References

1. Dergunova E.E., Morozova K.S. The art of LAN diagnostics. *Nauka i obrazovanie v sovremenom mire*, 2015, no. 6, pp. 16–17. (In Russ.)
2. Huang L. et al. Design and implementation of multicast routing system over SDN and sFlow. *8th IEEE International Conference on Communication Software and Networks (ICCSN)*, Beijing, 2016, pp. 524–529.
3. Gan'zha D. FLUKE NETWORKS prepares proposals for the Russian market. *Zhurnal setevykh reshenij LAN*, 2017, no. 10, pp. 2–9. (In Russ.)
4. Chikov A.E. Hardware and software for corporate network analysis. *Sistemy upravleniya, tehnicheskie sistemy: ustojchivost', stabilizatsiya, puti i metody issledovaniya*, 2016, pp. 258–262. (In Russ.)
5. Flow Analysis. Observer Analyzer integrates NetFlow // *Softing IT Networks GmbH*. URL: https://itnetworks.softing.com/fileadmin/media/documents/products/viavi/Analyzer/VIAVI_Observer-Analyzer-brochure_flow_analysis_Softing_IT_Networks_englisch.pdf (accessed: 01.10.2020).
6. Lawrimore J.H. et al. Quality control and processing of cooperative observer program hourly precipitation data. *Journal of Hydrometeorology*, 2020, vol. 21, no. 8, pp. 1811–1825.
7. Demaria E.M.C., Goodrich D.C., Kunkel K.E. Evaluating the reliability of the US cooperative observer program precipitation observations for extreme events analysis using the LTAR network. *Journal of Atmospheric and Oceanic Technology*, 2019, vol. 36, no. 3, pp. 317–332.
8. Slabicki M., Grochla K. Performance evaluation of CoAP, SNMP and NETCONF protocols in fog computing architecture. *NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium*, 2016, pp. 1315–1319.
9. Preobrazhenskij Yu.P. Problems of analyzing the health of computer networks. *Nauka molodyh – budushee Rossii*, 2019, pp. 141–144. (In Russ.)
10. Malahov S.V., Tarasov V.N., Kartashevskij I.V. Theoretical and experimental study of latency in software-configurable networks. *Infokommunikacionnye tehnologii*, 2015, vol. 13, no. 4, pp. 409–413. (In Russ.)
11. Shukhman A. et al. Development of network security tools for enterprise software-defined networks. *8th International conference on security of information and networks SIN*, Sochi, 2015, pp. 224–228. DOI: <https://doi.org/10.1145/2799979.2800009>
12. Tarasov V. et al. Analysis of intervals between traffic packets on the SDN networks depending on the TCP window size. *Problems of Infocommunications Science and Technology (PIC S&T). 2016 Third International Scientific-Practical Conference*, Kharkiv, Ukraine, 2016, pp. 15–17. DOI: <https://doi.org/10.1109/INFOCOMMST.2016.7905322>

Received 21.10.2020

УДК 519.872

АНАЛИЗ КОРРЕЛИРОВАННОЙ ОЧЕРЕДИ В СИСТЕМЕ G/G/1

Буранова М.А., Резяпкина М.И.

Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ

E-mail: buranova-ma@psuti.ru, rez8pkina@mail.ru

В статье рассматриваются коррелированные очереди, в которых время обслуживания пакета сильно коррелировано с временем между поступлениями пакетов из-за конечной пропускной способности входных и выходных каналов. Рассмотрены подходы к оценке параметров функционирования сетей с коррелированными очередями. Представлен аналитический метод для получения оценки времени ожидания пакета в сети на основе подхода с использованием копул. Исследованы корреляционные свойства реального видеотрафика, определены законы распределений для интервалов времени между пакетами и времен обработки пакетов, наиболее точную аппроксимацию дало распределение Вейбулла. Определено, что интервалы времени между пакетами и времена обработки

пакетов имеют значительную взаимную корреляцию. Получены оценки для среднего времени ожидания пакета в сети и для вариации задержки пакета при передаче в сети G/G/1.

Ключевые слова: качество обслуживания (*QoS – Quality of Service*), задержка, вариация задержки, коррелированная очередь

Введение

При проектировании инфокоммуникационных сетей одной из актуальных задач является точное прогнозирование параметров их функционирования, таких как оценка параметров качества обслуживания *QoS* (*Quality of Service*): задержка, изменение задержки передачи пакетов (джиттер), вероятность потерь пакетов. При решении данных задач возникают проблемы, обусловленные особенностями обрабатываемых в инфокоммуникационных сетях потоков, отличие потоков от пуассоновского, наличие свойств самоподобия, корреляционные связи внутри последовательностей обслуживаемых пакетов. Большинство работ по анализу инфокоммуникационных сетей и оценке параметров их функционирования основаны на предположении о независимости потоков [1–4], в то время как многочисленные исследования доказывают обратное [5–8].

Для моделирования систем обработки потоков в инфокоммуникационных сетях традиционно используют системы массового обслуживания, для аналитического описания которых необходимо определить последовательности интервалов времени между пакетами и времен обработки пакетов. Следует отметить, что современные инфокоммуникационные сети лучше описываются моделью очереди G/G/1. Имеются подтверждения о наличии зависимостей внутри данных последовательностей [3; 9; 10] и между ними [5–8]. Следует обратить внимание, что предположения о независимости могут привести к неоправданно оптимистичным прогнозам оценки параметров работы сетей.

