ABOUT CANONICAL REPRESENTATION OF THE BELLMAN FUNCTION IN AN OPTIMAL CONTROL AND MONITORING SYSTEMS FOR DYNAMIC NETWORKS

Yu. V. Solodyannikov
[1] SJC "Samara-Dialog", Samara, Russia

For dynamic network queues are posing problems of optimal control of complete and incomplete data, consider some types of functional, stated limitations on the state and management. Recorded Bellman equation for Markov network management tasks for complete data. The main result is to obtain a canonical representation of the solution of the Bellman equation for total Markov homogeneous Markov network management strategy. Examples of design problems of optimal control and optimal networking of information structures for some of modern telecommunications systems.

О КАНОНИЧЕСКОМ ПРЕДСТАВЛЕНИИ ФУНКЦИИ БЕЛЛМАНА В ЗАДАЧАХ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ И НАБЛЮДЕНИЯ ДЛЯ ДИНАМИЧЕСКИХ СЕТЕЙ

Ю. В. Солодянников[1]
[1] ЗАО "Самара-Диалог", Самара, Россия,
solo-dialog@mail.ru

Аннотация

Для динамической сети очередей даются постановки задач оптимального управления по полным и неполным данным, рассмотрены некоторые типы функционалов, сформулированы ограничения на состояния и управления. Записано уравнения Беллмана для задачи управления марковской сетью по полным данным. Основным результатом является получение канонического представления решения уравнения Беллмана для общей марковской сети с однородной марковской стратегией управления. Приводятся примеры решения задач синтеза оптимальных управлений и оптимальных сетевых информационных структур для некоторых современных систем телекоммуникаций.

1. Введение

В работе содержится обзор некоторых математических методов синтеза оптимальных управлений в марковских системах (СМО) и сетях (СеМО) массового обслуживания в условиях полных и неполных данных. Основные результаты были полученны ранее и изложены в [1, 2].

Задача ставится и решается как построение совместного управления сетевыми состоянием и наблюдениями, т. е. путем ответа на вопросы: что, где, когда и как измерять на сети при динамической маршрутизации.

2. Задача оптимального управления по полным данным на марковской сети

Вероятностная модель случайных процессов в сети задается на вероятностном пространстве $\langle \Omega, \mathfrak{F}, \mathbf{P} \rangle$ с потоком σ-алгебр $\mathfrak{F} = \{\mathfrak{F}_t\}_{t \geqslant 0}$.

Входные потоки требований описываются последовательностью $\{\tau^{ij}_{n_{ij}}\}$ моментов прихода на узел i с узла j. Точечные процессы моментов прихода считаем независимыми в совокупности пуассоновскими потоками. Пусть λ^j_i - интенсивность пуассоновского потока требований, назначенных узлу j.

Каждое требование, находящееся в сети в данный момент времени t, характеризуется парой $(i, j) : i \in I, j \in I$, i - узел, на котором находится требование в данный момент, j - узел, которому требование предназначено. Требование, передающееся по каналу (i, k), до момента конца передачи считается находящимся на узле i.

X^j_{it} - случайная величина длины очереди (в числе требований) на узле i с назначением в узел j в момент времени t.

Кроме пары (i, j) для характеризации типа очереди используется индекс предыстории $\overline{p} = (p_1, \ldots, p_L)$, где p_k - номера узлов, которые проходило требование, причем $p_L = i$. Уровень запоминания предыстории L может быть различным. Если действует ограничение $L = 1$(простая предыстория), то запоминается только узел, на котором требование находится. Если $L = 2$, то запоминается узел i (на котором находится требование) и узел, с которого данное требование пришло. Если $L \leq \infty$, то запоминается вся предыстория до уровня L.

Предполагается, что траектории случайных процессов $X^j_{\overline{p}t}$ непрерывны справа и имеют конечные пределы слева (класс D). Сделанные предположения относительно $X^j_{\overline{p}t}$ означают,что $X^j_{\overline{p}t}$ - неотрицательные скачкообразные процессы с траекториями класса D.

