S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 18, №2, стр. 1493–1505 (2021) УДК 621.394.74
DOI 10.33048/semi.2021.18.112 MSC 60K25

ЭРГОДИЧНОСТЬ ПО ХАРРИСУ РАЗДЕЛЕННОГО ПРОТОКОЛА УПРАВЛЕНИЯ ПЕРЕДАЧЕЙ ИНФОРМАЦИИ

С.Г. ФОСС, М.Г. ЧЕБУНИН

Abstract. Additive-increase multiplicative-decrease transmission control protocols are well known and have been studied in numerous papers. It is significantly more difficult to study systems of interacting protocols. We consider a queueing system where both the input intensity and the service intensity follow TCP protocols and the dynamics of the latter depends on both intensities. This type of stochastic system was proposed by Baccelli, Carofiglio and Foss in 2009, who have proved the positive recurrence of the underlying Markov chain and studied a number of statistical properties of the model. In this paper, we introduce a more general stochastic model and prove a stronger statement: the Harris ergodicity of the corresponding Markov chain.

Keywords: split TCP, Markov chains, Harris ergodicity.

1. Введение, описание модели и основной результат

Имеется большое количество математических и прикладных работ, посвящённых TCP протоколам, а также ряд обзорных публикаций на эту тему, см., например, работу [1] и обширный список литературы в ней. Однако число работ, изучающих системы, в которых функционирует несколько использующих TCP (и, возможно, взаимозависимых) потоков сообщений, очень ограничено, см., например, работу [2] и её список литературы.

Одну из таких систем мы и рассматриваем в данной работе. А именно, мы изучаем т. н. модель «split TCP» («разделенный протокол управления передачей информации»), предложенную в работе [2], в которой рассматривается

Foss, S.G., Chebunin, M.G., Harris ergodicity of a Split Transmission Control Protocol.

Работа выполнена при поддержке РНФ, проект № 17-11-01173-П.

Поступила 13 октября 2021 г., опубликована 2 декабря 2021 г.

3-компонентный марковский процесс в непрерывном времени, первая и вторая компонента которого следуют своим TCP протоколам, а третья компонента описывает накопленную работу, причём первая компонента функционирует автономно, а динамика второй компоненты зависит от первой и третьей компонент. Этот марковский процесс описывает работу открытой системы обслуживания (передачи данных), в которой интенсивность поступления извне требований $X(t)$ задаётся непрерывным слева марковским процессом, удовлетворяющим уравнению

$$dX(t) = adt - kX(t)M(dt), \tag{1}$$

где $a > 0$, $k \in (0,1)$ и $M(t)$ – пуассоновский процесс с параметром $\lambda_1 > 0$, а интенсивность передачи (ухода) сообщений $Y(t)$ зависит от значения накопленной в системе к моменту t незавершенной работы

$$Q(t) = \max\left(\sup_{0\le u\le t}\int_u^t (X(v)-Y(v))dv,\ Q(0)+\int_0^t (X(u)-Y(u))du\right). \tag{2}$$

А именно, если $Q(t) > 0$, то приращение интенсивности $Y(t)$ удовлетворяет уравнению

$$dY(t) = bdt - lY(t)N(dt), \tag{3}$$

а если $Q(t) = 0$, то уравнению

$$dY(t) = b\frac{X(t)}{Y(t)}dt - lY(t)N(dt). \tag{4}$$

Здесь $b > 0$, $l \in (0,1)$ и $N(t)$ – другой пуассоновский процесс с параметром $\lambda_2 > 0$, не зависящий от первого. Предложенная модель допускает такую интепретацию: имеется резервуар, в который жидкость поступает с интенсивностью $X(t)$ и покидает его с интенсивностью $Y(t)$, в зависимости от уровня его наполнения, т.е. величины $Q(t)$. Можно представить $X(t)$ и $Y(t)$ как регулируемые диаметры труб, по которым жидкость, соответственно, поступает в резервуар и покидает его, а $Q(t)$ как объём жидкости в резервуаре в момент времени t.

В работе [2] в случае $k = l = 1/2$ были найдены условия положительной возвратности марковского процесса $(X(t), Y(t), Q(t))$ и отмечено, что хвост распределения $Q(t)$ в стационарном режиме убывает не быстрее чем $e^{-\sqrt{x}}$, т. е. является «тяжелым». [1] Мы рассматриваем более общую ситуацию, в которой $k, l \in (0,1)$, и доказываем не только положительную возвратность марковского процесса, но и сходимость распределения $(X(t), Y(t), Q(t))$ к предельному в метрике полной вариации. А именно, справедлива следующая основная теорема.

Теорема 1. *Если*

$$al\lambda_2 < bk\lambda_1, \tag{5}$$

то существует распределение π в трёхмерном положительном ортанте, к которому распределение процесса $(X(t), Y(t), Q(t)$ сходится в метрике полной вариации, т.е.

$$\sup_A |\mathbb{P}((X(t), Y(t), Q(t)) \in A) - \pi(A)| \to 0 \quad \textit{при} \quad t \to \infty \tag{6}$$

[1] Говорят, что случайная величина ξ имеет тяжелый (правый) хвост распределения, если $\mathbb{E}e^{t\xi} = \infty$ для всех $t > 0$.

при любых конечных начальных значениях $(X(0), Y(0), Q(0))$. *Здесь супремум берётся по всем измеримым по Лебегу множествам A в трехмерном эвклидовом пространстве.*

Основным шагом в доказательстве этой теоремы является получение аналогичного утверждения для вложенной цепи Маркова. Пусть $W(t)$ – суперпозиция процессов $M(t)$ и $N(t)$, т. е. пуассоновский процесс, каждая точка которого независимо от других принадлежит либо процессу $M(t)$ (с вероятностью $\frac{\lambda_1}{\lambda_1+\lambda_2}$), либо процессу $N(t)$ (с вероятностью $\frac{\lambda_2}{\lambda_1+\lambda_2}$), и пусть $0 < T_1 < T_2 < \ldots$ – последовательные точки процесса $W(t)$. Обозначим через $(X_n, Y_n, Q_n) = (X(T_n+0), Y(T_n+0), Q(T_n+0))$, $n \geq 1$ значения процесса $(X(t), Y(t), Q(t))$ в соответственные вложенные моменты времени. Положим также $(X_0, Y_0, Q_0) = (X(0), Y(0), Q(0))$.

