ISSN 1992-6502 (Print) Вестник УГАТУ ISSN 2225-2789 (Online)
2024. Т. 28, № 1 (103). С. 92–98 http://journal.ugatu.su

УДК 004.6

Алгоритм анализа сетевого трафика в многокомпонентных структурно сложных информационных системах

Д.С. Алексеева[1], Н.А. Кононов[2], В.В. Антонов[3], Р.Р. Еникеев[4]

[1]ads.stat@mail.ru, [2]knnv.nkt@gmail.com, [3]antonov@bashkortostan.ru, [4]enikeevrr@mail.ru,

ФГБОУ ВО «Уфимский университет науки и технологий», г. Уфа, Россия

Поступила в редакцию 21.02.2024

Аннотация. В статье рассмотрена на возможность применения микросервисной архитектуры и микросервисного подхода для реализации компонента сложной системы, направленного на анализ сетевых потоков информации в целях идентификации угроз и повышения информационной безопасности. Рассмотрены базовые понятия информационной предметной области в контексте решаемой задачи. Сформированная функциональная модель проектируемого микросервиса позволяет рассмотреть взаимосвязи элементов процесса. Разработанная информационная модель на уровне абстракции отражает однородные связи общих характеристик, что дает возможность представить процесс в виде динамических моделей. Динамическая модель в нотации BPMN будет служить основой для графоаналитической модели проектируемого микросервиса. Используя графоаналитическую модель и пи-исчисления возможно сформировать математическую модель, открывающую возможность создания формального алгоритма и реализации с помощью языков высокого уровня. Также динамическую модель можно использовать для формирования кода с помощью *Java Workflow Tooling*. Сформированную математическую модель с применением пи-исчислений возможно использовать для определения метода обучения нейронной сети и построения датасета. Так как ядром сервиса является нейронная сеть, то открывается возможность динамического совершенствования средства анализа трафика.

Ключевые слова: сетевой трафик; анализ сетевого трафика; информационная безопасность; пи-исчисления; категориальный подход; абстрактное информационное моделирование; BPMN; алгоритм защиты информации.

ВВЕДЕНИЕ

Одним из важных трендов в разработке многокомпонентных структурно сложных систем является сервисная и микросервисная архитектура. Она позволяет разбить приложение на множество небольших сервисов, каждый из которых выполняет свою функцию. Такой подход упрощает разработку, тестирование и масштабирование приложения, а также повышает его отказоустойчивость и гибкость. Преимуществом микросервисной архитектуры является возможность использования различных технологий и языков программирования для каждого сервиса, что упрощает разработку и обеспечивает более эффективную работу системы в целом. Кроме того, такой подход позволяет легко добавлять новые сервисы и функциональность без необходимости переписывать всю систему.

Рассматривая микросервис формально, его можно представить, как несколько множеств. Как правило, каждый микросервис обеспечивает одну бизнес-функцию, обозначим ее как $BF_i \in <BF_1 \dots BF_n>$. Совокупность правил, принципов, зависимостей поведения объектов

предметной области реализуется через набор бизнес-правил, которые можно обозначить как $< BR_1 \dots BR_k >$. Как уже было отмечено ранее, бизнес-правила каждого микросервиса могут быть реализованы на различных языка программирования и на различных технологических платформах, обозначим их через множество $< T\_PL_1 , \dots , T\_PL_t >$. Следовательно, микросервис n, состоящий из m-элементов данных и k-элементов бизнес-правил, можно представить как $S_n = \{BF_n , < D_{n1} \dots D_{nm} >,\ T\_DB_n, < BR_{n1} \dots BR_{nk} >, T\_PL_t \}$. Информационная система некоторой организации W (IS_w) может быть описана как набор сервисов и связей ($C_1 \dots C_n$) между ними: $IS_n = \{< S_{w1} \dots S_{wo} >, < C_{w1} \dots C_{wp} >\}$, где o – количество сервисов; а p – количество связей между ними.

Технологически известно, что микросервис вполне может решать задачи, связанные с потоком информационных данных (сетевым трафиком). Передача данных в Сети происходит частями, называемыми пакетами. Трафик измеряется в стандартных единицах измерения цифровой информации (байтах, килобайтах). Сетевой трафик (СТ) делится на входящий и исходящий. Входящий трафик – тот трафик, который принимается устройством из Сети Интернет. Исходящий трафик – тот трафик, что передает устройство в Сеть. Анализ сетевого трафика основан на проверке данных, проходящих через узлы Сети.

