УДК 621.391.037.372

В. В. ПОПОВСКИЙ, д-р техн. наук, В. Ф. ОЛЕЙНИК, канд. техн. наук

ИСПОЛЬЗОВАНИЕ ТЕОРИИ ИГР ПРИ АНАЛИЗЕ ГРУППОВЫХ ВЗАИМОДЕЙСТВИЙ РАДИОЭЛЕКТРОННЫХ СРЕДСТВ

Многие задачи радиотехники и связи адекватно представляются в виде, который дает возможность использовать игровые методы решения, обладающие большой общностью [1; 2]. Охватывая различные стратегии (антагонистические, кооперативные или безразличные), эти методы позволяют решать практически весь цикл задач, отражающих динамику поведения. Это могут быть задачи по помехозащите, по обеспечению электромагнитной совместимости в группировке радиоэлектронных средств (РЭС) или доступа к общему ресурсу, по маршрутизации и др.

Рассмотрение дифференциальной игры двух лиц с противоположными интересами приводит к решению Нэша [2 – 4]. Однако условия равновесия Нэша существенно изменяются с увеличением числа участников игры. Особенный интерес при этом возникает к эффективности стратегий при различном уровне взаимной информированности игроков. Так, показано [5; 6], что увеличение объема информации приводит к увеличению объема затрат обеих конфликтующих сторон. В работе [7] показано, что увеличение объема информации для одной из сторон повышает ее шансы и наносит вред противоположной стороне.

Постановка задачи. Рассмотрим группировку РЭС с числом участников М. Это может быть, например, группировка сотовых или спутниковых РЭС или множество маршрутизаторов в телекоммуникационной сети, обеспечивающих обменность информацией. Любая такая современная группировка может успешно и устойчиво функционировать, если параметрами ее элементов централизованно или децентрализованно управляют. Управление может выполняться как состоянием всей или частью системы, так и наблюдением (базисом наблюдения). При первом варианте управление этого состояния $x(t)$ в непрерывной форме описывается дифференциальным уравнением [1]

$$\frac{d\vec{x}(t)}{dt} = F(t)\vec{x}(t) + \sum_j B_j(t)\vec{u}_j(t) + G\vec{w}(t), \tag{1}$$

где $F(t), B(t), G(t)$ – матрицы состояния, управления и возбуждения; $\vec{u}(t)$ – вектор управляемых параметров; $\vec{w}(t)$ – вектор виртуальных шумов возбуждения с уровнем $\mathrm{E}\left[w(k)w^T(k)\right] = V_w$.

Для дискретной системы аналог уравнения (1) имеет вид

$$\vec{x}(k+1) = \Phi(k)\vec{x}(k) + \sum_j B_j(k)\vec{u}_j(k) + G(k)\vec{w}(k). \tag{2}$$

Для принятия решения об управлении на следующем шаге каждый из М участников должен наблюдать (измерять) все доступные параметры состояния $\vec{x}(k)$. Уравнение наблюдения имеет вид

$$\vec{y}(k) = H(k)\vec{x}(k) + \vec{v}(k), \tag{3}$$

где $H(k)$ – масштабирующая матрица; $\vec{v}(k)$ – шум наблюдения с уровнем V_v.

Рассмотрим более подробно, каким образом параметры каждого из М игроков влияют на его индивидуальные и групповые показатели, как приобретаемая в процессе функционирования информация I_{xy} влияет на качество этого функционирования.

Решение игровой задачи. Исходя из (1) и (2) можно заключить, что каждый из участников может влиять на вектор $\vec{x}(t)$ через соответствующие матричные элементы F, B или G. Вместе с тем F и G на практике выбираются постоянными, ибо они задают системные свойства: скорость изменения состояния F_{ij} и уровень случайных влияний G_{ij}. Элементы матрицы $B_{ij}(k)$ являются именно теми элементами, которые зависят от целевой функции и сложившейся на данный момент ситуации, задаваемой уравнением наблюдения (3).