Теорема 2. *При выполнении условия* (5) *цепь Маркова* (X_n, Y_n, Q_n) *является эргодичной по Харрису и, в частности, существует вероятностная мера* π^* *такая, что*

$$\sup_A |\mathbb{P}((X_n, Y_n, Q_n) \in A) - \pi^*(A)| \to 0 \quad \textit{при} \quad t \to \infty \tag{7}$$

при любом начальном значении (X_0, Y_0, Q_0).

Замечание 1. *Отметим, что для доказательства теоремы 2 нужно построить, в частности, минорантную меру для трехмерной марковской цепи. И мы пришли к одному занимательному факту, с аналогами которого ранее не встречались. Обычно минорантная мера определяется на пространстве той же размерности, что и сам марковский процесс. Но оказывается, что в этой задаче нагляднее и проще построить минорантную меру в двумерном пространстве, с занулением третьей координаты* $Q(t)$.

Замечание 2. *Приводимое ниже доказательство теоремы 2 оказывается технически достаточно громоздким в рассматриваемом нами случае пуассоновских процессов* $M(t)$ *и* $N(t)$. *Поэтому мы не стремились приводить и доказывать наши утверждения в максимальной общности. Однако следует отметить, что аналогичный результат остаётся верным и при существенно более общих стохастических предположениях, при использовании практически той же схемы доказательства. Более подробно, вместо пуассоновского процесса* $W(t)$ *можно рассмотреть процесс восстановления, у которого распределение длин интервалов времени между моментами скачков имеет конечный второй момент и непрерывную компоненту, плотность распределения которой равномерно отделена от нуля в окрестности начала координат, и (независимо от всего остального) каждый момент скачка принадлежит либо процессу* $M(t)$ *с вероятностью* $p \in (0,1)$, *либо процессу* $N(t)$ *с вероятностью* $1-p$. *Условие* (5) *при этом приобретает следующий вид:*

$$a\left(\frac{1-pk}{pk} + \frac{\mathbb{E}T_1^2}{2(\mathbb{E}T_1)^2}\right) < b\left(\frac{1-(1-p)l}{(1-p)l} + \frac{\mathbb{E}T_1^2}{2(\mathbb{E}T_1)^2}\right).$$

В параграфе 2 мы приводим доказательство теоремы 2, а в параграфе 3 – доказательство теоремы 1. В Приложении собраны вспомогательные факты и утверждения: мы приводим комментарий о необходимости условия (5) для стабильности рассматриваемого марковского процесса, напоминаем определение эргодичности по Харрису, а также формулируем известные достаточные условия для положительной возвратности множества и теорему о сходимости в метрике полной вариации процесса в непрерывном времени, допускающего вложенную цепь Маркова, обладающую требуемыми условиями эргодичности.

2. Доказательство теоремы 2.

Повторим, что (по теореме о суперпозиции независимых пуассоновских процессов) можно считать, что мы наблюдаем общий пуассоновский процесс $W(t)$ с интенсивностью $\lambda_1 + \lambda_2$ и что каждый его скачок с вероятностью $\frac{\lambda_1}{\lambda_1+\lambda_2}$ относится к процессу $M(t)$, а с вероятностью $\frac{\lambda_2}{\lambda_1+\lambda_2}$ – к процессу $N(t)$.

Обозначим через $t_n = T_n - T_{n-1}$ длительности промежутков времени между скачками процесса $W(t)$ и положим $T_0 = 0$. Мы покажем, что последовательность $Z_n = (X_n, Y_n, Q_n)$ является однородной по времени цепью Маркова, которая положительно возвратна по Харрису (мы напоминаем определение этого понятия в Приложении). Всюду далее мы будем использовать стандартные обозначения

$$\mathbb{E}_z(\cdot) = \mathbb{E}(\cdot|Z_0 = z), \quad \text{и} \quad \mathbb{P}_z(\cdot) = \mathbb{P}(\cdot|Z_0 = z).$$

Нам потребуется ряд дополнительных обозначений. Пусть $\widehat{Y}(t)$ – вспомогательный автономный процесс, который следует стандартному TCP протоколу, т. е. процесс, который независимо от всего удовлетворяет уравнению (3) при всех значениях $Q(t)$ при $t \geq 0$. Такой процесс получается, если взять в качестве начального значение $Q(0) = \infty$ – тогда $Q(t) = \infty > 0$ при всех $t > 0$, из чего следует, что $Y(t) = \widehat{Y}(t)$ п. н. при всех $t \geq 0$. Пусть теперь $Y(t)$ – процесс, определяемый уравнениями (3)-(4) при тех же начальных значениях $X(0)$ и $Y(0)$, но при конечном начальном значении $Q(0)$. Тогда нетрудно видеть, что $Y(t) \leq \widehat{Y}(t)$ п. н. для всех $t \geq 0$, т. к. линейная скорость роста автономного процесса всегда не меньше линейной скорости процесса $Y(t)$ (см. уравнения (3) и (4)), поэтому при равных начальных значениях и в силу монотонности автономного процесса по начальному условию (действительно, при разных начальных данных $y_n \leq y'_n$ один процесс всегда мажорирует другой, т. е. $y_n + b(t - T_n) \leq y'_n + b(t - T_n)$ при $t \in (T_n, T_{n+1})$ и $y_{n+1} = (y_n + bt_{n+1})(1 - l) \leq y'_{n+1} = (y'_n + bt_{n+1})(1 - l)$, где $y_n = \widehat{Y}(T_n + 0)$), мы получаем требуемое.