ЦЕЛЬ ИССЛЕДОВАНИЯ

Постоянный мониторинг СТ используется для обнаружения аномалий, уязвимостей, диагностики работоспособности Сети, отслеживания подозрительной активности. Проводится, как правило, двумя способами: в режиме реального времени и в режиме просмотра отдельных временных промежутков. При первом подходе сетевой трафик анализируется на лету. При втором – трафик (его часть) записывается на диск и далее анализируется. Однако при этом подходе существуют определенные подводные камни: терабайтное дисковое пространство и функция, позволяющая отматывать данные в произвольный момент. Это необходимо в том случае, когда проблема безопасности повторяется нерегулярно и ее невозможно категорировать, а также когда обнаружено нарушение информационной безопасности. Работу микросервиса, выполняющего статистический анализ загрузки каналов, в общем виде можно представить следующим образом (рис. 1).

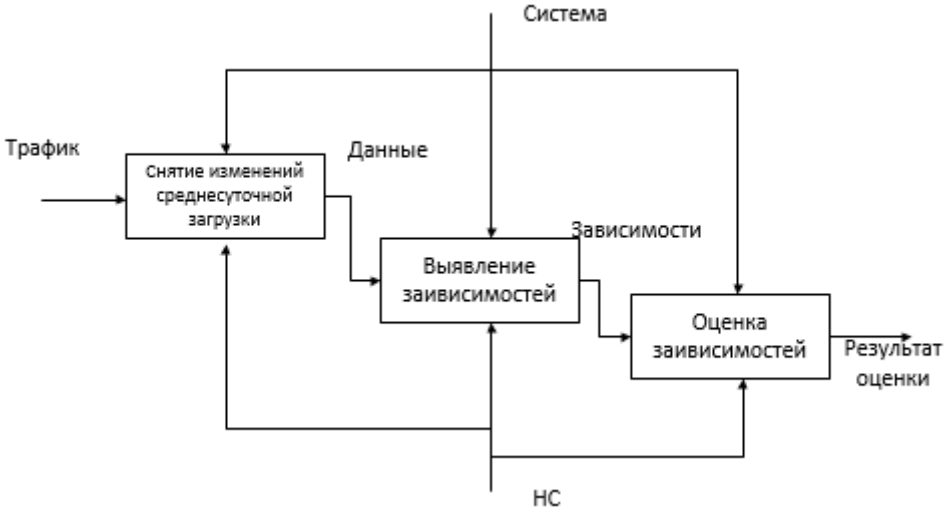

Рис. 1. Функциональная модель проектируемого микросервиса.

Микросервис выполняет мониторинг трафика (снятие его изменений), выявляет зависимости внутри дампа наблюдений и в конце выполняет оценку зависимостей. Снятие сведений об изменении трафика может выполняться с помощью протокола *Netflow*, разработанного *Cisco System* и фактически являющегося промышленным стандартом. Работа протокола представлена на рис. 2.

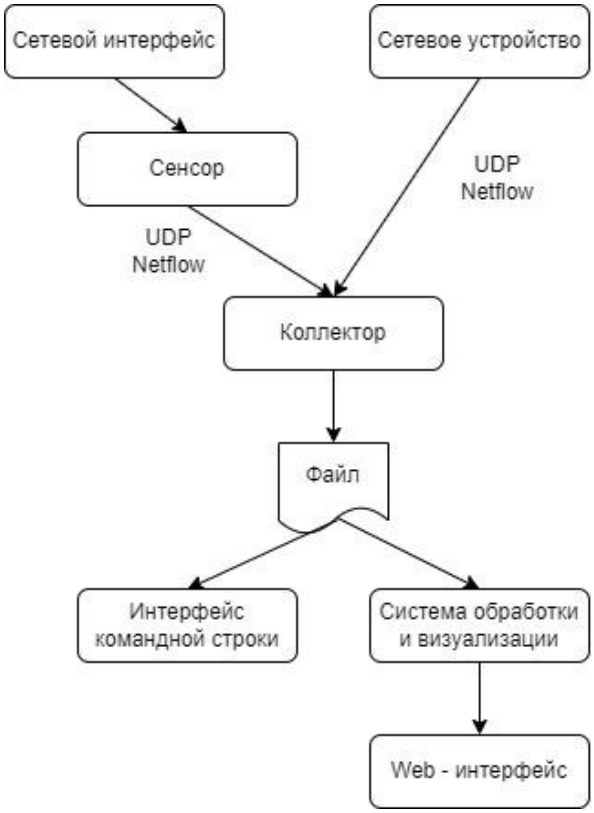

Рис. 2. Работа протокола.