Наличие той или иной информации для i-го игрока приводит, очевидно, к возрастанию уровня его соответствующего управления u_i, т. е. $B_i(k)B_i^T(k)$ может быть показателем степени его информированности. Более того, соотношение $B_i(k)B_i^T(k) > B_j(k)B_j^T(k)$ свидетельствует о том, что информированность i-го игрока выше, чем j-го. При этом i-й игрок обладает большей маневренностью, чем j-й, и занимает доминирующее положение.

Непосредственное использование уравнения (3) может отображать ситуацию и в отсутствие шумов наблюдения, при $\vec{v}(k) = 0$. Если же $\vec{v}(k) \neq 0$, необходима статистическая обработка $\vec{y}(k)$, которую следует осуществлять с использованием рекурсивных процедур, ибо выборочные интервальные оценки, например $\hat{x} = N^{-1}\sum_k^N x(k)$, для использования в динамических системах имеют ограниченную применимость. Оптимальной в смысле минимума среднего квадрата отклонения критерия

$$J = \min E(\vec{x}(k) - \hat{\vec{x}}(k))^2 \tag{4}$$

является процедура Калмана–Бьюси [1], которая имеет вид

$$\hat{x}(k) = \hat{x}(k/k-1) + K(k)\left[y(k) - H(k)x(k/k-1)\right], \tag{5а}$$

$$\hat{x}(k+1/k) = \Phi(k,k-1)\hat{x}(k) + \sum_j B_j u_j(k). \tag{5б}$$

Здесь $K(k) = V_{\tilde{x}}(k/k-1)H^T(HV_{\tilde{x}}(k/k-1)H^T(t)V_v(k))^{-1}$; $V_{\tilde{x}}(k+1/k) = \Phi(k)[I - K(k)H] \times \times V_{\tilde{x}}(k/k-1)\Phi^T(k) + V_W(k)$ – апостериорная дисперсия ошибки прогноза; $V_{\tilde{x}}(k)$ – апостериорная дисперсия ошибки оценки; $V_{\tilde{x}}(k) = \left[I - K(k)H(k)\right]V_{\tilde{x}}(k/k-1)$.

Ошибка оценки $V_{\tilde{x}}(k)$ и ошибка прогноза на один шаг $V_{\tilde{x}}(k+1/k)$ соответственно определяются соотношениями

$$V_{\tilde{x}}(k) = \mathrm{E}[x(k) - \hat{x}(k)][x(k) - \hat{x}(k)]^T,$$

$$V_{\tilde{x}}(k+1/k) = \mathrm{E}[x(k+1) - \hat{x}(k+1/k)][x(k+1) - \hat{x}(k+1/k)]^T.$$

Характерно, что если прогноз $\hat{x}(k+1/k)$ для различных участников игры определяется на одних и тех же шагах дискретизации Δt, то ошибки этого прогноза зависят от динамики того или иного игрока. Вместе с тем из определения значения прогноза (5б) следует, что его погрешности определяются значениями $\Phi(k,k-1) = \exp\{-\Delta t/\tau_{\text{кор}}\}$, где $\Delta t = (k+1) - k$ – шаг дискретизации; $\tau_{\text{кор}}$ – интервал корреляции процесса $x(k)$. Очевидно, для менее инерционных участников, для которых $\tau_{\text{кор}\,i} < \tau_{\text{кор}\,j}$, при одном и том же шаге Δt ошибка прогноза будет больше, чем для более инерционных. Соответственно ухудшена будет и оценка (5а).