Анализ данных проблем можно найти в [6; 7] и ссылках на них. В [6] авторы рассмотрели систему с пуассоновскими поступлениями, в которой время обслуживания B_n n -го пакета пропорционально времени A_n между прибытиями $(n - 1)$ -го и n -го пакетов, то есть $B_n = \zeta A_n$, где ζ – положительная постоянная. Было получено преобразование Лапласа распределения времени пребывания в системе, которое состоит из времени ожидания в очереди плюс время обработки пакета путем решения линейного функционального уравнения [7], полученного из уравнения эволюции системного времени. Однако они рассматривают только случай $\zeta \neq 1$ из-за технических трудностей, связанных с решением функционального уравнения.

Есть и другие работы [1; 8], основанные на моделях с учетом корреляций. Однако, как показано в [6], данные модели слабо учитывают особенности процессов прибытия и обслуживания и, следовательно, приводят к неточности в прогнозировании производительности системы. И их нельзя использовать для получения фактического времени ожидания.

Одна из самых ранних работ по систематическому исследованию вопроса о корреляции между временем обслуживания пакетов и временем между пакетами – это работа [1], в которой Л. Клейнрок изучал влияние коррелированных длин сообщений и времен между прибытиями в контексте модели сети очередей для сетей связи. Там же он показал сложность этой проблемы.

В [5] были показаны некоторые подходы к анализу функционирования сетей при наличии корреляций внутри последовательностей интервалов времени между пакетами и времен обработки пакетов, при этом корреляции между данными последовательностями учитываются только для оценки задержки, а для оценки такого важного параметра качества обслуживания в инфокоммуникационных сетях, как изменение задержки при передаче пакетов (джиттер, вариация задержки), не показан.

Более общие корреляции были рассмотрены в [11] с использованием двумерного экспоненциального распределения для характеристики корреляции между временем прибытия и обслуживанием. Впоследствии эта работа была расширена в нескольких статьях. В [12] было показано, что плотность задержки имеет гиперэкспоненциальное распределение, а в [13] изучалась чувствительность этого распределения к величине коэффициента корреляции. Период занятости исследовался в [14], а система с бесконечным числом серверов рассматривалась в [15].

Аналогично подходу в [5] рассмотрим канал связи с конечной пропускной способностью, соединяющий два узла – узел 1 и узел 2 (см. рисунок 1). Пакеты передаются от узла 1 к узлу 2. В такой системе время между прибытиями составляет два последовательных пакета в узле 2 – это сумма времени передачи (время обслуживания в узле 1) первого пакета и времени между двумя пакетами.

Здесь время между появлением пакетов определяется периодом времени между периодом

Рисунок 1. Корреляционная модель очереди

окончания передачи первого пакета и периодом начала передачи второго пакета (см. рисунок 1). Поскольку канал, соединяющий два узла, имеет конечную пропускную способность, скажем τ , канал не может отправлять количество пакетов от узла 1 к узлу 2 превышает τT , и, следовательно, для больших пакетов требуется более длительное время передачи в обоих узлах. Следовательно, время между прибытиями в узле 2 сильно коррелирует со временем обслуживания в узле 2. Определение задержки и изменения задержки при передаче пакетов по сети на основе данной модели представляется достаточно перспективным.

Модель коррелированной очереди

Анализируемая модель представляет собой систему очередей в узле 2, в которой хранятся пакеты, поступающие по входному каналу с конечной пропускной способностью к выходному каналу, имеющему одинаковую со входным каналом пропускную способность. Процесс поступления пакетов в очередь моделируется чередующимся процессом, имеющим два типа периодов: ON-периоды передачи (наличия пакетов) и OFF-периоды между передачей пакетов.

С точки зрения теории очередей можно представить последовательности ON-периода как последовательность времен обработки пакетов, последовательности OFF-периодов как последовательностей интервалов времени между пакетами. Периоды OFF фиксируют время между появлением последовательных пакетов в узле 1, а периоды ON фиксируют время передачи пакетов в узле 1, так что время между поступлениями двух последовательных пакетов в очереди состоит из периода OFF и ON.

Поскольку очередь в узле 2 может передавать пакет через свой выходной канал только после того, как получит последний бит пакета, при анализе можно предположить, что в конце каждого периода ON есть прибывший пакет, время

обслуживания которого является таким же, как и ON-период, который следует за ним. Кроме того, можно допустить длительность периода OFF равной нулю, так как в случае, если у узла 1 имеются пакеты для передачи или имеется пачка пакетов, он может непрерывно передавать их узлу 2.

При анализе данной модели воспользуемся подходом, показанным в [5] для аналогичного случая, где введено предположение, что количество последовательных периодов ON между двумя периодами OFF положительной длительности имеет некоторое произвольное распределение, а периоды OFF положительной длины распределены экспоненциально. Каждый период ON аппроксимируется гиперэкспоненциальным распределением.

Данная модель разработана для оценки средней задержки пакетов в сети. Изменение задержки может быть определено как вариация задержки пакета [10]. Вариация задержки пакета в потоке $\sigma(X) = \sqrt{D(X)}$, где $D(X)$ – дисперсия задержки пакетов. Если ввести обозначение $f_T(x)$ – плотность вероятностей случайной величины T (время ожидания пакета в сети) и с учетом, что задержка пакета определяется случайной величиной T , дисперсию задержки можно определить согласно выражению

$$\begin{aligned} D(X) &= \int_0^\infty [x - M(x)]^2 f_T(x) dx = \\ &= \int_0^\infty x^2 f_T(x) dx - \left(\int_0^\infty x f_T(x) dx \right)^2. \end{aligned} \quad (1)$$

Чтобы исследовать влияние корреляций на производительность системы, необходимо рассмотреть и проанализировать аналогичную очередь GI/G/1.

