

УДК 004.724  
DOI: 10.15827/0236-235X.123.507-513

Дата подачи статьи: 28.04.18  
2018. Т. 31. № 3. С. 507–513

## **Подходы к идентификации сетевых потоков и организации маршрутов трафика в виртуальном центре обработки данных на базе нейронной сети**

*И.П. Болодурина*<sup>1</sup>, д.т.н., профессор, зав. кафедрой, *prmat@mail.osu.ru*  
*Д.И. Парфенов*<sup>1</sup>, к.т.н., начальник отдела, *fdot\_it@mail.osu.ru*

<sup>1</sup> *Оренбургский государственный университет, г. Оренбург, 460018, Россия*

Ежегодно растет трафик, циркулирующий в сетях передачи данных, и на сегодняшний день его основу составляет трафик Big Data. Целью данного исследования является разработка новых методов маршрутизации трафика в наложенных сетях виртуальных центров обработки данных. Эффективное построение маршрутов в современных компьютерных сетях, обрабатывающих большие потоки данных, является одним из важнейших показателей работы центров обработки данных.

Для решения данной задачи авторами разработан ансамбль моделей, описывающий подход к построению адаптивных маршрутов в наложенных сетях виртуального центра обработки данных. Новизна предлагаемого решения заключается в разработке гибридного подхода, позволяющего на основе методов интеллектуального анализа данных осуществлять управление маршрутизацией в сети виртуального центра обработки данных с учетом данных о состоянии сетевых узлов, наложенных каналов связи и требований QoS, предъявляемых потоками трафика.

Предложенная в работе модель идентификации маршрутов позволяет решать такие задачи, как определение порядка использования цепочек наложенных каналов связи в сети виртуального центра обработки данных и установление правил для обеспечения качества обслуживания для критически важного трафика. На базе построенных моделей реализовано программно-алгоритмическое решение, представляющее собой модуль для контроллера программно-конфигурируемой сети.

Описанный в данной статье алгоритм позволяет получать наборы квазиоптимальных и оптимальных правил маршрутизации за полиномиальное время. Эффективность предложенного решения доказана путем экспериментального исследования с использованием реальной сетевой инфраструктуры виртуального центра обработки данных. Сопоставление результатов экспериментального исследования, проводившегося с использованием известных алгоритмов маршрутизации, показало не только сокращение времени отклика в сети, но и снижение нагрузки на сетевые узлы, обрабатывающие трафик.

**Ключевые слова:** *маршрутизация трафика, нейронные сети, виртуальный ЦОД, Big Data, QoS.*

В последние годы объем данных, которые требуется обрабатывать и анализировать, растет экспоненциально. Это обусловлено тем, что информация, раньше считавшаяся бесполезной, долго не хранилась и не обрабатывалась в автоматизированном режиме, в настоящий момент приобретает высокую ценность. Одним из примеров таких данных являются журналы доступа к информационным системам, события безопасности, возникающие в сегментах сети, и другие неструктурированные данные. На современном этапе развития информационных технологий появились средства, позволяющие хранить, обрабатывать и анализировать большие объемы данных в режиме реального времени. При этом с увеличением объема данных растет и интенсивность циркуляции трафика в инфокоммуникационных сетях. Таким образом, сегодня можно говорить о том, что основу современных конвергентных сетей составляет трафик Big Data.

Для эффективной работы с трафиком Big Data крупные коммерческие и государственные организации, в том числе промышленные предприятия, переносят свою ИТ-инфраструктуру в виртуальные центры обработки данных (ЦОД). При этом виртуализация компонентов ЦОД применяется как на прикладном уровне, так и на уровне сети. Важные бизнес-приложения и сервисы, а также обрабаты-

ваемые ими данные размещаются на базе мультиоблачных платформ, а в качестве базового компонента для передачи данных в виртуальном ЦОД применяется программно-конфигурируемая сеть [1, 2]. Такой подход позволяет обеспечить гибкое управление сетевыми ресурсами, а также дает возможность поддерживать необходимое качество обслуживания (QoS) в соответствии с решаемыми конечными пользователями задачами. Однако, кроме трафика Big Data, в конвергентных сетях, как правило, присутствуют и другие типы потоков данных. При этом каждый поток данных обладает индивидуальными характеристиками, на базе которых определяются требования QoS [3]. Принимая во внимание перечисленные особенности мультисервисных сетей ЦОД, проблема обеспечения качества обслуживания для трафика Big Data становится особенно актуальной [4]. Одним из подходов, позволяющих решить данную задачу, является организация адаптивной маршрутизации трафика [5–7]. В контексте виртуальной инфраструктуры он означает управление потоками данных, циркулирующими в сети ЦОД, на уровне каналов связи.

