САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ – ПРОЦЕССОВ УПРАВЛЕНИЯ

Кафедра математической теории игр и статистических решений

Гриних Александра Леонидовна

Выпускная квалификационная работа

«О временной состоятельности нормативных принципов оптимальности в динамических играх»

Научный руководитель,
доктор физ.-мат. наук,
профессор,
Петросян Л. А.

Рецензент,
Мельник А. В.

Санкт-Петербург
2019

Содержание

Введение.

В современном мире многие процессы взаимодействия людей можно описать теоретико-игровой моделью. Одной из основополагающих моделей теории игр является «дилемма заключённого». Она позволяет анализировать взаимодействие двух рациональных агентов в условиях, когда для достижения общей выгоды необходимо поступиться личными интересами (отказаться от выбора строго доминирующей стратегии для достижения Парето-оптимума). Для реализации многостороннего взаимодействия была реализована модель «дилеммы заключённого n лиц», которая впервые была рассмотрена Гамбургером (Hamburger H.) [7]. В ней были сохранены основные принципы взаимодействия, аналогичные классической модели.

Решение подобного рода задач заключается в нахождении равновесных стратегий поведения, а также иных принципов оптимальности в построенной модели. Кроме того, большое количество игроков делает эту задачу более интересной с точки зрения кооперативной теории игр, поскольку даже характиристическая функция выглядит менее тривиально, чем в двухагентной модели.

Эксперименты с частично кооперативным поведением в повторяющейся «дилемме заключённого n лиц» были описаны и проанализированы Страффином [14]. Поскольку взаимодействие лиц осуществляется многоэтапно, а каждый поступок накладывает отпечаток на дальнейшие взаимоотношения, следует рассматривать повторяющийся вариант модели. Ауманн [1] анализирует равновесное поведение в условиях неопределённого количества повторений данной игры.

В данной работе исследуется новый равновесный принцип поведения в условиях данной модели. Строится новая характеристическая функция Петросяна [13] для рассмотрения нормативных принципов оптимальности в динамической модели «дилеммы заключённого n лиц». В частности, находится подъядро Петросяна-Панкратовой [12] динамической игры, которое в т. ч. содержит вектор Шепли.

В первой главе приводится наиболее полное описание модели « дилемма заключённого n лиц», обобщающее уже существующие наработки в этой области, а также построена функция выигрышей игроков, выделены основные положения данной игры. Для наиболее полного понимания строятся таблицы соответствия общей функции выигрыша, выведенной в данной работе, с разными видами таблиц выигрышей, рассматриваемых в более ранней литературе.

В разделе 2 находится новое равновесие по Нэшу в конечной многошаговой игре и доказывается теорема об эффективном наказании при кооперативном поведении игроков в конечной повторяющейся и динамической играх, основанных на модели «дилемма заключённого n лиц». Приводится пример расчёта максимально необходимого количества шагов в повторяющейся и динамической игре для обеспечения эффективного наказания при использовании данной модели для трёх игроков.

В третьем разделе находится ядро динамической модели, а также, основываясь на построении новой характеристической функции Петросяна [13] для многошаговой динамической игры найдено подъядро динамической «дилеммы заключённого n лиц» и доказано, что оно обладает свойством сильной динамической устойчивости.

Последний раздел относится к поискам вектора Шепли в стохастических и динамических играх, основанных на модели «дилемма заключённого n лиц».

Обзор литературы.

Большинство исследований, направленных на изучение модели «дилеммы заключённых n лиц», можно разбить на четыре блока: теоретическое исследование одношаговой модели, анализ многошаговых и повторяющихся моделей игры, поиск эволюционно-устойчивых стратегий и эмпирические результаты.

В работе Страффина [14] вводится общее представление об одношаговой модели данной игры, выводятся некоторые принципы построения выигрышей игроков, такие как:

- D является доминирующей стратегией для всех игроков;
- если все игроки выберут стратегию D, это будет хуже для всех, чем если бы всеми игроками была выбрана стратегия C.

Приводится пример построения характеристической функции для одношаговой игры трёх лиц, соответствующей данной модели. Кроме того, исследуется частично кооперативное и некооперативное поведение групп в рассматриваемой модели. Приводятся примеры известных экономических вопросов, таких как «трагедия общин», которые могут рассматриваться с использованием данной модели.

Для обобщения результатов о равновесии в конечной модели, Кэрролом была введена функция ограниченной вероятности [3]. Он доказал теорему о некооперативном равновесии по Нэшу в повторяющейся «дилемме заключённого n лиц» с функцией ограниченной вероятности (аналогом конечной игры).

В работе Петросяна-Грауэр (2002г.) [10] рассматривается повторяющаяся дилемма и находится дисконтирующий фактор, обеспечивающий эффективное наказание в бесконечношаговой игре. Доказывается теорема о том, что при соблюдении определённых границ фактора дисконтирования существует сильное равновесие по Нэшу и это равновесие найдено. Доказывается, что для бесконечно повторяющейся игры «дилемма заключённого n лиц», С-ядро не пусто.

В статье Петросяна-Грауэр (2004г.) [11] строится характеристическая функция для повторяющейся модели «дилемма заключённого n лиц». Доказывается теорема о существовании равновесия в повторяющейся игре с выигрышами большими, чем при некооперативном поведении всех игроков, если выигрыши каждого из игроков при общем некооперативном поведении отрицательны.

Силагий (Szilagyi M. N.) [15] рассмотрел различные типы поведения игроков в повторяющейся игре, которые строятся в зависимости от статистики сыгранных этапов:

- Павловский: вероятность конкретной стратегии p на данном этапе меняется пропорционально поощрению и наказанию от предыдущих сыгранных профилей;
- Предсказуемая случайность: p - константа. Она не зависит от действий остальных игроков;
- Бухгалтерский: вероятность зависит от среднего поощрения за стратегии на предыдущих этапах;
- Конформистский: повторение действий большинства на предыдущем этапе;
- Жадный: постоянно выбирает стратегию «отклонение», для получения наибольшего выигрыша.

С помощью методов машинного обучения было рассмотрено изменение стратегий в игре, исходя из типов игроков.

В работе Такезава и Прайс [16] осуществляется поиск эволюционно-устойчивых и нейтрально-устойчивых стратегий. Проведены агентно-ориентированные симуляции динамики системы и стабильности кооперативных и разнородных популяций. В среднем, стратегии колеблятся от кооперативной до менее кооперативной, но никогда не концентрировалась на некооперативном равновесии по Нэшу.

Эксперимент над шестьюдесятью волонтёрами рассмотрен в статье Гоэринг Д. Й. и Кахан Й. П. (Goehring D. J. и Kahan J. P.) [6]. Игры в тройках, разбитых по половому признаку, показали, что альтруистичное поведение в модели «дилемма заключённого n лиц» равновероятно для представителей обоих полов, а зависит лишь от выигрыша от отклонения (разницы между выигрышами при кооперативном и девиантном поведении отдельного игрока).

Аналогичный эксперимент был рассмотрен в работе Фокса-Гайера [5]. Однако, им удалось отследить не только влияние пола, выигрыша от отклонения на альтруизм в модели «дилемма заключённого n лиц», но также и такого немаловажного фактора, как анонимность выбора и личности игрока для остальных участников, несмотря на то, что игра была статической. Т. е. наказание за отклонение не могло быть выражено в денежной форме, но всё равно некоторым образом влияло на выбор людей.

Бонацич (Bonacich P. at al.) [2] экспериментально показали, что в повторяющейся конечной модели «дилемма заключённого n лиц» вероятность увеличения количества кооперирующихся игроков на следующем этапе возрастёт, если их количество выросло хотя бы на одного на данном этапе по сравнению с предыдущим.

1. Описание модели «дилеммы заключённого для n лиц».

Рассмотрим игру «дилемма заключённого n лиц». В этой игре каждый из игроков имеет две доступные чистые стратегии: C («кооперироваться») и D («отклониться»).

Строится статическая игра Г. Множество всех игроков обозначим за N, мощность которого равна $|N| = n$. Множество чистых стратегий игроков x_i зададим как $x_i \in (C, D)\,,\ \forall\, i \in N$.

