Раздел II. Алгоритмы обработки информации

УДК 004.056.5 DOI 10.18522/2311-3103-2020-3-89-98

Н.В. Болдырихин, Ф.А. Алтунин, Д.А. Короченцев

ОСОБЕННОСТИ КЛАССИФИКАЦИИ ЗАШИФРОВАННОГО СЕТЕВОГО ТРАФИКА

В настоящее время растет интерес к задачам эффективного управления пакетными сетями: качеству обслуживания, обеспечению информационной безопасности, оптимизации использования программно-аппаратных ресурсов сети. Все эти задачи во многом опираются на анализ и классификацию сетевого трафика. Данный трафик неоднороден, как правило, имеет пульсирующий характер, трудно поддается прогнозированию, описывается математическим аппаратом случайных процессов. В разное время условия прохождения пакетов по одному и тому же пути могут значительно отличаться. Вместе с тем появляется значительное количество приложений, требовательных к задержкам и джиттеру. Задача администрирования в данном контексте состоит в правильной настройке узлов коммутации и маршрутизации. Классификация трафика позволяет идентифицировать пакеты различных приложений и служб и обеспечить их приоретизацию при передаче по сети. Например, трафик видеоконференций необходимо передавать в первую очередь, поскольку он очень чувствителен к задержкам и джиттеру, трафик данных можно передавать в последнюю очередь. Классификация трафика на сегодняшний день задача актуальная как с точки зрения администрирования сети, так и с точки зрения обеспечения её безопасности. Ввиду того, что большое количество приложений сейчас шифрует передаваемую информацию и просмотреть ее содержимое очень сложно, особый интерес представляет классификация трафика, которая позволяет по косвенным признакам определить аномалии в работе сети, признаки вторжения. В данной работе рассмотрены особенности решения задачи классификации зашифрованного трафика. Целью работы является исследование особенностей классификации зашифрованного трафика с использованием корреляционного анализа и алгоритма, основанного на разности интегральных площадей. Задачи исследования: – разработать алгоритм классификации трафика на основе корреляции и известными образцами; – разработать алгоритм, основанный на разности интегральных площадей под кривыми интенсивности трафика; – провести практическое исследование точности решения задачи классификации. В работе рассмотрена классификация трафика по трем группам: аудио, видео, данные. В результате выявлена достаточная точность корреляционного алгоритма при определении аудио и трафика данных. Для выявления видеотрафика лучше использовать алгоритм, основанный на разности интегральных площадей под кривыми интенсивности.

Информационная безопасность; классификация трафика; зашифрованный трафик; статистический метод; приложение; сеть связи.

N.V. Boldyrikhin, F.A. Altunin, D.A. Korochentsev

CLASSIFICATION FEATURES OF ENCRYPTED NETWORK TRAFFIC

Currently, there is growing interest in the tasks of efficient packet network management: quality of service, ensuring information security, optimization of the network hardware and software resources. All these tasks rely heavily on the analysis and classification of network traffic. This traffic is heterogeneous, as a rule, has a pulsating nature, difficult to predict and described by the mathematical apparatus of random processes. At different times, the conditions for passing

packets along the same path can vary significantly. At the same time, a significant number of applications are appearing requiring latency and jitter. The administration task in this context is to correctly configure the switching and routing nodes. Traffic classification allows you to identify packages of various applications and services and ensure their prioritization during transmission over the network. For example, video conferencing traffic needs to be transmitted first of all, since it is very sensitive to delays and jitter, data traffic can be transmitted last. The classification of traffic today is an urgent task both in terms of network administration and in terms of ensuring its security. Due to the fact that a large number of applications now encrypt the transmitted information and it is very difficult to view its contents, the traffic classification is of particular interest, which allows indirect signs to determine anomalies in the network, signs of intrusion. In this paper, we consider the features of solving the classification problem of encrypted traffic. The aim of the work is to study the classification features of encrypted traffic using correlation analysis and an algorithm based on the difference in integral areas. Research Objectives: – develop a traffic classification algorithm based on correlation and known patterns; – develop an algorithm based on the difference of the integral areas under the traffic intensity curves; – conduct a practical study of the accuracy of solving the classification problem. The work considers the classification of traffic into three groups: audio, video, data. As a result, a sufficient accuracy of the correlation algorithm in determining audio and data traffic was revealed. To identify video traffic, it is better to use an algorithm based on the difference of the integral areas under the intensity curves.