Проблема организации эффективной маршрутизации в виртуальной сети, как правило, заключается в том, что на базе традиционного ЦОД развернуто несколько облачных платформ [8]. Каждая из

таких платформ может использовать разные среды виртуализации с различными уровнями абстракции. В связи с этим потоки данных внутри ЦОД проходят через различные виртуальные интерфейсы, имеющие собственную задержку, что не позволяет оптимальным образом согласовывать потоки данных. Поэтому для эффективной маршрутизации трафика в виртуальном ЦОД необходимо разработать единую стратегию планирования маршрутов, которая не зависит от среды виртуализации и в то же время может обеспечить гарантированное качество обслуживания для циркулирующих потоков данных.

В рамках исследования сформулированы следующие требования в области маршрутизации трафика в программно-управляемой инфраструктуре виртуального ЦОД:

- поддержка многопутевой маршрутизации для обеспечения сбалансированного использования сети и повышения общего качества обслуживания;
- переход к моделям трафика, основанным на потоках и методах его маршрутизации;
- использование составных параметров, характеризующих каналы связи и позволяющих максимально учитывать требования к QoS в зависимости от класса идентифицированного трафика.

В настоящее время нет достаточно эффективных методов маршрутизации, адаптированных к потокам трафика Big Data в сети виртуального ЦОД [9]. Прежде всего это связано с проблемой идентификации конкретных приложений и сервисов в общем потоке трафика ЦОД. Существующие решения для идентификации потоков используют в качестве анализируемых параметров лишь незначительный ряд характеристик, например таких, как длина и заголовки сетевого пакета, IP-адрес источника и назначения, а также другие [10]. Это не позволяет достаточно точно идентифицировать приложения и сервисы, а значит, снижает эффективность решений, организующих управление маршрутизацией в сети.

Подобные проблемы наблюдаются также и у методов, используемых для организации традиционной маршрутизации. Существующие подходы не учитывают всю собираемую информацию о каналах связи. Так, важным фактором, влияющим на задержку, является время обработки трафика на сетевых узлах. Этот параметр напрямую зависит от загрузки интерфейсов и самого сетевого узла, обрабатывающего трафик, но при выборе маршрута он не учитывается.

Таким образом, в рамках исследования установлено, что существующие подходы к маршрутизации не учитывают состояние сетевых объектов, требования QoS, а также их коммуникационные схемы взаимодействия облачных приложений и сервисов, возникающие в процессе наложения виртуальных маршрутов на физические каналы связи.

## Методы и подходы

При решении поставленной задачи по организации адаптивной маршрутизации трафика в сети виртуального ЦОД в настоящем исследовании предлагается использовать методы интеллектуального анализа данных и машинного обучения. Выбор методов обусловлен высокой эффективностью их работы с неструктурированными и динамически изменяющимися и непрерывно поступающими от различных источников (сетевого оборудования, виртуальных и физических вычислительных узлов) потоками данных. Кроме этого, предлагаемые методы способны осуществлять быстрый анализ данных в режиме реального времени. Такой подход позволяет с минимальной задержкой принимать управленческие решения по организации маршрутов трафика в программно-конфигурируемой сети, что, в свою очередь, напрямую влияет на эффективность работы всей инфраструктуры виртуального ЦОД.

Одним из основных аспектов в рамках предлагаемого подхода является решение задачи идентификации маршрутов в наложенной сети виртуального ЦОД. В процессе исследования осуществлена декомпозиция задачи идентификации маршрутов на две последовательно связанные между собой подзадачи. Первая направлена на классификацию каналов связи в виртуальном ЦОД по ряду признаков, характерных для всех наложенных сетей. На втором этапе на основе полученных данных выполняется кластеризация маршрутов в соответствии с требованиями к QoS, предъявляемыми текущими потоками трафика. Агрегированные данные объединяются с информацией о текущих потоках трафика и подаются на вход нейронной сети для принятия решения о выборе подходящего маршрута. Таким образом, решение поставленной задачи по организации адаптивной маршрутизации позволяет определять приоритеты при формировании полосы пропускания потоков трафика, а также устанавливать правила по управлению сетью ЦОД.

