Annals of Operations Research
https://doi.org/10.1007/s10479-023-05231-7

ORIGINAL RESEARCH

A game theoretic framework for distributed computing with dynamic set of agents

Swapnil Dhamal[1] · Walid Ben-Ameur[1] · Tijani Chahed[1] · Eitan Altman[2,5] · Albert Sunny[3] · Sudheer Poojary[4]

Accepted: 6 February 2023

Abstract

We consider a distributed computing setting wherein a central entity seeks power from computational providers by offering a certain reward in return. The computational providers are classified into long-term stakeholders that invest a constant amount of power over time and players that can strategize on their computational investment. In this paper, we model and analyze a stochastic game in such a distributed computing setting, wherein players arrive and depart over time. While our model is formulated with a focus on volunteer computing, it equally applies to certain other distributed computing applications such as mining in blockchain. We prove that, in Markov perfect equilibrium, only players with cost parameters in a relatively low range which collectively satisfy a certain constraint in a given state, invest. We infer that players need not have knowledge about the system state and other players' parameters, if the total power that is being received by the central entity is communicated to the players as part of the system's protocol. If players are homogeneous and the system consists of a reasonably large number of players, we observe that the total power received by the central entity is proportional to the offered reward and does not vary significantly

✉ Swapnil Dhamal
swapnil.dhamal@gmail.com

Walid Ben-Ameur
walid.benameur@telecom-sudparis.eu

Tijani Chahed
tijani.chahed@telecom-sudparis.eu

Eitan Altman
eitan.altman@inria.fr

Albert Sunny
albert@iitpkd.ac.in

Sudheer Poojary
sudheer.poojary@gmail.com

1 Télécom SudParis, Institut Polytechnique de Paris, Evry, France

2 INRIA Sophia Antipolis Méditerranée, Valbonne, France

3 Indian Institute of Technology, Palakkad, Palakkad, India

4 Qualcomm India Pvt. Ltd., Bengaluru, India

5 LIA, Avignon University, Avignon, France

despite the players' arrivals and departures, thus resulting in a robust and reliable system. We then study by way of simulations and mean field approximation, how the players' utilities are influenced by their arrival and departure rates as well as the system parameters such as the reward's amount and dispensing rate. We observe that the players' expected utilities are maximized when their arrival and departure rates are such that the average number of players present in the system is typically between 1 and 2, since this leads to the system being in the condition of least competition with high probability. Further, their expected utilities increase almost linearly with the offered reward and converge to a constant value with respect to its dispensing rate. We conclude by studying a Stackelberg game, where the central entity decides the amount of reward to offer, and the computational providers decide how much power to invest based on the offered reward.

Keywords Game theory · Stochastic game · Markov perfect equilibrium · Stackelberg game · Distributed computing · Volunteer computing

1 Introduction

A distributed computing system could be viewed as several providers of computational power contributing to solve large problems. In certain applications, a common central entity coordinates and utilizes the provided computational power (e.g., volunteer computing (Sarmenta, 2001b; Anderson and Fedak, 2006)), while in other applications, the computational providers compete for being the first to solve a problem (e.g., mining in blockchain (Zheng and Xie, 2018)). Where a common central entity is involved, the computational providers typically contribute to the central entity's power, which in turn could use the combined power to either fulfil its own computational needs (e.g., mining blocks or running demanding programs) or distribute it to the next level of requesters of power (e.g., by a computing service provider to its customers in a utility computing model). Based on the returns that the central entity expects from completing the tasks for which the computational power is being sought, it would usually decide the compensation or reward to be dispensed to the providers. This reward would be distributed among the providers based on their respective contributions. Throughout the paper, we will refer to the central entity in charge of the distributed computing system as the *center*.

A computational provider incurs a certain cost per unit time for investing a certain amount of power. A higher power investment by a provider is likely to fetch it a higher reward while also increasing its incurred cost, thus resulting in a trade-off. In practice, most computational providers are neither present constantly for providing their power, nor do they invest a constant amount of power when they are present. In view of this, we consider two types of computational providers, namely, long-term stakeholders that invest a constant amount of power over time and providers that arrive and depart over time as well as vary their invested power. This consideration becomes of particular significance when providers of the latter type are strategic, aiming to maximize their respective utilities, and their arrival or departure affects the competition among them for obtaining the offered reward. Specifically, such providers can harness this knowledge in order to strategize on the power that is to be invested, based on the presence of the other providers and their invested power. We will refer to computational providers that arrive and depart over time and strategize on their invested power as *players*. In this paper, we formulate a stochastic game where the players arrive and depart during a run of volunteer computing. We hence analyze how the players would invest in an equilibrium, from which no provider would want to deviate unilaterally. As we shall see, while we formulate our model considering a volunteer computing setting, our model applies equally well to decentralized settings such as mining in blockchain.

In order to formulate our game and analyze its equilibrium, it is important to understand what a stochastic game is and which equilibrium notion we consider. Stochastic games (Shapley, 1953) are a multiagent generalization of Markov decision processes (MDPs). In MDP, a player's payoff and probabilistic state transitions depend on the current state and the player's strategy; while in a stochastic game, they additionally depend on the strategies of all the other players. Similar to MDP, a stochastic game continues until it reaches a terminal state if there exists any, or continues indefinitely in absence of a terminal state. The natural equilibrium notion that we consider is Markov perfect equilibrium (MPE) (Maskin and Tirole, 2001), which is pertinent to stochastic games. MPE could be viewed as an adaptation of subgame perfect Nash equilibrium. Similar to policy in MDP, a player's MPE policy is a mapping from the state space to the player's strategy space; it indicates the player's strategy when the system is in any given state. A player determines its strategy in each state by foreseeing its effects on the state transitions and the resulting utilities, as well as the strategies of all the other players in each state. Just as a player's Nash equilibrium strategy is a best response to the other players' Nash equilibrium strategies, a player's MPE policy is a best response to the other players' MPE policies.

As described above, this paper focuses on a distributed computing setting wherein the center offers reward in return for the power invested by computational providers, which comprises long-term stakeholders that invest a constant amount of power and a dynamic set of players that invest strategically. We now list our contributions along with an overview of our results, and then highlight how our studied problem, model, and results compare with those in the literature.

1.1 Our contributions and results

- We propose a stochastic game model that captures the arrival and departure of players and their strategic computational investments in a typical distributed computing system. We formulate our model based on a continuous time Markov chain framework, and hence obtain the utility function of a player while accounting for the state transitions and policies of all the players. We show that interestingly, a closed-form expression can be obtained for the utility function (Sect. 2).
- Through a game theoretic analysis, we determine the players' MPE policies and prove that only players with cost parameters in a relatively low range which collectively satisfy a certain constraint in a given state, invest. We infer that players need not have knowledge about the system state and other players' parameters, if the total power that is being received by the center is communicated to the players as part of the system's protocol (Sect. 3).
- Using extensive simulations and mean field approximation, we study the effects of the arrival and departure rates and other system parameters on players' utilities. We observe that if players are homogeneous, their expected utilities are the highest when the competition in the system is the least, i.e., when the system consists of one player. In line with this, the players' expected utilities are maximized when their arrival and departure rates are such that the average number of players present in the system is typically between 1 and 2, since this leads to the system being in the condition of least competition with

high probability. The utilities increase almost linearly with the offered reward and converge to a constant value with respect to its dispensing rate. We moreover observe that their individual investment is the highest when the system consists of two players, and thereafter decreases almost inversely proportionally to the number of players present in the system; the total power received by the center increases with the number of players and converges to an amount proportional to the offered reward (Sect. 4).
- We present a Stackelberg game, where the central entity as the leader decides the amount of reward to offer, and the computational providers as the followers decide how much power to invest based on the offered reward. The amount of reward determined by the central entity influences the total power invested by the providers, which in turn influences the central entity's own utility. We show that under practically reasonable assumptions, the central entity's utility is a concave function of the offered reward; we harness this fact to analytically determine the optimal amount of reward that the central entity should offer in order to maximize its utility (Sect. 5).

1.2 Related work

In the literature, stochastic games have been extensively studied in terms of theory (Goeree and Holt, 1999) as well as applicability in queuing systems (Altman, 1996) , multiagent reinforcement learning (Bowling and Veloso, 2000) , networks (Fu and Kozat, 2013) , and complex living systems (Bellomo, 2008) , among other applications. We briefly describe some of the works that are relevant to ours, and position our work with respect to them. Altman and Shimkin (1998) consider a processor-sharing service system where the service rate to individual customers decreases with an increase in the load. Based on the observed load, an arriving customer's decision comprises whether to join the shared system or to use a constant-cost alternative such as a personal computer. The authors show that if customers aim to minimize their individual service times, any Nash equilibrium consists of threshold decision rules, with a threshold on the queue length in the shared system. Nahir et al. (2012) consider a similar setup with the difference that customers consider using the system over a long time scale and for multiple jobs. We consider a reverse of this setup, wherein players provide computational resources instead of receiving them. One could observe the difference in the obtained results; while a player's Nash equilibrium strategy follows a thresholding policy with a threshold on the number of players present in the system in (Altman and Shimkin, 1998), it follows a policy in our case that is smooth and non-monotone with respect to the number of players present in the system (since a player's investment is the highest when the system consists of two players and thereafter decreases almost inversely proportionally to the number of players present in the system). Wang and Zhang (2013) investigate Nash equilibrium and socially optimal strategies in a queuing system, where reentering the system (i.e., becoming a repeated customer) is a strategic decision of the customers. Based on their observation of the system and the underlying reward-cost structure, customers could employ a pure strategy such as reentering or balking, or a mixed strategy such as reentering with a certain probability. In our model, whether to reenter the system is not explicitly a strategic decision, however, deciding to invest zero amount of power is practically equivalent to deciding to be absent from the system.

Hu and Wellman (2003) generalize single-agent Q-learning to a noncooperative multiagent context by updating the Q-function based on the presumption that agents choose Nash equilibrium actions. In the framework of general-sum stochastic games, their proposed method is shown to converge under highly restrictive assumptions, and it is observed that agents are more likely to reach a joint optimal path with Nash Q-learning than with single-agent Q-learning. In contrast, we determine closed-form expressions for the equilibrium strategies directly; this is possible because we are able to obtain closed-form expression for a player's utility given the system state and players' strategies. Hassin and Haviv (2002) propose a version of subgame perfect Nash equilibrium for games with homogeneous players wherein the system state indicates the number of players present in the system, and each player selects a strategy based on its private information regarding the system state. Further, there exist works which develop algorithms for computing reasonably good, not necessarily optimal, strategies in a state-learning setting (Jiang et al., 2014; Wang et al., 2018) . In contrast to these works, our work focuses on analytically deriving equilibrium strategies and moreover, their closed-form expressions, in a setting where players have knowledge regarding either the system state or the total power that is being received by the center. Note that while the assumption of state knowledge is perhaps strong in most general applications, the assumption of having knowledge regarding the total power that is being received by the center is justifiable in a volunteer computing setting, since the total power could be made a common knowledge by the center in order to exhibit its transparency and trustworthiness for attracting players to be part of its system (we shall subsequently provide details on such practical aspects).

As noted by (Abraham et al., 2006; Kwok et al., 2005) , distributed systems have been studied from the game theoretic perspective. Mengistu and Che (2019) and Zheng and Xie (2018) respectively present surveys on the challenges in volunteer computing and blockchain systems, which are two of most prominent examples of modern-day distributed systems. In the literature, studies considering strategic aspects in volunteer computing have primarily focused on load balancing (Murata et al., 2008; Al Ridhawi et al., 2021) and sabotage-tolerance (Sarmenta, 2001a; Watanabe et al., 2009) , while those in blockchain have focused on selfish mining (Eyal and Sirer, 2014; Sapirshtein et al., 2016; Kwon et al., 2017) and pooled mining (Lewenberg et al., 2015; Eyal, 2015) . The aforementioned works on distributed systems do not consider game theoretic aspects of investment, which is the focus of our paper. Among the few works that consider game theoretic aspects of investment, the closest to ours are (Dimitri, 2017) and (Altman et al., 2020) (whose utility model is based on that of (Dimitri, 2017)). A critical shortcoming of this utility model is that it does not explicitly account for time (as acknowledged in (Dimitri, 2017)); in particular, the cost incurred does not account for the time spent for mining. Apart from the difference in the utility formulation, a fundamental difference is the formulation and analysis of a stochastic game resulting from the arrival and departure of players, hence the difference in the equilibrium notion (Markov perfect equilibrium versus Nash equilibrium) and the sets of analyses and simulations studied.

To summarize, there exist game theoretic studies for distributed systems in the literature, of which we have listed the representative works above. However, the aspect of strategic investment of power by computational providers has not been well studied. Furthermore, to the best of our knowledge, this work is the first to study the game theoretic aspects of distributed computing when the set of players is dynamic. In addition to proposing and analyzing a stochastic game framework from the providers' perspective, we study a Stackelberg game that also considers the central entity's perspective.

Table 1 Notation

r	Expected reward dispensed per segment
β	Rate of dispensing reward
c_i	Cost incurred by player i when it invests unit power for unit time
λ_i	Arrival rate corresponding to player i
μ_i	Departure rate corresponding to player i
$\mathcal{U}$	Universal set of players
ℓ	Amount of power apart from that invested by the players
S	Set of players currently present in the system
$x_i^{(S)}$	Strategy of player i in state S
$\mathbf{x}^{(S)}$	Strategy profile of players in state S
$\mathbf{x}$	Policy profile
$R_i^{(S,\mathbf{x})}$	Expected utility of player i computed in state S under policy profile $\mathbf{x}$

2 Our model

We now model a distributed computing system wherein players arrive and depart over time as well as strategize on the amount of power to be invested, while receiving a certain reward for providing their computational power. While we formulate our model considering a volunteer computing setting, our model applies equally well to decentralized settings such as mining in blockchain, as we shall discuss later. Table 1 presents the notation that we follow throughout the paper.

2.1 Model formulation

Consider a center which seeks power from computational providers, so as to utilize it for completing certain computational tasks. These tasks would generally be computationally demanding such as mining blocks, running simulations with a very large number of iterations, or finding a good enough solution to an NP-hard problem with a very large search space using randomized search. The center expects certain returns from completing a task, and would dispense reward to the providers either after the completion of a task or after certain amount of time that is determined by the center. Based on the task to be completed and the returns expected, the center can typically determine the amount of reward that it can dispense and the amount of time after which it can dispense the reward. The reward that is to be dispensed is distributed among the providers based on their respective contributions.

