Counterfactual Regret Minimization in Potential Games

Sloan J. Cinkle scinkle1@jh.edu
Johns Hopkins University

Abstract

Game theory has emerged as a vital tool for analyzing and modeling strategic decision-making across various domains. A key component of this framework is the use of regret minimization techniques, such as counterfactual regret minimization (CFR), to develop optimal strategies for players in diverse game settings. This research investigates the limitations and potential improvements of CFR in multiplayer environments, with a specific focus on potential games and congestion games.

The primary objective of this study is to explore the convergence properties of CFR in multiplayer games, especially those characterized by multiple Nash equilibria. Our main hypothesis is that, in certain potential game settings, CFR based on the potential function will converge to a Nash equilibrium in multiplayer games, even when multiple equilibria are present. By evaluating the adaptability and effectiveness of CFR algorithms in more complex game situations, this research aims to contribute to the development of more robust strategies for solving strategic games in real-world applications.

Through a combination of theoretical and experimental approaches, this study provides insights into the performance of CFR in congestion games and its interaction with various equilibrium concepts. Our empirical results demonstrate the ability of CFR to identify and converge to Nash equilibria in single and multiple equilibrium scenarios. Furthermore, we discuss future research directions, which include the extension of CFR algorithms to more complex strategic settings and the integration of other regret minimization techniques for enhanced performance. Overall, this research sheds light on the applicability of CFR in multiplayer settings and offers valuable insights for the advancement of game-theoretic research.

1. Introduction

Game theory has become an essential tool for understanding and modeling strategic decision-making in a wide range of domains. One of the critical aspects of this framework is the development and application of regret minimization techniques, such as counterfactual regret minimization (CFR), to devise optimal strategies for players in various game settings. This research aims to investigate the limitations and potential improvements of CFR in multiplayer settings, with a particular focus on potential games and congestion games.

The primary objective of this study is to explore the convergence properties of CFR in multiplayer games, especially those characterized by multiple Nash equilibria. The main hypothesis of this research is that, in certain potential game settings, CFR based on the potential function will converge to a Nash equilibrium in multiplayer games, even when there are multiple equilibria. The hypothesis aims to evaluate the adaptability and effectiveness of CFR algorithms in more complex game situations and contribute to the development of more robust strategies for solving strategic games in real-world applications.

The paper is organized as follows: Section 1 provides this introduction and an overview of the paper's structure. Section 2 presents the literature review, which is divided into two subsections: regret minimization and congestion games. In Section 3, the theoretical approach is described, providing a comprehensive understanding of the connection between regret minimization and congestion games. Section 4 outlines the experimental methods employed in the study, including a detailed explanation of the specific algorithms and game settings. Section 5 presents the empirical results of the research, divided into two subsections: games with a single Nash equilibrium and games with multiple Nash equilibria. Finally, Section 6 concludes the paper, summarizing the main findings and their implications for the application of CFR in multiplayer games.

2. Literature Review

To provide a foundation for understanding the literature, we will first define some basic concepts in game theory. A game Γ contains a set of players $N = (1, \dots, n)$, where a player's strategy $\sigma_i \in \Sigma_i$ dictates how the player should act in a given situation. The set of all player strategies is denoted $\sigma = (\sigma_1, \dots, \sigma_n) \in \Sigma$, and the set of player strategies excluding that of player i is $\sigma_{-i} \in \Sigma_{-i}$. An infostate I is a player's information about the current state of the game, while a history h is a sequence of past actions that led to the current state of the game.

The expected value of a given strategy for player i is the sum of utilities for terminal states times the probability that the terminal state will be reached according to strategy σ. Formally, $u_i(\sigma) = \sum_{z \in Z} u_i(h)\pi^\sigma(h)$, where $\pi^\sigma(h)$ is called the reach probability of history h given that players act according to σ.

A Nash equilibrium is a stable set of strategies where no player has an incentive to unilaterally change their strategy, assuming that the other players continue to use their current strategies. Formally, a set of player strategies $\sigma = (\sigma_1, \sigma_2, \dots)$ is an ϵ-Nash equilibrium if

$$u_i(\sigma) + \epsilon \geq \max_{\sigma_i' \in \Sigma_i} u_i(\sigma_i', \sigma_{-i}) \quad \forall i \in N.$$

