

УДК 621.391

ОБЗОР ТЕХНОЛОГИЧЕСКИХ И ТЕОРЕТИЧЕСКИХ РЕШЕНИЙ В ОБЛАСТИ МАРШРУТИЗАЦИИ НА ОСНОВЕ КАЧЕСТВА ОБСЛУЖИВАНИЯ



[О.Ю. ЕВСЕЕВА](#),

[С.В. ГАРКУША](#)

Харьковский национальный
университет радиоэлектроники

Abstract - Under rising requirements for telecommunication network role of routing increases, because routing is an effective tool to ensure the specified end-to end quality of service (QoS). Nowadays the routing is transformed into QoS-based routing which focuses on the searching one or more paths for delivering packets where all described in the SLA requirements for quality of service must be satisfied. A set of technological and theoretical solutions is developed for solving QoS-based routing problem. About it several concepts deserves attention among which source routing, precomputed routing, route server concept. From theoretical point of view QoS-based routing problem can be solved by using graph theory approach or flow-based mathematical model of network. The graph theory approach is based on structure only and doesn't take into account flow properties of transmitted traffic. As a result it doesn't allow allocating network resources between different traffics and concordance between routing, accesses, queuing, and reservation. As a rule a flow-based approach solves QoS-based routing problem as mathematical programming problem with QoS-constrains. For example, it can be tensor constrains. The article contains possible lines of development of QoS-based routing that include conversion to dynamical metrics and dynamical models with QoS-constrains, not only load balancing but quality balancing, concordance between routing and another traffic control tasks.

Анотація – Проведено огляд рішень задач маршрутизації на основі якості обслуговування, що включає як протокольні варіанти, так і теоретичні рішення, які ще не отримали технологічної реалізації. Огляд охоплює комбінаторні методи маршрутизації, а також поточкові моделі та методи формалізації та розв'язання маршрутних задач.

Аннотация – Проведен обзор решений задач маршрутизации на основе качества обслуживания, включающий как протокольные варианты, так и теоретические решения, которые еще не получили технологической реализации. Обзор охватывает комбинаторные методы маршрутизации, а также потоковые модели и методы формализации и решения маршрутных задач.

Введение

С ростом требований к качеству обслуживания (Quality of Service) в современных телекоммуникационных сетях, развивающихся в соответствии с концепцией построения сетей следующего поколения (Next Generation Network, NGN), все больше внимания уделяется средствам маршрутизации. Причина этого состоит в том, что функционал DiffServ, основанный на приоритетной обработке пакетов на узлах (маршрутизаторах) телекоммуникационной сети (ТКС), способен повысить уровень QoS на отдельных элементах сети, в то время как именно маршрутизация является инструментом обеспечения требуемых значений ключевых межконцевых (end-to-end) QoS-показателей. Таким образом, технологии маршрутизации вышли за рамки

средств поиска лишь доступных маршрутов в сети; протокол маршрутизации в идеале должен обеспечивать расчет одного или нескольких путей доставки пакетов, вдоль которых будут выполняться все описанные в SLA-договоре требования к качеству обслуживания. Это является формулой концепции QoS-маршрутизации (QoS-based routing) [1, 2], определяющей направление развития моделей, методов, а впоследствии алгоритмов и протоколов маршрутизации.

Ввиду актуальности задач QoS-маршрутизации как с точки зрения теории, так и практики учеными и производителями сетевого оборудования предложен ряд разноплановых решений, отличающихся степенью учета состояния ТКС, вычислительной сложностью, уровнем гарантий относительно качества обслуживания, а в конечном итоге – предполагаемой областью применения. В этой связи целью данной статьи является обзор технологических и теоретических решений в области маршрутизации на основе качества обслуживания для сравнения их преимуществ и недостатков, а также определения наиболее перспективных путей развития данного, бесспорно важного для телекоммуникаций направления.

