

П.В.Голубцов, В.А.Любецкий

ИНФОРМАЦИОННАЯ СТРУКТУРА ДИНАМИЧЕСКИХ ИГР И ОБЪЕМ ДОСТУПНОЙ ИГРОКАМ ИНФОРМАЦИИ

Abstract. *Here the dynamic game with discrete time is generalized to a stochastic environment, in order to examine the implications of incomplete and asymmetric information. In this game the next state of a system and players' payoffs depend not only on its current state and players controls, but also on Markov stochastic elements. At each step of the game, the players both know current state of the system, and also have some (generally incomplete or delayed, and even asymmetric) knowledge of the current values of stochastic elements. The knowledge structure of each specific game version is held in common by the competitors. In the dynamic game each player sets its strategy, with the objective of maximizing the expected discounted sum of seasonal payoffs, and conditional on the extent of its current knowledge, and of the anticipated policy of its competitor. The implications of alternative knowledge structures are explored, through dynamic programming and simulation. Both information structures and various game parameters are varied continuously, to explore their interplay. Particular focus is on demonstrating the often unexpected, and sometimes counter-intuitive, effects that knowledge enrichment may have, in these incomplete information games.*

Ключевые слова: динамическая игра, оптимальная стратегия, равновесие по Нэшу, информационная структура игры, алгоритм, зависимость от информации.

1. Постановка задачи

Можно привести много примеров сложных информационных взаимодействий в различных системах – экономических, социальных, политических, технических, биологических. Последние, например, подробно обсуждаются в [1]. Информационное взаимодействие в системах может описываться разными математическими средствами и, в частности, такими, как стохастические динамические игры с равновесием по Нэшу. Выбор игроком на каждом шаге игры управляющего решения основан, прежде всего, на информации, доступной ему на тот момент. Характер и объем этой информации для каждого из участников игры называется *информационной структурой игры*.

Авторы предложили новые алгоритмы для нахождения равновесия по Нэшу. *На примере* хорошо известной задачи морского

рыболовства [2,3] авторы приводят имеющие общее значение и, как кажется, неожиданные выводы о зависимости результатов игры от доступной игрокам информации.

По необходимости мы начнем с краткого описания самой этой частной задачи, а затем перейдем к упомянутым выводам. В общей форме они будут представлены в другой публикации авторов.

Общий ресурс R (в этом примере – рыбы) в соответствии со значениями θ^α и θ^β случайного фактора θ делится на две части R^α и R^β , доступные для лова соответственно игроками α и β . После лова каждым из игроков в пределах его части ресурса остатки S^α и S^β рыбы соединяются в общий остаток $S = S^\alpha + S^\beta$ для воспроизводства из него на следующий год нового значения общего ресурса $R_+ = F(S)$. Решения игроков состоят в определении размера той части ресурса, которая не будет ими выловлена (в текущем году), т.е. $S^\alpha = p^\alpha \cdot R^\alpha$, $S^\beta = p^\beta \cdot R^\beta$. Тогда уловы игроков в текущем году соответственно равны $H^\alpha = (1 - p^\alpha) \cdot R^\alpha$ и $H^\beta = (1 - p^\beta) \cdot R^\beta$. В нашей модели учитывается зависимость себестоимости лова от объема доступного игроку ресурса а именно себестоимость лова у игрока α при данном объеме x ресурса в его части обратно пропорциональна этому объему и равна c^α/x , где c^α - коэффициент стоимости лова. Этот коэффициент можно представить себе как некоторую интегральную оценку качества и характеристики обслуживания рыболовного флота игрока α , если не указано иное, то принимается $c = 0.2$. Таким образом, себестоимость лова (за один год) игрока α равна $cost_\alpha = -c_\alpha \cdot \log(p^\alpha)$. Тогда его выигрыш (годовой доход) равен $v^\alpha = H^\alpha - cost_\alpha$. При величине H^α подразумевается коэффициент, переводящий объем улова в стоимость (у нас он равен 1). Аналогично для второго игрока β . Используется функция воспроизводства ресурса, которая учитывает возможность его полного уничтожения, если его объем упадет ниже некоторого критического уровня, т.е. должно выполняться $F(S) < S$ при достаточно малых S . После некоторых обоснований естественно выбрать $F(S) = 0.6 \cdot S + 1.8 \cdot S^2 - 1.6 \cdot S^3$. Кроме того, у нас значения случайного фактора θ равны $\theta_1 = 0.1$ или $\theta_2 = 0.9$, они независимы по времени, одинаково распределены и наступают с равными вероятностями 0.5; все коэффициенты, понижающие со временем годовые доходы, равны 0.9 (для обоих игроков); коэффициенты кооперации игроков составляют единичную матрицу (для чистого соперничества) и равны все 0.5 (для полной кооперации); выбирался достаточно длинный промежуток усреднения, равный 2000 лет. Чтобы плавно менять точность поступающей к игрокам информации о случайном параметре θ , использовались так называемые неточные измерения и менялся параметр точности измере-

