually. After the local agent stops paying, even the selected users stop

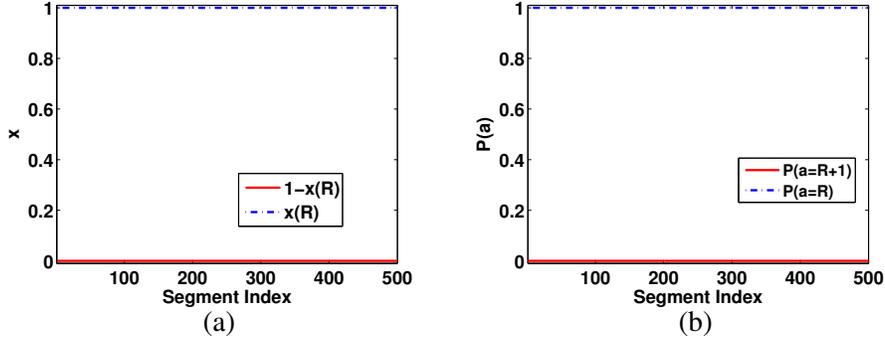


Fig. 9. The reputation and action distribution, when $y_{in} = 0.8$. (a) The reputation distribution x , where $x(R)$ is the probability that a user has reputation R . (b) $P(a = R)$ and $P(a = R + 1)$ are the percentages of users that use action R (coop

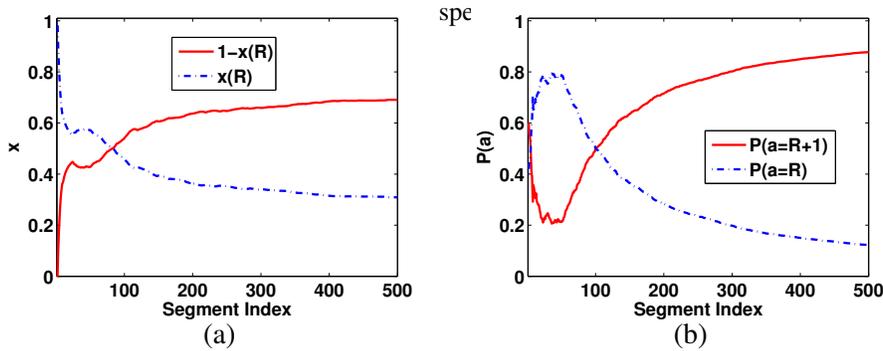


Fig. 10. The reputation and action distribution, when $y_{in} = 0.5$. (a) The reputation distribution x , where $x(R)$ is the probability that a user has reputation R . (b) $P(a = R)$ and $P(a = R + 1)$ are the percentages of users that use action R (cooperation) and $R + 1$ (non-cooperation), respectively, when their reputation is no less than $t_r - 1$. cooperation. Users cannot find enough cooperative users to cooperate with, and they all tend to not cooperate as shown in Fig. 10b. Thus, the probability of a user's reputation being R also drops as shown in Fig. 10a.

B. User Membership Dynamics

In the above simulations, we consider a fixed group of users' interaction. Once cooperation is initiated, they will cooperate until the end of this game. However, in practical video streaming system, users may join and leave the system from time to time. The membership dynamics may also affect their cooperation.

Consider a scenario where a group of existing users have been cooperating with each other with policy π_c and they all have reputation R . Then, some existing users leave, while several new users join the system with assigned reputation R . From the analysis in Section V-B1, if more than y_{in}^{min} percent of users in the newly formed group fully cooperate at all views and if this information is known to all users, then the game will converge to the full cooperation equilibrium and the cooperation is initiated in the next segment. Let $y_e \in [0, 1]$ denote the percentage of existing users after the membership update. If $y_e \geq y_{in}^{min}$, from their previous experience, existing users know that if they cooperate together, they

can gain positive payoffs and no one has incentive to unilaterally deviate from cooperation, even in the worst case scenario where newcomers do not cooperate at all. Thus, if existing users can communicate and confirm with each other that they will continue playing full cooperation, and if such information is publicly known to the newcomers, the game will converge to the full cooperation equilibrium and all users will cooperate. However, when $y_e < y_{in}^{min}$, even if all existing users continue cooperation, the full cooperation policy is no longer guaranteed to be the optimal strategy for them, and they will not confirm their cooperation with each other, which will also lead to non-cooperation of newcomers. Thus, cooperation of the network will be interrupted by this membership update.

In the following, we test how user membership dynamics affect user cooperation. We let $c = 0.25 \leq \bar{c}_2 = 0.7$ so that full cooperation is a optional equilibrium policy. At the beginning of the game, we select a large enough $y_{in} = 0.8$ to initiate user cooperation. For the membership dynamics, the initial number of users is 10. Users arrive the IMVS according to a Poisson process with an average arrival rate of λ users per segment duration. The sojourning period of each user follows an exponential distribution with an average of μ segments. Thus, a higher λ and a smaller μ result in more frequent membership update. In our simulations, we use batch join where new users can only join the streaming service at periodic moments, called batch moments. All new users coming between two neighboring batch moments will join and start receiving the streaming service at the same batch moment. In our simulations, the interval between neighboring batch moments is 30 segments, corresponding to a maximum of 10 seconds waiting time for a newly arrival user. For existing users, they can leave at any time instance. At each batch moment, the local agent will update y_e , and broadcast y_e and existing users' decisions to everyone in the network. In the following, we test different λ and μ , and study how y_e impacts user cooperation at each batch moment.

Fig. 11 shows the simulation results with low frequent membership update with $\lambda = 0.1$ and $\mu = 100$. Fig. 11a gives the number of users at each time instance. Fig. 11b shows y_{in}^{min} and y_e at each batch moment. Note that (37) includes the term \mathbb{G}_v , which is affected by the number of users. Thus, with user membership dynamics, y_{in}^{min} also changes. In Fig. 11b, since y_e is always higher than y_{in}^{min} , user cooperation will not be affected. Thus, from Fig. 11d users will always cooperate, and they will also maintain the reputation R as shown in Fig. 11c.

Fig. 12 shows the simulation results with high frequent membership update with $\lambda = 0.33$ and $\mu = 30$. Comparing Fig. 12a with Fig. 11a, we observe $\lambda = 0.33$ and $\mu = 30$ result in much more frequent membership update. From Fig. 12b, we observe that at the 90th segment, y_e is much smaller than y_{in}^{min} . Thus, user cooperation is interrupted after this batch moment. From Fig. 12d, we observe that users start

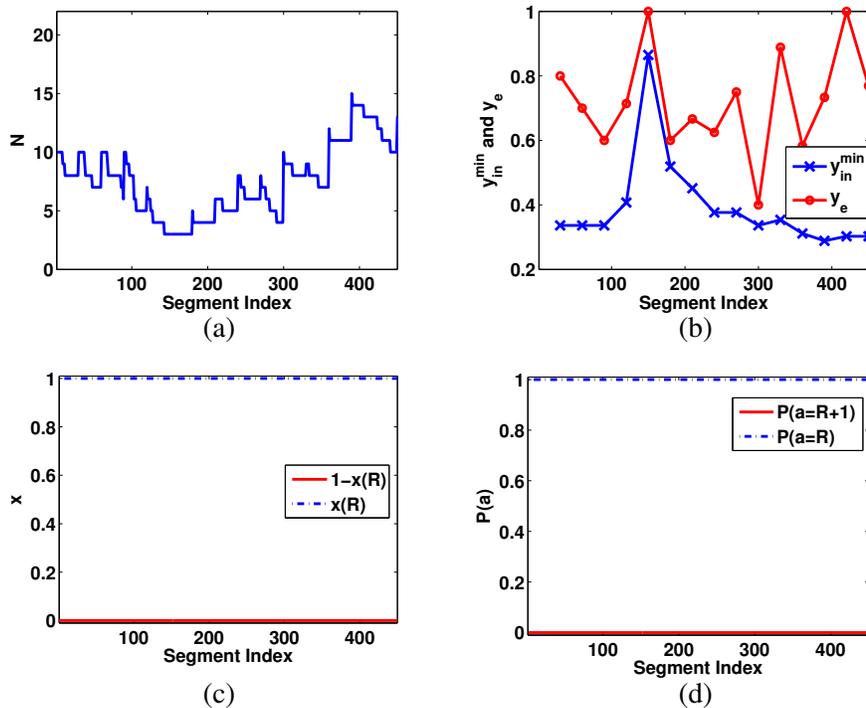


Fig. 11. The simulation results with low frequent membership update. $\lambda = 0.1$ and $\mu = 100$. (a) The number of users in the network. (b) y_{in}^{min} and y_e at each batch moment. (c) The reputation distribution. (d) The action distribution for users with reputation no less than $t_r - 1$.