Для упрощения дальнейшего изложения введём следующие вспомогательные параметры:

$$k' = 1 - k, \quad l' = 1 - l,$$
$$\alpha = \frac{k'\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2}, \quad \alpha' = \frac{\alpha}{\lambda_1 + \lambda_2}, \quad \beta = \frac{\lambda_1 + l'\lambda_2}{\lambda_1 + \lambda_2}, \quad \beta' = \frac{\beta}{\lambda_1 + \lambda_2},$$
$$a' = \frac{\alpha' a}{1 - \alpha} = \frac{\alpha a}{k\lambda_1}, \quad b' = \frac{\beta' b}{1 - \beta} = \frac{\beta b}{l\lambda_2},$$
$$C = a' + b', \quad C' = C + \frac{a}{(\lambda_1 + \lambda_2)^2}, \quad N_0 = 2[C' + 1] + 2,$$
$$\delta = b' - a' + (b - a)\frac{\mathbb{E}t_1^2}{2\mathbb{E}t_1} = \frac{b}{l\lambda_2} - \frac{a}{k\lambda_1}.$$

Заметим, что

$$\mathbb{E}_x X_n = \mathbb{E}_x \mathbb{E}(X_n|X_{n-1}) = \frac{k'\lambda_1}{\lambda_1 + \lambda_2}\left(\mathbb{E}_x X_{n-1} + \frac{a}{\lambda_1 + \lambda_2}\right)$$

$$+\frac{\lambda_2}{\lambda_1+\lambda_2}\left(\mathbb{E}_x X_{n-1}+\frac{a}{\lambda_1+\lambda_2}\right)=\alpha\mathbb{E}_x X_{n-1}+\alpha' a$$

$$=\alpha^n x+\alpha' a(1+\alpha+\alpha^2+\cdots+\alpha^{n-1})=\alpha^n x+a'(1-\alpha^n). \tag{8}$$

Аналогично

$$\mathbb{E}_y\widehat{Y}_n=\mathbb{E}_y\mathbb{E}(\widehat{Y}_n|\widehat{Y}_{n-1})=\frac{l'\lambda_2}{\lambda_1+\lambda_2}\left(\mathbb{E}_y\widehat{Y}_{n-1}+\frac{b}{\lambda_1+\lambda_2}\right)$$

$$+\frac{\lambda_1}{\lambda_1+\lambda_2}\left(\mathbb{E}_y\widehat{Y}_{n-1}+\frac{b}{\lambda_1+\lambda_2}\right)=\beta\mathbb{E}_y\widehat{Y}_{n-1}+\beta' b$$

$$=\beta^n y+\beta' b(1+\beta+\beta^2+\cdots+\beta^{n-1})=\beta^n y+b'(1-\beta^n). \tag{9}$$

Также введем следующие вспомогательные процессы:

$$\widehat{Q}(t)=Q_0+\int_0^t (X(u)-\widehat{Y}(u))du,$$

$$Q^X(t)=\int_0^t X(u)du, \qquad Q^{\widehat{Y}}(t)=\int_0^t \widehat{Y}(u)du.$$

Из определения процесса $Q(t)$ (см. (2)) получаем, что

$$Q(t)=\int_{\theta_t}^t (X(u)-Y(u))du+Q_0\mathbf{1}(\theta_t=0),$$

где $\theta_t=\inf(u<t : Q(u)>0\ \forall u\in(\theta_t,t))$. Следовательно $-Q^{\widehat{Y}}(t)\le\widehat{Q}(t)\le Q(t)\le Q_0+Q^X(t)$ п. н. при $t\ge 0$.

Положим $z=(x,y,q)$ и $L(z)=x+y+cq$, где c — некоторая положительная постоянная. Определим множество $B=\{z\ :\ L(z)\le N\}$. Для доказательства положительной возвратности множества B воспользуемся теоремой 1 из [5] (см. Приложение). Заметим, что $\sup_{z\in B}\mathbb{E}_z L(Z_n)<\infty$ для любого фиксированного n. Пусть $z\notin B$ и $N>>N_0$, тогда из (8) и (9) получаем, что

$$\Delta_z(n)=\mathbb{E}_z(L(Z_n)-L(z))\le\mathbb{E}_x(X_n-x)+\mathbb{E}_y(\widehat{Y}_n-y)+c(\mathbb{E}_z Q_n-q)$$

$$\le-(1-\alpha^n)x-(1-\beta^n)y+C+c(\mathbb{E}_z Q_n-q).$$

Выберем $n_0\ge 1$ такое, что $\alpha^n+\frac{1}{n(\lambda_1+\lambda_2)}<1/2$ и $\beta^n<1/2$ при всех $n\ge n_0$. Положим $c=1/n_1^2$ для некоторого $n_1\ge n_0$, тогда для всех $1\le n\le n_1$

$$c(\mathbb{E}_z\widehat{Q}_n-q)\le c(\mathbb{E}_z Q_n-q)\le c\mathbb{E}_z Q_n^X\le\frac{1}{n^2}\mathbb{E}(xT_n+aT_n^2/2)\le\frac{x}{n(\lambda_1+\lambda_2)}+\frac{a}{(\lambda_1+\lambda_2)^2}.$$

Следовательно, при всех $n_0\le n\le n_1$

$$\Delta_z(n)\le-x/2-y/2+C'.$$

Из условия теоремы вытекает, что $\delta>0$ (см. пояснения по условию (5) в Приложении), тогда