ния π , который определяет матрицу M переходного распределения наблюдаемой величины ξ для различных значений θ , следующим образом: по главной и вспомогательной диагоналям находится соответственно $(1 + \pi) / 2$ и $(1 - \pi) / 2$. Таким образом, если $\pi = 1$ (максимальная точность), то M - единичная матрица, что отвечает тождественному измерению. Если $\pi = 0$ (минимальная точность), то M - матрица из чисел 0.5, что отвечает независимости результатов измерения от состояния θ . Детали, относящиеся к этой модели, в части задачи рыболовства можно найти в [2,3], а в математической части – в [4]. Но все эти подробности мало существенны для выводов, которые составляют основное содержание настоящей статьи и приводятся ниже в параграфе 2.

Нами были выделены и рассматривались в компьютерных экспериментах, в частности, *следующие информационные структуры игры* (в скобках приводится название структуры и ее обозначение).

1. Игроки в равной мере имеют в каждый текущий момент t времени (как и во все предыдущие моменты) реализацию случайного фактора θ («текущая информация», обозначаемая Cur).

2. Игроки в равной мере знают только само вероятностное распределение случайного дискретного фактора θ («минимальная информация», обозначаемая Min).

3. Игроки имеют асимметричную информацию: первый имеет текущую информацию по пункту 1, а второй – лишь минимальную по пункту 2 («асимметричная информация», обозначаемая Cur-Min). Здесь определяется понятие точности поступающей игрокам информации.

4. Игроки имеют в равной мере неполную информацию: они получают реализацию случайной величины ξ измерения случайного фактора θ («неполная информация, полученная из измерений», обозначаемая Meas).

5. Игроки имеют асимметричную информацию: первый получает неполную информацию по пункту 4, а второй – лишь минимальную по пункту 2 («асимметричная неполная информация», обозначаемая Meas-Min).

6. Игроки ведут себя кооперативно в условиях одной из симметричных информационных структур 1, 2 или 4 (так как при кооперативном поведении представляется естественным, что игроки полностью обмениваются всей доступной им информацией).

Структуры из пунктов 1-5 соответствуют соперничеству игроков, а структуры из пункта 6 соответствуют их сотрудничеству.

2. Результаты компьютерного моделирования

1. *Влияние коэффициента стоимости лова для различных информационных структур.* Для простоты представления наших результатов будем считать здесь, что $c_\alpha = c_\beta$, обозначая далее этот общий коэффициент стоимости лова как c . Графики на рис.1 показывают *средние* (по всем годам лова) *доходы* в зависимости от коэффициента стоимости c лова для указанных там пяти информационных структур игры. Наиболее интересным в случаях, когда игроки соперничают, является резкое сокращение среднего дохода при снижении коэффициента стоимости c лова. Это связано с тем, что при низком значении c конкуренция носит весьма острый характер, что приводит к резкому снижению объема ресурса. Такой эффект не наблюдается, когда игроки сотрудничают и, таким образом, оказываются в состоянии поддерживать объем рыбы на достаточно высоком уровне. Кроме того, на рис. 1 видно, что информационное преимущество в случае информационно-асимметричной игры чрезвычайно выгодно первому игроку (Sig-Min 1). Более того, ему не выгодно делиться информацией с соперником и, таким образом, перейти к игре с полной симметричной информацией (Sig). Что касается второго игрока в асимметричном случае (Sig-Min 2), то он, как правило, предпочел бы получить дополнительную информацию, но не для низкого значения c коэффициента стоимости лова (<0.3), когда односторонний недостаток информации более выгоден даже для него, чем полная симметричная информация (Sig). Заметим, что при достаточно низких значениях коэффициента стоимости c лова (<0.35) симметричная минимальная информация (Min) становится предпочтительнее полной информации (Sig). Это, по-видимому, связано с тем, что при недостатке информации возрастает риск случайного уничтожения ресурса и, как следствие этого, поведение игроков становится более «осторожным». Как отмечалось, соперничество становится особенно разрушительным при низком коэффициенте стоимости c лова, когда можно легко уничтожить ресурс. Однако если игроки сотрудничают, то их доход становится заметно выше, особенно когда себестоимость лова мала. При этом они в состоянии поддерживать объем возобновляемого ресурса на довольно высоком уровне.