Since our study focuses on the setting where players stochastically arrive and depart over time, we can naturally formulate our model based on a continuous time Markov chain framework. As an overview, a state would correspond to the set of players present in the system, and the state transitions would comprise the arrival and departure of players and the dispensing of reward by the center. It is known that for preserving the Markov property that the past and future states be independent if conditioned on the current state, it is necessary that the time spent in each state has a memoryless property (i.e., the amount of additional time that would be spent in a state does not depend on the amount of time that has been already spent in the state) and is hence exponentially distributed. We shall see that this requirement is naturally satisfied in the studied setting. We now define the elements of the stochastic game that results from our modeling based on the continuous time Markov chain framework, namely, state space, reward, players' policies (i.e., their strategies corresponding to each state), state transitions and sojourn time corresponding to each state, and players' utility functions as computed in each state.

2.1.1 State space

Let $\mathcal{U}$ be the universal set of players, who arrive and depart over time. We consider a standard setting for modeling the arrivals and departures of players. A player j, who is not present in the system, arrives after time which is exponentially distributed with expected time $\frac{1}{\lambda_j}$. That is, λ_j is the arrival rate parameter corresponding to player j. A player can depart by shutting down its computer or by stopping/pausing its provision of the computational power to the volunteer computing system (e.g., for running its own computationally intensive tasks). A player j, who is present in the system, departs after time which is exponentially distributed with expected time $\frac{1}{\mu_j}$ (i.e., μ_j is the departure rate parameter corresponding to player j). The stochastic arrival and departure of players and the stochastic dispensing of reward, make the described process, a continuous time multi-state stochastic process. A state corresponds to the set of players present in the system, that is, the system is in state S if the set of players present in the system is S. Here, $S \subseteq \mathcal{U}$, i.e., $S \in 2^{\mathcal{U}}$. Throughout the paper, we unambiguously write $j \in \mathcal{U} \setminus S$ as $j \notin S$.

2.1.2 Reward

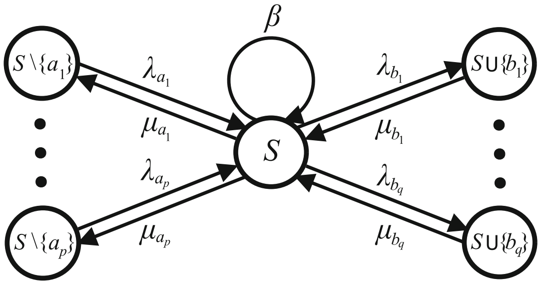
Let a *segment* be defined as the portion of time between consecutive instances of dispensation of reward. Consider that the length of a segment is exponentially distributed with rate parameter β. In other words, β is the rate of dispensing reward. Also, based on the returns it expects, consider that the center is willing to dispense an expected total reward of r per segment. As an example, if the center utilizes the received power for mining blocks, a segment can be imagined to be the time to mine a block, if the reward is dispensed when a block is mined. The center would determine how much reward it can dispense per block based on its expected returns. Since block mining typically is a memoryless process, the time taken to mine a block is exponentially distributed (see Appendix A for details); the memoryless property of mining and the exponential distribution of the mining time are well accepted conventions in the literature (Liu et al., 2019; Biais et al., 2019; Dimitri, 2017; Grunspan and Pérez-Marco, 2020).

Since the expected duration of a segment is $\frac{1}{\beta}$ and the expected total reward dispensed per segment is r, the reward can be spread over a segment in a continuous form such that the reward dispensed per unit time is $r\beta$. Consider that the reward at any time instant is allocated to the providers in proportion to their respective investments at that instant. Using the continuous form of the reward (i.e., $r\beta$ per unit time) and the aforementioned rule of allocating the reward, the center can maintain account of the reward amount that a player should receive and hence dispense it when the segment terminates.

2.1.3 Players' strategies and policies

The above-mentioned accounting in a continuous form lets players compute the amount of instantaneous reward they would receive if they invest a certain amount of power at that instant, without having to consider the history. We harness this *memoryless* property for

Fig. 1 A localized schema of the underlying continuous time Markov chain around a state S (here, $S = \{a_1, \ldots, a_p\}$ where $p = |S|$, and $\mathcal{U} \setminus S = \{b_1, \ldots, b_q\}$ where $q = |\mathcal{U}| - |S|$)



formulating our Markov decision process. In general, we consider that players are Markovian, that is, a player aims to maximize its expected utility from the current time onwards, without considering the history. A player can modulate its invested power at any time instant so as to maximize this utility.

Let the strategy of a player i indicating the amount of power that it would invest at time τ if the system is in state S, be denoted by $x_i^{(S,\tau)}$. As players are Markovian, a player has no incentive to change its investment amidst a state, if no other player changes its investment. Hence, we consider that no player changes its investment within a state, that is, $x_i^{(S,\tau)} = x_i^{(S,\tau')}$ for any τ, τ'. Thus, player i's investment strategy can now be written as a function of just the state, that is, $x_i^{(S)}$. For a state S where $j \notin S$, we have $x_j^{(S)} = 0$ by convention. A strategy profile of the players corresponding to a state S is a tuple comprising each player's strategy when the system is in state S; let it be denoted by $\mathbf{x}^{(S)} = (x_i^{(S)})_{i\in\mathcal{U}}$. As described earlier, a player's policy indicates its strategy when the system is in any given state. Let $\mathbf{x}_i = (x_i^{(S)})_{S\subseteq\mathcal{U}}$ denote the policy of player i. Similar to strategy profile, a policy profile of the players is a tuple comprising each player's policy; let us denote it by $\mathbf{x} = (\mathbf{x}_i)_{i\in\mathcal{U}}$. In addition, let ℓ be the amount of power received by the center apart from that invested by the players. This could be the center's own power or that invested by long-term stakeholders who invest a constant amount of power irrespective of the system state. Hence, the total amount of power received by the center in state S is $\sum_{j\in S} x_j^{(S)} + \ell$.

Now, since the portion of the reward that is allocated to a player at any given instant is proportional to its share of the total power received by the center at that instant, the reward allocated to player i per unit time when the system is in state S, is $\frac{x_i^{(S)}}{\sum_{j\in S} x_j^{(S)}+\ell} r\beta$. We denote by cost parameter c_i, the cost incurred by player i for investing unit amount of power for unit amount of time. So, the cost incurred by player i per unit time in state S is $c_i x_i^{(S)}$. Hence, we have that its profit per unit time in state S is $\frac{x_i^{(S)}}{\sum_{j\in S} x_j^{(S)}+\ell} r\beta - c_i x_i^{(S)}$.

2.1.4 State transitions and sojourn times

As mentioned earlier, the state transitions in the continuous time Markov chain underlying our model comprises the arrival and departure of players and the dispensing of reward by the center, and a transition occurs after time that is exponentially distributed with the corresponding rate parameter. Figure 1 presents a localized schema of the underlying chain showing the transitions from and into a state S along with their corresponding rate parameters. In general, the possible events that can occur in a state $S \in 2^{\mathcal{U}}$ are as follows:

1. the current segment ends with rate β and the system stays in state S for the next segment;

2. a player $j \notin S$ arrives with rate λ_j, and the system transits to state $S \cup \{j\}$;
3. a player $j \in S$ departs with rate μ_j, and the system transits to state $S \setminus \{j\}$.

We can understand a continuous time Markov chain as having two components, namely, (a) a parameter corresponding to each state specifying the distribution of the amount of time that would be spent in that state and (b) the jump chain describing the state transition probabilities (like in a discrete time Markov chain). It is clear that if the system is in state S, the amount of time until the occurrence of any of the above events is the minimum of the times until any of the above events occurs. Now, the minimum of exponentially distributed random variables, is another exponentially distributed random variable with rate which is the sum of the rates corresponding to the original random variables. So, the amount of time until the occurrence of any of the above events, is exponentially distributed with rate parameter $B^{(S)}$, where $B^{(S)} = \beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$. Hence, the *sojourn time* corresponding to state S for the current segment (i.e., the expected amount of time spent in state S until the occurrence of any of the above events) is $\frac{1}{B^{(S)}}$.

When any of the above events occurs, the system transits from state S according to the state transition probabilities. If an event occurs before any other event, the system transits to the target state corresponding to that event. It is known that the probability of an event occurring before any other event is equivalent to the corresponding exponentially distributed random variable being the minimum, which in turn, is proportional to its rate. Hence, if the system is in state S, the current segment ends before any arrival or departure event with probability $\frac{\beta}{B^{(S)}}$, a player $j \notin S$ arrives before any other event with probability $\frac{\lambda_j}{B^{(S)}}$, and a player $j \in S$ departs before any other event with probability $\frac{\mu_j}{B^{(S)}}$. Thus, the system can make the following transitions from a state $S \in 2^{\mathcal{U}}$:

1. the system advances to the next segment and stays in state S with probability $\frac{\beta}{B^{(S)}}$;
2. the system transits to state $S \cup \{j\}$ with probability $\frac{\lambda_j}{B^{(S)}}$;
3. the system transits to state $S \setminus \{j\}$ with probability $\frac{\mu_j}{B^{(S)}}$.

2.1.5 Utility function

As explained earlier, when the system is in state S, the profit made per unit time by player i is $\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}$, and the sojourn time in state S for the current segment is $\frac{1}{B^{(S)}}$. So, the net expected profit made by player i in state S before the system transits to another state or advances to the next segment, is $\frac{\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}}{B^{(S)}}$.

In economics, the utilities corresponding to future events are commonly considered to be discounted, that is, the utility corresponding to a future event is perceived to be lower at the present time as compared to at the time of its occurrence. In our model, this discounting could be owing to a number of reasons, one being the uncertainty regarding whether or not there would be a next segment. We consider that a player i perceives its utility to be discounted by a factor of $\delta \in [0, 1)$ for every future segment, where $\delta = 0$ means that the utility corresponding to only the current segment is valued.

Let $R_i^{(S,\mathbf{x})}$ denote the expected utility of player i as computed in state S. We now obtain an expression for $R_i^{(S,\mathbf{x})}$ using the above description, as summarized below:

1. the net expected profit made by player i for the current segment in state S before the system transits to another state, is $\frac{\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}}{B^{(S)}}$;
2. with probability $\frac{\beta}{B^{(S)}}$, the system stays in state S while advancing to the next segment, for which player i's expected utility is perceived as $\delta R_i^{(S,\mathbf{x})}$;
3. with probability $\frac{\lambda_j}{B^{(S)}}$, the system transits to state $S \cup \{j\}$ where player i's expected utility would be $R_i^{(S\cup\{j\},\mathbf{x})}$;
4. with probability $\frac{\mu_j}{B^{(S)}}$, the system transits to state $S \setminus \{j\}$ where player i's expected utility would be $R_i^{(S\setminus\{j\},\mathbf{x})}$.

Hence, player i's expected utility when the system is in state S can be recursively computed as:

$$R_i^{(S,\mathbf{x})} := \frac{\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}}{B^{(S)}} + \frac{\beta}{B^{(S)}} \cdot \delta R_i^{(S,\mathbf{x})} + \sum_{j \notin S} \frac{\lambda_j}{B^{(S)}} \cdot R_i^{(S\cup\{j\},\mathbf{x})} + \sum_{j \in S} \frac{\mu_j}{B^{(S)}} \cdot R_i^{(S\setminus\{j\},\mathbf{x})} \tag{1}$$

where $B^{(S)} = \beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$.

Note that while we formulated our model considering the use case of volunteer computing where the fraction of reward received is proportional to the invested power, it also applies to other applications such as mining in blockchain where the probability of winning the reward is proportional to the invested power. We refer the reader to Appendix A for a more detailed discussion. Furthermore, while we modeled the most general case of heterogeneous players, the cases of homogeneous players as well as multi-type players (which also have not been studied in the literature) are special cases of our model and analysis.

2.2 A closed-form expression for the expected utility

We now derive a closed-form expression for a player's expected utility from Eq. (1) which is recursive. Define an ordering $\mathcal{O}$ on sets which presents a one-to-one mapping from a set $S \subseteq \mathcal{U}$ to an integer between 1 and $2^{|\mathcal{U}|}$, both inclusive. Let $\mathbf{R}_i^{(\mathbf{x})}$ be the vector whose component $\mathcal{O}(S)$ is $R_i^{(S,\mathbf{x})}$. We now present the following convergence result and provide its proof in Appendix B. The proof is based on harnessing the fact that the transition matrix is strictly substochastic.

Lemma 1 *The recursive equation for $\mathbf{R}_i^{(\mathbf{x})}$, Eq. (1), converges for any policy profile $\mathbf{x}$.*

As the recursive equation for $R_i^{(S,\mathbf{x})}$ converges, the values of $R_i^{(S,\mathbf{x})}$ on both sides of Eq. (1) would be the same at convergence. Hence, bringing all terms containing $R_i^{(S,\mathbf{x})}$ to one side, we get that player i's expected utility as computed in state S is:

$$R_i^{(S,\mathbf{x})} = \frac{\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}}{D^{(S)}} + \sum_{j \notin S} \frac{\lambda_j}{D^{(S)}} \cdot R_i^{(S\cup\{j\},\mathbf{x})} + \sum_{j \in S} \frac{\mu_j}{D^{(S)}} \cdot R_i^{(S\setminus\{j\},\mathbf{x})} \tag{2}$$

where $D^{(S)} = (1-\delta)\beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$.

It is worth pointing out the change in the denominator from $B^{(S)}$ to $D^{(S)}$ where β is multiplied by a factor of $(1-\delta)$.

In order to derive a closed-form expression for the expected utility, define a matrix $\mathbf{W}$ of size $2^{|\mathcal{U}|} \times 2^{|\mathcal{U}|}$. When referring to element $W(\mathcal{O}(S), \mathcal{O}(S'))$, we use the shorthand $W(S, S')$ as it does not introduce any ambiguity. Let the elements of $\mathbf{W}$ be:

$$\begin{aligned} &\text{for } j \notin S: \; W(S, S \cup \{j\}) = \frac{\lambda_j}{D^{(S)}}, \\ &\text{for } j \in S: \; W(S, S \setminus \{j\}) = \frac{\mu_j}{D^{(S)}}, \\ &\text{and all other elements of } \mathbf{W} \text{ are } 0. \end{aligned} \tag{3}$$

Since $\beta > 0$ and $\delta < 1$, we have $D^{(S)} > \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$. Hence, the sum of the elements in each row of $\mathbf{W}$ is less than 1. That is, $\mathbf{W}$ is strictly substochastic.