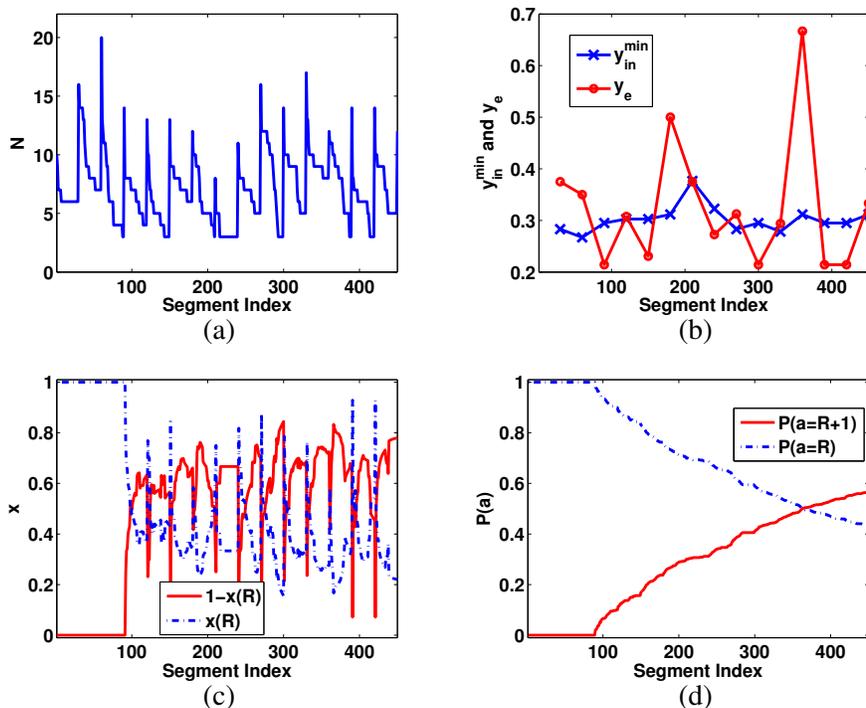


Fig. 12. The simulation results with high frequent membership update. $\lambda = 0.33$ and $\mu = 30$. (a) The number of users in the network. (b) y_{in}^{min} and y_e at each batch moment. (c) The reputation distribution. (d) The action distribution for users with reputation no less than $t_r - 1$.

to play non-cooperatively after the 90th segment, and the probability of a user's reputation being R also drops as shown in Fig. 12c. Since user cooperation has been interrupted, even if we have $y_e > y_{in}^{min}$ at a later batch time, users still do not cooperate. In Fig. 12c, we observe that after segment 90, $\mathbf{x}(R)$ fluctuates, and it increases at each batch moment, and then drops before the next batch moment. This is because at each batch moment, new users who are assigned with the highest reputation $R = 10$ joins the system, which will increase $\mathbf{x}(R)$. However, since they do not cooperate, the probability of their reputations being R decreases rapidly.

To overcome the cooperation interruption due to membership dynamics, one possibility is to let the local agent resume the PfC scheme: it first reset all users' reputation to R , randomly selects $y_{in} \geq y_{in}^{min}$ percent users, and pays for their cooperation to initiate user cooperation again.

VII. CONCLUSION

In this work, we propose an IMVS system that supports cooperative view switching. To stimulate user cooperation, we model user interaction as an indirect reciprocity game. From the game analysis, we observe that users cooperate at some views but not others. Since peers can predict their future view navigation paths probabilistically, a peer likely to enter a view switching path not requiring others' help will receive low utility from cooperation, and thus has less incentive to cooperate. Furthermore, we observe that more number of reputation levels provide more incentive for user cooperation, and thus should be used. In addition, we observe that the game may have multiple equilibria with different cooperation levels. To initiate user cooperation, we propose a PfC scheme. Finally, we study how user membership dynamics affect user cooperation. We observe that as long as the percentage of existing users is higher than a predetermined threshold, users will continue cooperation. Otherwise, the PfC scheme should be used to initiate user cooperation.

REFERENCES

- [1] T. Fujii, K. Mori, K. Takeda, K. Mase, M. Tanimoto, and Y. Suenaga, "Multipoint measuring system for video and sound—100 camera and microphone system," in *IEEE International Conference on Multimedia and Expo*, Toronto, Canada, July 2006.
- [2] G. Cheung, A. Ortega, and N. Cheung, "Interactive streaming of stored multiview video using redundant frame structures," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 744–761, March 2011.
- [3] H. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "Insights into pplive: A measurement study of a large-scale P2P IPTV system," in *IPTV workshop in conjunction with WWW2006*, May 2006.
- [4] X. Liao, H. Jin, Y. Liu, L. M. Ni, and D. Deng, "Anysee: Peer-to-Peer live streaming," *IEEE INFOCOM 2006*, pp. 1–10, April 2006.

- [5] X. Zhang, J. Liu, B. Li, and T. P. Yum, "CoolStreaming/DONet: A data-driven overlay network for Peer-to-Peer live media streaming," *IEEE INFOCOM 2005*, vol. 3, pp. 2102–2111, Mar. 2005.
- [6] N.-M. Cheung, A. Ortega, and G. Cheung, "Distributed source coding techniques for interactive multiview video streaming," in *27th Picture Coding Symposium*, Chicago, IL, May 2009.
- [7] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transaction on Information Theory*, vol. 19, pp. 471–480, 1973.
- [8] A. D. Wyner and Jacob Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, 1976.
- [9] Z. Liu, G. Cheung, and Y. Ji, "Distributed source coding for WWAN multiview video multicast with cooperative peer-to-peer repair," *IEEE International Conference on Communications*, June 2011.
- [10] Z. Liu, Y. Shen, S. Panwar, K. Ross, and Y. Wang, "Using layered video to provide incentives in P2P live streaming," *ACM SigComm Workshop on P2P Streaming and IP-TV*, pp. 311–316, Aug. 2007.
- [11] Z. Liu, Y. Shen, K. W. Ross, S. S. Panwar, and Y. Wang, "Substream trading: Towards an open P2P live streaming system," *IEEE International Conference on Network Protocols, 2008*, pp. 94–103, Oct. 2008.
- [12] W. Lin, H. V. Zhao, and K. J. R. Liu, "Incentive cooperation strategies for Peer-to-Peer live streaming social networks," *IEEE Transaction on Multimedia*, vol. 11, no. 3, pp. 396–412, April 2009.
- [13] F. Pianese, D. Perino, J. Keller, and E. W. Biersack, "PULSE: An adaptive, incentive-based, unstructured P2P live streaming system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1645–1660, Dec. 2007.
- [14] M. K. H. Yeung and Y. K. Kwok, "Game-theoretic scalable Peer-to-Peer media streaming," *IEEE International Symposium on Parallel and Distributed Processing, 2008*, pp. 1–11, April 2008.
- [15] G. Tan and S. A. Jarvis, "A payment-based incentive and service differentiation mechanism for peer-to-peer streaming broadcast," *International Workshop on Quality of Service (IWQoS)*, June 2006.
- [16] D. Guo and Y. Kwok, "A new auction based approach to efficient P2P live streaming," *2011 IEEE 17th International Conference on Parallel and Distributed Systems*, pp. 573–580, Dec. 2011.
- [17] C. Wang, H. Wang, Y. Lin, and S. Chen, "A lightweight currency-based P2P VoD incentive mechanism," *2010 Third International Joint Conference on Computational Science and Optimization (CSO)*, vol. 1, pp. 272–276, may 2010.
- [18] A. Habib and J. Chuang, "Incentive mechanism for peer-to-peer media streaming," *Internation Workshop on Quality of Service (IWQoS)*, pp. 171–180, June 2004.
- [19] Y. Zhang and Mihaela van der Schaar, "Peer-to-Peer multimedia sharing based on social norms," *Image Communication*, vol. 27, no. 5, pp. 383–400, May 2012.
- [20] M. Gupta, P. Judge, and M. Ammar, "A reputation system for peer-to-peer networks," *13th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, pp. 144–152, 2003.
- [21] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in P2P networks," *12th International Conference on World Wide Web*, pp. 640–651, 2003.
- [22] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, JohnWiley and Sons, 1994.
- [23] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291–1298, 2005.
- [24] Y. Sun, Z. Han, and K. J. R. Liu, "Defense of trust management vulnerabilities in distributed network," *IEEE Comm. Magazine*, vol. 46, no. 2, pp. 112–119, Feb. 2008.
- [25] D. Blackwell, "Discounted dynamic programming," *Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 226–235, Feb. 1965.

- [26] R. Givan, T. Dean, and M. Greig, "Equivalence notions and model minimization in Markov decision processes," *Artificial Intelligence*, vol. 147, no. 1-2, pp. 163–223, 2003.