Обслуживание в каналах описывается в терминах точечных процессов [3]. Время обслуживания требований в канале (i, j) является случайной величиной, распределенной экспоненциально со средним значением $1/\mu_{ij}$. Времена обслуживания одного требования в различных каналах независимы. Величина μ_{ij} является характеристикой канала и называется его пропускной способностью.

Задача оптимального управления ставится в классической форме. Имеется функционал вида

$$\Phi(u) = \mathbf{E}^u \left\{ c_0(T, X_T) + \int\limits_0^T c(t, X_t, u_t) dt \right\} \ , \tag{1}$$

где $\mathbf{E}^u$ - символ математического ожидания, $c_0(\cdot)$ и $c(\cdot)$ - непрерывные неотрицательные функции своих аргументов. Функция $c(\cdot)$ называется функцией цены, функционал (1) - функционалом стоимости, а время T — горизонтом оптимального управления.

Задача оптимального управления состоит в нахождении управляющего процесса u_t, реализации которого удовлетворяют некоторой системе ограничений $u_t \in U$ и который доставляет минимум (или максимум) функционалу (1).

3. Уравнение Беллмана

Функция Беллмана $V(t, x)$, называемая иногда также функцией цены, определяется как

$$V(t, x) = \min_{\{u_s \in U\}_{t \le s < T}} \mathbf{E}^u \left\{ c_0(T, X_T) + \int\limits_t^T c(s, X_s, u_s) ds \mid X_{t-} = x \right\} \ . \tag{2}$$

Вывод уравнения Беллмана для марковской СеМО имеется в [4], а само это уравнение в случае простой предыстории имеет вид:

$$-\frac{\partial V}{\partial t}(t, x) = \min_{u \in U} \{ c(t, x, u) + \sum_{i \in I} \sum_{j \in I} u_{0i}^j \lambda_i^j [V(t, x + e_{(i)}^j) - V(t, x)] 1(Q_i < N) +$$

$$+ \sum_{k \in I} \sum_{i \in I, i \ne k} \sum_{j \in D^-(i), j \ne i} u_{ij}^k [V(t, x + e_j^k - e_i^k) - V(t, x)] 1((Q_j < N_j) \wedge (x_i^k > 0)) +$$

$$+ \sum_{k \in I} u_{k0}^k [V(t, x - e_k^k) - V(t, x)] 1(x_k^k > 0) \}. \tag{3}$$

В уравнении (3) $e_{\overline{p}}^k$ - совокупность величин, индексированных соответствующим множеством индексов, причем $e_{\overline{p}}^k = 1$ для индексов $\overline{p}, k$ и $e_{\overline{p}}^k = 0$ для остальных.

Начальные условия уравнений для (3):

$$V(T, x) \equiv c_0(T, x) \ . \tag{4}$$

4. О каноническом представлении функции цены

Идея о существовании некого канонического представления для функции цены, являющейся решением задачи (3)-(4), впервые изложена в работах автора [3] и [4]. Существование такого представления, во-первых, ограничивает поиск решения уравнения Беллмана некоторым узким классом функций, во-вторых, позволяет делать выводы о свойствах (в частности, асимптотических) функции цены. Во многих случаях знание канонического представления позволяет получить функцию цены в явном виде, не прибегая к непосредственному решению уравнения Беллмана.

Сформулируем результат для модели с простой предысторией и однородной стратегией управления в виде следующего утверждения.

Теорема 1. *Пусть $\hat{u}_t = \hat{u}(X_t)$ — оптимальная однородная марковская стратегия, отвечающая минимуму функционала накопленной задержки*

$$\Phi(u) = \mathbf{E}^u \int\limits_0^T \sum_{i\in I} \sum_{j\in I} X_{it}^j dt$$

в задаче оптимального управления на марковской сети с простой предысторией. Тогда функция Беллмана, являющаяся решением задачи Коши (3)-(4), имеет представление

$$\begin{aligned} V(t,x) = \tau \sum_{i\in I}\sum_{j\in I} x_i^j + \sum_{i\in I}\sum_{j\in I} \lambda_i^j \hat{u}_{0i}^j(x)[\tau \circledast p_i^{N_i}(\tau,x)] - \\ - \sum_{j\in I} \hat{u}_{j0}^j(x)[\tau \circledast p_j^{+}(\tau,x)]\ , \end{aligned} \tag{5}$$

