

© 2017 г. В.В. КОКШЕНЕВ, канд. техн. наук (vladimir_fin@mail.ru)
П.А. МИХЕЕВ, канд. техн. наук (doka.patrick@gmail.com),
С.П. СУЩЕНКО, д-р техн. наук (ssp.inf.tsu@gmail.com)
(Томский государственный университет)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ БЫСТРОДЕЙСТВИЯ СЕЛЕКТИВНОГО И ГРУППОВОГО РЕЖИМОВ ПОВТОРНОЙ ПЕРЕДАЧИ ТРАНСПОРТНОГО ПРОТОКОЛА

Предложена модель виртуального соединения, управляемого транспортным протоколом в режимах селективного и группового отказа в виде цепи Маркова с дискретным временем, учитывающая влияние протокольных параметров размера окна и длительности тайм-аута ожидания подтверждений, вероятности искажения сегментов в отдельных звеньях тракта передачи данных на пропускную способность транспортного соединения. Проведен анализ зависимости пропускной способности управляющей процедуры от протокольных параметров, уровня ошибок в каналах связи, длительности круговой задержки. Предложен метод выбора протокольных параметров.

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

1. Введение

Важнейшим показателем качества взаимодействия сетевых приложений и используемого программно-технического обеспечения компьютерных сетей является пропускная способность транспортных соединений. Данная операционная характеристика в значительной мере определяется транспортным протоколом и его параметрами — шириной окна и длительностью тайм-аута [1, 2]. Моделирование абонентского соединения и анализ его потенциальных возможностей выполнялся, например, в [2–6]. Но результаты получены только для однозвенного тракта передачи данных [2–5] либо только при существенных ограничениях на параметры протокола [6]. Современные транспортные протоколы содержат большое разнообразие механизмов управления перегрузками [7]. Имеется широкий спектр исследований [7–21] в области управления протокольными параметрами транспортного протокола с целью предупреждения и обхода перегрузок, ориентированный на построение моделей диагностики перегрузки по различным индикаторам [7] и адаптацию протокольных параметров к изменяющейся сетевой нагрузке и связности, уровню потерь, активности взаимодействующих абонентов и др. При этом реализация управляющих механизмов для обхода перегрузок опирается на доступную полосу пропускания транспортных соединений при текущих и прогнозируемых к изменению значениях протокольных параметров. Однако все еще не изучены потенциальные возможности транспортного протокола, отсутствуют аналитические зависимости влияния протокольных параметров, режимов повтора утерянных сегментов, характеристик тракта передачи данных, конкурентного соперничества абонентских сеансов связи за полосу пропускания

разделяемых сетевых каналов на результирующие операционные характеристики транспортного соединения. Не исследовано влияние соотношений между длительностью круговой задержки и протокольными параметрами на пропускную способность тракта передачи данных, управляемого транспортным протоколом. Кроме того, процессы передачи данных в компьютерных сетях имеют существенно дискретный характер [22], обусловленный конвейерным механизмом переноса в сети сегментов ограниченных размеров и применением алгоритмов с решающей обратной связью на различных уровнях сетевой архитектуры, однако большинство результатов [2–4, 7–20] основано на моделях с непрерывным временем, что приводит к сужению области их использования.

В данной статье предложена математическая модель процесса передачи данных в фазе информационного переноса в виде цепи Маркова с дискретным временем (раздел 2), в аналитическом виде найдены стационарные распределения вероятностей состояний для режимов селективного и группового отказов [2] (разделы 3, 4), получены аналитические соотношения для пропускной способности, на основе которых выполнен сравнительный анализ потенциальных возможностей транспортного соединения в различных режимах отказа (раздел 5), получены выражения для рациональных значений протокольных параметров: ширины окна и длительности тайм-аута (раздел 6).

2. Модель транспортного соединения

Рассмотрим процесс переноса данных между абонентами транспортного протокола, основанного на алгоритме с решающей обратной связью [2] и функционирующего в режиме селективного или группового отказов. Примером семейства таких надежных протоколов является доминирующий в современных компьютерных сетях протокол TCP [1]. В режиме селективного отказа повторной передаче от источника подлежат только не принятые получателем сегменты, а в режиме группового отказа — все, начиная с первого не принятого [2]. Полагаем, что взаимодействующие абоненты имеют неограниченный поток данных для передачи, а обмен выполняется протокольными блоками данных транспортного протокола (сегментами) одинаковой длины. Подтверждения получателя о корректности приема данных переносятся в сегментах встречного потока. Считаем, что участки переприема вдоль тракта передачи данных имеют одинаковое быстродействие в обоих направлениях, а длительность цикла передачи сегмента в отдельном звене составляет t . В общем случае длина пути от источника до адресата, переносящего информационный поток, и длина обратного пути, по которому передаются подтверждения на принятые сегменты могут быть различными. Полагаем, что длина тракта передачи данных, выраженная в количестве участков переприема, в прямом направлении равна $D_{\Pi} \geq 1$. Обратный тракт, по которому доставляются сквозные подтверждения отправителю о корректности приема последовательности сегментов данных, имеет длину $D_{\circ} \geq 1$. Заданы вероятности искажения сегмента в каналах связи для прямого $R_{\Pi}(d)$, $d = \overline{1, D_{\Pi}}$ и обратного — $R_{\circ}(d)$, $d = \overline{1, D_{\circ}}$, направлений передачи каждого участка переприема. Тогда достоверности передачи сегментов данных вдоль тракта от источника до адресата и обратно составят $F_{\Pi} = \prod_{d=1}^{D_{\Pi}} (1 - R_{\Pi}(d))$ и $F_{\circ} = \prod_{d=1}^{D_{\circ}} (1 - R_{\circ}(d))$

