МОДЕЛИРОВАНИЕ СЕАНСОВ СВЯЗИ ЦЕПЯМИ МАРКОВА

С. П. Сущенко, В. В. Кокшенев

Томский государственный университет
Томск, Россия
e-mail: ssp.inf.tsu@gmail.com

Предложена модель виртуального соединения, управляемого транспортным протоколом в режимах селективного и группового режима повтора в виде цепи Маркова с дискретным временем, учитывающая влияние протокольных параметров размера окна и длительности тайм-аута ожидания сквозных квитанций, вероятности искажения пакетов в отдельных звеньях тракта передачи данных на пропускную способность виртуального соединения. Проведен анализ зависимости пропускной способности управляющей процедуры от протокольных параметров, уровня ошибок в каналах связи, длины тракта передачи данных.

Ключевые слова: транспортный протокол, тракт передачи данных, цепь Маркова, быстродействие виртуального соединения, размер окна, длительность тайм-аута.

В работе обобщаются результаты, полученные в [1], и развивается анализ, выполненный в [2]. Отметим, что до сих пор отсутствуют аналитические модели транспортного протокола, позволяющие функционально связать его быстродействие и другие операционные характеристики с протокольными параметрами, длиной тракта передачи и уровнем потерь в каналах связи. Рассмотрим процесс переноса данных между абонентами транспортного протокола, функционирующего в режиме селективного или группового отказа. Считаем, что участки переприема вдоль тракта передачи данных имеют одинаковое быстродействие в обоих направлениях. Полагаем, что длина тракта передачи данных, выраженная в количестве участков переприема, равна *D*. Обратный тракт, по которому доставляются сквозные подтверждения отправителю о корректности приема последовательности сегментов данных, также имеет длину *D*. Заданы вероятности искажения протокольного блока в канале связи для прямого $R_{\text{п}}(d), d=\overline{1,D}$ и обратного – $R_{\text{о}}(d), d=\overline{1,D}$ направлений передачи каждого участка переприема. Тогда достоверность передачи сегментов данных вдоль тракта от источника до адресата и обратно составят $F_n=\prod_{d=1}^{D}(1-R_{\text{п}}(d))$ и $F_o=\prod_{d=1}^{D}(1-R_{\text{о}}(d))$ соответственно. Передача данных на каждом участке переприема выполняется в соответствии с алгоритмом решающей обратной связи, время обработки пакетов в узлах тракта одинаково и длительность цикла передачи протокольного блока в отдельном звене составляет *t*. Взаимодействующие абоненты имеют неограниченный поток данных для передачи, а обмен выполняется информационными пакетами одинаковой длины. Подтверждения получателя о корректности приема данных переносятся в информационных пакетах встречного потока. Полагаем, что потерь кадров из-за отсутствия буферной памяти в узлах тракта не происходит. Управление потоком данных реализуется механизмом окна с протокольным параметром $W \geq 1$. Процесс информационного переноса в виртуальном соединении, управляемом транспортным протоколом, может быть описан Марковским процессом с дискретным временем (с длительностью такта t) в силу того, что время между получениями сквозных подтверждений имеет геометрическое распределение с параметром F_o. Область возможных состояний определяется длительностью тайм-аута ожидания сквозного подтверждения S, выраженной в циклах t. Размер тайм-аута связан с длиной тракта и шириной окна неравенствами $S > W$, $S \geq 2D$. Состояниям цепи Маркова $i = \overline{0, W}$ соответствует размер очереди переданных, но не подтвержденных данных в источнике потока, а состояниям $i = \overline{W+1, S-1}$ – время, в течение которого отправитель не активен и ожидает получение квитанции о корректности приема переданной последовательности из W протокольных блоков данных. Из нулевого состояния в $(2D-1)$-е источник продвигается с каждым тактом t с вероятностью детерминированного события. В состояниях $i \geq 2D-1$ после истечения очередного дискретного цикла t к отправителю начинают прибывать сквозные квитанции и, в зависимости от результатов доставки, отправитель передает новые протокольные блоки (при положительной квитанции), либо повторно – искаженные. Завершение цикла пребывания в состоянии $2D-1$ соответствует времени доведения первого протокольного блока до адресата и получения на него сквозной квитанции. Дальнейший рост номера состояния происходит с вероятностью искажения сквозной квитанции $1-F_o$ в обратном тракте. В состояниях $i \geq 2D-1$ в режиме селективного повтора получение сквозной квитанции порождает переход в $(2D-1)$-е состояние при $W \geq 2D$ или в состояние $(2D+W-2-i)$-е при $W \leq 2D$. В режиме группового повтора для состояний $i \geq 2D-1$ возврат в состояния $2D-1$ (при $W \geq 2D$) или $2D+W-2-i$ (при $W \leq 2D$) происходит при получении квитанции только в случае успешной доставки данных адресату, дошедших к моменту $i-2D+1$ до получателя, в противном случае следует возврат в нулевое состояние, поскольку очередь переданных, но не подтвержденных данных в этот момент обнуляется. Переходные вероятности цепи Маркова π_{ij} из исходного состояния i в результирующее j, описывающие динамику процесса передачи информационного потока в режиме селективного отказа, имеют вид:

$$\pi_{ij} = \begin{cases} 1,\ j = i+1, i = \overline{0, 2D-2}; \\ 1-F_o,\ j = i+1, i = \overline{2D-1, S-2}; \\ F_o,\ j = 2D-1, i = \overline{2D-1, W-1}, W \geq 2D; \\ F_o,\ j = 2D+W-2-i, i = \overline{2D-1, 2D+W-3}, W \leq 2D; \\ F_o,\ j = 2D+W-2-i, i = \overline{W, 2D+W-3}, W \geq 2D; \\ F_o,\ j = 0, i = \overline{2D+W-2, S-2}; \\ 1,\ j = 0, i = S-1. \end{cases}$$

Разнообразие вида решения системы уравнений равновесия для вероятностей состояний цепи Маркова определяется соотношениями между протокольными параметрами W, S и длиной тракта D. Поскольку длительность тайм-аута должна превышать ширину окна и быть не короче двойной длины пути передачи, то выделяются четыре варианта решения для различных областей изменения указанных параметров.

В случае интервальных ограничений на протокольные параметры $1 \leq W \leq 2D$, $\max\{W+1, 2D\} \leq S \leq 2D+W-1$ решение системы уравнений локального равновесия для вероятностей состояний определится зависимостями:

$$P_i = P_0, i = \overline{0,2D+W-1-S}; \quad P_i = P_0(1-F_o)^{2D+W-1-S-i}, i = \overline{2D+W-S, W-1};$$
$$P_i = P_0(1-F_o)^{2D-S}, i = \overline{W, 2D-2}; \quad P_i = P_0(1-F_o)^{i+1-S}, i = \overline{2D-1, S-1};$$
$$P_0 = \frac{F_o(1-F_o)^S}{(1-F_o)^{2D}[F_o(2D-W-1)+2]+(1-F_o)^S[F_o(2D+W-S+1)-2]}.$$
При ограничениях $1 \le W \le 2D, S \ge 2D+W-1$ вероятности состояний принимают вид:
$$P_i = P_0(1-F_o)^{-i}, i = \overline{0, W-1}; \quad P_i = P_0(1-F_o)^{-W+1}, i = \overline{W, 2D-2};$$
$$P_i = P_0(1-F_o)^{i-2D-W+2}, i = \overline{2D-1, S-1};$$
$$P_0 = \frac{F_o(1-F_o)^{W-1}}{2+F_o(2D-W-1)-(1-F_o)^W-(1-F_o)^{S-2D+1}}.$$
Если ширина окна W превалирует над двойной длиной тракта передачи данных ($W \ge 2D$), а область значений длительности тайм-аута S ограничена интервалом $W+1 \le S \le 2D+W-1$, то решение системы уравнений равновесия преобразуется к:
$$P_i = P_0, i = \overline{0,2D+W-S-1}; \quad P_i = P_0(1-F_o)^{2D+W-1-S-i}, i = \overline{2D+W-S, 2D-2};$$
$$P_i = P_0(1-F_o)^{i+1-S}, i = \overline{2D-1, S-1};$$
$$P_0 = \frac{F_o(1-F_o)^S}{(1-F_o)^{2D}+(1-F_o)^{W+1}+(1-F_o)^S[F_o(2D+W-S+1)-2]}.$$
Для протокольных параметров, связанных с длиной тракта неравенствами вида $W \ge 2D$ и $S \ge 2D+W-1$ вероятности P_i определяются соотношениями:
$$P_i = P_0(1-F_o)^{-i}, i = \overline{0, 2D-2}; \quad P_i = P_0(1-F_o)^{i-2D-W+2}, i = \overline{2D-1, S-1};$$
$$P_0 = \frac{F_o(1-F_o)^S}{1-(1-F_o)^W+(1-F_o)^{W-2D+1}[1-(1-F_o)^{S-W}]}.$$
Из приведенных решений нетрудно видеть, что они «сшиваются» на границах областей изменения протокольных параметров ширины окна W и длительности тайм-аута S. Важнейшей операционной характеристикой протокола является его пропускная способность, определяемая параметрами тракта передачи данных, накладными расходами и особенностями протокольных процедур управления передачей [2]. Для селективной процедуры отказа нормированное быстродействие определяется соотношением:
$$Z_c(W,S,D,F_n,F_o) = F_n F_0 \left\{ \sum_{i=2D-1}^{2D+W-2} (i-2D+2)P_i + W \sum_{i=2D+W-1}^{S-1} P_i \right\}.$$
С учетом вариативности выражений для вероятностей состояний цепи Маркова отсюда при различных связях между протокольными параметрами и длиной тракта получаем четыре функциональные зависимости показателя быстродействия. В случае интервальных ограничений на протокольные параметры $1 \le W \le 2D, \max\{W+1, 2D\} \le S \le 2D+W-1$ имеем:
$$Z_c^{(1)}(W,S,D,F_n,F_o) = \frac{F_n\left\{(1-F_o)^{2D}-(1-F_o)^{S+1}[1+(S-2D+1)F_o]\right\}}{(1-F_o)^{2D}[F_o(2D-W-1)+2]+(1-F_o)^S[F_o(2D+W-S+1)-2]}.$$
При $1 \le W \le 2D, S \ge 2D+W-1$ приходим к:
$$Z_c^{(2)}(W,S,D,F_n,F_o) = \frac{F_n\left\{1-(1-F_o)^W - WF_o(1-F_o)^{S-2D+1}\right\}}{2+F_o(2D-W-1)-(1-F_o)^W-(1-F_o)^{S-2D+1}}.$$