где $p_i^{N_i}(\eta,x) = \mathbf{P}\{\sum\limits_{j\in I} x_{i\eta}^j < N_i \mid X_0 = x,\ u_\eta = \hat{u}(x_\eta)\}$, $p_j^{+}(\eta,x) = \mathbf{P}\{x_{j\eta}^j > 0 \mid X_0 = x,\ u_\eta = \hat{u}(x_\eta)\}$, $\tau = T - t$.

Рассуждения, позволяющие получить каноническое представление функции цены для марковской сети и функционала накопленной задержки, аналогичны рассуждениям при выводе формулы Литтла для произвольной СМО [5]. Доказательство теоремы 1 имеется в [1].

Иногда представление, аналогичное (5), можно получить и для случаев, когда оптимальная марковская стратегия не является однородной. Как правило, упрощения в таких случаях определяются ограничениями, накладываемыми на управления в соответствии со спецификой задачи.

5. Жидкостная аналогия

Рассмотрим жидкостную модель СМО с управляемым выходным потоком. Пусть имеем воронку бесконечной емкости с регулятором, который управляет потоком, вытекающим из воронки. В воронку заливается жидкость с постоянной скоростью $\dfrac{da(t)}{dt} = \lambda > 0$, а вытекает со скоростью $\dfrac{db(t)}{dt} = u(t)$. Будем предполагать, что скорость истечения ограничена интервалом $0 \leqslant u(t) \leqslant \mu$, где $\mu > 0$. Ограничимся рассмотрением "эргодического" случая с $\lambda/\mu = \rho < 1$.

Пусть в начальный момент времени $t = 0$ в воронке находился начальный объем жидкости $x_0 \geqslant 0$. Общее количество $b(t)$ вытекшей за время t жидкости не может превысить суммы поступившего количества $a(t)$ и начального количества x_0. Общее количество жидкости, содержащейся в воронке в момент t, определяется дифференциальным уравнением

$$\frac{dx(t)}{dt} = \lambda - u(t)$$

с начальным условием $x(0) = x_0$. Очевидно, $x(t) \geqslant 0$, и поскольку скорость истечения жидкости из воронки при $x(t) = 0$ не может превышать скорости притока λ, то управляемая скорость истечения должна подчиняться системе ограничений

$$u(t) \in U(x, \lambda, \mu) = \begin{cases} [0, \mu], \ x(t) > 0 \ , \\ [0, \lambda], \ x(t) = 0 \ . \end{cases}$$

Поставим задачу оптимального управления, состоящую в отыскании оптимальной стратегии управления $\tilde{u}(t, x(t))$, подчиняющейся указанной системе ограничений и доставляющей минимум функционалу

$$\tilde{J}_0 = \int\limits_0^T x(t)dt \ .$$

Функция Беллмана, определяемая как

$$\tilde{V}(t, x) = \min_{u \in U(x,\lambda,\mu)} \{ \int\limits_t^T x(s)ds \mid x(t) = x \} \ , \tag{6}$$

является решением уравнения Беллмана

$$-\frac{\partial \tilde{V}}{\partial t}(t, x) = \min_{u \in U(x,\lambda,\mu)} \{ x + \lambda \frac{\partial \tilde{V}}{\partial x}(t, x) - u \frac{\partial \tilde{V}}{\partial x}(t, x) \}$$

с начальным условием

$$\tilde{V}(T, x) \equiv 0 \ .$$

Теорема 2. *При $\rho < 1$ оптимальное управление в задаче не зависит от времени и имеет представление*

$$\tilde{u}(t, x) = \lambda 1(x = 0) + \mu 1(x > 0) = \begin{cases} \lambda \ , \ x = 0 \ , \\ \mu \ , \ x > 0, \end{cases}$$

а функция цены имеет представление:

$$\tilde{V}(t,x) = x\tau + \frac{\lambda\tau^2}{2} - [\lambda 1(x=0) + \mu 1(x>0)]\tau \circledast \left[1 - \mathfrak{u}_{-1}\left(\tau - \frac{x}{\mu-\lambda}\right)\right] \ . \quad (7)$$

Заметим, что формула (7) является "жидкостным" вариантом канонического представления (5) и легко получается из определения функции цены (6) и расчета количеств поступившей и вытекшей жидкости. Такой подход с использованием канонического представления функции цены может быть применен и к другим самым различным объектам сетевой структуры живой и неживой природы, например управление водохранилищами, управление кровообращением и др.