2.1. Regret Minimization

Regret minimization is a concept in game theory that focuses on minimizing average overall regret for each player over T iterations. Let $\pi^\sigma(h)$ be the probability of history h occuring when players choose actions according to policy σ, and define $\pi^\sigma(I) = \sum_{h \in I} \pi^\sigma(h)$ as the probability of reaching a particular infostate I given σ. Then $\pi_i^\sigma(I)$ is the probability of reaching I when player i plays according to σ, and $\pi_{-i}^\sigma(I)$ is the probability of reaching I when all players except player i play according to σ. The average overall regret of player i at time T is defined by

$$R_i^T = \frac{1}{T} \max_{\sigma_i^* \in \Sigma_i} \sum_{t=1}^{T} \pi_{-i}^{\sigma^t}(I)(u_i(\sigma_i^*, \sigma_{-i}^t) - u_i(\sigma^t)), \tag{1}$$

and the average strategy of player i for an infostate I over T iterations is

$$\bar{\sigma}_i^T(I)(a) = \frac{\sum_{t=1}^{T} \pi_i^{\sigma^t}(I)\sigma^t(I)(a)}{\sum_{t=1}^{T} \pi_i^{\sigma^t(I)}}.$$

The measurement of overall average regret in Equation (1) applies for normal-form games, where players decide their strategies before any actions are taken. However, it does not carry over to sequential games, where players determine their optimal strategies at each infostate.

One approach to addressing regret in sequential games is Counterfactual Regret Minimization (CFR) (Zinkevich et al., 2007). In a sequential game, $u_i(\sigma, I)$ is the expected utility of player i at infostate I. Let $\sigma|I \to a$ be an identical strategy to σ except that player i always chooses action a at infostate I. The immediate counterfactual regret of player i at infostate I on iteration T is defined by

$$R^T_{i,imm}(I) = \frac{1}{T} \max_{a \in A(I)} \sum_{t=1}^{T} \pi^{\sigma^t}_{-i}(I)(u_i(\sigma^t|_{I \to a}, I) - u_i(\sigma^t, I)).$$

Let $R^{T,+}_{i,imm}(I) = \max(R^T_{i,imm}(I), 0)$. Zinkevich et al. (2007) provides a formal proof that $R^{T,+}_{i,imm}(I)$ is an upper bound on R^T_i. Moreover, the average overall regret of player i can be calculated for an action a in infostate I using the equation

$$R^T_i(I, a) = \frac{1}{T} \sum_{t=1}^{T} \pi^{\sigma^t}_{-i}(I)(u_i(\sigma^t|_{I \to a}, I) - u_i(\sigma^t, I)).$$

Let $R^{T,+}_i(I, a) = \max(R^T_i(I, a), 0)$. Player i can now minimize R^T_i by selecting actions according to the policy

$$\sigma^{T+1}_i(I)(a) = \begin{cases} \frac{R^{T,+}_i(I,a)}{\sum_{a \in A(I)} R^{T,+}_i(I,a)} & \text{if} \sum_{a \in A(I)} R^{T,+}_i(I, a) > 0 \\ \frac{1}{|A(I)|} & \text{otherwise} \end{cases}$$

Brown (2020) gives a formal proof that, in two-player zero-sum games, if $\frac{R^T_i}{T} < \epsilon_i$ for all $i \in N$ at time T, then the average strategy $\bar{\sigma}^T$ is an ϵ-Nash equilibrium. However, Abou Risk et al. (2010) discusses the difficulties of applying CFR to multiplayer games. Specifically, in a three-player game, even when one player chooses a Nash equilibrium strategy, the other two players can still choose a sequence of actions to reduce the utility of the first player. Therefore, there is no theoretical guarantee that CFR will converge to an ϵ-Nash equilibrium in multiplayer games.

There has been a great deal of research to improving the convergence of CFR in multiplayer settings, most of which focus on reducing the number of strategies that an agent considers at a given infostate. Heads-up experts used by Abou Risk et al. (2010) apply expert domain knowledge to guide player strategies whenever one of the three players "folds" in a game of poker. This alleviates the need for the remaining players to find their equilibrium, and focuses on the strategy spaces of the first player who dropped out. Brown and Sandholm (2016) explore warm starting, a technique for improving the convergence to an equillibrium at the start of the training process by incorporating either external domain knowledge or coarse abstractions to improve the convergence to restrict the search space during early iterations. Strategy stitching is a technique reviewed by Brown and Sandholm (2014) that transfers regrets on actions in one setting to warm start the convergence to an equilibrium in a different setting that has a similar structure.

Hartley et al. (2017) and Tian et al. (2020) explore the use of traditional CFR in a multiplayer setting, but limit their research to a 2v2 game scenario where players on the same team have a shared utility function. While this approach does technically apply CFR to a multiplayer game, it essentially simplifies into a two-player game when considering the Nash equilibrium solution. In light of these limitations and the advancements made in applying CFR to multiplayer settings, it is worth considering alternative classes of games that are guaranteed to converge to Nash equilibria, even in multiplayer settings.

2.2. Congestion Games

Potential games are a class of games introduced by Rosenthal (1973) that are defined by the existence of a potential function that encapsulates the total utilities of all players. The potential function $P : \Sigma \to \mathbb{R}$ must satisfy the equality

$$P(\sigma_i', \sigma_{-i}^t) - P(\sigma^t) = u_i(\sigma_i', \sigma_{-i}^t) - u_i(\sigma^t) \quad \forall \sigma_i, \sigma_i' \in \Sigma_i \tag{2}$$

for every $i \in N$ and for every $\sigma_{-i} \in \Sigma_{-i}$.