Рассмотрим коррелированную очередь узла в сети, где процесс ввода в очередь управляется чередующимися процессами ON и OFF, а пропускная способность входного и выходного каналов одинакова. Как упоминалось в разделе 1, перио-

ды ON фиксируют время между появлением пакетов, а периоды OFF фиксируют время передачи пакетов на предыдущем узле в сети. Поскольку предполагается, что система очередей может обслуживать пакет только после того, как получит весь пакет, очевидно, что всегда будет новый пакет в конце каждого периода ON.

Так как пропускная способность входного канала и выходного канала одинакова, время обслуживания пакетов равно периодам ON, за которыми они следуют. В конце каждого периода ON начинается либо период OFF, либо другой период ON, так как допускаются периоды ON с нулевой длиной. Следует обратить внимание, что последовательные периоды ON захватывают последовательные передачи пакетов в предыдущем узле сети. Важно, что после каждого периода OFF всегда начинается новый период ON.

При анализе предполагается, что периоды ON генерируются в соответствии с гиперэкспоненциальным распределением, имеющим M экспоненциальных составляющих, то есть функция плотности $f(x)$ периодов ON определяется выражением

$$f(x) = \sum_{i=0}^{M-1} \theta_i \delta_i e^{-\delta_i x}, \quad \sum_{i=0}^{M-1} \theta_i = 1, \quad \theta_i \geq 0.$$

В конце каждого периода ON либо период OFF начинается с вероятностью $q = 1 - p$, либо другой период ON начинается с вероятностью p . Тогда общее число периодов ON между двумя последовательными периодами OFF, длина которых больше, чем ноль, имеет распределение, отличное от экспоненциального. Предполагается, что периоды OFF положительной длины генерируются в соответствии с экспоненциальным распределением с параметром β .

Пусть I_n и J_n ($n \geq 1$) обозначают длительность n -го периодов ON и OFF соответственно. Пусть начальные периоды ON – вложенные точки и пусть W_n обозначает незавершенную работу системы, определяемую количеством работы в системе в n -й вложенной точке ($n \geq 0$). Время 0 – это нулевая вложенная точка, и мы предполагаем, что первый период ON начинается в момент времени 0. Тогда W_n удовлетворяет следующему уравнению:

$$W_0 = 0, \quad W_{n+1} = \left[(W_n - I_{n+1})^+ + I_{n+1} - J_{n+1} \right]^+, \quad n \geq 0, \quad (2)$$

где $A^+ = \max(A, 0)$. Пусть S_{n+1} обозначает тип экспоненциального распределения, которое принимает I_{n+1} . Тогда $S_{n+1} = i$ с вероятностью θ_i , $0 \leq i \leq M-1$, и, следовательно, I_{n+1} соответству-

ет экспоненциальному распределению с параметром δ_i .

Пусть $\mathcal{W}_n(s)$ – преобразование Лапласа W_n для $R(s) \geq 0$. Тогда, приняв, что $S_{n+1} = i$, получим

$$\begin{aligned} \mathcal{W}_{n+1}(s) = & \sum_{i=0}^{M-1} \theta_i P\{W_n = 0\} E[e^{(-s[I_{(n+1)} - J_{(n+1)})^+}] | S_{(n+1)} = i] + \\ & + \sum_{i=0}^{M-1} \theta_i \int_{\omega=0^+}^{\infty} dW_n(\omega) \times \\ & \times \left\{ \int_{x=0}^{\omega} \delta_i e^{(-\delta_i x)} E[e^{(-s(\omega - J_{(n+1)})^+}] dx + \right. \\ & \left. + \int_{x=\omega}^{\infty} \delta_i e^{(-\delta_i x)} E[e^{(-s(x - J_{(n+1)})^+}] dx \right\}. \end{aligned} \quad (3)$$

Наблюдая за тем, что каждый следующий период ON наступает после периода ON $n+1$ с вероятностью p (в этом случае мы имеем $J_{n+1} = 0$) и что период OFF наступает после периода ON $n+1$ с вероятностью $q = 1 - p$ (в этом случае длина J_{n+1} соответствует экспоненциальному распределению с параметром β), получаем, что

$$\mathcal{W}(s) = \sum_{l=0}^{\infty} \prod_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(s + \sum_{j=0}^{k-1} \delta_{i_j} \right) c, \quad (4)$$

где c – константа (с учетом, что все задействованные параметры заданы), определяемая по формуле

$$\begin{aligned} c = \mathcal{W}(\beta) - \sum_{i=0}^{M-1} \theta_i \frac{\beta}{\delta_i + \beta} \mathcal{W}(\delta_i + \beta); \\ a_i(s) = \theta_i \frac{\beta - ps}{q(\delta_i + \beta)}. \end{aligned}$$

Предполагая, что используется k экспонент в гиперэкспоненциальном распределении, β – параметр распределения, получим

$$\prod_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(s + \sum_{j=0}^{k-1} \delta_{i_j} \right) c.$$

В итоге из (4) следует окончательное выражение для $\mathcal{W}(s)$

$$\mathcal{W}(s) = \frac{\sum_{l=0}^{\infty} \prod_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(s + \sum_{j=0}^{k-1} \delta_{i_j} \right)}{\prod_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(s + \sum_{j=0}^{k-1} \delta_{i_j} \right)}. \quad (5)$$

Среднее время ожидания пакета в очереди можно определить как первую производную $\mathcal{W}(s)$ при условии $s = 0$

$$E[W] = \frac{\sum_{l=1}^{\infty} \sum_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(\sum_{j=0}^{n-1} \delta_{i_j} \right) \sum_{i_k=0}^{M-1} \theta_{i_k} (p\delta_{i_k} + \beta)}{q \sum_{j=0}^k \delta_{i_j}} \cdot \frac{1}{\sum_{l=0}^{\infty} \prod_{k=0}^{l-1} \sum_{i_k=0}^{M-1} a_{i_k} \left(\sum_{j=0}^{k-1} \delta_{i_j} \right)}. \quad (6)$$