 ISSN 0485-8972. Радиотехника. 2003. Вып. 133

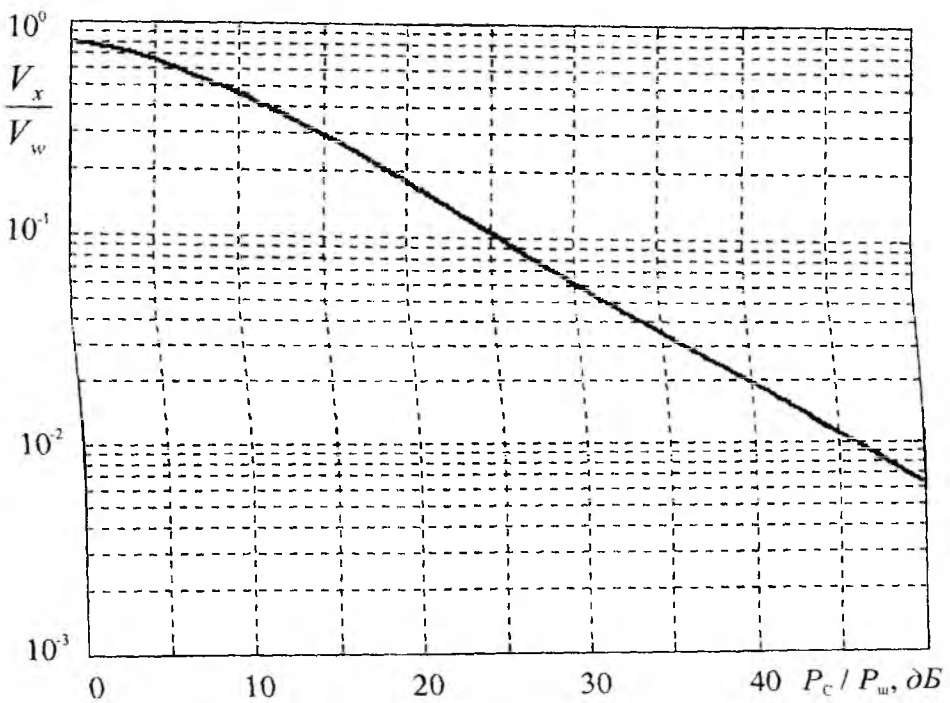
Рассмотрим другие зависимости, связанные с соотношением уровней сигналов P_c и помех $P_п \overset{\Delta}{=} V_v(k)$. Уровень сигнала P_c при нулевом среднем $x(k)$ определяется дисперсией самого процесса $x(k)$ или значением V_w.

При известном отношении уровней сигнал-шум в канале наблюдения отношение апостериорной и априорной дисперсий легко вычислить по формуле [1]

$$V_{\tilde{x}}(k)/V_w(k) = 2/(1+(1+P_c/P_ш)^{1/2})^{-1}. \tag{6}$$

Из (6) видно, что с увеличением $P_c/P_ш$ апостериорная дисперсия пропорционально уменьшается (рисунок), соответственно точность оценки возрастает.

С повышением точности оценки, или же соответственно с уменьшением $V_{\tilde{x}}(k)$, обеспечивается минимум критерия J_i определяемого из (4) при меньших его значениях, что улучшает качество стратегии i-го игрока.



Найдем закон оптимального управления $u(k)$. Из теории оптимального управления известно, что значение $u_{\text{opt}}(k)$ можно отыскать, если минимизировать гамильтониан вдоль оптимальной траектории [1]. Оптимальность уравнения для i-го игрока задается критерием

$$J_i = E\left\{\sum_{k=0}^{N-1}\left[\vec{x}^T(k)P_i\vec{x}(k) + \sum_{j=1}^{M}\vec{u}_j^T(k)Q_{ij}\vec{u}_j(k)\right] + \vec{x}^T(N)P_i x(N)\right\}. \tag{7}$$

Слагаемые функционала (7) имеют конкретную физическую интерпретацию. Последнее слагаемое обеспечивают минимум разбросов состояния $\vec{x}$ относительно оптимальной траектории на конечном участке траектории $x(N)$. Первое слагаемое в квадратных скобках обеспечивает минимизацию разбросов состояния на всем протяжении траектории движения системы. Сумма слагаемых, содержащих Q_{ij}, минимизирует затраты на управление, что особенно важно, когда для управления расходуется горючее или энергия.