Information security; traffic classification; encrypted traffic; statistical method; application; communication network.

Введение. В настоящее время сети связи стали неотъемлемой частью жизни общества. Очевидно, что в таких сетях задача управления и контроля информационных потоков приобретает высокую значимость [1–6]. Неотъемлемой частью такого управления и контроля является классификация трафика [7–15].

Изначально классификация трафика использовалась администраторами для анализа информационных потоков с целью повышения управляемости сети, эффективности использования каналов связи, однако в настоящее время задача актуальна и в области информационной безопасности [16–18].

В общем случае задача классификации сетевого трафика может быть сформулирована как вероятностное отнесение трафика к определенному типу (например, трафик данных, видео, аудио и т.д.), либо определение соответствия трафика какому-то конкретному программному объекту, или классу объектов (например, программы рассылки спама) [7–15].

Существует несколько подходов к решению задачи классификации трафика: классификация, основанная на портах, классификация, основанная на полезной нагрузке, статистические методы анализа. Каждый из подходов обладает своими достоинствами и недостатками [7–15].

Эффективность классических методов в последнее время невелика. Значительная часть существующего вредоносного программного обеспечения (сетевые черви, троянские программы, программы рассылки спама и т.д.) генерирует сетевой трафик, который зачастую бывает зашифрованным [13, 18–20]. Так же многие легальные приложения генерируют огромное количество шифрованного трафика, содержимое которого невозможно проконтролировать при отсутствии ключей шифрования. Отчасти решить эту задачу можно с помощью различных статистических методов классификации трафика.

Классификация трафика на основе корреляции. Каждый тип трафика имеет свои индивидуальные особенности, поэтому может быть классифицирован с использованием корреляционного анализа. В данной работе рассматривается 3 типа трафика: видео (video), аудио (audio) и данные (data). На рис. 1–3 приведены примеры зависимости интенсивностей соответствующих видов трафика от времени.

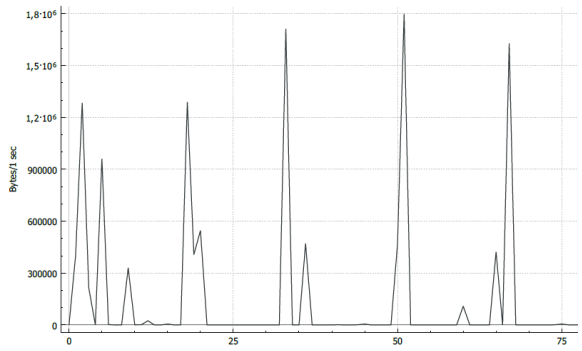

Рис. 1. Интенсивность трафика, создаваемого при просмотре видео на странице YouTube

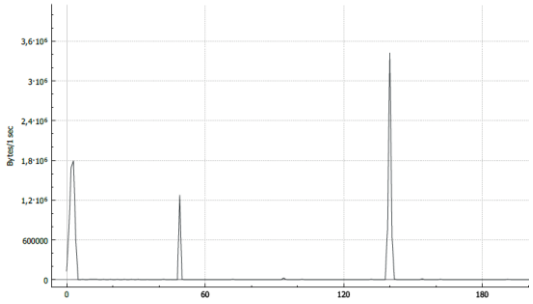

Рис. 2. Интенсивность трафика, создаваемого при прослушивании музыки на странице Яндекс Музыка

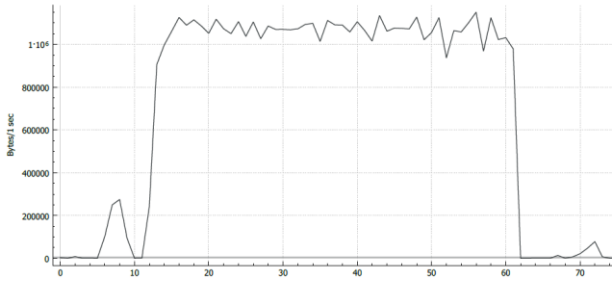

Рис. 3. Интенсивность трафика, создаваемого при скачивании файла

Суть алгоритма заключается в сборе некоторого количества образцов каждого из типов трафика. Эти образцы представляются в виде векторов интенсивности трафика и хранятся в соответствующей базе данных. Для классификации какого-либо неизвестного образца рассчитывается коэффициент корреляции относительно каждого известного образца из базы данных и вычисляется его среднее значение по каждой группе образцов.