APPENDIX A

PROOF OF PROPOSITION 1

Proof: As discussed in Section III-B, one-shot deviation of a given action policy means that a user takes a different action rather than the one defined in the action policy only for the current response to a request, but still follows the given action policy in the future responses. Assume that a user receives a request at reputation r and view v . Given the action policy π , the lifetime utility is defined in (14). If the user takes one-shot deviation to action $a'_{r,v} \neq a_{r,v}$ when responding to this request but follows π in the future, his one-shot deviation lifetime utility is

$$W_{r,v}^{a'_{r,v},\pi} = U_{r,v}^{a'_{r,v}} + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') \left[(1 - P_{r \rightarrow 1}^{a'_{r,v}}) W_{\min\{r+1, R\}, v'}^\pi + P_{r \rightarrow 1}^{a'_{r,v}} W_{1, v'}^\pi \right]. \quad (38)$$

Comparing (14) and (38), one-shot deviation to $a'_{r,v}$ gives a different expected short-term utility $U_{r,v}^{a'_{r,v}}$ and a different reputation transition probability $P_{r \rightarrow 1}^{a'_{r,v}}$. We then use one-shot deviation principle to prove this proposition by contradiction.

Assume that π is an equilibrium policy where action $t_r + 1 \leq a_{r,v} \leq R$ is chosen for some reputation level $r \geq t_r - 1$ and some view $v \in \mathcal{V}$. We first derive his lifetime utility $W_{r,v}^\pi$ if he follows the policy π . From Section III, for a user at view v with reputation $r \geq t_r - 1$, by taking action $a_{r,v} \in \pi$, his expected immediate cost is $C^{a_{r,v}} = c \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j)$, and following (6), his reputation transitions to 1 with probability $P_{r \rightarrow 1}^{a_{r,v}} = \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j)$. Thus, the expected short-term payoff is $U_{r,v}^{a_{r,v}} = -c \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) + (1 - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j)) G_{\min\{r+1, R\}, v} + \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) G_{1, v}$. Substitute $U_{r,v}^{a_{r,v}}$ and $P_{r \rightarrow 1}^{a_{r,v}}$ into (14), we have

$$\begin{aligned} W_{r,v}^\pi &= -c \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) + \left(1 - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \right) G_{\min\{r+1, R\}, v} + \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) G_{1, v} \\ &\quad + \eta^L \left(1 - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \right) \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{\min\{r+1, R\}, v'}^\pi + \eta^L \left(\sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \right) \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{1, v'}^\pi. \end{aligned}$$

We then analyze the one-shot deviation to $a'_{r,v} = t_r$ or $a'_{r,v} = R + 1$. With one-shot deviation to $a'_{r,v} = t_r$ for the current response, the user's expected immediate cost is $C^{t_r} = c \sum_{r_j=t_r}^R \mathbf{x}(r_j)$ and his reputation falls to 1 with probability $P_{r \rightarrow 1}^{t_r} = 0$. Thus, the expected short-term payoff becomes $U_{r,v}^{t_r} = -c \sum_{r_j=t_r}^R \mathbf{x}(r_j) + G_{\min\{r+1, R\}, v}$, and his life time utility becomes

$$W_{r,v}^{a'_{r,v}=t_r, \pi} = -c \sum_{r_j=t_r}^R \mathbf{x}(r_j) + G_{\min\{r+1, R\}, v} + \eta^L \sum_{v'=1}^M \mathbf{T}^L(v, v') W_{\min\{r+1, R\}, v'}^\pi.$$

Comparing $W_{r,v}^{a'_{r,v}=t_r,\pi}$ with $W_{r,v}^\pi$, we have

$$\begin{aligned} W_{r,v}^\pi - W_{r,v}^{a'_{r,v}=t_r,\pi} &= \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \underbrace{\left\{ c - G_{\min\{r+1,R\},v} + G_{1,v} - \eta^L \sum_{v'=1}^M \mathbf{T}^L(v,v') \left[W_{\min\{r+1,R\},v'}^\pi - W_{1,v'}^\pi \right] \right\}}_{\triangleq \nabla} \\ &= \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \nabla. \end{aligned}$$

Similarly, by one-shot deviation to $a'_{r,v} = R+1$ and refusing to help in the current response, his cost is zero, and his reputation falls to 1 with probability $P_{r \rightarrow 1}^{R+1} = \sum_{r_j=t_r}^R \mathbf{x}(r_j)$. Thus, the expected short-term payoff is $U_{r,v}^{R+1} = (1 - \sum_{r_j=t_r}^R \mathbf{x}(r_j))G_{\min\{r+1,R\},v} + \sum_{r_j=t_r}^R \mathbf{x}(r_j)G(1,v)$. His life-time utility is

$$\begin{aligned} W_{r,v}^{a'_{r,v}=R+1,\pi} &= \left(1 - \sum_{r_j=t_r}^R \mathbf{x}(r_j) \right) G_{\min\{r+1,R\},v} + \sum_{r_j=t_r}^R \mathbf{x}(r_j)G_{1,v} + \eta^L \left(\sum_{r_j=t_r}^R \mathbf{x}(r_j) \right) \sum_{v'=1}^M \mathbf{T}^L(v,v') W_{1,v'}^\pi \\ &\quad + \eta^L \left(1 - \sum_{r_j=t_r}^R \mathbf{x}(r_j) \right) \sum_{v'=1}^M \mathbf{T}^L(v,v') W_{\min\{r+1,R\},v'}^\pi, \end{aligned}$$

and when comparing $W_{r,v}^{a'_{r,v}=R+1,\pi}$ with $W_{r,v}^\pi$, we have

$$W_{r,v}^\pi - W_{r,v}^{a'_{r,v}=R+1,\pi} = - \sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) \nabla. \quad (39)$$

Note that $\sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \geq 0$ and $\sum_{r_j=a_{r,v}}^R \mathbf{x}(r_j) \geq 0$. Therefore, when $\nabla < 0$, we have $W_{r,v}^\pi - W_{r,v}^{a'_{r,v}=t_r,\pi} \leq 0$, and one-shot deviation to $a'_{r,v} = t_r$ gives a higher lifetime utility. When $\nabla > 0$, $W_{r,v}^\pi - W_{r,v}^{a'_{r,v}=R+1,\pi} \leq 0$, and π cannot resist one-shot deviation to $a'_{r,v} = R+1$. This contradicts the assumption that π is an equilibrium policy, and proves that a policy π with $t_r + 1 \leq a_{r,v} \leq R$ for $r \leq t_r - 1$ cannot be an equilibrium policy. ■

APPENDIX B

PROOF OF PROPOSITION 2

Proof: Following the definition in 1, for a given view v , to prove all states in $\mathcal{S}_{\bar{\mathcal{R}},v}$ have bisimilarity relationship, we first show that for any two states (r,v) and (r',v) with $r, r' \in \bar{\mathcal{R}}$ and the same action a , they have the same probability to transition to another state subspace. We then show that they give the same expected short-term utility.

Given the updated reputation space $\{\{1\}, \{2\}, \dots, \{t_r - 2\}, \bar{\mathcal{R}}\}$ and the updated state space $\{\mathcal{S}_{r,v}\}_{r \in 1, \dots, \bar{\mathcal{R}}, v \in \mathcal{V}}$, we first show that the two states (r,v) and (r',v) have the same probability to transition to another state

subspace. From the discussion in Section III, the reputation and view transition probabilities are independent, i.e., $P_{(r,v) \rightarrow (r'',v')}^a = P_{r \rightarrow r''}^a \mathbf{T}^L(v, v')$. Therefore, we only need to consider the reputation transition probability, that is, $\sum_{r'' \in \mathcal{R}_j} P_{r \rightarrow r''}^a = \sum_{r'' \in \mathcal{R}_j} P_{r' \rightarrow r''}^a$ where $\mathcal{R}_j \in \{\{1\}, \{2\}, \dots, \{t_r - 2\}, \bar{\mathcal{R}}\}$.

- Note that from (6), the probability that the reputation transitions to 1 depends only on the action. Thus, given the same action a , we have $P_{r \rightarrow 1}^a = P_{r' \rightarrow 1}^a$ for all $r, r' \in \bar{\mathcal{R}}$.
- Note that given $r \geq t_r - 1$, from the social norm in (3), the reputation r can only be updated to either 1 or $\min\{r+1, R\} \in \bar{\mathcal{R}}$. So for $r, r' \in \bar{\mathcal{R}}$, we have $P_{r \rightarrow r''}^a = P_{r' \rightarrow r''}^a = 0$ for $2 \leq r'' \leq t_r - 2$.
- Given that the reputation r can only be updated to either $\min\{r+1, R\} \in \bar{\mathcal{R}}$ or 1, we have $\sum_{r'' \in \bar{\mathcal{R}}} P_{r \rightarrow r''}^a = P_{r \rightarrow \min\{r+1, R\}} = 1 - P_{r \rightarrow 1}^a$ and $\sum_{r'' \in \bar{\mathcal{R}}} P_{r' \rightarrow r''}^a = P_{r' \rightarrow \min\{r'+1, R\}} = 1 - P_{r' \rightarrow 1}^a$. Thus, $\sum_{r'' \in \bar{\mathcal{R}}} P_{r \rightarrow r''}^a = \sum_{r'' \in \bar{\mathcal{R}}} P_{r' \rightarrow r''}^a$.