1. Если $x\ge N_0$, то

$$\Delta_z(n_0)\le-x/2+C'\le-N_0/2+C'\le-1.$$

2. Если $y\ge N_0$, то

$$\Delta_z(n_0)\le-y/2+C'\le-N_0/2+C'\le-1.$$

3. Если $x < N_0$ и $y < N_0$, то тогда $q > (N-2N_0)n_1^2$. Из (8) и (9) следует, что

$$\sup_{x,y\le N_0}(|\mathbb{E}_x X_n - a'| + |\mathbb{E}_y \widehat{Y}_n - b'|) \le \alpha^n(N_0+a') + \beta^n(N_0+b') = h(n).$$

Так как $h(n)$ монотонно убывает, то можно выбрать n_1 такое, что $h(n_1) \le \delta/2$ для всех $n \ge n_1$. Затем можно выбрать $n_2 > n_1$ такое, что

$$\begin{aligned}
C' + c\mathbb{E}_z(\widehat{Q}_{n_2} - \widehat{Q}_{n_1}) &= C' + c\mathbb{E}_z\left(\int_{T_{n_1}}^{T_{n_2}}(X(u) - \widehat{Y}(u))du\right)\\
&= C' + c\sum_{i=n_1}^{n_2-1}\mathbb{E}_z\left(\int_{T_i}^{T_{i+1}}(X_i - \widehat{Y}_i + (a-b)(u-T_i))du\right)\\
&= C' + c\sum_{i=n_1}^{n_2-1}\mathbb{E}_z\left((X_i - \widehat{Y}_i)t_{i+1} + (a-b)\frac{t_{i+1}^2}{2}\right)\\
&= C' + c\mathbb{E}t_1\cdot\sum_{i=n_1}^{n_2-1}\left((\mathbb{E}_x X_i - \mathbb{E}_y\widehat{Y}_i) + (a-b)\frac{\mathbb{E}t_1^2}{2\mathbb{E}t_1}\right)\\
&\le C' + c\mathbb{E}t_1\cdot\sum_{i=n_1}^{n_2-1}\left((a'-b') + \delta/2 + (a-b)\frac{\mathbb{E}t_1^2}{2\mathbb{E}t_1}\right)\\
&= C' - \frac{\delta(n_2-n_1-1)\mathbb{E}t_1}{2n_1^2} < -2.
\end{aligned}$$

Определим $\eta = \inf\{t \ge 0 : Q(t) = 0\})$ и заметим, что п. н.

$$\begin{aligned}
Q_n &= Q_n(\mathbf{1}(\eta \ge T_n) + \mathbf{1}(\eta < T_n)) = \widehat{Q}_n\mathbf{1}(\eta \ge T_n) + Q_n\mathbf{1}(\eta < T_n)\\
&= \widehat{Q}_n + (Q_n - \widehat{Q}_n)\mathbf{1}(\eta < T_n) \le \widehat{Q}_n + (Q_n^X + Q_n^{\widehat{Y}})\mathbf{1}(\eta < T_n).
\end{aligned}$$

Так как $q > (N-2N_0)n_1^2$, то при $N \to \infty$

$$\begin{aligned}
\mathbb{P}_z(\eta < T_{n_2}) &\le \mathbb{P}_z(Q_{n_2}^{\widehat{Y}} > q) \le \frac{\mathbb{E}_z Q_{n_2}^{\widehat{Y}}}{q} \le \frac{1}{q}\mathbb{E}_z(N_0 T_{n_2} + bT_{n_2}^2/2)\\
&\le \frac{1}{q}\left(\frac{n_2N_0}{\lambda_1+\lambda_2} + \frac{bn_2^2}{(\lambda_1+\lambda_2)^2}\right) \to 0.
\end{aligned}$$

Заметим, что $\mathbb{E}_z((Q_{n_2}^X + Q_{n_2}^{\widehat{Y}})\mathbf{1}(\eta < T_{n_2})) \to 0$ при $N \to \infty$, т. к.

$$\mathbb{E}_z(Q_{n_2}^X + Q_{n_2}^{\widehat{Y}}) \le \frac{2n_2N_0}{\lambda_1+\lambda_2} + \frac{(a+b)n_2^2}{(\lambda_1+\lambda_2)^2}.$$

Следовательно, при достаточно большом N можно утверждать, что

$$\begin{aligned}
\Delta_z(n_2) &= -(1-\alpha^{n_2})x - (1-\beta^{n_2})y + C + c\mathbb{E}_z(Q_{n_2} - q)\\
&\le -(1-\alpha^{n_2})x - (1-\beta^{n_2})y + C + c\mathbb{E}_z(\widehat{Q}_{n_2} \pm \widehat{Q}_{n_1} - q) + c\mathbb{E}_z((Q_{n_2}^X + Q_{n_2}^{\widehat{Y}})\mathbf{1}(\eta < T_{n_2}))\\
&\le -x/2 - y/2 + C' + c\mathbb{E}_z(\widehat{Q}_{n_2} - \widehat{Q}_{n_1}) + c\mathbb{E}_z((Q_{n_2}^X + Q_{n_2}^{\widehat{Y}})\mathbf{1}(\eta < T_{n_2}))\\
&\le -2 + c\mathbb{E}_z((Q_{n_2}^X + Q_{n_2}^{\widehat{Y}})\mathbf{1}(\eta < T_{n_2})) < -1.
\end{aligned}$$

Положим

$$g(z) = n_0\mathbf{1}(x \ge N_0 \vee y \ge N_0) + n_2\mathbf{1}(x < N_0 \wedge y < N_0),$$

тогда при достаточно большом N

$$\Delta_z(g(z)) = \mathbb{E}_z(L(Z_{g(z)}) - L(z)) = \mathbb{E}\mathbb{E}_z(L(Z_{g(z)}) - L(z)|g(z)) < -1.$$

Далее, используя теорему 29 из [6] (см. Приложение), покажем положительную возвратность множества

$$G = [0, \varepsilon_1] \times [v\varepsilon_2, V\varepsilon_2] \times 0,$$

где $V > v > \frac{\varepsilon_1}{\varepsilon_2} > 0$, а ε_1, ε_2 — произвольные (сколь угодно малые) положительные числа. Без ограничения общности мы будем предполагать что $G \subset B$, в противном случае мы всегда можем выбрать другое $N > \varepsilon_1 + V\varepsilon_2$. Положим $\tau_G = \min\{n : Z_n \in G\}$ и найдем такое натуральное число $\mathbf{N}$, что $\inf_{z\in B} \mathbb{P}_z(\tau_G \le \mathbf{N}) > 0$.