После того как получен набор данных о сетевом трафике, происходит выявление зависимостей, что является прямой задачей регрессионного анализа. Далее при оценке зависимостей пользователь получает информацию о сетевых аномалиях.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На рис. 3 представлена информационная модель проектируемого микросервиса на концептуальном уровне абстракции.

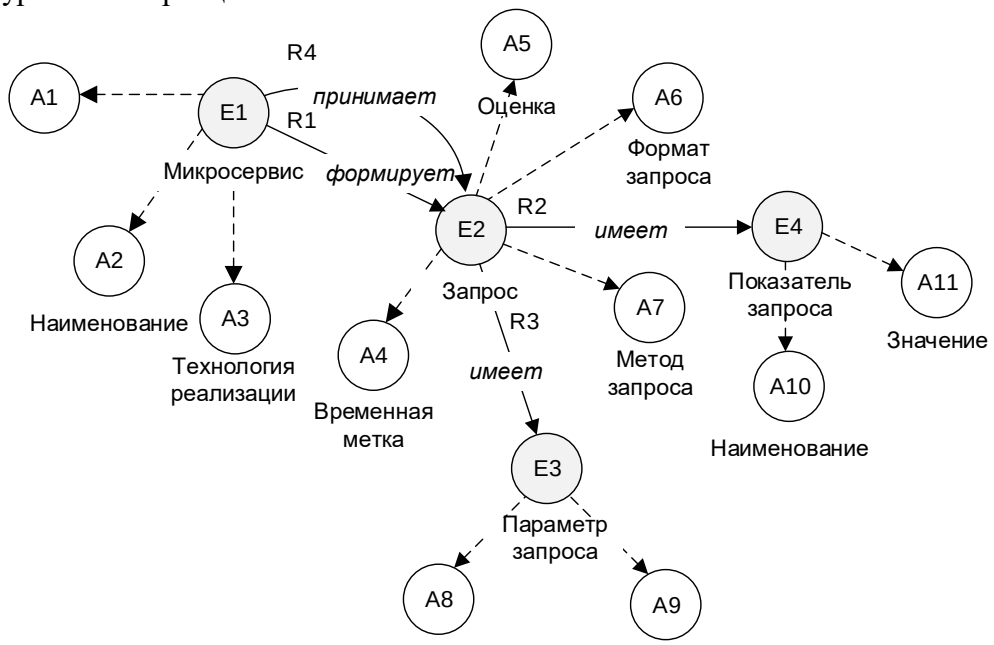

Рис. 3. Информационная модель проектируемого микросервиса анализа сетевого трафика на концептуальном уровне абстракции.

Модель разработана с применением категориального подхода. Для графического представления формальной модели применены элементы теории графов. В качестве вершин рассматриваются сущности предметной области ($E = \langle E_1, \ldots, E_n \rangle$, где n – количество сущностей) и ее атрибуты ($A = \langle A_1, \ldots, A_m \rangle$, где m – количество атрибутов), а в качестве вершин семантические связи ($R = \langle R_1, \ldots, R_t \rangle$, где t – количество связей между ними.

В качестве математической основы принимается формула вида пи-исчислений. Для того чтобы моделировать процесс, с математической точки зрения, на основе пи-исчисления, требуется провести графоаналитическое моделирование процесса. Исходный процесс, описанный на рис. 4 в форме динамической модели в нотации *BPMN*, дополнен, и разработана графоаналитическая модель, представленная на рис. 5.