Учитывая, что общую сетевую задержку в соответствии с моделью коррелированной очереди рисунок 1 и согласно выражению (2) можно записать как

$$Z = (W - I)^+ + I,$$

где I – это общая случайная величина для I_n , можно получить преобразование Лапласа для Z , аналогично подходу для $W(s)$ в виде

$$\begin{aligned} \mathcal{Z}(s) &= E \left[e^{-s[(W-I)^+ + I]} \right] = \\ &= \frac{\beta - s}{\beta - ps} \mathcal{W}(s) + \frac{sq}{\beta - ps} c, \\ E[Z] &= E[W] + \frac{q(1-c)}{\beta}, \end{aligned} \quad (7)$$

где q – вероятность, с которой в конце каждого периода ON начинается период OFF.

Данный подход показывает возможные пути определения параметров качества обслуживания в системе с наличием корреляции. При этом очевидно, что выражения (7) и (8) довольно сложны при использовании для численной оценки задержки в очереди и в системе в целом. Кроме того, данный подход не позволяет определить такой важный параметр, как изменение задержки, который очень важен при оценке качества функционирования современных инфокоммуникационных сетей.

Поэтому рассмотрим более оптимальный подход (с точки зрения удобства применения), основанный на показанном в [16] определении двумерных плотностей вероятностей с использованием копул. Как было показано выше, при определении времени ожидания пакета в сети и вариации задержки в сети фактически задача сводится к определению плотности вероятностей $f_T(x)$.

В [2–4] показано, что если учесть предположение Линдли [1], заключающееся в том, что $(i+1)$ -й пакет не будет ждать в очереди при выполнении условия, что интервал времени между приходом $(i+1)$ -го и i -го пакета больше, чем время задержки i -го пакета в узле сети, и если принять, что T_i – время задержки i -го пакета в узле сети, которое определяется в виде $T_i = S_i + R_i$, где R_i – время между поступлениями пакетов и S_i – время обслуживания пакета, то время ожидания пакета в очереди равняется

$$W_{i+1} = \begin{cases} 0, & \text{если } R_{i+1} \geq T_i, \\ W_i + S_i - R_{i+1}. & \end{cases}$$

Джиттер при этом определяется как

$$J_{i+1} = \begin{cases} |S_{i+1} - T_i|, & \text{если } R_{i+1} \geq T_i, \\ |S_{i+1} - R_{i+1}|. & \end{cases}$$

Определим функции плотности вероятностей рассматриваемых случайных величин: $f_R(y)$ – функция плотности вероятности для времени между поступлениями; $f_S(z)$ – функция плотности вероятности для времени обслуживания; $f_T(x)$ – функция плотности вероятности для времени ожидания в сети.

Законы распределения случайных величин S и R можно определить, анализируя реализацию случайного процесса, определяя законы распределения для времени обслуживания и для времени между поступлениями. Распределение случайной величины T – непростая задача, а для случая зависимых случайных величин требует особого подхода. Заметим, что

$$T = S + R. \quad (8)$$

Тогда плотность распределения случайной величины T [17]

$$f_T(x) = \int_0^{\infty} f_{S,R}(y, x-y) dy. \quad (9)$$

Или

$$f_T(x) = \int_0^{\infty} f_{S,R}(x-z, z) dz. \quad (10)$$

Определить двумерную плотность вероятностей можно с использованием копул на основе результатов теоремы Склара [18], методика применения такого подхода представлена в [16].

Согласно показанному в [17; 19] получим

$$f_{S,R}(y, z) = c(F_R(y), F_S(z)) f_R(y) f_S(z), \quad (11)$$

где

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$$

– плотность (производная) копулы, для нашего случая $u = F_R(y)$, $v = F_S(z)$ – функции распределения случайных величин.

Из работ, посвященных теории копул, следует, что для различных классов (типов распределений) следует использовать различные копулы. Анализ имеющихся подходов [16] показал, что наиболее удобной для использования в нашем случае является копула Фарли – Гумбеля – Моргенштерна (Farlie–Gumbel–Morgenstern copula), имея достаточно простую функциональную форму, она позволяет моделировать умеренные зависимости.

Копулу Фарли – Гумбеля – Моргенштерна [17–21] можно представить в виде:

Рисунок 2. Гистограмма измерений интервалов времени между поступлениями пакетов. Распределение Weibull, параметры: $a = 0,005$; $b = 1$

$$C(u, v) = uv \left[1 + \theta(1 - u^p)(1 - v^p) \right], \quad (12)$$

где θ – параметр копулы $\theta = \theta(R)$, R – коэффициент корреляции,

Учтем, что параметр θ должен удовлетворять следующим условиям:

$$-\left(\max(1, p^2)\right)^2 \leq \theta < \frac{1}{p},$$

где p – целое число ($p = 2, 3, \dots$).

Параметр θ определяется согласно выражению

$$\rho_s(u, v) = 3\theta \left(\frac{p}{p+2} \right)^2,$$

где $\rho_s(u, v)$ – коэффициент корреляции Спирмена.

При $p = 2$ из (11) получим

$$f_{s,R}(y, z) = \left[1 + \theta(1 - 2u)(1 - 2v) \right] \times f_R(y) f_s(z). \quad (13)$$

При известном значении $f_T(x)$ достаточно просто определить среднее время обработки (задержку) пакетов в сети и вариацию задержки пакетов в сети.