6. Пример синтеза оптимального управления роутером

Рассмотрим стационарный пуассоновский поток требований интенсивности $\lambda > 0$. Требования обслуживаются пулом из n приборов, перед каждым из которых допускается неограниченная очередь. В момент поступления каждое из требований некоторым устройством с вероятностью $u_i \in [0,1], \sum\limits_{i=1}^{n} u_i = 1$ направляется в очередь i-го прибора. Время обслуживания на i - м приборе экспоненциально распределенно с интенсивностью обслуживания $\mu_i > 0$.

Обозначим $X_t = (X_{1t}, \dots, X_{nt})$ n-компонентный скачкообразный процесс длин очередей с множеством состояний $x = (x_1, \dots, x_n)$, $x \in [\mathbb{Z}_0^\infty]^n$. При сделанных предположениях каждый из приборов роутера в отдельности представляет собой СМО с ограничением нагрузки, а процессы X_{it} есть управляемые марковские процессы.

Поставим задачу нахождения оптимальной марковской стратегии управления $\hat{u}(t,x) = \hat{u}(t, X_t) \mid X_t = x$, доставляющей минимум функционалу накопленной задержки.

$$J(u) = \mathbf{E}^u \int_0^T \sum_{i=1}^{n} X_{it} dt \ . \quad (8)$$

Уравнение Беллмана (3) для данного случая имеет вид

$$-\frac{\partial V}{\partial t}(t,x) = \min_{u \in U} \left\{ \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \lambda u_i [V(t, x+e_i) - V(t,x)] + \right.$$
$$\left. + \sum_{i=1}^{n} \mu_i [V(t, x-e_i) - V(t,x)] 1(x_i > 0) \right\} \quad (9)$$

с начальным условием $V(T,x) \equiv 0$ и множеством ограничений

$$U = \left\{ u_i \in [0,1], \sum_{i=1}^{n} u_i = 1 \right\} \ . \quad (10)$$

Выражение для оптимальной стратегии управления через функцию цены находится исходя из линейности по u_i выражения под знаком минимума в (9) и линейности системы ограничений (10). Обозначим через $J_0(t,x) = \{j : j = \arg\min_{1\le i\le n} \Delta_i^+(t,x)\}$ заведомо непустое множество индексов i, для которых достигается минимальное значение разностных производных функции цены $\Delta_i^+(\tau,x) = V(t,x+e_i) - V(t,x)$ в точке (τ,x), $\tau = T-t$, а через $j_0(t,x)$ — произвольно выбранный элемент множества $J_0(t,x)$. Оптимальное управление записывается в виде

$$\hat{u}_i(t,x) = 1(i = j_0(t,x)) \ .$$

При получении явного выражения для функции цены используем ее каноническое представление определенное теоремой 1. Поскольку в векторе оптимального управления согласно (10) только один компонент (с индексом $j_0(x)$) равен 1, а остальные равны 0, то выражение для функции цены при замене переменных $\tau = T-t$, $z = \theta - t$, $y = s-t$ примет вид

$$V(t,x) = \tau \sum_{j=1}^{n} x_j + \lambda \frac{\tau^2}{2} - \sum_{j=1}^{n} \mu_j [\tau \circledast p_j^+(\tau,x)] \ ,$$

где $p_j^+(\eta,x) = \mathbf{P}\{x_{j\eta} > 0 \mid x_0 = x,\ u_\eta = \hat{u}(\eta,x_\eta)\}$ — зависящие от оптимального управления вероятности ненулевых длин очередей в момент времени η при условии, что в нулевой момент времени вектор длин очередей был равен x.