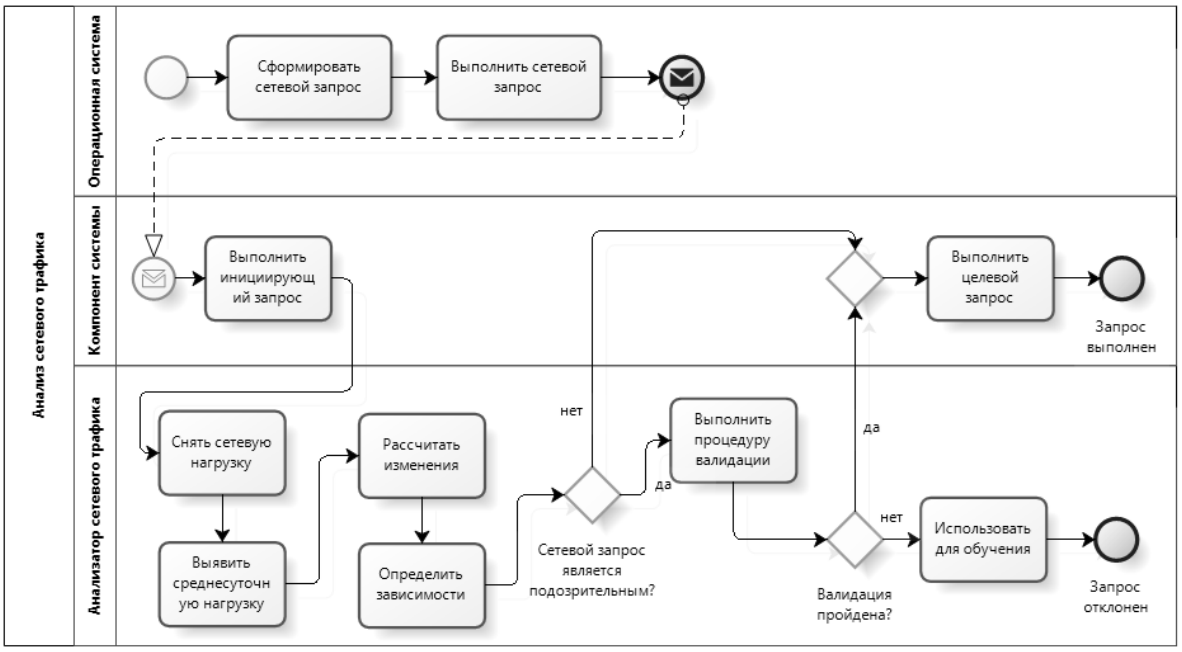

Рис. 4. Описание предлагаемого процесса в форме динамической модели в нотации BPMN.

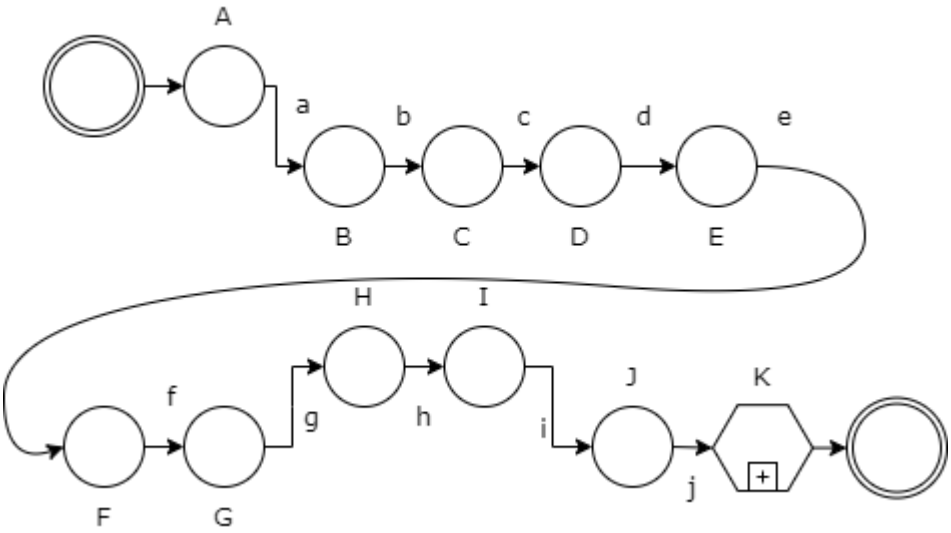

Рис. 5. Предлагаемый процесс в виде графоаналитической модели.

Принято, что переходы обозначаются прописными буквами, элементы – строчными. В табл. 1 приведено описание конструкций, отраженных на рис. 5.

Таблица 1

Описание конструкций

Конструкция	Наименование пути	Обозначение в терминах пи-исчисления	
		Краткое	Расширенное
1	2	3	4
A→B	a	A|B	$A = \tau_a.\bar{a} < x > .0$ $B = \mathrm{a(x)}.\tau_B.B'$
B→C	b	B|C	$B = \tau_b.\bar{b} < x > .0$ $C = b(\mathrm{x}).\tau_C.C'$
C→D	c	C|D	$C = \tau_c.\bar{c} < x > .0$ $D = \mathrm{c(x)}.\tau_D.D'$
D→E	d	D|E	$D = \tau_d.\bar{d} < x > .0$ $E = \mathrm{d(x)}.\tau_E.E'$
E→F	e	E|F	$E = \tau_e.\bar{e} < x > .0$ $F = o(x).\tau_F.F'$
F→G	f	F|G	$F = \tau_f.\bar{f} < x > .0$ $G = p(x).\tau_G.G'$
G→H	g	G|H	$G = \tau_g.\bar{g} < x > .0$ $H = q(x).\tau_H.H'$
H→I	h	H|I	$H = \tau_h.\bar{h} < x > .0$ $I = r(x).\tau_I.I'$
I→J	i	I|J	$I = \tau_i.\bar{\imath} < x > .0$ $J = r(x).\tau_J.J'$
J→K	j	J|K	$J = \tau_j.\bar{\jmath} < x > .0$ $K = r(x).\tau_K.K'$