С технической и программной сторон для организации эффективного сбора данных о состоянии физических и наложенных каналов связи в сети виртуального ЦОД и анализа их использования в рамках исследования предложен гибридный подход, основанный на совместном применении технологии самоорганизующихся интеллектуальных программных агентов и виртуализации сетевых функций. Это позволяет не только эффективно собирать данные о состоянии каналов, но и предварительно обрабатывать данные непосредственно на сетевых узлах. В рамках предложенного подхода программные агенты обмениваются данными внутри группы о состоянии каналов и доступных маршрутах, а управляющий агент передает обработанную и проанализированную информацию в сжатом виде на контроллер сети.

Такой подход сокращает накладные расходы на интеллектуальную инфраструктуру управления адаптивной маршрутизацией за счет двух ключевых факторов: объединение источников информации и узлов анализа в рамках одного объекта, что оптимизирует количество используемых вычислительных ресурсов; передача на контроллер сети уже обработанных данных, что сокращает объем технической информации, циркулирующей по сети ЦОД.

Таким образом, научная новизна предлагаемого решения заключается в разработке гибридного подхода, позволяющего на основе методов интеллектуального анализа данных управлять маршрутизацией в сети виртуального ЦОД с учетом данных о состоянии сетевых узлов, каналов связи, в том числе наложенных сетей, и требований QoS, предъявляемых потоками трафика.

### Модель идентификации каналов связи в наложенной сети виртуального ЦОД

Модель идентификации каналов связи в наложенной сети является основой для организации адаптивной маршрутизации для облачных приложений и сервисов, расположенных в виртуальном ЦОД. Разработанная модель поэтапно идентифицирует все каналы связи путем решения задач классификации и кластеризации. Опишем каждый из перечисленных элементов.

Так, на этапе классификации наложенных каналов связи в рамках настоящего исследования предлагается использовать метод обучения без учителя. Это позволяет не только идентифицировать новые каналы, возникающие динамически в сети виртуального ЦОД, но и изменять принятую ранее классификацию в зависимости от текущей загруженности сетевых узлов и топологии наложенной сети. Кроме того, использование метода обучения без учителя не нуждается в начальной ручной разметке входных данных, он основывается только на подобии между классифицируемыми каналами связи.

Формально задача классификации наложенных каналов связи определяется следующим образом. Пусть дано множество каналов связи  $Nv = \{nv_1, \dots, nv_n\}$ , в котором каждый канал передачи данных  $nv_i$  характеризуется множеством атрибутов  $A = \{a_{1,i}, \dots, a_{i,p}\}$ , где  $p$  – количество атрибутов. Прежде чем классифицировать каналы связи, важно установить их атрибуты. Количество атрибутов определим по результатам анализа свойств, влияющих на качество обслуживания, что также играет ключевую роль при выборе маршрута. В качестве атрибутов будем использовать следующий набор характеристик, позволяющих однозначно идентифицировать наложенный канал связи в виртуальном ЦОД: задержка канала, jitter, процент потери пакетов, пропускная способность, узлы агрегации канала (для

каждого канала связи предусмотрены два узла агрегации: источник и назначение), текущая интенсивность потока передаваемых данных, используемые правила QoS. Для более точной идентификации каналов с целью построения оптимальных маршрутов можно использовать и другие статистические характеристики, определенные в Service Level Agreement (SLA), на базе которого в виртуальном ЦОД формируются требования к качеству обслуживания в сети. Определим также множество классов каналов связи  $C = \{c_1, \dots, c_k\}$ . Тогда для решения задачи классификации необходимо определить такое отображение  $f: Nv \rightarrow C$ , чтобы наложенный канал сети ЦОД  $nv_i$  соответствовал только одному классу  $c_i$ . Количество классов может быть чрезмерно большим, поэтому осуществляется переход ко второму этапу задачи идентификации каналов связи – к кластеризации.

Цель кластеризации – построение оптимального разбиения объектов на группы, то есть разбиение  $n$  классов на  $Cn$  кластеров. При этом для кластеризации объектов выбирается метрика подобия  $d(k_j, k_p)$ .

Формально задачу кластеризации каналов связи в сети ЦОД можно сформулировать следующим образом. Пусть даны множества наложенных каналов связи  $Nv = \{nv_1, \dots, nv_n\}$  и желаемое количество кластеров  $Cn$ . Требуется определить отображение  $f: Nv \rightarrow \{k_1, \dots, k_{Cn}\}$ , такое, чтобы каждый из каналов связи был отнесен только к одному кластеру  $k_i$ ,  $1 \leq i \leq Cn$ , при условии, что  $D = \bigcup_{j=1}^{Cn} k_j$  и  $k_i \cap k_j = \emptyset, \forall i \neq j$ . Оптимальность проведенной кластеризации определим как требование минимизации среднеквадратичной ошибки разбиения.