I. Обзор технологических решений в области QoS-маршрутизации

В настоящее время в сетях IP, составляющих транспортную основу NGN, используется достаточно широкий перечень протоколов динамической маршрутизации: протокол маршрутной информации (Routing Information Protocol, RIP), протокол маршрутизации внутреннего шлюза (Interior Gateway Routing Protocol, IGRP), расширенный протокол маршрутизации внутреннего шлюза (Enhanced IGRP, EIGRP), протокол взаимодействия промежуточных систем (Intermediate System-to-Intermediate System, IS-IS), открытый протокол кратчайшего пути (Open Shortest Path First, OSPF), протокол пограничного шлюза (Border Gateway Protocol, BGP) [3 – 9] и др.

Средства поддержки качества обслуживания в современных протоколах маршрутизации в последнее время достаточно сильно эволюционировали, прежде всего, за счет пересмотра метрик, используемых при выборе маршрута. Во-первых, происходит отказ от топологических метрик – числа переприемов (hops), характерного для протокола RIP, и осуществляется переход на QoS-метрики, основанные на учете основных показателей качества обслуживания: скорости передачи, средних задержек, уровня потерь в трактах передачи сети. В результате маршруты передачи пакетов того или иного трафика прокладываются с учетом QoS-показателей вдоль них. Во-вторых, все больше протоколов поддерживают так называемые комбинированные (комбинированные) метрики, в рамках которых учитываются одновременно несколько основных QoS-показателей. Так, в рамках протокола EIGRP в качестве метрики интерфейса используется следующая величина [5, 9]:

$$M_{EIGRP} = 256 \left[K_1 M_b + \frac{K_2 M_b}{256 - M_l} + K_3 M_d \right] \frac{K_5}{M_r + K_4}, \quad (1)$$

где K_1, K_2, K_3, K_4, K_5 – постоянные коэффициенты; M_b, M_d, M_r, M_l – отдельные метрики маршрута, причем M_b – величина обратной логической пропускной способности маршрута; M_d – кумулятивная логическая задержка маршрута, M_r и M_l – надежность и загрузка маршрута соответственно.

Весовые коэффициенты $K_i, i = \overline{1,5}$ (1) являются административно настраиваемыми параметрами, по умолчанию их значения равны $K_1=K_3=1, K_2=K_4=K_5=0$. В результате формула (1) принимает вид (часть $\frac{K_5}{M_r + K_4}$ не используется):

$$M_{EIGRP} = 256[K_1M_b + K_3M_d] . \quad (2)$$

Однако существующие на сегодняшний день протоколы не укладываются в полной мере в рамки концепции QoS-маршрутизации, поскольку в них заложена идея поиска кратчайшего пути в одной, пусть даже композитной, метрике без учета достигаемых при этом значений других метрик [3 – 9]. Так, например, несмотря на то, что в рамках EIGRP при выборе маршрута может учитываться множество показателей качества обслуживания (задержка, загрузка, надежность, пропускная способность), использование комбинированной метрики (1) вовсе не гарантирует наилучших значений, например, задержки, вдоль найденного маршрута. Более того, задержка, используемая в EIGRP, является административно назначаемым при настройке интерфейса параметром и в общем случае может не соответствовать той величине задержки, которая на самом деле имеет место на данном интерфейсе.