Используя рис. 5 и табл. 1, процесс можно представить в виде следующих математических моделей: общей математической модели, представленной в формуле 1, и расширенной математической модели, представленной в формуле 2.

$$P = \mathrm{A|B|C|D|E|F|G|H|I|J|K} + 0. \tag{1}$$

$$\begin{gathered} P = \tau_a.\bar{a} < x > .0\big|\mathrm{a(x)}.\tau_B.B' + \tau_b.\bar{b} < x > .0\big|b(\mathrm{x}).\tau_C.C' + \\ +\tau_c.\bar{c} < x > .0\big|\mathrm{c(x)}.\tau_D.D' + \tau_d.\bar{d} < x > .0\big|\mathrm{d(x)}.\tau_E.E' + \tau_e.\bar{e} < x > .0\ |o(x).\tau_F.F' + \\ + \tau_f.\bar{f} < x > .0\ |p(x).\tau_G.G' + \tau_g.\bar{g} < x > .0\ |q(x).\tau_H.H' + \tau_h.\bar{h} < x > .0\ |r(x).\tau_I.I' + \\ + \tau_i.\bar{\imath} < x > .0\ |r(x).\tau_J.J' + \tau_j.\bar{\jmath} < x > .0\ |r(x).\tau_K.K' + 0. \end{gathered} \tag{2}$$

ВЫВОДЫ

Таким образом, в работе рассмотрено проектирование средства сетевого анализа трафика как автоматизированного средства выявления аномалий, которое дает возможность снизить факторы нарушения информационной безопасности в контексте структурно сложных многокомпонентных информационных систем. Результаты исследования подтверждены функциональной моделью процесса, динамической моделью в нотации *BPMN*, а также информационной моделью проектируемого микросервиса анализа сетевого трафика как компоненты

системы на концептуальном уровне абстракции. Авторами предложно решение, в основе которого лежит нейронная сеть, которая позволит выполнять динамическое совершенствование средства анализа трафика. С помощью математической модели на основе пи-исчислений можно определить обучения нейронной сети и построить набор данных (*dataset*). *BPMN* – модель процесса, предложенная в статье, может быть преобразована в код с применением инструментального средства разработки *Java Workflow Tooling*.

СПИСОК ЛИТЕРАТУРЫ

1. **Винокуров А.** Принципы организации учёта IP-трафика // [Электронный ресурс] URL: https://habr.com/ru/articles/136844/ (дата обращения: 10.01.2023) [Vinokurov A. *Accounting management principles of IP-traffic*. Electronic resource: https://habr.com/ru/articles/136844 (accessed 01.10.2023).]

2. **Конюхов А.** NetFlow, Cisco и мониторинг трафика // [Электронный ресурс]. URL: https://habr.com/ru/articles/175359/ (дата обращения: 12.01.2023) [Konyukhov A. NetFlow, *Cisco and traffic monitoring*. Electronic resource: https://habr.com/ru/articles/175359/ (accessed 01.12.2023).]

3. **Рудь В.** Концептуальное логическое и физическое моделирование данных // [Электронный ресурс] URL: https://marcus-aurelius.ru/articles/Data_modeling.html (дата обращения: 28.01.2023) [Rud' V. *Conceptual logical and physical data modeling*. Electronic resource: https://marcus-aurelius.ru/articles/Data_modeling.html (accessed 01.28.2023).]

4. **Черниговский А.В., Кривов М.В., Истомин А.Л.** Исследование и выбор математической модели сетевого трафика // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2020. №3 (132). С. 84-99 [Chernigovskiy A.V., Krivov M.V., Istomin A.L. *Investigating network traffic and selecting a matching mathematical model*. Herald of the Bauman Moscow State Technical University, Series Instrument Engineering, 2020, no. 3 (132), pp. 84-99 (in Russian).]