Предположим, что процесс стартует из некоторой точки $z \in B$. Тогда с положительной вероятностью (отделённой от нуля равномерно по всем $z \in B$) осуществляется следующая последовательность событий: в течение малого времени происходит некоторое число скачков процесса X, позволяющее ему оставаться в малой окрестности нуля; после этого в течение некоторого равномерно ограниченно времени скачков не поисходит и процесс Q зануляется; затем в течение малого времени происходит некоторое количество скачков процесса Y, после чего его сначение оказывается ниже уровня $V\varepsilon_2$; наконец, в течение некоторого времени скачков опять не происходит и Y попадает в интервал $[v\varepsilon_2, V\varepsilon_2]$.

Опишем эту последовательность действий более подробно. Итак, пусть процесс стартует из некоторой точки $z \in B$. Тогда:

1) Предположим что за $\varepsilon_1/2a$ единиц времени произошло

$$N_1 = [\log_{k'}(\varepsilon_1/2N)] + 1$$

скачков процесса W (здесь $[x]$ — целая часть числа x) и, более того, более того, все эти скачки относились только к процессу X (вероятность этого $(\lambda_1/(\lambda_1 + \lambda_2))^{N_1} > 0$). Тогда в момент времени $\varepsilon_1/2a$ процесс X будет ниже уровня ε_1.

2) Далее, мы будем предполагать, что в течение каждого промежутка времени длины $\varepsilon' = \varepsilon_1 k/a$ будет происходить один скачок процесса W и, более того, он будет относиться к процессу X и происходить во второй половине интервала времени, т. е. в промежутке от $\varepsilon'/2$ до ε'. Таким образом, мы можем добиться того, что $X(t) \in [\varepsilon_1 kk'/2, \varepsilon_1]$ в течение промежутка времени любой требуемой нам длительности.

3) Так как $Q(\varepsilon_1/2a) \le q + x + \varepsilon_1/2 < n_1 N + \varepsilon_1$, то как максимум через $\sqrt{2(n_1 N + \varepsilon_1)/(b - \varepsilon_1)}$ единиц времени можно добиться обнуления процесса Q с использованием скачков процесса X. Далее, предположим, что в течение следующих $\varepsilon' N_2$ единиц времени происходит N_2 процесса X, где

$$N_2 = [\sqrt{2(n_1 N + \varepsilon_1)/(b - \varepsilon_1)}/\varepsilon'] + 1.$$

Тогда в момент времени $\varepsilon_1/2a + \varepsilon' N_2$ процесс будет находиться в множестве

$$[0, \varepsilon_1] \times (\varepsilon_1, N + b(\varepsilon_1/2a + \varepsilon' N_2)] \times 0.$$

4) Далее, при необходимости, мы будем предполагать, что за ε' единиц времени произошло $N_3' + 1$ скачка, один из которых произошел в промежутке от $\varepsilon'/2$ до ε' и относился к процессу X, а все остальные относились к процессу Y, где N_3' определяется как минимальное количество скачков, после которых

процесс Y будет ниже уровня $V\varepsilon_2$ и выше уровня ε_1. Ясно, что в таком случае $N_3' \le N_3$, где

$$N_3 = [\log_{l'}(V\varepsilon_2/(N + b(\varepsilon_1/2a + (N_2+1)\varepsilon'))] + 1.$$

5) После этого, предполагая, что происходят только скачки процесса X, мы подождем конечное время, не превышающее $(v\varepsilon_2)^2/b\varepsilon_1 kk'$, до попадания в множество G. Ясно, что для этого потребуется $N_4' + 1$ скачка процесса X, где первые N_4' скачка происходят за $N_4'\varepsilon'$ единиц времени, а последний – за малое время сразу после попадания процесса Y выше уровня $v\varepsilon_2$, так чтобы Y не успел выйти выше уровня $V\varepsilon_2$. Заметим, что в таком случае $N_4' \le N_4$, где

$$N_4 = [(v\varepsilon_2)^2/b\varepsilon_1 kk'\varepsilon'] + 1.$$

В итоге мы можем утверждать, что если вложенная цепь Маркова стартует из любой точки множества B, то с равномерно положительной вероятностью не более чем за $\mathbf{N} = N_1 + N_2 + N_3 + N_4 + 2$ шагов она окажется во множестве G.

Далее мы построим минорантную меру для цепи Маркова. Стартуя из G, мы будем предполагать, что первый скачок происходит у процесса Y, а второй у процесса X (это событие происходит с вероятностью $\frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} > 0$). Положим $R(t) = Y^2(t)/2$, тогда

$$dR(t) = bX(t)dt - \frac{l}{2}R(t)N(dt).$$

Положим $l'' = 1 - l/2$ и будем предполагать, что $R(t) > X^2(t)/2$ до момента T_2, тогда

$$X(T_2) = (x + aT_2)k', \quad R(T_1) = \left(r + bxT_1 + \frac{abT_1^2}{2}\right)l'',$$