Therefore, we prove that all states in $\mathcal{S}_{(\bar{\mathcal{R}}, v)}$ have the same probability to transition to another state subspace.

To show that given the same action a , the two states (r, v) and (r', v) have the same expected short-term utility $U_{r,v}^a = -c \sum_{r''=a}^R \mathbf{x}(r'') + (1 - P_{r \rightarrow 1}^a) G_{\min(r+1, R), v} + P_{r \rightarrow 1}^a G_{1, v}$, first note that both the probability to upload ($\sum_{r''=a}^R \mathbf{x}(r'')$) and the probability for the reputation to transition to 1 ($P_{r \rightarrow 1}^a$) depend on the action a only. Thus, the same action taken at (r, v) and (r', v) introduces the same cost term and the same reputation transition probability. Second, from the analysis in Section III-A2, there is only one term in the expected short-term gain $G_{\min(r+1, R), v}$ that depends on the reputation $\min\{r+1, R\}$, that is, the term $I[a_{r_k, v_k(t+l)} \leq \min\{r+1, R\}]$ in (10). From Theorem 1, the action $a_{r_k, v_k(t+l)}$ can only be t_r or $R+1$. In addition, given $t_r - 1 \leq r, r' \leq R$, we have $t_r - 1 \leq \min\{r+1, R\}, \min\{r'+1, R\} \leq R$. Therefore, $I[a_{r_k, v_k(t+l)} \leq \min\{r+1, R\}] = I[a_{r_k, v_k(t+l)} \leq \min\{r'+1, R\}]$ and $G_{\min(r+1, R), v} = G_{\min(r'+1, R), v}$. It proves that $U_{r,v}^a = U_{r',v}^a$.

In summary, given view v and action a , the above proves that any two states in $\mathcal{S}_{\bar{\mathcal{R}}, v}$ have the same probability to transition to another state subspace and the same expected short-term utility. Thus, all states in $\mathcal{S}_{\bar{\mathcal{R}}, v}$ have bisimilarity relationship and can be aggregated into one space. This completes the proof. ■

APPENDIX C

PROOF OF (22)

We will show below that in the single anchor view IMVS, the expected short-term gain in (11) can be rewritten as

$$G_{r',v} = \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k) I[a_{v_k} \leq r'] \right) g_v.$$

Proof: From the discussion in Section III, if the helper k has reputation $r_k < t_r - 1$, he takes action $R+1$ and does not cooperate. That is, $I[a_{r_k, v_k(t+l)} \leq r'] = 0$ with $a_{r_k, v_k(t+l)} = R+1$ when $r_k < t_r - 1$ in (10). When he has reputation $r_k \geq t_r - 1$, his action depends on his view only. That is, $a_{r_k, v_k(t+l)} = a_{v_k(t+l)}$ for all $r_k \geq t_r - 1$. Thus, we can first rewrite $P[\mathbb{H}_h | \mathbb{H}_1(v')]$ in (10) as

$$P[\mathbb{H}_h | \mathbb{H}_1(v')] = \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k(t+l)=1}^M \mathbf{p}_{v'}(v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \right), \quad (40)$$

where $v_k(t)$ is the helper's view at time t . We then focus on the second term in (40), and prove that $\left(\sum_{v_k(t+l)=1}^M \mathbf{p}_{v'}(v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \right)$ are the same for all views v' in a single anchor view IMVS.

First note that in the single anchor view IMVS, as shown in Section IV-A1, the steady state view distribution is symmetric with respect to the anchor view σ . Also, as will be shown in Appendix D, the maximum expected short-term gain $\{g_v\}$ is symmetric with respect to the anchor view too. With homogeneous users with the same cost to upload a frame, it is expected that the equilibrium policy will be symmetric with respect to the anchor view, that is, $a_v = a_{M+1-v}$ for all $1 \leq v \leq M$.

From (10), we have

$$\begin{aligned} & \sum_{v_k(t+l)=1}^M \mathbf{p}_{v'}(v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \\ &= \sum_{v_k(t+l)=1}^M \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{T}(v'', v_k(t+l)) \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})} I[a_{v_k(t+l)} \leq r'] \\ &= \sum_{v'' \in \mathcal{V}_{v'}} \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})} \left(\sum_{v_k(t+l)=1}^M \mathbf{T}(v'', v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \right). \end{aligned} \quad (41)$$

For the term $\left(\sum_{v_k(t+l)=1}^M \mathbf{T}(v'', v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \right)$ in (41), following the discussion in Section IV, if v'' is the anchor view σ , we have $\mathbf{T}(v'', v_k(t+l)) = \mathbf{v}(v_k(t+l))$ where \mathbf{v} is the steady state view distribution with $\mathbf{v}(\sigma) = P_a$ and $\mathbf{v}(v) = \frac{(1-P_a)}{M-1}$ for all other views. Thus, we have

$$\begin{aligned} & \sum_{v_k(t+l)=1}^M \mathbf{T}(\sigma, v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \\ &= \sum_{v_k(t+l)=1}^M \mathbf{v}(v_k(t+l)) I[a_{v_k(t+l)} \leq r'] \\ &= \sum_{v_k(t+l)=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} I[a_{v_k(t+l)} \leq r'] + P_a I[a_\sigma \leq r']. \end{aligned} \quad (42)$$

Here, the last equality is due to the fact that $\mathbf{v}(v)$ and a_v are symmetric with respect to the anchor view σ .

If v'' is a normal view, for example, a normal view in the left normal view set with $v'' < \sigma$, in one segment, the helper k transitions to the anchor view with probability P_a and to each normal view in the left normal view set with probability $\frac{2(1-P_a)}{M-1}$. Thus, we have

$$\begin{aligned}
& \sum_{v_k(t+l)=1}^M \mathbf{T}(v'', v_k(t+l))I[a_{v_k(t+l)} \leq r'] \\
&= \sum_{v_k(t+l)=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} I[a_{v_k(t+l)} \leq r'] + P_a I[a_\sigma \leq r'_i] \\
&= \sum_{v_k(t+l)=1}^M \mathbf{v}(v_k(t+1))I[a_{v_k(t+l)} \leq r']. \tag{43}
\end{aligned}$$

The analysis of the scenario where v'' is a normal view in the right normal view set is similar and thus omitted.

To summarize, for all views, the term $\left(\sum_{v_k(t+l)=1}^M \mathbf{T}(v'', v_k(t+l))I[a_{v_k(t+l)} \leq r']\right)$ are the same and equals to $\sum_{v_k(t+l)=1}^M \mathbf{v}(v_k(t+1))I[a_{v_k(t+l)} \leq r']$ from (42). Based on the above analysis, we have

$$\begin{aligned}
& \sum_{v_k(t+l)=1}^M \mathbf{p}_{v'}(v_k(t+l))I[a_{v_k(t+l)} \leq r'] \\
&= \sum_{v'' \in \mathcal{V}_{v'}} \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})} \sum_{v_k(t+l)=1}^M \mathbf{v}(v_k(t+l))I[a_{v_k(t+l)} \leq r'] \\
&= \sum_{v_k=1}^M \mathbf{v}(v_k)I[a_{v_k} \leq r'], \tag{44}
\end{aligned}$$

which is not related to v' .

Thus, $P[\mathbb{H}_h | \mathbb{H}_1(v')]$ in (40) can be rewritten as

$$P[\mathbb{H}_h | \mathbb{H}_1(v')] = \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k)I[a_{v_k} \leq r'] \right), \tag{45}$$

and the expected short-term gain $G_{r',v}$ in (11) can be rewritten as

$$\begin{aligned}
G_{r',v} &= \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k)I[a_{v_k} \leq r'] \right) \\
&\quad \times \sum_{l=1}^L \eta^l \sum_{v'=1}^M P[\mathbb{H}_1(v')] P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v] \\
&= \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k)I[a_{v_k} \leq r'] \right) g_v. \tag{46}
\end{aligned}$$

where g_v is defined in (12). This completes the proof. ■

APPENDIX D

PROOF OF PROPOSITION 3

Here, we show that in a single anchor view IMVS with large number of views (e.g., $M \geq 30$), all views in $\underline{\mathcal{V}}$, their corresponding g_v 's are approximately the same. For views in $v \in \bar{\mathcal{V}}$, their corresponding g_v 's are also approximately the same.

Proof: First, following (12), we define

$$z(v, l) \triangleq \eta^l \sum_{v'=1}^M P[\mathbb{H}_1(v')] P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v] \quad (47)$$

and g_v can be rewritten as $g_v = \sum_{l=1}^L z(v, l)$, where $z(v, l)$ is the expected gain received at the l th segment after the view switching from view v if helpers always help.