### Модель организации маршрутов трафика

Следующий этап решения задачи после идентификации каналов связи – определение оптимального маршрута для каждого класса трафика и генерация соответствующих правил на контроллер сети. Полученные из модели идентификации данные о каналах связи представим в виде графа, иллюстрирующего топологию наложенных сетей и позволяющего определить возможные маршруты для каждого класса согласно коммуникационной схеме взаимодействия облачных приложений и сервисов. В данном графе каждый узел представляет собой элемент сети, отвечающий за маршрутизацию и передачу данных, а ребра описывают сетевые соединения между ними. При этом веса ребер рассчитываются с использованием интегральной оценки на базе введенных ранее показателей, характеризующих каналы связи в сети ЦОД. Таким образом, на базе данного описания мы получаем полное представление о состоянии сети в виртуальном ЦОД. Сам граф можно представить в сле-

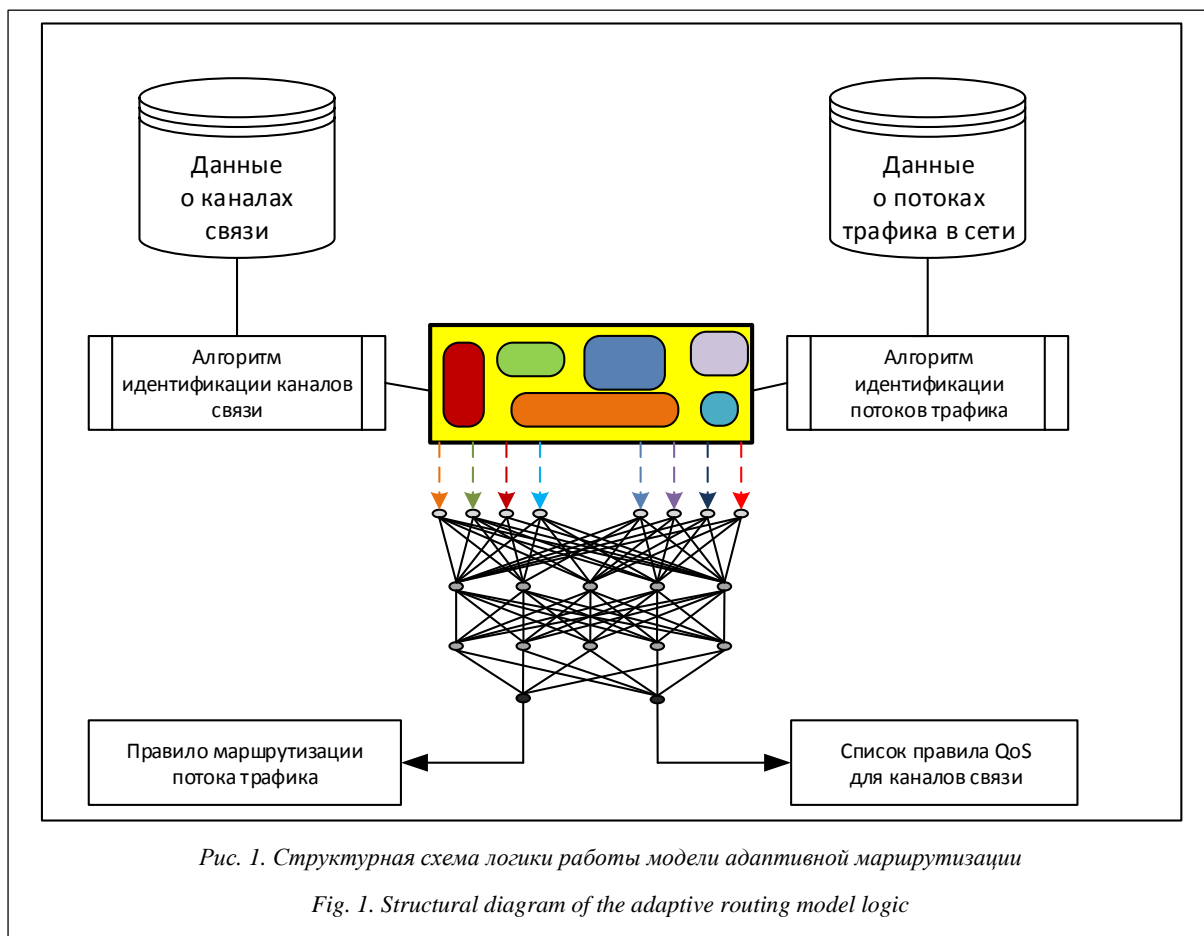
дующем виде:  $RDC = (V, E)$ , где  $V = Nc \cup Nd \cup Nr$  обозначает множество вершин, включающих три типа сетевых объектов виртуального ЦОД, а именно  $Nc$  – множество вершин, задействованных при текущей маршрутизации трафика,  $Nd$  – множество сетевых объектов, задействованных при перестроении маршрутов трафика,  $Nr$  – узлы, не входящие в маршрут, достигающий заданного узла.