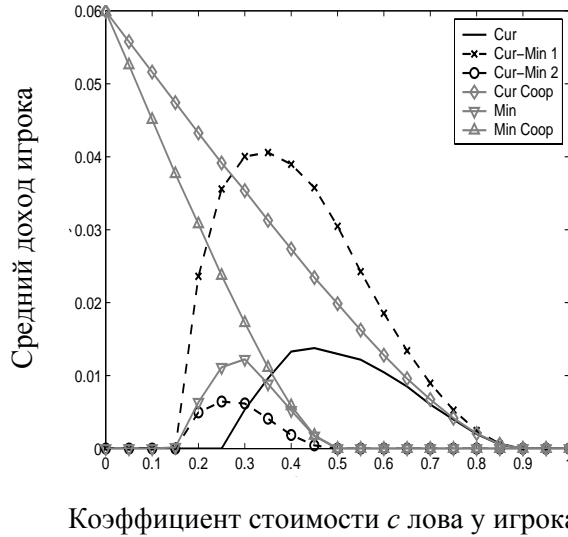


Рис. 1. Используются сокращения: Cur – оба игрока имеют одну и ту же текущую информацию (включая все предшествующие моменты времени), и показан выигрыш 1-го игрока, 2-й игрок имеет такой же выигрыш; Cur-Min i – игроки имеют асимметричную информацию: игрок 1 обладает текущей информацией, а игрок 2 минимальной ($i=1$ или $i=2$, показаны выигрыши обоих игроков); Cur-Coop – совместная текущая информация и кооперативное поведение игроков; Min – оба игрока имеют в равной мере минимальную информацию; Min-Coop – совместная минимальная информация и кооперативное поведение игроков.

2. *Информация доступна игрокам с некоторой точностью.* В этом наборе численных экспериментов (рис. 2) информация о текущем значении θ получается из неточных измерений случайного параметра θ . Точность измерения (точность информации) является параметром, изменяющимся от 0 (отсутствие информации) до 1 (полная информация). Когда игроки сотрудничают (Meas-Coop), ситуация вполне естественная: чем точнее информация, тем выше доход. Суммарный доход в такой кооперативной игре всегда выше, чем суммарные доходы в некооперативных играх. Как видно на рис. 2, в некооперативных играх с симметричной (Meas) и асимметричной (Meas-Min) информационными структурами, повышение точности информации до определенного уровня выгодно обоим игрокам и даже игроку, которому вовсе не доступны результаты измерений (Meas-Min 2). Однако дальнейшее повышение точности информации приводит к значительному снижению доходов. По-видимому, это объясняется тем, что

стратегии игроков становятся более агрессивными. При низкой точности измерений ситуация, когда второй игрок также имеет доступ к результатам измерений (Meas), более выгодна для него, чем одностороннее полное отсутствие у него этой информации (Meas-Min 2). Однако начиная с определенного уровня точности измерений (>0.38), такая дополнительная информация становится невыгодной для него.