Другой проблемой, связанной с применением существующих протоколов маршрутизации, является несбалансированное использование сетевых ресурсов. В основу всех перечисленных протоколов маршрутизации положены алгоритмы нахождения кратчайшего пути в графе (Дийкстры или Беллмана-Форда) [3 – 9], результатом работы которых является один, кратчайший в выбранной метрике, путь. Поскольку ни в одном из упомянутых выше протоколов в метрику не входит доступная (не задействованная) пропускная способность трактов передачи, всегда для заданной пары адресатов в качестве кратчайшего будет выбираться один и тот же путь, что в конечном итоге приводит к его перегрузке даже при наличии свободных обходных маршрутов. Исключение составляет алгоритм Constrained Shortest Path First (CSPF), применяемый в сетях MPLS TE для расчета пути коммутации меток (Label Switching Path, LSP) и учитывающий при этом текущую загруженность трактов передачи и атрибуты нового LSP. Целесообразным решением в данной ситуации является многопутевая маршрутизация с балансировкой нагрузки не только между путями равной стоимости, что заложено во всех протоколах маршрутизации, но и между путями неравной стоимости [10], что присутствует только в протоколах IGRP, EIGRP и внутридоменной версии BGP – iBGP (interior BGP). Как правило, балансировка между маршрутами с различной стоимостью в рамках данных протоколов требует от администраторов сети дополнительных настроек и зачастую ими не используется.

Вычисление маршрута в рассмотренных выше протоколах маршрутизации как внутренних, так и внешних (межсетевых или междоменных), реализуется распределенно и в момент поступления трафика в сеть, т.е. по требованию. Однако, как отмечается в ряде работ [11 – 15], при переходе к QoS-маршрутизации заслуживают внимания следующие маршрутные концепции, во-первых, концепция централизованного вычисления путей в рамках маршрутизации от источника (Source routing) – концепция сервера маршрутов (Route Server, RS) [11 – 12], во-вторых, концепция предвычисления путей (Precomputed routing, PR) [11, 13 – 15]. При решении задач внутренней (в пределах одного домена или сети одного провайдера) маршрутизации функции сервера маршрутов в зависимости от сетевой архитектуры возлагаются на различные устройства: элемент вычисления путей PCE (Path Computation Element) в домене MPLS [16 – 19]; брокер пропускной способности BB (Bandwidth Broker) в DiffServ-доме IP-сети [11, 20 – 22] или иное специализированное устройство, например, сервер маршрутизации Routing and Traffic Engineering Server (RATES) для сетей MPLS [12, 23] или управляющая маршрутами BGP платформа Routing Control Platform (RCP) [24]. Наличие единого центра управления в сети (доме) позволяет не только решить задачу маршрутизации оптимально с точки зрения всей сети и ее текущего состояния, но и обеспечивает решение задач управления доступом и резервирования ресурсов, что немаловажно для обеспечения гарантированного QoS в целом. Масштабируемость в условиях централизованного принятия решения сервером RS обеспечивается за счет предвычисления путей и для MPLS – за счет предварительного установления пути коммутации меток LSP [11].

Много рациональных идей было заложено в протокол маршрутизации PNNI (Private Network-to-Network Interface), используемый в прежде достаточно перспективной технологии ATM (Asynchronous Transfer Mode). Данный протокол реализует стратегию маршрутизации «от источника», поддерживая композитную метрику по ключевым показателям качества обслуживания, к которым, прежде всего, относятся доступная скорость передачи ячеек (Available Cell Rate, AvCR); максимальная задержка передачи ячеек (Maximum Cell Transfer Delay, MaxCTD); процент потерь (Cell Loss Ratio, CLR) – среднее количество потерянных во время передачи ячеек; разброс задержки (Cell Delay Variation, CDV); максимальная скорость передачи (Maximum Cell Rate, MaxCR) [25]. Однако, наряду с традиционными достоинствами с точки зрения поддержки функциональности качества обслуживания, данный маршрутизирующий протокол является вдобавок и сигнальным протоколом, отвечающим за установление соединения и обеспечение QoS-гарантий на основе резервирования ресурсов. Интеграция в рамках одного протокола функций маршрутизации и сигнализации заметно повышает его возможности с точки зрения обеспечения качества обслуживания. Напомним, что в IP-сетях задачи резервирования ресурсов несколько отделены от протоколов маршрутизации и возложены на специальный протокол RSVP (Resource ReSerVation Protocol).