$$R(T_2) = R(T_1) + b(x + aT_1)(T_2 - T_1) + \frac{ab(T_2 - T_1)^2}{2}.$$

Положим $\varepsilon_1 = \varepsilon,\ \varepsilon_2 = \varepsilon\sqrt{b/a},\ v = \sqrt{H-1},\ V = \sqrt{H+1}$, для некоторых $\varepsilon > 0$ и $H \ge 1 + 9a/bl''$. Предположим, что $t_1 = \phi \in [0, \varepsilon/a]$ и $t_2 = \psi \in [0, \varepsilon/a]$, тогда $X^2(t)/2 < 9\varepsilon^2/2 \le R(t)$ и

$$\begin{aligned} X(\phi,\psi,x) &= (x + a(\phi+\psi))k', \\ R(\phi,\psi,x,r) &= \left(r + bx\phi + \frac{ab}{2}\phi^2\right)l'' + b(x + a\phi)\psi + \frac{ab}{2}\psi^2 \\ &= rl'' + bx(\phi+\psi) + \frac{ab}{2}(\phi+\psi)^2 - (bx\phi + \frac{ab}{2}\phi^2)\frac{l}{2}. \end{aligned}$$

Заметим, что введенные выше функции возрастают по всем переменным. Якобиан данного преобразования строго положителен (при фиксированных положительных x и ϕ) и равен $J(\phi) = ablk'(x + a\phi)/2$. Ясно, что совместная плотность вектора (ϕ,ψ) отделена от нуля на множестве $[0,\varepsilon/a]^2$. Положим $(u_1, u_2) \equiv (X(\phi,\psi,x), R(\phi,\psi,x,r))$, тогда

$$\begin{aligned} f_{X,R}(u_1,u_2) &= \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}\frac{f_{t_1,t_2}(\phi,\psi)}{|J(\phi)|} = \frac{2\lambda_1\lambda_2 e^{-(\lambda_1+\lambda_2)(\phi+\psi)}}{ablk'(x+a\phi)} \\ &\ge \frac{\lambda_1\lambda_2 e^{-2(\lambda_1+\lambda_2)\varepsilon/a}}{ablk'\varepsilon} \equiv p > 0, \end{aligned}$$

для всех $(\phi, \psi) \in [0, \varepsilon/a]^2$ и $(x, \sqrt{2r}, 0) \in G$. Определим множество

$$D \equiv [X(0,0,\varepsilon), X(\frac{\varepsilon}{a}, \frac{\varepsilon}{a}, 0)] \times [R(0,0,0, \frac{b\varepsilon^2(H+1)}{2a}), R(\frac{\varepsilon}{a}, \frac{\varepsilon}{a}, \varepsilon, \frac{b\varepsilon^2(H-1)}{2a})]$$
$$= [\varepsilon k', 2\varepsilon k'] \times \left[\frac{b\varepsilon^2(H+1)l''}{2a}, \frac{b\varepsilon^2(H-1)l''}{2a} + \frac{b\varepsilon^2(4-3l/4)}{a}\right].$$

Тогда для любых $(u_1, u_2) \in D$ плотность $f_{X,R}(u_1, u_2) \geq p > 0$. Так как плотность случайной величины $Y(T_2)^2/2$ равномерно отделена от нуля на множестве D, случайная величина $Y(T_2)$ положительная, то плотность $Y(T_2)$ тоже отделена от нуля на отрезке

$$\left[\varepsilon\sqrt{\frac{b(H+1)l''}{a}}, \varepsilon\sqrt{\frac{b(H-1)l''}{2a} + \frac{2b(4-3l/4)}{a}}\right].$$

Следовательно мы можем утверждать, что найдется положительная постоянная p', которая минорирует плотность вектора (X, Y) на множестве

$$D' \equiv [\varepsilon k', 2\varepsilon k'] \times \left[\varepsilon\sqrt{\frac{b(H+1)l''}{a}}, \varepsilon\sqrt{\frac{b(H-1)l''}{2a} + \frac{2b(4-3l/4)}{a}}\right].$$

Тогда в качестве минорантной меры мы можем выбрать меру

$$\mu(\cdot) = p' \frac{\lambda(\cdot \cap D')}{\lambda(D')},$$

где λ – мера Лебега.

Заметим, что пересечение множеств $G \cap D'$ непусто и, более того, имеет положительную μ-меру. Тогда для любого целого $i \geq 1$ существует положительное число δ_i такое, что

$$\mathbb{P}(\tau_G(\mu) = i) > \mu(G \cap D')\mathbb{P}(T_i < \delta_i)(\lambda_1/(\lambda_1 + \lambda_2))^i > 0.$$

Следовательно, если цепь Маркова стартует из случайного начального состояния, имеющего распределение μ, то наибольший общий делитель моментов возвращения в множество G равен единице, что завершает доказательство эргодичности по Харрису вложенной цепи Маркова.

3. Доказательство теоремы 1

Мы привели выше доказательство эргодичности по Харрису цепи Маркова в дискретном времени. Аналогичный результат справедлив и в случае непрерывного времени. Для его доказательства достаточно воспользоваться теоремой 3 работы [3] (стр. 78) или воспользоваться результатами параграфа 3 главы 7 книги [4].

Мы воспользуемся теоремой 3 из [3] (для удобства читателя мы приводим это утверждение в Приложении) и заметим, что нами в предыдущем параграфе уже была установлена справедливость первых двух условий этой теоремы, поэтому нам осталось проверить лишь условие 3. А оно также выполнено, так как T_1 имеет абсолютно непрерывное распределение и

$$\mathbb{P}(\hat{\tau} > t) = \int \mu(dy)\mathbb{P}(T_{\tau_G(y)} > t) > \int_{G \cap D'} \mu(dy)\mathbb{P}(T_{\tau_G(y)} > t, \tau_G(y) = 1)$$
$$> \mu(G \cap D')\mathbb{P}(T_1 > t, T_1 < \delta_1) > 0$$

при некотором $\delta_1 > 0$.