We first prove that $\{z(v, l)\}$ are symmetric with respect to the anchor view σ , that is, $z(v, l) = z(M+1-v, l)$ for all $v = 1, \dots, M$. Let us consider the view pair $(v, M+1-v)$. For any $v' \in \mathcal{V}$, from (8) and (9), it is easy to show that the two probabilities $P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v]$ and $P[v_i(t+l) = M+1-v', v_i(t+l-1) \notin \mathcal{V}_{M+1-v'} | v_i(t) = M+1-v]$ (the probability that a user starts at view $v/M+1-v$ at time t and switches to $v'/M+1-v'$ at time $t+l$ from a non-adjacent view) are the same, and the probabilities that there is at least one user who can help are also the same ($P[\mathbb{H}_1(v')] = P[\mathbb{H}_1(M+1-v')]$). Therefore, from (47), we have $z(v, l) = z(M+1-v, l)$ and $z(v, l)$ is symmetric with respect to the anchor view σ . Since $g_v = \sum_{l=1}^T z(v, l)$, g_v is also symmetric with respect to the center anchor view.

In the following, we first show that for all $l \geq 2$, $z(v, l)$ is the same for all $v \in \mathcal{V}$. Therefore, for g_v of different view v , the difference is caused by $z(v, 1)$ (i.e., $l = 1$). In the second step, we show that views in the same subset $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$), the corresponding $z(v, 1)$'s are approximately the same.

- We first show that for $l \geq 2$, $z(v, l)$ is the same for all $v \in \mathcal{V}$.

Substitute (8) in Section III into (47), we first rewrite $z(v, l)$ ($l \geq 2$) as

$$\begin{aligned} z(v, l) &= \eta^l \sum_{v'=1}^M P[\mathbb{H}_1(v')] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}^{l-1}(v, v'') \mathbf{T}(v'', v') \right) \\ &= \eta^l \sum_{v'=1}^M P[\mathbb{H}_1(v')] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \sum_{\tilde{v}=1}^M \mathbf{T}(v, \tilde{v}) \mathbf{T}^{l-2}(\tilde{v}, v'') \mathbf{T}(v'', v') \right) \\ &= \eta \sum_{\tilde{v}=1}^M \mathbf{T}(v, \tilde{v}) \left\{ \eta^{(l-1)} \sum_{v'=1}^M P[\mathbb{H}_1(v')] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}^{l-2}(\tilde{v}, v'') \mathbf{T}(v'', v') \right) \right\} \\ &= \eta \sum_{\tilde{v}=1}^M \mathbf{T}(v, \tilde{v}) z(\tilde{v}, l-1). \end{aligned} \quad (48)$$

Therefore, if v is the anchor view, we have $\mathbf{T}(\sigma, \tilde{v}) = \mathbf{v}(\tilde{v})$ and thus,

$$z(\sigma, l) = \eta \sum_{\tilde{v}=1}^M \mathbf{v}(\tilde{v}) z(\tilde{v}, l-1) = \sum_{\tilde{v}=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} z(\tilde{v}, l-1) + P_a z(\sigma, l-1). \quad (49)$$

Here, the last equality is due to the fact that $\mathbf{v}(\tilde{v}) = \frac{(1-P_a)}{M-1}$ for all $\tilde{v} \neq \sigma$ and $\{z(\tilde{v}, l-1)\}$ are symmetric.

If v is a normal view, e.g., a normal view in the left side normal view set, after one segment, a user will only transition to the anchor view with probability P_a and to each normal view in the left side normal view set with probability $\frac{2(1-P_a)}{M-1}$. Thus, we have

$$z(v, l) = \sum_{\tilde{v}=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} z(\tilde{v}, l-1) + P_a z(\sigma, l-1) = z(\sigma, l). \quad (50)$$

The analysis is the same if v is a normal view in the right side normal view set. Thus, for $l \geq 2$, $z(v, l)$ is the same for all $v \in \mathcal{V}$.

• Second, we show that for views in $\underline{\mathcal{V}} = \{\sigma-1, \sigma, \sigma+1\}$, the corresponding $z(v, 1) = \eta \sum_{v' \notin \mathcal{V}_v} P [\mathbb{H}_1(v')] \mathbf{T}(v, v')$'s are approximately the same with a large M .

Note that given the view transition matrix \mathbf{T} as in (21), it is easy to show that $P [\mathbb{H}_1(v')]$ are the same for all $v' \in \{1, 2, \dots, \sigma-2, \sigma+2, \dots, M\}$ and equal to $1 - \left[P_a + \frac{M-7}{M-1}(1-P_a) \right]^N$.

When v is the anchor view σ , we have

$$\begin{aligned} z(\sigma, 1) &= \eta \sum_{v'=1}^{\sigma-2} \frac{(1-P_a)}{M-1} P [\mathbb{H}_1(v')] + \eta \sum_{v'=\sigma+2}^M \frac{(1-P_a)}{M-1} P [\mathbb{H}_1(v')] \\ &= \eta(\sigma-2) \frac{2(1-P_a)}{M-1} P [\mathbb{H}_1(1)]. \end{aligned} \quad (51)$$

Similarly, for $v = \sigma - 1$, we have

$$z(\sigma-1, 1) = \eta \sum_{v'=1}^{\sigma-3} \frac{2(1-P_a)}{M-1} P [\mathbb{H}_1(v')] = \eta(\sigma-3) \frac{2(1-P_a)}{M-1} P [\mathbb{H}_1(1)]. \quad (52)$$

Therefore, we have

$$\begin{aligned} z(\sigma, 1) - z(\sigma-1, 1) &= \eta \frac{2(1-P_a)}{M-1} P [\mathbb{H}_1(1)] \\ \text{and } \frac{z(\sigma, 1) - z(\sigma-1, 1)}{z(\sigma, 1)} &= \frac{1}{\sigma-2} = \frac{5}{M-2}. \end{aligned} \quad (53)$$

In a high dimension IMVS system with a large number of views (e.g., $M \geq 30$) and a large probability to transition to the anchor frame (e.g., $P_a \geq 0.5$), the difference between $z(\sigma, 1)$ and $z(\sigma-1, 1)$ is small and decreases as M increases.

Due to the symmetry of the IVMS system, $z(\sigma+1, 1) = z(\sigma-1, 1)$. Thus, all three views in $\underline{\mathcal{V}}$ have approximately the same $z(v, 1)$ with a large number of views.

- With the same analysis as above, we can show that all views in $\bar{\mathcal{V}}$ have approximately the same $z(v, 1)$ when M and P_a are large.

From all the above analysis, in the single anchor view IMVS, when the number of view and the probability to transition to the anchor view are large, g_v 's are approximately the same for all views in $\bar{\mathcal{V}}$; and similarly, all views in $\underline{\mathcal{V}}$ have approximately the same maximum expected short-term gain g_v . This completes the proof. ■

APPENDIX E

PROOF OF PROPOSITION 4

Here, we prove that in the single anchor view IMVS, with the state classification $\{\mathcal{S}_{r,v}\}_{r=1,\dots,\bar{R},v\in\{\underline{\mathcal{V}},\bar{\mathcal{V}}\}}$, all states in the same subspace have bisimilarity relationship and can be aggregated as one state.

Proof: With the above state classification, we first show that all states in the same subspace have the same probability to transition to another subspace, and then show that they have the same expected short-term utility. Then, from Definition 1, they have bisimilarity relationship and can be aggregated as one state.

- We first show that for any (r, v) and (r, v') from $\mathcal{S}_{r,\underline{\mathcal{V}}}$ ($\mathcal{S}_{r,\bar{\mathcal{V}}}$) with $v \neq v'$, they have the same probability to transition to another state subspace $\mathcal{S}_{r',\underline{\mathcal{V}}}$ (or $\mathcal{S}_{r',\bar{\mathcal{V}}}$) for $r' \in \{1, 2, \dots, t_r - 2, \bar{R}\}$. That is, $\sum_{s \in \mathcal{S}_{r',\underline{\mathcal{V}}}} P_{(r,v) \rightarrow s}^a = \sum_{s \in \mathcal{S}_{r',\underline{\mathcal{V}}}} P_{(r,v') \rightarrow s}^a$ and $\sum_{s \in \mathcal{S}_{r',\bar{\mathcal{V}}}} P_{(r,v) \rightarrow s}^a = \sum_{s \in \mathcal{S}_{r',\bar{\mathcal{V}}}} P_{(r,v') \rightarrow s}^a$.

