

Оптимизация управления передачей данных по флуктуирующему каналу связи при неточной информации о его состоянии¹

Н. А. Кузнецов^{*,**,a}, Д. В. Мясников^{*,b}, К. В. Семенихин^{*,**,***,c}

^{*}Институт радиотехники и электроники им. В.А. Котельникова РАН, Москва, Россия,

^{**}Московский физико-технический институт (государственный университет),
Долгопрудный, Московская область, Россия,

^{***}Московский авиационный институт (национальный исследовательский университет),
Москва, Россия

e-mail: ^akuznetsov@cplire.ru, ^bdmitry.myasnikov89@gmail.com, ^csiemenkv@gmail.com

Поступила в редколлегию 24.05.2018

Аннотация—Рассматривается задача оптимизации передачи пакетов по каналу связи, состояние которого описывается ненаблюдаемым марковским процессом. Моделью передатчика служит одноканальная система массового обслуживания с конечной очередью и интенсивностью обработки в виде произведения управляемой скорости передачи и коэффициента, зависящего от состояния канала. На вход поступает нестационарный пуассоновский поток пакетов. Переполнение очереди приводит к потерям, а ухудшение связи — к замедлению интенсивности обработки. Управление выбирается из условия минимума среднего числа потерь с учетом ограничения на энергозатраты передатчика на конечном промежутке времени. В расширенной задаче указан явный вид оптимальной стратегии. На ее основе предложено несколько способов формирования управления по неполной информации. Для этого используется два варианта оценки скрытого состояния канала: оценка фильтрации и оценка по текущему состоянию очереди. Полученные стратегии анализируются по результатам численного эксперимента.

КЛЮЧЕВЫЕ СЛОВА: система массового обслуживания, передача данных, оптимальное управление, расширенная задача, управление по неполной информации, оценка фильтрации.

1. ВВЕДЕНИЕ

Для распределенных инфокоммуникационных систем насущной проблемой является отсутствие полной информации о текущем состоянии соединений между различными узлами, загруженности вычислительных ресурсов и даже исправности отдельных элементов сети. Для оценки качества сетевого взаимодействия можно использовать агрегированные статистические данные, отражающие время кругового обращения пакетов и ошибки при их передаче [1, 2]. Несмотря на то, что эта информация не может полностью характеризовать состояние сети в целом, она позволяет выработать гибкие и эффективные механизмы преодоления коллизий при передаче данных по зашумленным каналам связи [3–5]. Тот же вывод можно сделать в отношении систем массового обслуживания (СМО): наличие лишь частичной информации о состоянии системы не является принципиальным ограничением для успешного управления ее ресурсами [6, 7].

Для учета неопределенности входных потоков запросов и случайности процессов передачи данных описание инфокоммуникационных систем, как правило, строится на базе методологии

¹ Работа выполнена при поддержке Российского научного фонда (проект № 16-11-00063).

марковских цепей с дискретным или непрерывным временем. К настоящему времени для решения задач управления и оптимизации в таких системах сформировались два принципиально разных подхода.

Первый из них основан на исследовании марковского процесса принятия решений в стационарном режиме, исключая оптимизацию затрат на вывод системы из неблагоприятного начального состояния [8–10]. При наличии частичной информации управление ищется в виде функции только от текущего значения наблюдаемых компонент вектора состояния без учета предыстории. Для устранения проблем, связанных с отсутствием выпуклости соответствующих оптимизационных моделей, традиционно используют смешанные, т.е. рандомизованные стратегии. Но эффект от них достигается только на значительном промежутке времени, на порядок превышающем время, необходимое для перехода к стационарному режиму. Поэтому применение соответствующих методов возможно только для инфокоммуникационных систем, которые функционируют в течение длительного времени при сохранении стационарности входных потоков и характеристик внешних возмущений.

При использовании второго подхода система рассматривается на конечном горизонте с целью оптимизации интегральных или терминальных характеристик качества ее функционирования [6, 11, 12]. Поэтому для синтеза соответствующих решений используются методы оптимального стохастического управления. Они особенно эффективны при наличии небольшого числа агрегированных состояний, к которым можно свести сложные модели функционирования распределенных инфокоммуникационных систем. При наличии неточной информации о системе искомое управление учитывает предшествующее ее развитие в форме оценок ненаблюдаемых компонент. Эти оценки — суть апостериорные вероятности, которые находятся в результате решения соответствующих уравнений фильтрации. Еще одно важное достоинство данного подхода — включение нестационарности входных потоков и учет изменчивости условий внешней среды. Важность этих условий ярко проявляется в задаче управления передачей данных между узлами распределенной роботизированной системы, выполняющей свое задание автономно в случайно меняющейся обстановке [13, 14].

В данной работе рассматривается задача оптимизации передачи однопакетных данных (пакетов) по каналу связи, состояние которого неизвестно и подвержено случайным изменениям. Процесс поступления пакетов описывается нестационарным пуассоновским потоком, а процесс передачи — одноканальной конечной СМО с интенсивностью обработки в виде произведения управляемой скорости передачи и коэффициента, зависящего от состояния канала связи. Переполнение очереди ведет к потерям поступающих пакетов, а ухудшение связи — к замедлению итоговой скорости пересылки. Цель управления — выбрать скорость передачи из условия минимума среднего числа потерь с учетом ограничения на энергозатраты передатчика на фиксированном интервале времени. Управление строится на основе информации о текущей загрузке СМО. Эта информация рассматривается в качестве косвенных наблюдений за неизвестным состоянием канала. В работе предлагается несколько стратегий, использующих оптимальное управление по полной информации и оптимальный фильтр состояния канала связи. Сравнение предложенных стратегий проводится в ходе численного эксперимента.

2. ПОСТАНОВКА ЗАДАЧИ

Для описания сети передачи данных введем пару марковских процессов X_t, θ_t , определенных на конечном промежутке времени $[0, T]$. Процесс X_t описывает состояние одноканальной СМО с нестационарным пуассоновским входным потоком, имеющим заданную интенсивность $\alpha(t)$, а θ_t — ненаблюдаемое состояние флуктуирующего канала связи. СМО определяет состояние очереди пакетов, предназначенных для дальнейшей отсылки на базовую станцию с передатчика, установленного на борту беспилотного аппарата. Допустим, что N — наибольшее

число пакетов, которое может вмещать система, а K — число агрегированных состояний канала связи. Следуя методологии стохастического анализа [11], возьмем в качестве множеств состояний

$$S^X = \{e_0, e_1, \dots, e_N\} \subset \mathbb{R}^{N+1} \quad \text{и} \quad S^\theta = \{f_1, \dots, f_K\} \subset \mathbb{R}^K, \quad (1)$$

где e_n и f_k — единичные векторы-столбцы.

Так как X_t — процесс рождения и гибели, его модель полностью определяется двумя характеристиками: интенсивностью прихода пакета $a_{n,n+1} = \alpha(t)$, $n = 0, 1, \dots, N-1$ и интенсивностью обработки $a_{n,n-1} = \langle d(t), \theta \rangle \mu$, $n = 1, \dots, N$, где μ — управляемая скорость передачи данных, а $\langle d(t), \theta \rangle$ — множитель, уменьшающий эту скорость из-за ожидания подтверждения об успешной отправке или сообщения о потере (скобка $\langle \cdot, \cdot \rangle$ обозначает скалярное произведение). Компоненты вектора $d(t) \in \mathbb{R}^K$ удовлетворяют естественному правилу: чем лучше состояние канала, тем меньше в среднем тратится время на доставку пакета. Поэтому будем считать, что номера состояний f_1, \dots, f_K выбраны согласно соотношению $1 \geq d_1(t) > \dots > d_K(t) > 0$.

Таким образом, определен генератор $A(t, \theta, \mu) = \{a_{n,p}(t, \theta, \mu)\}_{n,p=0,1,\dots,N}$, иначе говоря, матрица интенсивностей переходов марковского процесса X_t .

Генератор процесса θ_t обозначим через $B(t) = \{b_{k,l}(t)\}_{k,l=1,\dots,K}$. Интенсивности переходов предполагаются зависящими от времени для того, чтобы учесть вызванное перемещением передатчика изменение условий соединения.

Для обоих процессов будет использоваться мартингалное разложение [11]:

$$dX_t = A^*(t, \theta_t, \mu_t) X_t dt + dM_t^X, \quad d\theta_t = B^*(t) \theta_t dt + dM_t^\theta, \quad (2)$$

где A^* и B^* — транспонированные матрицы, а M_t^X, M_t^θ — пара векторных квадратично интегрируемых мартингалов, ортогональных друг другу. Условие ортогональности двух мартингалов означает, что их скачки, или что то же самое, скачки процессов X_t, θ_t происходят в разные моменты времени.

Назовем *управлением по полной информации* случайный процесс μ_t , если он:

- (i) принимает (при каждом t) значения из промежутка $[\underline{m}(t), \overline{m}(t)]$ с известными границами $0 < \underline{m}(t) \leq \overline{m}(t) < \infty$;
- (ii) имеет кусочно-непрерывные траектории;
- (iii) является предсказуемым относительно фильтрации $\mathcal{G} = \{\mathcal{G}_t\}_{t \geq 0}$, порожденной процессами X_t и θ_t .

Аналогично этому процесс μ_t — есть *управление по неполной информации*, если выполнены условия (i) и (ii), а предположение (iii) заменено следующим:

- (iv) процесс μ_t предсказуем относительно фильтрации $\mathcal{X} = \{\mathcal{X}_t\}_{t \geq 0}$, порожденной только процессом X_t .

Классы управлений по полной и соответственно по неполной информации обозначаются буквами \mathcal{C} и \mathcal{I} . Предположение $\mu \in \mathcal{C}$ означает, что μ_t полностью определяется эволюцией состояний СМО и канала связи вплоть до текущего момента t : $\{X_s, \theta_s: s \in (0, t)\}$. Однако в случае $\mu \in \mathcal{I}$ мы исходим из того, что состояние канала θ_t недоступно прямому измерению, поэтому в каждый момент времени управление μ_t функционально зависит только от предыстории состояния СМО $\{X_s, s \in (0, t)\}$.