Предположим, что функции выигрыша каждого из игроков H_i, $\forall\, i \in N$, линейно зависят от числа игроков, выбирающих стратегию «кооперироваться» и стратегии этого (i-ого) игрока:

$$H_i\left(x_1, \ldots, x_i, \ldots, x_n\right) = \begin{cases} C_i\left(x\right) = a_1 x + b_1,\ \forall\ x \in \left(0, n\right], \text{если } x_i = C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i\left(x\right) = a_2 x + b_2,\ \forall\ x \in \left[0, n\right), \text{если } x_i = D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Эта модель для конкретных значений параметров $a_1,\ a_2,\ b_1,\ b_2$ рассматривалась во многих исследованиях [1–7, 10–11, 14–19]. Функция выигрышей игроков $H_i\left(x_1,\ \ldots,\ x_i,\ \ldots,\ x_n\right)$ при этом должна удовлетворять следующим предположениям:

1. $D_i(x-1) > C_i(x), \forall x \in [1, n-1]$, как и в случае со стандартной дилеммой заключённых для двух игроков, в модели для n лиц стратегия

D строго доминирует стратегию C для каждого из игроков. Поскольку в первом случае, $D_i(x-1)$, по крайней мере i-ый игрок выбирает стратегию «отклоняться» (D), а во втором, $C_i(x)$, как минимум i-ый игрок играет стратегию «кооперироваться» (C), это неравенство рассматривается только для количества кооперирующихся игроков (x) от одного до $(n-1)$.

2. $C_i(n) > D_i(0)$, а, значит, ситуация выбора стратегии «кооперироваться» (C) всеми игроками оптимальна по Парето в отличие от ситуации выбора всеми стратегии «отклоняться» (D), или, что то же самое, $H_i(C,\ldots,C) > H_i(D,\ldots,D)$.

3. $D_i(x) \geqslant D_i(0)$ и $C_i(x) \geqslant C_i(1)$, что говорит о том, что выигрыш игроков от выбора той или иной стратегии при кооперации x игроков по крайней мере не хуже, чем при отсутствии кооперирующихся игроков. Необходимо учитывать, что при выборе стратегии кооперации i-м игроком, как минимум этот игрок уже будет кооперироваться, т. е. $x \neq 0$. При этом для $D_i(x)$ количество кооперирующихся игроков не может равняться n, поскольку i-ый игрок играет стратегию «отклоняться» (D), а, значит, $x \neq n$.

4. $C_i(x) = C_j(x)$ и $D_i(x) = D_j(x)$, иными словами, выигрыши игроков зависят лишь от стратегии этого игрока и количества кооперирующихся игроков, а именно игроков, выбирающих стратегию «кооперироваться» (C), но не зависят от личности этого игрока, другими словами, все игроки симметричны.

Коэффициенты a_1, a_2, b_1, b_2 подбираются с учётом перечисленных четырёх предположений, при этом, зачастую $a_1 = a_2$, а $b_2 > b_1$, что, как видно, не противоречит ни одному из четырёх выдвинутых ранее предположений, а также позволяет обеспечить возможность рассматривать модель для больших величин n.

Для сопоставления полученных в данной работе результатов и уже существующей литературы по данной теме стоит описать, каким образом находить рассматриваемые в функции выигрыша коэффициенты a_1, a_2, b_1, b_2 в уже реализованных играх, которые были описаны в существующей на данный момент литературе.

В части из этих исследований [10], [14], [16] выигрыши игроков модели «дилемма заключённых n лиц» записываются в виде открытой части таблицы (она может быть продолжена для увеличивающегося количества игроков), где по горизонтали отражено количество кооперирующихся игроков, за исключением данного (i-ого), а каждой строке соответствует одна из чистых стратегий $x_i \in \{C, D\}$ этого (i-ого) игрока. Данная форма записи выглядит следующим образом (см. Таблицу 1):

Таблица 1. Дилемма заключённого n лиц (по диагонали количество кооперирующихся игроков без учёта данного)

	0	**1**	**2**	**3**	**4**
C	$a_1 + b_1$	$2a_1 + b_1$	$3a_1 + b_1$	$4a_1 + b_1$	$5a_1 + b_1$
D	b_2	$a_2 + b_2$	$2a_2 + b_2$	$3a_2 + b_2$	$4a_2 + b_2$

где на месте выражений из коэффициентов a_1, a_2, b_1, b_2 обычно стоят конкретные числа. Составлением системы уравнений можно привести этот вариант записи выигрышей к обобщённой системе, предлагаемой в начале данной главы.

Кроме того, в некоторых из исследований [2], [5], [6] также принята иная табличная форма записи выигрыша i-ого игрока, где по горизонтали учитывается общее количество кооперирующихся игроков, в том числе включая данного (i-ого) игрока (см. Таблица 2):

Таблица 2. Дилемма заключённого n лиц (по диагонали количество кооперирующихся игроков с учётом рассматриваемого)

	0	**1**	**2**	**3**	**4**
C		$a_1 + b_1$	$2a_1 + b_1$	$3a_1 + b_1$	$4a_1 + b_1$
D	b_2	$a_2 + b_2$	$2a_2 + b_2$	$3a_2 + b_2$	$4a_2 + b_2$

где выигрыши игрока при кооперации сдвинуты на одну ячейку вправо по сравнению с предыдущим вариантом записи без учёта i-ого игрока (см. Таблицу 1).

Оба этих варианта записи выигрышей игроков делают процесс поиска общих решений, принципов оптимальности и прочего трудоёмким, поскольку делают необходимым первоначальных пересчёт выигрышей всех игроков до необходимого количества.

Лишь в некоторых случаях, при малых n, обычно не превышающих 3 игроков, используют матричную форму записи выигрышей игроков [17], но в этом случае перевод в обобщённую функцию выигрыша не представляет никаких сложностей. В дальнейшем матричная форма записи может позволить развивать модель «дилеммы заключённых n лиц» в сторону несимметричности игроков, оставляя при этом соблюдение первых трёх общих положений для данной игры. Особенно актуально это для кооперативной теории игр, поскольку это позволит избежать тривиальных решений для вектора Шэпли и иных возможных дележей.

Обобщённая форма записи функции выигрыша игроков позволит упростить дальнейшие расчёты, связанные с вычислением характеристической функции динамической игры, а также C-ядра, D-подъядра и вектора Шепли для стохастической и динамической модели «динамической модели дилеммы заключённых n лиц».

2. Определение эффективного наказания.

2.1. Модель кооперации в повторяющейся игре. Рассмотрим повторяющуюся игру Γ, которая повторяется K раз. В данной повторяющейся игре построим некоторый способ поведения, который с одной стороны обеспечивает игрокам высокие выигрыши, а с другой – устойчив относительно отклонения коалиций, т. е. являлся бы сильным равновесием по Нэшу [10], или устойчив относительно индивидуальных отклонений.

Предположим, что первые $K-k^*$ шагов все игроки выбирают стратегию C. На каждом из этих шагов суммарный выигрыш игроков будет равен

$$A\left(N\right)=\sum_{i=1}^{n}H_i\left(C,\ldots,C\right)=\sum_{i=1}^{n}H_i\left(\overline{x}\right).$$

Поскольку стратегия D является строго доминирующей на каждом этапе игры, согласно первому положению о выигрышах в данной модели, каждый из игроков имеет склонность к отклонению в том смысле, что его индивидуальный выигрыш увеличится в независимости от того, какие стратегии выбрали остальные участники. Значит, отклониться выгодно каждому игроку, поскольку его выигрыш при кооперативном поведении остальных игроков составит

$$\overline{V}\left(i\right)=H_i\left(\overline{x}||x_i\right)>H_i\left(\overline{x}\right)=A\left(i\right).$$

Соответственно, необходимо предусмотреть «наказание» для отклоняющихся игроков, обеспечивающее кооперативное поведение на данном этапе. Выберем k^* таким образом, чтобы отклонение от кооперации любой коалиции на первых $K-k^*$ шагах было бы не выгодно этой коалиции в том смысле, что при отклонении сумма выигрышей всех игроков, входящих в коалицию, строго уменьшится.

Теорема 1. *Ситуация в повторяющейся «дилемме заключённого для n лиц», в которой все игроки выбирают стратегию C на первых $K - k^*$ шагах, а затем переключаются на стратегию D, является сильным равновесием на первых $K - k^*$ шагах, и устойчивой, относительно индивидуальных отклонений, на последних k^* шагах.*

Под сильным равновесием в Γ_k здесь понимается n-мерный профиль стратегий $x^*(\cdot) = (x_1^*(\cdot), \ldots, x_n^*(\cdot))$ такой, что

$$\sum_{i \in M} H_i(x^*(\cdot)) \geqslant \sum_{i \in M} H_i(x^*(\cdot) \| x_M(\cdot))$$

для всех $M \subset N$, $x_M(\cdot) \in \prod_{j \in M} X_j$ [10].