Анализ статистических характеристик видеотрафика

В качестве анализируемого трафика можно использовать видеотрафик, зарегистрированный на уровне пользователя, как один из наиболее характерных для обработки в современных инфокоммуникационных сетях. Для регистрации потока наиболее подходящим является анализатор пакетов Wireshark. Все пакеты собираются в режиме реального времени и предоставляются в удобном для чтения формате. Программа поддерживает очень мощную систему фильтрации.

Файл трассировки Wireshark (trace file) содержит необходимые для анализа работы сети характеристики трафика: интервалы времени между

Рисунок 3. Гистограмма измерений длин пакетов. Распределение Weibull (3P), параметры: $a = 94$; $b = 1$; $c = -4$

пакетами и длины пакетов (времена обработки пакетов).

Для аналитической оценки проанализируем статистические характеристики собранного трафика. Статистический критерий согласия Колмогорова – Смирнова – это один из наиболее распространенных и удобных инструментов для проверки гипотезы о предполагаемом законе распределения некоторой выборки случайных величин. Для данного исследования необходимо определить законы распределения случайных величин: интервалов времени между пакетами и длин пакетов (времен обработки пакетов).

С помощью Easyfit Professional проанализируем классы распределений процессов поступления интервалов времени между пакетами и длительности пакетов. При аппроксимации снятых параметров трафика в программе Easyfit Professional получаем законы распределения на рисунках 2 и 3.

Следовательно, интервалы времени между пакетами подчиняются закону Вейбулла с параметрами $a = 0,005$; $b = 1$ (см. рисунок 2), а длины пакетов – Вейбулла (Weibull, 3P) с параметрами $a = 94$; $b = 1$; $c = -4$ (см. рисунок 3).

Функция распределения и плотность распределения Вейбулла имеет вид

$$F(x) = 1 - \exp \left[- \left(\frac{x}{a} \right)^b \right], \quad (14)$$

$$f(x) = \left(\frac{b}{a} \right) \left(\frac{x}{a} \right)^{(b-1)} \exp \left[- \left(\frac{x}{a} \right)^b \right]. \quad (15)$$

Аналогично для трехпараметрического распределения Вейбулла (Weibull) функция распределения и плотность распределения имеют вид

$$F(x) = 1 - \exp \left[- \left(\frac{x - c}{a} \right)^b \right], \quad (16)$$

Рисунок 4. График зависимости коэффициентов взаимной корреляции временных интервалов времен обработки пакетов и времен между поступлениями пакетов (ON- и OFF-периодов)

$$f(x) = \left(\frac{b}{a}\right) \left(\frac{x-c}{a}\right)^{(b-1)} \exp\left[-\left(\frac{x-c}{a}\right)^b\right], \quad (17)$$

где a – параметр масштаба, характеризующий степень растянутости кривой распределения вдоль оси x и связанный со средним значением случайной величины; b – параметр формы; c – параметр сдвига, являющийся минимальным возможным значением случайной величины x .

Анализ законов распределений для интервалов времени между пакетами и времен обработки пакетов показал принадлежность к классу распределений с тяжелыми хвостами, это позволяет сделать вывод о том, что для моделирования системы обработки данного потока наиболее точной является модель G/G/1.

При определении параметров функционирования необходимо учитывать наличие зависимостей между временными интервалами, между пакетами и временами обработки пакетов.

Анализ взаимной корреляции интервалов времени между пакетами и времен обработки пакетов определяется согласно выражению для оценки корреляции двух случайных величин ξ и η по формуле:

$$R(\eta, \xi) = \frac{E(\xi\eta) - E(\xi)E(\eta)}{\sigma_\eta\sigma_\xi} = \frac{E\left(\xi - (E(\xi))\right)E\left(\eta - E(\eta)\right)}{\sigma_\eta\sigma_\xi}, \quad (19)$$

где $E(\xi)$, $E(\eta)$ – математическое ожидание случайной величины ξ и η соответственно; σ_η , σ_ξ – среднее квадратическое отклонение случайных величин ξ и η .

Из анализа графика зависимости коэффициентов взаимной корреляции (рисунок 4) можно сделать вывод о наличии достаточно сильной корреляции между интервалами временем об-

работки пакетов и интервалами времен между поступлениями пакетов. Это позволяет сделать предположение о зависимости данных случайных величин. Из чего следует обоснованность применения методики оценки задержки в системе с коррелированной очередью. Но для оценки параметра θ из выражения (12) необходимо определить коэффициент корреляции Спирмена.

Метод ранговой корреляции Спирмена позволяет определить силу и направление корреляционной связи между двумя признаками или двумя профилями признаков как

$$\rho_s = 1 - 6 \frac{\sum d_i}{n(n^2 - 1)}, \quad (20)$$

где для случайных величин ξ и η $d_i = \xi_i - \eta_i$, n – объем выборки. Для рассматриваемого потока коэффициент корреляции Спирмена $\rho_s = 0,46$.