Let $\mathbf{Z}_i^{(\mathbf{x})}$ be the vector whose component $\mathcal{O}(S)$ is $Z_i^{(S,\mathbf{x})}$, where

$$Z_i^{(S,\mathbf{x})} = \left(\frac{r\beta}{\sum_{j \in S} x_j^{(S)} + \ell} - c_i \right) \frac{x_i^{(S)}}{D^{(S)}} \tag{4}$$

Note that Eq. (2) can be written in matrix form as $\mathbf{R}_i^{(\mathbf{x})} = \mathbf{W}\mathbf{R}_i^{(\mathbf{x})} + \mathbf{Z}_i^{(\mathbf{x})}$, which gives $(\mathbf{I} - \mathbf{W})\mathbf{R}_i^{(\mathbf{x})} = \mathbf{Z}_i^{(\mathbf{x})}$. Since $\mathbf{W}$ is strictly substochastic, we obtain the following result presenting a closed-form expression for the expected utility.

Proposition 1 $\mathbf{R}_i^{(\mathbf{x})} = (\mathbf{I} - \mathbf{W})^{-1}\mathbf{Z}_i^{(\mathbf{x})}$.

While a general analysis of the concerned stochastic game when considering arbitrary forms of $\mathbf{W}$ and $\mathbf{Z}_i^{(\mathbf{x})}$ may not be tractable, we shall show that the analysis turns out to be tractable for the proposed model.

3 Analysis of Markov perfect equilibrium

It is known that in a finite player game with a finite state space and finite action spaces, if the horizon is either finite or infinite with the utility function being continuous at infinity, Markov perfect equilibrium (MPE) is guaranteed to exist (Maskin and Tirole, 2001). However, our considered game has infinite action spaces in each state and so, it cannot be inferred whether an MPE exists. In this section, we analyze MPE for our considered game, thus showing its existence, and hence discuss its properties. Recall that a player's MPE policy is a best response to the other players' MPE policies. Let the equilibrium utility of player i as computed in state S (while foreseeing the effects of its actions on the state transitions and the resulting utilities, as well as the MPE policies of other players) be denoted by $\hat{R}_i^{(S,\mathbf{x})}$. A general approach for determining an optimal policy in a single-agent MDP is using policy-value iterations to reach a fixed point. We can determine MPE in a similar way. In particular, for maximizing $\hat{R}_i^{(S,\mathbf{x})}$, we could assume that we have optimized for other states and use those values to find an optimizing $\mathbf{x}$ for maximizing $\hat{R}_i^{(S,\mathbf{x})}$. It is worth noting that for our model, we could determine the fixed point directly since we have a closed-form expression for vector $\mathbf{R}_i^{(\mathbf{x})}$ in terms of policy profile $\mathbf{x}$ (Proposition 1). Now, from Eq. (2), the Bellman equations over states $S \in 2^{\mathcal{U}}$ for player i can be written as:

$$\hat{R}_i^{(S,\mathbf{x})} = \max_{\mathbf{x}} \left\{ \frac{\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r\beta - c_i x_i^{(S)}}{D^{(S)}} + \sum_{j \notin S} \frac{\lambda_j}{D^{(S)}} \cdot \hat{R}_i^{(S \cup \{j\},\mathbf{x})} + \sum_{j \in S} \frac{\mu_j}{D^{(S)}} \cdot \hat{R}_i^{(S \setminus \{j\},\mathbf{x})} \right\}$$

where $D^{(S)} = (1-\delta)\beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$.

We now determine the MPE investment policy of each player, that is, the investment strategy of each player for each state, in MPE.

Proposition 2 *In MPE, a player i invests $x_i^{(S)} = \max\left\{\psi^{(S)}\left(1 - \frac{c_i \psi^{(S)}}{r\beta}\right), 0\right\}$, where $\psi^{(S)} = \sum_{j \in S} x_j^{(S)} + \ell = r\beta \frac{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j}}{2\sum_{j\in\hat{S}} c_j}$. Here, $\hat{S}$ is the maximal set of players $j \in S$ which collectively satisfy the constraints $c_j < \frac{r\beta}{\psi^{(S)}}$. Set $\hat{S}$ can be constructed iteratively by adding players j from set $S \setminus \hat{S}$ one at a time, in ascending order of c_j, until when adding a new player p to $\hat{S}$ violates the constraint $c_p < \frac{2\sum_{j\in\hat{S}} c_j}{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j}}$.*

Proof Recall that since $\mathbf{W}$ is a strictly substochastic matrix, $(\mathbf{I} - \mathbf{W})^{-1} = \lim_{t\to\infty} \sum_{\eta=0}^{t-1} (\mathbf{W})^\eta$. Since all the elements of $\mathbf{W}$ are non-negative, all the elements of $(\mathbf{W})^\eta$ also are non-negative for any natural number η, and hence all the elements of $(\mathbf{I}-\mathbf{W})^{-1}$ are non-negative. Also, since $\mathbf{R}_i^{(\mathbf{x})} = (\mathbf{I}-\mathbf{W})^{-1}\mathbf{Z}_i^{(\mathbf{x})}$ (Proposition 1) and since $\mathbf{W}$ is independent of $x_i^{(S)}$, maximizing the components of $\mathbf{Z}_i^{(\mathbf{x})}$ (namely, $Z_i^{(S,\mathbf{x})}$) individually with respect to $x_i^{(S)}$ would essentially maximize all the elements of $\mathbf{R}_i^{(\mathbf{x})}$. Recall that

$$Z_i^{(S,\mathbf{x})} = \left(\frac{\beta}{\sum_{j\in S} x_j^{(S)} + \ell} r - c_i\right) \frac{x_i^{(S)}}{D^{(S)}}.$$

where $D^{(S)} = \beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$.

It can be shown that $Z_i^{(S,\mathbf{x})}$ is a concave function w.r.t. $x_i^{(S)}$ (the second derivative is $\frac{-2r\ell\beta}{(\sum_{j\in S} x_j^{(S)}+\ell)^3 D^{(S)}}$). The first order condition $\frac{dZ_i^{(S,\mathbf{x})}}{dx_i^{(S)}} = 0$ gives

$$x_i^{(S)} = \left(\sum_{j\in S} x_j^{(S)} + \ell\right)\left(1 - \frac{c_i}{r\beta}\left(\sum_{j\in S} x_j^{(S)} + \ell\right)\right).$$

Let $\psi^{(S)} = \sum_{j\in S} x_j^{(S)} + \ell$. As $x_i^{(S)}$ is non-negative, we have

$$x_i^{(S)} = \max\left\{\psi^{(S)}\left(1 - \frac{\psi^{(S)}}{r\beta} c_i\right), 0\right\}. \tag{5}$$

Let $\hat{S} = \{j \in S : x_j^{(S)} > 0\}$. We later show how to determine set $\hat{S}$. Summing the above over all players in S and then adding ℓ on both sides, we get

$$\sum_{j\in S} x_j^{(S)} + \ell = \psi^{(S)}\left(|\hat{S}| - \frac{\psi^{(S)}}{r\beta}\sum_{j\in\hat{S}} c_j\right) + \ell.$$

Substituting $\sum_{j\in S} x_j^{(S)} + \ell$ as $\psi^{(S)}$, we get

$$\frac{1}{r\beta}\sum_{j\in\hat{S}} c_j \left(\psi^{(S)}\right)^2 - (|\hat{S}|-1)\psi^{(S)} - \ell = 0.$$

Note that if $\hat{S} = \emptyset$ (that is, $x_j^{(S)} = 0, \forall j \in S$), we have $|\hat{S}| = 0$ and $\sum_{j \in \hat{S}} c_j = 0$, in which case we obtain the trivial result $\psi^{(S)} = \ell$. Hence, consider $|\hat{S}| > 0$ and $\sum_{j \in \hat{S}} c_j > 0$. Solving the above equation for positive value of $\psi^{(S)}$, we get

$$\psi^{(S)} = r\beta \frac{|\hat{S}| - 1 + \sqrt{(|\hat{S}| - 1)^2 + \frac{4\ell}{r\beta} \sum_{j \in \hat{S}} c_j}}{2 \sum_{j \in \hat{S}} c_j}.$$

Substituting this expression for $\psi^{(S)}$ in Eq. (5) gives the MPE strategy of player i in state S.

So, $x_i^{(S)} > 0$ iff $c_i < \frac{2 \sum_{j \in \hat{S}} c_j}{|\hat{S}| - 1 + \sqrt{(|\hat{S}| - 1)^2 + \frac{4\ell}{r\beta} \sum_{j \in \hat{S}} c_j}}$. In other words, $i \in \hat{S}$ iff $c_i < \frac{2 \sum_{j \in \hat{S}} c_j}{|\hat{S}| - 1 + \sqrt{(|\hat{S}| - 1)^2 + \frac{4\ell}{r\beta} \sum_{j \in \hat{S}} c_j}}$. Now, it is mathematically possible for $\hat{S}$ to consist of players with higher cost parameters while excluding players with lower cost parameters (e.g., consider $\ell \to 0$, $S = \{1, 2, 3\}$, $c_1 = 1$, $c_2 = 2$, $c_3 = 4$; here $\hat{S}$ could be any of $\{1, 2\}, \{1, 3\}, \{2, 3\}$). However, since we are examining MPE, given such a set $\hat{S}$, a non-investing player with a lower cost parameter could unilaterally deviate to invest, which would hence lower the threshold cost parameter, thus compelling a previously investing player with a higher cost parameter to not invest. Hence, the constraint implies that if player i invests, then player j with $c_j < c_i$ also invests. So, there exists a threshold player $\hat{i}$ such that any player j with $c_j > c_{\hat{i}}$ would not invest. Hence, set $\hat{S}$ can be constructed iteratively (initiating from an empty set) by adding players j from set $S \setminus \hat{S}$ one at a time, in ascending order of c_j, until the above constraint is violated for the cost parameter of the newly added player. □

3.1 Practical aspects

We now briefly discuss certain practical aspects of our model and the result. We consider that a player can modulate its invested power as and when the system changes its state. As this may not be feasible every time in practice, the power can be modulated by a pre-configured automated software on the player's machine. The player can strategically devise its policy, that is, how much power to invest when the system is in a given state.

From Proposition 2, it can be seen that a player is not required to have knowledge about the arrival and departure rates, for determining its MPE policy. This is owing to the fact that a player's MPE utility $R_i^{(S,\mathbf{x})}$ computed in state S is a linear combination with constant non-negative weights, of $Z_i^{(S',\mathbf{x})}$ over all states S', which are mutually independent (that is, the value of $Z_i^{(S',\mathbf{x})}$ in a given state S' does not depend that in another state).

Furthermore, from Proposition 2, it may seem that in order to determine its MPE policy, a player is required to have knowledge about the system state and other players' cost parameters. However, note that if the total power $\psi^{(S)}$ that is being received by the center is known, player i's MPE investment $x_i^{(S)} = \max\left\{\psi^{(S)}\left(1 - \frac{c_i \psi^{(S)}}{r\beta}\right), 0\right\}$ does not require knowledge about the system state and other players' cost parameters.

With regard to the players having knowledge about the state that the system is in or the total power that is being received by the center, the state or the total power could be made a common knowledge by the center in order to exhibit its transparency and trustworthiness, so as to attract players to be part of its distributed system. A parallel to this can be drawn in the context of certain blockchain mining pools where a real-time dashboard shows information about the total power being invested by the pool's members. Furthermore, we shall see in

Sect. 4 that in MPE, a player invests more power when there are less players present in the system. Since this ensures that the center receives a decent amount of power even when there are less players present in the system, the center itself has an ulterior motive for making the state a common knowledge. Alternatively, players themselves could form a group wherein they share information about their arrivals and departures. The power invested by players not belonging to the group could then be thought of as being part of ℓ.

The above justifications are relevant when we aim to determine the precise and accurate investments in MPE by considering players to be heterogeneous. The assumption of players' parameters being a common knowledge can be bypassed if we consider players to be homogeneous. Similarly, if neither the system state nor the total power being received by the center is a common knowledge, a mean field approach could be employed wherein we study a player's investment in the 'average state'. We shall have a more elaborate discussion on these in the next section.

4 Sensitivity analysis

We inferred in the previous section that for players to determine their MPE policies, they are not required to have knowledge about the arrival and departure rates. However, it can be seen from Eq. (2) and Proposition 1 that the players' utilities would depend on these rates. Hence, in this section, we study the effects of these rates, as well as the other system parameters, on the utilities in MPE.

From Proposition 1, we can see that computing expected utility involves the computation of $(\mathbf{I} - \mathbf{W})^{-1}$, which in general, is arguably infeasible to obtain analytically as well as computationally for practical values of $|\mathcal{U}|$ since the number of states would be $2^{|\mathcal{U}|}$. So, for simplification, consider that the players are homogeneous; let their common arrival rate, departure rate, and cost parameter be λ, μ and c, respectively. With this simplification, from the system's perspective, the states corresponding to the players' sets can be mapped to their cardinalities. From a particular player i's perspective, the players' sets can be mapped to their cardinalities while also capturing whether they contain player i. Hence, the state space comprises the empty set, the universal set, and two states each (capturing whether or not the set contains player i) for all the other $|\mathcal{U}| - 1$ cardinalities. So, the total number of states is $2|\mathcal{U}|$, as opposed to $2^{|\mathcal{U}|}$ in the general case.

It can be seen that in the homogeneous case, if the number of players present in the system is s, the collective constraint on the cost parameters presented in Proposition 2 for players to invest can be written as $c < \frac{2sc}{s-1+\sqrt{(s-1)^2+\frac{4\ell}{r\beta}sc}}$. This simplifies to $c < \frac{r\beta}{\ell}$. If this constraint is not satisfied, no player invests, which would not be of interest. Hence, we consider that the values of parameters r, c, β, ℓ are such that $c < \frac{r\beta}{\ell}$.