К сожалению, на шагах $[K - k^*, K]$ некоторым коалициям будет выгодно отклонение от стратегии D в том смысле, что их выигрыши могут быть увеличены, поскольку, по крайней мере профиль стратегий «кооперироваться всем игрокам» для максимальной коалиции приносит Парето-оптимальные выигрыши, в отличие от профиля «отклоняться всем игрокам», согласно второму положению о построении выигрышей в модели «дилемма заключённого n лиц». Однако, если на последних шагах игроки будут придерживаться стратегии D, то, поскольку ситуация $(D, \ldots, D)$ является равновесием по Нэшу в одношаговой игре, индивидуальное отклонение невыгодно.

Доказательство. Построим ситуацию равновесия в дилемме заключённого n лиц, которая удовлетворяет условию сильного равновесия по Нэшу, т. е. устойчива относительно отклонения коалиции любой мощности.

Примем за $\overline{X}_i$ стратегию, при которой, как только игрок получает выигрыш меньший, чем при полной кооперации, переключается на игру «против каждого отклонившегося». Поскольку выигрыши от выбора стратегий D или C имеют положительную зависимость от количества игроков, выбирающих стратегию C, единственный способ минимизировать выигрыш каждого из отклонившихся игроков, это выбирать стратегию D. Примечательно, что в данной игре стратегии игры «против всех» и «против каждого отклонившегося» [13] совпадают. Важен лишь сам факт отклонения хотя бы одного игрока, но не важно, какой именно это игрок или каких размеров коалиция при этом отклонилась. Таким образом, пусть все n игроков на каждом шаге l^* из первых $K - k^*$ шагов играют триггерную стратегию, т. е.:

$$\overline{x}_i = \begin{cases} C, & \text{если } x_j = C,\ \forall\ j \in \{N \setminus i\}, \forall l < l^*, \\ D, & \text{иначе.} \end{cases}$$

Попытаемся отыскать такое число шагов k^*, которое может гарантировать кооперацию, а, значит, будет достаточным, чтобы эффективно «наказать» отклонившуюся коалицию.

Рассмотрим коалицию $M \subseteq N$, мощность которой равна $|M| = m$. Согласно нашим предположениям, на начальном этапе все n игроков играют стратегию C («кооперироваться»). В связи с этим, выигрыш коалиции M в данных условиях равен

$$A(M) = (a_1 n + b_1) m = a_1 nm + b_1 m.$$

Данная коалиция в целях максимизации суммы выигрышей игроков, входящих в неё, может быть заинтересована в отклонении от стратегии $\overline{X}_M$. Положим, что остальные игроки $N \backslash M$, продолжают играть стратегию $\overline{X}_{N \backslash M}$, а, значит, на этапе отклонения $\overline{x}_i = C$, для всех $i \in \{N \setminus M\}$. Обозначим максимальный доход, который получает коалиция M при отклонении на данном шаге,

$$\overline{W}(M) = \max_{x_M \in \prod_{i \in M} x_i} \sum_{i \in M} h_i \left(x_M, \overline{x}_{N \setminus M}\right).$$

В нашем случае это сумма выигрышей D_M для m игроков при равенстве $n-m$ количества игроков, играющих стратегию «кооперироваться» (D)

$$\overline{W}(M) = \sum_{i \in M} D_i (n-m)$$

А поскольку все игроки симметричны, согласно четвёртому положению о выигрышах игроков в рассматриваемой модели, то

$$\overline{W}(M) = mD_i (n-m) = m (a_2 (n-m) + b_2) = a_2 nm - a_2 m^2 + b_2 m.$$

Необходимо отметить, что на данном шаге получают выигрыш больший, чем при кооперативном поведении, только те коалиции $\overline{M}$, чья мощность не превышает значения $\overline{m}$, при котором $\overline{W}(M) > A(M)$

$$a_2 nm - a_2 m^2 + b_2 m > a_1 nm + b_1 m.$$

Значит, мощность отклоняющейся коалиции для получения выгоды от отклонения должна удовлетворять следующему условию

$$\overline{m} < \frac{a_2 n + b_2 - a_1 n - b_1}{a_2}.$$

Следовательно, нам достаточно доказать, что наказание будет эффективно для коалиций с мощностью меньше, чем $\overline{m}$.

Поскольку все члены коалиции M сыграют стратегию D, согласно триггерной стратегии, на каждом последующем шаге все члены коалиции $\{N \setminus \{M\}\}$ будут играть «против каждого отклонившегося», а следовательно, стратегию «отклоняться» (D). В таких условиях, решение о последующей стратегии коалиции M необходимо принимать, исходя из функции, максимизирующей сумму математических ожиданий выигрышей игроков из M:

$$\overline{V}\left(M\right) = \max_{x_M}\left[\sum_i h_i\left(x_M, \overline{x}_{N\setminus M}\right)\right].$$

Поскольку $(n-m)$ игроков на данном этапе будут играть стратегию D, данное значение можно вычислить как

$$\overline{V}\left(M\right) = \max\left\{\sum_{i\in M} D_i\left(0\right), \sum_{i\in M} C_i\left(m\right)\right\}.$$

Симметричность игроков даёт нам возможность вычислить это значение, как

$$\overline{V}\left(M\right) = \max\left\{m\left(a_1 m + b_1\right), m b_2\right\}.$$

Таким образом, получаем, что на этапе «наказания» выигрыш отклонившейся коалиции на каждом шаге будет равен

$$\overline{V}\left(M\right) = \begin{cases} a_1 m^2 + b_1 m, & \text{если } m \geqslant \frac{b_2 - b_1}{a_1}, \\ b_2 m, & \text{иначе.} \end{cases}$$

Определим количество шагов, гарантирующих кооперацию игроков, а, значит, позволяющих уменьшить суммарный выигрыш коалиции за весь период игры таким образом, чтобы он был меньше, чем при использовании триггерной стратегии всеми игроками на всех этапах игры

$$k^* \left(\overline{V}\left(M\right) - A\left(M\right)\right) \leqslant A\left(M\right) - \overline{W}\left(M\right),$$

т. е. при $k^* \geqslant \frac{\overline{W}(M)-A(M)}{A(M)-\overline{V}(M)}$ наказание эффективно.

Поскольку выражение для $\overline{V}\left(M\right)$ различно для нескольких вариантов мощности коалиции, найдём k^* для каждого из них.

Для коалиции M, количество членов которой меньше, чем $\frac{b_2-b_1}{a_1}$, получим следующее количество шагов

$$\underline{k}^* \left(a_1 n + b_1 - b_2\right) > a_2 n - a_2 m + b_2 - a_1 n - b_1.$$

В таком случае, при количестве игроков в коалиции $m < \frac{b_2-b_1}{a_1}$, им будет невыгодно отклоняться на первых $K - \underline{k}^*$ шагах:

$$\underline{k}^* > \frac{a_2 n - a_2 m + b_2 - a_1 n - b_1}{a_1 n + b_1 - b_2}.$$

Теперь определим, какое количество шагов коалиция с мощностью $\frac{b_2-b_1}{a_1} \leqslant m < \frac{a_2 n + b_2 - a_1 n - b_1}{a_2}$ гарантированно не будет отклоняться. Необходимо помнить, что коалиции с мощностью $m \geqslant \frac{a_2 n + b_2 - a_1 n - b_1}{a_2}$ не имеет смысла отклоняться даже на последнем шаге игры. Если оставшееся до окончания игры количество шагов будет удовлетворять условию, что

$$\overline{k}^* \left(a_1 n m + b_1 m - a_1 m^2 - b_1 m\right) > a_2 n m - a_2 m^2 + b_2 m - a_1 n m - b_1 m.$$

то коалиции с мощностью $\frac{b_2-b_1}{a_1} \leqslant m < \frac{a_2n+b_2-a_1n-b_1}{a_2}$ также будет невыгодно отклонение.

Тогда, при мощности $\frac{b_2-b_1}{a_1} \leqslant m < \frac{(a_2n+b_2-a_1n-b_1)n}{a_2}$ получим

$$\overline{k}^* \geqslant \frac{a_2nm - a_2m^2 + b_2m - a_1nm - b_1m}{a_1nm - a_1m}.$$

Поскольку для рассмотренных интервалов количество шагов, необходимых для обеспечения кооперации, уменьшается с увеличением мощности коалиции, мы можем определить пороговое значение k^*, обеспечивающее эффективность наказания для любых коалиций $\max(k^*) = k^*(1)$:

$$\max(k^*) = \begin{cases} \frac{a_2(n-1)+b_2-a_1n-b_1}{a_1(n-1)}, & \text{если } \frac{b_2-b_1}{a_1} \leqslant 1, \\ \frac{a_2(n-1)+b_2-a_1n-b_1}{a_1n+b_1-b_2}, & \text{иначе.} \end{cases}$$

Теорема доказана.