5. **Гетьман А.И., Евстропов Е.Ф., Маркин Ю.В.** Анализ сетевого трафика в режиме реального времени: обзор прикладных задач, подходов и решений // [Электронный ресурс]. URL: https://www.ispras.ru/preprints/docs/prep_28_2015.pdf (дата обращения: 03.02.2023) [Get'man A.I., Evstropov E.F., Markin Y.V. *Wirespeed network traffic analysis: survey of applied problems, approaches and solutions*. Electronic resource: https://www.ispras.ru/preprints/docs/prep_28_2015.pdf (accessed 03.02.2023).]

6. **Артемов В.В.** Классификация сетевого трафика // Молодой ученый. 2022. № 26 (421). С. 7-9 [Artemov V.V. *Network traffic classification* // Young Scientist. 2022. No. 26 (421). P. 7-9 (in Russian).]

7. **Методология функционального** моделирования IDEF0: руководящий документ. Москва: ИПК Издательство стандартов, 2000. [*Functional modeling IDEF0 methodology: governing document*. Moscow: Izdatelstvo Standartov, 2000 (in Russian).]

8. **Справочник по символу BPMN 2.0** // [Электронный ресурс] URL https://camundarus.ru/bpmn/reference/ (дата обращения: 18.04.2023) [*Guide to BPMN 2.0 symbols*. Electronic resource: https://camundarus.ru/bpmn/reference/ (accessed 04.18.2023).]

9. **Пи-исчисление** // [Электронный ресурс] URL: https://ru.wikipedia.org/wiki/Пи-исчисление (дата обращения: 30.04.2023) [*Pi-calculus*. Electronic resource: https://ru.wikipedia.org/wiki/Пи-исчисление (accessed 04.30.2023).]

ОБ АВТОРАХ

АЛЕКСЕЕВА Дарья Сергеевна, аспирант кафедры автоматизированных систем управления, УУНиТ, Уфа.

КОНОНОВ Никита Алексеевич, аспирант кафедры автоматизированных систем управления, УУНиТ, Уфа.

АНТОНОВ Вячеслав Викторович, профессор кафедры автоматизированных систем управления, УУНиТ, Уфа.

ЕНИКЕЕВ Рустем Радомирович, доцент кафедры автоматизированных систем управления, УУНиТ, Уфа.

METADATA

Title: Network traffic analysis algorithm in multicomponent structurally complex information systems.

Authors: D.S. Alekseeva [1], N.A. Kononov[2], V.V. Antonov[3], R.R. Enikeev[4]

Affiliation: Ufa University of Science and Technology, Ufa, Russia.

Email: [1]ads.stat@mail.ru, [2]knnv.nkt@gmail.com, [3]antonov.v@bashkortostan.ru, [4]enikeevrr@mail.ru

Language: Russian.

Source: Vestnik UGATU (scientific journal of Ufa University of Science and Technology), vol. 28, no.1 (103), pp. 92-98, 2024. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Abstract: The article discusses the possibility of using a microservice architecture and a microservice approach to implement a component of a complex system aimed at analyzing network information flows in order to identify threats and improve information security. The basic concepts of the information domain in the context of the problem being solved are considered. The formed functional model of the designed microservice allows us to consider the interrelationships of the process elements. The developed information model at the level of abstraction reflects homogeneous connections of common characteristics, which makes it possible to represent the process in the form of dynamic models. The dynamic model in BPMN notation will serve as the basis for the graphoanalytic model of the designed microservice. Using a graphoanalytic model and pi-calculations, it is possible to form a mathematical model that opens up the possibility of creating a formal algorithm and implementing it using high-level languages. It is also possible to use the dynamic model to generate code using Java Workflow Tooling. The formed mathematical model with the use of pi-calculus can be used to determine the method of training a neural network and building a data set. Since the core of the service is a neural network, it opens up the possibility of dynamic improvement of the traffic analysis tool.

Key words: network traffic, network traffic analysis, information security, pi-calculus, categorical approach, abstract information modeling, BPMN, information security algorithm

About authors:

ALEKSEEVA Darya Sergeevna, Postgraduate student at the Department of Automated Control Systems, Ufa, Russia.

KONONOV Nikita Alekseevich, Postgraduate student at the Department of Automated Control Systems, Ufa, Russia.

ANTONOV Vyacheslav Viktorovich, Professor at the Department of Automated Control Systems, Ufa, Russia.

ENIKEEV Rustem Radomirovich, Associate Professor at the Department of Automated Control Systems, Ufa, Russia.