Рассмотрим два показателя качества управления сетью передачи данных:

$$J_0[\mu] = \int_0^T \mathbf{E}\{\mathbf{I}\{X_t = N\}\alpha(t) + \mathbf{I}\{X_t > 0\}\langle \ell(t), \theta_t \rangle \mu_t\} dt, \quad (3)$$

$$J_1[\mu] = \int_0^T \mathbf{E}\{p_t\} dt, \quad (4)$$

где $J_0[\mu]$ — среднее количество потерянных пакетов, а $J_1[\mu]$ — энергия сигнала, генерируемого передатчиком.

При переполнении очереди, т.е. при $X_t = N$, все входящие пакеты теряются. Этот факт учитывает первое слагаемое в (3), где $\mathbf{I}\{\dots\}$ обозначает индикатор. Второе слагаемое описывает интенсивность случайных потерь на пути от передатчика к базовой станции. Интенсивность этих потерь предполагается пропорциональной скорости передачи для того, чтобы учесть естественное требование: чем больше посылаются данных, тем больше их теряется. Коэффициент этой пропорциональности определяется состоянием канала связи, причем в хорошем состоянии потерь в среднем меньше. Поэтому соответствующие коэффициенты, т.е. компоненты вектора $\ell(t)$, удовлетворяют неравенствам $0 \leq \ell_1(t) \leq \dots \leq \ell_K(t) \leq 1$.

Вид функционала (4) означает, что энергозатраты передатчика определяются мощностью сигнала p_t . Согласно теории информации [15] пропускная способность канала равна максимальной скорости передачи данных μ_t , которая может быть достигнута при заданной мощности сигнала p_t . Для гауссовского стационарного канала указанные величины связаны соотношениями

$$\mu_t = \frac{w}{H} \ln \left\{ 1 + \frac{p_t}{w \langle \nu(t), \theta_t \rangle} \right\} \quad \text{и} \quad p_t = w \langle \nu(t), \theta_t \rangle (\exp\{H \mu_t / w\} - 1). \quad (5)$$

Здесь $\langle \nu(t), \theta_t \rangle$ — интенсивность шума, зависящая от состояния канала связи. Компоненты вектора $\nu(t) \in \mathbb{R}^K$ удовлетворяют неравенствам $0 < \nu_1(t) \leq \dots \leq \nu_K(t) < \infty$ в силу того, что шумовая нагрузка на канал тем слабее, чем лучше его состояние. Коэффициент w определяет ширину спектра сигнала. Параметр H — энтропия наугад взятого пакета (она равна $L \ln 2$ в том случае, если пакеты могут быть произвольными фиксированной длины L бит). Если размерностью μ_t считать «число пакетов в секунду», то единицей измерения для w будет «Гц», для p_t — «Вт», для $\nu_k(t)$ — «Вт/Гц».

Далее без ограничения общности можно положить $w = 1$, т.е. в качестве единицы измерения времени взята величина, обратная ширине спектра сигнала. Поэтому функционал энергозатрат принимает вид

$$J_1[\mu] = \int_0^T \mathbf{E}\{\langle \nu(t), \theta_t \rangle (\exp\{H \mu_t\} - 1)\} dt. \quad (6)$$

Допустим, что на мощность сигнала есть технологические ограничения

$$0 < \underline{p} \leq p_t \leq \bar{p} < \infty. \quad (7)$$

Чтобы они были заведомо выполнены, предположим, что введенные ранее границы скорости передачи данных взяты равными

$$\underline{m}(t) = \frac{1}{H} \ln \left\{ 1 + \frac{\underline{p}}{\nu_1(t)} \right\}, \quad \bar{m}(t) = \frac{1}{H} \ln \left\{ 1 + \frac{\bar{p}}{\nu_K(t)} \right\}. \quad (8)$$

Задача оптимального управления по полной информации принимает вид:

$$J_0[\mu] \rightarrow \min_{\mu \in \mathcal{C}} \quad \text{при ограничении} \quad J_1[\mu] \leq \bar{J}_1, \quad (9)$$

где \bar{J}_1 — заданная верхняя граница. Она будет назначаться из диапазона $(T\underline{p}, T\bar{p})$, поскольку при $\bar{J}_1 < T\underline{p}$ энергетическое ограничение $J_1[\mu] \leq \bar{J}_1$ нарушается даже для минимальной загрузки канала, а при $\bar{J}_1 > T\bar{p}$, наоборот, ограничение оказывается лишним, коль скоро оно выполняется для максимальной загрузки.

Итак, цель решения задачи (9) — определить управление $\hat{\mu}$, дающее на фиксированном промежутке времени T минимальное в среднем количество потерянных пакетов на основе полной информации с учетом лимита энергии сигнала, генерируемого передатчиком.

Оптимальное управление по неполной информации определяется аналогично (9) с заменой \mathcal{C} на класс \mathcal{I} .

3. ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ПО ПОЛНОЙ ИНФОРМАЦИИ

Для построения оптимального управления по полной информации $\hat{\mu}$ введем расширенный функционал, соответствующий задаче с ограничением (9):

$$L[\mu, \lambda] = J_0[\mu] + \lambda(J_1[\mu] - \bar{J}_1), \quad \mu \in \mathcal{C}, \quad \lambda \geq 0. \quad (10)$$

Предлагаемый метод построения управления $\hat{\mu}$, оптимального при наличии ограничения, основан на принципе двойственной оптимизации [12] и состоит из следующих шагов:

а) нахождение оптимального управления в расширенной задаче:

$$\tilde{\mu}(\lambda) \in \arg \min_{\mu \in \mathcal{C}} L[\mu, \lambda], \quad \lambda \geq 0; \quad (11)$$

б) решение двойственной задачи

$$\hat{\lambda} \in \arg \max_{\lambda \geq 0} L[\tilde{\mu}(\lambda), \lambda]; \quad (12)$$

в) определение искомого управления $\hat{\mu} = \tilde{\mu}(\hat{\lambda})$.

Для синтеза управления (11) будем следовать изложенной в [7] схеме.

Определим сначала *составной процесс*, описывающий состояние сети передачи данных в целом, в виде тензорного произведения $X_t \otimes \theta_t$. Тензорное произведение двух векторов $x \in \mathbb{R}^{N+1}$ и $y \in \mathbb{R}^K$ отождествляется с матрицей $x \otimes y = xy^*$ размера $(N+1) \times K$. Пространство матриц $\mathbb{R}^{(N+1) \times K}$ наделяется скалярным произведением $\langle \pi, \varphi \rangle = \text{tr}[\pi\varphi^*]$, где $\pi, \varphi \in \mathbb{R}^{(N+1) \times K}$. В частности $\langle \pi, x \otimes y \rangle = \langle x, \pi y \rangle$.

Поскольку скачки X_t, θ_t происходят в разные моменты времени, по формуле Ито имеем $d(X_t \otimes \theta_t) = dX_t \otimes \theta_{t-} + X_{t-} \otimes d\theta_t$. С учетом (2) получаем стохастическое дифференциальное уравнение

$$d(X_t \otimes \theta_t) = \{A^*(t, \theta_t, \mu_t)(X_t \otimes \theta_t) + (X_t \otimes \theta_t)B(t)\} dt + dM_t, \quad (13)$$

где M_t — матричный квадратично интегрируемый мартингал, определяемый уравнением $dM_t = dM_t^X \otimes \theta_{t-} + X_{t-} \otimes dM_t^\theta$. Тем самым составной процесс $X_t \otimes \theta_t$ является управляемым марковским процессом со значениями в множестве $\mathbb{S}^X \otimes \mathbb{S}^\theta$. Его генератор можно найти из мартингального разложения (13).

Для этого обозначим через $C(t, m)$ генератор составного процесса, соответствующего постоянному управлению $\mu_t = m$, где m — фиксированное положительное число. По определению он представляет собой линейный оператор, действующий в пространстве $\mathbb{R}^{(N+1) \times K}$. Вид сопряженного оператора $C^*(t, m)$ следует из (13):

$$C^*(t, m)[e_n \otimes f_k] = A^*(t, f_k, m)(e_n \otimes f_k) + (e_n \otimes f_k)B(t). \quad (14)$$

Квадратные скобки в левой части используются для указания аргумента, к которому применяется линейный оператор. Тогда нетрудно записать вид исходного оператора

$$(C(t, m)[\varphi])_{n,k} = (A(t, f_k, m)\varphi + \varphi B^*(t))_{n,k} \quad (15)$$

с учетом того, что $(\dots)_{n,k}$ обозначает операцию взятия в матрице элемента с индексами (n, k) , а φ — произвольная матрица из $\mathbb{R}^{(N+1) \times K}$.