Корреляция высчитывается по следующей формуле:

$$r_{xy} = \frac{\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{m}(x_i - \bar{x})^2 \sum_{i=1}^{m}(y_i - \bar{y})^2}}, \quad (1)$$

где x_i, y_i – координаты векторов неизвестного и известного образцов соответственно, а $\overline{x}$, $\overline{y}$ – их математические ожидания.

Значительное превышение среднего коэффициента корреляции по одной из групп в сравнении с остальными считается основанием для отнесения неизвестного образца к соответствующему типу трафика. Под термином «значительное превышение» может пониматься любое разумное пороговое значение, например, 50 %.

На рис. 4 приведен алгоритм классификации, разработанный в рамках данной работы.

В рамках натурного эксперимента, проведенного авторами, использовалось по 20 образцов каждого типа трафика, с которыми сравнивались «неизвестные» образцы трафика.

Для иллюстрации работы алгоритма проверялись образцы каждого типа. Заведомо известно: к какому типу трафика относится в классифицируемый образец.

Ниже представлена табл. 1, в которой описаны результаты расчетов корреляции для классифицируемого образца трафика, заведомо известного как data.

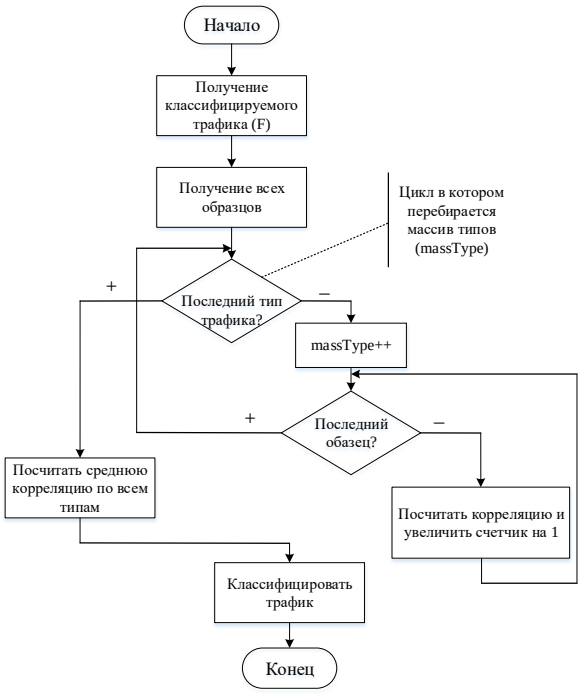

Рис. 4. Алгоритм классификации трафика на основе корреляции

Таблица 1

Средняя корреляция для классифицируемого образца типа data

video	audio	data
-0,126	-0,29	0,623

Из табл. 1 видно, что средний коэффициент корреляции классифицируемого образца типа data по группе образцов data превысил аналогичный показатель приблизительно в 4,94 и 2,14 раза соответственно для группы video и audio.

Далее в табл. 2 и 3 представлены аналогичные расчеты для классифицируемых образцов audio и video.

Таблица 2

Средняя корреляция для классифицируемого образца типа audio

video	audio	data
0,122	0,485	-0,337

Из табл. 2 видно, что средний коэффициент корреляции классифицируемого образца типа audio по группе образцов audio превысил аналогичный показатель приблизительно в 3,98 и 1,44 раза соответственно для группы video и data.

Таблица 3

Средняя корреляция для классифицируемого образца типа video

video	audio	data
0,053	0,082	-0,036

Из табл. 3 видно, что средний коэффициент корреляции классифицируемого образца типа video по группе образцов video превысил аналогичный показатель приблизительно в 1,55 раза для группы data, однако оказался ниже в 1,47 раза по сравнению с группой audio.