4. Приложение

4.1. **О необходимости условия** (5) **в теореме 1.** Поясним, почему условие (5) теоремы 1 является необходимым для стабильной работы системы (заметим, что данное условие удобно также формулировать в виде $\delta := b/l\lambda_2 - a/k\lambda_1 > 0$). Пусть $(X^0, \widehat{Y}^0)$ – друмерное предельное распределение, к которому сходятся в метрике полной вариации при $n \to \infty$ распределения $(X_n, \widehat{Y}_n)$. Тогда при $t \to \infty$

$$\frac{Q^X(t)}{t} \sim \frac{\sum \int_{T_{i-1}}^{T_i} (X_i + au)du}{n} \frac{n}{t} \to \frac{\mathbb{E} \int_0^{t_1} (X^0 + au)du}{\mathbb{E}t_1}$$
$$= \mathbb{E}X^0 + a\frac{\mathbb{E}t_1^2}{2\mathbb{E}t_1} = \mathbb{E}X^0 + \frac{a}{(\lambda_1 + \lambda_2)}.$$

Предположим, что в начальный момент времени процесс стартовал с начального значения с очень большой координатой Q_0 и достаточно малыми другими координатами. Тогда к течение длительного времени процессы $X(t)$ и $Y(t)$ развиваются независимо/автономно, достаточно быстро каждый из них сходится к своему предельному/стационарному распределению и развивается в нём в течение длительного времени. Поэтому для стабильной работы системы нужно чтобы предел от $Q^X(t)/t$ при $t \to \infty$ был не больше предела от $Q^{\widehat{Y}}(t)/t$. Следовательно, с необходимостью получаем, что должно выполняться следующее неравенство

$$(\lambda_1 + \lambda_2)\mathbb{E}X^0 + a < (\lambda_1 + \lambda_2)\mathbb{E}\widehat{Y}^0 + b.$$

Так как X^0, $\widehat{Y}^0$ – стационарные распределения, то аналогично (8) и (9), получаем

$$\mathbb{E}X^0 = \alpha\mathbb{E}X^0 + a\alpha', \quad \mathbb{E}\widehat{Y}^0 = \alpha\mathbb{E}\widehat{Y}^0 + b\beta'.$$

Следовательно, можно будет переписать неравенство следующим образом

$$(\lambda_1 + \lambda_2)\frac{a\alpha'}{1-\alpha} + a = \frac{a}{1-\alpha} < \frac{b}{1-\beta} = (\lambda_1 + \lambda_2)\frac{b\beta'}{1-\beta} + b \iff \delta > 0.$$

4.2. **Эргодичность по Харрису.** Пусть $\{X_n\}$ — однородная по времени цепь Маркова, принимающая значения в польском пространстве $\mathcal{X}$. Для некоторого фиксированного множества $V \in \mathcal{B}_{\mathcal{X}}$ определим случайную величину $\tau_V(x) = \min\{k \geq 1 : X_k(x) \in V\}$, являющуюся временем первого попадания из состояния x в множество V (здесь $\tau_V(x) = \infty$, если $X_k(x) \notin V$ при всех $k \geq 1$).

Определение 1. Цепь Маркова $\{X_n\}$ в $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ называется *харрисовой* (или неприводимой по Харрису), если существуют множество $V \in \mathcal{B}_{\mathcal{X}}$, вероятностная мера μ на $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ и числа $m \geq 1$, $p \in (0,1)$ такие, что

$$\text{(I)} \quad \mathbb{P}(\tau_V(x) < \infty) = 1 \ \forall x \in \mathcal{X}; \quad \sup_{x \in V} \mathbb{E}\tau_V(x) < \infty;$$

$$\text{(II)} \quad \inf_{x \in V} \mathbb{P}(X_m \in B | X_0 = x) \geq p\mu(B) \ \forall B \in \mathcal{B}_{\mathcal{X}}.$$

Положим $\tau_V(\mu) = \{k \geq 1 : X_k(\mu) \in V\}$. Очевидно, что $\tau_V(\mu)$ в силу (I) есть собственная случайная величина. Обозначим через $\mathcal{K}$ множество возможных значений $\tau_V(\mu)$, т. е. значений k_i, при которых $\mathbb{P}(\tau_V(\mu) = k_i) > 0$. Введём дополнительно условие *непериодичности* этой марковской цепи: существуют $i \geq 1$ и $k_1, k_2, \ldots, k_i \in \mathcal{K}$ такие, что

$$\text{(III)} \quad \text{НОД}\{m + k_1, m + k_2, \ldots, m + k_i\} = 1,$$

где под НОД понимается «наибольший общий делитель». Ясно, что условие (III) с необходимостью следует из первых двух условий, если $m = 1$. Однако это, вообще говоря, не так при $m > 1$.

Определение 2. Пусть выполнены условия (I), (II), (III). Тогда цепь Маркова $\{X_n\}$ называется *эргодичной по Харрису.*

4.3. **Достаточные условия для положительной возратности множества.** Мы приведём формулировки известных утверждений, используемых при доказательстве положительной возвратности множества G в процессе доказательства теоремы 2. Доказательства этих утверждений можно найти, к примеру, в [5] и [6].

Пусть $\{X_n\}$ – однородная по времени цепь Маркова, принимающая значения в польском пространстве $\mathcal{X}$. Пусть $L : \mathcal{X} \to \mathbb{R}_+$– неотрицательная измеримая тестовая функция. Пусть $g : \mathcal{X} \to \mathbb{N}$ – некоторая другая измеримая функция, принимающая целочисленные значения, а $h : \mathcal{X} \to \mathbb{R}$ – ещё одна измеримая функция, принимающая вещественные значения. Функция $g(x)$ используется для определения интервала времени, на котором берется приращение тестовой функции, а функция $(-h)$ – для оценки среднего этого приращения за $g(x)$ единиц времени.