В области внешней (межсетевой или междоменной) маршрутизации можно

выделить два направления развития. Первое связано с использованием протокола пограничного шлюза BGP и его QoS-расширения [8, 26]. Как известно, BGP является протоколом маршрутизации дистанционно-векторного типа, в котором решение о направлении продвижения пакета принимается каждым узлом в отдельности (подход hop-by-hop) [8 – 9]. Отсутствие на маршрутизаторе информации о текущем состоянии всей сети не позволяет в условиях применения BGP получить в конечном итоге маршрут, оптимальный с точки зрения всей сети.

Другой подход к межсетевой маршрутизации заключается в использовании маршрутизации от источника [12, 23, 27]. Хотя маршрутизация от источника имеет свои сложности, связанные, например, с необходимостью сбора подробной информации о топологии сети, именно она потенциально способна обеспечить расчет такого маршрута, вдоль которого QoS-требования гарантировано выполняются. Масштабируемость в этом случае обеспечивается за счет иерархического представления топологической информации [27].

Таким образом, хотя маршрутизация потенциально является одним из наиболее действенных механизмов в плане предоставления качества обслуживания, на практике в рамках существующих протоколов маршрутизации подобные функции остаются нереализованными в полной мере. В связи с переходом к сетям NGN вопросы совершенствования алгоритмов и протоколов являются особо актуальными. В целом на основании выявленных недостатков существующих протоколов и с учетом перспективных концепций можно сформулировать требования к протоколу маршрутизации в современных ТКС. Во-первых, протокол маршрутизации должен отвечать концепции QoS-маршрутизации, т.е. оперировать не только номинальными, но и доступными сетевыми ресурсами, информация о которых должна обеспечиваться путем их постоянного мониторинга, а также обеспечивать маршрут не просто минимальной стоимости, а такой, вдоль которого гарантированно выполняются требования приложений к качеству их обслуживания как по показателям сетевой производительности, так и (в идеале) по показателям воспринимаемого качества обслуживания – QoE (Quality of experience). Во-вторых, протокол маршрутизации должен обеспечивать сбалансированное использование сетевых ресурсов (Load-Balancing Routing), что требует перехода от однопутевых к многопутевым стратегиям. В-третьих, задача маршрутизации должна решаться согласованно (в комплексе) с задачами управления доступом и резервирования ресурсов [28 – 29].

Все перечисленное в совокупности нацеливает на реализацию в рамках перспективных протоколов маршрутизации концепции сервера маршрутов в сочетании с концепцией предвычисления путей и иерархическим представлением топологической информации как средств обеспечения масштабируемости, что в целом требует пересмотра теоретических решений, закладываемых в их основу, и ориентирует на поиск новых, более конструктивных методов управления трафиком в современных мультисервисных ТКС.

II. Обзор теоретических решений в области QoS-маршрутизации

Анализ комбинаторных решений задач QoS-маршрутизации

Представление прикладной задачи как комбинаторной задачи и решение ее путем направленного перебора является одним из распространенных подходов в процессе анализа и синтеза ТКС. Так, как уже было отмечено, основу существующих протоколов маршрутизации составляют различные алгоритмы кратчайшего пути в графе. Прежде всего, это алгоритмы Дейкстры и Беллмана-Форда, которые обеспечивают нахождение дерева кратчайших в выбранной метрике путей от узла-источника ко всем остальным узлам, и вычислительная сложность которых приемлема для реализации в реальном масштабе времени [3, 7, 30 – 31].