Теперь схему построения оптимального управления в расширенной задаче можно свести к последовательности нескольких шагов:

1) указать представление расширенного функционала

$$L[\mu, \lambda] = \int_0^T \mathbb{E} \langle F(t, \mu_t, \lambda), X_t \otimes \theta_t \rangle dt \quad (16)$$

с помощью подходящей матричной функции $F(t, m, \lambda) = \{F_{n,k}(t, m, \lambda)\}$;

2) задать матричную функцию $W(\cdot) = \{W_{n,k}(\cdot)\}$

$$W(t, \varphi, m, \lambda) = C(t, m)[\varphi] + F(t, m, \lambda); \quad (17)$$

3) решить параметрическую задачу минимизации

$$\tilde{\mathbf{m}}(\cdot) = \{\tilde{m}_{n,k}(\cdot)\}: \quad \tilde{m}_{n,k}(t, \varphi, \lambda) \in \arg \min_{m \in [\underline{m}(t), \bar{m}(t)]} W_{n,k}(t, \varphi, m, \lambda); \quad (18)$$

4) определить матричную функцию $\varphi(t, \lambda) = \{\varphi_{n,k}(t, \lambda)\}$ в результате интегрирования системы обыкновенных дифференциальных уравнений

$$\dot{\varphi}_{n,k} = -W_{n,k}(t, \varphi, \tilde{m}_{n,k}(t, \varphi, \lambda), \lambda), \quad t \in [0, T], \quad (19)$$

с терминальным условием $\varphi|_{t=T} = 0$;

5) получить искомое управление и соответствующее значение критерия

$$\tilde{\mu}_t(\lambda) = \langle \tilde{\mathbf{m}}(t, \varphi(t, \lambda), \lambda), X_{t-} \otimes \theta_{t-} \rangle, \quad L[\tilde{\mu}(\lambda), \lambda] = \langle \varphi(0, \lambda), \pi(0) \rangle, \quad (20)$$

где $\pi(0)$ — матрица вероятностей начального состояния сети, т.е. $\pi(0) = \mathbb{E}\{X_0 \otimes \theta_0\}$.

Все матричные функции $F(\cdot)$, $W(\cdot)$, $\varphi(\cdot)$, $\tilde{\mathbf{m}}(\cdot)$, $\pi(t) = \mathbb{E}\{X_t \otimes \theta_t\}$ имеют одну и ту же размерность $(N+1) \times K$.

Перейдем к реализации представленной схемы. Начнем с подынтегральной функции в (16)

$$F_{n,k}(t, m, \lambda) = \mathbb{I}\{n = N\}\alpha(t) + \mathbb{I}\{n > 0\}\ell_k(t) m + \lambda \nu_k(t)(e^{Hm} - 1) - \lambda \bar{J}_1/T. \quad (21)$$

С учетом (15) и (21) функция (17) теперь вполне определена.

Для решения задачи минимизации (18) потребуется явная зависимость от параметра m . Поэтому раскроем выражение (15)

$$(C(t, m)[\varphi])_{n,k} = \sum_{p=0}^N a_{n,p}(t, f_k, m) \varphi_{p,k} + \sum_{l=1}^K \varphi_{n,l} b_{k,l}(t),$$

где первое слагаемое имеет вид

$$\begin{cases} -\alpha(t) \varphi_{0,k} + \alpha(t) \varphi_{1,k}, & n = 0; \\ d_k(t)m \varphi_{n-1,k} - (d_k(t)m + \alpha(t)) \varphi_{n,k} + \alpha(t) \varphi_{n+1,k}, & 0 < n < N; \\ d_k(t)m \varphi_{N-1,k} - d_k(t)m \varphi_{N-1,N}, & n = N. \end{cases}$$

Используя (21) и заменяя не зависящие от m слагаемые многочлином, получаем

$$W_{n,k}(t, \varphi, m, \lambda) = \begin{cases} \lambda \nu_k(t) e^{Hm} + \dots, & n = 0; \\ -\{d_k(t)(\varphi_{n,k} - \varphi_{n-1,k}) - \ell_k(t)\}m + \lambda \nu_k(t) e^{Hm} + \dots, & n > 0. \end{cases}$$

Тогда при $\lambda > 0$, то искомая стратегия (18) находится как

$$\tilde{m}_{n,k}(t, \varphi, \lambda) = \begin{cases} \underline{m}(t), & n = 0; \\ m^o\left(\frac{d_k(t)(\varphi_{n,k} - \varphi_{n-1,k}) - \ell_k(t)}{\lambda \nu_k(t)}; \underline{m}(t), \overline{m}(t)\right), & n > 0, \end{cases} \quad (22)$$

где $m^o(c; \underline{m}, \overline{m})$ — точка минимума функции $-cm + e^{Hm}$ на отрезке $[\underline{m}, \overline{m}]$. Эта точка определена единственным образом:

$$m^o(c; \underline{m}, \overline{m}) = \begin{cases} \underline{m}, & c \leq H e^{H\underline{m}}, \\ \ln(c/H)/H, & H e^{H\underline{m}} \leq c \leq H e^{H\overline{m}}, \\ \overline{m}, & c \geq H e^{H\overline{m}}. \end{cases} \quad (23)$$

Если же $\lambda = 0$, то при $n = 0$ можно положить $\tilde{m}_{n,k}(t, \varphi, \lambda) = \underline{m}(t)$, а при $n > 0$ стратегия $\tilde{m}_{n,k}(t, \varphi, \lambda)$ будет соответственно нижней границей, верхней границей или произвольной точкой отрезка $[\underline{m}(t), \overline{m}(t)]$ в зависимости от того, какое из трех условий выполнено: коэффициент $d_k(t)(\varphi_{n,k} - \varphi_{n-1,k})$ меньше, больше и равен параметру $\ell_k(t)$.

Теперь благодаря теореме 3 из [12] получаем следующий факт.

Теорема 1. При $\lambda > 0$ оптимальное управление в расширенной задаче (11) имеет вид (20), где стратегия $\tilde{m}_{n,k}(t, \varphi, \lambda)$ однозначно определяется формулами (22) и (23), а функция $\varphi(t, \lambda)$ находится из решения дифференциальной системы (19) с правой частью, задаваемой выражением (17) с учетом (15) и (21).

Чтобы пояснить структуру оптимальной стратегии (22) необходимо сначала объяснить смысл решения $\varphi_{n,k}(t, \lambda)$ дифференциальной системы (19). Согласно методу динамического программирования величина $\varphi_{n,k} = \varphi_{n,k}(t, \lambda)$ равна оптимальному значению функционала $L[\mu, \lambda]$ при рассмотрении задачи (11) на промежутке $[t, T]$ с начальным условием $X_t = e_n$, $\theta_t = f_k$. Тогда разность $\varphi_{n,k} - \varphi_{n-1,k} > 0$ характеризует выгоду от уменьшения загрузки СМО.

Если для простоты положить $\ell_k = 0$, то в относительных величинах эта выгода будет описываться коэффициентом $c = (\varphi_{n,k} - \varphi_{n-1,k}) \cdot d_k / (\lambda \nu_k)$. В этом выражении отношение $d_k / (\lambda \nu_k)$ тем меньше, чем хуже состояние канала связи f_k или больше вес энергетического критерия λ . Поэтому величина c будет малой при плохом качестве связи или жестких энергетических ограничениях. При этих предположениях согласно (23) оптимальная скорость передачи назначается минимально возможной. Если же, наоборот, состояние связи хорошее, а энергетический критерий не является доминирующим показателем, то в (23) реализуется третий случай, когда оптимальная скорость передачи устанавливается на максимальном уровне \bar{m} . Второй случай из описанных в (23) соответствует пограничной ситуации.

Вернемся к условию $\ell_k > 0$. Оно означает наличие случайных потерь, не связанных с состоянием СМО. Увеличение процента потерь ℓ_k ведет к сокращению величины c , определяющей выигрыш от снижения загрузки СМО.

Осталось рассмотреть случай $\lambda = 0$, в котором ограничение на энергию сигнала отсутствует. Структура решения $\tilde{m}_{n,k}(t, \varphi, 0)$ остается в целом прежней за исключением того, что переход между двумя крайними стратегиями \underline{m} и \bar{m} реализуется в виде скачка, поскольку пограничная ситуация определяется уравнением $d_k(\varphi_{n,k} - \varphi_{n-1,k}) = \ell_k$, а не диапазоном (как в случае $\lambda > 0$).

Теперь решение задачи оптимального управления по полной информации получается с помощью теоремы 4 из [12]. Ее применение основано на непрерывной зависимости оптимальной стратегии (22) от множителя $\lambda > 0$.

Теорема 2. *Если стратегия $\underline{m}(t)$ удовлетворяет ограничению $J_1[\underline{m}] < \bar{J}_1$, то решение $\hat{\lambda}$ двойственной задачи (12) существует и не превосходит*

$$\bar{\lambda} = J_0[\underline{m}] / (\bar{J}_1 - J_1[\underline{m}]). \quad (24)$$

Если $\hat{\varphi}(t)$ — решение дифференциальной системы (19) при $\lambda = \hat{\lambda}$, то с помощью стратегии $\tilde{\mathbf{m}} = \{\tilde{m}_{n,k}\}$, определенной в (22)–(23), оптимальное управление в задаче (9) имеет вид

$$\hat{\mu}_t = \langle \tilde{\mathbf{m}}(t, \hat{\varphi}(t), \hat{\lambda}), X_{t-} \otimes \theta_{t-} \rangle. \quad (25)$$

Полученный результат является следствием того, что пара $(\hat{m}, \hat{\lambda})$ образует седловую точку функционала $L[\mu, \lambda]$, т.е.

$$L[\hat{\mu}, \lambda] \leq L[\hat{\mu}, \hat{\lambda}] \leq L[\mu, \hat{\lambda}] \quad \forall \mu \in \mathcal{C} \quad \forall \lambda \geq 0. \quad (26)$$

Из левого неравенства следуют ограничение $J_1[\hat{\mu}] \leq \bar{J}_1$ и условие дополняющей нежесткости $\hat{\lambda}(J_1[\hat{\mu}] - \bar{J}_1) = 0$. Отсюда вытекает $L[\hat{\mu}, \hat{\lambda}] = J_0[\hat{\mu}]$, что с учетом правого неравенства дает соотношение $\hat{\lambda} \leq \bar{\lambda}$. Функция, максимизируемая в (12), является вогнутой, как минимум линейных. Поэтому множитель $\hat{\lambda}$ можно определить методом деления отрезка пополам или методом золотого сечения на отрезке $[0, \bar{\lambda}]$.