Для получения замкнутого выражения для функции цены осталось записать формулы для $p_j^+(\tau,x)$. А именно, для $j = j_0(x)$ будем иметь

$$p_j^+(\tau,x) = 1 - e^{-(\lambda+\mu_j)\tau}[\rho_j^{-x_j/2} I_{x_j}(2\tau\sqrt{\lambda\mu_j}) +$$

$$+ \rho_j^{-(x_j+1)/2} I_{x_j+1}(2\tau\sqrt{\lambda\mu_j}) + (1-\rho_j) \sum_{i=x_j+2}^{\infty} \rho_j^{-i/2} I_i(2\tau\sqrt{\lambda\mu_j})] \ ,$$

где $\rho_j = \lambda/\mu_j$. Для $j \ne j_0(x)$ будем иметь

$$p_j^+(\tau,x) = 1 - \sum_{i=x_j}^{\infty} \frac{(\mu_j\tau)^i}{i!} e^{-\mu_j\tau} \ .$$

Переходя к этапу синтеза оптимального управления, запишем выражения для разностных производных функции цены $\Delta_j^+(\tau,x)$. Обозначим через $F_k(s,a,b)$ и $f_k(s,a,b)$ соответственно функцию распределения и плотность распределения обобщенного интервала занятости в СМО $M/M/1$ с параметром входного потока a и параметром времени обслуживания b. Тогда для $j = j_0(x)$ будем иметь

$$\Delta_j^+(\tau,x) = \int_0^{\tau} (1 - F_{x_j}(s,\lambda,\mu_j)) ds \ ,$$

а для $j \neq j_0(x)$

$$\Delta_j^+(\tau, x) = \int\limits_0^\tau (1 - F_{x_j}(s, 0, \mu_j))ds \ .$$

Плотность распределения $f_k(s, a, b)$ для $a > 0$ известна:

$$f_k(s, a, b) = e^{-(a+b)s}(k+1)s^{-1}(a/b)^{-(k+1)/2} I_{k+1}(2s\sqrt{ab}) \ ,$$

а для $a = 0$ она очевидно выражается формулой

$$f_k(s, 0, b) = \frac{(bs)^k}{k!} b e^{-bs} \ .$$

Оптимальная стратегия управления состоит в направлении пришедшего в момент времени t из внешней среды требования в очередь к прибору, для которого минимально значение разностной производной $\Delta_j^+(\tau, x)$ при условии, что вектор длин очередей в этот момент точно известен и равен x. Эта оптимальная стратегия, являясь марковской, вообще говоря, не является однородной. Однако, исходя из асимптотических свойств $\Delta_j^+(\tau, x)$, при больших τ можно показать, что оптимальные управления зависят от времени лишь вблизи горизонта оптимального управления, т. е. для значений t, близких к T.

7. Заключение

В работе рассмотрены актуальные вопросы совместного оптимального управления и наблюдения для управляемых объектов сетевой структуры. При этом учитываются особенности скачкообразных процессов как в измерениях, так и в управлении. Такая общая постановка задачи управления и наблюдения применительно к СМО и СеМО позволяет с единых позиций формулировать и решать соответствующие оптимизационные задачи.

ЛИТЕРАТУРА

1. Солодянников Ю.В. Управление и наблюдение для динамических сетей массового обслуживания. I // АиТ. 2014. Т. 75. No.3 С. 14-45.
2. Солодянников Ю.В. Управление и наблюдение для динамических сетей массового обслуживания. II // АиТ. 2014. Т. 75. No.5 С. 91-114.
3. Солодянников Ю.В. Некоторые вопросы сетеметрии//Теоретические проблемы вычислительных сетей. Научный Совет по комплексной проблеме "Кибернетика" АН СССР, Куйбыш. гос. ун-т. Куйбышев, 1986. С. 74–102.
4. Солодянников Ю.В. О статистике систем и сетей массового обслуживания//Проблемы устойчивости стохастических моделей. Тр. X Всесоюз. семинара. Куйбыш. гос. ун-т. Куйбышев, 1987. С. 101–116.
5. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979.