Monderer and Shapely (1996b) shows that every finite potential game possesses a pure-strategy Nash equilibrium at

$$P(\sigma) \geq \max_{\sigma_i' \in \Sigma_i} P(\sigma_i', \sigma_{-i}).$$

Additionally, every finite potential game has the finite improvement property, which means that in a potential game Γ, there exists a path of strategies $\gamma = (\sigma^0, \sigma^1, \dots)$ for player i where $\sigma^k = (\sigma_i', \sigma_{-i}^{k-1})$ for any $\sigma_i' \in \Sigma_i$ where $\sigma_i' \neq \sigma_i^{k-1}$. If $u_i(\sigma^k) > u_i(\sigma^{k-1})$ for all $k \geq 1$, then γ is called an improvement path. If all improvement paths in Γ are finite, then Γ has the finite improvement property.

It follows from Monderer and Shapley (1996a) that potential games have the Fictitious Play property. This means that, for any iterative self-play algorithm where agents choose their best response to the other players' strategies, the average strategies chosen by all agents will converge to a Nash equilibrium as the number of iterations goes to infinity Berger (2007).

Monderer and Shapely (1996b) also define congestion games, a type of potential game defined by $N = \{1, \dots, n\}$ players and $M = \{1, \dots, m\}$ facilities. Each $\sigma_i \in \Sigma_i$ is a nonempty subset of facilities used by player i. The cost of each facility to each user is $c_a(\kappa_a(\sigma))$, where $\kappa_a(\sigma)$ is the number of users of facility a. The potential function of a congestion game is given as

$$P(\sigma) = \sum_{a \in \bigcup_{i=1}^{n} \sigma_i} \left(\sum_{k=1}^{\kappa_a(\sigma)} c_a(k) \right).$$

The literature shows that every congestion game is a potential game and that every finite potential game is isomporhic to a congestion game.

3. Theoretical Approach

The problem of solving Nash equilibria in multiplayer games would be alleviated by focusing the research solely on congestion games. Since all players have a shared utility function, any two players have no incentive to deviate from their Nash equilibrium strategies once the third player converges. Additionally, whereas Monderer and Shapely (1996b) and Holzman and Law-Yone (1997) define congestion games in the normal-form, there may be additional benefits of representing a congestion game as a sequential game for use in CFR. Players would take turns scheduling actions that are automatically carried out after all players have scheduled their actions for a given round. Specifically, if the players have information about how many other players occupy or have occupied their node at a given time, it could result in better decision-making and coordination betwen players.

From Equation (1) and Equation (2), it is trivial to see how the potential function can be incorporated into regret minimization. The average overall regret of player i at iteration T in a potential game is

$$R_i^T = \frac{1}{T} \max_{\sigma_i^* \in \Sigma_i} \sum_{t=1}^{T} \pi_{-i}^{\sigma^t}(I)(P(\sigma_i^*, \sigma_{-i}^t) - P(\sigma^t)),$$

and this can be measured at each infostate I with

$$R_{i,imm}^T(I) = \frac{1}{T} \max_{a \in A(I)} \sum_{t=1}^{T} \pi_{-i}^{\sigma^t}(I)(P(\sigma^t|_{I \to a}, I) - P(\sigma^t, I)).$$

Theorem 1 (Convergence) *Counterfactual regret minimization is converges to a Nash equilibrium in potential games under specific conditions.*

Proof For CFR to converge to a Nash equilibrium in potential games, the following conditions must hold:

1. There exists a potential function for the game such that the potential function is monotonically increasing as the players' strategies approach a Nash equilibrium.

2. At each iteration, players update their strategies using the CFR algorithm with the objective of minimizing their counterfactual regret with respect to the potential function.

Assuming these conditions hold, as players minimize their counterfactual regret, the potential function's value will increase monotonically. From the finite improvement property established by Monderer and Shapely (1996b), it must converge to a maximum value. At this maximum value, no player can deviate from their current strategy and improve their utility, given the strategies of the other players. This implies that the players' strategies have converged to a Nash equilibrium. ■

Note that there is no theoretical guarantee on the bound of the Nash equilibrium for CFR in potential games. The original CFR algorithm was designed for zero-sum games, where one player's maximum utility is equal to the other players' minimum utility. In potential games, this property does not hold, as all players share the potential function.

In this research, we found it beneficial to make certain adjustments to the CFR algorithm when applying it to potential games. Firstly, we subtracted the minimum observed potential as of iteration T from the actual potential value of a terminal state. This adjustment can help mitigate the differences between a zero-sum game and a potential game by creating a relative measure of the potential function.