Классификация трафика на основе разности интегральных площадей под кривыми интенсивности. Вышеприведенный вычислительный эксперимент показал, что для решения проблемы классификации видеотрафика необходимо воспользоваться другим методом, например, приведенном в [21] (для решения другой задачи), который подразумевает сравнение площадей под кривыми интенсивности трафика на заданном временном интервале. Для проверки эффективности данного метода в качестве классифицируемых образцов использовалось три образца видеотрафика.

В табл. 4 представлены значения площадей под кривой для классифицируемых образцов, полученные путем интегрирования на заданном интервале времени

$$S_{Vi} = \int_{t_1}^{t_2} I_i(t)dt,\ i = 1,2,3, \tag{2}$$

где $I_i(t)$ – интенсивность трафика i-го образца, $t_1 = 0,\ t_2 = 799.$

Таблица 4

Площади классифицируемых образцов

S_{V1}, Bit	S_{V2}, Bit	S_{V3}, Bit
$3{,}635 * 10^4$	$2{,}645 * 10^4$	$3{,}883 * 10^4$

Для проведения классификации использовались средние значения площадей под кривой (табл. 5) для каждой группы образцов из базы данных, определяемых выражениями

$$\overline{S}_V = \frac{\sum_{j=1}^{20} \int_{t_1}^{t_2} I_{Vj}(t)dt}{20}, \tag{3}$$

$$\overline{S}_A = \frac{\sum_{j=1}^{20} \int_{t_1}^{t_2} I_{Aj}(t)dt}{20}, \tag{4}$$

$$\bar{S}_D = \frac{\sum_{j=1}^{20} \int_{t_1}^{t_2} I_{Dj}(t)dt}{20}, \quad (5)$$

где $\bar{S}_V, \bar{S}_A, \bar{S}_D$ – средние площади для video, audio и data типов трафика соответственно; $I_{Vj}(t), I_{Aj}(t), I_{Dj}(t)$ – интенсивности трафика для каждого образца, хранимого в базе данных для групп video, audio и data соответственно. В табл. 5 приведены результаты расчетов средних площадей.

Таблица 5

Средние площади для каждого типа трафика

$\bar{S}_V$, Bit	$\bar{S}_A$, Bit	$\bar{S}_D$, Bit
$3{,}174 * 10^4$	$1{,}307 * 10^4$	$1{,}555 * 10^5$

Суть классификации трафика состоит в определении разности площади классифицируемого образца и средней площади каждого типа трафика. Для одной из групп образцов эта разность будет иметь наименьшее абсолютное значение. Именно к этой группе следует отнести классифицируемый образец.

В табл. 6 представлены разности между площадями образцов видеотрафика и средними площадями по группам.

Таблица 6

Разности площадей (абсолютные значения)

	$\bar{S}_V$, Bit	$\bar{S}_A$, Bit	$\bar{S}_D$, Bit
S_{V1}, Bit	$4{,}605*10^3$	$2{,}327*10^4$	$1{,}192*10^5$
S_{V2}, Bit	$5{,}292*10^3$	$1{,}338*10^4$	$1{,}129*10^5$
S_{V2}, Bit	$7{,}094*10^3$	$2{,}576*10^4$	$1{,}167*10^5$

Данные расчеты доказывают эффективность метода площадей для классификации видеотрафика, поскольку разности площадей по группе video в разы меньше разностей по группам audio и data.

Заключение. В результате проведенных исследований выяснилось, что алгоритм классификации трафика на основе корреляции позволяет с достаточным качеством отличить трафик данных от трафика видео и аудио. Для данного типа трафика коэффициент корреляции по своей группе значительно превысил аналогичный показатель для других групп: в 4,94 и 2,14 раза соответственно для группы видео и аудио. Учитывая тот факт, что эти коэффициенты корреляции являются усредненными по результатам двадцати опытов, можно заключить, что данный алгоритм может успешно применяться на практике.

Аналогичным образом дела обстоят и с трафиком аудио, который можно легко отличить от видеотрафика и трафика данных. По результатам эксперимента средний коэффициент корреляции по своей группе превысил аналогичный показатель приблизительно в 3,98 и 1,44 раза соответственно для групп видео и данные.