С появлением концепции NGN и смещением акцентов при оценке алгоритмов управления трафиком на их возможности по поддержке функций QoS графо-комбинаторные модели и методы были существенно пересмотрены. В целом описанные в литературе подходы могут быть разделены на два класса: модели и методы QoS-маршрутизации и модели и методы k-путевой маршрутизации, которые, за исключением единичных предложений, почти не пересекаются [32 – 33]

Концепция QoS-маршрутизации требует определения такого пути (путей), между заданной парой узлов-адресатов, вдоль которого будут выполняться требования одновременно по нескольким QoS-показателям (метрикам) [32, 34]. (Здесь и в дальнейшем будет рассматриваться только задача одноадресной доставки информации.) В рамках подобной маршрутизации требования к QoS-показателям вдоль пути выступают в качестве ограничений на этапе его поиска, задача которого в общем случае может быть сформулирована двояко: как задача поиска пути с ограничениями (Multi-Constrained Path, MCP) или как задача поиска оптимального пути с ограничениями (Multi-Constrained Optimal Path, MCOP) [32, 35]. И в том и в другом случае использование одновременно двух и более метрик независимо от их типа (аддитивных, мультипликативных) ведет к тому, что задача нахождения пути становится NP-полной и может быть точно решена путем полного перебора [34 – 35]. В этой связи в литературе предложено множество эвристических алгоритмов, обладающих полиномиальным временем решения [36 – 42].

В целом на основании описанных в литературе моделей и методов в направлении снижения вычислительной сложности задач MCP и MCOP можно указать три основных подхода [34, 35, 43]:

1. Переход от множества простых, одиночных метрик, каждая из которых соответствует одному QoS показателю, к одной композитной (комбинированной), представляющей собой некоторую функцию отдельных метрик.
2. Поочередное использование метрик, т.е. поиск всех путей, удовлетворяющих требованиям в первой метрике, затем поиск среди найденных, но уже с использованием второй метрики и т.д.

3. Замена всего множества метрик одной, той, которая лежит в основе каждой из отдельных метрик и определяет их значения.

Так, в рамках первого подхода вместо r аддитивных метрик в качестве комбинированной стоимостной функции пути p предлагается использовать следующие:

$$l(p) = \sum_{i=1}^r d_i w_i(p) \quad [44],$$

$$l(p) = \sum_{i=1}^r \max(w_i(p), L_i) \quad [44],$$

$$l(p) = \left[\sum_{i=1}^r \left(\frac{w_i(p)}{L_i} \right)^q \right]^{1/q}, \quad q > 1 \quad [45],$$

$$l(p) = \max_{1 \leq i \leq r} \left(\frac{w_i(p)}{L_i} \right) \quad [46 - 47],$$

где d_i – весовые коэффициенты различных метрик; L_i – допустимое значение метрики w_i .

Наличие единственной метрики позволяет применить известные алгоритмы поиска кратчайшего пути и их всевозможные расширения [3, 34, 35, 48]. В целом задача конструирования метрик как одиночных, так и комбинированных представляет собой отдельный предмет исследований [49 – 50].

Примером второго подхода является описанный в работе [51] алгоритм решения задачи поиска пути, удовлетворяющего требованиям по пропускной способности и по задержке. Алгоритм предусматривает два этапа: на первом этапе из графа удаляются все ветви, пропускные способности которых не отвечают выдвинутым требованиям, а на втором этапе в графе, полученном после удаления части ветвей, при помощи алгоритма Дijkstra ищется путь с минимальной задержкой. Другим примером является алгоритм Delay-Cost-Constrained Routing (DCCR) [35, 52], в котором на первом этапе в исходном графе генерируется множество путей, удовлетворяющих требованиям по задержке, а на втором этапе среди этого множества, начиная с пути с минимальной задержкой, выбирается один, удовлетворяющий еще и требованию по стоимости. К этой же группе относятся часто встречающиеся в литературе алгоритмы Widest Shortest Path и Shortest Widest Path [51, 53 – 54], где первый обеспечивает поиск пути с наибольшей пропускной способностью среди множества путей с минимальной стоимостью, а второй – поиск пути с наименьшей стоимостью среди множества путей с максимальной пропускной способностью.