Another modification we made was to initialize the average overall regret of each player at each information set to 1 rather than 0. This adjustment can be beneficial when the potential function is nonpositive, because it provides an initial incentive for exploration and helps players avoid getting stuck in suboptimal strategies.

4. Experimental Methods

This section outlines the experimental methods employed in the study to evaluate the performance of Counterfactual Regret Minimization (CFR) in multiplayer congestion games. Two congestion games with three players each were developed, with varying numbers of Nash equilibria. Two versions of CFR were applied to each game: a zero-sum adjusted version and a potential-based version. Additionally, the impact of extra information in players' infostates was explored by implementing sequential versions of the games.

4.1. Congestion Games

Two congestion games were designed, each consisting of four nodes (A, B, C, and D), where A is the starting node and D is the terminal node. The routes in each game are AB, AC, BC, BD, and CD. Both games are modified versions of the two-player game presented by Roughgarden (2010), however, this research focuses specifically on a multiplayer setting.

- **3-Player Game (Single Nash Equilibrium):** In this game, the unique Nash equilibrium strategy is for all players to travel on the route ABCD. The congestion map for this game can be found in Figure 1, and the associated player utilities and potential payoffs are shown in Table 1.
- **3-Player Game (Multiple Nash Equilibria):** In this game, the Nash equilibrium strategy is for two players to take route AB and one player to take route CD. The congestion map for this game can be found in Figure 2, and the associated player utilities and potential payoffs are shown in Table 2.

4.2. CFR Algorithms

Two versions of CFR were developed for the experiments:

- **Zero-Sum Adjusted CFR:** In this version, players' utilities at each terminal state were adjusted to be zero-sum by subtracting the average utility for all players from each player's utility.
- **Potential-Based CFR:** In this version, modifications to the CFR algorithm were made as described in Section 3, focusing on the potential function and the initialization of immediate counterfactual regret.

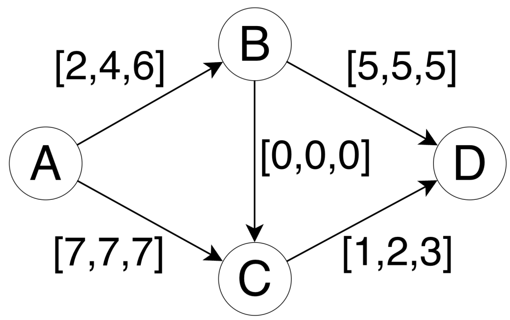


Figure 1: 3-player congestion game with a single Nash equilibrium.

(*a*) Player utilities.

	ABD	ABCD	ACD
	(11, 11, 11)	(11, 7, 11)	(9, 8, 9)
ABD	(11, 11, 7)	(11, 8, 8)	(9, 9, 6)
	(9, 9, 8)	(9, 6, 9)	(7, 9, 9)
	(7, 11, 11)	(8, 8, 11)	(6, 9, 7)
ABCD	(8, 11, 8)	(9, 9, 9)	(7, 10, 7)
	(6, 7, 9)	(7, 7, 10)	(5, 10, 10)
	(8, 9, 9)	(9, 6, 9)	(9, 9, 7)
ACD	(9, 9, 6)	(10, 7, 7)	(10, 10, 5)
	(9, 7, 9)	(10, 5, 10)	(10, 10, 10)

(*b*) Potential function.

	ABD	ABCD	ACD
	27	23	24
ABD	23	20	21
	24	21	24
	23	20	21
ABCD	20	18	19
	21	19	22
	24	21	24
ACD	21	19	22
	24	22	27

Table 1: Payoffs of the 3-player congestion game with a single Nash equilibrium.

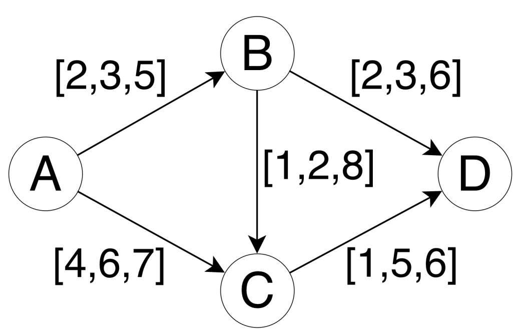


Figure 2: 3-player congestion game with many Nash equilibria.

(*a*) Player utilities.

	ABD	ABCD	ACD
ABD	(11, 11, 11) (8, 8, 7) (6, 6, 5)	(8, 7, 8) (7, 12, 12) (5, 9, 9)	(6, 5, 6) (5, 9, 9) (9, 12, 12)
ABCD	(7, 8, 8) (12, 7, 12) (9, 5, 9)	(12, 12, 7) (19, 19, 19) (11, 11, 10)	(9, 9, 5) (11, 10, 11) (4, 11, 11)
ACD	(5, 6, 6) (9, 5, 9) (12, 9, 12)	(9, 9, 5) (10, 11, 11) (11, 4, 11)	(12, 12, 9) (11, 11, 4) (13, 13, 13)

(*b*) Potential function.