Область значений протокольных параметров $W \ge 2D$, $W+1 \le S \le 2D+W-1$ приводит к соотношению:

$$Z_c^{(3)}(W,S,D,F_n,F_o)=\frac{F_n\left\{(1-F_o)^{2D}-(1-F_o)^{S+1}[1+(S-2D+1)F_o]\right\}}{(1-F_o)^{2D}+(1-F_o)^{W+1}+(1-F_o)^{S}[F_o(2D+W-S+1)-2]}.$$

Для значений $W \ge 2D$, $S \ge 2D+W-1$ пропускная способность выразится зависимостью:

$$Z_c^{(4)}(W,S,D,F_n,F_o)=\frac{F_n\left\{1-(1-F_o)^{W}-WF_o(1-F_o)^{S-2D+1}\right\}}{1-(1-F_o)^{W}+(1-F_o)^{W-2D+1}[1-(1-F_o)^{S-W}]}.$$

Из полученных соотношений нетрудно убедиться в совпадении выражений на границах областей изменения протокольных параметров:

$$Z_c^{(1)}(W,S=2D+W-1,D,F_n,F_o)=Z_c^{(2)}(W,S=2D+W-1,D,F_n,F_o),$$

$$Z_c^{(1)}(W=2D,S,D,F_n,F_o)=Z_c^{(3)}(W=2D,S,D,F_n,F_o),$$

$$Z_c^{(1)}(W=2D,S=2D+W-1,D,F_n,F_o)=Z_c^{(4)}(W=2D,S=2D+W-1,D,F_n,F_o),$$

$$Z_c^{(2)}(W=2D,S,D,F_n,F_o)=Z_c^{(3)}(W=2D,S,D,F_n,F_o),$$

$$Z_c^{(2)}(W=2D,S=2D+W-1,D,F_n,F_o)=Z_c^{(4)}(W=2D,S=2D+W-1,D,F_n,F_o),$$

$$Z_c^{(3)}(W,S=2D+W-1,D,F_n,F_o)=Z_c^{(4)}(W,S=2D+W-1,D,F_n,F_o).$$

Рассмотрим режим группового повтора. Переходные вероятности цепи Маркова, описывающие динамику очереди протокольных блоков в источнике ожидающих сквозного подтверждения, задаются следующим образом:

$$\pi_{ij}=\begin{cases}1,\ j=i+1, i=\overline{0,2D-2};\\ 1-F_o,\ j=i+1, i=\overline{2D-1,S-2};\\ F_oF_n^{i-2D+2},\ j=2D-1, i=\overline{2D-1,W-1}, W\ge 2D;\\ F_o(1-F_n^{i-2D+2}),\ j=0, i=\overline{2D-1,W-1}, W\ge 2D;\\ F_oF_n^{i-2D+2},\ j=2D+W-2-i, i=\overline{2D-1,2D+W-3}, W\le 2D;\\ F_o(1-F_n^{i-2D+2}),\ j=0, i=\overline{2D-1,2D+W-3}, W\le 2D;\\ F_o,\ j=2D+W-2-i, i=\overline{W,2D+W-3}, W\ge 2D;\\ F_o,\ j=0, i=\overline{2D+W-2,S-2};\\ 1,\ j=0, i=S-1.\end{cases}$$