Третий подход отражен, например, в работе [55], где показано, что для случая использования механизма обслуживания очередей WFQ такие метрики как скорость передачи, межконцевая задержка, джиттер и используемый объем буферной памяти не являются независимыми. Все они представляют собой функции от зарезервированной пропускной способности, выбранного пути и характеристик трафика, а потому исходная задача нахождения пути с множеством ограничений на основании установленной взаимосвязи может быть решена при помощи модифицированного алгоритма Беллмана-Форда.

В работах [34, 48, 56] показано, что задача МСР может быть решена за полиномиальное время, если $(r-1)$ из r метрик будут представлены целыми, ограниченными сверху числами, для чего предлагается сначала аппроксимировать метрики, являющихся на самом деле действительными числами, а затем применить расширенные алгоритмы Дейкстры и Беллмана-Форда (Extended Dijkstra shortest path, EDSP, Extended Bellman-Ford, EBF).

Основные графокомбинаторные алгоритмы, решающие задачи МСР и МСОР в контексте QoS-маршрутизации, приведены в табл. 1 [32 – 35, 43, 57 – 59]. Перечисленные алгоритмы используют различные метрики, постановки задачи, ориентируются на различные способы реализации (распределенный, hop-by-hop, например, DCUR, TBR, или от источника, by source, например H_MCOF, LARAC), однако практически все они, за исключением алгоритма A*Prune [59], являются однопутевыми.

С другой стороны, представленные в литературе, например [60 – 63], алгоритмы k -путевой маршрутизации обеспечивают поиск k наилучших только в одной единственной метрике; учет требований в других метриках не производится. В этом проявляется один из главных недостатков графокомбинаторного подхода – его ограниченность в согласованном решении задач сбалансированной загрузки сети и выполнения QoS требований. Кроме того графовые модели по своей сути ориентированы на однопродуктовые двухполюсные сети и их расширение на случай множества продуктов и пар адресатов (многопродуктовые многополюсные сети) приводит к сложностям как описательного, так и вычислительного характера. В рамках подобных моделей остается нерешенным один из важнейших для предоставления услуг QoS вопрос распределения ресурсов вдоль найденного пути, что является необходимым условием для достижения согласованного решения задач маршрутизации, резервирования ресурсов и управления доступом. Успешное их решение невозможно без учета наряду со структурными характеристиками сети параметров информационного трафика [64].

Таким образом, несмотря на использование моделей и методов кратчайшего пути в рамках существующих протоколов управления трафиком, данный подход в свете перехода к мультисервисным сетям NGN с предоставлением широкого спектра услуг гарантированного качества не является перспективным и может рассматриваться как промежуточный этап на пути перехода к более сложным, но более адекватным потоковым моделям ТКС.