Определение совместной плотности распределений интервалов времени между пакетами и длин пакетов видеотрафика

Зная законы распределения случайных величин R и S , учитывая (11) и (13), можно определить совместную плотность распределения данных случайных величин с учетом, что они являются зависимыми

$$f_{S,R}(y, z) = [1 + \theta(1 - 2u)(1 - 2v)] \times f_R(y) f_S(z). \quad (21)$$

Основываясь на том, что $f_R(y)$ – функция плотности вероятности для времени между поступлениями, $f_S(z)$ – функция плотности вероятности для времени обслуживания, а R и S – случайные величины, для которых необходимо определить совместную плотность распределения, в (21) будем использовать функцию распределения (14), (16) вместо u и v и плотности

распределения (15) и (17) вместо $f_R(y)$ и $f_S(z)$ из выражения (13). Тогда можем получить следующее выражение:

$$f_{S,R}(y,z) = \left[1 + \theta \left(1 - 2 \left[1 - \exp \left[- \left(\frac{y}{a_1} \right)^{b_1} \right] \right] \right) \right] \times \left[1 - 2 \left(1 - \exp \left[- \left(\frac{(x-y)-c}{a_2} \right)^{b_2} \right] \right) \right] \times \left(\frac{b_1}{a_1} \right) \left(\frac{y}{a_1} \right)^{(b_1-1)} \exp \left[- \left(\frac{y}{a_1} \right)^{b_1} \right] \times \left(\frac{b_2}{a_2} \right) \left(\frac{(x-y)-c}{a_2} \right)^{(b_2-1)} \exp \left[- \left(\frac{(x-y)-c}{a_2} \right)^{b_2} \right].$$

Задача, которую необходимо решить в рамках данного исследования, – определение суммы двух случайных величин: интервалов времени между пакетами и времен обработки пакетов, согласно (9)–(10) при условии $b_1 = b_2 = 1$:

$$\begin{aligned} f_T(x) &= \int_0^\infty f_{S,R}(y, x-y) dy = \\ &= \int_0^\infty \left[1 + \theta \left(1 - 2 \left[1 - \exp \left[- \left(\frac{y}{a_1} \right) \right] \right] \right) \right] \times \\ &\times \left[1 - 2 \left(1 - \exp \left[- \left(\frac{(x-y)-c}{a_2} \right) \right] \right) \right] \times \\ &\times \left(\frac{1}{a_1} \right) \exp \left[- \left(\frac{y}{a_1} \right) \right] \times \\ &\times \left(\frac{1}{a_2} \right) \exp \left[- \left(\frac{(x-y)-c}{a_2} \right) \right] dy = \\ &= (3\theta - 1) \frac{a_1 a_2}{(a_2 - a_1)^2} \exp \left(- \frac{x-c}{a_2} \right) + \\ &+ 2\theta \left(\frac{1}{(a_2 - a_1)} + \frac{1}{(a_2 - 2a_1)} \right) \exp \left(- \frac{2a_1(x-c)}{a_1 a_2} \right) + \\ &+ 2\theta \left(\frac{1}{(2a_2 - a_1)} \right) \exp \left(- \frac{a_1(x-c)}{a_1 a_2} \right). \end{aligned}$$

Полученное выражение для $f_T(x)$ позволяет оценить как среднее время ожидания пакета в сети или среднюю задержку пакета в сети, так и вариацию задержки пакета в сети. Среднее значение задержки представляет собой

$$E[T] = \int_0^\infty x f_T(x) dx =$$

$$\begin{aligned} &= (3\theta - 1) \frac{a_1 a_2^3}{(a_2 - a_1)^2} \exp \left(- \frac{c}{a_2} \right) + \\ &+ 2\theta \left(\frac{1}{(a_2 - a_1)} + \frac{1}{(a_2 - 2a_1)} \right) \left(\frac{a_1 a_2}{2a_1} \right)^2 \exp \left(\frac{2a_1 c}{a_1 a_2} \right) + \\ &+ 2\theta \left(\frac{1}{(2a_2 - a_1)} \right) \left(\frac{a_1 a_2}{a_1} \right)^2 \exp \left(\frac{a_1 c}{a_1 a_2} \right). \end{aligned}$$

Вариацию задержки можно определить согласно (1) как

$$\begin{aligned} \sigma(X) &= \sqrt{D(X)} = \\ &= \sqrt{\int_0^\infty x^2 f_T(x) dx - \left(\int_0^\infty x f_T(x) dx \right)^2}. \end{aligned}$$

Введем вспомогательную величину A согласно следующим обозначениям:

$$\begin{aligned} A &= \int_0^\infty x^2 f_T(x) dx = \\ &= (3\theta - 1) \frac{2a_1 a_2^4}{(a_2 - a_1)^2} \exp \left(- \frac{c}{a_2} \right) + \\ &+ 4\theta \left(\frac{1}{(a_2 - a_1)} + \frac{1}{(a_2 - 2a_1)} \right) \left(\frac{a_1 a_2}{2a_1} \right)^3 \exp \left(\frac{2a_1 c}{a_1 a_2} \right) + \\ &+ 4\theta \left(\frac{1}{(2a_2 - a_1)} \right) \left(\frac{a_1 a_2}{a_1} \right)^3 \exp \left(\frac{a_1 c}{a_1 a_2} \right). \end{aligned}$$

Тогда вариация задержки примет окончательный вид: $\sigma(X) = \sqrt{A - (E[T])^2}$.

Численные результаты

Учитывая полученные в результате статистического анализа законы распределения интервалов времени между пакетами и длин пакетов, принимая во внимание параметры данных распределений, получим:

– для интервалов времени между пакетами $a = 0,005$; $b = 1$;

– для длин пакетов: $a = 94$; $b = 1$; $c = -4$.

При $c_s = 0,46$ получим следующие выражения для оценки времени ожидания пакетов в сети $E[T] = 104$ мс;

Для вариации задержки получим $\sigma(X) = 84$ мс.

Выводы

В работе проанализированы методики оценки параметров качества обслуживания пакетов в сети при наличии корреляций в очереди между интервалами времени между пакетами и длинами пакетов. Определено, что при наличии корреля-

ции достаточно эффективным является подход, основанный на использовании копул.

При анализе корреляционных свойств интервалов времени между пакетами и длин пакетов, характеризующих обрабатываемый трафик, определено наличие сильных корреляций. Получены оценки средней задержки в сети и вариации задержки в сети при наличии корреляций между интервалами времени между пакетами и длинами пакетов.