Let us first understand the expected utility of a player i in the absence of state transitions (i.e., $\lambda = \mu = 0$) and only the current segment is considered (i.e., $\delta = 0$). For the state corresponding to the players' set having cardinality s and containing player i, let $x^{(s)}$ denote the player's MPE strategy and $V^{(s)}$ denote the aforementioned utility when $\lambda = \mu = 0$ and $\delta = 0$. Using Eq. (2), we can see that $V^{(s)} = \frac{\frac{x^{(s)}}{sx^{(s)}+\ell}r\beta - cx^{(s)}}{\beta}$ (note that $V^{(s)}$ is conceptually different from $Z_i^{(S,\mathbf{x})}$ [Eq. (4)] since in the latter, the arrival and departure rates as well as δ

are not 0). Now, from Proposition 2, we have that

$$x^{(s)} = \frac{r\beta\rho^{(s)}}{2sc}\left(1 - \frac{\rho^{(s)}}{2s}\right) \quad \text{and} \quad sx^{(s)} + \ell = \frac{r\beta\rho^{(s)}}{2sc} \tag{6}$$

where $\rho^{(s)} = s - 1 + \sqrt{(s-1)^2 + \frac{4\ell}{r\beta}sc}$.

Hence, $V^{(s)} = r\left(1 - \frac{\rho^{(s)}}{2s}\right)^2 = \frac{r}{4s^2}\left(s + 1 - \sqrt{(s-1)^2 + \frac{4\ell}{r\beta}sc}\right)^2$. So, the expected utility of player i in the current segment, when there are s players present in the system without transiting to another state, is

$$\frac{r}{4s^2}\left(s + 1 - \sqrt{(s-1)^2 + \frac{4\ell}{r\beta}sc}\right)^2 = V^{(s)}, \text{ if player } i \text{ is present} \\ \text{and } \; 0, \text{ if player } i \text{ is absent} \tag{7}$$

We now proceed to analyzing the expected utility for the general homogeneous case in the presence of state transitions. In what follows, we drop player i's specification in the notation since it does not introduce ambiguity in the homogeneous case.

4.1 Formulation of different types of expected utilities

Recall that a state captures the number of players present and whether or not the given player is present in the system. While a player can compute its expected utility in each state, it may not always be the case that a player knows the current state. However, the player can always compute the following types of expected utilities:

(a) $R_{\ni}$—the conditional expected utility given that it is currently present;
(b) $R_{\not\ni}$—the conditional expected utility given that it is currently absent;
(c) $\langle R \rangle$—the overall expected utility without having to know whether or not it is currently present in the system.

We now formulate the aforementioned types of expected utilities. In what follows, let $N = |\mathcal{U}|$. Let $R_{\ni}^{(s)}$ and $R_{\not\ni}^{(s)}$ denote the expected utilities as computed in the states corresponding to the given player being present and absent, respectively, when the number of players present in the system is s. We can hence write Eq. (2) for these two types of states, given the number of players present (s), as:

$$\text{For } s \in \{1, \ldots, N\}: \; R_{\ni}^{(s)} = \frac{\beta V^{(s)}}{D^{(s)}} + \frac{(N-s)\lambda}{D^{(s)}} R_{\ni}^{(s+1)} + \frac{(s-1)\mu}{D^{(s)}} R_{\ni}^{(s-1)} + \frac{\mu}{D^{(s)}} R_{\not\ni}^{(s-1)} \tag{8}$$

$$\text{For } s \in \{0, \ldots, N-1\}: \; R_{\not\ni}^{(s)} = \frac{(N-s-1)\lambda}{D^{(s)}} R_{\not\ni}^{(s+1)} + \frac{\lambda}{D^{(s)}} R_{\ni}^{(s+1)} + \frac{s\mu}{D^{(s)}} R_{\not\ni}^{(s-1)} \tag{9}$$

where $D^{(s)} = (1-\delta)\beta + (N-s)\lambda + s\mu$. Note that the second term vanishes in Eq. (8) for $s = N$, while the last term vanishes in Eq. (9) for $s = 0$.

Let $\langle R^{(s)} \rangle$ be the expected utility of the given player computed when the number of players present is s, without having to know whether or not the player is currently present. Given that the number of players present is s, without any additional information, the given player would be present with probability $\frac{s}{N}$ and absent with probability $\frac{N-s}{N}$. So, we have

$$\langle R^{(s)} \rangle = \frac{s}{N} R_{\ni}^{(s)} + \frac{N-s}{N} R_{\not\ni}^{(s)} \tag{10}$$

Now, in order to derive $R_\ni$, $R_{\not\ni}$ and $\langle R\rangle$, while the given player need not know the number of players present, it should have the probability distribution over the number of players present in the system. For deducing this distribution, we harness the fact that the underlying stochastic arrival and departure process resembles an *Engset's* system (Cohen, 1957) in queueing theory, which concerns a finite population size as in our model. Given population size N, arrival rate λ and departure rate μ, the probability $\mathbb{P}^N_{\lambda,\mu}(s)$ that the number of players present in the system is s, is:

$$\mathbb{P}^N_{\lambda,\mu}(s) = \binom{N}{s}\left(\frac{\lambda}{\lambda+\mu}\right)^s\left(\frac{\mu}{\lambda+\mu}\right)^{N-s}$$

It is known that the probability of a given player being present, or alternatively the fraction of time for which a given player is present, is $\frac{\lambda}{\lambda+\mu}$. The expected number of players present is hence $\frac{\lambda}{\lambda+\mu}N$. The mean duration of a full Engset cycle is $\frac{1}{\lambda}+\frac{1}{\mu}$.

As earlier, given that the number of players present is s, a given player would be present with probability $\frac{s}{N}$ and absent with probability $\frac{N-s}{N}$. So, the probability that the system consists of s players, with the given player present is $\mathbb{P}^N_{\lambda,\mu}(s)\frac{s}{N}$, and that with the given player absent is $\mathbb{P}^N_{\lambda,\mu}(s)\frac{N-s}{N}$. Hence, the overall probability of the given player being present is $\sum_s \mathbb{P}^N_{\lambda,\mu}(s)\frac{s}{N}$, and that of being absent is $\sum_s \mathbb{P}^N_{\lambda,\mu}(s)\frac{N-s}{N}$. Thus, we have

$$R_\ni = \frac{\sum_s R^{(s)}_\ni \mathbb{P}^N_{\lambda,\mu}(s)\frac{s}{N}}{\sum_s \mathbb{P}^N_{\lambda,\mu}(s)\frac{s}{N}} \tag{11}$$

$$R_{\not\ni} = \frac{\sum_s R^{(s)}_{\not\ni} \mathbb{P}^N_{\lambda,\mu}(s)\frac{N-s}{N}}{\sum_s \mathbb{P}^N_{\lambda,\mu}(s)\frac{N-s}{N}} \tag{12}$$

$$\langle R\rangle = \sum_s \mathbb{P}^N_{\lambda,\mu}(s)\langle R^{(s)}\rangle = \sum_s R^{(s)}_\ni \mathbb{P}^N_{\lambda,\mu}(s)\frac{s}{N} + \sum_s R^{(s)}_{\not\ni}\mathbb{P}^N_{\lambda,\mu}(s)\frac{N-s}{N}$$

As the probabilities of the given player being present and absent are also given by $\frac{\lambda}{\lambda+\mu}$ and $\frac{\mu}{\lambda+\mu}$, respectively, we can also write

$$\langle R\rangle = \frac{\lambda}{\lambda+\mu}R_\ni + \frac{\mu}{\lambda+\mu}R_{\not\ni} \tag{13}$$

While the above expressions can be evaluated numerically, they are not easy to analyze or get insights into. With the aim of obtaining simplified expressions for $R_\ni$, $R_{\not\ni}$ and $\langle R\rangle$, albeit approximate, we present a mean field approach.

4.2 A mean field approach

In our proposed mean field approach, we consider only two states, namely, $S_\ni$ and $S_{\not\ni}$ corresponding to given player being present and absent, respectively. The system is considered to invariably comprise an average number of players (say n), which is not affected by the arrival or departure of the given player. Being an Engset's system, we have that $n = \frac{\lambda}{\lambda+\mu}N$. From Eq. (7), the expected utility of player i in the current segment when there are s players present in the system without transiting to another state, is $V^{(s)}$ if the player is present when the utility is being computed, and 0 otherwise. So, in the mean field approach, the given player's utility would be $V^{(n)} = \frac{r}{4n^2}\left(n+1-\sqrt{(n-1)^2+\frac{4\ell}{r\beta}nc}\right)^2$ if computed in state

$S_{\ni}$, and 0 if computed in state $S_{\not\ni}$. Note that since $V^{(s)}$ is defined for $s \in \{1, \dots, N\}$, the interpolation $V^{(n)}$ holds valid for $n \in [1, N]$.

For computing the expected utility in state $S_{\ni}$ providing an approximation to $R_{\ni}$, the events that we need to account for are: (a) the player departing with rate μ, thus transiting the system to state $S_{\not\ni}$, in which the expected utility computed would be $R_{\not\ni}$, and (b) the segment terminating with rate β, in which case the system stays in state $S_{\ni}$ for the next segment where the expected utility would be perceived as $\delta R_{\ni}$. While in state $S_{\ni}$ for the current segment, the net expected profit made per unit time is $\beta V^{(n)}$ and the sojourn time is $\frac{1}{\beta+\mu}$. So, the net expected profit made in state $S_{\ni}$ before the system transits to $S_{\not\ni}$ or advances to the next segment, is $\frac{\beta V^{(n)}}{\beta+\mu}$. On similar lines, we can express the expected utility in state $S_{\not\ni}$ providing an approximation to $R_{\not\ni}$. Thus, we have:

$$\begin{aligned} R_{\ni} &\approx \frac{\beta V^{(n)}}{\beta+\mu} + \frac{\beta}{\beta+\mu}\delta R_{\ni} + \frac{\mu}{\beta+\mu} R_{\not\ni} \\ R_{\not\ni} &\approx \frac{\beta}{\beta+\lambda}\delta R_{\not\ni} + \frac{\lambda}{\beta+\lambda} R_{\ni} \end{aligned} \tag{14}$$

where $n = \frac{\lambda}{\lambda+\mu}N$. Solving the above two equations, we get the following closed-form expressions:

$$R_{\ni} \approx \frac{V^{(n)}}{1-\delta}\left(\frac{(1-\delta)\beta+\lambda}{(1-\delta)\beta+\lambda+\mu}\right) \tag{15}$$

$$R_{\not\ni} \approx \frac{V^{(n)}}{1-\delta}\left(\frac{\lambda}{(1-\delta)\beta+\lambda+\mu}\right) \tag{16}$$

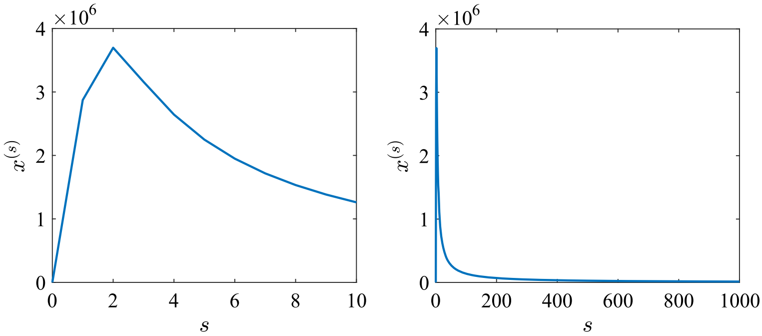
Also, since we know from Eq. (13) that $\langle R\rangle = \frac{\lambda}{\lambda+\mu}R_{\ni} + \frac{\mu}{\lambda+\mu}R_{\not\ni}$, we have:

$$\langle R\rangle \approx \frac{V^{(n)}}{1-\delta}\left(\frac{\lambda}{\lambda+\mu}\right) \tag{17}$$

Note that the above expression for $\langle R\rangle$ can also be viewed as $\frac{n}{N}V^{(n)}(1+\delta+\delta^2+\dots)$. This is consistent with our understanding that the mean field approach simplifies the system to contain the average number of players n, where the probability of the given player being present is $\frac{n}{N}$, thus rendering an expected utility of $\frac{n}{N}V^{(n)}$ for a segment. Considering also all the future segments and the discount factor associated with them, we obtain the aforementioned expression for $\langle R\rangle$. Further, it can be seen from Eqs. (15), (16) and (17) that $R_{\not\ni} < \langle R\rangle < R_{\ni}$, which is intuitive. As a concluding remark to the described mean field approach, we highlight that larger values of $n = \frac{\lambda}{\lambda+\mu}N$ would lead to better approximations, owing to the underlying assumption that the arrival or departure of the given player does not affect the number of players present in the system.

4.3 Effects of parameters on players' utilities

We now study the effects of the arrival and departure rates and other system parameters on the aforementioned different types of expected utilities of a player. In this study, we consider the use case of Bitcoin mining pools which are a form of volunteer computing systems. For supplementary details, a discussion on the applicability of our proposed model to mining in blockchain (e.g., Bitcoin mining) is provided in Appendix A. In this use case, the individual miners in the mining pool constitute the set of players $\mathcal{U}$, and the amount of power apart from that invested by this set of players (i.e., this mining pool) is ℓ. Further, the mining of a block



(a) $N = 10$ (b) $N = 1000$

Fig. 2 MPE investment of a player as a function of the number of players present in the system

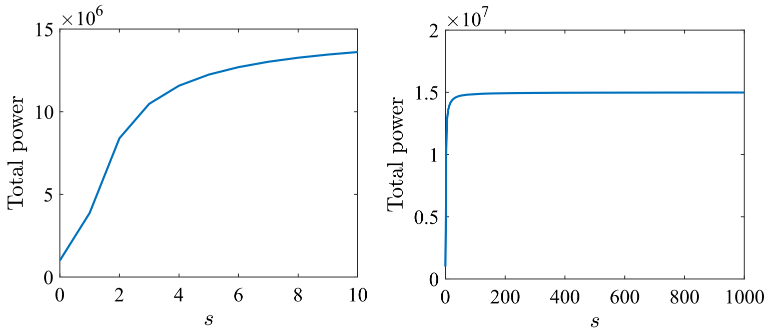
corresponds to a segment. With this in view, we now present the parameters' values that we consider in our numerical study with the help of practical references.

As of 2022, the amount of offered reward for successfully mining a block is 6.25 Bitcoins (Conway, 2022) ; this approximately translates to $\$3 \times 10^5$. The Bitcoin mining complexity is set with a target of finding new blocks once every 10 minutes on average (Conway, 2022) ; this translates to 6 blocks hour^{-1} on average. Hence, the amount of reward offered for a segment (i.e., mining of a block) is $\$3 \times 10^5$, and its dispensing rate is 6 hour^{-1}. For electricity costs, we consider the rate to be \$0.12 per kWh (Alves, 2022; Wong and McArdle, 2020) . Since the number of individual players (i.e., miners) in a mining pool is usually in the order of thousands, we consider $N = 1000$ (Bitpanda, 2021) . For our numerical study, we consider $\ell = 10^6$. Hence, unless specified otherwise, we consider $r = 3 \times 10^5$, $\beta = 6$, $c = 0.12$, $N = 10^3$, and $\ell = 10^6$. In what follows, we consider $\delta = 0$, that is, players consider the expected utility corresponding to only the current segment. The results for other values of $\delta \in (0, 1)$ are just scaled versions of the results for $\delta = 0$ and are qualitatively very similar.