соответственно. Считаем, что потеря сегментов из-за отсутствия буферной памяти в узлах тракта не происходит. Управление потоком данных реализуется механизмом скользящего окна [1, 2] с протокольным параметром ширины окна $W \geq 1$. 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_0 . Данная модель является обобщением формализаций процесса передачи данных, предложенных в [3–6], на случай транспортного соединения произвольной длины. Область возможных состояний цепи Маркова определяется длительностью тайм-аута ожидания сквозного подтверждения S , выраженной в количестве циклов продолжительности t . Размер тайм-аута связан с длиной тракта и шириной окна неравенствами $S > W$, $S \geq D_{\text{п}} + D_0$. Очевидно, что сумму длин прямого и обратного трактов можно интерпретировать как круговую задержку $D = D_{\text{п}} + D_0$, выраженную в длительностях t (без учета потерь протокольных блоков при передаче вдоль тракта). Состояниям цепи Маркова $i = \overline{0, W}$ соответствует размер очереди переданных, но не подтвержденных сегментов в источнике потока, а состояниям $i = \overline{W + 1, S - 1}$ — время, в течение которого отправитель не активен и ждет подтверждения о корректности приема переданной последовательности из W сегментов. Из нулевого состояния в $(D - 1)$ -е источник продвигается с каждым тактом t с вероятностью детерминированного события. В состояниях $i \geq D - 1$ после истечения очередного дискретного цикла t к отправителю начинают прибывать подтверждения и в зависимости от результатов доставки отправитель передает новые сегменты (при положительном подтверждении) либо повторно — искаженные. Завершение цикла пребывания в состоянии $D - 1$ соответствует времени доведения первого сегмента до адресата и получения на него подтверждения. Дальнейший рост номера состояния происходит с вероятностью искажения подтверждения $1 - F_0$ в обратном тракте. В состояниях $i \geq D - 1$ в режиме селективного отказа получение подтверждения порождает переход в $(D - 1)$ -е состояние при $W \geq D$ или в состояние $D + W - 2 - i$ при $W \leq D$. В режиме группового отказа для исходных состояний $i \geq D - 1$ возврат в состояния $D - 1$ (при $W \geq D$) или $D + W - 2 - i$ (при $W \leq D$) происходит при получении подтверждения только в случае успешной доставки данных адресату, дошедших к моменту $i - D + 1$ до получателя, в противном случае следует возврат в нулевое состояние, поскольку очередь переданных, но неподтвержденных сегментов в этот момент обнуляется. В силу того что в состояниях $i \geq W$ источник приостанавливает отправку сегментов, получение подтверждений в состояниях $i = \overline{W, D + W - 3}$ приводит к переходу в состояния $D + W - 2 - i$, а из состояний $i = \overline{D + W - 2, S - 2}$ — в нулевое. Это справедливо для селективного отказа, а в режиме группового отказа указанные изменения состояний реализуются при поступлении положительных подтверждений. В состоянии $S - 1$ истекает тайм-аут ожидания подтверждения от получателя о корректности принятых сегментов и происходит безусловный переход в нулевое состояние во всех режимах отказа.

3. Вероятности состояний для режима селективного отказа

Переходные вероятности π_{ij} из исходного состояния i в результирующее j цепи Маркова, описывающей процесс передачи информационного потока в режиме селективного отказа, имеют вид

$$\pi_{ij} = \begin{cases} 1, & i = \overline{0, D-2}; \quad j = i+1; \\ 1 - F_0, & i = \overline{D-1, S-2}; \quad j = i+1; \\ F_0, & i = \overline{D-1, W-1}; \quad W \geq D; \quad j = D-1; \\ F_0, & i = \overline{D-1, D+W-3}; \quad W \leq D; \quad j = D+W-2-i; \\ F_0, & i = \overline{W, D+W-3}; \quad W \geq D; \quad j = D+W-2-i; \\ F_0, & i = \overline{D+W-2, S-2}; \quad j = 0; \\ 1, & i = S-1; \quad j = 0. \end{cases}$$

Разнообразие вида решения системы уравнений равновесия для вероятностей состояний цепи Маркова определяется соотношениями между протокольными параметрами W , S и общей длиной тракта D . Поскольку длительность тайм-аута должна превышать ширину окна и быть не короче круговой задержки ($S \geq D$), то выделяются четыре варианта решения для различных областей изменения протокольных параметров.

Для протокольных параметров, связанных с общей длиной тракта неравенствами вида

$$(1) \quad W \geq D, \quad S \geq D + W - 1,$$

система уравнений равновесия записывается следующим образом:

$$(2) \quad P_0 = P_{S-1} + F_0 \sum_{i=D+W-2}^{S-2} P_i;$$

$$(3) \quad P_i = P_{i-1} + F_0 P_{D+W-2-i}, \quad i = \overline{1, D-2};$$

$$(4) \quad P_{D-1} = P_{D-2} + F_0 \sum_{i=D-1}^{W-1} P_i;$$

$$(5) \quad P_i = P_{i-1}(1 - F_0), \quad i = \overline{D, S-1}.$$

С учетом условия нормировки решение данной системы определяется соотношениями:

$$(6) \quad \begin{aligned} P_i &= P_0(1 - F_0)^{-i}, \quad i = \overline{0, D-2}; \\ P_i &= P_0(1 - F_0)^{i-D-W+2}, \quad i = \overline{D-1, S-1}; \\ P_0 &= \frac{F_0(1 - F_0)^{W-1}}{1 - (1 - F_0)^W + (1 - F_0)^{W-D+1}[1 - (1 - F_0)^{S-W}]}. \end{aligned}$$

Если ширина окна W превалирует над общей длиной тракта передачи данных, а область значений длительности тайм-аута S имеет интервальные ограничения

$$(7) \quad W \geq D, \quad W + 1 \leq S \leq D + W - 1,$$

то уравнения (2), (3) для состояний $i = \overline{0, D - 2}$ преобразуются к

$$(8) \quad \begin{aligned} P_0 &= P_{S-1}; \\ P_i &= P_{i-1}, \quad i = \overline{1, D + W - S - 1}; \\ P_i &= P_{i-1} + F_0 P_{D+W-2-i}, \quad i = \overline{D + W - S, D - 2}, \end{aligned}$$

и вероятности состояний цепи Маркова принимают вид

$$(9) \quad \begin{aligned} P_i &= P_0, \quad i = \overline{0, D + W - S - 1}; \\ P_i &= P_0(1 - F_0)^{D+W-1-S-i}, \quad i = \overline{D + W - S, D - 2}; \\ P_i &= P_0(1 - F_0)^{i+1-S}, \quad i = \overline{D - 1, S - 1}; \\ P_0 &= \frac{F_0(1 - F_0)^S}{(1 - F_0)^D + (1 - F_0)^{W+1} + (1 - F_0)^S [F_0(D + W - S + 1) - 2]}. \end{aligned}$$