Для определения оптимальных маршрутов в рамках построенной модели будем использовать нейронные сети. Они легко работают в распределенных системах и позволяют распараллеливать процессы на множество узлов, снижая тем самым вычислительную нагрузку, а также подстраивать весовые коэффициенты, основываясь на исходных данных. Это помогает сделать выбор ключевых характеристик менее субъективным. Для поиска оптимального маршрута в настоящем исследовании предлагается использовать гибридную нейронную сеть, состоящую из самоорганизующейся сети Кохонена и многослойного персептрона. Преимуществом сети Кохонена является способность идентифицировать новые кластеры. Обученная сеть распознает кластеры в данных обучения и назначает все данные тому или иному из них. Если сеть затем встречает набор данных, который отличается от любого из известных образцов, он будет независимо определять новый кластер элементов. Струк-

турная схема логики работы модели адаптивной маршрутизации показана на рисунке 1.

В качестве исходных данных для построения карты Кохонена в части разбиения каналов связи на множества используются данные, полученные из модели идентификации каналов связи виртуального ЦОД. Каждый кластер формирует числовые показатели интенсивности потоков, проходящих через канал данного типа. Фактически отдельные элементы карты Кохонена представляют собой группу наложенных каналов связи, идентичных по своим характеристикам. Входной вектор для самоорганизующейся сети Кохонена содержит следующие компоненты:  $x_1$  – узел назначения,  $x_2$  – узел источника,  $x_3$  – jitter,  $x_4$  – пропускная способность канала,  $x_5$  – процент потери пакетов,  $x_6$  – загрузка источника,  $x_7$  – количество переданных байтов по каналу,  $x_8$  – загрузка узла назначения.

После кластеризации с использованием на самоорганизующейся сети Кохонена данные о потоке трафика, для которого необходимо осуществить поиск подходящего маршрута, и наложенных каналах связи ЦОД подаются на вход многослойного персептрона. Обученная поиску маршрутов нейронная сеть анализирует данные о состоянии сети виртуального ЦОД с учетом требований поступившего потока трафика. На выходе нейронной сети формируются два вектора. Первый вектор



формирует правила для маршрутизации трафика, передаваемые на контроллер, второй – список правил QoS для каналов связи, попавших в маршрут для текущего потока трафика.

**Программно-алгоритмическая реализация адаптивной маршрутизации**

На базе разработанных моделей спроектирован модуль адаптивной маршрутизации, позволяющий оптимизировать потоки трафика в сети виртуального ЦОД. Схема логики взаимодействия функциональных элементов модуля представлена на рисунке 2.

Разработанный модуль в своей работе использует приведенный далее обобщенный алгоритм выбора маршрутов для потоков трафика, циркулирующих по наложенным сетям внутри виртуального ЦОД.

*Шаг 1.* Идентификация наложенных сетей виртуального ЦОД путем классификации по характерным признакам, представленным в модели канала связи.

*Шаг 2.* На основе полученных данных кластеризация каналов по признакам QoS с использованием самоорганизующейся карты Кохонена.

*Шаг 3.* На основе данных о текущем потоке трафика поиск оптимального маршрута, удовлетворяющего требованиям QoS и SLA, с использованием нейронной сети.

*Шаг 4.* Если найдены подходящие маршруты, ранжирование по признаку загруженности каналов и сетевых устройств.

*Шаг 4.1.* Для наименее нагруженного маршрута внесение правил через контроллер программно-конфигурируемой сети для текущего потока трафика.

*Шаг 5.* Если маршрут не найден, поиск наименее нагруженной цепочки каналов среди всех имеющихся наложенных каналов до узла назначения.

*Шаг 5.1.* Оценка возможностей принятия трафика найденной цепочкой наложенных каналов.

*Шаг 5.2.* Если канал не удовлетворяет одному из условий, полное освобождение его от другого трафика, используя шаги 1–4.

*Шаг 6.* Применение к цепочке каналов политики QoS на контроллере программно-конфигурируемой сети и маршрутизация текущего потока через него.