Вместе с тем, следует отметить, что классифицировать видеотрафик при помощи данного алгоритма не получилось. Значение среднего коэффициента корреляции видеотрафика по своей группе образцов превысило аналогичный показатель приблизительно в 1,55 раза для трафика данных, однако оказалось ниже в 1,47 раза по сравнению со средней корреляцией для группы образцов аудиотрафика.

Для выявления видеотрафика следует использовать другой подход, например, использующий разности интегральных площадей под кривыми интенсивности, приведенный в данной статье. Данный метод показал высокую эффективность в рамках исследования, проводимого для трех различных образцов видеотрафика, которые сравнивали с существующей базой образцов. Данные эксперимента показали превышение разности интегральных площадей под кривыми интенсивности в среднем в 21,22 раза для трафика данных и в 3,73 раза для аудиотрафика по сравнению со своей группой образцов. Это свидетельствует о хорошей результативности данного подхода.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Мироненко А.Д., Сидоренко В.В., Ни Ю.А.* Аналитический обзор методов мониторинга сети // Актуальные вопросы науки и техники: Сб. статей. – Пенза: 2020. – С. 58-61.
2. *Болдырихин Н.В., Короченцев Д.А., Манакова А.Н., Качнов С.А.* Особенности использования динамических моделей при описании сетевого оборудования в задачах мониторинга // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2019. – С. 305-311.
3. *Болдырихин Н.В., Рыбалко И.П., Сосновский И.А., Гирин И.С.* Применение динамических моделей при описании сетевого оборудования в задачах мониторинга // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2017. – С. 70-73.
4. *Кносаль В.М., Алтунин Ф.А., Давыдов Р.В., Болдырихин Н.В.* Анализ математических моделей сетей связи, используемых в задачах мониторинга // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2017. – С. 245-248.
5. *Буковшин В.А., Болдырихин Н.В.* Сравнительное исследование технологий анализа интенсивности сетевого трафика // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2019. – С. 104-107.
6. *Татарникова Т.М.* Статистические методы исследования сетевого трафика // Информационно-управляющие системы. – 2018. – № 5 (96). – С. 35-43.
7. *Алтунин Ф.А., Кносаль В.М., Давыдов Р.В., Болдырихин Н.В.* Анализ методов классификации трафика // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2017. – С. 23-27.
8. *Callado A., Kamienski C., Szabo G., Gero B., Kelner J., Fernandes S., Sadok D.* A Survey on Internet Traffic Identification // Communications Surveys & Tutorials, IEEE. – 3rd Quarter 2009. – Vol. 11, Issue 3. – P. 37-52.
9. *Risso F., Baldi M., Morandi O., Baldini A., Monclus P.* Lightweight, payload-based traffic classification: An experimental evaluation // in Proc. IEEE ICC. – 2008. – P. 5869-5875.
10. *Усовик С.В., Воронин А.В.* Алгоритм классификации трафика телекоммуникационной сети // Информационные системы и технологии. – 2011. – № 1 (63). – С. 107-110.
11. *Кузьмин В.В.* Классификация и идентификация трафика в мультисервисной сети оператора связи // Современные проблемы науки и образования. – 2014. – № 5. – С. 231.
12. *Маркович Н.М.* Анализ видео трафика: классификация и оценивание потерь передачи // Идентификация систем и задачи управления (SICPRO ‘09): Тр. VIII международной конференции. – М., 2009. – С. 1035-1058.
13. *Болдырихин Н.В., Алтунин Ф.А.* Анализ методов классификации трафика при решении задач обеспечения информационной безопасности сетей связи // Актуальные проблемы науки и техники: Матер. национальной научно-практической конференции. – Ростов-на-Дону, 2019. – С. 356-357.
14. *Василишин Н.С., Ушаков И.А., Котенко И.В.* Исследование алгоритмов анализа сетевого трафика с использованием технологий больших данных для обнаружения компьютерных атак // Аллея науки. – 2018. – Т. 3, № 6 (22). – С. 1012-1021.
15. *Платонова А.В., Белоусов А.А.* Исследование методов статистического анализа для поиска аномалий в сетевом трафике // XIV Королёвские чтения: Сб. трудов международной молодежной научной конференции. – Самара, 2017. – С. 85-86.