4. УПРАВЛЕНИЕ ПО НЕПОЛНОЙ ИНФОРМАЦИИ

Для синтеза алгоритма управления сетью передачи данных в отсутствии полной информации о состоянии канала связи θ_t необходимо сначала решить задачу фильтрации

$$\vartheta_t = \mathbf{E}\{\theta_t \mid \mathcal{X}_t\}. \quad (27)$$

Вектор ϑ_t представляет собой оптимальную в среднем квадратическом оценку скрытого состояния канала по текущим наблюдениям $\{X_s: s \leq t\}$ за состоянием СМО. В силу предположения $\theta_t \in S^\theta$ компоненты вектора ϑ_t являются условными вероятностями

$$\langle \vartheta_t, f_k \rangle = P\{\theta_t = f_k \mid \mathcal{X}_t\}. \quad (28)$$

Уравнения для искомой оценки фильтрации (27) выведены в [11, Спар. 7] в форме уравнения Закаи для ненормированной оценки $\tilde{\vartheta}_t$, которая после нормировки приводит к искомой

$$\vartheta_t = \tilde{\vartheta}_t / \langle \tilde{\vartheta}_t, e \rangle, \quad (29)$$

где e — вектор, имеющий соответствующую размерность и состоящий только из единиц. Упомянутое уравнение имеет вид

$$d\tilde{\vartheta}_t = B^*(t)\tilde{\vartheta}_t dt + \sum_{k=1}^K \left(\langle \gamma^{-1} A_0^*(t, f_k, \mu_t) X_{t-} - e, d\bar{N}_t \rangle + [\gamma N - \langle a(t, f_k, \mu_t), X_t \rangle] dt \right) \langle \tilde{\vartheta}_{t-}, f_k \rangle f_k, \quad (30)$$

где γ — произвольный положительный параметр, имеющий размерность интенсивности; $A_0 \in \mathbb{R}^{(N+1) \times (N+1)}$ — матрица, полученная из A в результате заполнения главной диагонали нулями, т.е. $A_0 = A - \text{diag}[a_{0,0}, \dots, a_{N,N}]$; $a \in \mathbb{R}^{N+1}$ — вектор-столбец, составленный из интенсивностей выхода $a_n = -a_{n,n}$ из состояния e_n , $n = 0, 1, \dots, N$; $\bar{N}_t \in \mathbb{R}^{N+1}$ — векторный процесс, p -я компонента которого равна количеству переходов в состояние e_p за время t , $p = 0, 1, \dots, N$.

Введем дополнительные обозначения: пусть $T_1 < T_2 < \dots$ — последовательность моментов переключений процесса X_t , причем $T_0 = 0$.

Процесс решения дифференциально-разностного уравнения (30) можно представить в виде описанной ниже рекуррентной процедуры.

Алгоритм 1. Выполнить следующие шаги:

- 0) взять вектор $v^{(0)}$ равным начальному распределению $\pi^\theta(0) = E\theta_0$ и положить $m = 0$;
- 1) если $X_t \equiv e_n$ при $t \in (T_m, T_{m+1})$, то определить $\vartheta_t = y(t) / \langle y(t), e \rangle$ через решение системы линейных дифференциальных уравнений

$$\dot{y} = (B^*(t) - \text{diag}[a_n(t, f_1, \mu_t), \dots, a_n(t, f_K, \mu_t)] + \gamma N \cdot I) y, \quad (31)$$

которая интегрируется на интервале (T_m, T_{m+1}) с начальным условием $y(T_m) = v^{(m)}$, где I — единичная матрица соответствующей размерности;

- 2) если в момент T_{m+1} для процесса X_t произошел переход $e_n \rightarrow e_p$, то положить

$$v^{(m+1)} = \gamma^{-1} \text{diag}[a_{n,p}(t, f_1, \mu_t), \dots, a_{n,p}(t, f_K, \mu_t)] y(t-) \Big|_{t=T_{m+1}}; \quad (32)$$

- 3) нормировать $v^{(m+1)} \stackrel{\text{def}}{=} v^{(m+1)} / \langle v^{(m+1)}, e \rangle$ и положить $\vartheta_{T_{m+1}} = v^{(m+1)}$;
- 4) если $T_{m+1} > T$, то закончить итерации, иначе увеличить m на единицу и перейти к шагу 1.

Итак, описанный выше алгоритм вычисления нормированной оценки ϑ_t содержит три принципиально важных шага: между скачками состояний СМО оценка определяется из решения системы линейных дифференциальных уравнений (31) с точностью до нормировки, в момент переключения состояния СМО происходит пересчет оценки и на шаге 3 производится нормировка вектора оценки. Нормировка вектора $v^{(m+1)}$ нужна для того, чтобы дифференциальная

система (31) решалась с нормированным начальным условием. В противном случае функция $y(t)$, вычисляемая в алгоритме, совпадала бы с решением $\tilde{\vartheta}_t$ уравнения Закаи (30), о котором известно, что $\tilde{\vartheta}_t \rightarrow 0$ п.н. при $t \rightarrow \infty$, причем на каждой траектории это происходит с экспоненциальной скоростью. Однако данное обстоятельство делает невозможным практическую реализацию нормировки вектора ϑ_t , который уже при сравнительно небольших t оказывается неотличимым от машинного нуля. Кроме того, важно отметить, что параметр γ не влияет на значение нормированной оценки ϑ_t , он лишь необходим для соблюдения физической размерности.

Наряду с оценкой состояния канала $\vartheta_t = \mathbf{E}\{\theta_t | \mathcal{X}_t\}$, использующей всю предысторию динамики СМО $\mathcal{X}_t = \sigma\{X_s : s \leq t\}$, рассмотрим также оценку $\mathbf{E}\{\theta_t | X_t\}$. Она использует только текущее значение процесса X_t и представляет собой вектор, составленный из условных вероятностей $\mathbf{P}\{\theta_t = f_k | X_t\}$, $k = 1, \dots, K$. Они выражаются через элементы матрицы совместного распределения $\pi(t) = \mathbf{E}\{X_t \otimes \theta_t\}$:

$$\mathbf{E}\{\langle \theta_t, f_k \rangle | X_t = e_n\} = \pi_{n,k}(t) / \sum_{l=1}^K \pi_{n,l}(t), \quad n = 0, 1, \dots, N.$$

С учетом того, что $\pi_{n,k}(t) = \langle \pi^*(t)e_n, f_k \rangle$, окончательно получаем

$$\mathbf{E}\{\theta_t | X_t\} = \pi^*(t)X_t / \langle \pi^*(t)X_t, e \rangle. \quad (33)$$

Теперь на основе описанной в теореме 2 оптимальной стратегии с полной информацией

$$\hat{\mathbf{m}}(t) = \{\hat{m}_{n,k}(t)\}, \quad \hat{\mu}_t = \langle \hat{\mathbf{m}}(t), X_{t-} \otimes \theta_{t-} \rangle, \quad (34)$$

укажем несколько способов построения управлений, использующих вместо истинного состояния канала его оценки (27) и (33). Вариант управления будем обозначать двумя буквами. Первая буква соответствует используемой оценке: «f» — для оценки фильтрации (filter); «c» — для оценки, основанной на текущем состоянии (current state). Вторая буква обозначает операцию, выполняемую над апостериорным распределением: «e» — усреднение (expectation); «m» — наиболее правдоподобное значение (mode).

Итак, рассмотрим четыре варианта управления по неполной информации:

1) f.e.-управление представляет собой усреднение управления по полной информации $\hat{\mu}_t$ относительно апостериорного распределения (27) ненаблюдаемого состояния канала с учетом всей предыстории \mathcal{X}_t , т.е.

$$\mu_t^{f.e.} = \mathbf{E}\{\hat{\mu}_t | \mathcal{X}_t\} = \langle \hat{\mathbf{m}}(t), X_{t-} \otimes \vartheta_t \rangle, \quad (35)$$

где ϑ_t определяется уравнениями (29), (30) и находится в соответствии с алгоритмом 1, в которых управление μ_t выбрано равным $\hat{\mu}_t$;

2) f.m.-управление использует тот же объем информации, но вместо усреднения стратегий $\hat{m}_{n,k}(t)$ по $k = 1, \dots, K$ выбирается только одна из них, соответствующая номеру κ_t наиболее правдоподобного состояния канала

$$\mu_t^{f.m.} = \langle \hat{\mathbf{m}}(t), X_{t-} \otimes f_{\kappa_t} \rangle, \quad \kappa_t = \arg \max_{k=1, \dots, K} \langle \vartheta_t, f_k \rangle; \quad (36)$$

3) c.e.-управление основано на усреднении относительно апостериорного распределения (33), которое определяется только по текущему значению X_t :

$$\mu_t^{c.e.} = \mathbf{E}\{\hat{\mu}_t | X_t\} = \langle \hat{\mathbf{m}}(t), X_t \otimes \mathbf{E}\{\theta_t | X_t\} \rangle; \quad (37)$$

4) с.м.-управление использует моду указанного распределения, т.е.

$$\mu_t^{c.m.} = \langle \widehat{\mathbf{m}}(t), X_t \otimes f_{\kappa_t} \rangle, \quad \kappa_t = \arg \max_{k=1, \dots, K} \langle E\{\theta_t | X_t\}, f_k \rangle. \quad (38)$$

Сравнение качества предложенных управлений будет проведено по результатам численного эксперимента (см. следующий раздел).

Важно отметить, что при использовании f.e.- или f.m.-управления вместо явного представления функционала (16) доступна лишь его потраекторная оценка

$$\mathcal{L}[\mu, \lambda] = \int_0^T \langle F(t, \mu_t, \lambda), X_t \otimes \vartheta_t \rangle dt.$$

Для получения приемлемой оценки необходимо многократно смоделировать динамику управляемой СМО совместно с решением уравнения фильтрации, а затем взять среднее арифметическое полученных потраекторных оценок.