При ограничениях

$$(10) \quad 1 \leq W \leq D, \quad S \geq D + W - 1$$

уравнения равновесия (3), (4) переписываются в виде:

$$P_i = P_{i-1} + F_0 P_{D+W-2-i}, \quad i = \overline{1, W - 1}; \quad P_i = P_{i-1}, \quad i = \overline{W, D - 1}.$$

Вероятности состояний при этом имеют подмножество ($i = \overline{W, D - 1}$) значений, инвариантных к номеру состояния:

$$(11) \quad \begin{aligned} P_i &= P_0(1 - F_0)^{-i}, \quad i = \overline{0, W - 1}; \quad P_i = P_0(1 - F_0)^{-W+1}, \quad i = \overline{W, D - 1}; \\ P_i &= P_0(1 - F_0)^{i-D-W+2}, \quad i = \overline{D, S - 1}; \\ P_0 &= \frac{F_0(1 - F_0)^{W-1}}{2 + F_0(D - W - 1) - (1 - F_0)^W - (1 - F_0)^{S-D+1}}. \end{aligned}$$

В случае интервальных ограничений на оба протокольных параметра

$$(12) \quad 1 \leq W \leq D, \quad \max\{W + 1, D\} \leq S \leq D + W - 1$$

уравнение (2) принимает вид (8), а уравнения (3), (4) преобразуются к следующим:

$$\begin{aligned} P_i &= P_{i-1}, \quad i = \overline{1, D + W - S - 1, \overline{W, D - 1}}; \\ P_i &= P_{i-1} + F_0 P_{D+W-2-i}, \quad i = \overline{D + W - S, W - 1}. \end{aligned}$$

Решение системы уравнений локального равновесия определится зависимостями с двумя подмножествами ($i = \overline{0, D + W - S - 1}, \overline{W, D - 1}$) значений вероятностей состояний, не зависящих от номера состояния:

$$\begin{aligned}
 P_i &= P_0, \quad i = \overline{0, D + W - 1 - S}; \\
 P_i &= P_0(1 - F_o)^{D+W-1-S-i}, \quad i = \overline{D + W - S, W - 1}; \\
 (13) \quad P_i &= P_0(1 - F_o)^{D-S}, \quad i = \overline{W, D - 1}; \\
 P_i &= P_0(1 - F_o)^{i+1-S}, \quad i = \overline{D, S - 1}; \\
 P_0 &= \frac{F_o(1 - F_o)^S}{(1 - F_o)^D [F_o(D - W - 1) + 2] + (1 - F_o)^S [F_o(D + W - S + 1) - 2]}.
 \end{aligned}$$

Таким образом, стационарное распределение вероятностей состояний цепи Маркова при различных связях между шириной окна W , длительностью тайм-аута S и общей длиной тракта передачи данных D в (1), (7), (10) и (12) определяется соотношениями (6), (9), (11) и (13) соответственно. Из указанных решений нетрудно видеть, что они «сшиваются» на границах областей изменения протокольных параметров ширины окна W и длительности тайм-аута S . При минимальной длительности тайм-аута ($S = D$) состояния цепи Маркова равновероятны и инвариантны к размеру окна: $P_i = \frac{1}{D}, i = \overline{0, D - 1}$. Если $S = D + 1$, то вероятности состояний определяются двумя множествами с равномерным рассеянием значений:

$$\begin{aligned}
 P_i &= \frac{1 - F_o}{1 + D - F_o W}, \quad i = \overline{0, W - 2}, D; \\
 P_i &= \frac{1}{1 + D - F_o W}, \quad i = \overline{W - 1, D - 1}.
 \end{aligned}$$

Для абсолютно надежного обратного тракта передачи данных ($F_o = 1$) вся вероятностная масса либо равномерно рассеяна по состояниям $i = \overline{W - 1, D - 1}$ ($P_i = \frac{1}{1+D-W}, i = \overline{W - 1, D - 1}; 1 \leq W \leq D$), либо сосредоточена в состоянии $D - 1$ ($P_{D-1} = 1, W \geq D$).

4. Вероятности состояний для режима группового отказа

Рассмотрим режим группового отказа. Переходные вероятности цепи Маркова, описывающей динамику очереди сегментов в источнике, ожидающих подтверждения, задаются следующим образом

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

Как и в случае режима селективного отказа, решение системы уравнений равновесия в режиме группового отказа для различных областей допустимых значений протокольных параметров имеет четыре аналитических варианта.

Для протокольных параметров, ограниченных только снизу (1), система уравнений локального равновесия имеет вид:

$$(14) \quad P_0 = P_{S-1} + F_o \left\{ \sum_{i=D-1}^{D+W-3} P_i (1 - F_{\Pi}^{i-D+2}) + \sum_{i=D+W-2}^{S-2} P_i \right\};$$

$$(15) \quad P_i = P_{i-1} + F_o P_{D+W-2-i} F_{\Pi}^{D+W-2-i}, \quad i = \overline{1, D-2};$$

$$(16) \quad P_{D-1} = P_{D-2} + F_o \sum_{i=D-1}^{W-1} P_i F_{\Pi}^{i-D+2};$$

$$(17) \quad P_i = P_{i-1} (1 - F_o), \quad i = \overline{D, S-1}.$$

Вероятности состояний, найденные отсюда, определяются соотношениями:

$$(18) \quad \begin{aligned} P_i &= P_0 \frac{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1-i}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}}, \quad \Phi = F_{\Pi} (1 - F_o), \quad i = \overline{0, D-2}; \\ P_i &= P_0 \frac{(1 - \Phi)(1 - F_o)^{i-D+1}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}}, \quad i = \overline{D-1, S-1}; \\ P_0 &= F_o (1 - \Phi) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}] / \left\{ (D-1)(1 - F_{\Pi}) F_o (1 - \Phi) + \right. \\ &\quad \left. + F_{\Pi} F_o^2 (\Phi^{W-D+1} - \Phi^W) + (1 - \Phi)^2 [1 - (1 - F_o)^{S-D+1}] \right\}. \end{aligned}$$