Таблица 1. Графокомбинаторные алгоритмы однопутевой маршрутизации

Алгоритм (название первоисточника)	Тип ограничений	Временная сложность
<i>DCUR (Delay Constrained Unicast Routing)</i>	Ограничения на задержку, оптимизация стоимости	$O(m^3)$
<i>LARAC (Lagrange Relaxation based Aggregated Cost)</i>	Ограничения на задержку, оптимизация стоимости	$O(n^2 \log^4 n)$
<i>DCCR (Delay-Cost-Constrained Routing)</i>	Ограничения на задержку, оптимизация стоимости	$O(kn \log(km) + k^2 m)$
<i>SSR+ DCCR (Search Space Reduction + DCCR)</i>	Ограничения на задержку, оптимизация стоимости	$O(xn \log(m) + kn \log(km) + k^2 m)$
<i>Wang-Crowcroft's algorithm</i>	Ограничения на ПС и задержку	$O(m \log(m) + n)$
<i>Ma-Steenkiste's algorithm</i>	Ограничения на ПС	$O(m \log(m) + n)$
	Ограничения на ПС и задержку	$O(kmn)$
<i>Guerin-Orda's algorithm</i>	Ограничения на ПС	$O(m \log(m) + n)$
	Ограничения на задержку	Полиномиальная
<i>Chen-Nahrstedt's algorithm</i>	Ограничения на ПС и стоимость	$O(xmn)$
<i>Wang-Crowcroft's algorithm</i>	Оптимизация ПС	$O(mn)$
<i>H_MCOP (Heuristic algorithm for Multi-Constrained Optimal Path)</i>	Ограничения на r аддитивных метрик, оптимизация стоимости	$O(m \log(m) + rn)$
<i>Sun-Landgendorfer's algorithm</i>	Ограничения на задержку, оптимизация стоимости	$O(m)$
<i>Distributed Jaffe's algorithm</i>	Ограничения на 2 аддитивные метрики	$O(m^5 b \log(bm))$
<i>Shin-Chou's algorithm</i>	Ограничения на задержку	$O(n)$
<i>TAMCRA (Tunable Accuracy Multiple Constraints Algorithm)</i>	Ограничения на r аддитивных метрик	$O(km \log(km) + k^3 rn)$
<i>SAMCRA (Self-Adaptive Multiple Constraints Algorithm)</i>	Ограничения на r аддитивных метрик	$O(km \log(km) + k^2 rn)$
<i>Randomized Algorithm</i>	Ограничения на r метрик, оптимизация числа переприемов	$O(rm \log(m) + rn)$
<i>EBF (Extended Bellman-Ford)</i>	Ограничения на задержку и стоимость	$O(xmn)$
<i>EDSP (Extended Dijkstra shortest path)</i>	Ограничения на время и стоимость	$O(x^2 m^2)$
<i>DSA (Delay Scaling Algorithm)</i>	Ограничения на задержку, оптимизация стоимости	$O((n + m \log(m))x/\epsilon)$
<i>A*Prune</i>	Ограничения на r аддитивных метрик, оптимизация стоимости	$O(xq(r+h+\log q))$
<i>TBR (Ticket Based Probing)</i>	Ограничения на задержку или ПС, оптимизация стоимости	$O(n)$
<i>LPH (Limited Path Heuristic)</i>	Ограничения на r метрик	$O(x^2 mn)$
<i>LCDC (Least-cost Least Delay)</i>	Оптимизация стоимости и задержки	$O(m^2)$
Примечание: ПС – пропускная способность; m – число узлов сети; n – число трактов передачи в сети; k, x – константы алгоритмов; b – максимальный вес дуги в сети; h – максимальное число переприемов, q – число сохраняемых маршрутов.		

Анализ потоковых решений задач QoS-маршрутизации

В рамках потоковых моделей задача управления трафиком, в т.ч. маршрутизации, зачастую формулируется в виде задачи математического программирования: линейного [3, 4, 65 – 68], нелинейного [3, 69 – 72], целочисленного [66, 73], смешанного [74] с использованием терминологии распределения потока на графах. Обязательными компонентами такой постановки задачи являются целевая функция, которая в зависимости от вкладываемого физического смысла подлежит минимизации или максимизации; условия сохранения потока; условия, связанные с ограниченностью сетевых ресурсов, прежде всего пропускной способности трактов передачи (условия отсутствия перегрузки) и условия неотрицательности потока. В зависимости от уровня рассмотрения потока и физического смысла, вкладываемого в используемые переменные, каждое из перечисленных условий конкретизируется. Например, в работах [65, 75 – 77] они принимают следующий вид:

условия сохранения потока

$$\begin{cases} \sum_{j:(i,j) \in \mathcal{E}} x_{ij}^k - \sum_{j:(j,i) \in \mathcal{E}} x_{ji}^k = 0, & k \in \mathcal{K}, i \neq s_k, t_k, \\ \sum_{j:(i,j) \in \mathcal{E}} x_{ij}^k - \sum_{j:(j,i) \in \mathcal{E}} x_{ji}^k = 1, & k \in \mathcal{K}, i = s_k, \\ \sum_{j:(i,j) \in \mathcal{E}} x_{ij}^k - \sum_{j:(j,i) \in \mathcal{E}} x_{ji}^k = -1, & k \in \mathcal{K}, i = t_k; \end{cases} \quad (3)$$