We first study how a player's investment strategy is influenced by the number of players present in the system. Recall that since the considered parameters' values satisfy the constraint $c < \frac{r\beta}{\ell}$, a player would invest a positive amount of power if it is present. Figure 2 presents a player's MPE investment strategy as a function of the number of players present in the system. When $s = 0$, trivially $x^{(s)} = 0$ as no player is present. When $s = 1$, the player which is present faces no competition from any strategic agent; the only competition it faces is that due to ℓ which is a constant. However, when $s = 2$, it transforms into a game with two strategic agents, which is why the players are compelled to invest more power until they settle at their equilibrium investments. In order to understand the nature of the plot for $s > 2$, recall from Eq. (6) that $sx^{(s)} + \ell = \frac{r\beta\rho^{(s)}}{2sc}$, where $\rho^{(s)} = s - 1 + \sqrt{(s-1)^2 + \frac{4\ell}{r\beta}sc}$. So, we have $x^{(s)} = \frac{1}{s}\left(\frac{r\beta\rho^{(s)}}{2sc} - \ell\right)$. It can be easily shown that for $s > 2$, we have $\frac{dx^{(s)}}{ds} < 0$, implying that $x^{(s)}$ is a monotone decreasing function of s for $s > 2$. This explains the peak at $s = 2$. Beyond a certain value of s, that is, $s >> \frac{4\ell c}{r\beta}$, we have $\rho^{(s)} \approx 2s$. So, we have

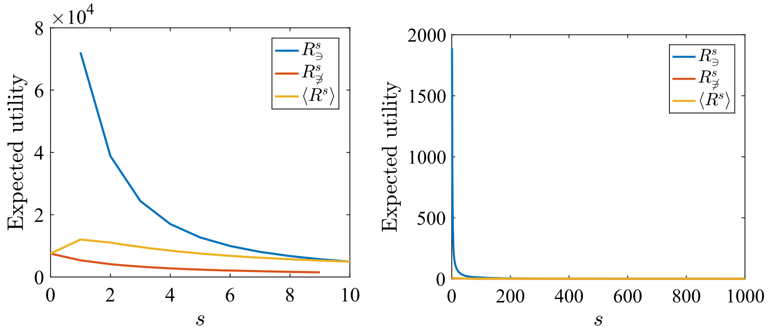
$$x^{(s)} \approx \frac{1}{s}\left(\frac{r\beta}{c} - \ell\right), \quad \text{for } s >> \frac{4\ell c}{r\beta} \tag{18}$$

In other words, beyond a certain value of s, $x^{(s)}$ is approximately inversely proportional to s.



(a) $N = 10$ (b) $N = 1000$

Fig. 3 Total power received by the center as a function of the number of players present in the system

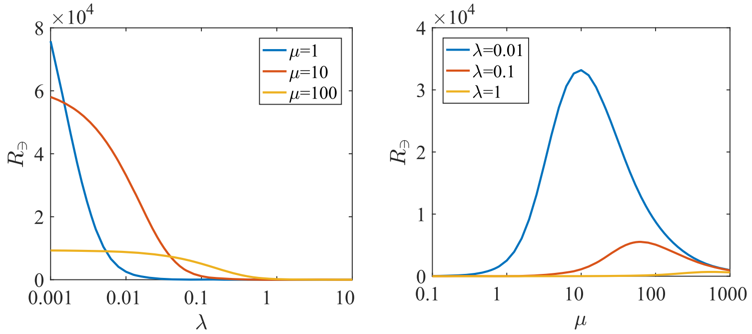


(a) $N = 10$ (b) $N = 1000$

Fig. 4 Expected utility of a player as computed when a certain number of players are present in the system

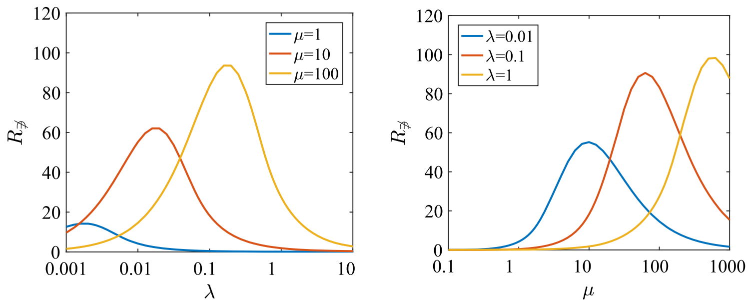
An intuition for the above observation is that when more players are present, their aggregate power can dominate power ℓ even if they do not invest large amounts of power individually, in which case the reward received by a player is almost inversely proportional to the number of players present. The reduced reward per unit amount of power invested is then compensated by the player by reducing its power invested so as to reduce the cost incurred. This strategy where players invest more when there are less players present, is beneficial for the center for continuing to receive a decent amount of power even when less players are present in the system. Figure 3 presents the effect of the number of players present on the total amount of power received by the center. The nature of this plot can again be understood from Eq. (6). In particular, it can be seen that as s grows to a large number, the total power converges to $\frac{r\beta}{c}$ (i.e., 1.5×10^7 for the considered parameters' values). Intuitively, as the expected reward being dispensed is bounded, it is natural that the total power received by the center would also be bounded.

Figure 4 presents how the number of players present affects the different types of expected utilities of a player, when we consider $\lambda = 1$, $\mu = 4$. Owing to their definitions, the plots for $R^{(s)}_{\ni}$ and $R^{(s)}_{\not\ni}$ do not have values at $s = 0$ and $s = N$, respectively. A first observation is that for a given value of s, the utility computed when the player is present is higher than when it is absent; this is true as the constraint $c < \frac{r\beta}{\ell}$ is satisfied, investing always fetches some reward. It can be seen that the plots of both $R^{(s)}_{\ni}$ and $R^{(s)}_{\not\ni}$ decrease monotonically. If



(a) w.r.t. λ for different values of μ (b) w.r.t. μ for different values of λ

Fig. 5 Effect of arrival and departure rates on a player's expected utility if the player is present in the system when the utility is being computed

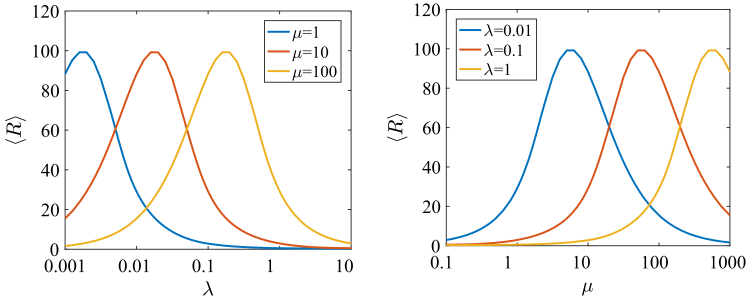


(a) w.r.t. λ for different values of μ (b) w.r.t. μ for different values of λ

Fig. 6 Effect of arrival and departure rates on a player's expected utility if the player is absent from the system when the utility is being computed

the given player is present, it is clearly advantageous to have less players present so that the player receives a larger share of the reward. If the player is absent, it is again advantageous to have less players present so that when the player arrives, it is likely to receive a larger share of the reward. Note that the number of players present would likely change by the time the player arrives, but the change around a smaller number of players is beneficial as compared to that around a larger number. Recall that $\langle R^{(s)}\rangle = \frac{s}{N} R_{\ni}^{(s)} + \frac{N-s}{N} R_{\not\ni}^{(s)}$, which justifies why its plot drifts away from $R_{\not\ni}^{(s)}$ and towards $R_{\ni}^{(s)}$ as s increases. When the utility is computed at $s = 0$, the reward is not received by any of the players until a player arrives. We can see that the plot for $\langle R^{(s)}\rangle$ has a peak at $s = 1$; this is where the system has the least competition (the only competition is due to ℓ) and the player is present with probability $\frac{s}{N}$. Also, note that $\frac{s}{N} V^{(s)}$ can be viewed as a myopic form of $\langle R^{(s)}\rangle$; it can be shown that $\frac{s}{N} V^{(s)}$ peaks at $s = 1$, which gives an idea for the peak at $s = 1$ for $\langle R^{(s)}\rangle$.

Figures 5, 6 and 7 illustrate the effects of arrival and departure rates on $R_{\ni}$, $R_{\not\ni}$ and $\langle R\rangle$, respectively. It can be seen that the plot of $R_{\ni}$ with respect to λ is monotone decreasing like a reverse sigmoid function, while all other plots are bell-shaped. If the player is present when the utility is being computed, it is advantageous if not many players arrive so that the reward is shared among less players; this explains the monotone decreasing plots in Fig. 5a. On the other hand, referring to Fig. 5b, as μ increases, it may be beneficial early on as more players



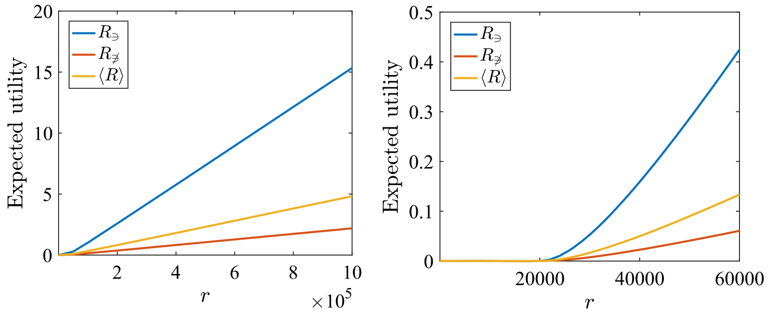
(a) w.r.t. λ for different values of μ (b) w.r.t. μ for different values of λ

Fig. 7 Effect of arrival and departure rates on a player's expected utility if the player is agnostic about its presence in the system when the utility is being computed

are likely to depart, thus resulting in the reward being shared among less players. However, beyond a certain value of μ, the disadvantage due to the increasing probability of the player itself departing and staying out of the system dominates the advantage of less players being present, which explains its non-monotonic nature. In Fig. 6a, the player is absent when the utility is being computed and so, an increase in λ would increase its probability of arriving and receiving a share of the reward while staying in the system. However, as λ increases, the number of players present is also likely to be higher and beyond a certain value, its disadvantage would dominate the advantage of the increases probability of the given player being present. In Fig. 6b, for an increase in μ initially, it is likely that when the given player arrives, the competition would be low, which would aid in obtaining a higher reward. However, beyond a certain μ, the fraction of time spent in the system once the player arrives, would be low and dominate the effect of less competition. On similar lines, the non-monotonicity in Fig. 7 can be explained as due to the trade-off between a lower competition owing to less players being present and a higher probability of the player itself being absent.

It can be seen that in the bell-shaped plots, the peaks occur when the average number of players $\frac{\lambda}{\lambda+\mu}N$ is between 1 and 2, typically close to 2. This is in line with our earlier observation that a player's overall expected utility is the highest when the system consists of one player, because if the average number of players in the system is such, the system consists of one player with high probability. Note, however, that if the average number of players in the system is close to 1, the probability of the system being in the state corresponding to zero players would be high, which would be highly disadvantageous to the player. Hence, though a player's overall expected utility is the highest when the system consists of one player, having a system with only one player on average is counterproductive.

It is also interesting to see that the plots of $R_{\ni}$ achieve a higher maxima for lower values of λ and μ (in Figs. 5a, b respectively), while the plots of $R_{\not\ni}$ achieve a higher maxima for higher values of λ and μ (in Figs. 6a, b respectively). This can be explained through the relation between the mean duration of a full Engset cycle (i.e., $\frac{1}{\lambda}+\frac{1}{\mu}$) and the expected time for the segment to terminate (i.e., $\frac{1}{\beta}$). Note that after a full Engset cycle, each player arrives and departs once on average and so, the effect of the given player being present or absent, diminishes. Hence, given the expected time after which the segment would terminate, if the player is present, it is beneficial if the Engset cycle takes a larger time portion of the segment, so as to take more advantage of its initial status of being present. This explains why the maxima of $R_{\ni}$ corresponding to lower values of the parameters are higher, since they result



(a) macro view (b) micro view

Fig. 8 Effect of r on a player's expected utility

in a longer Engset cycle. For the exact opposite reason, the maxima of $R_{\not\ni}$ corresponding to higher values of the parameters are higher, since they result in a shorter Engset cycle.

In Fig. 7, the plots of $\langle R\rangle$ for different values of λ and μ are just shifted versions of each other. Diving deeper, it can be observed that the plots only depend on the ratio $\frac{\lambda}{\mu}$ and not the individual values of λ and μ (which is why the peaks also are achieved for the same ratio $\frac{\lambda}{\mu}$). This is intuitive since if the player is agnostic about its presence in the system, it is unlike the dynamics for $R_{\ni}$ and $R_{\not\ni}$ where the values of λ and μ in relation to β could be of importance. The only critical factor would be the fraction of time that player would be present during the segment, without knowing whether or not it is currently present, and this depends only on the ratio (since $\frac{\lambda}{\lambda+\mu} = \frac{\frac{\lambda}{\mu}}{\frac{\lambda}{\mu}+1}$). This can also be understood from the mean field expression for $\langle R\rangle$ [Eq. (17)] whose value depends only on the ratio $\frac{\lambda}{\mu}$, as against those for $R_{\ni}$ and $R_{\not\ni}$ [Eqs. (15) and (16)] whose values are influenced by the individual values of λ and μ and not just their ratio.

We now study the effects of other parameters, namely, r, β and ℓ on the different types of expected utilities of a player. While studying the effect of a parameter, we consider the values of the other parameters to be as mentioned earlier. Further, in order to observe the asymmetry of $R_{\ni}$ and $R_{\not\ni}$ around $\langle R\rangle$, we consider $\lambda \neq \mu$, in particular, $\lambda = 1$, $\mu = 4$.