Замечание. Как уже писалось ранее в главе 1, зачастую [1, 2, 5, 6, 9, 10] в играх, основанных на данной модели, коэффициенты a_1, a_2, b_1, b_2 удовлетворяют дополнительным условиям, что $a_1 = a_2$ и, тогда, $b_2 > b_1$. В этом случае,

$$a_1x + b_1 > a_2(x-1) + b_2,$$

учитывая первое положение о выигрышах игроков.

Используя равенство $a_1 = a_2$, можем переписать, как

$$a_1(x - (x-1)) > b_2 - b_1.$$

А, поскольку $b_2 > b_1$

$$\frac{a_1}{b_2 - b_1} > 1.$$

Т. е. для большинства исследуемых на данный момент в литературе игр количество шагов, обеспечивающих эффективное наказание, будет равно

$$k^* = \frac{a_2\left(n-1\right) + b_2 - a_1 n - b_1}{a_1 n + b_1 - b_2}.$$

Пример 1. Построим многошаговую повторяющуюся игру трёх лиц Γ_k, соответствующую рассматриваемой модели. Запишем данную игру в матричной форме, где первый игрок — строчный, второй — столбцовый, третий — страничный. Порядок выигрышей соответствует номерам игроков (см. таблицы 3,4).

Таблица 3. Дилемма заключённого n лиц (третий игрок выбирает чистую стратегию C)

C	C	D
C	(6, 6, 6)	(4, 12, 4)
D	(12, 4, 4)	(8, 8, 2)

Таблица 4. Дилемма заключённого n лиц (третий игрок выбирает чистую стратегию D)

D	C	D
C	(4, 4, 12)	(2, 8, 8)
D	(8, 2, 8)	(4, 4, 4)

Представим выигрыши каждого игрока при выборе каждой из стратегий, как функцию от числа игроков, выбирающих стратегию C, следующим образом:

$$H_i\left(x_1,x_2,x_3\right)=\begin{cases} C_i\left(x\right)=2x,\ \forall\ x\in\left(0,3\right],\text{если } x_i=C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C; \\ D_i\left(x\right)=4x+4,\ \forall\ x\in\left[0,3\right),\text{если } x_i=D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Для данной игры, поскольку $\frac{b_2-b_1}{a_1}>1$, количество шагов, гарантирующих кооперацию, равно

$$k^*=\frac{4*(3-1)+4-2*3-0}{2*3+0-4}=3,$$

т. е. достаточно трёх шагов, чтобы эффективно «наказать» отклонившуюся коалицию.

Как итог, нами был получен новый равновесный профиль стратегий «всем игрокам играть триггерную стратегию $\overline{x}$ до момента $K-k^*$, а затем переключаться на стратегию D». Эта равновесная ситуация является сильным равновесием по Нэшу на первых $K-k^*$ шагах. На последних k^* шагах равновесие будет устойчивым только относительно индивидуальных отклонений, поскольку возможно отклонение коалиции $\underline{M}$, мощность $\underline{m}$ которой удовлетворяет условию

$$b_2\underline{m}<C_i\left(M\right)\underline{m},$$

Таким образом, для коалиции $\underline{M}$

$$\underline{m} > \frac{b_2 - b_1}{a_1}.$$

выигрыш на последних k^* может быть увеличен, что, к тому же, приведёт к повышению выигрыша остальных $N \setminus M$ игроков, поскольку выигрыши каждого отдельного игрока имеют прямую положительную зависимость от количества игроков, выбирающих стратегию поведения D.

В данной многошаговой игре Γ_k найдено пороговое значение k^* так, что на $(K - k^*)$ шагах отклонение от ситуации $(C, \ldots, C)$ не выгодно ни для какой коалиции игроков, а на шагах, начиная с $K - k^*$, отклонение от стратегии D невыгодно отдельным игрокам и, кроме того, коалициям, чья мощность не превышает значения .

2.2. Модель кооперации в динамической игре.

Пусть имеется конечный набор из f игр, удовлетворяющих модели «дилемма заключённого n лиц»: $(\gamma_1, \ldots, \gamma_f)$. На каждом из K_f шагов разыгрывается одна из этих игр. Определим количество шагов k_f^*, необходимых для эффективного наказания в полученной игре Γ_f.

Функции выигрыша игроков в каждой из данного набора f игр зададим как

$$H_i^{\gamma_j}(x_1, \ldots, x_i, \ldots, x_n) = \begin{cases} C_i^{\gamma_j}(x) = a_1^{\gamma_j} x^{\gamma_j} + b_1^{\gamma_j},\ \forall\ x^{\gamma_j} \in (0, n]\,, \text{если } x_i^{\gamma_j} = C \\ \text{и известно, что } x^{\gamma_j} \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i^{\gamma_j}(x^{\gamma_j}) = a_2^{\gamma_j} x + b_2^{\gamma_j},\ \forall\ x^{\gamma_j} \in [0, n)\,, \text{если } x_i^{\gamma_j} = D \\ \text{и известно, что } x^{\gamma_j} \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Построим выигрыш от отклонения для коалиции M в данной игре, используя эти коэффициенты,

$$\overline{W}^{\gamma_{j^*}}(M) - A^{\gamma_{j^*}}(M) = a_2^{\gamma_{j^*}} nm - a_2^{\gamma_{j^*}} m^2 + b_2^{\gamma_{j^*}} m - a_1^{\gamma_{j^*}} nm + b_1^{\gamma_{j^*}} m.$$

Наибольший выигрыш от отклонения

$$\max_{m \in (1,n)} \frac{\left(a_2^{\gamma_{j^*}}(n-m) + b_2^{\gamma_{j^*}}\right) m - \left(a_1^{\gamma_{j^*}} n + b_1^{\gamma_{j^*}}\right) m}{m}$$

получит коалиция M мощностью $|M| = 1$, поскольку функция выигрыша имеет прямую положительную связь с количеством игроков, выбирающих стратегию C.

Эффект от наказания, соответственно будет равен разности между выигрышем при полной кооперации и выигрышем, который гарантированно может обеспечить себе отклонившаяся коалиция

$$A^{\gamma_j}(M) - \overline{V}^{\gamma_j}(M) = a_1^{\gamma_j} nm + b_1^{\gamma_j} m - \max\left\{a_1^{\gamma_j} m^2 + b_1^{\gamma_j} m; b_2^{\gamma_j} m\right\},$$

при этом необходимо понимать, что это разные этапы игры, т. е. на этапе наказания $j > j^*$.

Наименьший эффект от наказания в таком случае будет достигаться, если разница выигрышей при отклонении и кооперативном поведении будет максимальна на этапе отклонения, когда

$$\max_{j\in[1,f]}\left(W^{\gamma_j}(1) - A^{\gamma_j}(1)\right),$$

а разница между выигрышем отдельного игрока при кооперативном поведении всех игроков и значением игры для этого игрока на этапе наказания будет минимальна

$$\min_{j\in[1,f]}\left(A^{\gamma_j}(1) - V^{\gamma_j}(1)\right).$$

В этом случае, можем определить пороговое значение количества шагов для эффективного наказания, когда выигрыш от отклонения будет меньше чем суммарное наказание за k^* шагов, т. е.

$$\max_{j\in[1,f]}\left(W_{\gamma_j}\left(1\right)-A_{\gamma_j}\left(1\right)\right)\leqslant k_f^*\min_{j\in[1,f]}\left(A_{\gamma_j}\left(1\right)-V_{\gamma_j}\left(1\right)\right).$$

В полученной многошаговой игре это достигается, когда, учитывая имеющийся набор из f игр, удовлетворяющих модели «дилемма заключённого n лиц»,

$$\max_{j\in[1,f]}\left(\left(a_2^{\gamma_j}n-a_2^{\gamma_j}+b_2^{\gamma_j}\right)-\left(a_1^{\gamma_j}+b_1^{\gamma_j}\right)\right)\leqslant k_f^*\min_{i\in[1,f]}\left(\left(a_2^{\gamma_j}n-a_2^{\gamma_j}+b_2^{\gamma_j}\right)-b_2^{\gamma_j}\right).$$

Соответственно, ни одной коалиции невыгодно будет отклониться от кооперативного поведения, если до конца игры останется шагов больше, чем

$$k_f^*\geqslant\frac{\max_{j\in[1,f]}\left(a_2^{\gamma_j}n-a_2^{\gamma_j}+b_2^{\gamma_j}-a_1^{\gamma_j}-b_1^{\gamma_j}\right)}{\min_{i\in[1,f]}\left(a_2^{\gamma_j}n-a_2^{\gamma_j}\right)}.$$

3. Подъядро игры «дилемма заключенного n лиц».