Области значений протокольных параметров (7) приводят к следующему изменению уравнений (14), (15):

$$(19) \quad \begin{aligned} P_0 &= P_{S-1} + F_o \sum_{i=D-1}^{S-2} P_i (1 - F_{\Pi}^{i-D+2}); \\ P_i &= P_{i-1}, \quad i = \overline{1, D+W-S-1}; \\ P_i &= P_{i-1} + F_o P_{D+W-2-i} F_{\Pi}^{D+W-2-i}, \quad i = \overline{D+W-S, D-2}. \end{aligned}$$

Решение результирующей системы уравнений равновесия принимает вид:

$$(20) \quad \begin{aligned} P_i &= P_0, \quad i = \overline{0, D+W-S-1}; \\ P_i &= P_0 \frac{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1-i}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}}, \quad i = \overline{D+W-S, D-2}; \\ P_i &= P_0 \frac{(1 - \Phi)(1 - F_o)^{i-D+1}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}}, \quad i = \overline{D-1, S-1}; \\ P_0 &= F_o (1 - \Phi) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}] / \left\{ (D+W-S) F_o (1 - \Phi) \times \right. \\ &\quad \times [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}] + (S-W-1) F_o (1 - F_{\Pi}) (1 - \Phi) + \\ &\quad \left. + F_{\Pi} F_o^2 (\Phi^{W-D+1} - \Phi^{S-D}) + (1 - \Phi)^2 [1 - (1 - F_o)^{S-D+1}] \right\}. \end{aligned}$$

При ограничениях вида (10) уравнения равновесия (15), (16) исходной системы уравнений равновесия преобразуются к виду:

$$\begin{aligned} P_i &= P_{i-1} + F_o P_{D+W-2-i} F_{\Pi}^{D+W-2-i}, \quad i = \overline{1, W-1}; \\ P_i &= P_{i-1}, \quad i = \overline{W, D-1}. \end{aligned}$$

Решение полученной системы определяется зависимостями:

$$\begin{aligned} P_i &= P_0 \frac{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1-i}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}}, \quad i = \overline{0, W-1}; \\ P_i &= P_0 \frac{1 - \Phi}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}}, \quad i = \overline{W, D-2}; \\ (21) \quad P_i &= P_0 \frac{(1 - \Phi)(1 - F_o)^{i-D+1}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}}, \quad i = \overline{D-1, S-1}; \\ P_0 &= F_o(1 - \Phi) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}] / \left\{ (W-1)F_o(1 - F_{\Pi})(1 - \Phi) + \right. \\ &\quad \left. + F_{\Pi} F_o^2 \Phi(1 - \Phi^{W-1}) + (D-W)F_o(1 - \Phi)^2 + \right. \\ &\quad \left. + (1 - \Phi)^2 [1 - (1 - F_o)^{S-D+1}] \right\}. \end{aligned}$$

Значения протокольных параметров, имеющие интервальные ограничения (12), изменяют уравнение (14) до (19), а уравнения (15), (16) до:

$$\begin{aligned} P_i &= P_{i-1}, \quad i = \overline{1, D+W-S-1, W, D-1}; \\ P_i &= P_{i-1} + F_o P_{D+W-2-i} F_{\Pi}^{D+W-2-i}, \quad i = \overline{D+W-S, W-1}. \end{aligned}$$

И обеспечивают решение вида:

$$\begin{aligned} P_i &= P_0, \quad i = \overline{0, D+W-S-1}; \\ P_i &= P_0 \frac{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1-i}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}}, \quad i = \overline{D+W-S, W-1}; \\ P_i &= P_0 \frac{1 - \Phi}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}}, \quad i = \overline{W, D-2}; \\ (22) \quad P_i &= P_0 \frac{(1 - \Phi)(1 - F_o)^{i-D+1}}{1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}}, \quad i = \overline{D-1, S-1}; \\ P_0 &= F_o(1 - \Phi) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}] / \left\{ (D+W-S)F_o(1 - \Phi) \times \right. \\ &\quad \times [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}] + (D-W)F_o(1 - \Phi)^2 + \\ &\quad \left. + (S-D)F_o(1 - F_{\Pi})(1 - \Phi) + F_{\Pi} F_o^2 (1 - \Phi^{S-D}) + \right. \\ &\quad \left. + (1 - F_o)(1 - \Phi)^2 [1 - (1 - F_o)^{S-D}] \right\}. \end{aligned}$$

В итоге для четырех областей совместных значений ширины окна и длительности тайм-аута в (1), (7), (10) и (12) стационарное распределение вероятностей состояний цепи Маркова для режима группового отказа определяется соотношениями (18), (20), (21) и (22) соответственно. При $F_o = 1$ вероятности состояний цепи Маркова преобразуются к виду:

$$P_i = P_0, \quad i = \overline{0, W-2};$$

$$P_i = \frac{P_0}{1 - F_{\Pi}}, \quad i = \overline{W-1, D-1};$$

$$P_0 = \frac{1 - F_{\Pi}}{D - F_{\Pi}(W-1)}.$$

Если в приведенных соотношениях для P_i положить $F_{\Pi} = 1$, то ожидаемо получим зависимости для селективного режима отказа. При длительности тайм-аута, равной круговой задержке ($S = D$), состояния цепи Маркова, как и в случае режима селективного отказа, равновероятны и инвариантны к размеру окна: $P_i = \frac{1}{D}$, $i = \overline{0, D-1}$. Для $S = D + 1$ вероятности состояний определяются двумя равномерно распределенными областями значений:

$$P_i = P_0, \quad i = \overline{0, W-2};$$

$$P_i = P_0 \frac{1 - \Phi}{1 - F_{\Pi} + F_{\Pi} F_o \Phi}, \quad i = \overline{W-1, D-1};$$

$$P_D = P_0 \frac{(1 - \Phi)(1 - F_o)}{1 - F_{\Pi} + F_{\Pi} F_o \Phi}.$$

5. Сравнительный анализ пропускной способности режимов отказа

Важнейшей операционной характеристикой протокола является его пропускная способность, определяемая параметрами тракта передачи данных, накладными расходами и особенностями протокольных процедур управления передачей [1, 2]. Нормированное быстродействие транспортного соединения определяется средним числом доставленных получателю неискаженных сегментов (с учетом режима отказа [2]) за среднее время между двумя последовательными поступлениями подтверждений [3–6]. Поскольку время между приходами подтверждений распределено по геометрическому закону с параметром F_o , то среднее время между приходами подтверждений в длительностях цикла t составит $\bar{T} = 1/F_o$. Тогда для селективной процедуры отказа пропускная способность определится зависимостью

$$Z_c(W, S) = F_o \left\{ \sum_{i=D-1}^{D+W-2} (i - D + 2) F_{\Pi} P_i + W F_{\Pi} \sum_{i=D+W-1}^{S-1} P_i \right\}.$$