From Section III, the reputation and view transition probabilities are independently. Thus, for $(r, v) \in \mathcal{S}_{r,\underline{\mathcal{V}}}$ (or $(r, v) \in \mathcal{S}_{r,\bar{\mathcal{V}}}$), we have $\sum_{s \in \mathcal{S}_{r',\underline{\mathcal{V}}}} P_{(r,v) \rightarrow s}^a = P_{r \rightarrow r'}^a \sum_{v'' \in \underline{\mathcal{V}}} \mathbf{T}^L(v, v'') = P_{r \rightarrow r'}^a (P_a + \frac{2(1-P_a)}{M-1})$. Similarly, for another state (r, v') in the same subspace $\mathcal{S}_{r,\underline{\mathcal{V}}}$ (or $\mathcal{S}_{r,\bar{\mathcal{V}}}$), we also have $\sum_{s \in \mathcal{S}_{r',\underline{\mathcal{V}}}} P_{(r,v') \rightarrow s}^a = P_{r \rightarrow r'}^a (P_a + \frac{2(1-P_a)}{M-1})$. Therefore, both states (r, v) and (r, v') have the same probability to transition to another state subspace $\mathcal{S}_{r',\underline{\mathcal{V}}}$. Similarly, we can prove that (r, v) and (r, v') have the same probability to transition to another subspace $\mathcal{S}_{r',\bar{\mathcal{V}}}$.

- We then show that with the same action a , the two states (r, v) and (r, v') in the same subspace $\mathcal{S}_{r,\underline{\mathcal{V}}}$ (or $\mathcal{S}_{r,\bar{\mathcal{V}}}$) have the same expected short-term utility. That is, $U_{r,v}^a = U_{r,v'}^a$ for any $a \in \mathcal{A}$.

From the discussion in Section III-B2, when a user has reputation $r < t_r - 1$, he does not help anyone and also no one helps him. Thus, we have $U_{r,v}^a = U_{r,v'}^a = 0$. When $r \geq t_r - 1$, the short-term utility is $U_v^a = -c \sum_{r_j=a}^R \mathbf{x}(r_j) + (1 - P_{\bar{R} \rightarrow 1}^a) G_{\min(r+1, R), v} + P_{\bar{R} \rightarrow 1}^a G_{1, v}$. Note that the first term $-c \sum_{r_j=a}^R \mathbf{x}(r_j)$ depends on the action a only, and thus, the same action taken at (r, v) and (r, v') introduces the same cost.

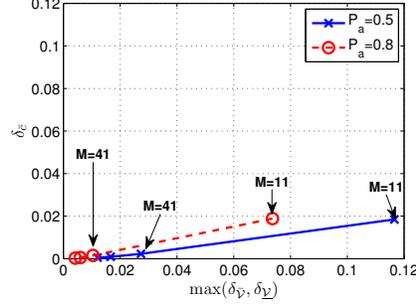


Fig. 13. $\delta_{\bar{c}}$ versus $\max(\delta_{\bar{\mathcal{V}}}, \delta_{\underline{\mathcal{V}}})$.

In addition, from (22), $G_{1,v} = \left(\sum_{r_k=t_r-1}^R \mathbf{x}(r_k) \right) \left(\sum_{v_k=1}^M \mathbf{v}(v_k) I[a_{v_k} \leq 1] \right) g_v = 0$, since no one will help a user with reputation 1. Similarly, $G_{1,v'} = 0$.

When comparing the two terms $G_{\min\{r+1,R\},v}$ and $G_{\min\{r+1,R\},v'}$, note that from (22), the only difference between the two terms are g_v and $g_{v'}$. From Proposition 3, with a large number of views and a large probability to transition to the anchor view, $g_v \approx g_{v'}$ when v and v' belong to the same view subset (either $\underline{\mathcal{V}}$ or $\bar{\mathcal{V}}$). Therefore, $G_{\min\{r+1,R\},v} \approx G_{\min\{r+1,R\},v'}$ and $U_{r,v}^a \approx U_{r,v'}^a$.

In summary, with a large M and a small difference between $G_{\min\{r+1,R\},v}$ and $G_{\min\{r+1,R\},v'}$, all states in the same subspace have the same probability to transition to other subspaces and receive approximately the same expected short-term utility.

From Definition 1, states with the same transition probability and expected short-term utility have bisimilarity relationship and can be aggregated as one state. In the following, for states with the same transition probability and approximately the same expected short-term utility, we study the impact of the difference in the expected short-term utility on the state aggregation and the equilibrium analysis. As an example, we consider the R -level ($R \geq 3$) reputation system and the full cooperation policy, and the analysis for other scenarios are similar.

Given the state partition $\{\mathcal{S}_{r,v}\}_{r=1,\dots,\bar{R},v=\underline{\mathcal{V}},\bar{\mathcal{V}}}$, if we aggregate all states in the same partition as one state, from the equilibrium analysis in Section IV-A4 and IV-A5, we can derive the condition for full cooperation policy to be a Nash equilibrium as $c \leq \bar{c}_1$ following Proposition 6.a,

When we do not aggregate the states in $\{\mathcal{S}_{r,v}\}_{r=1,\dots,\bar{R},v=\underline{\mathcal{V}},\bar{\mathcal{V}}}$ into one state, we can still derive the condition for full cooperation policy to be a Nash equilibrium as $c \leq \bar{c}_2$ following the analysis in Section IV-B for the general IMVS (with any number of anchor views).

To analyze how the difference in the expected short-term utility affects state aggregation and the full cooperation equilibrium analysis, we compare the above two thresholds \bar{c}_1 and \bar{c}_2 and define $\delta_{\bar{c}} = \frac{|\bar{c}_1 - \bar{c}_2|}{\bar{c}_2}$. When $\delta_{\bar{c}}$ is smaller, the difference between the two results (with and without state aggregation) is smaller.

Fig. 13 shows an example with $R = 10$, and the rest of the simulation setup is the same as Fig. 5. From Fig. 13, $\delta_{\bar{c}}$ decreases linearly as $\max(\delta_{\bar{v}}, \delta_{\underline{v}})$ decreases, and it drops below 1% when $\max(\delta_{\bar{v}}, \delta_{\underline{v}}) \leq 0.04$. Therefore, for single anchor view IMVS with a large number of views and a high probability to transition to the anchor view (e.g., $M \geq 31$ and $P_a \geq 0.5$), for all states in the same subspace $\mathcal{S}_{r,\underline{v}}$ ($\mathcal{S}_{r,\bar{v}}$), the impact of the difference in $U_{r,v}^a$ on the equilibrium analysis is negligible, and we can aggregate all states in $\mathcal{S}_{r,\underline{v}}$ ($\mathcal{S}_{r,\bar{v}}$) into one state to simplify the analysis. This completes the proof. ■

APPENDIX F

PROOF OF PROPOSITION 5

Proof: In this proof, we will examine each given policy and study when they are equilibriums. For each policy π , we first assume that all users use this policy and study the corresponding stationary reputation distribution \mathbf{x} . Then, following the one-shot deviation principle, we exam whether a user has incentive to unilaterally deviate to any one-shot deviation. Note that in the 2-level reputation system, we have $t_r = 2$ (i.e., all users have reputation belonging to $\bar{\mathcal{R}} = \{1, 2\}$), and $\gamma = \eta^{(t_r-1)L} = \eta^L$. Substitute them into (23), the lifetime utility of policy π for $v \in \{\underline{v}, \bar{v}\}$ is

$$\begin{aligned} W_v^\pi &= U_v^{a_v} + \left[\eta^L(1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v}) + \eta^L P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v} \right] [\mathbf{v}(\underline{v})W_{\underline{v}}^\pi + \mathbf{v}(\bar{v})W_{\bar{v}}^\pi] \\ &= U_v^{a_v} + \eta^L [\mathbf{v}(\underline{v})W_{\underline{v}}^\pi + \mathbf{v}(\bar{v})W_{\bar{v}}^\pi]. \end{aligned} \quad (54)$$

Similarly, with the one-shot deviation to a'_v , the lifetime utility becomes

$$W_v^{a'_v, \pi} = U_v^{a'_v} + \eta^L [\mathbf{v}(\underline{v})W_{\underline{v}}^\pi + \mathbf{v}(\bar{v})W_{\bar{v}}^\pi]. \quad (55)$$

Therefore, $W_v^\pi - W_v^{a'_v, \pi} = U_v^{a_v} - U_v^{a'_v}$, and we only need to compare the expected short-term payoffs when examine each policy.

- $\pi = \{a_{\underline{v}}, a_{\bar{v}}\} = \{2, 2\}$: By solving (20), we have $y = \mathbf{x}(2) = 1$ and $\mathbf{x}(1) = 0$. At view $v \in \{\underline{v}, \bar{v}\}$, with action 2, a user will help upload with probability 1. His reputation will be lowered to 1 with probability $P_{\bar{\mathcal{R}} \rightarrow 1}^2 = 0$, and he is a beneficial user with probability 1. Since other users all have reputation no less than $t_r - 1 = 1$, and use the policy $\{2, 2\}$, he can receive others' help whenever he needs in the next L segment, and $G_{2,v} = g_v$. Thus, the expected immediate payoff is $U_v^{a_v=2} = -c + g_v$.