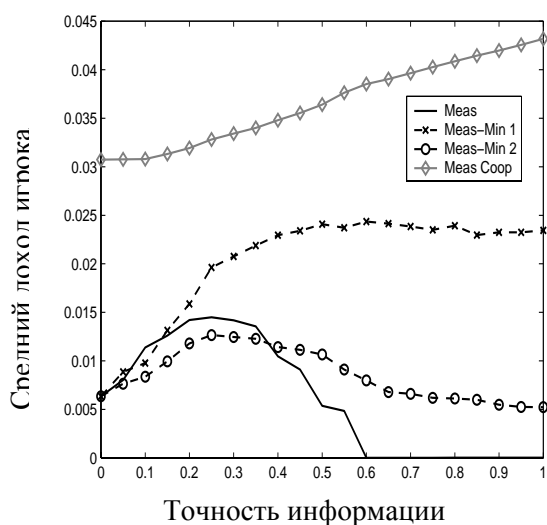


Рис. 2. Влияние точности поступающей информации. Здесь Meas означает, что оба игрока имеют одинаковую неполную информацию; Meas-Min i – первый игрок получает неполную информацию, а второй лишь минимальную, $i=1$ или $i=2$; Meas-Coop – кооперация в случае одинаковой неполной информации у игроков.

3. *Информационная асимметрия при наличии асимметрии в природных условиях.* В этом эксперименте (рис. 3) природные условия несколько благоприятствуют второму игроку. А именно θ принимает значения 0.1 и 0.8 с равными вероятностями (так что доступный ресурс первого игрока составляет 0.1 или 0.8 от общего ресурса, в то время как второй игрок получает, соответственно 0.9 или 0.2 от всего ресурса). Таким образом, в среднем ресурс у первого игрока ниже, чем у второго. Как и следовало ожидать, если игроки обладают одинаковой информацией, доход второго игрока (Meas 2) всегда будет выше дохода первого игрока (Meas 1). Однако когда первый игрок (Meas-Min 1) обладает серьезным

информационным преимуществом и точность информации >0.3 , то это перевесит природное преимущество второго игрока (Meas-Min 2). С другой стороны, сумма доходов игроков будет максимальной при условии кооперации, когда игроки обмениваются информацией и максимизируют их суммарный доход. Но и в этом случае вследствие природной асимметрии их непосредственные доходы не будут равны (Coop 1 и Coop 2). Такого рода сотрудничество можно рассматривать как «проявление доброй воли». Конечно, игроки могут договориться о перераспределении их суммарного дохода, что приведет к «побочным платежам» одного игрока другому. Случай равного распределения доходов также показан на рис. 3 (Coop).

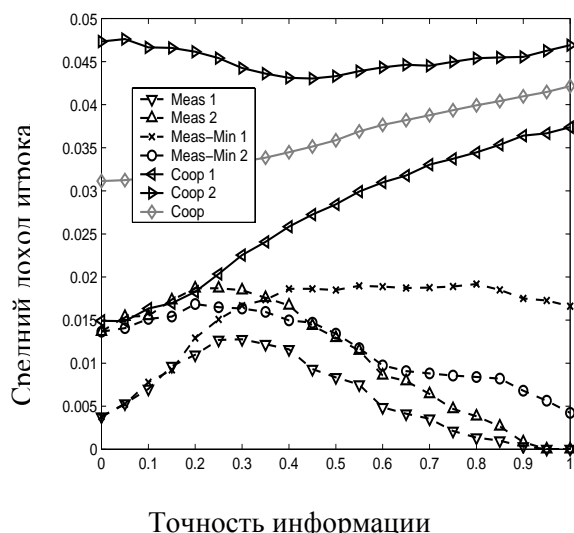


Рис. 3. Влияние точности поступающей информации при асимметричных природных условиях. Здесь Meas i означает, что оба игрока получают одинаковую неполную информацию; Meas-Min i – первый игрок получает неполную информацию, а второй лишь минимальную; Coop i – кооперация при наличии одинаковой неполной информации (кроме реальных доходов игроков в Coop 1 и Coop 2 указан и их усредненный доход Coop); $i=1$ или $i=2$.

4. *Влияние самой природной изменчивости.* Теперь (рис. 4) мы меняем степень природной изменчивости $\langle \theta_1, \theta_2 \rangle$ коэффициента θ . Точнее, θ принимает одно из двух значений: θ_1 или $\theta_2 = 1 - \theta_1$, где θ_1 может быть любым числом от 0 до 0.5. При $\theta_1=0.5$ изменчи-

вость отсутствует ($\theta=0.5$ и постоянно), при $\theta_i=0$ изменчивость максимальна (θ принимает значения 0 или 1). Возрастание дохода в кооперативной игре с полной информацией (Cur-Coop), а также при возрастании изменчивости θ (и убывании θ_i) вполне естественно. Действительно, при высокой вариабельности θ практически весь ресурс направляется в один из двух потоков, и это приводит к снижению средней себестоимости лова. Поскольку эта кооперативная игра полностью симметрична, средние годовые доходы игроков одинаковы. Похоже, что возрастание дохода в некооперативной игре при уменьшении θ_i от 0.5 до 0.4 может иметь такое же объяснение. Однако при малых θ_i роль соперничества (особенно при полной информации Cur) становится доминирующей. Действительно, если весь ресурс попадает в один поток, то соответствующий флот может практически полностью выловить его при сравнительно низких затратах.