Figure 8 illustrates the effect of reward parameter r on the different of types of expected utilities. The constraint for a player to invest: $c < \frac{r\beta}{\ell}$, can be rewritten as $r > \frac{c\ell}{\beta}$ (i.e., 20000 for the considered parameters' values). So, a player invests only when the value of r is higher than this threshold, as can be seen from the micro view of the plot. In order to explain the effect of r on $R_{\ni}$, $R_{\not\ni}$ and $\langle R\rangle$, we refer to their mean field equations, namely, Eqs. (15), (16) and (17). Here, $V^{(n)} = \frac{r}{4n^2}\left(n+1-\sqrt{(n-1)^2+\frac{4\ell}{r\beta}nc}\right)^2$, where $n = \frac{\lambda}{\lambda+\mu}N$ (i.e., 200 for the considered parameters' values, which is large enough for the mean field approach to be a good approximation of the system). In order to see how accurately the mean field expressions approximate the values of the different types of expected utilities, we present the % errors corresponding to them in Fig. 9 (the error plots for $R_{\ni}$, $R_{\not\ni}$ and $\langle R\rangle$ almost coincide for the considered range of r). Owing to the reasonably good approximation, we can harness the mean field expressions owing to their simplified forms, for getting insights into our observations. For small values of r, it can be shown that $V^{(n)}$ and hence $R_{\ni}$, $R_{\not\ni}$ and $\langle R\rangle$, are convex in r, which is evident from the micro view of the plot in Fig. 8. For large values of r, we have $V^{(n)} \approx \frac{r}{n^2}$, that is, $V^{(n)}$ is close to linear in r. Hence, $R_{\ni}$, $R_{\not\ni}$ and $\langle R\rangle$ are also

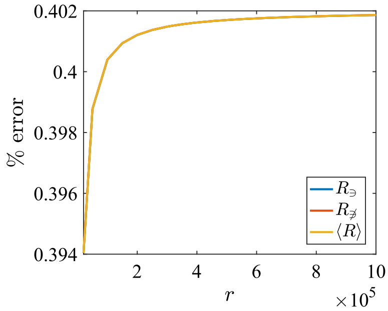
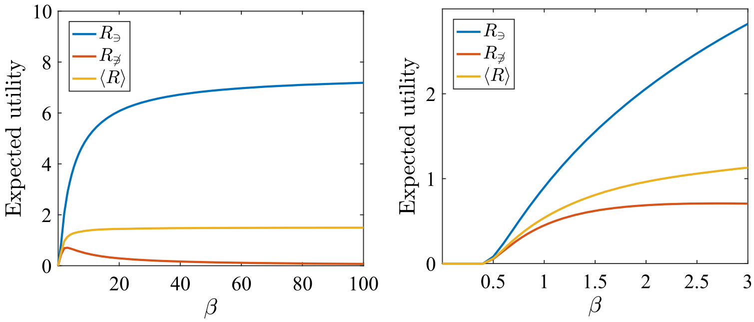


Fig. 9 Error in expected utility as a function of r if computed using mean field approach



(a) macro view (b) micro view

Fig. 10 Effect of β on a player's expected utility

close to linear in r for large values of r; their slopes can be deduced approximately from their mean field equations to be $\frac{1}{n^2(1-\delta)}\left(\frac{(1-\delta)\beta+\lambda}{(1-\delta)\beta+\lambda+\mu}\right)$, $\frac{1}{n^2(1-\delta)}\left(\frac{\lambda}{(1-\delta)\beta+\lambda+\mu}\right)$ and $\frac{1}{n^2(1-\delta)}\left(\frac{\lambda}{\lambda+\mu}\right)$, respectively (here, $\delta = 0$).

Figure 10 illustrates the effect of parameter β on $R_{\ni}$, $R_{\not\ni}$ and $\langle R \rangle$. The constraint for a player to invest can be written as $\beta > \frac{c\ell}{r}$ (i.e., 0.4 for the considered parameters' values), which can be seen from the micro view of the plot. It can be seen that $V^{(n)}$ increases monotonically with β. So, from Eqs. (15) and (17), it is clear that $R_{\ni}$ and $\langle R \rangle$ also increase monotonically with β, which we also observe in the plot. However, in the case of $R_{\not\ni}$ [Eq. (16)], the increase in $V^{(n)}$ dominates the decrease in $\frac{\lambda}{(1-\delta)\beta+\lambda+\mu}$ for very small values of β, while the domination reverses beyond a certain value of β, thus explaining its non-monotonicity. As β grows to large values, we can see that $V^{(n)}$ converges to approximately $\frac{r}{n^2}$. Hence, as can be deduced from the mean field equations, for large values of β, $R_{\ni}$, $R_{\not\ni}$ and $\langle R \rangle$ converge to approximately $\frac{r}{n^2(1-\delta)}$, 0 and $\frac{r}{n^2(1-\delta)}\left(\frac{\lambda}{\lambda+\mu}\right)$, respectively (i.e., 7.5, 0 and 1.5 for the considered parameters' values).

Figure 11 presents the effect of the value of ℓ on a player's expected utility. From the constraint for a player to invest, the threshold below which the value of ℓ should be, is $\frac{r\beta}{c}$ (i.e., 1.5×10^7 for the considered parameters' values); this can be observed from the plot. It can be shown that $V^{(n)}$ is a convex and decreasing function of ℓ. This, along with the mean field equations, implies that all the considered types of expected utilities: $R_{\ni}$, $R_{\not\ni}$ and $\langle R \rangle$, should be convex and decreasing in ℓ, which is what we see in the plot.

It can be observed from Figs. 8, 10 and 11 that the plot corresponding to $\langle R \rangle$ is closer to $R_{\not\ni}$ than to $R_{\ni}$, and it is easy to see why. From Eq. (13), $\langle R \rangle = \frac{\lambda}{\lambda+\mu}R_{\ni} + \frac{\mu}{\lambda+\mu}R_{\not\ni}$. As we

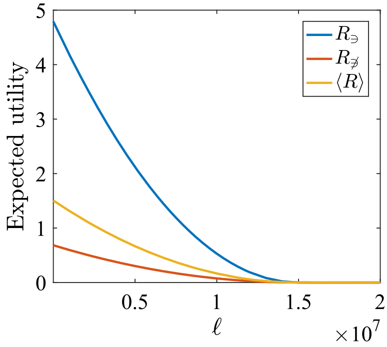


Fig. 11 Effect of ℓ on a player's expected utility

consider $\lambda = 1$ and $\mu = 4$ for these plots, we have that $\langle R \rangle$ gives weightage to $R_{\not\ni}$ that is four times of that given to $R_{\ni}$.

4.4 Practical interpretation of our results

We now present some practical interpretation of our results sequentially, in the context of the considered use case of Bitcoin mining pools and the corresponding set of parameters' values.

- If the system consists of a certain reasonably large number of players, the total power received by the center would not change significantly if the number of players changes to a certain extent; hence the system would be quite robust. Moreover, the total power would be proportional to the offered reward, thus giving the center adequate control with regard to the amount of power it would receive. Furthermore, an increase in the number of players would facilitate load balancing in terms of power share since the power invested individually by players would gradually decrease.
- Players' expected utilities are maximized when their arrival and departure rates are such that the average number of players present in the system is between 1 and 2 (typically close to 2). Hence, if players can strategize on their arrival and departure rates, a mediator could suggest rates to the players at which they should arrive and depart, and thus make the system socially optimal for the players (wherein the sum of players' utilities or alternatively, the expected utility of each player, is maximized).
- As long as the amount of offered reward per segment is higher than a certain value, players' utilities increase almost linearly with it. In this range, for the considered use case, a player's overall expected utility would increase by approximately 0.5 cents ($0.005) for every thousand units of increase in the offered reward. Hence, the center should take this effect into account while setting the reward so that players are incentivized enough to provide their power to the center by receiving a certain utility (especially if there is a competing center who also seeks power by offering a competitive reward amount).
- Changing the rate of dispensing reward, if it is beyond a certain value, would not change the players' utilities significantly. Hence, if the rate of dispensing reward is already reasonably high, the center could change it so as to increase its own utility resulting from any external factors dependent on this rate, without having to consider the effects on players' utilities.
- Since ℓ is the amount of (stable) power apart from that invested by the players, our results showcase how stable firms (albeit non-strategic) could influence the players' investment decisions and utilities. Players' utilities decrease almost linearly in ℓ over a wide range;

in this range, for the considered use case, a player's overall expected utility would drop by approximately 15 cents (\$0.15) for every million units of increase in ℓ. Moreover, if the center could control ℓ, it should be set not so high that players are discouraged from investing.

5 A Stackelberg game for determining optimal reward from the center's perspective

Till now, we have studied the game from the players' perspective while considering that the center is a non-strategic agent. In particular, we assumed that the reward being offered by the center is not dependent on the other parameters (namely, N, β, c, ℓ, λ and μ) and the players' strategic considerations. Since the center receives a certain amount of power in exchange of a certain amount of offered reward, and the received power can be used by the center for completing some task in order to obtain a certain amount of returns, there is every reason for the center to be strategic so as to balance the trade-off between the received power and the offered reward. Hence, we now expand our setting to incorporate the center also as an agent who can strategize on the amount of reward it has to offer. Specifically, we consider that the center's own returns depend on the total amount of power it receives, and the cost it incurs depends on the amount of reward it offers. We first observe how the offered reward influences the total power that is received by the center in a state.

Recall from Eq. (6) that when the offered reward is r and the number of players present is s (where $s \in \{1, \ldots, N\}$), we have that the total power, say $\psi_r^{(s)}$, equals $sx^{(s)} + \ell = \frac{r\beta\rho^{(s)}}{2sc}$, where $\rho^{(s)} = s-1+\sqrt{(s-1)^2 + \frac{4\ell}{r\beta}sc}$. It can be easily shown that $\frac{d\psi_r^{(s)}}{dr} > 0$ and $\frac{d^2\psi_r^{(s)}}{dr^2} < 0$. That is, the total power received by the center in a state is a monotone increasing concave function of r. Furthermore, for large values of r, it is clear that $\psi_r^{(s)}$ is close to linear in r. This is illustrated in Fig. 12 for different values of s (the parameters' values are same as those considered in Sect. 4). Note that in the general case where players could have different cost parameters, it is not clear from Proposition 2 how the total power received by the center in a state would change with a change in r, since the expressions for determining the set of investing players in any given state as well as their invested power are convoluted. However, we show that in MPE, the total power received by the center in any given state is a monotone increasing piecewise-concave (and close to piecewise-linear ramp) function of r. We provide the details in Appendix C.

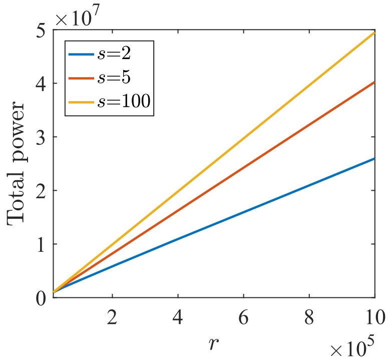


Fig. 12 Effect of the reward parameter r on the total power received by the center when the number of players present in the system is s

Now that we have shown the total power $\psi_r^{(s)}$ when there are s (where $s \in \{1, \ldots, N\}$) players present, to be a monotone increasing concave function of r, we proceed to analyze the nature of the expected total power received by the center as a whole. Let $\mathcal{T}(r)$ be the expected total power received by the center over all possible values of s, if it offers an expected reward of r per segment. Thus, we have that $\mathcal{T}(r) = \sum_{s=0}^{N} \mathbb{P}_{\lambda,\mu}^{N}(s)\psi_r^{(s)}$. Since $\psi_r^{(0)} = \ell$ by convention, we get:

$$\mathcal{T}(r) = \mathbb{P}_{\lambda,\mu}^{N}(0)\ell + \sum_{s=1}^{N} \mathbb{P}_{\lambda,\mu}^{N}(s)\psi_r^{(s)}$$

As $\psi_r^{(s)}$ is a monotone increasing concave function of r for $s \in \{1, \ldots, N\}$, their weighted sum (with positive weights) added to a constant, $\mathcal{T}(r)$, is also a monotone increasing concave function of r. Since the expected total power received (per unit time) by the center is $\mathcal{T}(r)$ and the expected duration of a segment is $\frac{1}{\beta}$, the total power received by the center over an entire segment is $\frac{\mathcal{T}(r)}{\beta}$, which also is a monotone increasing concave function of r.

We now study a Stackelberg game with the center as the leader and the computational providers as the followers, wherein the center decides the amount of reward to offer and the computational providers decide how much power to invest based on the offered reward. We model the center's utility as the difference between some relevant function of the total power received over a segment, and a function of the offered reward. Let $f(\cdot)$ denote the returns obtained by the center as a function of the total power received over a segment. In most practical applications, the returns would follow the law of diminishing returns and so, the returns would be a concave function in the total power received over a segment. Now, if $f(\cdot)$ is (weakly) concave and non-decreasing, and we already know that $\frac{\mathcal{T}(r)}{\beta}$ is a monotone increasing concave function of r, we have that the composition $f\left(\frac{\mathcal{T}(r)}{\beta}\right)$ is also a monotone increasing concave function of r. That is, the returns obtained by the center would be a concave function of r. Let $g(\cdot)$ denote the cost incurred by the center as a function of the offered reward. Thus, we can model the center's utility for a segment, say $U(r)$, to be:

$$U(r) = f\left(\frac{\mathcal{T}(r)}{\beta}\right) - g(r)$$

It is reasonable to consider the cost incurred due to dispensing the reward to be a monotone increasing linear or convex function of r. Hence, we have that the center's utility for a segment is a concave function of r, and hence can be maximized using calculus in order to obtain the optimal value of r. Note that if the future segments are also considered with a discount factor of $\delta \in [0, 1)$, the above expression would be multiplied by a factor of $1 + \delta + \delta^2 + \ldots = \frac{1}{1-\delta}$, which does not affect the optimal r.