Таким образом, основной задачей разработанного программно-алгоритмического решения является поиск оптимальных маршрутов, проходящих через наименее нагруженные узлы, позволяющие максимизировать производительность работы сети с учетом существующих потоков приложений и сервисов и их требований к задержкам, возникающим в процессе работы в виртуальном ЦОД.

**Экспериментальное исследование**

Исследование работы модуля адаптивной маршрутизации для контроллера сети виртуального ЦОД выполнялось на базе лаборатории облачных вычислений Оренбургского государственного университета. В качестве экспериментального

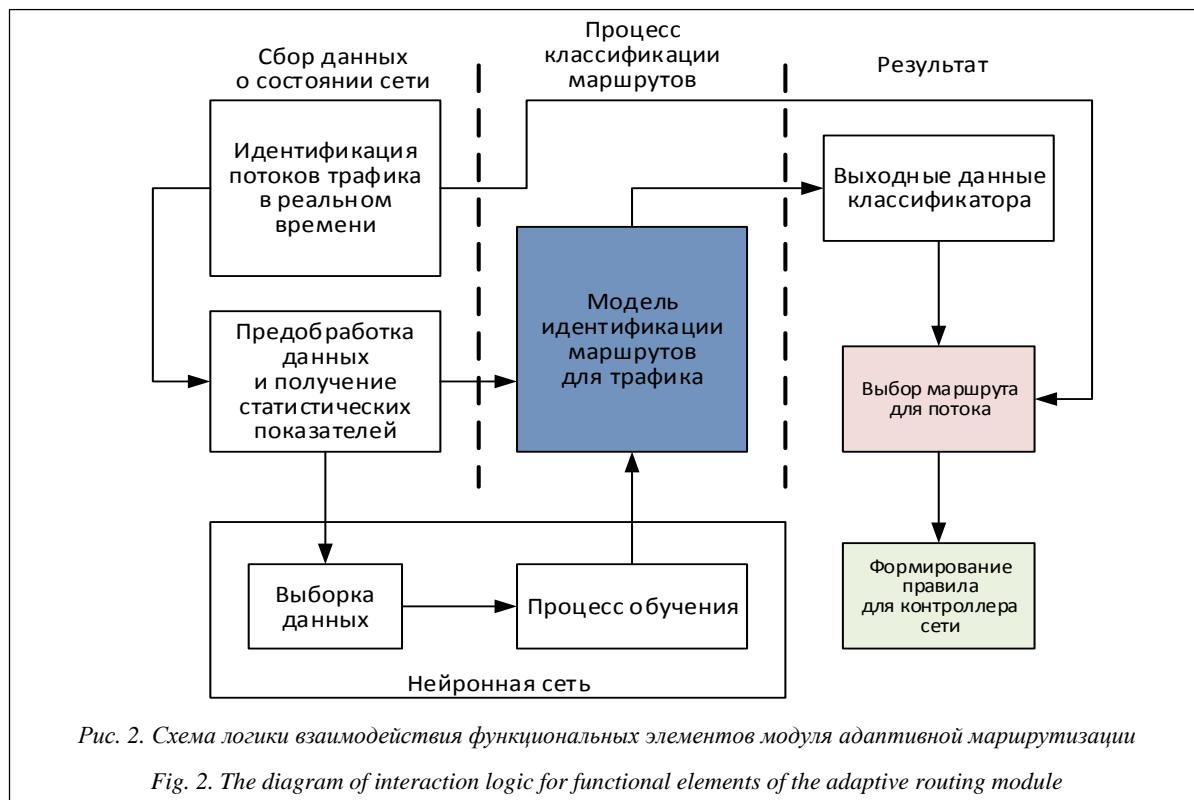


Рис. 2. Схема логики взаимодействия функциональных элементов модуля адаптивной маршрутизации

Fig. 2. The diagram of interaction logic for functional elements of the adaptive routing module

**Результаты экспериментального исследования**  
**The results of the experimental study**

Алгоритм сетевой маршрутизации	Измеряемый параметр	Число сетевых объектов ЦОД						
		10	25	50	100	250	500	1000
Алгоритм Дейкстры	Ресурсоемкость	75	680	2160	8450	13250	17850	340500
	Время отклика сети, мс	0.129	0.193	0.250	3.579	5.629	23.378	46.507
Алгоритм Йена	Ресурсоемкость	146	1104	1802	7230	12119	15214	113815
	Время отклика сети, мс	0.033	0.120	0.237	0.592	1.784	6.105	25.542
Алгоритм адаптивной маршрутизации	Ресурсоемкость	25	90	450	680	890	1250	3400
	Время отклика сети, мс	0.016	0.046	0.078	0.260	0.355	4.503	15.538