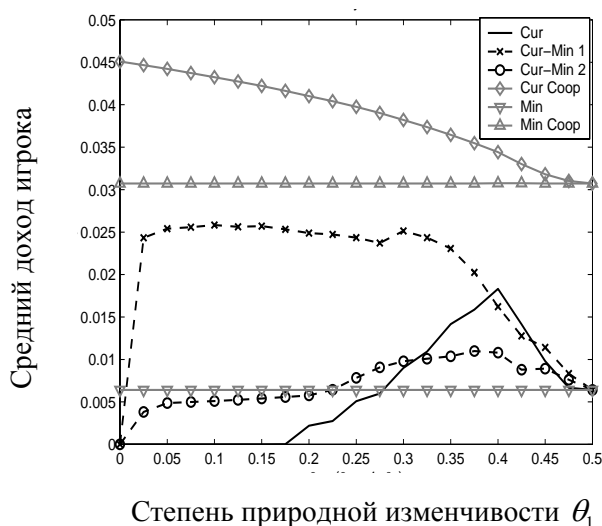


Рис. 4. Cur означает текущую информацию; Cur-Min i – асимметричную информацию, когда игрок 1 обладает текущей информацией, а игрок 2 минимальной, $i=1$ или $i=2$; Cur-Coop – кооперацию с текущей информацией; Min – минимальную информацию; Min-Coop – кооперацию с минимальной информацией.

5. *Переход игроков от соперничества к сотрудничеству.* Наконец, опишем простой способ включения «кооперации» в модель игры.

На рис. 5 представлена зависимость среднего дохода игрока от «степени сотрудничества» игроков для случаев текущей или минимальной информации. Нулевая степень кооперации d игроков соответствует поведению чистого соперничества, когда каждый игрок максимизирует только свой доход, в то время как степень кооперации 1 означает, что оба игрока максимизируют полный доход. При промежуточных степенях кооперации каждый игрок максимизирует выпуклую комбинацию своего дохода и дохода соперника. Точнее, коэффициенты кооперации определяются через степень кооперации, обозначаемую d , следующим образом: по главной и вспомогательной диагоналям находятся соответственно значения $1 - (d/2)$ и $d/2$. Таким образом, когда d меняется от 0 до 1, коэффициенты кооперации меняются от «отсутствия кооперации» до «полной кооперации». На рис. 5 хорошо видно, что дополнительная информация выгодна для игроков только в том случае, когда уровень их кооперации достаточно высок. При низком уровне их кооперации и особенно в ситуации соперничества дополнительная информация приводит к критическому сокращению ресурса и нулевому доходу.

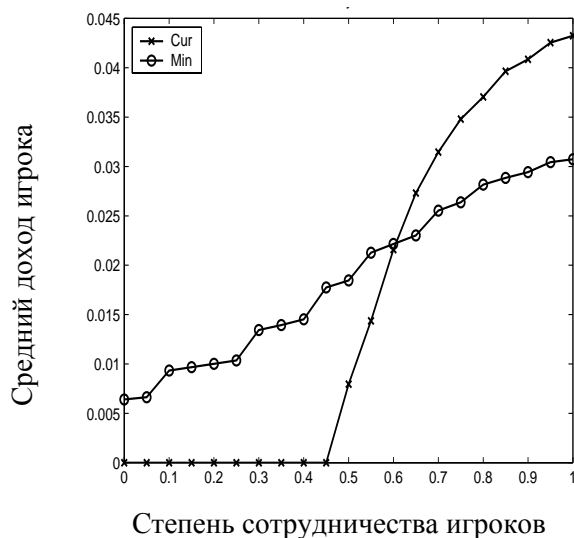


Рис. 5. Влияние степени сотрудничества при текущей и минимальной информации. Cur – текущая, Min – минимальная информации.