Рассмотрим игру Γ_f. Пусть

$$V^{\gamma_j}(N) = \max_{x_1,\ldots,x_i,\ldots,x_n} \sum_{i\in N} H_i^{\gamma_j}(x_1,\ldots,x_i,\ldots,x_n)$$

одинаково для всех $\gamma_j, j \in [1, f]$.

Определение 1. C-ядром кооперативной игры Γ_f называется набор возможных дележей $(\alpha_1,\ldots,\alpha_n)$ выигрыша $V(N)$, которые не могут улучшить ни одна коалиция и ни один индивидуальный игрок, т. е. выполняются условия:

- Парето-эффективности, т. е. $\sum_{i=1}^{n} \alpha_i = V(N)$;
- коалиционной рациональности, или $\sum_{i=1}^{s} \alpha_i \geqslant V(S)\ \forall\ S \subset N$;
- индивидуальной рациональности, а, значит, $x_i \geqslant V(i)$, $\forall i \in N$ [8].

Построим C-ядро для игры Γ_f. Значение игры для любой из коалиций на каждом из этапов игры не зависит от их выигрыша на любом другом этапе. В таком случае, можем считать, что математическое ожидание значения игры для коалиции $S,\ S \subset N$

$$V^{\Gamma_f}(S) = K_f \sum_{j=1}^{f} p_j V^{\gamma_j}$$

Рассмотрим две коалиции S и : $S \cup M = N$, $S \cap M = \varnothing$. Их мощности равны, соответственно $|S| = s$, $|M| = m$. Докажем, что

$$V^{\gamma_j}(S) + V^{\gamma_j}(M) \leqslant V^{\gamma_j}(N).$$

Найдём значения каждой из этих коалиций.

$$V^{\gamma_j}(S)=\max_{r\in[0,s]}\left(r\left(a_1^{\gamma_j}r+b_1^{\gamma_j}\right)+(s-r)\left(a_2^{\gamma_j}r+b_2^{\gamma_j}\right)\right);$$

$$V^{\gamma_j}(M)=\max_{t\in[0,m]}\left(t\left(a_1^{\gamma_j}t+b_1^{\gamma_j}\right)+(m-t)\left(a_2^{\gamma_j}t+b_2^{\gamma_j}\right)\right).$$

Значение максимальной коалиции равно

$$V^{\gamma_j}(N)=\max_{u\in[0,n]}\left(u\left(a_1^{\gamma_j}u+b_1^{\gamma_j}\right)+(n-u)\left(a_2^{\gamma_j}u+b_2^{\gamma_j}\right)\right).$$

Максимум последней функции достигается при $u=\frac{\left(2a_1^{\gamma_j}-a_2^{\gamma_j}\right)n+\left(a_2^{\gamma_j}-b_2^{\gamma_j}+b_1^{\gamma_j}-a_1^{\gamma_j}\right)}{2a_1^{\gamma_j}-2a_2^{\gamma_j}}$. Максимумы значений остальных коалиций достигаются при $r\leqslant u$ и $t\leqslant u$. Поскольку $D_i^{\gamma_j}$ и $C_i^{\gamma_j}$ линейно зависят от количества кооперирующихся игроков и, кроме того, $s+m=n$, очевидно, что $V^{\gamma_j}(S)+V^{\gamma_j}(M)\leqslant V^{\gamma_j}(N)$.

Тогда С-ядром «дилеммы заключённого n лиц» будет делёж:

$$\begin{cases}\sum_{i=1}^{N}\alpha_i=\max_{u\in[0,n]}\left(u\left(a_1^{\gamma_j}u+b_1^{\gamma_j}\right)+(n-u)\left(a_2^{\gamma_j}u+b_2^{\gamma_j}\right)\right);\\ \sum_{i=1}^{S}\alpha_i\geqslant\max_{r\in[0,s]}\left(r\left(a_1^{\gamma_j}r+b_1^{\gamma_j}\right)+(s-r)\left(a_2^{\gamma_j}r+b_2^{\gamma_j}\right)\right),\forall S\subset N.\end{cases}$$

Пусть $|N|=n=3$.

Теорема 2. *В динамической игре Γ_f, каждый шаг которой удовлетворяет модели «дилемма заключённого n лиц» для трёх игроков, D-подъядром Петросяна-Панкратовой [12] , принадлежащим C-ядру, является набор дележей:*

$$\alpha_i \geqslant K_f \max_{j\in[1,f]} b_2^{\gamma_j},$$

$$\alpha_i+\alpha_j \geqslant K_f \max\left(\max_{j\in[1,f]}\left(4a_1^{\gamma_j}+2b_1^{\gamma_j}\right);\max_{j\in[1,f]}\left(a_1^{\gamma_j}+b_1^{\gamma_j}+a_2^{\gamma_j}+b_2^{\gamma_j}\right);\max_{j\in[1,f]}\left(2b_2^{\gamma_j}\right)\right),$$

$$\alpha_i+\alpha_j+\alpha_k = K_f \max_{j\in[1,f]}\left(9a_1^{\gamma_j}+3b_1^{\gamma_j}\right),\ \forall\ i,\ j,\ k\in N$$

Доказательство. Поскольку равновесием по Нэшу в доминирующих стратегиях для одношаговой игры $\gamma_j, \forall j \in f$ является ситуация $(D, \ldots, D)$, значение игры для одного игрока в игре γ_j составит

$$V^{\gamma_j}(i) = H_i^{\gamma_j}(D,D,D) = b_2^{\gamma_j},\ \forall\ i \in N.$$

Все игроки симметричны. Следовательно, значением игры для двух игроков будет максимум трёх сумм $\sum_{i\in S} H_i^{\gamma_j}(D,D,D)$, $\sum_{i\in S} H_i^{\gamma_j}(C,D,D)$ или $\sum_{i\in S} H_i^{\gamma_j}(C,C,D)$, где $S=\{1,2\}$. Т. е.,

$$V^{\gamma_j}(1,2) = V^{\gamma_j}(1,3) = V^{\gamma_j}(2,3) = max\left(4a_1^{\gamma_j}+2b_1^{\gamma_j}; a_1^{\gamma_j}+b_1^{\gamma_j}+a_2^{\gamma_j}+b_2^{\gamma_j}; 2b_2^{\gamma_j}\right)$$

Поскольку стратегия «кооперироваться» (C) для всех игроков оптимальна по Парето в отличие от стратегии «отклоняться» (D),

$$V^{\gamma_j}(N) = \max\left\{9a_1^{\gamma_j}+3b_1^{\gamma_j}; 4a_1^{\gamma_j}+2b_1^{\gamma_j}+2a_2^{\gamma_j}+b_2^{\gamma_j}; a_1^{\gamma_j}+b_1^{\gamma_j}+2a_2^{\gamma_j}+2b_2^{\gamma_j}\right\}$$

Пусть

$$W(S) = \max_{j\in f} V(\gamma_j, S), S \subset N$$

Обозначим за $\overline{\gamma}_j$ подыгру игры Γ_f, которая начинается с этапа j, где $j \in [1, K_f]$. Т.е. подыгра $\overline{\gamma}_1$ совпадает с игрой Γ_f.

Тогда построим новую характеристическую функцию $\overline{W}\left(\Gamma_f, S\right)$, такую, что:

$$\overline{W}\left(\overline{\gamma}_j, S\right) = \left(K_f - j + 1\right) W\left(S\right),$$

где $j \in [1, K_f]$ – это номер шага реализуемой игры.

Поскольку $V^{\gamma_j}\left(N\right) = \max_{x_1,\ldots,x_i,\ldots,x_n} \sum_{i \in N} H_i^{\gamma_j}\left(x_1, \ldots, x_i, \ldots, x_n\right)$ одинаково для всех $\gamma_j, j \in [1, f]$, можем получить множество дележей, принадлежащих C-ядру [12]:

$$\sum_{i \in M} \alpha_i^{\gamma_f} \geqslant \overline{W}\left(i\right), \text{ где } \overline{W}\left(i\right) = K_f \max_{j \in [1,f]} V^{\gamma_j}\left(M\right).$$

$$\alpha_i + \alpha_j \geqslant \overline{W}\left(S\right), \text{ где } |S| = 2 \text{ и}$$

$$\overline{W}\left(S\right) = K_f \max\left(\max_{j \in [1,f]}\left(4a_1^{\gamma_j} + 2b_1^{\gamma_j}\right); \max_{j \in [1,f]}\left(a_1^{\gamma_j} + b_1^{\gamma_j} + a_2^{\gamma_j} + b_2^{\gamma_j}\right); \max_{j \in [1,f]}\left(2b_2^{\gamma_j}\right)\right).$$

$$\sum_{i \in N} \alpha_i^{\gamma_f} = \overline{W}\left(N\right), \text{ где } \overline{W}\left(N\right) = K_f V^{\gamma_j}\left(N\right), \forall j \in [1, f].$$

Теорема доказана

Кроме того, данное подъядро обладает свойством сильной динамической устойчивости.