	ABD	ABCD	ACD
ABD	21 17 15	17 21 18	15 18 20
ABCD	17 21 18	21 33 24	18 24 25
ACD	15 18 20	18 24 25	20 23 29

Table 2: Payoffs of the 3-player congestion game with many Nash equilibria.

To test the convergence of both CFR algorithms, the zero-sum adjusted CFR was run until it converged for the game with a single Nash equilibrium. The potential-based CFR was then run for the same number of iterations. For the 3-player game with multiple equilibria, a finite number of iterations were run for both algorithms, as there is no guarantee of convergence in either case.

4.3. Sequential Games

In a normal-form game, the players' strategies are determined before their first move is made. Sequential versions of the congestion games were also developed, in which players have additional information about the number of other players sharing the same node at each decision point. Both CFR algorithms were applied to these sequential games, and the same process for testing convergence was followed as in the normal-form games.

4.4. Evaluation Metrics

For all four scenarios (each of the two games with and without information), the performance of the CFR algorithms was evaluated using the following metrics:

- Average overall regret for the zero-sum adjusted and potential-based cases.
- Jensen-Shannon (JS) divergence (Menéndez et al., 1997) between the average strategies of all players at the two main decision points (AB or AC and BC or BD).
- Average overall regret and average strategies of the players at the two main decision points.

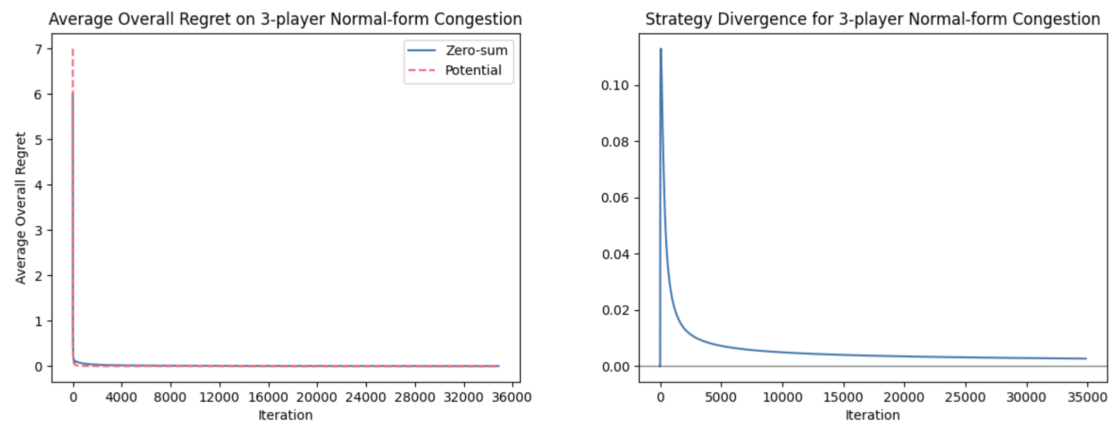
4.5. Monte Carlo CFR

The implementation of CFR is based on a Monte Carlo approach, where only one terminal state is reached on each iteration of the algorithm. Although this method requires more iterations to converge compared to traditional CFR, it has been shown to produce average strategies that are closer to Nash equilibrium than traditional CFR.

5. Empirical Results

5.1. Single Nash Equilibrium Games

For the normal-form 3-player congestion game with a single Nash equilibrium, zero-sum CFR converged at $\epsilon = 0.01$ after $T = 34,852$ iterations. Potential CFR was run for the same number of iterations. The average overall regret in the zero-sum case was 0.0001 for player 1, 0.0004 for player 2, and 0.0033 for player 3, while in the potential case, it was 0.0002. Average overall regret and JS divergence of strategies for the zero-sum and potential case are shown in Figure 3, and the average overall strategies are displayed in Table 3.



(a) Average overall regret. (b) JS divergence of strategies.

Figure 3: Average overall regret and strategy divergence for the 3-player normal-form congestion game with a single Nash equilibrium.

(a) Zero-sum strategies.

Player	**AB**	**BC**
1	0.9782	0.9727
2	0.9920	0.9892
3	0.9989	0.9988

(b) Potential strategies.