Литература

1. Kleinrock L. Queueing Systems: Volume I – Theory. New York: Wiley, 1975. 432 p.
2. Dbira H., Girard A., Sanso B. Calculation of packet jitter for non-poisson traffic // *Annals of Telecommunications*. 2016. Vol. 71, no. 5–6. P. 223–237.
3. Kartashevskii V.G., Buranova, M.A. Analysis of packet jitter in multiservice network // 5th International Scientific-Practical Conference Problems of Infocommunications Science and Technology, PIC S and T 2018 – Conference Proceedings. 2018. P. 797–802. DOI: <https://doi.org/10.1109/INFOCOMMST.2018.8632085>
4. Карташевский В.Г., Буранова М.А. Моделирование джиттера пакетов при передаче по мультисервисной сети // *Информационные технологии и телекоммуникации*. 2019. Т. 17, № 1. С. 34–40.
5. Hwang G.U., Sohraby K. Performance of correlated queues: the impact of correlated service and inter-arrival times // *Performance Evaluation*. 2004. Vol. 55, no. 1–2. P. 129–145. DOI: [https://doi.org/10.1016/S0166-5316\(03\)00102-0](https://doi.org/10.1016/S0166-5316(03)00102-0)
6. Analysis of a correlated queue in a communication system / I. Cidon [et al.] // *IEEE Trans. Inform. Theory*. 1993. No. 39 (2). P. 456–465.
7. On queues with interarrival times proportional to service times / I. Cidon [et al.] // *INFOCOM'93*. 1993. P. 308–313.
8. Elwalid A.I., Mitra D., Stern T.E. Statistical multiplexing of Markov modulated sources: theory and computational algorithms // 13th International Teletraffic Congress, Copenhagen. 1991. P. 495–500.
9. Шелухин О.И., Осин А.В., Смольский С.М. Самоподобие и фракталы // *Телекоммуникационные приложения*. М.: Физматлит, 2008. 368 с.
10. Kartashevskiy I., Buranova M. Calculation of Packet Jitter for Correlated Traffic // *Lecture Notes in Computer Science*. 2019. Vol. 11660. P. 610–620. DOI: https://doi.org/10.1007/978-3-030-30859-9_53
11. Conolly B.W., Choo Q.H. The waiting time process for a generalized correlated queue with exponential demand and service // *SIAM Journal on Applied Mathematics*. 1979. No. 37 (2). P. 263–275.
12. Hadidi N. Queues with partial correlation // *SIAM Journal on Applied Mathematics*. 1981. No. 40 (3). P. 467–475.
13. Hadidi N. Further results on queues with partial correlation // *Operations Research*. 1985. № 33. P. 203–209.
14. Langaris C. Busy-period analysis of a correlated queue with exponential demand and service // *Journal of Applied Probability*. 1987. № 24. P. 476–485.
15. Langaris C. A correlated queue with innitely many servers // *Journal of Applied Probability*. 1986. № 23. P. 155–165.
16. Карташевский И.В. Использование копул в статистическом анализе телекоммуникационного трафика // *Инфокоммуникационные технологии*. 2016. Т. 14, № 4. С. 405–412.
17. Фантаццини Д. Моделирование многомерных распределений с использованием копула-функций // *Прикладная эконометрика*. 2011. № 3 (23). С. 98–132.
18. Farlie D.G.J. The performance of some correlation coefficients for a general bivariate distribution // *Biometrika*. 1960. No. 47. P. 307–323.
19. Gumbel E.J. Bivariate exponential distributions // *Journal of the American Statistical Association*. 1960. No. 55. P. 698–707.
20. Morgenstern D. Einfache Beispiele zweidimensionaler Verteilungen // *Mitteilungsblatt für Mathematische Statistik*. 1956. No. 8. P. 234–235.
21. Пеникас Г.И. Модели «копула» в приложении к задачам финансов // *Журнал новой экономической ассоциации*. 2010. № 7 (7). С. 24–44.

Получено 04.12.2020

Буранова Марина Анатольевна, к.т.н., доцент кафедры информационной безопасности (ИБ) Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). 443010, Российская Федерация, г. Самара, ул. Л. Толстого, 23. Тел. +7 846 332-49-15. E-mail: buranova-ma@psuti.ru

Резяпкина Магдалина Игоревна, магистр кафедры ИБ ПГУТИ. 443010, Российская Федерация, г. Самара, ул. Л. Толстого, 23. Тел. +7 917 169-09-06. E-mail: rez8pkina@mail.ru

ANALYSIS OF THE CORRELATED QUEUE IN THE G/G/1 SYSTEM

Buranova M.A., Rezyapkina M.I.

*Povolzhskiy State University of Telecommunications and Informatics, Samara, Russian Federation
E-mail: mburanova@yandex.ru, rez8pkina@mail.ru*

This article discusses correlated queues in which the packet service time is strongly correlated with the time between packets due to the finite throughput of the input and output channels. Approaches to assessing the parameters of the functioning of networks with correlated queues are considered. An analytical method for obtaining an estimate of the packet waiting time in the network based on the copula approach is presented. The correlation properties of real video traffic are investigated, the distribution laws for the time intervals between packets and the packet processing times are determined. The Weibull distribution gave the most accurate approximation. It has been determined that the time intervals between packets and the processing times of the packets have significant cross-correlation. Estimates are obtained for the average latency of a packet in the network and for the variation of the packet delay during transmission in the G/G/1 network.