Определение 2. Подъядро $\overline{D}\left(\overline{\gamma}_1\right)$ называется сильно динамически устойчивым в игре Γ_f, если

- $\overline{D}\left(\gamma_{l+1}\right) \neq \varnothing, j \in [1, f]$;
- для всех дележей $\alpha^{\overline{\gamma}_1} \in \overline{D}\left(\overline{\gamma}_1\right)$, существует такая процедура распределе-

ния дележа $\beta = \left(\beta_1, \ldots, \beta_j, \ldots, \beta_{K_f}\right)$, что $\overline{D}\left(\overline{\gamma}_1\right) \supset \sum_{j=1}^{l} \beta_j \oplus \overline{D}\left(\gamma_{l+1}\right)$.

Здесь символ $\oplus$ определяется, как

$$\beta_j \oplus \overline{D}\left(\gamma_{l+1}\right) = \left\{\beta_j \oplus \overline{d}\left(\gamma_{l+1}\right) : \overline{d}\left(\gamma_{l+1}\right) \in \overline{D}\left(\gamma_{l+1}\right)\right\} \text{ [20]}.$$

Поскольку в каждой из f возможных игр γ_j, которые возможны на каждом из K_f этапов игры Γ_f $W(N)$ равны, то $\overline{W}(N) = K_f W(N)$.

Тогда, мы можем определить правило рапределения дележа, как частное от деления дележа $\alpha_i^{\Gamma_f}$ в игре Γ_f для каждого игрока $i, i \in [1, N]$ на количество шагов в этой игре K_f.

$$\beta_{i1} = \frac{\alpha_i^{\overline{\gamma}_1}}{K_f}.$$

В этом случае, на любом из остальных этапов игры, правило распределения дележа можно задать, как

$$\beta_{ij} = \frac{\alpha_i^{\overline{\gamma}_j}}{K_f - j + 1},$$

где $j \in [1, K_f]$.

Поскольку при пересмотре распределения дележей на любом из K_f этапов игры Γ_f, получим, что

$$\sum_{j\in[1,K_f]} \sum_{i\in S} \frac{\alpha_i^{\overline{\gamma}_j}}{K_f - j + 1} = \sum_{j=1}^{l} \sum_{i\in S} \frac{\alpha_i^{\overline{\gamma}_j}}{K_f - j + 1} + \sum_{j=l+1}^{K_f} \sum_{i\in S} \frac{\alpha_i^{\overline{\gamma}_j}}{K_f - j + 1} \geqslant$$

$$\geqslant lW(S) + (K_f - l)W(S) = \overline{W}(S).$$

Следовательно, D-подъядро получается сильно динамически устойчиво.

Пример 2. Построим многошаговую динамическую игру трёх лиц Γ_f, соответствующую рассматриваемой модели, учитывая, что максимальный выигрыш для коалиции из всех возможный игроков равен для каждой из рассматриваемых одношаговых этапов игры. Запишем данную игру в матричной форме, где первый игрок — строчный, второй — столбцовый, третий — страничный. Порядок выигрышей соответствует номерам игроков (см. таблицы 5–10).

Таблица 5. γ_1 (третий игрок выбирает чистую стратегию C)

C	C	D
C	(100, 100, 100)	(90, 115, 90)
D	(115, 90, 90)	(100, 100, 90)

Таблица 6. γ_1 (третий игрок выбирает чистую стратегию D)

D	C	D
C	(90, 90, 115)	(80, 100, 100)
D	(100, 80, 100)	(85, 85, 85)

Таблица 7. γ_2 (третий игрок выбирает чистую стратегию C)

C	C	D
C	(100, 100, 100)	(50, 105, 50)
D	(105, 50, 50)	(90, 90, 0)

Таблица 8. γ_2 (третий игрок выбирает чистую стратегию D)

D	C	D
C	(50, 50, 105)	(0, 90, 90)
D	(90, 0, 90)	(75, 75, 75)

Таблица 9. γ_3 (третий игрок выбирает чистую стратегию C)

C	C	D
C	(100, 100, 100)	(90, 110, 90)
D	(110, 90, 90)	(98, 98, 80)

Таблица 10. γ_3 (третий игрок выбирает чистую стратегию D)

D	C	D
C	(90, 90, 110)	(80, 98, 98)
D	(98, 80, 98)	(86, 86, 86)

Для определённости, зададим количество шагов для рассматриваемого примера, на которых разыгрываются с равной вероятностью каждая из этих игр, как $K_f = 5$, т. е. игра разыгрывается 5 этапов.

Определим значения каждой из коалиций для всех из этих f игр (см. таблицу 11):

Таблица 11. Значения игр γ_1–γ_3 для коалиций

M	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
V^{γ_1}	85	85	85	**180**	**180**	**180**	**300**
V^{γ_1}	75	75	75	150	150	150	**300**
V^{γ_1}	**86**	**86**	**86**	**180**	**180**	**180**	**300**

Тогда подъядром D C-ядра игры Γ_f для трёх игроков является множество:

$$\alpha_i \geqslant 430,\ \forall\ i \in N$$

$$\alpha_i + \alpha_j \geqslant 900,\ \forall\ i,\ j \in N,\ i \neq j$$

$$\alpha_i + \alpha_j + \alpha_k = 1500,\ \forall\ i,\ j,\ k \in N,\ i \neq j \neq k$$

Таким образом, D-подъядру C-ядра для динамической «дилеммы заключённого n игроков» принадлежат такие возможные дележи, как $(430, 470, 600)$, $(500, 500, 500)$ $(430, 535, 535)$ и другие. При этом, если к тому же на каждом шаге делёж игроков будет равен, например, $\frac{\alpha_i}{5}$, то становится понятно, что найденное подъядро D обладает свойством сильной динамической устойчивости.

4. Вектор Шепли для «дилеммы заключенного n лиц».

4.1. Вектор Шепли стохастической «дилеммы заключенного n лиц».

Определение 3. Вектором Шепли для игры Γ_f называется делёж $(\alpha_1, \ldots, \alpha_n)$ выигрыша $V(N)$, при котором каждый игрок получает среднее значение маржинального вклада в наибольший выигрыш максимальной коалиции N [9]. Это объём ожидаемого дополнительного вклада игрока при каждом из $n!$ очерёдностей формирования коалиции N при одинаковых вероятностях каждого из этих способов построения коалиции N.

Вектор Шепли удовлетворяет следующим аксиомам [18]:

- Эффективности: игроки распределяют между собой весь наибольший выигрыш максимальной коалиции

$$\sum_{i \in N} Sh_i(V) = V(N);$$

- Симметрии: если игроки i и j симметричны относительно V, тогда

$$Sh_i(V) = Sh_j(V);$$

- Аддитивности: сумма значений Шэпли для двух игр равна значению Шепли для суммы двух игр, т. е.

$$Sh_S(V) + Sh_S(W) = Sh_S(V + W);$$

- Болвана: если i-ый игрок является болваном, т. е. $V(S \cup i) - V(S) = 0$ для всех $S \subset N$, то

$$Sh_i(V) = 0$$

Рассмотрим игру Γ_f. Определим для каждой одношаговой игры γ_j, $j \in [1, f]$, значение игры для максимальной коалиции

$$V^{\gamma_j}(N) = \max_{x_1,\ldots,x_i,\ldots,x_n} \sum_{i\in N} H_i^{\gamma_j}(x_1,\ldots,x_i,\ldots,x_n).$$

Допустим, мощность отклонившейся коалиции S_j равна $|S_j| = s_j$. Тогда мощность кооперирующейся коалиции будет равна, соответственно $|N \setminus S_j| = n - s_j$. Выигрыш максимальной коалиции достигает своего максимума, когда выигрыш от отклонения каждого дополнительного игрока

$$D_i^{\gamma_j}(x = n-(s+1)) - C_i^{\gamma_j}(x = n-s) = \left(a_2^{\gamma_j}(n-s-1) + b^{\gamma_j}2\right) - \left(a_1^{\gamma_j}(n-s) + b_1^{\gamma_j}\right)$$

станет меньше суммы потерь всех остальных игроков, в т. ч. тех, кто уже отклонился

$$\sum_{i=1}^{n-s_j-1} \left(C_i^{\gamma_j}(x^*) - C_i^{\gamma_j}(x^{**})\right) + \sum_{i=1}^{s_j} \left(D_i^{\gamma_j}(x^*) - D_i^{\gamma_j}(x^{**})\right),$$

где $x^* = n - s_j$, а $x^{**} = n - s - 1$.