The above analysis and procedure holds for any form of the utility function as long as functions $f(\cdot)$ and $g(\cdot)$ satisfy their respective conditions, which are natural in most real-world applications. In order to conduct a more precise analysis so as to derive concrete value of optimal r, we consider a specific form of the utility function. Specifically, we consider $g(r) = r$ and $f\left(\frac{\mathcal{T}(r)}{\beta}\right) = \alpha\left(\frac{\frac{\mathcal{T}(r)}{\beta}}{\frac{\mathcal{T}(r)}{\beta} + \frac{k}{\beta}}\right)$, where k could be viewed as the power invested by the other agents competing against the center for completing the concerned task (for which the center is expected to obtain returns; this task could be mining a block, for instance). That is, we have

$$U(r) = \alpha\left(\frac{\mathcal{T}(r)}{\mathcal{T}(r) + k}\right) - r$$

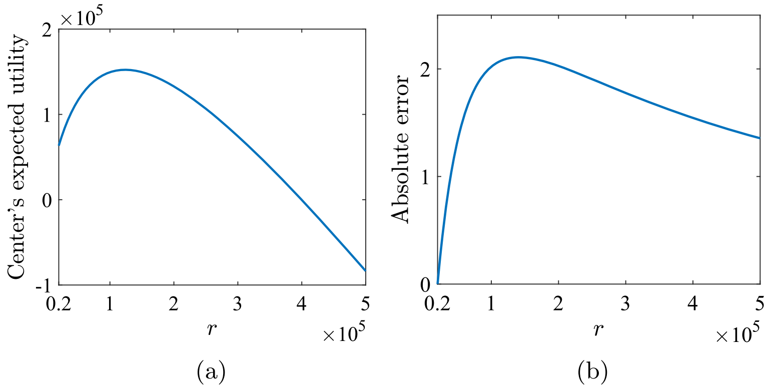


Fig. 13 **a** Effect of the reward parameter r on the center's expected utility and **b** absolute error in the center's expected utility if computed using mean field approach

For a numerical understanding of this utility function, consider $N = 1000, \beta = 6, c = 0.12, \ell = 10^6, \lambda = 1, \mu = 4, \alpha = 5 \times 10^5$, and $k = 5 \times 10^6$. Figure 13a illustrates how the center's expected utility $U(r)$ is influenced by the value of r (the values of r start from 0.2×10^5 owing to the constraint for the players to invest: $r > \frac{c\ell}{\beta}$). The plot is a concave function, as expected. It can be seen that if the value of r is set to be higher than a certain value, $U(r)$ the expected utility could be negative. The optimal value of the reward parameter for this utility function with this set of parameters is observed to be at $r = 123564.38$.

We can alternatively obtain the optimal r analytically by finding the solutions to $\frac{dU(r)}{dr} = 0$. For analytical tractability, we consider the mean field approach to approximate $\mathcal{T}(r)$, where we consider the total power in the 'average state' instead of the expected total power over all possible states. Since the average number of players in the Engset's system is $n = \frac{\lambda}{\lambda+\mu} N$, we consider $\mathcal{T}(r) \approx \psi_r^{(n)} = \frac{r\beta}{2nc}\left(n - 1 + \sqrt{(n-1)^2 + \frac{4\ell}{r\beta}nc}\right)$. As the original expression is valid for $s \in \{1, \dots, N\}$, the interpolated mean field expression holds for $s \in [1, N]$. Hence, the center's utility function can be approximated as $U(r) \approx f\left(\frac{\psi_r^{(n)}}{\beta}\right) - g(r)$. Considering the specific form of $U(r)$ mentioned above, we have $U(r) \approx \alpha\left(\frac{\psi_r^{(n)}}{\psi_r^{(n)}+k}\right) - r$. Figure 13b presents the absolute error that would be incurred if this mean field expression is used instead of the original expression for computing the center's expected utility (we show absolute and not relative error because of $U(r)$ taking value 0 and around in a certain range of r). We can infer that the mean field approach provides the value of $U(r)$ very accurately. Furthermore, taking the derivative of the mean field expression of $U(r)$ and equating it to zero, gives the real-valued solution: $r = 123564.56$. It is worth highlighting that this value is almost equal to the actual value of optimal r mentioned earlier (i.e., 123564.38).

6 Conclusion and future work

This work studied strategic investments in a typical distributed computing system, while capturing the arrival and departure of players. On formulating the utility function and deriving its closed-form expression, we determined the players' investments for the different states in Markov perfect equilibrium (MPE). In MPE, in a given state, only players with cost parameters in a relatively low range that collectively satisfy a certain constraint in that state, invest. We inferred that players need not have knowledge about the system state and other players' parameters, if the total power that is being received by the center is communicated to the players as part of the system's protocol.

Using simulations, we studied the effects of the number of players present in the system, the arrival and departure rates, and other system parameters. We first studied how the number of players present in the system affects their individual investments and the total power received by the center. We observed that individual investment is the highest when in a minimal game (i.e., with two players present) and decreases almost inversely proportionally to the number of players present. While the total power received by the center in a state increases with the number of players, it converges to an amount proportional to the offered reward. Hence, if the system consists of a reasonably large number of players, the total power received by the center does not vary significantly despite the players' arrivals and departures, thus resulting in a system that is robust and reliable.

We then studied the effects of the arrival and departure rates on a player's expected utility depending on whether or not the player is present at the time of computing it. We observed that typically a lower arrival rate and a higher departure rate, up to a certain extent, are beneficial for a player as the competition would be kept low with less players among whom reward would be shared. However, beyond that extent, these rates are detrimental as the player itself would likely stay absent for a significant amount of time and lose out on the reward. The dynamics of the system are such that the relation between the average durations of an Engset cycle and a segment is critical, when a player computes its utility knowing whether or not it is present in the system. However, this relation is immaterial if the player is agnostic about its presence, since the computed utility would depend only on the ratio between the arrival and departure rates. A general observation was that a player's expected utility is maximized when the average number of players present is between 1 and 2, and typically close to 2, since this leads to the system being in the condition of least competition (i.e., consisting of only one player) with high probability.

We also studied how a player's utility is influenced by the system parameters (namely, r, β and ℓ). Each of these parameters follows a thresholding criterion, which determines whether a player invests and obtains a positive utility. Broadly, we observed that the different types of utilities increase almost linearly with r, converge to constant values with respect to β, and decrease as a convex function in ℓ. We presented insights for these observations using a mean field approach, which provided simplified and analytically explainable expressions as well as highly accurate approximations.

We concluded by studying a Stackelberg game where the center determines the reward to be offered, which influences the total power invested by the players, which in turn influences the center's own utility. We showed that the expected total power received by the center is a monotone increasing concave function of the reward parameter. We hence formulated the center's utility function and showed it to be concave under practically reasonable assumptions. Using the mean field approach, we were able to analytically find the optimal reward parameter with very high accuracy.

We believe that our model enables us to lay a game theoretic foundation for analyzing strategic investments in distributed computing and take a first step towards solving a challenging problem, which leaves ample scope for it to be developed further. In order to develop a more sophisticated stochastic model, one could obtain real-world data concerning the arrivals and departures of players. From the perspective of mechanism design, it would be interesting to design incentives so as to elicit the true cost parameters of the players. Alternatively, one could devise a method for deducing these latent variables, namely, cost parameters, from the observed players' actions and game situations. It would be interesting to analyze the game under bounded rationality. Another promising possibility is to incorporate state-learning in our model. Among other future directions, one is to study the game by accounting for possibility of players forming coalitions.

Acknowledgements This work was partly supported by CEFIPRA grant No. IFC/DST-Inria-2016-01/448 "Machine Learning for Network Analytics". The contribution of Eitan Altman was supported by LION "Learning In Operations and Networks", a joint Indo-French research team between INRIA and IIT Bombay. We are thankful to the anonymous reviewers for their comments which led to improvements across all sections of the paper.

Appendix A: Other applications of the proposed model

We now briefly discuss the applicability of our proposed model and the utility function given by Eq. (1), to distributed computing applications apart from volunteer computing, such as mining in blockchain. Mining relies on a proof-of-work procedure (Nakamoto, 2008) wherein players (termed *miners*) collect data that is to be encapsulated in a block and repeatedly compute hashes on potential solutions from a very large search space. A player is said to have mined a block and is rewarded a monetary amount, say r, if the player is the first to find one of the solutions that generates a hash value satisfying certain constraints. The algorithms employed for finding such a solution are typically based on randomized search over an exponentially large search space. The time required to find a solution in such a large search space is independent of the search space explored thus far, resulting in the search being practically memoryless. Now, if a continuous random variable has the memoryless property over the set of reals, it is necessarily exponentially distributed. Hence, the time required to find a solution and hence mine a given block is exponentially distributed, whose rate parameter can be considered to be β (i.e., the expected time is $\frac{1}{\beta}$). In Bitcoin mining, the expected time to mine a block is set at 10 minutes.

Now, consider that a player i invests computational power of $x_i^{(S)}$ to mine when the system is in state S, and let ℓ be the amount of power that is invested by large mining firms over a large period of time (i.e., irrespective of the system state). It is known that if the number of solutions is ξ, the distance of the probability of a player finding a solution before others, from being proportional to the player's invested power, is $\tilde{O}(1/\xi)$ (Zeng and Zuo, 2019). As ξ is typically large in mining, the probability of a player being the first to mine a block at any given time is proportional to its invested power at that time. Hence, the probability of player i being the first to mine the block in state S and winning the reward of r, is $\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell}$.

The possible events that can occur in state S are similar to what we discussed for volunteer computing, namely, the current block getting mined (in place of current segment ending) with rate β, a player $j \notin S$ arriving with rate λ_j, and a player $j \in S$ departing with rate μ_j. In the event of the block getting mined, player i receives a reward of $\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r$ in expectation, and the system stays in the same state S for the next block for which i's expected utility is perceived as $\delta R_i^{(S,\mathbf{x})}$. Since the sojourn time in state S for the current block is $\frac{1}{B^{(S)}} = (\beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j)^{-1}$, the corresponding expected cost incurred is $\frac{c_i x_i^{(S)}}{B^{(S)}}$. Hence, player i's expected utility as computed in state S is:

$$R_i^{(S,\mathbf{x})} := \frac{\beta}{B^{(S)}} \cdot \left(\frac{x_i^{(S)}}{\sum_{j \in S} x_j^{(S)} + \ell} r + \delta R_i^{(S,\mathbf{x})} \right) - \frac{c_i x_i^{(S)}}{B^{(S)}}$$
$$+ \sum_{j \notin S} \frac{\lambda_j}{B^{(S)}} \cdot R_i^{(S \cup \{j\},\mathbf{x})} + \sum_{j \in S} \frac{\mu_j}{B^{(S)}} \cdot R_i^{(S \setminus \{j\},\mathbf{x})}$$

It is worth highlighting that the mathematical form of $R_i^{(S,\mathbf{x})}$ is the same as Eq. (1) and so, our analysis and results will hold also for such other applications.

Apart from the above application of block mining at individual level, our model can be applied to mining pools as well, especially those which offer pay-per-share payouts. In general, since most applications involving distributed computing share the same underlying concepts, our model is applicable to a wide variety of applications.

Appendix B: Convergence of expected utility

Let $\mathbf{M}$ be the state transition matrix, among the states corresponding to the set of strategic players present in the system. In what follows, instead of writing $M(\mathcal{O}(S), \mathcal{O}(S'))$, we simply write $M(S, S')$ since it does not introduce any ambiguity. So, the elements of $\mathbf{M}$ are:

$$M(S, S) = \frac{\delta\beta}{B^{(S)}} ,$$
$$\text{for } j \notin S : M(S, S \cup \{j\}) = \frac{\lambda_j}{B^{(S)}} ,$$
$$\text{for } j \in S : M(S, S \setminus \{j\}) = \frac{\mu_j}{B^{(S)}} ,$$
$$\text{and all other elements of } \mathbf{M} \text{ are } 0.$$

Here, $B^{(S)} = \beta + \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$. Since $\beta > 0$, we have $B^{(S)} > \sum_{j \notin S} \lambda_j + \sum_{j \in S} \mu_j$. Hence, $\mathbf{M}$ is strictly substochastic (sum of the elements in each of its rows is less than 1).

Let $\mathbf{F}_i^{(\mathbf{x})}$ be the vector whose component $\mathcal{O}(S)$ is $F_i^{(S,\mathbf{x})}$, where

$$F_i^{(S,\mathbf{x})} = \left(\frac{r\beta}{\sum_{j \in S} x_j^{(S)} + \ell} - c_i \right) \frac{x_i^{(S)}}{B^{(S)}}$$

We now provide a proof of Lemma 1, which states that the recursive equation for $\mathbf{R}_i^{(\mathbf{x})}$, Eq. (1), converges for any policy profile $\mathbf{x}$.

Proof Let $\mathbf{R}_{i\langle t \rangle}^{(\mathbf{x})} = (R_{i\langle t \rangle}^{(1,\mathbf{x})}, \ldots, R_{i\langle t \rangle}^{(2^{|\mathcal{U}|},\mathbf{x})})^T$, where t is the iteration number and $(\cdot)^T$ stands for matrix transpose. The iteration for the value of $\mathbf{R}_{i\langle t \rangle}^{(\mathbf{x})}$ starts at $t = 0$; we examine if it converges when $t \to \infty$. Now, the expression for the expected utility in all states can be written in matrix form and then solving the recursion, as

$$\mathbf{R}_{i\langle t \rangle}^{(\mathbf{x})} = \mathbf{M}\mathbf{R}_{i\langle t-1 \rangle}^{(\mathbf{x})} + \mathbf{F}_i^{(\mathbf{x})} = (\mathbf{M})^t \, \mathbf{R}_{i\langle 0 \rangle}^{(\mathbf{x})} + \left(\sum_{\eta=0}^{t-1} (\mathbf{M})^\eta \right) \mathbf{F}_i^{(\mathbf{x})}.$$

Now, since $\mathbf{M}$ is strictly substochastic, its spectral radius is less than 1. So when $t \to \infty$, we have $\lim_{t\to\infty}(\mathbf{M})^t = \mathbf{0}$. Since $\mathbf{R}^{(\mathbf{x})}_{i\langle 0\rangle}$ is a finite constant, we have $\lim_{t\to\infty}(\mathbf{M})^t\mathbf{R}^{(\mathbf{x})}_{i\langle 0\rangle} = \mathbf{0}$. Further, $\lim_{t\to\infty}\sum_{\eta=0}^{t-1}(\mathbf{M})^\eta = (\mathbf{I}-\mathbf{M})^{-1}$ (Hubbard and Hubbard, 2015) . This implicitly means that $(\mathbf{I}-\mathbf{M})$ is invertible. Hence,

$$\begin{aligned}\lim_{t\to\infty}\mathbf{R}^{(\mathbf{x})}_{i\langle t\rangle} &= \lim_{t\to\infty}(\mathbf{M})^t\,\mathbf{R}^{(\mathbf{x})}_{i\langle 0\rangle} + \left(\sum_{\eta=0}^{\infty}(\mathbf{M})^\eta\right)\mathbf{F}^{(\mathbf{x})}_i \\ &= \mathbf{0} + (\mathbf{I}-\mathbf{M})^{-1}\mathbf{F}^{(\mathbf{x})}_i.\end{aligned}$$

□

Note also that Proposition 1 can be proved alternatively along the same line as the above proof of Lemma 1, by having $\mathbf{W}$ in place of $\mathbf{M}$ and $\mathbf{Z}^{(\mathbf{x})}_i$ in place of $\mathbf{F}^{(\mathbf{x})}_i$.