16. *Буковшин В.А., Болдырихин Н.В.* Современные проблемы информационной безопасности // Современные материалы, техника и технология: Сб. статей. – Курск, 2018. – С. 47-52.
17. *Буковшин В.А., Болдырихин Н.В.* Кибербезопасность как неотъемлемая часть информационного мира // Современные материалы, техника и технология: Сб. статей. – Курск, 2018. – С. 52-55.
18. *Татаринов А.А., Болдырихин Н.В.* Анализ методов обнаружения вредоносного программного обеспечения на основе поведенческих признаков // Национальная безопасность России: актуальные аспекты: Сб. статей. – СПб., 2020. – С. 18-22.
19. *Тюрин К.А., Болдырихин Н.В.* Алгоритм вероятностной идентификации пользователей сети // Молодой исследователь Дона. – 2016. – № 2 (2). – С. 81-86.
20. *Харитонова Н.В., Негрышева Я.В., Болдырихин Н.В.* Использование технологии VPN в корпоративных сетях // Тр. Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону, 2013. – С. 250-252.
21. *Асриянц С.В., Селёва А.В., Болдырихин Н.В.* Идентификация объекта наблюдения на основе истории его местоположений // Advances in Science and Technology: Сб. статей IX международной научно-практической конференции. – М., 2017. – С. 64-67.

REFERNCES

1. *Mironenko A.D., Sidorenko V.V., Ni Yu.A.* Analiticheskiy obzor metodov monitoringa seti [Analytical review of network monitoring methods], *Aktual'nye voprosy nauki i tekhniki: Sb. statey* [Current issues of science and technology: Collection of articles]. Penza: 2020, pp. 58-61.
2. *Boldyrikhin N.V., Korochentsev D.A., Manakova A.N., Kachnov C.A.* Osobennosti ispol'zovaniya dinamicheskikh modeley pri opisanii setevogo oborudovaniya v zadachakh monitoringa [Features of using dynamic models when describing network equipment in monitoring tasks], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Don, 2019, pp. 305-311.
3. *Boldyrikhin N.V., Rybalko I.P., Sosnovskiy I.A., Girin I.S.* Primenenie dinamicheskikh modeley pri opisanii setevogo oborudovaniya v zadachakh monitoringa [Application of dynamic models in the description of network equipment in monitoring tasks], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Don, 2017, pp. 70-73.
4. *Knosal' V.M., Altunin F.A., Davydov R.V., Boldyrikhin N.V.* Analiz matematicheskikh modeley setey svyazi, ispol'zuemykh v zadachakh monitoringa [Analysis of mathematical models of communication networks used in monitoring tasks], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Don, 2017, pp. 245-248.
5. *Bukovshin V.A., Boldyrikhin N.V.* Sravnitel'noe issledovanie tekhnologiy analiza intensivnosti setevogo trafika [Comparative research of technologies for analyzing the intensity of network traffic], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Don, 2019, pp. 104-107.
6. *Tatarnikova T.M.* Statisticheskie metody issledovaniya setevogo trafika [Statistical methods of network traffic research], *Informatsionno-upravlyayushchie sistemy* [Information and control systems], 2018, No. 5 (96), pp. 35-43.
7. *Altunin F.A., Knosal' V.M., Davydov R.V., Boldyrikhin N.V.* Analiz metodov klassifikatsii trafika [Analysis methods for traffic classification], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Donu, 2017, pp. 23-27.
8. *Callado A., Kamienski C., Szabo G., Gero B., Kelner J., Fernandes S., Sadok D.* A Survey on Internet Traffic Identification, *Communications Surveys & Tutorials, IEEE,* 3rd Quarter 2009, Vol. 11, Issue 3, pp. 37-52.
9. *Risso F., Baldi M., Morandi O., Baldini A., Monclus P.* Lightweight, payload-based traffic classification: An experimental evaluation, *in Proc. IEEE ICC,* 2008, pp. 5869-5875.