Найдём мощность отклонившейся коалиции s_j, при которой достигается максимальное значение каждой из f возможных реализаций γ_j этапов игры Γ_f:

$$a_1^{\gamma_j}n - a_1^{\gamma_j}s_j - a_1^{\gamma_j} + a_2^{\gamma_j}s_j > a_2^{\gamma_j}n - a_2^{\gamma_j}s_j - a_2^{\gamma_j} + b_2^{\gamma_j} - a_1^{\gamma_j}n - a_1^{\gamma_j}s_j - b_1^{\gamma_j}.$$

В таком случае, мощность отклонившейся коалиции, при которой достигается максимум выигрыша коалиции N для γ_j,

$$s_j = \left[\frac{\left(2a_1^{\gamma_j} - a_2^{\gamma_j}\right) n + \left(a_2^{\gamma_j} - b_2^{\gamma_j} + b_1^{\gamma_j} - a_1^{\gamma_j}\right)}{\left(2a_1^{\gamma_j} - 2a_2^{\gamma_j}\right)} \right].$$

Тогда значение игры для максимальной коалиции на каждом из возможных этапов игры $\gamma_j, j \in [1, f]$ можно найти, как

$$V^{\gamma_j}(N) = \left(a_1^{\gamma_j}(n - s_j) + b_1^{\gamma_j}\right)(n - s_j) + \left(a_2^{\gamma_j}(n - s_j) + b_2^{\gamma_j}\right) s_j.$$

Найдём значение для каждой из статических игр γ_j, которые разыгрываются на K_f шагах многошаговой игры Γ_f

$$V^{\gamma_j}(N) = \left(a_1^{\gamma_j} - a_2^{\gamma_j}\right) s_j^2 + \left(a_2^{\gamma_j} n + b_2^{\gamma_j} - a_1^{\gamma_j} n - b_1^{\gamma_j}\right) s_j + a_1^{\gamma_j} n^2 + b_1^{\gamma_j} n,$$

где $s_j = \left[\frac{\left(2a_1^{\gamma_j} - a_2^{\gamma_j}\right) n + \left(a_2^{\gamma_j} - b_2^{\gamma_j} + b_1^{\gamma_j} - a_1^{\gamma_j}\right)}{\left(2a_1^{\gamma_j} - 2a_2^{\gamma_j}\right)} \right]$.

Поскольку игроки симметричные, а каждая из f игр имеют одинаковые вероятности появления на каждом шаге K_f игры Γ_f, вектор Шепли представляет собой делёж математического ожидания максимального значения полученной игры Γ_f поровну между всеми n игроками.

$$Sh_i(\Gamma_f) = \sum_{j=1}^{f} \frac{\left(a_1^{\gamma_j} - a_2^{\gamma_j}\right) s_j^2 + \left(a_2^{\gamma_j} n + b_2^{\gamma_j} - a_1^{\gamma_j} n - b_1^{\gamma_j}\right) s_j + a_1^{\gamma_j} n^2 + b_1^{\gamma_j} n}{nf} K_f,$$

для $\forall i \in N$, $s_j = \left[\frac{\left(2a_1^{\gamma_j} - a_2^{\gamma_j}\right) n + \left(a_2^{\gamma_j} - b_2^{\gamma_j} + b_1^{\gamma_j} - a_1^{\gamma_j}\right)}{\left(2a_1^{\gamma_j} - 2a_2^{\gamma_j}\right)} \right]$.

В случае, если каждая из f игр имеют разные вероятности p_j появления на каждом шаге игры, математическое ожидание вектора Шепли в этом случает будет равно, соответственно

$$Sh_i\left(V^{\Gamma_f}\right)=\sum_{j=1}^{f}\frac{\left(\left(a_1^{\gamma_j}-a_2^{\gamma_j}\right)s_j^2+\left(a_2^{\gamma_j}n+b_2^{\gamma_j}-a_1^{\gamma_j}n-b_1^{\gamma_j}\right)s_j+a_1^{\gamma_j}n^2+b_1^{\gamma_j}n\right)p_j}{n}K_f,$$

для $\forall i\in N,\ s_j=\left[\frac{\left(2a_1^{\gamma_j}-a_2^{\gamma_j}\right)n+\left(a_2^{\gamma_j}-b_2^{\gamma_j}+b_1^{\gamma_j}-a_1^{\gamma_j}\right)}{\left(2a_1^{\gamma_j}-2a_2^{\gamma_j}\right)}\right].$

4.2. Вектор Шепли динамической «дилеммы заключенного n лиц».

Построим динамическую игру Γ_2, в которой то, какая игра будет сыграна на следующем этапе, не является случайным событием, а напрямую зависит от стратегий игроков. На каждом из K_f этапов разыгрывается одна из двух возможных игр γ_1 или γ_2, каждая из которых удовлетворяет статической модели «дилеммы заключённых n лиц», причём последовательность этих игр зависит от профиля стратегий, разыгранного на предыдущем этапе, следующим образом:

- Разыгрывается статическая игра γ_1, если на предыдущем этапе игры Γ_2 $(n-x) \geqslant x$, т. е. количество игроков, выбравших стратегию «кооперироваться» (C), не превосходит количество игроков, сыгравших стратегию «отклоняться» (D).
- Разыгрывается игра γ_2 на первом этапе, либо если на предыдущем этапе игры Γ_2 $(n-x) < x$, а, следовательно, количество игроков, выбравших стратегию «кооперироваться» (C), превысило количество игроков, сыгравших стратегию «отклоняться» (D).

Пусть выигрыш в игре γ_1 для каждого из игроков определяется следующим образом:

$$H_i^{\gamma_1}(x_1,\dots,x_i,\dots,x_n) = \begin{cases} C_i^{\gamma_1}(x) = a_1^{\gamma_1}x + b_1^{\gamma_1},\ \forall\ x \in (0,n]\,, \text{если } x_i = C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i^{\gamma_1}(x) = a_2^{\gamma_1}x + b_2^{\gamma_1},\ \forall\ x \in [0,n)\,, \text{если } x_i = D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

где индекс γ_1 при всех коэффициентах означает, что на данном этапе многошаговой динамической игры Γ_2 разыгрывается статическая игра γ_1.

В игре γ_2, соответственно, определим выигрыш каждого из n игроков, как

$$H_i^{\gamma_2}(x_1,\dots,x_i,\dots,x_n) = \begin{cases} C_i^{\gamma_2}(x) = a_1^{\gamma_2}x + b_1^{\gamma_2},\ \forall\ x \in (0,n]\,, \text{если } x_i = C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i^{\gamma_2}(x) = a_2^{\gamma_2}x + b_2^{\gamma_2},\ \forall\ x \in [0,n)\,, \text{если } x_i = D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Наложим дополнительные ограничения на выигрыши игроков:

- $C_i^{\gamma_1}(x) < C_i^{\gamma_2}(x)$;
- $D_i^{\gamma_1}(x) > D_i^{\gamma_2}(x)$;
- $a_1^{\gamma_1} = a_2^{\gamma_1}$;

- $a_1^{\gamma_2} = a_2^{\gamma_2}$.

В этом случае, поскольку коэффициенты при количестве кооперирующихся игроков равны, максимум суммы выигрышей всех игроков в обеих возможных играх достигается при $x = n$. Таким образом,

$$V^{\gamma_1}(N) = a_1^{\gamma_1} n^2 + b_1^{\gamma_1} n;$$

$$V^{\gamma_2}(N) = a_1^{\gamma_2} n^2 + b_1^{\gamma_2} n;$$

$$V^{\Gamma_f}(N) = (K_f - 1)\left(a_1^{\gamma_1} n^2 + b_1^{\gamma_1} n\right) + \left(a_1^{\gamma_2} n^2 + b_1^{\gamma_2} n\right).$$

В таком случае, поскольку должны выполняться аксиомы симметрии и эффективности, вектор Шепли будет равен

$$Sh_i\left(V^{\Gamma_f}\right) = (K_f - 1)\left(a_1^{\gamma_1} n + b_1^{\gamma_1}\right) + \left(a_1^{\gamma_2} n + b_1^{\gamma_2}\right).$$

Пример 3. Рассмотрим пример динамической игры, удовлетворяющей рассматриваемому варианту модели «дилеммы заключённого n лиц», где на каждом из K_f этапов будут происходить «переключения» между этими двумя играми, γ_1 и γ_2, за счёт изменения профилей стратегий для достижения максимума суммы выигрыша всех игроков.