Будем считать, что $\sup_x L(x) = \infty$. Предположим, что

(L1) функция h ограничена снизу: $\inf_{x\in\mathcal{X}} h(x) > -\infty$

(L2) функция h принимает при больших $L(x)$ только положительные значения, отделённые от нуля, т. е. $\underline{\lim}_{L(x)\to\infty} h(x) > 0$

(L3) функция g локально ограничена сверху: $\sup_{L(x)\le N} g(x) < \infty$ для всех $N > 0$

(L4) функция g ограничена сверху функцией h при больших значениях $L(x)$, т. е. $\overline{\lim}_{L(x)\to\infty} g(x)/h(x) < \infty$

Для любого измеримого множества $B \subseteq \mathcal{X}$ определим

$$\tau_B = \inf\{n > 0 : X_n \in B\}$$

— первый момент возвращения в множество B, если цепь Маркова из него стартует, либо первый момент попадания в это множество, если стартует извне. Множество B называется возвратным, если $P_x(\tau_B < \infty) = 1$ при всех $x \in B$. Множество называется положительно возвратным, если $\sup_{x\in B} \mathbb{E}_x\tau_B < \infty$.

Теорема 3. *(Теорема 1 из* [5]*) Предположим, что снос L за $g(x)$ шагов оценивается сверху функцией $-h$, т.е.*

$$\mathbb{E}_x\left[V\left(X_{g(x)}\right) - V(X_0)\right] \le -h(x),$$

где функции L, g, h удовлетворяют условиям (L1) – (L4). *Пусть*

$$\tau \equiv \tau^{(N)} = \inf\{n > 0 : L(X_n) \le N\}$$

Тогда найдётся $N_0 > 0$, такое, что при всех $N > N_0$ и при любом $x \in \mathcal{X}$, имеем $\mathbb{E}_x\tau < \infty$. Более того, $\sup_{L(x)\le N} \mathbb{E}_x\tau < \infty$.

Теорема 4. *(Теорема 29 из* [6]*) Пусть* $\{X_n\}$ *— однородная по времени цепь Маркова со значениями в измеримом пространстве* $\mathcal{X}$*. Пусть* $B \subseteq \mathcal{X}$ *— подмножество, являющееся положительно возвратным для этой цепи. Пусть* D *— другое непустое множество в пространстве* $\mathcal{X}$ *и* $\tau_D = \min\{n : X_n \in D\}$*. Предположим, что* $\sup_{x\in D} \mathbb{E}_x \tau_B < \infty$ *и что найдётся такое натуральное число* N*, что* $\inf_{x\in B} \mathbb{P}_x(\tau_D \le N) > 0$*. Тогда множество* D *тоже положительно возвратно.*

4.4. **Достаточные условия для сходимости случайного процесса в метрике полной вариации.** Пусть $X = \{X(t) = X(t,x), t \in [0,\infty)\}$, $X(x,0) = x$ — произвольный случайный процесс со значениями в $\mathcal{X}$. Один из естественных подходов к изучению условий эргодичности процесса X связан с построением так называемых «вложенных» последовательностей, эргодичность которых может быть установлена. Вложенными обычно называют последовательности, образованные значениями процесса в некоторые марковские моменты времени. Пусть при $n \to \infty$

$$0 = T_0 < T_1 < T_2 < \dots < T_n < \dots, \ T_n \to \infty \text{ п. н.}$$

— некоторая случайная последовательность марковских моментов. Естественно ожидать, что при достаточно широких предположениях из эргодичности последовательности $X_n = X(T_n)$ будет следовать и эргодичность процесса X.

Теорема 5. *(Теорема 3, глава 7, см.* [3]*) Пусть процесс* X *допускает вложенную цепь Маркова* X_n *и выполняются следующие условия:*

1) X_n *удовлетворяет условиям (I)–(III);*

2) $\sup_{x\in\mathcal{X}} \mathbb{E}(T_1|X(0) = x) < \infty$*;*

3) распределение случайной величины $\hat\tau$ *имеет абсолютно непрерывную компоненту, где*

$$\mathbb{P}(\hat\tau > t) = \int \mu(dy)\mathbb{P}(T_{\tau_V(y)} > t|X(0) \in dy).$$

Тогда распределение процесса X *сходится к предельному распределению в метрике полной вариации.*

References

[1] A.Al. Hanbali, E. Altman, P. Nain, *A survey of TCP over ad hoc networks*, IEEE Communications, Surveys and Tutorials, **7**:3 (2005), 22–36.

[2] F. Baccelli, G. Carofiglio, S. Foss, *Proxy caching in split TCP: dynamics, stability and tail asymptotics*, In Y. Bertot (ed.) et al., *From semantics to computer science. Essays in honour of Gilles Kahn*, Cambridge University Press, Cambridge, 437–464, 2009. Zbl 1184.68017

[3] A. Borovkov, S. Foss, *Stochastically recursive sequences and their generalizations.* Sib. Adv. Math., **2**:1 (1992), 16–81. Zbl 0848.60025

[4] S. Asmussen, *Applied probability and queues.* Springer, New York, 2003. Zbl 1029.60001

[5] S. Foss, T. Konstantopoulos, *An overview of some stochastic stability methods.* J. Oper. Res. Soc. Japan., **47**:4 (2004), 275–303. Zbl 1134.93412

[6] S. Foss, N. Chernova, *Stability of random processes*, Novosibirsk State University, 2020.

Sergey Georgievich Foss
Novosibirsk State University,
2, Pirogova str.,
Novosibirsk, 630090, Russia
Heriot-Watt University,
Edinburgh, EH14 4AS, UK
Email address: sergueiorfoss25@gmail.com

Mikhail Georgievich Chebunin
Novosibirsk State University,
2, Pirogova str.,
Novosibirsk, 630090, Russia
Karlsruhe Institute of Technology,
Institute of Stochastics,
Karlsruhe, 76131, Germany.
Email address: chebuninmikhail@gmail.com