Player	**AB**	**BC**
1	1.0000	1.0000
2	1.0000	1.0000
3	1.0000	1.0000

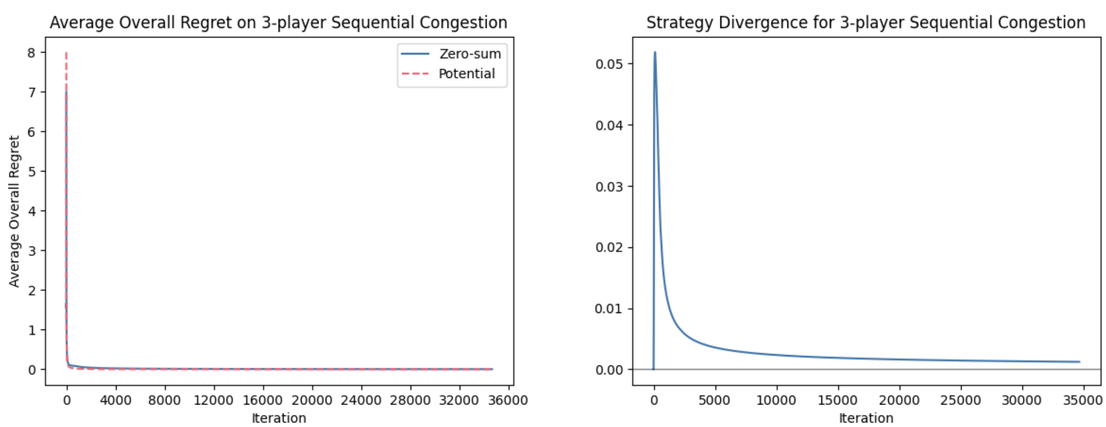
Table 3: Average strategies of each player in the 3-player normal-form congestion game with a single Nash equilibrium.

Although zero-sum CFR found the Nash-equilibrium strategies, potential CFR identified a stronger strategy profile. In the zero-sum case, the average overall strategy had each player performing optimal actions AB and BC with a probability between 97 and 100%. In contrast, the potential case showed these strategies with probabilities much closer to 100%.

Graphs depicting average overall regret indicate similarities between both algorithms, with potential CFR exhibiting lower regret than the sum of player regrets in some iterations. Initially, both algorithms had a 50% strategy, but the JS divergence increased before quickly converging to 0.

In the sequential 3-player congestion game with a single Nash equilibrium, zero-sum CFR reached convergence at $\epsilon = 0.01$ after $T = 34,665$ iterations, with potential CFR run for the same duration. The average overall regret was 0.0002 for player 1, 0.0005 for player 2, and 0.0033 for player 3 in the zero-sum case and 0.0004 in the potential case. The results of both normal-form and sequential games were similar, making it difficult to determine a superior performer. Average overall regret and JS divergence of strategies for the zero-sum and potential case are shown in Figure 4, and the average overall strategies are displayed in Table 4.

In the sequential 3-player congestion game with a single Nash equilibrium, zero-sum CFR ran for $T = 34665$ iterations until convergence at $\epsilon = 0.01$, and potential CFR for the same number of iterations. In the zero-sum case, average overall regret for each player was measured at $[0.0002, 0.0005, 0.0033]$, and in the potential case, average overall regret equalled 0.0004.



(a) Average overall regret. (b) JS divergence of strategies.

Figure 4: Average overall regret and strategy divergence for the 3-player sequential congestion game with a single Nash equilibrium.

5.2. Multiple Nash Equilibria Games

In the normal-form 3-player congestion game with multiple Nash equilibria, both zero-sum CFR and potential CFR were executed for $T = 16,000$ iterations. The average overall regret for each player in the zero-sum case was $[0.0000, 0.0000, 0.2605]$, while in the potential case, it was 0.0004. Average overall regret and JS divergence of strategies for the zero-sum and potential case are shown in Figure 5, and the average overall strategies are displayed in Table 5. Zero-sum CFR could not find a Nash equilibrium in this game, as player 1 and (*a*) Zero-sum strategies.

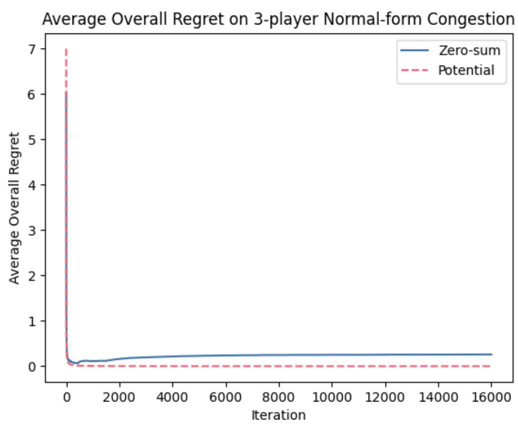
Player	$\mathbf{A_2B}$	$\mathbf{B_2C}$
1	0.9748	0.9762
2	0.9872	0.9926
3	0.9974	0.9991

(*b*) Potential strategies.