В игре γ_1 выигрыши игроков зададим, как

$$H_i^{\gamma_1}(x_1,\ldots,x_i,\ldots,x_n) = \begin{cases} C_i^{\gamma_1}(x) = 1750x + 500,\ \forall\, x \in (0,4]\,, \text{если } x_i = C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i^{\gamma_1}(x) = 2000x + 3000,\ \forall\, x \in [0,4)\,, \text{если } x_i = D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Соответственно выигрыши каждого из игроков на этапах, где будет разыгрываться игра γ_1 составят

$$H_i^{\gamma_2}(x_1,\ldots,x_i,\ldots,x_n) = \begin{cases} C_i^{\gamma_1}(x) = 1500x,\ \forall\, x \in (0,4]\,, \text{если } x_i = C \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C, \\ D_i^{\gamma_1}(x) = 3500x + 5000,\ \forall\, x \in [0,4)\,, \text{если } x_i = D \\ \text{и известно, что } x \text{ игроков выбрали} \\ \text{стратегию } C. \end{cases}$$

Определим максимальную сумму выигрышей игроков в каждой из этих статических игр, γ_1 и γ_2, при всех возможных комбинациях кооперирующихся игроков x, где $x \in [0,n]$ (см. таблицу 12).

Таблица 12. Сумма выигрышей в динамической модели «дилеммы заключённых n лиц» при всех возможных значениях x

x	**0**	**1**	**2**	**3**	**4**
$\sum_{i \in N} H_i^{\gamma_1}$	12 000	17 250	22 000	26 250	**30 000**
$\sum_{i \in N} H_i^{\gamma_2}$	20 000	27 000	**30 000**	29 000	24 000

при этом достаточно лишь найти эти значения для каждого из возможных объёмов коалиции без идентификации того, кто входит в каждую из коалиций, поскольку игроки симметричны согласно последнему положению функции выигрыша игроков в модели «дилемма заключённого n лиц».

В этом случае, для достижения максимального суммарного выигрыша, в игре γ_2 количество кооперирующихся игроков должно быть равно двум, что не превышает количества отклонившихся игроков, которое тоже равно 2, поэтому на следующем этапе игры Γ_2 будет разыгрываться игра γ_1. В ней максимальный суммарный выигрыш достигается при $x = 4$, а, значит, количество игроков, выбирающих стратегию «кооперироваться» (C), превысит количество игроков, играющих стратегию «отклоняться» (D), которое равно 0, следовательно, на следующем после игры γ_1 этапе динамическая игра Γ_2 будет снова переключаться на γ_2.

Симметричность игроков, согласно аксиоме симметрии приводит к тому, что значение Шепли в этом случае будет равно для всех игроков, а аксиома эффективности позволяет с точностью найти вектор Шепли без использования обшей формулы, поскольку необходимо поровну разделить максимальный суммарный выигрыш, достигнутый в целом за всю игру Γ_2, на общее количество игроков n. Таким образом, вектор Шепли в этом случае будет равен

$$Sh_i\left(V^{\Gamma_f}\right) = \frac{\left[\frac{K_f}{2}\right] V^{\gamma_1} + \left(K_f - \left[\frac{K_f}{2}\right]\right) V^{\gamma_2}}{n}.$$

В нашем случае, это будет равно:

$$Sh_i\left(V^{\Gamma_f}\right) = \frac{30\ 000\left[\frac{K_f}{2}\right] + 30\ 000\left(K_f - \left[\frac{K_f}{2}\right]\right)}{4} = 7\ 500 K_f.$$

Вывод.

Таким образом, в данной дипломной работе была наиболее общим образом описана игра «дилемма заключённого n лиц». Проведён анализ нового равновесного принципа поведения игроков в условиях динамической модели игры.

Определена функция выигрышей игроков данной модели, зависящая от количества кооперирующихся игроков и, соответственно, стратегии данного игрока. Наиболее полно сформулированы отличительные свойства модели игры для каждого из игроков. Построена характеристическая функция для динамической модели. Построены таблицы сопоставления обобщённой функции выигрыша игроков, сформулированной в данной модели, с различными вариантами частичной записи выигрышей игроков, предложенными в различных источниках.

Выведено ещё одно коалиционно-устойчивое на первых $K - k^*$ шагах и индивидуально устойчивое на последних k^* шагах решение в конечношаговой повторяющейся игре, каждый шаг которой удовлетворяет рассматриваемой модели. Была доказана теорема об эффективном наказании при кооперативном поведении игроков в повторяющейся игре. Найдено оптимальное наказание в динамической модели «дилемы заключённого n лиц».

Построено C-ядро динамической модели. Определено D-подъядро Петросяна–Панкратовой [12] для многошаговой динамической игры «дилеммы заключённого n лиц». Доказана его динамическая устойчивость в условиях рассматриваемой модели.

Осуществлён поиск нормативных принципов оптимальности, таких, как, например вектор Шепли в стохастической и динамической играх, основанных на модели «дилемма заключённого n лиц».

Литература

1. Aumann R. J. Acceptable points in general cooperative n-person games // Contributions to the Theory of Games (AM-40). 1959. No 4. P. 287–324.

2. Bonacich P. et al. Cooperation and group size in the n-person prisoners' dilemma //Journal of Conflict Resolution. 1976. No 20(4). P. 687–706.

3. Carroll J. W. Iterated N-player prisoner's dilemma games // Philosophical Studies. 1988. No 53(3). P. 411–415.

4. Essam E. L. S., Elshobaky E. M., Soliman K. M. Two population three-player prisoner's dilemma game //Applied Mathematics and Computation. 2016. No 277. P. 44–53.

5. Fox J., Guyer M. «Public» choice and cooperation in n-person prisoner's dilemma //Journal of Conflict Resolution. 1978. No 22(3). P. 469–481.

6. Goehring D. J., Kahan J. P. The uniform n-person prisoner's dilemma game: construction and test of an index of cooperation //Journal of Conflict Resolution. 1976. No 20(1). P. 111-128.

7. Hamburger H. N-person prisoner's dilemma // Journal of Mathematical Sociology. 1973. No 3. P. 27–48.

8. Kannai Y. The core and balancedness //Handbook of game theory with economic applications. 1992. Vol. 1. P. 355-395.

9. Moulin H. Fair division and collective welfare // MIT press. 2004. Vol. 1.

10. Petrosjan L. A., Grauer L. V. Strong Nash equilibrium in multistage games // International Game Theory Review. 2002. No 4(2). P. 255—264.

11. Petrosjan L. A., Grauer L. V. Multistage games // Journal of applied mathematics and mechanics. 2004. Vol. 68. No 4. P. 597–605.

12. Petrosjan L. A., Pankratova Y. B., New characteristic function for multistage dynamic games // Vestnik of Saint Petesrburg University, Applied mathematics. Computer Science. Control Processes. 2018. No 14(4). P. 316–324.

13. Petrosyan L. et al. Strong Strategic Support of Cooperation in Multistage Games //International Game Theory Review (IGTR). 2019. No 21(1). P. 1–12.

14. Straffin P. D. Game Theory and Strategy. Washington: The Mathematical Association of America, 1993. 244 p.

15. Szilagyi M. N. An investigation of N-person Prisoners' Dilemmas //Complex Systems. 2003. No 14(2). P. 155–174.

16. Takezawa M., Price M. E. Revisiting "The Evolution of Reciprocity in Sizable Groups": Continuous reciprocity in the repeated n-person prisoner's dilemma //Journal of theoretical biology. 2010. Vol. 264. No 2. P. 188–196.

17. Van Lange P. A. M., Liebrand W. B. G., Kuhlman D. M. Causal attribution of choice behavior in three N-person prisoner's dilemmas //Journal of Experimental Social Psychology. 1990. Vol. 26. No 1. P. 34–48.

18. Winter E. The shapley value //Handbook of game theory with economic applications. 2002. No 3. P. 2025–2054.

19. Yao X. Evolutionary stability in the n-person iterated prisoner's dilemma // BioSystems. 1996. Vol. 37. No 3. P. 189–197.

20. Петросян Леон Аганесович, Громова Екатерина Викторовна Сильно динамически устойчивое кооперативное решение в одной дифференциальной игре управления вредными выбросами // Управление большими системами: Сборник трудов. 2015. Т. 55. С. 140—159.