С учетом вариативности выражений для вероятностей состояний цепи Маркова отсюда при различных связях между протокольными параметрами и

круговой задержкой получаем функциональные зависимости данного показателя:

$$(23) \quad Z_c(W, S) = \begin{cases} \frac{F_{\Pi} \{ (1 - F_o)^D - (1 - F_o)^{S+1} [1 + (S - D + 1) F_o] \}}{(1 - F_o)^D [F_o(D - W - 1) + 2] + (1 - F_o)^S [F_o(D + W - S + 1) - 2]}, \\ W < D, D \leq S \leq D + W - 1; \\ \frac{F_{\Pi} \{ 1 - (1 - F_o)^W - W F_o (1 - F_o)^{S-D+1} \}}{2 + F_o(D - W - 1) - (1 - F_o)^W - (1 - F_o)^{S-D+1}}, \\ W < D, S \geq D + W - 1; \\ \frac{F_{\Pi} \{ (1 - F_o)^D - (1 - F_o)^{S+1} [1 + (S - D + 1) F_o] \}}{(1 - F_o)^D + (1 - F_o)^{W+1} + (1 - F_o)^S [F_o(D + W - S + 1) - 2]}, \\ W \geq D, W + 1 \leq S \leq D + W - 1; \\ \frac{F_{\Pi} \{ 1 - (1 - F_o)^W - W F_o (1 - F_o)^{S-D+1} \}}{1 - (1 - F_o)^W + (1 - F_o)^{W-D+1} [1 - (1 - F_o)^{S-W}]}, \\ W \geq D, S \geq D + W - 1. \end{cases}$$

Пропускная способность транспортного соединения для режима группового отказа с учетом повторной передачи всех сегментов начиная с первого неприятого [2] задается зависимостью

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

С точностью до сомножителя P_0 отсюда получаем:

$$(24) \quad Z_g(W, S) = \begin{cases} \frac{P_0 F_{\Pi} \{ 1 - F_{\Pi} - (1 - \Phi)(1 - F_o)^{S-D+1} + F_{\Pi} F_o \Phi^{S-D+1} \}}{(1 - F_{\Pi}) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{S-D}]}, \\ \max\{W + 1, D\} \leq S \leq D + W - 1; \\ \frac{P_0 F_{\Pi} \{ (1 - F_{\Pi})(1 - \Phi^W) - (1 - \Phi)(1 - F_{\Pi}^W)(1 - F_o)^{S-D+1} \}}{(1 - F_{\Pi}) [1 - F_{\Pi} + F_{\Pi} F_o \Phi^{W-1}]}, \\ S \geq D + W - 1. \end{cases}$$

Нетрудно убедиться, что при $F_{\Pi} = 1$ из данного соотношения, устраняя неопределенность вида $0/0$, получаем зависимость для режима селективного отказа. Неограниченный рост ширины окна ($W \rightarrow \infty$), а следовательно, и длительности тайм-аута S , обеспечивает инвариантную к круговой задержке и достоверности передачи данных в обратном тракте зависимость пропускной способности в селективном режиме отказа, определяемую только достоверностью передачи сегментов в прямом направлении $Z_c(\infty, \infty) = F_{\Pi}$. То же самое имеет место для селективного режима отказа, абсолютно надежного обратного сквозного канала ($F_o = 1$) и ширины окна не меньшей круговой задержки ($W \geq D$). При $F_o = 1$ и $W \leq D$ пропускная способность транспортного соединения в селективном режиме отказа (23) обратно пропорциональна разности

общей длины пути и ширины окна, сниженной на единицу:

$$Z_c(W \leq D, S) = \frac{F_{\Pi}}{D - W + 1}.$$

При $F_o = 1$ для различных размеров окна пропускная способность режима группового отказа (24) составит:

$$Z_g(W, S) = \begin{cases} \frac{F_{\Pi}}{D - W + 1 + (W - 1)(1 - F_{\Pi})}, & W \leq D; \\ \frac{F_{\Pi}}{1 + (D - 1)(1 - F_{\Pi})}, & W \geq D. \end{cases}$$

При единичной ширине окна ($W = 1$) селективный и групповой режимы отказа совпадают и пропускная способность примет вид

$$Z(1, S) = \begin{cases} \frac{F_{\Pi} \{1 - (1 - F_o)^{S-D+1} [1 + (S - D + 1)F_o]\}}{2 + F_o(D - 2) + (1 - F_o)^{S-D} [F_o(D + 2 - S) - 2]}, & D \leq S \leq D + W - 1; \\ \frac{F_{\Pi} F_o [1 - (1 - F_o)^{S-D+1}]}{1 + F_o(D - 1) - (1 - F_o)^{S-D+1}}, & S \geq D + W - 1. \end{cases}$$

При минимальных размерах окна ($W = 1$) и тайм-аута ($S = D$) пропускная способность протокола составит

$$Z(1, D) = \frac{F_{\Pi} F_o}{D}.$$

То же самое соотношение имеет место в случае $W < D$ и $S = D$ для селективного и группового режимов отказа. Из чего следует, что тайм-аут минимальной длительности обеспечивает инвариантности протокола к режиму отказа, а пропускной способности — к ширине окна. При $W = 1$ и неограниченной длительности тайм-аута ($S = \infty$) пропускная способность управляющей процедуры преобразуется к

$$Z(1, \infty) = \frac{F_{\Pi} F_o}{1 + F_o(D - 1)}.$$

В режиме группового отказа предельные возможности управляющей процедуры, соответствующие неограниченной ширине окна ($W \rightarrow \infty$), определяются зависимостью

$$Z_g(\infty, \infty) = \frac{F_{\Pi} F_o}{1 - F_{\Pi}(1 - F_o) + F_o(D - 1)(1 - F_{\Pi})}.$$