стенда использовалась сеть, состоящая из аппаратных OpenFlow-коммутаторов, виртуальных машин под управлением Linux с головным контроллером NOX. Для оценки устойчивости работы с потоками трафика, ориентированными на Big Data, проведен ряд экспериментальных исследований с последовательным увеличением числа сетевых объектов ЦОД от 10 до 1 000. Для создания наложенных сетей с таким количеством устройств использовались функциональные возможности виртуальных сетевых функций, позволяющих на базе среды виртуализации запускать программные аналоги физических узлов. Для этого на экспериментальном стенде использовались десять вычислительных узлов, на каждом из которых создавались от 10 до 100 сетевых объектов с использованием технологии виртуализации сетевых функций и среды эмуляции Mininet. В таблице представлены результаты исследования в сравнении с двумя наиболее известными алгоритмами маршрутизации трафика – классическим алгоритмом Дейкстры и алгоритмом, применяемым для многопутевой маршрутизации (алгоритм Йена).

Исследования показали, что предлагаемое программно-алгоритмическое решение позволяет не только снизить ресурсоемкость маршрутизации при значительном количестве устройств, но и сократить время отклика в сети по сравнению с традиционно используемыми алгоритмами.

### Заключение

Таким образом, в настоящем исследовании предложено решение, основанное на ансамбле моделей, включающем в себя модель идентификации каналов связи в наложенной сети виртуального ЦОД и модель организации маршрутов трафика. Представленные решения оформлены в виде алгоритма и реализованы в виде модуля адаптивной маршрутизации для программно конфигурируемой сети. Разработанное в рамках исследования программное средство позволяет улучшить время отклика на 25 % – один из ключевых показателей, влияющих на качество обслуживания в сети ЦОД. В сравнении с известными алгоритмами маршрути-

зации предлагаемое программно-алгоритмическое решение обладает меньшей ресурсоемкостью, что также снижает накладные расходы на сетевую инфраструктуру.

*Работа выполнена при поддержке РФФИ (научные проекты №№ 16-07-01004 и 16-37-60086) и гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – кандидатов наук (МК-1624.2017.9).*

### Литература

1. Решетников В.Н., Болодурин И.П., Парфенов Д.И. Моделирование размещения сервис-ориентированных приложений в программно-управляемой инфраструктуре виртуального центра обработки данных // Программные продукты и системы. 2016. № 4. С. 15–22.
2. Herguner K., Kalan R.S., Cetinkaya C., Sayit M. Towards QoS-aware routing for DASH utilizing MPTCP over SDN. Proc. IEEE Conf. NFV-SDN, 2017, pp. 1–6.
3. Kim T.H., Kim S., Lee K.Y., Park S.Y. A QoS assured network service chaining algorithm in network function virtualization architecture. Proc. 15th IEEE/ACM Intern. Sympos. on Cluster, Cloud and Grid Computing, 2015, pp. 1221–1224.
4. Yisheng L., Yanjie D., Wenwen K., Zhengxi L., Fei-Yue W. Traffic flow prediction with big data: a deep learning approach. Proc. IEEE Transactions on Intelligent Transportation Systems, 2015, vol. 16, no. 2, pp. 865–873.
5. Chemeritsky E., Smelansky R. On QoS management in SDN by multipath routing. Proc. 1st Intern. Sc. and Tech. Conf. (Modern Networking Technologies), 2014, pp. 1–6.
6. Tomovic S., Radusinovic I., Prasad N. Performance comparison of QoS routing algorithms applicable to large-scale SDN networks. Proc. Intern. Conf. on Computer as a Tool, EUROCON-2015, 2015, pp. 1–6.
7. Layuan L., Chunlin L. A QoS multicast routing protocol for dynamic group topology. Information Sciences, 2005, vol. 169, no. 1-2, pp. 113–130.
8. Носков А.Н., Манов И.А. Разработка механизма адаптивной маршрутизации трафика в программно-конфигурируемых сетях // Моделирование и анализ информационных систем. 2015. Вып. 22. С. 521–532.
9. Bolodurina I.P., Parfenov D.I. Dynamic routing algorithms and methods for controlling traffic flows of cloud applications and services. Bulletin of the South Ural State Univ. Series: Computational Mathematics and Software Engineering. 2017. vol. 6, no. 2. pp. 84–98.
10. Шелухин О.И., Симонян А.Г., Ванюшина А.В. Влияние структуры обучающей выборки на эффективность классификации приложений трафика методами машинного обучения // T-COMM: Телекоммуникации и транспорт. 2017. Вып. 11. С. 25–31.