Пусть для каждого из М игроков получено решение, минимизирующее показатель (7), т.е. известен набор стратегий $\{g_1^*, ..., g_M^*\}$, называемых решением Нэша [3; 4]:

$$J_i(g_1^*, ..., g_M^*) \le J_j, \forall g_j. \tag{8}$$

С учетом критерия (7) траекторию управляемой системы (1) можно проанализировать в обратном времени, начиная с последнего шага N_1 где $L_i(N) = P_i$ [6]:

$$\begin{gathered} L_i(k) = P_i + \Phi^T[(I + \sum_j B_j B_j^T L_j(k+1))^{-1}]^T [L_i(k+1) + \\ + \sum_j L_j(k+1) B_j Q_{ij} B_j^T L_j(k+1)][I + \sum_j B_j B_j^T L_j(k+1)]^{-1} \Phi \end{gathered} \tag{9}$$

Можно показать [1; 2], что оптимальное управление

$$u^*(k) = D(k)\hat{\hat{x}}(k), \tag{10}$$

удовлетворяющее условию (8), существует и единственно, если для каждого шага k существует обратное значение элемента в квадратных скобках, входящего в выражение

$$D(k) = -B_i^T L_i(k+1)\left[I + \sum_j B_j(k) B_j^T(k) L_j(k+1)\right]^{-1} F(k). \tag{11}$$

Затраты i-го игрока на k-м шаге, определяемые как $J_i(k)$, если суммирование производится от k до $N-1$, представляется так:

$$J_i(k) = E\left[\hat{\hat{x}}^T(k) L_i(k)\hat{\hat{x}}(k)\right] + T_i(k),$$

где $T_i(k) = Tr\{P_i V_{\tilde{x}}(k) + \Phi^T(k) L_i(k+1)\Phi(k) V_{\tilde{x}}(k) + L_i(k+1)(V_w(k) - V_{\tilde{x}}(k+1)\} + T_i(k+1)$, (12)

$$T_i(N) = Tr\{P_i V_{\tilde{x}}(N)\}. \tag{13}$$

Из рассмотрения (11) следует, что значения $D_i(k)$ не зависят от результатов наблюдений (3). Это может быть лишь в том случае, когда именно эти значения $D_i(k)$ и должны быть задействованы в данной игре Нэша при условии, что все игроки обладают полной информацией о состоянии и используют стратегии замкнутого контура управления без задержек. Таким образом, затраты $J_i(k)$ полностью определяются соотношением величин V_w и V_x, соответственно характеризующих априорные дисперсии полезного оцениваемого сигнала и инерционного i-го игрока, которая характеризуется значением $\Phi_i(k, k-1)$ заключенной в значениях матрицы состояния $\Phi(k)$.

Условия существования обратного значения в выражении (10) достигаются, когда при $B_j = b_j B$

$$I + BB^T \Sigma_j b_j L_j(k+1) \ge 1. \tag{14}$$

Можно показать [6], что если $X > Y \ge 0$, то $X^{-1} > X^{-1} Y X^{-1}$. При этом, если величина $B_j B_j^T$ не вырождена, из (9) получим неравенство

$$\|\Phi(k)\|^2 \sigma_{\max}(A) < \lambda_{\min}(BB^T + BB^T \sum_j b_j^2 P_j BB^T), \tag{15}$$

где $$A = \begin{bmatrix} \|BB^T\| & \|BQ_{12}B^T\| & \dots & \|BQ_{1m}B^T\| \\ \|BQ_{21}B^T\| & \|BB^T\| & \dots & \|BQ_{2m}B^T\| \\ \dots & \dots & \dots & \dots \\ \|BQ_{m1}B^T\| & \|BQ_{m2}B^T\| & \dots & \|BB^T\| \end{bmatrix};$$

$\sigma_{\max}(A)$ – наибольшее сингулярное значение А; $\lambda_{\min}(BB^T + BB^T \sum_j b_j^2 P_j BB^T)$ – наименьшее собственное значение матрицы в скобках, соответствующее наименьшему значению энтропии.

Отметим, что энтропия, как мера неопределенности, для случайного гауссова процесса полностью определяется значением его дисперсии:

$$H_x = 0{,}5\log 2\pi e\sigma_x^2, \tag{16}$$

где σ_x^2 – значение дисперсии.