Если же $\mu_t = \langle \mathbf{m}(t), X_t \rangle$ — это с.е.- или с.м.-управление, где $\mathbf{m}(t)$ — соответствующая стратегия, записанная в виде вектора с компонентами $m_n(t)$, $n = 0, 1, \dots, N$, то для функционала $L[\mu, \lambda]$ можно получить представление

$$L[\mu, \lambda] = \int_0^T \sum_{n=0}^N \sum_{k=1}^K F_{n,k}(t, m_n(t), \lambda) \pi_{n,k}(t) dt. \quad (39)$$

Набор вероятностей $\{\pi_{n,k}(t)\}$ определяется в результате интегрирования системы дифференциальных уравнений Колмогорова

$$\dot{\pi}_{p,l} = \sum_{n=0}^N \pi_{n,l} a_{n,p}(t, f_l, m_n(t)) + \sum_{k=1}^K \pi_{p,k} b_{k,l}(t), \quad (40)$$

где пара индексов (p, l) пробегает множество $\{0, 1, \dots, N\} \times \{1, \dots, K\}$.

5. ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Сеть передачи данных рассматривается на временном горизонте $T = 100$ с. Предположим, что СМО вмещает максимально $N = 8$ пакетов, а канал имеет два возможных состояния: f_1 — «хорошая связь» и f_2 — «плохая связь». Известно, что пакеты имеют фиксированный размер в 1500 байт, откуда $H = 8 \cdot 1500 \cdot \ln 2 \approx 8317$. Сигнал передается в спектральном диапазоне $w = 10^6$ Гц при наличии шума, интенсивность которого равна $\nu_1 = 10^{-14}$ или $\nu_2 = 10^{-6}$ Вт/Гц в зависимости от состояния канала [16]. Во время пересылки пакетов с передатчика на базовую станцию при хорошей связи задержек и потерь не происходит, а при плохой связи скорость передачи замедляется в два раза и 10% пакетов теряется. Эти условия можно кратко записать в виде

$$d_1 = 1, \quad d_2 = 1/2, \quad \ell_1 = 0, \quad \ell_2 = 0,1.$$

Управляемая скорость передачи данных μ_t должна находиться в пределах от $\underline{m} = 0,5$ до $\overline{m} = 8$ пакетов в секунду. Это означает, что мощность сигнала p_t меняется в диапазоне от $\underline{p} = 0,042 \cdot 10^{-9}$ до $\overline{p} = 0,069$ Вт, а отношение «сигнал/шум»

$$\text{SNR}_t = p_t / \langle \nu(t), \theta_t \rangle = \exp\{H \mu_t / w\} - 1$$

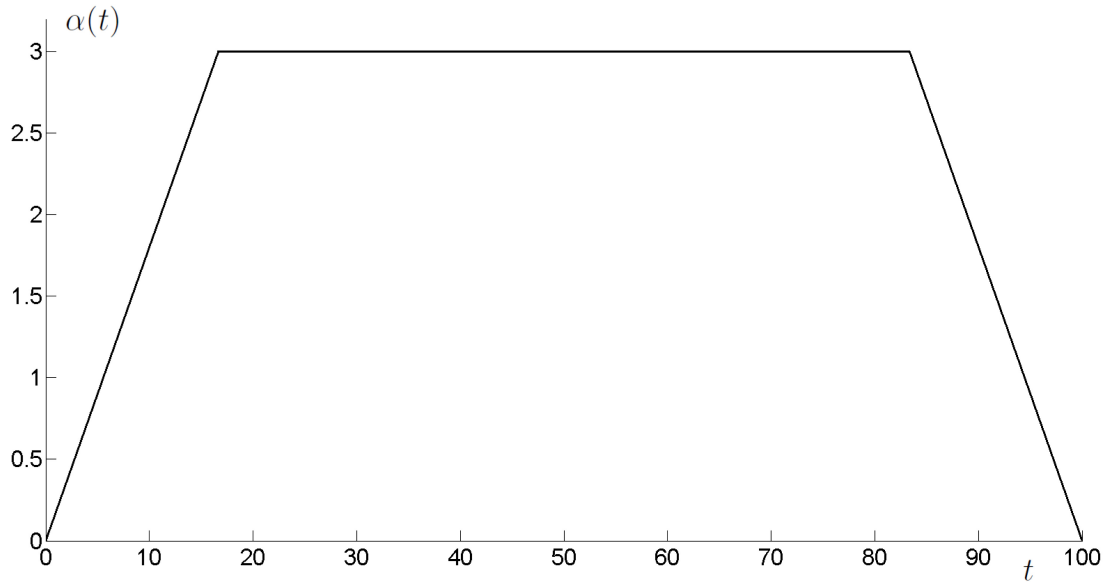


Рис. 1. Интенсивность входного потока $\alpha(t)$.

удовлетворяет $0,0042 \leq \text{SNR}_t \leq 0,0688$.

Интенсивность входного потока изображена на рис. 1. Промежуток времени от 15 до 85 с соответствует рабочему режиму, в котором передатчик должен переслать на базовую станцию максимальный объем информации.

Интенсивности переходов между состояниями канала связи представлены на рис. 2. Минимальные и максимальные значения интенсивностей отличаются более чем на порядок, а смена режимов между ними происходит резко и почти синхронно. Это позволяет достаточно точно восстановить истинное состояние канала θ_t . Например, на промежутке $[0, 20]$ большую часть времени будет иметь место $\theta_t = f_2$. Аналогично, на $[22, 38]$ в основном будет верно $\theta_t = f_1$ и т.д.

Согласно теореме 2 было найдено оптимальное управление

$$\hat{\mu}_t = \hat{m}_{n,k}(t) \quad \text{при} \quad X_{t-} = e_n, \quad \theta_{t-} = f_k,$$

удовлетворяющее ограничению на энергию сигнала $J_1[\mu] \leq \bar{J}_1$, где $\bar{J}_1 = 2$ Дж. Это ограничение оказалось активным, т.е. $J_1[\hat{\mu}] = \bar{J}_1$. Среднее число потерь получилось равным $J_0[\hat{\mu}] = 46,6$.

Стратегия $\hat{m}_{n,1}(t)$, соответствующая хорошему состоянию связи f_1 , оказалась тривиальной. Это означает, что при простом СМО мощность сигнала выбирается минимальной $p_t = \underline{p}$, а при наличии хотя бы одного пакета скорость передачи назначается максимальной, поэтому

$$\hat{m}_{0,1} \equiv \underline{m} \quad \text{и} \quad \hat{m}_{n,1} \equiv \bar{m} \quad \text{при} \quad n > 0. \quad (41)$$

Вид оптимальной стратегии $\hat{m}_{n,2}(t)$ при плохом качестве связи f_2 изображен на рис. 3. Зависимость от загрузки СМО в данной стратегии проявляется только на промежутках, где большую часть времени канал должен находиться в плохом состоянии (ср. с изображенной на рис. 2 интенсивностью перехода $b_{1,2}(t)$). На этих промежутках наблюдается естественная закономерность: чем больше пакетов в очереди, тем больше оптимальная скорость передачи данных. На интервалах, соответствующих хорошей связи, скорость передачи данных назначается минимально возможной, кроме ситуации переполнения очереди $n = 8$.

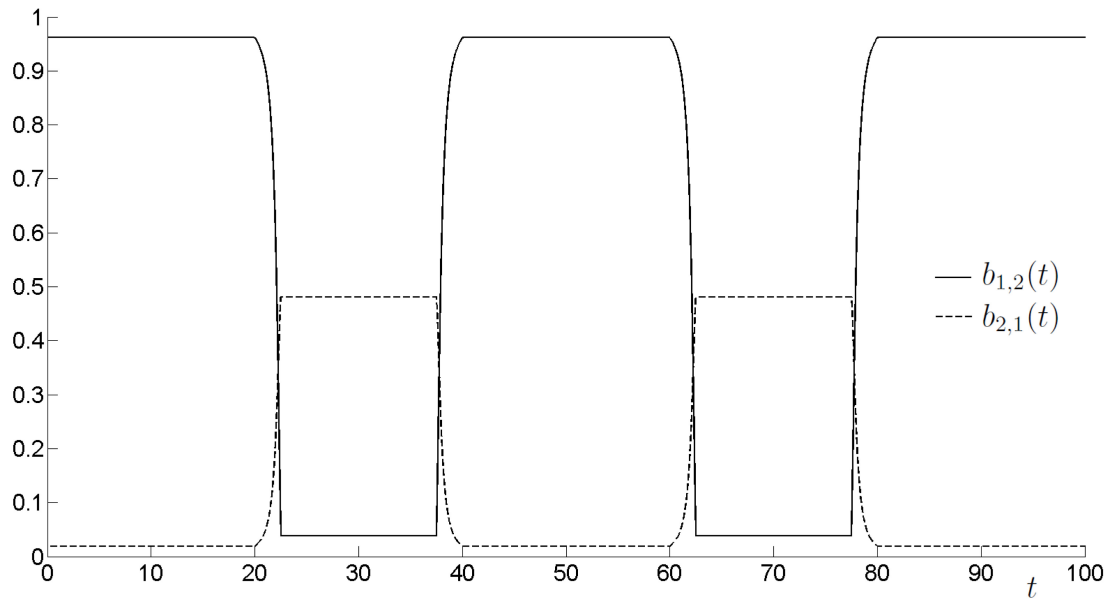


Рис. 2. Интенсивность переходов между состояниями канала связи $b_{1,2}(t)$, $b_{2,1}(t)$.