Отсюда нетрудно видеть значительную зависимость пропускной способности от круговой задержки. При одинаковом качестве прямого и обратного трактов передачи данных ($F_{\Pi} = F_o = F$) потенциальная пропускная способность режима группового отказа

$$Z_g(\infty, \infty) = \frac{F^2}{1 + F(D - 2)(1 - F)}$$

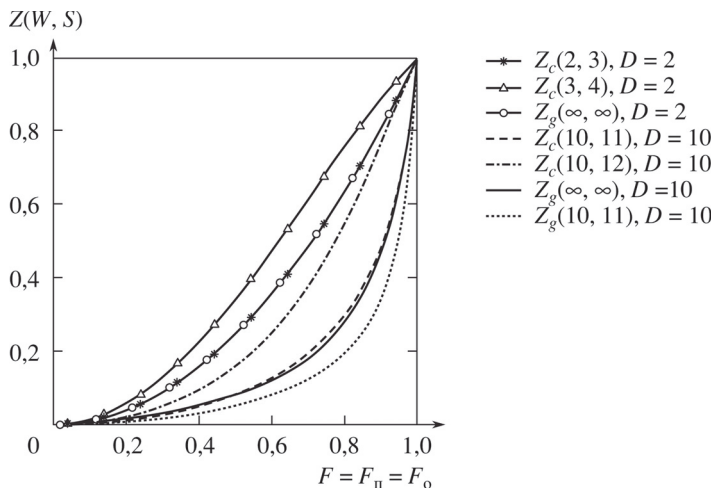


Рис. 1. Сравнительные зависимости быстродействия транспортного соединения в различных режимах повтора от достоверности передачи данных.

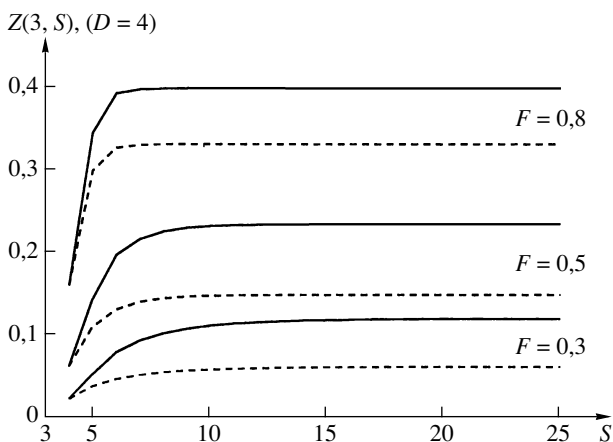


Рис. 2. Зависимость пропускной способности транспортного соединения от размера тайм-аута при $F_{\pi} = F_{\circ} = F$. Сплошные линии соответствуют селективному режиму, а штриховые — групповому.

в режиме селективного отказа практически достигается уже при наборе протокольных параметров $W = D$ и $S = D + 1$ (см. рис. 1):

$$Z_c(D, D + 1) = \frac{F^2(3 - 2F)}{1 + D(1 - F)}.$$

При этом для однозвенного прямого и обратного тракта ($D = 2$) пропускная способность обоих режимов отказа с указанными протокольными параметрами совпадает на всей области изменения достоверности передачи данных F , а для многозвенного тракта ($D \geq 3$) равенство пропускных способностей имеет место в точках $F = 0; 0,5; 1$. В области $F \in (0; 0,5)$ режим группового от-

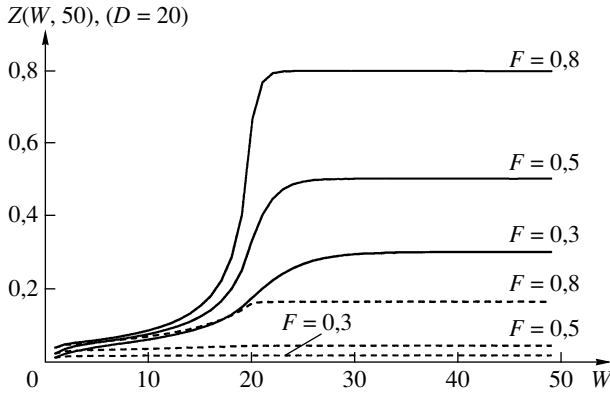


Рис. 3. Зависимость пропускной способности транспортного соединения от ширины окна при $F_{\text{п}} = F_{\text{о}} = F$. Сплошные линии соответствуют селективному режиму, а штриховые — групповому.

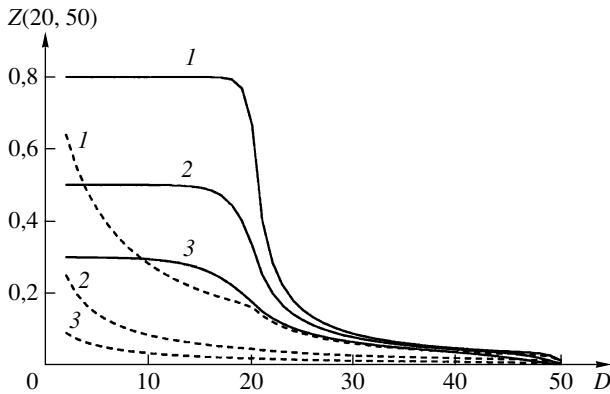


Рис. 4. Зависимость пропускной способности транспортного соединения от длительности круговой задержки при $F_{\text{п}} = F_{\text{о}} = F$. Сплошные линии соответствуют селективному режиму, а штриховые — групповому. Графики 1 соответствуют $F = 0,8$, графики 2 — $F = 0,5$ и графики 3 — $F = 0,3$.