### The approaches to identification of network flows and organization of traffic routes in a virtual data center based on a neural network

I.P. Bolodurina<sup>1</sup>, Dr.Sc. (Engineering), Professor, Head of Chair, prmat@mail.osu.ru  
D.I. Parfenov<sup>1</sup>, Ph.D. (Engineering), Head of Department, fdot\_it@mail.osu.ru

<sup>1</sup> Orenburg State University, Orenburg, 460018, Russian Federation

**Abstract.** Every year traffic in data transmission networks grows. Currently, their basis is Big Data traffic. The purpose of this study is to develop new methods for routing traffic in superimposed networks of virtual data centers. The effective construction of routes in modern computer networks that process Big Data flows is one of the most important indicators of the operation of data centers.

In order to solve this problem, the authors developed an ensemble of models describing the approach to constructing adaptive routes in superimposed networks of a virtual data center. The novelty of the proposed solution is to develop a hybrid approach that allows (based on methods of data mining) managing routing in a virtual data center network, taking into account data on the status of network nodes, superimposed communication channels and QoS requirements imposed by traffic flows.

The proposed route identification model allows solving such problems as determining the order of using chains of superimposed communication channels in a virtual data center network and establishing rules for providing quality of service for mission-critical traffic. A software-algorithmic solution is implemented based on the built models. It is a module for a controller of a software-configurable network.

The algorithm described in the paper allows obtaining sets of quasi-optimal and optimal routing rules in polynomial time. The effectiveness of the proposed solution is proved by experimental research using a real network infrastructure of a virtual data center. Comparison of the results of an experimental study carried out using known routing algorithms showed a reduction in response time in the network, as well as a reduction in the load on network nodes processing traffic.

**Keywords:** traffic routing, neural networks, virtual data center, Big Data, QoS.

**Acknowledgements.** The work has been supported by the Russian Foundation for Basic Research (scientific projects no. 16-07-01004 and 16-37-60086) and a grant of the President of the Russian Federation for state support of young Russian scientists who are candidates of science (MK-1624.2017.9).

#### References

1. Reshetnikov V.N., Bolodurina I.P., Parfenov D.I. Modeling of placing service-oriented applications in a software-defined infrastructure of the virtual data center. *Software & Systems*. 2016, vol. 4, pp. 15–22 (in Russ.).
2. Herguner K., Kalan R.S., Cetinkaya C., Sayit M. Towards QoS-aware Routing for DASH Utilizing MPTCP over SDN. *IEEE Conf. on NFV-SDN*. 2017, pp. 1–6.
3. Kim T.H., Kim S., Lee K.Y., Park S.Y. A QoS assured network service chaining algorithm in network function virtualization architecture. *Proc. 15th IEEE/ACM Intern. Symp. on Cluster, Cloud and Grid Computing*. 2015, pp. 1221–1224.
4. Yisheng L., Yanjie D., Wenwen K., Zhengxi L., Fei-Yue W. Traffic flow prediction with big data: a deep learning approach. *IEEE Trans. on Intelligent Transportation Systems*. 2015, vol. 16, no. 2, pp. 865–873.
5. Chemeritsky E., Smelansky R. On QoS management in SDN by multipath routing. *Proc. 1st Intern. Sci. and Tech. Conf. (Modern Networking Technologies)*. 2014, pp. 1–6.
6. Tomovic S., Radusinovic I., Prasad N. Performance comparison of QoS routing algorithms applicable to large-scale SDN networks. *Intern. Conf. on Computer as a Tool (EUROCON)*. 2015, pp. 1–6.
7. Layuan L., Chunlin L. A QoS multicast routing protocol for dynamic group topology. *Information Sciences*. 2005, vol. 169, no. 1-2, pp. 113–130.
8. Noskov A.N., Manov I.A. Development of an adaptive routing mechanism in software-defined networks. *Modeling and Analysis of Information Systems*. 2015, vol. 22, no. 4, pp. 521–532 (in Russ.).
9. Bolodurina I.P., Parfenov D.I. Dynamic Routing Algorithms and Methods for Controlling traffic Flows of Cloud Applications and Services. *Bulletin of the South Ural State Univ. Ser. Computational Mathematics and Software Engineering*. 2017, vol. 6, no. 2, pp. 84–98.
10. Shelukhin O.I., Simonyan A.G., Vanyushina A.V. Influence of training sample structure on traffic application efficiency classification using machine-learning methods. *T-Comm*. 2017, vol. 11, no. 2, pp. 25–31 (in Russ.).