As discussed in Section III, action $t_r = 2$ dominates action 1, we only need to exam the one-shot deviation to $R + 1 = 3$. By taking action 3, he will help to upload with probability 0, and his reputation falls to 1 with probability $P_{\bar{\mathcal{R}} \rightarrow 1}^3 = 1$. Thus, he cannot receive others' help, and the expected immediate gain 0. Therefore, the expected short-term payoff by taking action $R + 1$ is $U_v^{a'_v=3} = 0$.

Comparing $U_v^{a_v=2}$ and $U_v^{a'_v=3}$, we have $U_v^{a_v=2} - U_v^{a'_v=3} = g_v - c$. Given that $g_{\bar{v}} > g_{\underline{v}}$, if $g_{\underline{v}} \geq c$, then $U_v^{a_v=2} - U_v^{a'_v=3} \geq 0$ for all $v \in \{\underline{v}, \bar{v}\}$, and thus, $\{a_{\underline{v}}, a_{\bar{v}}\} = \{2, 2\}$ is an equilibrium policy.

- $\pi = \{a_{\underline{v}}, a_{\bar{v}}\} = \{3, 2\}$: Similar to the above analysis, by solving (20), we have $y = \mathbf{x}(2) = \frac{1}{1+\mathbf{v}(\underline{v})}$ and $\mathbf{x}(1) = \frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})}$. We then first exam the one-shot deviation principle at view \underline{v} . By taking action $a_{\underline{v}} = 3$, he will help upload with probability 0. His reputation falls to 1 with probability $P_{\bar{R} \rightarrow 1}^3 = \frac{1}{1+\mathbf{v}(\underline{v})}$ and he is a beneficial user with probability $\frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})}$. Since other users all have reputation no less than $t_r - 1 = 1$ and they only cooperate at \bar{v} with probability $\mathbf{v}(\bar{v})$, thus, his expected short-term gain is $\frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})} G_{2,\underline{v}} = \frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})} \mathbf{v}(\bar{v}) g_{\underline{v}}$. Then, his expected short-term payoff is $U_{\underline{v}}^{a_{\underline{v}}=3} = \frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})} \mathbf{v}(\bar{v}) g_{\underline{v}}$.

Since action $t_r = 2$ dominates action 1, we only need to study the one-shot deviation to $a'_{\underline{v}} = 2$. By taking action $a'_{\underline{v}} = 2$, he will help upload with probability $\frac{1}{1+\mathbf{v}(\underline{v})}$. His reputation will be lowered to 1 with probability $P_{\bar{R} \rightarrow 1}^2 = 0$, and he is a beneficial user with probability 1. Thus, he will receive the expected short-term gain $g_{\underline{v}} \mathbf{v}(\bar{v})$, and his expected immediate payoff is $U_{\underline{v}}^{a'_{\underline{v}}=2} = -\frac{c}{1+\mathbf{v}(\underline{v})} + g_{\underline{v}} \mathbf{v}(\bar{v})$. Compare $U_{\underline{v}}^{a_{\underline{v}}=3}$ and $U_{\underline{v}}^{a'_{\underline{v}}=2}$, and we have

$$U_{\underline{v}}^{a_{\underline{v}}=3} - U_{\underline{v}}^{a'_{\underline{v}}=2} = \frac{\mathbf{v}(\underline{v})}{1+\mathbf{v}(\underline{v})} g_{\underline{v}} \mathbf{v}(\bar{v}) - \left(-\frac{c}{1+\mathbf{v}(\underline{v})} + g_{\underline{v}} \mathbf{v}(\bar{v}) \right) = \frac{c - g_{\underline{v}} \mathbf{v}(\bar{v})}{1+\mathbf{v}(\underline{v})}. \quad (56)$$

Thus, to resist one-shot deviation at \underline{v} , we should have $c - g_{\underline{v}} \mathbf{v}(\bar{v}) \geq 0$.

We then examine the one-shot deviation principle at view \bar{v} . Following a similar procedure to the above analysis we have

$$U_{\bar{v}}^{a_{\bar{v}}=2} - U_{\bar{v}}^{a'_{\bar{v}}=3} = \frac{g_{\bar{v}} \mathbf{v}(\bar{v}) - c}{1+\mathbf{v}(\bar{v})}. \quad (57)$$

Thus, to resist one-shot deviation at \bar{v} , we should have $g_{\bar{v}} \mathbf{v}(\bar{v}) - c \geq 0$.

To summarize, only when $g_{\bar{v}} \mathbf{v}(\bar{v}) \geq c \geq g_{\underline{v}} \mathbf{v}(\bar{v})$, we have both $U_{\underline{v}}^{a_{\underline{v}}=3} - U_{\underline{v}}^{a'_{\underline{v}}=2} \geq 0$ and $U_{\bar{v}}^{a_{\bar{v}}=2} - U_{\bar{v}}^{a'_{\bar{v}}=3} \geq 0$, and $\{a_{\underline{v}}, a_{\bar{v}}\} = \{3, 2\}$ is an equilibrium policy.

- $\pi = \{a_{\underline{v}}, a_{\bar{v}}\} = \{3, 3\}$: In this case, $y = \mathbf{x}(2) = 0.5$ and $\mathbf{x}(1) = 0.5$. Since no user cooperates, no user can gain from others' help and $G_{2,v} = G_{1,v} = 0$, while taking action $t_r = 2$ and cooperating with beneficial users only introduces a cost due to helping upload with probability 0.5. Thus, using action $R + 1 = 3$ and playing non-cooperatively is a dominant strategy. Therefore, $\{a_{\underline{v}}, a_{\bar{v}}\} = \{3, 3\}$ is always an equilibrium policy.

- $\pi = \{a_{\underline{v}}, a_{\bar{v}}\} = \{2, 3\}$: This action policy is symmetric with $\{a_{\underline{v}}, a_{\bar{v}}\} = \{3, 2\}$ that we discussed earlier. Thus, the cost range for $\{a_{\underline{v}}, a_{\bar{v}}\} = \{2, 3\}$ being an equilibrium policy can be symmetrically written as $g_{\underline{v}} \mathbf{v}(\underline{v}) \geq c \geq g_{\bar{v}} \mathbf{v}(\underline{v})$. However, since $g_{\bar{v}} > g_{\underline{v}}$, this range is empty, and thus, $\{a_{\underline{v}}, a_{\bar{v}}\} = \{2, 3\}$ cannot be an equilibrium policy. ■

APPENDIX G

PROOF OF PROPOSITION 6

Proof: In Section IV-A5, we prove Proposition 6.a. In this appendix, we prove Proposition 6.b to 6.d.

First, note that (23) gives the lifetime utility with aggregated views, and it is a linear system that can be easily solved. The solution to (23) is

$$\begin{cases} W_{\underline{\mathcal{V}}}^{\pi} = \frac{U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}} - \mathbf{v}(\bar{\mathcal{V}})[((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}} - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}}]}{1 - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\mathbf{v}(\bar{\mathcal{V}}) - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\mathbf{v}(\underline{\mathcal{V}})}, \\ W_{\bar{\mathcal{V}}}^{\pi} = \frac{U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}} + \mathbf{v}(\underline{\mathcal{V}})[((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}} - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}}]}{1 - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\bar{\mathcal{V}}}})\mathbf{v}(\bar{\mathcal{V}}) - ((1-P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\eta - \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_{\underline{\mathcal{V}}}})\mathbf{v}(\underline{\mathcal{V}})}. \end{cases} \quad (58)$$

Similar to the analysis in Section IV-A5 and Appendix F, we then derive the lifetime utility with a one-shot deviation to a'_v at view $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, and we have

$$W_v^{a'_v, \pi} = U_{\underline{\mathcal{V}}}^{a'_v} + \left[\eta^L (1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v}) + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v} \right] \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right]. \quad (59)$$

Comparing W_v^{π} with its one-shot deviation $W_v^{a'_v, \pi}$, we have

$$\begin{aligned} W_v^{\pi} - W_v^{a'_v, \pi} &= U_{\underline{\mathcal{V}}}^{a_v} - U_{\underline{\mathcal{V}}}^{a'_v} + \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi} \right] \\ &\quad \times \left\{ \left[\eta^L (1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v}) + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a_v} \right] - \left[\eta^L (1 - P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v}) + \gamma P_{\bar{\mathcal{R}} \rightarrow 1}^{a'_v} \right] \right\}. \end{aligned} \quad (60)$$

When $W_v^{\pi} - W_v^{a'_v, \pi} \geq 0$, the policy can resist one-shot deviation to a'_v and is an equilibrium. Since action t_r and $R+1$ dominate all other actions, we only need to study the one-shot deviation to t_r or $R+1$.

• Proof of Proposition 6.b:

Assume that all users follow the policy $\pi = \{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$, we first solve (20) and find the stationary reputation distribution. We have $y = \frac{\sqrt{(1+\mathbf{v}(\underline{\mathcal{V}}))^2 + 4(t_r-2)\mathbf{v}(\underline{\mathcal{V}})} - (1+\mathbf{v}(\underline{\mathcal{V}}))}{2(t_r-2)\mathbf{v}(\underline{\mathcal{V}})}$, and $\mathbf{x}(r) = \frac{1-y}{t_r-1}$ for $1 \leq r \leq t_r - 1$.