10. *Usovik S.V., Voronin A.V.* Algoritm klassifikatsii trafika telekommunikatsionnoy seti [The algorithm for classifying the traffic of a telecommunication network], *Informatsionnye sistemy i tekhnologii* [Information systems and technologies], 2011, No. 1 (63), pp. 107-110.
11. *Kuz'min V.V.* Klassifikatsiya i identifikatsiya trafika v mul'tiservisnoy seti operatora svyazi [Classification and identification of traffic in the operator's multiservice network], *Sovremennye problemy nauki i obrazovaniya* [Modern problems of science and education], 2014, No. 5, pp. 231.
12. *Markovich N.M.* Analiz video trafika: klassifikatsiya i otsenivanie poter' peredachi [Video traffic analysis: classification and estimation of transmission losses], *Identifikatsiya sistem i zadachi upravleniya (SICPRO '09): Tr. VIII mezhdunarodnoy konferentsii* [Identification of systems and management tasks (SICPRO '09): Proceedings of the VIII international conference]. Moscow, 2009, pp. 1035-1058.
13. *Boldyrikhin N.V., Altunin F.A.* Analiz metodov klassifikatsii trafika pri reshenii zadach obespecheniya informatsionnoy bezopasnosti setey svyazi [Analysis of traffic classification methods for solving problems of information security of communication networks], *Aktual'nye problemy nauki i tekhniki: Mater. natsional'noy nauchno-prakticheskoy konferentsii* [Actual problems of science and technology: Materials of the national scientific and practical conference]. Rostov-on-Don, 2019, pp. 356-357.
14. *Vasilishin N.S., Ushakov I.A., Kotenko I.V.* Issledovanie algoritmov analiza setevogo trafika s ispol'zovaniem tekhnologiy bol'shikh dannykh dlya obnaruzheniya komp'yuternykh atak [Research of algorithms for analyzing network traffic using big data technologies for detecting computer attacks], *Alleya nauki* [Alley of science], 2018, Vol. 3, No. 6 (22), pp. 1012-1021.
15. *Platonova A.V., Belousov A.A.* Issledovanie metodov statisticheskogo analiza dlya poiska anomaliy v setevom trafike [Research of statistical analysis methods for searching for anomalies in network traffic], *XIV Korolevskie chteniya: Cb. trudov mezhdunarodnoy molodezhnoy nauchnoy konferentsii* [XIV Royal readings: proceedings of the international youth scientific conference]. Samara, 2017, pp. 85-86.
16. *Bukovshin V.A., Boldyrikhin N.V.* Sovremennye problemy informatsionnoy bezopasnosti [Modern problems of information security], *Sovremennye materialy, tekhnika i tekhnologiya: Sb. statey* [Modern materials, technique and technology: Collection of articles]. Kursk, 2018, pp. 47-52.
17. *Bukovshin V.A., Boldyrikhin N.V.* Kiberbezopasnost' kak neot"emlemaya chast' informatsionnogo mira [Cybersecurity as an integral part of the information world], *Sovremennye materialy, tekhnika i tekhnologiya: Sb. statey* [Modern materials, technique and technology: Collection of articles]. Kursk, 2018, pp. 52-55.
18. *Tatarinov A.A., Boldyrikhin N.V.* Analiz metodov obnaruzheniya vredonosnogo programmnogo obespecheniya na osnove povedencheskikh priznakov [Analysis of methods for detecting malicious software based on behavioral characteristics], *Natsional'naya bezopasnost' Rossii: aktual'nye aspekty: Sb. statey* [National security of Russia: current aspects: Collection of articles]. Saint Petersburg, 2020, pp. 18-22.
19. *Tyurin K.A., Boldyrikhin N.V.* Algoritm veroyatnostnoy identifikatsii pol'zovateley seti [Algorithm for probabilistic identification of network users], *Molodoy issledovatel' Dona* [Young don Explorer], 2016, No. 2 (2), pp. 81-86.
20. *Kharitonova N.V., Negrysheva Ya.V., Boldyrikhin N.V.* Ispol'zovanie tekhnologii VPN v korporativnykh setyakh [Using VPN technology in corporate networks], *Tr. Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki* [Proceedings of the North Caucasus branch of the Moscow technical University of communications and Informatics]. Rostov-on-Don, 2013, pp. 250-252.
21. *Asriyants S.V., Seleva A.V., Boldyrikhin N.V.* Identifikatsiya ob"ekta nablyudeniya na osnove istorii ego mestopolozheniy [Identification of an object of observation based on the history of its locations], *Advances in Science and Technology: Sb. statey IX mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Advances in Science and Technology: Collection of articles of the IX international scientific and practical conference]. Moscow, 2017, pp. 64-67.