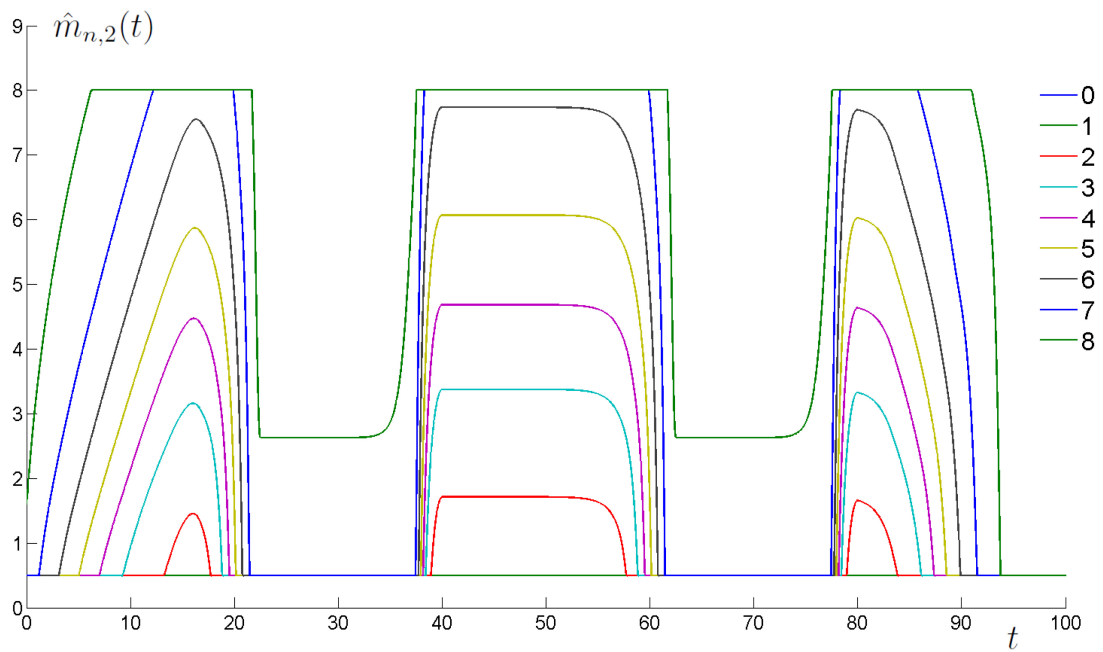


Рис. 3. Оптимальная стратегия $\{\hat{m}_{n,2}(t)\}_{n=0,\dots,s}$ передачи данных в плохом состоянии канала при активном ограничении на энергию сигнала.

Если верхнюю границу \bar{J}_1 увеличить в полтора раза до 3 Дж, то энергетическое ограничение будет выполнено с запасом $J_1[\hat{\mu}] = 2,549$ Дж. В этом случае множитель λ равен нулю и соответствующее оптимальное управление $\hat{\mu}_t$ минимизирует только функционал потерь. В сравнении с предыдущим случаем, где ограничение на энергию сигнала было существенным, среднее число потерянных пакетов оказалось значительно меньше $J_0[\hat{\mu}] = 32,6$. Оптимальная скорость передачи данных $\hat{m}_{n,2}(t)$ теперь имеет кусочно-постоянный вид, так как реализуются только две крайние альтернативы \underline{m} или \bar{m} (см. рис. 4). От числа пакетов n зависит лишь время, в течение которого скорость передачи равна верхней границе.

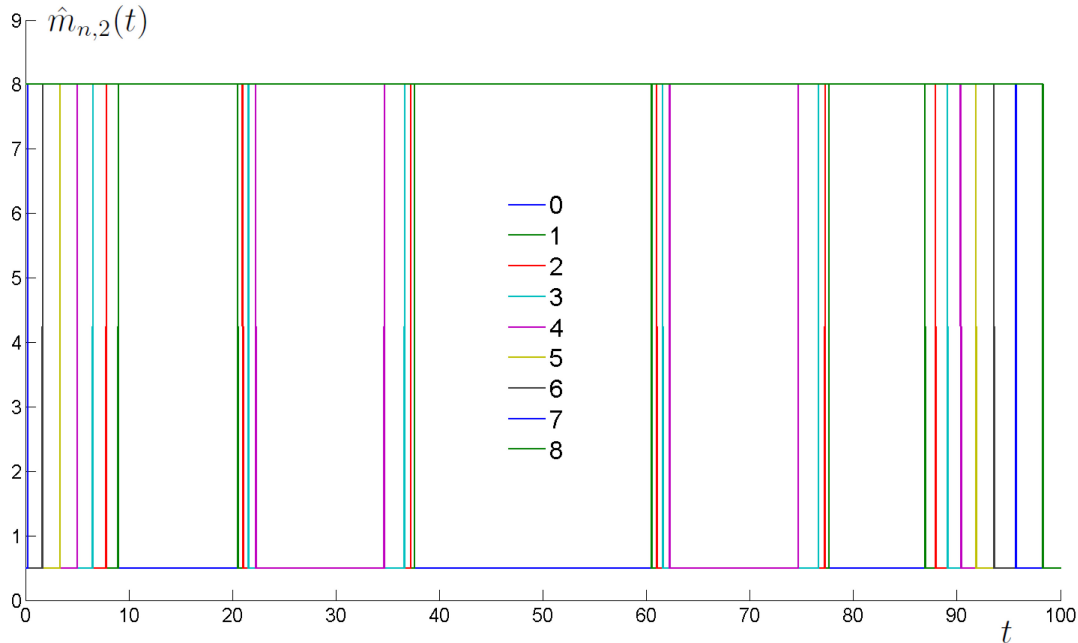


Рис. 4. Оптимальная стратегия $\{\hat{m}_{n,2}(t)\}_{n=0,\dots,8}$ передачи данных в плохом состоянии канала при пассивном ограничении на энергию сигнала.

Поскольку оба описанных решения оптимальны в задаче с расширенным функционалом $L[\mu, \lambda]$, их структура существенно зависит от наличия или отсутствия слагаемого в виде энергетического критерия (эти два варианта отвечают случаям $\lambda > 0$ и $\lambda = 0$ соответственно). При имеющемся наборе параметров отношение «сигнал/шум» значительно меньше единицы: $\text{SNR}_t \approx H\mu_t/w < 0,1$. Поэтому энергетический критерий (6) близок к функционалу, который можно получить предельным переходом при $w \rightarrow \infty$:

$$J_1^\infty[\mu] = \int_0^T \mathbf{E}\{H\langle\nu(t), \theta_t\rangle\mu_t\} dt.$$

Следовательно, в случае широкополосного сигнала и небольшой скорости передачи μ_t энергетический критерий можно считать линейным по μ_t .

Если же считать спектральный диапазон w конечным, но таким что $H\bar{m}/w \ll 1$, исходный функционал (6) приобретает смысл регуляризующего слагаемого в задаче с двумя плоскими (т.е. линейными по управлению) критериями $J_0[\mu]$ и $J_1^\infty[\mu]$. Тем самым рассматриваемая оптимизационная модель приводит к новой интерпретации метода регуляризации, предложенного ранее в [17].

Теперь перейдем к обсуждению результатов нахождения управлений по неполной информации.

На рис. 5 изображены три кривые, каждая из которых составлена из значений векторного критерия $(J_0[\mu], J_1[\mu])$, когда μ пробегает семейство управлений, параметризованных множителем $\lambda \geq 0$. Первое семейство определяется управлением $\tilde{\mu}(\lambda)$, оптимальным в расширенной задаче (11). Его структура описана в теореме 1. Два других семейства задаются управлениями по неполной информации: $\mu^{c.e.}(\lambda)$ и $\mu^{c.m.}(\lambda)$. Они используют только текущее состояние СМО и определяются согласно (37) и (38) с заменой $\hat{\mu}$ на $\tilde{\mu}(\lambda)$.

Нижняя кривая представляет собой границу множества достижимости, т.е. множества значений векторного критерия при различных управлениях по полной информации $\mu \in \mathcal{C}$. Вертикальный и горизонтальный участки соответствуют значениям $J_0 = 32,6$ и $J_1 = 0,292$ Дж. Они получены в результате минимизации на классе \mathcal{C} этих функционалов по отдельности. Остальная часть границы множества достижимости описывается строго выпуклой кривой. Это верно несмотря на то, что ее большая часть похожа на прямолинейный отрезок. Поэтому в рассматриваемой оптимизационной модели не пришлось прибегать к дополнительной процедуре регуляризации.

Если зафиксировать верхнюю границу на уровне $\bar{J}_1 = 2$ Дж, то за счет выбора λ можно указать стратегии с неполной информацией $\mu^{c.e.} = \mu^{c.e.}(\lambda)$ и $\mu^{c.m.} = \mu^{c.m.}(\lambda)$, такие что $J_1[\mu^{c.e.}] = J_1[\mu^{c.m.}] = \bar{J}_1$. Тогда среднее число потерь оказывается равным $J_0[\mu^{c.e.}] = 57,2$ и $J_0[\mu^{c.m.}] = 60,4$, что значительно больше, чем в случае оптимального управления по полной информации $J_0[\hat{\mu}] = 46,6$.

Если проанализировать операции, выполняемые над апостериорным распределением при формировании управления по неполной информации, то усреднение выглядит предпочтительнее в сравнении с выбором наиболее правдоподобного значения. Судя по графикам на рис. 5, эта закономерность нарушается только при заниженной границе $\bar{J}_1 \approx 0,6$ Дж.