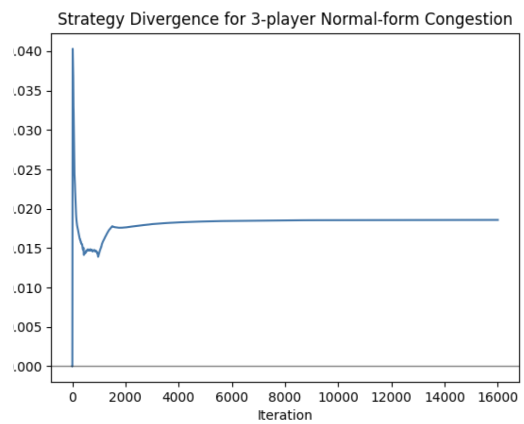
Player	$\mathbf{A_2B}$	$\mathbf{B_2C}$
1	1.0000	1.0000
2	1.0000	1.0000
3	1.0000	1.0000

Table 4: Average strategies of each player in the 3-player sequential congestion game with a single Nash equilibrium.

player 2 were indecisive between actions. Although player 3's strategy aligned with one of the game's Nash equilibria, player 1 and player 2 had 50% strategy probabilities at the starting node. In the potential case, players 1 and 2 opted for route ABD, while player 3 chose ACD, which is consistent with the game's Nash equilibria strategy. The average overall regret in the zero-sum case did not converge, while the potential CFR displayed a continual decrease over all iterations. The JS divergence at crucial decision points underlines the better performance of potential CFR. It is worth noting that player 3's strategy at node BC is irrelevant, as it always prefers node C over node B.



(*a*) Average overall regret.



(*b*) JS divergence of strategies.

Figure 5: Average overall regret and strategy divergence for the 3-player normal-form congestion game with many Nash equilibria.

For the sequential 3-player congestion game with multiple Nash equilibria, both zero-sum CFR and potential CFR were run for $T = 16,000$ iterations. The average overall regret for each player in the zero-sum case was $[0.0001, 0.0001, 0.2355]$, while in the potential case, it was 0.0008. Average overall regret and JS divergence of strategies for the zero-sum and potential case are shown in Figure 6, and the average overall strategies are displayed in (*a*) Zero-sum strategies.

Player	**AB**	**BC**
1	0.5018	0.4993
2	0.5002	0.4996
3	0.9997	0.0013

(*b*) Potential strategies.

Player	**AB**	**BC**
1	1.0000	0.0000
2	1.0000	0.0000
3	0.0000	0.2197

Table 5: Average strategies of each player in the 3-player normal-form congestion game with many Nash equilibria.

Table 6. Once again, the results of the normal-form and sequential games were similar, with no discernible differences in performance.

(*a*) Zero-sum strategies.

Player	$\mathbf{A_2B}$	$\mathbf{B_1C}$
1	0.5019	0.4990
2	0.5002	0.4996
3	0.9993	0.0011

(*b*) Potential strategies.

Player	$\mathbf{A_2B}$	$\mathbf{B_1C}$
1	1.0000	0.0000
2	1.0000	0.0000
3	0.0000	0.4288

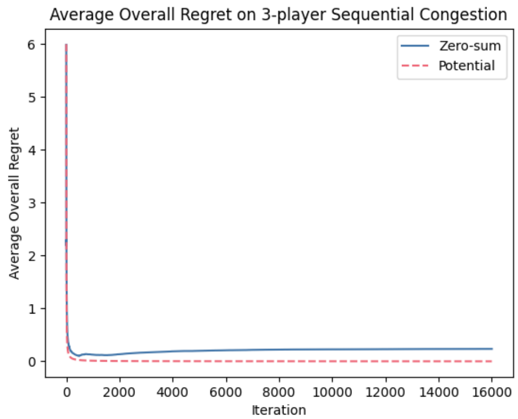
Table 6: Average strategies of each player in the 3-player sequential congestion game with many Nash equilibria.

6. Conclusion

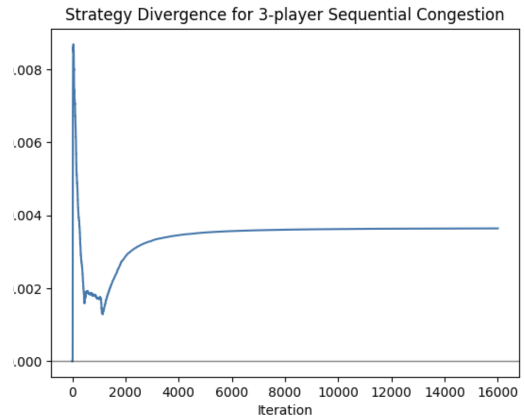
In this research, the performance of Counterfactual Regret Minimization (CFR) has been investigated in multiplayer congestion games. Two congestion games were designed, with the first having a single Nash equilibrium and the second having multiple Nash equilibria. Two versions of CFR were applied in each case: a zero-sum adjusted version and a potential-based version. Sequential versions of the games were also developed to examine the impact of extra information on player strategies.