Так же, как и в случае режима селективного отказа, решение системы уравнений равновесия для различных областей допустимых значений протокольных параметров имеет четыре аналитических варианта. Множества значений, ограниченных снизу и сверху неравенствами $1 \le W \le 2D$ и $\max\{W+1,2D\} \le S \le 2D+W-1$, приводят к решению вида:

$$P_i=P_0, i=\overline{0,2D+W-S-1};\quad P_i=P_0\frac{1-F_n+F_nF_o[F_n(1-F_o)]^{W-1-i}}{1-F_n+F_nF_o[F_n(1-F_o)]^{W-1}}, i=\overline{2D+W-S,W-1};$$

$$P_i=P_0\frac{1-F_n(1-F_o)}{1-F_n+F_nF_o[F_n(1-F_o)]^{W-1}}, i=\overline{W,2D-2};$$

$$P_i=P_0\frac{[1-F_n(1-F_o)](1-F_o)^{i-2D+1}}{1-F_n+F_nF_o[F_n(1-F_o)]^{W-1}}, i=\overline{2D-1,S-1}.$$

Ограничения $1 \le W \le 2D, S \ge 2D+W-1$ обеспечивают следующее решение:

$$P_i = P_0 \frac{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1-i}}{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}}, i=\overline{0, W-1};$$
$$P_i = P_0 \frac{1-F_n(1-F_o)}{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}}, i=\overline{W, 2D-2};$$
$$P_i = P_0 \frac{[1-F_n(1-F_o)](1-F_o)^{i-2D+1}}{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}}, i=\overline{2D-1, S-1}.$$

Области изменения протокольных параметров $W \ge 2D$, $W+1 \le S \le 2D+W-1$ дают зависимости:

$$P_i = P_0, i=\overline{0, 2D+W-S-1};\ P_i = P_0 \frac{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1-i}}{1-F_n + F_n F_o [F_n(1-F_o)]^{S-2D}}, i=\overline{2D+W-S, 2D-2};$$
$$P_i = P_0 \frac{[1-F_n(1-F_o)](1-F_o)^{i-2D+1}}{1-F_n + F_n F_o [F_n(1-F_o)]^{S-2D}}, i=\overline{2D-1, S-1}.$$

Для протокольных параметров, ограниченных только снизу ($W \ge 2D$, $S \ge 2D+W-1$), получаем:

$$P_i = P_0 \frac{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1-i}}{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}}, i=\overline{0, 2D-2};$$
$$P_i = P_0 \frac{[1-F_n(1-F_o)](1-F_o)^{i-2D+1}}{1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}}, i=\overline{2D-1, S-1}.$$

Выражения для P_0 здесь не выписываем в силу их громоздкости. Если в приведенных соотношениях положить $F_n = 1$, то ожидаемо получим зависимости для селективного режима повтора. Пропускная способность транспортного соединения для режима группового отказа задается следующей зависимостью:

$$Z_г(W,S,D,F_n,F_o) = F_n F_0 \left\{ \sum_{i=2D-1}^{2D+W-2} \frac{1-F_n^{i-2D+2}}{1-F_n} P_i + \frac{1-F_n^W}{1-F_n} \sum_{i=2D+W-1}^{S-1} P_i \right\}.$$

В том случае если сквозной тайм-аут имеет интервальные ограничения ($\max\{W+1, 2D\} \le S \le 2D+W-1$), то для пропускной способности с точностью до сомножителя P_0 получаем:

$$Z_г^{(1)}(W,S,D,F_n,F_o) = \frac{P_0 F_n \left\{1-F_n - [1-F_n(1-F_o)](1-F_o)^{S-2D+1} + F_n F_o [F_n(1-F_o)]^{S-2D+1}\right\}}{(1-F_n)[1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}]}.$$

Для наиболее важного случая длительности тайм-аута, ограниченной снизу ($S \ge 2D+W-1$), и произвольной ширины окна пропускная способность выразится зависимостью:

$$Z_г^{(2)}(W,S,D,F_n,F_o) = \frac{P_0 F_n \left\{(1-F_n)[1-[F_n(1-F_o)]^W] - [1-F_n(1-F_o)](1-F_o)^{S-2D+1}(1-F_n^W)\right\}}{(1-F_n)[1-F_n + F_n F_o [F_n(1-F_o)]^{W-1}]}.$$