We first consider the view set $\underline{\mathcal{V}}$, and derive a user's expected short-term utility $U_{\underline{\mathcal{V}}}^a = -c \sum_{r=a}^R \mathbf{x}(r) + (1 - P_{r \rightarrow 1})^1 G_{\min\{r+1, R\}, \underline{\mathcal{V}}} + P_{r \rightarrow a}^a G_{1, \underline{\mathcal{V}}}$ when he follows the policy $\pi = \{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$. For a user at view $\underline{\mathcal{V}}$ with reputation no less than $t_r - 1$, by taking action $a_{\underline{\mathcal{V}}} = R+1$, he will help upload with probability 0, and therefore the cost term ($c \sum_{r=a}^R \mathbf{x}(r)$) in $U_{\underline{\mathcal{V}}}^a$ is 0. Furthermore, after this action, his reputation falls to 1 with probability $P_{\bar{\mathcal{R}} \rightarrow 1}^{R+1} = y$, and from (22), $G_{1, \underline{\mathcal{V}}} = 0$ since no one is willing to help a user with reputation 1. We then derive $G_{\min\{r+1, R\}, \underline{\mathcal{V}}}$, his expected short-term utility if his reputation stays at $\bar{\mathcal{R}}$. Since other users only cooperate when they have reputation no less than $t_r - 1$ and are at view $\bar{\mathcal{V}}$, from (22), we have $G_{\min\{r+1, R\}, \underline{\mathcal{V}}} = [y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}})g_{\underline{\mathcal{V}}}$. Thus, his expected short-term payoff is $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=R+1} = (1-y)g_{\underline{\mathcal{V}}}[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}})$.

If the user takes the one-shot deviation to $a'_{\underline{v}} = t_r$ at view \underline{v} , he will help to upload with probability y and cost term is $-cy$. Also, with action $a'_{\underline{v}} = t_r$, the probability that his reputation falls to 1 is $P_{\bar{R} \rightarrow 1}^{t_r} = 0$, and same as the above analysis, at reputation \bar{R} , his expected short-term gain is $G_{\min(r+1, R), \underline{v}} = g_{\underline{v}}[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})$. Thus, by taking one-shot deviation to $a'_{\underline{v}} = t_r$ at view \underline{v} , his expected short-term payoff $U_{\underline{v}}^{a'_{\underline{v}}=t_r} = -yc + g_{\underline{v}}[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})$.

Substitute the above terms $U_{\underline{v}}^{a'_{\underline{v}}=t_r}$, $P_{\bar{R} \rightarrow 1}^{t_r}$, $U_{\underline{v}}^{a_{\underline{v}}=R+1}$ and $P_{\bar{R} \rightarrow 1}^{R+1}$ into (58) and (60), and we have

$$W_{\underline{v}}^{\pi} - W_{\underline{v}}^{a'_{\underline{v}}=t_r, \pi} = y \left\{ c - \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\underline{v}}}{1 - \eta^L + \eta^L y - y\gamma} \right\}, \quad (61)$$

where $D_{\underline{v}} = \{\mathbf{v}(\bar{V})(\eta^L - \gamma)g_{\bar{v}} + [1 - \gamma - (\eta^L - \gamma)\mathbf{v}(\bar{V})]g_{\underline{v}}\}$. Thus, to resist the one-shot deviation at view \underline{v} , we need $c - \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\underline{v}}}{1 - \eta^L + \eta^L b - b\gamma} \geq 0$.

When he is at view \bar{v} , by taking action $a_{\bar{v}} = t_r$ defined in the policy, he will help upload with probability y and the cost term is $-cy$. Also, his reputation falls to 1 with probability $P_{\bar{R} \rightarrow 1}^{t_r} = 0$. Similar to the above analysis, his expected short-term gains are $G_{1, \bar{v}} = 0$ and $G_{\min(r+1, R), \bar{v}} = g_{\bar{v}}[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})$, and his expected short-term payoff is $U_{\bar{v}}^{a_{\bar{v}}=t_r} = -yc + g_{\bar{v}}[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})$.

For the one-shot deviation to $a'_{\bar{v}} = R + 1$ at view \bar{v} , following the same analysis as above, we have

$$W_{\bar{v}}^{\pi} - W_{\bar{v}}^{a'_{\bar{v}}=R+1, \pi} = y \left\{ \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\bar{v}}}{1 - \eta^L + \eta^L y - y\gamma} - c \right\}, \quad (62)$$

where $D_{\bar{v}} = \{[1 - \gamma - (\eta^L - \gamma)\mathbf{v}(\underline{v})(1 - y)]g_{\bar{v}} + (\eta^L - \gamma)\mathbf{v}(\underline{v})(1 - y)g_{\underline{v}}\}$. Therefore, to resist the one-shot deviation at view \bar{v} , we need $\frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\bar{v}}}{1 - \eta^L + \eta^L y - y\gamma} - c \geq 0$.

Note that

$$\begin{aligned} D_{\bar{v}} - D_{\underline{v}} &= \{(1 - \gamma) - (\eta^L - \gamma)[(1 - y) + y\mathbf{v}(\bar{V})]\}(g_{\bar{v}} - g_{\underline{v}}) \\ &> (\eta^L - \gamma)[y - y\mathbf{v}(\bar{V})](g_{\bar{v}} - g_{\underline{v}}) > 0. \end{aligned} \quad (63)$$

Thus, if

$$\frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\bar{v}}}{1 - \eta^L + \eta^L y - y\gamma} \geq c \geq \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{V})D_{\underline{v}}}{1 - \eta^L + \eta^L y - y\gamma}, \quad (64)$$

we have both $W_{\underline{v}}^{\pi} - W_{\underline{v}}^{a'_{\underline{v}}=t_r, \pi} \geq 0$ and $W_{\bar{v}}^{\pi} - W_{\bar{v}}^{a'_{\bar{v}}=R+1, \pi} \geq 0$, and therefore, $\{a_{\underline{v}}, a_{\bar{v}}\} = \{R + 1, t_r\}$ is an equilibrium policy.

- **Proof of Proposition 6.c:**

Given the policy $\{a_{\underline{v}}, a_{\bar{v}}\} = \{R + 1, R + 1\}$, if $t_r = 2$, we have $y = 0.5$ and $\mathbf{x}(1) = 0.5$. If $t_r \geq 3$, we have $y = \frac{\sqrt{t_r-1}-1}{t_r-2}$ and $\mathbf{x}(r) = \frac{1-y}{t_r-1}$ for $1 \leq r \leq t_r - 1$. However, no matter which t_r is used, since no user cooperates, no user can gain from others' help and $G_{r,v} = 0$ for all r and v . However, playing

cooperatively with action t_r only introduces a cost due to helping upload with probability y . Thus, using action $R + 1$ and playing non-cooperatively is a dominant strategy, from which no one will deviate.

• Proof of Proposition 6.d:

Given the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R+1\}$, note that it is symmetric with respect to the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$ that we discussed earlier. Thus, the cost range for $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R+1\}$ being an equilibrium policy can be symmetrically written as

$$\frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\underline{\mathcal{V}})E_{\underline{\mathcal{V}}}}{1 - \eta^L + \eta^L y - y\gamma} \geq c \geq \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\underline{\mathcal{V}})E_{\bar{\mathcal{V}}}}{1 - \eta^L + \eta^L y - y\gamma}, \quad (65)$$

where $E_{\underline{\mathcal{V}}} = [1 - \gamma - (\eta^L - \gamma)\mathbf{v}(\bar{\mathcal{V}})(1 - y)]g_{\underline{\mathcal{V}}} + (\eta^L - \gamma)\mathbf{v}(\bar{\mathcal{V}})(1 - y)g_{\bar{\mathcal{V}}}$

and $E_{\bar{\mathcal{V}}} = \mathbf{v}(\underline{\mathcal{V}})(\eta^L - \gamma)g_{\underline{\mathcal{V}}} + [1 - \gamma - (\eta^L - \gamma)\mathbf{v}(\underline{\mathcal{V}})]g_{\bar{\mathcal{V}}}$. Comparing $E_{\underline{\mathcal{V}}}$ and $E_{\bar{\mathcal{V}}}$, we have

$$\begin{aligned} E_{\bar{\mathcal{V}}} - E_{\underline{\mathcal{V}}} &= \{(1 - \gamma) - (\eta^L - \gamma)[(1 - y) + y\mathbf{v}(\underline{\mathcal{V}})]\}(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) \\ &> (\eta^L - \gamma)[y - y\mathbf{v}(\underline{\mathcal{V}})](g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) > 0. \end{aligned} \quad (66)$$

Thus, the cost range in (65) is empty, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R + 1, t_r\}$ is not an equilibrium policy. ■