Keywords: *quality of service (QoS), delay, delay variation, correlated queue*

DOI: 10.18469/ikt.2020.18.4.05

Buranova Marina Anatolievna, Povolzhskiy State University of Telecommunications and Informatics, 23, L. Tolstoy Street, Samara, 443010, Russian Federation; Associate Professor of Information Security Department, PhD in Technical Science. Tel. +7 846 339-11-67. E-mail: buranova-ma@psuti.ru

Rezyapkina Magdalina Igorevna, Povolzhskiy State University of Telecommunications and Informatics, 23, L. Tolstoy Street, Samara, 443010, Russian Federation; Master's Degree student of Information Security Department. Tel. +7 917 169-09-06. E-mail: rez8pkina@mail.ru

References

1. Kleinrock L. *Queueing Systems: Volume I – Theory*. New York: Wiley, 1975, 432 p.
2. Dbira H., Girard A., Sanso B. Calculation of packet jitter for non-poisson traffic. *Annals of Telecommunications*, 2016, vol. 71, no. 5–6, pp. 223–237.
3. Kartashevskii V.G., Buranova, M.A. Analysis of packet jitter in multiservice network. *5th International Scientific-Practical Conference Problems of Infocommunications Science and Technology, PIC S and T 2018 – Conference Proceedings*, 2018, pp. 797–802. DOI: <https://doi.org/10.1109/INFOCOMMST.2018.8632085>
4. Kartashevskij V.G., Buranova M.A. Modeling packet jitter during transmission over a multiservice network. *Informatsionnye tehnologii i telekommunikatsii*, 2019, vol. 17, no. 1, pp. 34–40. (In Russ.)
5. Hwang G.U., Sohraby K. Performance of correlated queues: the impact of correlated service and inter-arrival times. *Performance Evaluation*, 2004, vol. 55, no. 1–2, pp. 129–145. DOI: [https://doi.org/10.1016/S0166-5316\(03\)00102-0](https://doi.org/10.1016/S0166-5316(03)00102-0)
6. Cidon I. et al. Analysis of a correlated queue in a communication system. *IEEE Trans. Inform. Theory*, 1993, no. 39 (2), pp. 456–465.
7. Cidon I. et al. On queues with interarrival times proportional to service times. *INFOCOM'93*, 1993, pp. 308–313.
8. Elwalid A.I., Mitra D., Stern T.E. Statistical multiplexing of Markov modulated sources: theory and computational algorithms. *13th International Teletraffic Congress*, Copenhagen, 1991, pp. 495–500.

9. Sheluhin O.I., Osin A.V., Smol'skij S.M. Self-Similarity and Fractals. *Telecommunication Applications*. Moscow: Fizmatlit, 2008, 368 p. (In Russ.)
10. Kartashevskiy I., Buranova M. Calculation of Packet Jitter for Correlated Traffic. *Lecture Notes in Computer Science*, 2019, vol. 11660, pp. 610–620. DOI: https://doi.org/10.1007/978-3-030-30859-9_53
11. Conolly B.W., Choo Q.H. The waiting time process for a generalized correlated queue with exponential demand and service. *SIAM Journal on Applied Mathematics*, 1979, no. 37 (2), pp. 263–275.
12. Hadidi N. Queues with partial correlation. *SIAM Journal on Applied Mathematics*, 1981, no. 40 (3), pp. 467–475.
13. Hadidi N. Further results on queues with partial correlation. *Operations Research*, 1985, no. 33, pp. 203–209.
14. Langaris C. Busy-period analysis of a correlated queue with exponential demand and service. *Journal of Applied Probability*, 1987, no. 24, pp. 476–485.
15. Langaris C. A correlated queue with innitely many servers. *Journal of Applied Probability*, 1986, no. 23, pp. 155–165.
16. Kartashevskij I.V. Use of copulas in statistical analysis of telecommunication traffic. *Infokommunikacionnye tehnologii*, 2016, vol. 14, no. 4, pp. 405–412. (In Russ.)
17. Fantatstini D. Modeling multivariate distributions using copula functions. *Prikladnaja ekonometrika*, 2011, no. 3 (23), pp. 98–132. (In Russ.)
18. Farlie D.G.J. The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, 1960, no. 47, pp. 307–323.
19. Gumbel E.J. Bivariate exponential distributions. *Journal of the American Statistical Association*, 1960, no. 55, pp. 698–707.
20. Morgenstern D. Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsblatt für Mathematische Statistik*, 1956, no. 8, pp. 234–235.
21. Penikas G.I. Copula models as applied to finance problems. *Zhurnal novoj ekonomicheskoy assotsiatsii*, 2010, no. 7 (7), pp. 24–44. (In Russ.)

Reseived 04.12.2020

ТЕХНОЛОГИИ КОМПЬЮТЕРНЫХ СИСТЕМ И СЕТЕЙ

УДК 004.056

A BRIEF OVERVIEW OF DATA HIDING METHODS IN DIGITAL IMAGES

Verdiyev S.G., Naghiyeva A.F.

Azerbaijan Technological University, Ganja, Azerbaijan Republic

E-mail: info_tel@inbox.ru

The Internet and wireless communications have created new opportunities for the mass exchange of multimedia information. This has created new challenges in terms of the security and storage of the information transmitted through global and local networks. This article is devoted to the study of steganography methods of safety, in particular, those based on LSB concealment, in the spatial area of digital images. The classification of steganography methods has given based on the works of the last 5 years. The most popular and widespread image steganography methods have been investigated, describing their merits and challenges. The most commonly used terminology was identified and various methods of data concealment are presented. The visual and statistical steganalysis methods for computing the statistical difference between the input image and the stegoimage is also considered. The article may be useful to researchers in developing new, more effective steganography methods of information hiding.

Keywords: *information security, steganography, information hiding, image steganography, steganalysis*