The results demonstrated that both zero-sum and potential CFR algorithms could identify Nash equilibrium strategies in the game with a single Nash equilibrium. However, the potential-based CFR algorithm exhibited better performance in terms of identifying stronger strategy profiles and achieving lower average overall regret compared to the zero-sum adjusted CFR algorithm. The potential-based CFR algorithm also outperformed the zero-sum adjusted CFR algorithm in the game with multiple Nash equilibria, as it successfully identified one of the game's Nash equilibria strategies while the zero-sum adjusted CFR algorithm could not.

In both normal-form and sequential games, the potential-based CFR algorithm consistently displayed lower average overall regret compared to the zero-sum adjusted CFR



(a) Average overall regret.



(b) JS divergence of strategies.

Figure 6: Average overall regret and strategy divergence for the 3-player sequential congestion game with many Nash equilibria.

algorithm. The results also indicated that the inclusion of extra information in players' infostates did not lead to discernible differences in performance between the normal-form and sequential games.

This research has demonstrated the potential of the potential-based CFR algorithm in solving multiplayer congestion games. While the zero-sum adjusted CFR algorithm also shows promise, the potential-based algorithm offers a more robust approach, as evidenced by its ability to identify stronger strategy profiles and achieve lower average overall regret in both single and multiple Nash equilibrium scenarios. As multiplayer games continue to pose challenges for game theory, the potential-based CFR algorithm offers a promising avenue for future research and practical applications.

By expanding our understanding of how CFR can be effectively applied in multiplayer settings and potential games, we open new avenues for tackling complex strategic problems in diverse fields, such as economics, political science, and artificial intelligence. The successful convergence of CFR algorithms in games with multiple equilibria highlights their potential to handle intricate scenarios that may have previously seemed unmanageable.

6.1. Future Research

It was observed that there does not appear to be a significant difference when incorporating additional information in the players' infostate. Future research could explore the incorporation of other types of information into the players' infostate that could provide additional benefits. By identifying and incorporating relevant information, it may be possible to further enhance the performance of regret minimization algorithms and develop more effective strategies for solving strategic games.

One of the advantages of CFR is that it can be applied to games with imperfect information. Beier et al. (2004) studies the feasibility of computing Nash equilibria in imperfect-information congestion games. Following the success of CFR in imperfect information games, it may be more suitable to model a congestion game with randomized start states for each player. Using the potential function in CFR may allow it to account for imperfect information within a potential game.

The method of game traversal in this research is based on MCCFR (Burch et al., 2014), but there exist other variations of CFR that could prove to converge faster for large-scale potential games that may be seen in real-world situations. A possible direction for future research is to explore other improvments on the CFR algorithm, such as CFR-D (Burch et al., 2014) or CFR-AVG (Tammelin, 2014) in the potential setting.

References

N. Abou Risk, D. Szafron, et al. Using counterfactual regret minimization to create competitive multiplayer poker agents. In *AAMAS*, pages 159–166, 2010.

R. Beier, A. Czumaj, P. Krysta, and B. Vöcking. Computing equilibria for congestion games with (im)perfect information. In *SODA*, pages 746–755. Citeseer, 2004.

U. Berger. Brown's original fictitious play. *Journal of Economic Theory*, 135(1):572–578, 2007.

N. Brown. Equilibrium finding for large adversarial imperfect-information games. *PhD thesis*, 2020.

N. Brown and T. Sandholm. Regret transfer and parameter optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

N. Brown and T. Sandholm. Strategy-based warm starting for regret minimization in games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.

N. Burch, M. Johanson, and M. Bowling. Solving imperfect information games using decomposition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.

M. Hartley, S. Zheng, and Y. Yue. Multi-agent counterfactual regret minimization for partial-information collaborative games. In *31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.

R. Holzman and N. Law-Yone. Strong equilibrium in congestion games. *Games and Economic Behavior*, 21(1-2):85–101, 1997.

M. L. Menéndez, J. A. Pardo, L. Pardo, and M. C. Pardo. The jensen-shannon divergence. *Journal of the Franklin Institute*, 334(2):307–318, 1997.

D. Monderer and L. Shapely. Potential games. *Games and Economic Behavior*, 14(1): 124–143, 1996b.

D. Monderer and L. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996a.

R. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

T. Roughgarden. Computing equilibria: a computational complexity perspective. *Economic Theory*, pages 193–236, 2010.

O. Tammelin. Solving large imperfect information games using cfr+. *arXiv preprint arXiv:1407.5042*, 2014.

Y. Tian, Q. Gong, and Y. Jiang. Joint policy search for multi-agent collaboration with imperfect information. *Advances in Neural Information Processing Systems*, 33:19931–19942, 2020.

M. Zinkevich, M. Johanson, M. Bowling, and C. Piccione. Regret minimization in games with incomplete information. *Advances in Neural Information Processing Systems*, 20, 2007.