Статью рекомендовал к опубликованию д.т.н. А.В. Елисеев.

Болдырихин Николай Вячеславович – Донской государственный технический университет; e-mail: boldyrikhin@mail.ru; 344065, г. Ростов-на-Дону, пер. Днепровский, 116К, кв. 111; тел.: +79043442295; кафедра кибербезопасности информационных систем; к.т.н.; доцент.

Короченцев Денис Александрович – e-mail: mytelefon@mail.ru; 344038, г. Ростов-на-Дону, пр. Михаила Нагибина, 29, кв. 24; тел.: +79034895173; кафедра кибербезопасности информационных систем; зав. кафедрой; к.т.н.

Алтунин Федор Александрович – ООО «Яндекс.Маркет Лаб»; email: altuninf@gmail.com; Москва, Ленинский пр-т 129, корп. 3, кв. 119; тел.: +79889496866; инженер по тестированию.

Boldyrikhin Nikolay Vyacheslavovich – Don State Technical University; e-mail: boldyrikhin@mail.ru; 116K, per. Dneprovsky, apt. 111, Rostov-on-Don, 344065, Russia; phone: +79043442295; the department of cybersecurity of information systems; cand. of eng. sc.; associate professor.

Korochentsev Denis Aleksandrovich – e-mail: mytelefon@mail.ru; 29, Mikhail Nagibin Ave., apt. 24, Rostov-on-Don, 344038, Russia; phone: +79034895173; the department of cybersecurity of information systems; head of the department; cand. of eng. sc.

Altunin Fedor Aleksandrovich – Yandex.Market Lab LLC; e-mail: altuninf@gmail.com; 129, Leninsky Prospect building. 3, apt. 119, Moscow, Russia; phone: +79889496866; testing engineer.

УДК 004.432.4 DOI 10.18522/2311-3103-2020-3-98-111

И.И. Левин, А.И. Дордопуло, И.В. Писаренко, Д.В. Михайлов

ПРЕДСТАВЛЕНИЕ ГРАФОВ С АССОЦИАТИВНЫМИ ОПЕРАЦИЯМИ НА ЯЗЫКЕ ПРОГРАММИРОВАНИЯ SET@L

Как правило, информационный граф с ассоциативными операциями реализуется в виде последовательной («голова/хвост») или параллельной («разбиение пополам») топологии, причем обе структуры содержат одинаковое число операционных вершин. Редукционные преобразования графов с представленными топологиями при недостатке вычислительного ресурса не обеспечивают создание эффективной ресурсонезависимой программы: вариант «разбиение пополам» характеризуется нерегулярной межитерационной коммутацией, а структура «голова/хвост» – увеличенной скважностью данных при редукции. В данной статье предлагается преобразовать топологию графа с ассоциативными операциями в один из комбинированных вариантов с последовательными и параллельными фрагментами вычислений, синтезированный в соответствии с заданным вычислительным ресурсом. Это позволяет повысить удельную производительность вычислений при редукции. Модифицированная топология включает изоморфные подграфы с топологией «разбиение пополам», содержащие максимальное число аппаратно реализуемых операционных вершин, а обработка промежуточных данных осуществляется по принципу «голова/хвост». Вычислительная структура для рассмотренной топологии имеет минимальную латентность и состоит из одного базового подграфа и одной вершины, в которую редуцируется блок обработки промежуточных данных с топологией «голова/хвост». Разработан алгоритм, позволяющий в зависимости от доступного аппаратного ресурса перейти от базового последовательного варианта реализации к различным комбинированным топологиям вплоть до предельного случая топологии «разбиение пополам». Поскольку традиционные методы параллельного программирования могут описать множество топологий только в виде набора отдельных подпрограмм, для создания ресурсонезависимого описания графов с ассоциативными операциями предлагается использовать язык архитектурно-независимого программирования Set@l. Принципы построения топологий «голова/хвост» и «разбиение пополам» описаны в виде признаков