3. Описание идеи алгоритмов. Здесь используются некоторые стандартные обозначения из теории дискретных марковских процессов управления (см. например, [4]). Для случая Сиг авторы предлагают по аналогии с динамическим программированием ввести функцию

$$V^\# = V^\alpha_t(R_t, v_t, \langle P^\alpha_t, Q^\alpha_{t+1} \rangle, \langle P^\beta_t, Q^\beta_{t+1} \rangle),$$

описывающую выигрыш игрока α , который соответствует произвольным стратегиям P^α_t и P^β_t в момент t и оптимальным «хвостовым» последовательностям Q^α_t и Q^β_t . А также аналогичную функцию для игрока β . Тогда оптимальные стратегии в момент t могут быть получены путем нахождения точек равновесия по Нэшу при всех возможных значениях R_t и v_t для функций $V^\#$ игроков α и β относительно переменных $\langle p^\alpha_t, p^\beta_t \rangle$. Таким образом, равновесные по Нэшу стратегии могут быть построены рекурсивно.

В случае неточной информации авторы предлагают следующую поточечную конструкцию для получения стратегий, обеспечивающих равновесие по Нэшу в функциональных пространствах. Иными словами, авторы предлагают способ перехода от задачи построения равновесия по Нэшу в функциональном пространстве к такой же задаче в конечномерном арифметическом пространстве. Определим

$$W^\alpha_t = v^\alpha_t(R, v, P^\alpha(R, \xi), P^\beta(R, \xi)) + \gamma_\alpha U^{\alpha}_{t+1}(\rho_t(R, v, P^\alpha(R, \xi), P^\beta(R, \xi)))$$

и аналогичную функцию для игрока β . Предположим, что для некоторого фиксированного ξ стратегии $P^{\hat{\alpha}}$ и $P^{\hat{\beta}}$ обеспечивают равновесие по Нэшу для функций

$$V^{-\alpha}(R, \xi, P^\alpha, P^\beta) = E_{v|\xi} W^\alpha_t(R, v, \xi, P^\alpha, P^\beta)$$

и аналогично для β . Это означает, что для всех R и ξ и произвольных стратегий P^α и P^β выполняется

$$V^{-\alpha}(R, \xi, P^{\hat{\alpha}}, P^{\hat{\beta}}) \geq V^{-\alpha}(R, \xi, P^\alpha, P^\beta)$$

и аналогично для β . Отсюда получаем

$$U^{-\alpha}(R, P^{\hat{\alpha}}, P^{\hat{\beta}}) = E_\xi V^{-\alpha}(R, \xi, P^{\hat{\alpha}}, P^{\hat{\beta}}) \geq E_\xi V^{-\alpha}(R, \xi, P^\alpha, P^\beta) = U^{-\alpha}(R, P^\alpha, P^{\hat{\beta}})$$

и аналогично для β . Таким образом, исходная задача построения стратегий, обеспечивающих равновесие по Нэшу, сводится к аналогичной задаче для функций $V^{-\alpha}$ и $V^{-\beta}$, а для этой второй задачи стратегии, обеспечивающие равновесие по Нэшу, вычисляются «поточечно». Подробно наши алгоритмы изложены в [5].

ЛИТЕРАТУРА

1. Кузнецов Н.А., Любецкий В.А., Чернавский А.В. Информационные взаимодействия, 1: допсихический уровень // Информационные процессы. 2003. №1 (Электронная версия: <http://www.jip.ru/>).
2. Clark C.W. (1980). Restricted access to common-property fishery resources: a game-theoretic analysis // Dynamic Optimization and Mathematical Economics. New York: Plenum. P. 117-132.
3. Levhari, D., Mirman L.J. The great fish war: an example using a dynamic Cournot-Nash solution // Bell Journal of Economics, 1980. V. 11, P. 322-344.
4. Hernandez-Lerma O., Lasserre J.B. Discrete-Time Markov Control Processes. Basic Optimality Criteria. New York: Springer, 1996.
5. Голубцов П.В., Любецкий В.А. (2003). Стохастические динамические игры с информацией различного типа // Проблемы передачи информации. РАН, том 39, вып. 3, С. 40-71.