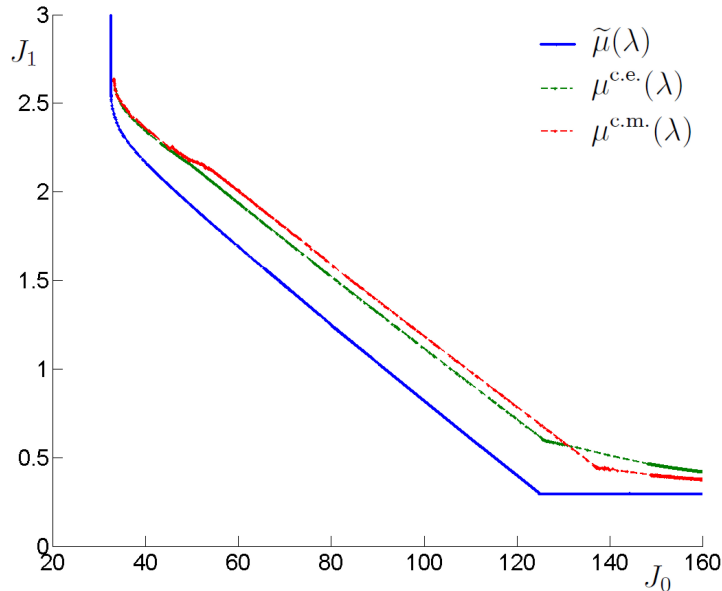


Рис. 5. Значения функционалов на семействах управлений $\tilde{\mu}(\lambda)$, $\mu^{c.e.}(\lambda)$, $\mu^{c.m.}(\lambda)$

Теперь рассмотрим результаты применения ф.е.-управления, которое согласно (35) формируется с использованием оценки фильтрации ϑ_t состояния ненаблюдаемого канала θ_t . При

множителе $\lambda = \hat{\lambda}$, соответствующем границе $\bar{J}_1 = 2$ Дж, управление $\mu_t^{\text{f.e.}}$ будет иметь вид

$$\mu_t^{\text{f.e.}} = \hat{m}_{n,1} \langle \vartheta_t, f_1 \rangle + \hat{m}_{n,2}(t) \langle \vartheta_t, f_2 \rangle \quad \text{при} \quad X_{t-} = e_n, \quad (42)$$

где $\hat{m}_{n,1}$ — стратегия (41) для «хорошего» состояния канала f_1 , а $\hat{m}_{n,2}(t)$ — стратегия для «плохого» состояния f_2 (см. рис. 3). Каждая из этих двух стратегий берется с весом, равным апостериорной вероятности соответствующего состояния $\langle \vartheta_t, f_k \rangle = P\{\theta_t = f_k \mid \mathcal{X}_t\}$, $k = 1, 2$. Следовательно, управление (42) однозначно определяется одной из этих вероятностей. На рис. 6 изображены траектории индикатора «плохого» состояния $\langle \theta_t, f_2 \rangle$ и его оценки $\langle \vartheta_t, f_2 \rangle$.

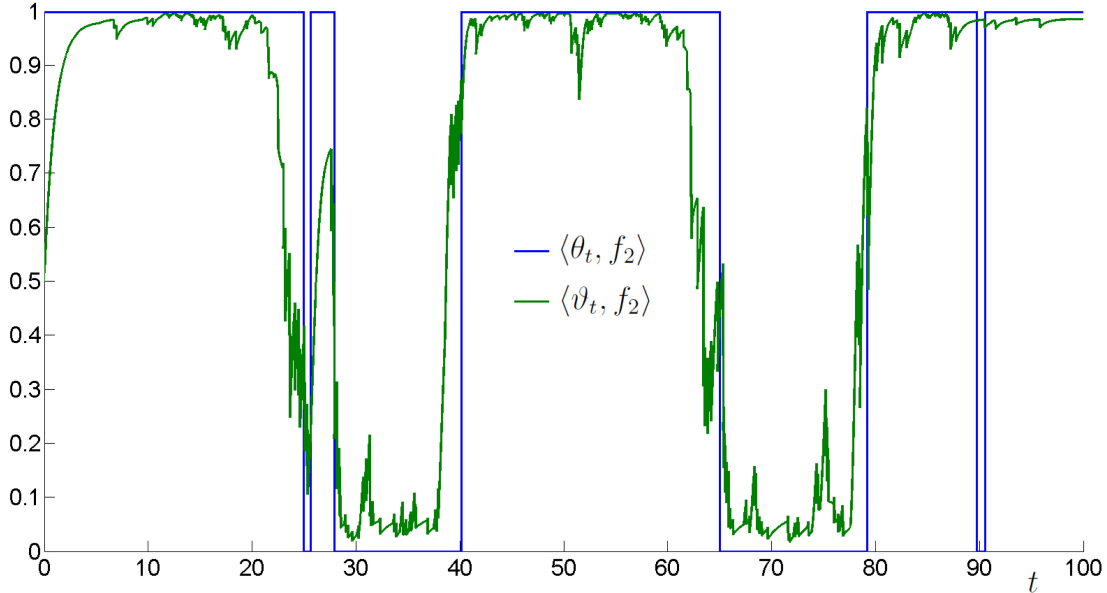


Рис. 6. Апостериорная вероятность «плохого» состояния канала $\langle \vartheta_t, f_2 \rangle$ в сравнении с индикатором этого состояния $\langle \theta_t, f_2 \rangle$

Для определения качества управления, использующего оценку фильтрации, понадобятся статистические оценки функционалов (3), (4). Сначала рассмотрим их потраекторные оценки

$$\begin{aligned} \mathcal{J}_0^{(j)}[\mu] &= \int_0^T (\mathbb{I}\{X_t^{(j)} = N\} \alpha(t) + \mathbb{I}\{X_t > 0\} \langle \ell(t), \vartheta_t^{(j)} \rangle \mu_t^{(j)}) dt, \\ \mathcal{J}_1^{(j)}[\mu] &= \int_0^T w \langle \nu(t), \vartheta_t^{(j)} \rangle (\exp\{H \mu_t^{(j)} / w\} - 1) dt, \end{aligned}$$

вычисляемые по одной траектории $X_t^{(j)}$ процесса X_t , где $\vartheta_t^{(j)}$ и $\mu_t^{(j)}$ — соответствующие реализации оценки фильтрации ϑ_t и управления μ_t . Если в результате моделирования получен пучок из M траекторий, то для функционалов можно определить выборочные оценки

$$\mathcal{J}_0[\mu] = \frac{1}{M} \sum_{j=1}^M \mathcal{J}_0^{(j)}[\mu], \quad \mathcal{J}_1[\mu] = \frac{1}{M} \sum_{j=1}^M \mathcal{J}_1^{(j)}[\mu].$$

Определим управление $\mu^{\text{f.e.}}(\lambda)$ аналогично (42) с заменой $\hat{m}_{n,k}(t)$ на стратегии, определяющие оптимальное управление в расширенной задаче $\tilde{\mu}(\lambda)$. Рис. 7 дает представление о качестве

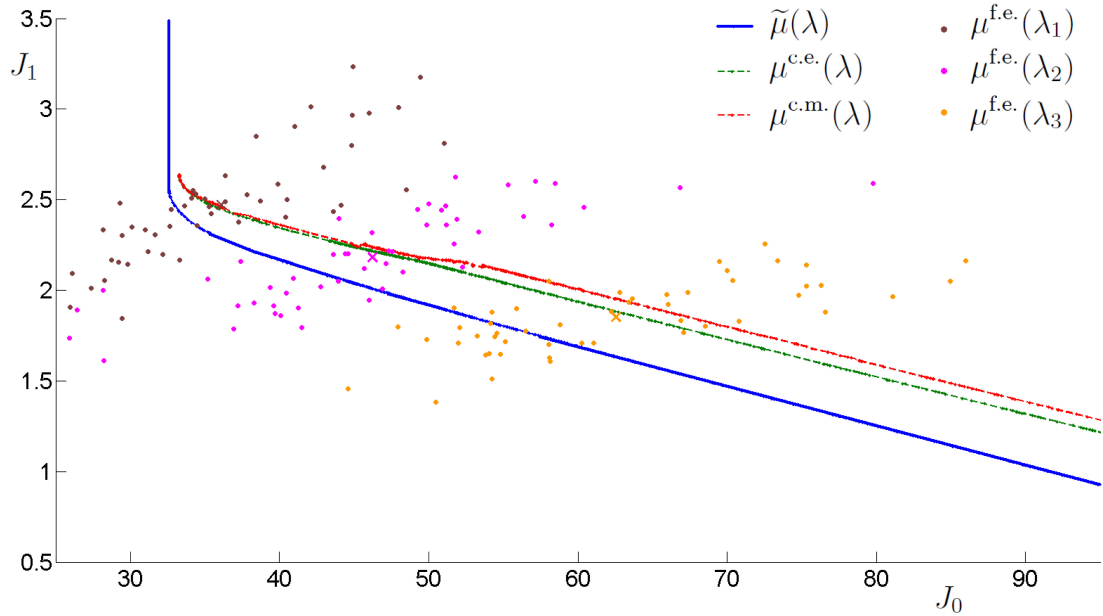


Рис. 7. Значения функционалов на семействах управлений $\tilde{\mu}(\lambda)$, $\mu^{c.e.}(\lambda)$, $\mu^{c.m.}(\lambda)$ (кривые) в сравнении с потраекторными оценками (точки) и выборочными оценками (крестики), полученными на управлении $\mu^{f.e.}(\lambda)$ при $\lambda = \lambda_1, \lambda_2, \lambda_3$

f.e.-управления в сравнении с рассмотренными ранее стратегиями $\tilde{\mu}(\lambda)$, $\mu^{c.e.}(\lambda)$, $\mu^{c.m.}(\lambda)$. Помимо кривых, соответствующих этим параметрическим семействам, на рис. 7 изображены статистические оценки функционалов, вычисленные на управлении $\mu^{f.e.}(\lambda)$ для трех значений параметра λ : $\lambda_1 = 20$, $\lambda_2 = 41$, $\lambda_3 = 45$. Потраекторные оценки представлены $M = 50$ точками, а выборочные оценки — крестиками.

Сильный разброс по J_0 объясняется тем, что переполнение очереди происходит достаточно редко, а вес соответствующего слагаемого $I\{X_t = N\}$, наоборот, значителен. При ужесточении требований на энергию сигнала, т.е. при увеличении параметра λ от λ_1 к λ_3 , среднее число потерь возрастает. Эта закономерность близка к зависимости с.е.- и с.м.-управлений, определяемых только текущим состоянием процесса X_t .