Нетрудно убедиться, что при $F_n = 1$ из данных соотношений получаем зависимости для режима селективного отказа. Неограниченный рост ширины окна ($W \to \infty$), а следовательно и длительности тайм-аута S, приводит к инвариантной к длине тракта зависимости пропускной способности в селективном режиме повтора и определяется только достоверностью передачи сегментов в прямом тракте. То же самое имеет место для селективного режима отказа, абсолютно надежного обратного канала ($F_o = 1$) и ширины окна не меньшей удвоенной длины пути ($W \ge 2D$):

$$Z_c^{(4)}(W=\infty,S,D,F_n,F_o)=Z_c^{(3,4)}(W\ge 2D,S,D,F_n,F_o=1)=F_n.$$

При $F_o=1$ и $W\le 2D$ быстродействие транспортного соединения в селективном режиме повтора обратно пропорционально различию удвоенной длины пути и ширины окна:

$$Z_c^{(1,2)}(W\le 2D,S,D,F_n,F_o=1)=\frac{F_n}{2D-W+1}.$$

В режиме группового отказа предельные возможности управляющей процедуры, соответствующие неограниченному размеру окна ($W\to\infty$), определяются зависимостью:

$$Z_г^{(2)}(W=\infty,S,D,F_n,F_o)=\frac{F_nF_o}{1-F_n(1-F_o)+(2D-1)F_o(1-F_n)}.$$

а при $F_o=1$ для различных размеров окна быстродействие составит:

$$Z_г^{(2)}(W\ge 2D,S,D,F_n,F_o)=\frac{F_n}{1+(2D-1)(1-F_n)},$$

$$Z_г^{(1)}(W\le 2D,S,D,F_n,F_o)=\frac{F_n}{2D-W+1+(W-1)(1-F_n)}.$$

Для старт-стопного протокола ($W=1$) селективный и групповой режимы повтора совпадают и пропускная способность примет вид:

$$Z_{cc}(W=1,S,D,F_n,F_o)=\frac{F_n\{1-(1-F_o)^{S-2D+1}[1+(S-2D+1)F_o]\}}{2+F_o(2D-2)+(1-F_o)^{S-2D}[F_o(2D+2-S)-2]},2D\le S\le 2D+W-1;$$

$$Z_{cc}(W=1,S,D,F_n,F_o)=\frac{F_nF_o[1-(1-F_o)^{S-2D+1}]}{1+F_o(2D-1)-(1-F_o)^{S-2D+1}},S\ge 2D+W-1.$$

При минимальной длительности тайм-аута ($S=2D$) быстродействие старт-стопного протокола составит:

$$Z_{cc}(W=1,S=2D,D,F_n,F_o)=\frac{F_nF_o}{2D}.$$

То же самое соотношение имеет место в случае $W<2D$ и $S=2D$ для селективного и группового режима повторной передачи. Из чего следует, что тайм-аут минимальной длительности при $W<2D$ сводит оба режима отказа к старт-стопной процедуре и обеспечивает инвариантность пропускной способности к ширине окна. При неограниченной длительности тайм-аута ($S=\infty$) пропускная способность старт-стопной процедуры преобразуется к:

$$Z_{cc}(W=1,S=\infty,D,F_n,F_o)=\frac{F_nF_o}{1+F_o(2D-1)}.$$

Следует отметить значительную зависимость индекса быстродействия в режиме группового отказа от длины транспортного соединения, что при конвейерной интерпретации тракта передачи данных легко объясняется необходимостью перезагрузки всего или части конвейера в случае потери хотя бы одного протокольного блока в отправленной последовательности (повторной передачи всех блоков, начиная с первого потерянного).

ЛИТЕРАТУРА

1. *Кокшенев, В. В.* Анализ быстродействия асинхронной процедуры управления звеном передачи данных / В. В. Кокшенев, С. П. Сущенко // Вычислительные технологии, 2008. Т. 15. Спец. вып. № 5. С. 61–65.

2. *Кокшенев, В. В.* Анализ селективного режима отказа транспортного протокола в нагруженном тракте передачи данных / В. В. Кокшенев, П. А. Михеев, С. П. Сущенко // Вестник ТГУ. Серия управление, вычислительная техника и информатика. 2013. № 3(24). С. 78–94.