Оценивая последовательность $x(k)$, приобретаем определенный уровень информации [5]:

$$I_{xy} = H_x - H_{x/y}, \tag{17}$$

где $H_{x/y}$ – условная энтропия, уровень неопределенности информации о процессе $x(k)$ при условии наблюдения (3) и оценки (5), т.е. подтверждается тот факт, что условная апостериорная энтропия $H_{x/y}$ всегда меньше априорной H_x.

Анализ игровой задачи. Для начала рассмотрим ситуацию, когда система состоит из двух игроков, управления наблюдения (3) которых принимают вид

$$y_1(k) = H_1(k)x_1(k) + V_1(k), \quad y_2(k) = H_2(k)x_2(k) + V_2(k), \tag{18}$$

где индексы 1 и 2 относятся к упомянутым игрокам. Можно утверждать, что при равной динамике, при $\tau_{\text{кор}1} = \tau_{\text{кор}2}$ апостериорные дисперсии $V_{\tilde{x}1}(k)$ и $V_{\tilde{x}2}(k)$ будут разниться настолько, насколько разнятся их априорные дисперсии V_w и уровни шумов наблюдения $V_v(k)$ в полосе приема сигналов $x(t)$. Очевидно, в силу разности $\tau_{\text{кор}}$ будут различаться и их ошибки прогноза. При этом, если $\tau_{\text{кор}2} > \tau_{\text{кор}1}$, то

$$V_{\tilde{x}2}(k/k-1) - V_{\tilde{x}1}(k/k-1) \geq V_{\tilde{x}2}(k) - V_{\tilde{x}1}(k). \tag{19}$$

Соответственно из (16) и (17) следует, что информация $I_2 \leq I_1$. Из соотношения (19) можно сделать вывод о том, что второй игрок, у которого апостериорный прогноз и оценка хуже, приобретает меньше информации и по качеству функционирования проигрывает первому игроку.

Рассмотрим далее кооперативную игру, когда игроки обладают общими стратегиями, а их потери (8) являются общими. Очевидно, при этом дополнительный выигрыш одного из участников уменьшает общие потери, т.е. дополнительная информация, попадающая в систему за счет тех или иных параметров и процессов, в кооперативной игре приносит общую пользу. С другой стороны, наличие в коалиции игрока, обладающего малой динамикой, с большими значениями априорной и апостериорной дисперсии при высоком уровне шума наблюдения приводит к большим общим потерям.

Таким образом, игровая задача может быть рассмотрена в рамках методов переменных состояний при векторном представлении уравнений состояния и наблюдения. Параметры этих уравнений полностью определяют содержательность задачи.

Список литературы: 1. *Сейдж Э.П.* Теория оценивания и ее применение в связи и управлении / Э.П. Сейдж, Дж. Л. Мелса; Пер. с англ. под ред. Б.Р. Левина. М.: Связь, 1976. 496 с. 2. *Олейник В.Ф.* Методы коллективного принятия решений в задачах ЭМС систем подвижной связи // Изв. вузов. Радиоэлектроника. 2002. № 6. 3. *Мулен Э.* Теория игр: Пер с англ. М.: Наука, 1985. 200 с. 4. *Горелик В.А.* Анализ конфликтных ситуаций в системах управления / В.А. Горелик, М.А. Горелов, А.Ф. Кононенко. М.: Радио и связь, 1991. 288 с. 5. *Вилкас Э.Й.* Оптимальность в играх и решениях. М.: Наука, 1990. 256 с. 6. *Ho Y.C.* A simple example on informatioveness and performance / Y.C. Ho, J. Blay // J. Optimiz. Theory Appl. 1973. Vol. II, N 4. 7. *Walsh P.M.* An example of the impact of information structure on decentralized multicriterion control problems / P.M. Walsh, J.B. Crus // IEEE Trans. Autom. Control. 1979. Vol. AC-24, N 12.

Харьковский национальный

университет радиоэлектроники,

Центр «Укрчастотнадзор», г. Киев

Поступила в редколлегию 10.06.2003