каза несущественно превышает селективный отказ по показателю пропускной способности, а в области $F \in (0,5; 1)$ — селективный отказ незначительно превосходит групповой по указанному критерию (рис. 1). С ростом длительности тайм-аута S точка F промежуточного равенства пропускной способности различных режимов отказа смещается влево (к нулевому значению), а область превосходства и абсолютное преимущество селективного отказа по показателю пропускной способности увеличивается (см. рис. 1). Отметим, однако, что при высоком уровне достоверности передачи данных ($F \rightarrow 1$) различие между значениями пропускной способности режимов отказа быстро снижается. Характерная зависимость пропускной способности транспортного соединения от длительности сквозного тайм-аута имеет вид кривой с насыщением (см. рис. 2) для любого режима отказа. Зависимость пропускной способности тракта передачи данных от размера окна имеет пологий участок

роста при ширине окна, меньшей круговой задержки ($W < D$), с дальнейшим ростом W пропускная способность резко увеличивается в окрестности длительности круговой задержки до насыщения, соответствующего тому или иному режиму отказа (см. рис. 3). С ростом длины тракта передачи данных (а следовательно, и длительности круговой задержки) его пропускная способность снижается, существенно деградируя после того как круговая задержка превысит размер окна ($D > W$). Характерная зависимость представлена на рис. 4.

6. Выбор размера окна и длительности тайм-аута

Поскольку зависимость функции пропускной способности (23) и (24) от протокольных параметров имеет вид кривых с насыщением без особых точек, то рациональные значения размера окна W_0 и длительности тайм-аута S_0 могут быть выбраны по заданному уровню пропускной способности от потенциальных возможностей транспортного соединения в различных режимах отказа, достигаемых при неограниченных значениях протокольных параметров. Очевидно, что решение необходимо искать в зоне насыщения показателя пропускной способности, соответствующей области изменения протокольных параметров с ограничениями снизу (1). Желаемый уровень от потенциальной пропускной способности в произвольном режиме повтора $Z(\infty, \infty)$ при этом необходимо задавать для каждой координаты функции пропускной способности. Так как протокольные параметры связаны неравенством $S > W$, то процедура выбора разбивается на два этапа. На первом — по заданному уровню $y_W < 1$ определяется рациональная ширина окна W_0 из условия

$$(25) \quad Z(W_0, \infty) = y_W Z(\infty, \infty),$$

а на втором — рациональная длительность тайм-аута S_0 при рассчитанном на первом этапе значении W_0 по заданному уровню $y_S < 1$ из уравнения

$$(26) \quad Z(W_0, S_0) = y_S Z(W_0, \infty) = y_W y_S Z(\infty, \infty).$$

Здесь коэффициенты y_W и y_S , задающие уровни достигаемой пропускной способности по координатам W и S , в соответствии с условием (1) имеют ограничения снизу:

$$y_W \geq \frac{Z(D, \infty)}{Z(\infty, \infty)}, \quad y_S \geq \frac{Z(W_0, D + W_0 - 1)}{Z(W_0, \infty)}.$$

Уравнения (25), (26) являются критериями последовательного выбора ширины окна и длительности тайм-аута соответственно. Из данных уравнений для селективного режима отказа с учетом (23) получаем аналитические зависимости рациональных значений протокольных параметров:

$$W_0 = \left\lceil \frac{1}{\ln(1 - F_0)} \ln \frac{(1 - y_W)(1 - F_0)^{D-1}}{y_W + (1 - y_W)(1 - F_0)^{D-1}} \right\rceil,$$

$$S_0 = D - 1 + \left\lceil \frac{1}{\ln(1 - F_0)} \ln \frac{(1 - y)[1 - (1 - F_0)^{W_0}] - y(1 - F_0)^{W_0 - D + 1}}{W_0 F_0 - y} \right\rceil,$$

где $\lfloor \dots \rfloor$ — означает округление до большего целого, $y = y_W y_S$. В случае группового режима отказа расчет параметров согласно (25), (26) с учетом (24) выполняется по следующим соотношениям:

$$W_o = \left\lfloor \frac{1}{\ln \Phi} \times \right. \\ \times \ln \frac{(1 - y_W) \Phi^{D-1} (1 - \Phi) [1 - \Phi + (D - 1)(1 - F_n) F_o]}{\Phi^{D-1} (1 - \Phi) [1 - \Phi + (D - 1)(1 - F_n) F_o] + y_W F_n (1 - F_o)^2 (1 - \Phi^{D-1})} \left. \right\rfloor, \\ S_o = D - 1 + \left\lfloor \frac{1}{\ln(1 - F_o)} \ln(1 - F_n) \times \right. \\ \times \frac{(1 - y - \Phi^{W_o})(1 - \Phi) [1 - \Phi + (D - 1)(1 - F_n) F_o] - y F_n (1 - F_o)^2 (\Phi^{W_o - D + 1} - \Phi^{W_o})}{(1 - F_n^{W_o})(1 - \Phi) [1 - \Phi + (D - 1)(1 - F_n) F_o] - y(1 - F_n)(1 - \Phi)^2} \left. \right\rfloor.$$

7. Заключение

В статье предложена модель процесса переноса сегментов данных в транспортном соединении, управляемом надежным транспортным протоколом с подтверждением данных, принятых получателем, в режимах селективного и группового отказов. Математическая модель основана на описании очереди переданных, но не подтвержденных сегментов данных цепью Маркова с конечным числом состояний и дискретным временем. Получены стационарные распределения состояний марковской цепи для различных областей изменения размера окна и длительности тайм-аута. Найдены аналитические выражения для пропускной способности транспортного соединения в различных режимах отказа. Предложен метод выбора параметров размера окна и длительности тайм-аута. Для тракта передачи данных с одинаковым уровнем достоверности доставки сегментов в прямом и обратном направлениях ($F_n = F_o$) из анализа численных результатов установлено, что потенциально достижимая пропускная способность широко используемого на практике режима группового отказа, соответствующая неограниченным значениям параметров протокола ($W = S = \infty$), в режиме селективного отказа практически достигается при ширине окна, совпадающей с длительностью круговой задержки ($W = D$), и минимальной длительности тайм-аута ($S = W + 1$). Обнаружена существенная зависимость пропускной способности в режиме группового отказа от общей длины транспортного соединения D , что при конвейерной интерпретации тракта передачи данных легко объясняется необходимостью перезагрузки всего или части транспортного соединения в случае потери хотя бы одного сегмента в отправленной последовательности (повторной передачи всех сегментов начиная с первого потерянного). В целом, пропускная способность при любом режиме отказа в значительной мере определяется соотношением между шириной окна и длительностью круговой задержки. В качестве направлений дальнейшего развития исследований следует выделить задачу анализа пропускной способности транспортных соединений абонентов, имеющих хотя бы часть общего сетевого маршрута и соперничающих за доступную полосу пропускания разделяемых участков переприема. Кроме

того, важным представляется поиск области целесообразности применения протокольных процедур с интегрированным механизмом опережающего исправления ошибок (Forward Error Correction) [23], основанного на передаче последовательности групп из информационных и дополнительных (избыточных) сегментов, допускающих восстановление на принимающей стороне при искажении подмножества сегментов группы. Дополнительно к этому актуальным является исследование влияния на задержку мультисегментных абонентских сообщений конвейерного эффекта, проявляющегося при передаче в многозвенных транспортных соединениях.