Appendix C: Effect of offered reward on total power received by the center in a state

Here, we discuss the general case where players could have different cost parameters. Note from Proposition 2 that since the set of investing players in any given state could change with r, it is not even clear whether the total power received by the center in a state would increase monotonically with r. In particular, we need to inspect whether the total power received by the center could decrease when the set of investing players expands owing to the increased reward. We show the following result.

Proposition 3 *If players invest as per Proposition 2, the total power received by the center in any given state is a monotone increasing continuous function of the reward parameter.*

Proof Recall that in a state S, $\psi^{(S)} = r\beta\frac{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}}c_j}}{2\sum_{j\in\hat{S}}c_j}$, where $\hat{S}\subseteq S$ is the set of investing players. It is clear that for a given set of investors $\hat{S}$, $\psi^{(S)}$ increases monotonically with r. As r varies, set $\hat{S}$ may change, thus changing the values of $|\hat{S}|$ as well as $\sum_{j\in\hat{S}}c_j$. In order to show a monotonic increase of $\psi^{(S)}$ with r despite any changes in set $\hat{S}$, we need to show that at any value of r where players get added to $\hat{S}$, the value of $\psi^{(S)}$ does not decrease (i.e., either increases or stays the same). Without loss of generality, consider that only one player gets added at any such value of r. In what follows, we show continuity at values of r where the set of investing players changes.

Consider a value of r such that the set of investing players is $\hat{S}\setminus\{i\}$ when the reward parameter is infinitesimally lower than r, while it is $\hat{S}$ (i.e., player i gets added to the set of investing players) when the reward parameter is infinitesimally higher than r. At this value of r, let $\underline{\psi}^{(S)}$ be the limit of $\psi^{(S)}$ from the left and $\overline{\psi}^{(S)}$ be its limit from the right. We will now show that $\underline{\psi}^{(S)} = \overline{\psi}^{(S)}$.

Since player i barely satisfies the cost constraint at this value of r, we have (the following equality is in limit): $c_i = \frac{2\sum_{j\in\hat{S}}c_j}{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}}c_j}}$. So, the limit of $\psi^{(S)}$ from the right is

$$\overline{\psi}^{(S)} = r\beta\frac{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}}c_j}}{2\sum_{j\in\hat{S}}c_j} = \frac{r\beta}{c_i}. \tag{C1}$$

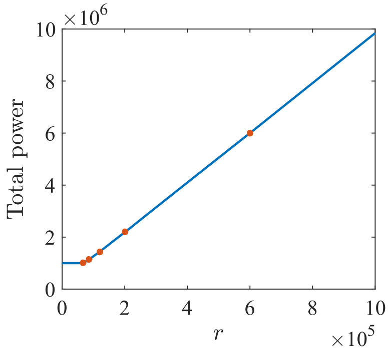


Fig. 14 Effect of the reward parameter r on the total power received by the center in a state

Now, $c_i = \frac{2\sum_{j\in\hat{S}} c_j}{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j}}$ is equivalent to

$$r = \frac{\ell}{\beta} \cdot \frac{c_i^2}{\sum_{j\in\hat{S}\setminus\{i\}} c_j - c_i(|\hat{S}|-2)}. \tag{C2}$$

This gives us an expression for r at which the set of investing players expands from $\hat{S}\setminus\{i\}$ to $\hat{S}$.

Now, the limit of $\psi^{(S)}$ from the left is $\underline{\psi}^{(S)} = r\beta\frac{|\hat{S}|-2+\sqrt{(|\hat{S}|-2)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}\setminus\{i\}} c_j}}{2\sum_{j\in\hat{S}\setminus\{i\}} c_j}$.

Let $\underline{\psi}^{(S)} = r\beta y$, where $y = \frac{|\hat{S}|-2+\sqrt{(|\hat{S}|-2)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}\setminus\{i\}} c_j}}{2\sum_{j\in\hat{S}\setminus\{i\}} c_j}$. This, in conjunction with Eq. (C2), gives

$$y^2 \sum_{j\in\hat{S}\setminus\{i\}} c_j - y(|\hat{S}|-2) = \left(\frac{1}{c_i}\right)^2 \sum_{j\in\hat{S}\setminus\{i\}} c_j - \frac{1}{c_i}(|\hat{S}|-2).$$

It can be easily seen that the above equation is satisfied when the value of y is $\frac{1}{c_i}$, and since y has a unique value from its definition, we must have $y = \frac{1}{c_i}$. Hence, from the above and Eq. (C1), we have $\underline{\psi}^{(S)} = r\beta y = \frac{r\beta}{c_i} = \overline{\psi}^{(S)}$. This completes the proof. □

Figure 14 presents representative plots showing the effect of the reward parameter r on the total power received by the center in a given state S. We consider the following values for the purpose of visualization (the plots for any other values follow similar behavior): $\beta = 6$, $\ell = 10^6$, $|S| = 5$, and $\{c_i\}_{i\in S} = \{0.40, 0.45, 0.50, 0.55, 0.60\}$. We vary the value of r from 0 up to 10^6 with a resolution of 10^3. As r increases, the set of investing players expands (which is intuitive and also can be seen from the proof of Proposition 3). In the plots, the points at which a previously non-investing player turns into an investing player are marked by red dots. It can be seen that with an increase in r, the total power increases similar to a piecewise-linear ramp function.

Recall that in a state S, the investing players $i \in \hat{S}$ collectively satisfy: $c_i < \frac{2\sum_{j\in\hat{S}} c_j}{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j}}$. For low values of r, the threshold is too low for the players' cost parameters to satisfy; hence no strategic players invest and the total power equals ℓ (this is the base of the ramp function). For values of r which attract investments, the term $\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j$ is of a similar order as $|\hat{S}|$ or lower (this can be seen from the critical value of r derived in Eq. (C2), which consequently results in $\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j$ being upper bounded by $4|\hat{S}|$). From Proposition 2, the total power received by the center in state S is $\psi^{(S)} = r\beta\frac{|\hat{S}|-1+\sqrt{(|\hat{S}|-1)^2+\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j}}{2\sum_{j\in\hat{S}} c_j}$. It can be seen that within any range of r wherein $\hat{S}$ does not change, $\psi^{(S)}$ increases as a concave function of r. It can also be seen that in general, $\psi^{(S)}$ is not differentiable with respect to r at the breakpoints where set $\hat{S}$ changes. Thus, $\psi^{(S)}$ increases in r as a piecewise-concave function. Moreover, due to the suppressed nature of the term $\frac{4\ell}{r\beta}\sum_{j\in\hat{S}} c_j$ in $\psi^{(S)}$ for values of r which attract investments, the increase in $\psi^{(S)}$ with r is close to being linear within any range of r wherein $\hat{S}$ does not change. Hence, the increase in $\psi^{(S)}$ with respect to r would be typically close to a piecewise-linear ramp function.

References

Abraham, I., Dolev, D., Gonen, R., & Halpern, J. (2006). Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In *ACM symposium on principles of distributed computing* (pp. 53–62). ACM.

Al Ridhawi, I., Aloqaily, M., & Jararweh, Y. (2021). An incentive-based mechanism for volunteer computing using blockchain. *ACM Transactions on Internet Technology (TOIT), 21*(4), 1–22.

Altman, E. (1996). Non zero-sum stochastic games in admission, service and routing control in queueing systems. *Queueing Systems, 23*(1–4), 259–279.

Altman, E., & Shimkin, N. (1998). Individual equilibrium and learning in processor sharing systems. *Operations Research, 46*(6), 776–784.

Altman, E., Reiffers, A., Menasché, D. S., Touati, C., & El-Azouzi, R. (2020). Blockchain competition between miners: A game theoretic perspective. *Frontiers in Blockchain, 2*, 26.

Alves, B. (2022). Average retail price of electricity to the residential sector in the United States in selected years from 1975 to 2021. Statista. https://www.statista.com/statistics/200199/residential-sector-electricity-prices-in-the-us-since-1975/.

Anderson, D. P., & Fedak, G. (2006). The computational and storage potential of volunteer computing. In *Sixth IEEE international symposium on cluster computing and the grid (CCGRID'06)* (pp. 73–80). IEEE.

Bellomo, N. (2008). *Modeling complex living systems: A kinetic theory and stochastic game approach*. Berlin: Springer.

Biais, B., Bisiere, C., Bouvard, M., & Casamatta, C. (2019). The blockchain folk theorem. *The Review of Financial Studies, 32*(5), 1662–1715.

Bitpanda. (2021). What is the purpose of mining pools and how do they work? Bitpanda. https://www.bitpanda.com/academy/en/lessons/what-is-the-purpose-of-mining-pools-and-how-do-they-work/.

Bowling, M., & Veloso, M. (2000). An analysis of stochastic game theory for multiagent reinforcement learning. Technical report, Carnegie-Mellon University Pittsburgh Pennsylvania School of Computer Science Technical Report No. CMU-CS-00-165

Cohen, J. (1957). The generalized Engset formulae. *Philips Telecommunication Review, 18*(4), 158–170.

Conway, L. (2022). What is Bitcoin halving? Definition, how it works, why it matters. Investopedia. https://www.investopedia.com/bitcoin-halving-4843769.

Dimitri, N. (2017). Bitcoin mining as a contest. *Ledger, 2*, 31–37.

Eyal, I. (2015). The miner's dilemma. In *2015 IEEE symposium on security and privacy* (pp. 89–103). IEEE.

Eyal, I., & Sirer, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. In *International conference on financial cryptography and data security* (pp. 436–454). Springer.

Fu, F., & Kozat, U. C. (2013). Stochastic game for wireless network virtualization. *IEEE/ACM Transactions on Networking, 21*(1), 84–97.

Goeree, J. K., & Holt, C. A. (1999). Stochastic game theory: For playing games, not just for doing theory. *Proceedings of the National Academy of Sciences, 96*(19), 10,564-10,567.

Grunspan, C., & Pérez-Marco, R. (2020). The mathematics of Bitcoin. *EMS Newsletter, 115*, 31–37.

Hassin, R., & Haviv, M. (2002). Nash equilibrium and subgame perfection in observable queues. *Annals of Operations Research, 113*(1–4), 15–26.

Hu, J., & Wellman, M. (2003). Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research, 4*(Nov), 1039–1069.
Hubbard, J. H., & Hubbard, B. B. (2015). *Vector calculus, linear algebra, and differential forms: A unified approach*. Matrix Editions.
Jiang, C., Chen, Y., Yang, Y. H., Wang, C. Y., & Liu, K. R. (2014). Dynamic Chinese restaurant game: Theory and application to cognitive radio networks. *IEEE Transactions on Wireless Communications, 13*(4), 1960–1973.
Kwok, Y. K., Song, S., & Hwang, K. (2005). Selfish grid computing: Game-theoretic modeling and NAS performance results. In *IEEE international symposium on cluster computing and the grid* (pp 1143–1150). IEEE.
Kwon, Y., Kim, D., Son, Y., Vasserman, E., & Kim, Y. (2017). Be selfish and avoid dilemmas: Fork after withholding (faw) attacks on bitcoin. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 195–209).
Lewenberg, Y., Bachrach, Y., Sompolinsky, Y., Zohar, A., & Rosenschein, J. S. (2015). Bitcoin mining pools: A cooperative game theoretic analysis. In *International conference on autonomous agents and multiagent systems (AAMAS)* (pp. 919–927). IFAAMAS.
Liu, Z., Luong, N. C., Wang, W., Niyato, D., Wang, P., Liang, Y. C., & Kim, D. I. (2019). A survey on blockchain: A game theoretical perspective. *IEEE Access, 7*, 47615–47643.
Maskin, E., & Tirole, J. (2001). Markov perfect equilibrium: I. observable actions. *Journal of Economic Theory, 100*(2), 191–219.
Mengistu, T. M., & Che, D. (2019). Survey and taxonomy of volunteer computing. *ACM Computing Surveys (CSUR), 52*(3), 1–35.
Murata, Y., Inaba, T., Takizawa, H., & Kobayashi, H. (2008). Implementation and evaluation of a distributed and cooperative load-balancing mechanism for dependable volunteer computing. In *2008 IEEE international conference on dependable systems and networks with FTCS and DCC (DSN)* (pp. 316–325). IEEE.
Nahir, A., Orda, A., & Raz, D. (2012). Workload factoring with the cloud: A game-theoretic perspective. In *IEEE international conference on computer communications (INFOCOM)* (pp. 2566–2570). IEEE.
Nakamoto, S. (2008). A peer-to-peer electronic cash system. Bitcoin.org. https://bitcoin.org/bitcoin.pdf.
Sapirshtein, A., Sompolinsky, Y., & Zohar, A. (2016). Optimal selfish mining strategies in Bitcoin. In *International conference on financial cryptography and data security* (pp. 515–532). Springer.
Sarmenta, L. F. (2001). Sabotage-tolerance mechanisms for volunteer computing systems. In *Proceedings first IEEE/ACM international symposium on cluster computing and the grid* (pp. 337–346). IEEE.
Sarmenta, L. F. G. (2001). Volunteer computing. Ph.D. thesis, Massachusetts Institute of Technology.
Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences, 39*(10), 1095–1100.
Wang, C. Y., Chen, Y., & Liu, K. R. (2018). Game-theoretic cross social media analytic: How Yelp ratings affect deal selection on Groupon? *IEEE Transactions on Knowledge and Data Engineering, 30*(5), 908–921.
Wang, J., & Zhang, F. (2013). Strategic joining in M/M/1 retrial queues. *European Journal of Operational Research, 230*(1), 76–87.
Watanabe, K., Fukushi, M., & Horiguchi, S. (2009). Collusion-resistant sabotage-tolerance mechanisms for volunteer computing systems. In *2009 IEEE international conference on e-business engineering* (pp. 213–218). IEEE.
Wong, P., & McArdle, P. (2020). Average U.S. residential monthly electricity price, consumption, and bill (2009–2019). U.S. Energy Information Administration. https://www.eia.gov/todayinenergy/detail.php?id=46276.
Zeng, Y., & Zuo, S. (2019). The Matthew effect in computation contests: High difficulty may lead to 51% dominance? In *The world wide web conference* (pp. 2281–2289). ACM.
Zheng, Z., & Xie, S. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services, 14*(4), 352–375.