Чтобы сравнить качество различных стратегий при наличии активного ограничения $J_1[\mu] = \bar{J}_1$, была взята верхняя граница, совпадающая с выборочной оценкой функционала $\bar{J}_1 = \mathcal{J}_1[\mu^{f.e.}] = 1,85$ Дж на управлении $\mu^{f.e.} = \mu^{f.e.}(\lambda_3)$. В результате получились следующие значения:

$$J_0[\tilde{\mu}] = 52,8, \quad \mathcal{J}_0[\mu^{f.e.}] = 62,6, \quad J_0[\mu^{c.e.}] = 64,2, \quad J_0[\mu^{c.m.}] = 67,5.$$

Каждое из этих управлений определяется своим значением параметра λ , поэтому зависимость от него опущена. Серьезный выигрыш по числу потерь имеет управление $\tilde{\mu}$, которое использует полную информацию X_t, θ_t . В случае ненаблюдаемого состояния канала связи θ_t управление $\mu^{f.e.}$, использующее предысторию состояния СМО \mathcal{X}_t , имеет лишь небольшое преимущество над стратегиями $\mu^{c.e.}$ и $\mu^{c.m.}$, зависящими от текущего состояния X_t . То же соотношение сохраняется и при других вариантах границы \bar{J}_1 . Отсутствие решающего превосходства стратегии типа f.e. над с.е. или с.м. является следствием достаточно точной априорной информации о состоянии канала.

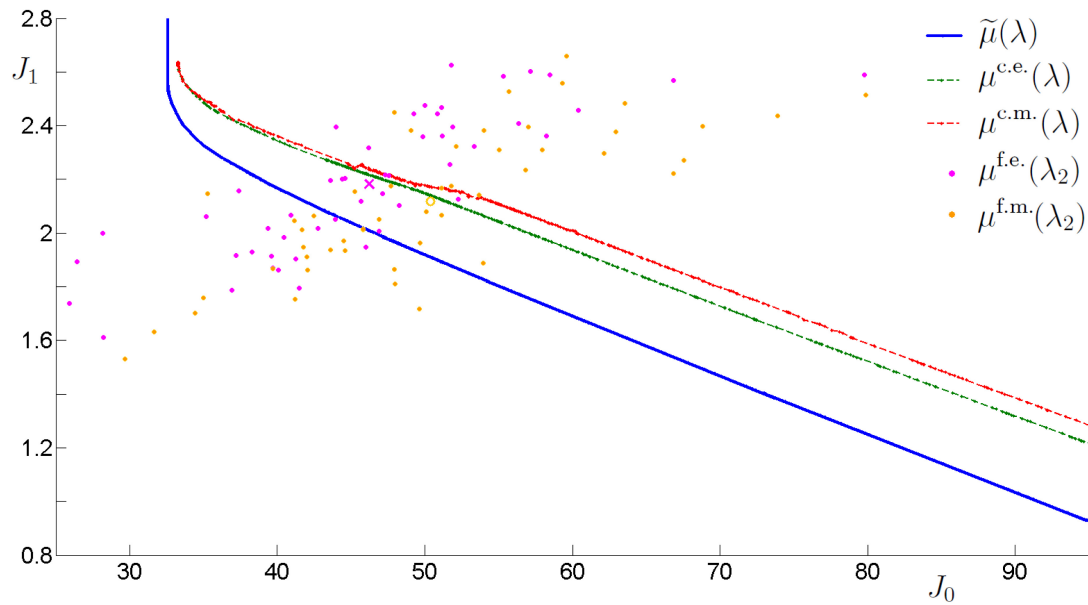


Рис. 8. Значения функционалов на семействах управлений $\tilde{\mu}(\lambda)$, $\mu^{c.e.}(\lambda)$, $\mu^{c.m.}(\lambda)$ (кривые) в сравнении с потраекторными оценками (точки) и выборочными оценками (крестик и кружочек), полученными на управлении $\mu^{f.e.}(\lambda_2)$ и $\mu^{f.m.}(\lambda_2)$

На рис. 8 представлены результаты сравнения f.e.- и f.m.-управлений, построенных на основе одного и того же управления по полной информации $\tilde{\mu}(\lambda)$ при $\lambda = \lambda_2$. Точки определяют потраекторные оценки функционалов, а крестик и кружочек — выборочные оценки на управлениях $\mu_t^{f.e.}(\lambda_2)$ и $\mu_t^{f.m.}(\lambda_2)$ соответственно. С точки зрения минимизации потерь более выгодной операцией оказывается усреднение, применяемое в f.e.-управлении. Однако используемый в f.m.-стратегии максимум апостериорной вероятности приводит к более экономному расходу энергии.

6. ЗАКЛЮЧЕНИЕ

В статье исследована задача оптимизации передачи данных (однотипных пакетов) по каналу связи, состояние которого подвержено ненаблюдаемым случайным изменениям. Для моделирования процесса передачи данных используется одноканальная конечная СМО с управляемой интенсивностью обработки, зависящей от состояния канала. Управление выбирается из условия минимума среднего числа потерь пакетов с учетом ограничения на энергию сигнала на конечном промежутке времени. Для синтеза управления по полной информации определены явные выражения, описывающие оптимальную стратегию в задаче с расширенным функционалом. На основе полученного решения предложено несколько способов формирования управления по неполной информации. Соответствующие стратегии задействуют два варианта оценки ненаблюдаемого состояния канала по загрузке СМО (оценка фильтрации и оценка по текущему состоянию очереди) и две операции над апостериорным распределением (усреднение и мода). Результаты численного моделирования показывают, что управление по полной информации значительно лучше стратегий, использующих лишь частичную информацию о состоянии сети передачи данных. Если сравнить варианты управления по неполной информации, то можно утверждать, что использование предыстории в виде оценки фильтрации дает лишь небольшое преимущество над стратегиями, основанными на текущем состоянии СМО. Кроме

того, с точки зрения минимизации числа потерь усреднение апостериорного распределения оказывается предпочтительнее взятия наиболее правдоподобного значения.

СПИСОК ЛИТЕРАТУРЫ

1. Борисов А. В. Применение алгоритмов оптимальной фильтрации для решения задачи мониторинга доступности удаленного сервера. *Информ. и ее примен.* 2014, т. 8, № 3, стр. 34–50.
2. Борисов А. В., Босов А. В., Миллер Г. Б. Моделирование и мониторинг состояния VoIP-соединения. *Информ. и ее примен.* 2016, т. 10, № 2, стр. 2–13.
3. Welzl M. *Network congestion control: Managing Internet traffic*. New York: J. Wiley & Sons, 2005.
4. Миллер Б. М., Авраченко К. Е., Степанян К. В., Миллер Г. Б. Задача оптимального стохастического управления потоком данных по неполной информации. *Проблемы передачи информации*. 2005, т. 41, № 2, стр. 89–110.
5. Миллер Б. М., Миллер Г. Б., Семенихин К. В. Оптимизация выбора каналов связи при передаче потока данных с учетом потерь. *Автоматика и телемеханика*. 2018, № 1, стр. 84–99.
6. Rieder U., Winter J. Optimal control of Markovian jump processes with partial information and applications to a parallel queueing model. *Math. Meth. Oper. Res.* 2009, vol. 70, pp. 567–596.
7. Кузнецов Н. А., Мясников Д. В., Семенихин К. В. Оптимизация двухфазной системы массового обслуживания и ее применение к управлению передачей данных между двумя агентами робототехнической системы. *Информационные процессы*. 2017, т. 17, № 1, стр. 19–42.
8. Hordijk A., Spieksma F. Constrained admission control to a queueing system. *Advances in Applied Probability*. 1989, vol. 21, no. 2, pp. 409–431.
9. Kitaev M. Y., Rykov V. V. *Controlled queueing systems*. Boca Raton: FL: CRC, 1995.
10. Altman E. *Constrained Markov decision processes*. Boca Raton, FL: Chapman & Hall/CRC, 1999.
11. Elliott R. J., Aggoun L., Moore J. B. *Hidden Markov models. Estimation and control*. New York: Springer-Verlag, 2008.
12. Миллер Б. М., Миллер Г. Б., Семенихин К. В. Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений. *Автоматика и телемеханика*. 2011, № 2, стр. 111–130.
13. Ho D.-T., Grotli E. I., Sujit P. B., Johansen T. A., Sousa J. B. Optimization of wireless sensor network and UAV data acquisition. *J. Intelligent & Robotic Systems*. 2015, vol. 78, no. 1, pp. 159–179.
14. Miller B., Miller G., Semnikhin K. Optimization of the data transmission flow from moving object to nonhomogeneous network of base stations. *Proc. 20th IFAC World Congress (IFAC'2017)*. Toulouse, France: 2017, pp. 6334–6339.
15. Cover T. M., Thomas J. A. *Elements of Information Theory*. New York: J. Wiley & Sons, 1991.
16. Rappaport T. S. *Wireless Communications: Principles and Practice. 2nd edition*. Upper Saddle River: Prentice Hall, 2002.
17. Миллер Б. М., Миллер Г. Б., Семенихин К. В. Регуляризация задачи оптимального управления марковским процессом с конечным числом состояний при наличии ограничений. *Автоматика и телемеханика*. 2016, № 9, стр. 96–123.

Optimal Data Transmission Control through a Fluctuating Channel with Unknown State

N. A. Kuznetsov, D. V. Myasnikov, K. V. Semenikhin

We consider optimization problem for data packets transmission through a communication channel governed by a hidden Markov process. The transmitter is modelled as a single-channel finite-buffer queuing system with non-stationary Poisson arrivals. The service rate is proportional to the controlled transmission rate with channel-dependent factor. Buffer overflow leads to packet losses, whereas channel state worsening results in lower service rate. The goal of the optimization problem is to minimize average losses under constraint on the transmitter energy consumption. The exact form of the optimal policy is presented for the augmented control problem. Several control policies with incomplete information are proposed on the basis of the optimal control and hidden state estimates. We consider two estimators based on the optimal filtering equations and the current queue state. Computer simulation results are presented to compare the control policies under consideration.

KEYWORDS: queuing system, data transmission, optimal control, augmented problem, control with incomplete information, filtering estimate.