СПИСОК ЛИТЕРАТУРЫ

1. *Fall K., Stevens R.* TCP/IP Illustrated, Volume 1: The Protocols (2nd Edition). Addison-Wesley Professional Computing Series, 2012.
2. *Богуславский Л.Б.* Управление потоками данных в сетях ЭВМ. М.: Энергоатомиздат, 1984.
3. *Boguslavskij L.B., Gelenbe E.* Analytical Models of Data Link Control Procedures in Packet-switching Computer Networks // Autom. Remote Control. 1980. V. 41. No. 7. P. 1033–1042.
4. *Gelenbe E., Labetoulle J., Pujolle G.* Performance Evaluation of the HDLC Protocol // Comput. Networks. 1978. V. 2. Iss. 4/5. P. 409–415.
5. *Кокшениев В.В., Сущенко С.П.* Анализ быстродействия асинхронной процедуры управления звеном передачи данных // Вычислительные технологии. 2008. Т. 15. Спец. выпуск № 5. С. 61–65.
6. *Кокшениев В.В., Михеев П.А., Сущенко С.П.* Анализ селективного режима отказа транспортного протокола в нагруженном тракте передачи данных // Вестн. ТГУ. Сер. Управление, вычислит. техника и информатика. 2013. № 3 (24). С. 78–94.
7. *Calleari C., Giordano S., Pagano M., et al.* A Survey of Congestion Control Mechanisms in Linux TCP // Communications in Computer and Information Science: Distributed Computer and Communication Networks 17th Int. Conf., DCCN 2013, Moscow, Russia, October 7–10, 2013. Revised Selected Papers. Vishnevsky V.; Kozыrev D.; Larionov A. (Eds.). 2014. P. 28–42.
8. *Vasenin V.A., Simonova G.I.* Mathematical Models of Traffic Control in Internet: New Approaches Based of TCP/AQM Schemes // Autom. Remote Control. 2005. V. 66. No. 8. P. 1274–1286.
9. *Bogoyavlenskaya O.Yu.* Analysis of the Random Flow Generated by the Feedback Transport Protocol in a Data Transfer Network // Autom. Remote Control. 2003. V. 64. No. 12. P. 1882–1889.
10. *Bogoyavlenskaya O.Yu.* Probabilistic Model of the Algorithms of Distributed Control Protocol in the Internet Network // Autom. Remote Control. 2009. V. 70. No. 1. P. 107–117.
11. *Mikadze I.S., Khocholava V.V.* On a Model of Information Transmission Through Unreliable Communication Channel // Autom. Remote Control. 2004. V. 65. No. 8. P. 1250–1254.
12. *Arvidsson A., Krzesinski A.* A Model of a TCP Link // Proc. 15th Int. Teletraffic Congr. Specialist Seminar. 2002.
13. *Altman E., Avrachenkov K., Barakat C.* A Stochastic Model of TCP/IP with Stationary Random Loss // Comput. Communication Review. 2000. V. 30. No. 4. P. 231–242.

14. *Olsen Y.* Stochastic Modeling and Simulation of the TCP Protocol // Uppsala Dissertations in mathematics 28. 2003.
15. *Kassa D.F.* Analytic Models of TCP Performance // PhD Thesis, University of Stellenbosch. 2005.
16. *Bogoiavlenskaia O.* Discrete Model of TCP Congestion Control Algorithm with Round Dependent Loss Rate // Internet of Things, Smart Spaces, and Next Generation Networks and Systems. 2015. V. 9247 of the series Lect. Notes in Computer Science. P. 190–197.
17. *Giordano S., Pagano M., Russo F., et al.* Modeling TCP Startup Performance // J. Math. Sci. 2014. V. 200. Iss. 4. P. 424–431.
18. *Kravets O.Ya.* Mathematical Modeling of Parameterized TCP Protocol // Autom. Remote Control. 2013. V. 74. No. 7. P. 1218–1224.
19. *Wang J., Wen J., Han Y., et al.* Achieving High Throughput and TCP Reno Fairness in Delay-based TCP over Large Networks // Frontiers Comput. Sci. 2014. V. 8. Iss. 3. P. 426–439.
20. *Nikitinskiy M.A., Chalyy D.Ju.* Performance Analysis of Trickles and TCP Transport Protocols under High-load Network Conditions // Automat. Control Comput. Sci. 2013. V. 47. No. 7. P. 359–365.
21. *Kokshenev V.V., Suschenko S.P.* Analytical Model of the TCP Reno Congestion Control Procedure through a Discrete-Time Markov Chain // Communications in Computer and Inform. Sci.: Distributed Computer and Communication Networks 17th Int. Conf., DCCN 2013, Moscow, Russia, October 7–10, 2013. Revised Selected Papers. Vishnevsky V.; Kozyrev D.; Larionov A. (Eds.) 2014. P. 124–135.
22. *Ivanovskii V.B.* Properties of Output Flows in Discrete Queueing Systems // Autom. Remote Control. 1984. V. 45. No. 11. Part 1. P. 1413–1419.
23. *Lundqvist H., Karlsson G.* TCP with End-to-End FEC // Communications, 2004 Int. Zurich Seminar on. 2004. P. 152–156.

Статья представлена к публикации членом редколлегии В.М. Вишневым.

Поступила в редакцию 02.11.2015