условия неотрицательности потока

$$0 \leq x_{ij}^k \leq 1; \quad (4)$$

условия отсутствия перегрузки

$$\sum_{k \in \mathcal{K}} d_k x_{ij}^k \leq c_{ij} \alpha, \quad 0 \leq \alpha \leq 1, (i, j) \in \mathcal{E}, \quad (5)$$

где управляющей переменной служит величина x_{ij}^k , которая характеризует долю k -го трафика, протекающего в тракте $(i, j) \in \mathcal{E}$; \mathcal{K} – множество трафиков в сети; \mathcal{E} – множество трактов передачи сети; s_k, t_k – соответственно узел-источник и узел-получатель k -го трафика; d_k – интенсивность k -го трафика; c_{ij} – пропускная способность тракта передачи (i, j) ; α – динамически управляемый порог максимальной загрузки каналов ТКС.

Принципиальные отличия в постановку задачи управления и соответственно получаемый результат вносит используемый критерий оптимальности, что определяет тип задачи математического программирования (линейное или нелинейное) и, следовательно, методы ее решения.

Наиболее распространенными в классе задач линейного программирования являются следующие критерии оптимальности:

1) минимум стоимостной функции [3, 67 – 68]

$$\sum_{k \in \mathcal{K}} \sum_{(i,j) \in \mathcal{E}} w_{ij} x_{ij}^k \rightarrow \mathbf{min} \quad \text{или} \quad \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}_k} w_p x_p^k \rightarrow \mathbf{min}, \quad (6)$$

где w_{ij} , w_p – стоимость передачи единицы потока по тракту передачи $(i, j) \in \mathcal{E}$ и по пути $p \in \mathcal{P}_k$ соответственно; \mathcal{P}_k – множество путей передачи k -го трафика.

2) минимум вероятности отказа в обслуживании для вновь поступающего трафика или максимум суммарного обслуженного потока [66, 68, 78]:

$$\sum_{k \in \mathcal{K}} d_k \rightarrow \mathbf{max} \quad \text{или} \quad \sum_{k \in \mathcal{K}} d_k b_k \rightarrow \mathbf{max}, \quad (7)$$

где b_k – прибыль, получаемая от обслуживания потока k -трафика;

3) минимум максимального по сети коэффициента использования каналов [4, 65, 79 – 80] (или минимум фактора перегрузки сети [81])

$$\mathbf{min}_{x_{ij}^k} \alpha, \quad (8)$$

где $\alpha = \mathbf{max}_{(i,j) \in \mathcal{E}} (K_{ij}^{ucn})$, $K_{ij}^{ucn} = \frac{x_{ij}^k}{c_{ij}} = \frac{1}{c_{ij}} \sum_{k \in \mathcal{K}} x_{ij}^k$ – коэффициент использования тракта передачи (i, j) .

Относительно перечисленных критериев сделаем ряд замечаний. Использование стоимостного критерия (6) в рамках задачи целочисленного программирования ($x_{ij}^k \in \{0, 1\}$) в однопродуктовом случае представляет собой формализацию задачи кратчайшего пути. Зачастую использование критерия (6) сопровождается введением дополнительного ограничения, связанного с требованием сбалансированного использования сетевых ресурсов [80]

$$K_{ij}^{ucn} \leq \beta, \quad (9)$$

где β – максимальный допустимый порог для коэффициента использования канальных ресурсов сети.

Критерий (7) согласно теореме Форда-Фалкерсона о максимальном потоке и минимальном разрезе [82] позволяет выявить наиболее нагруженные участки сети и зачастую используется именно с этой целью на промежуточных этапах эвристических алгоритмов маршрутизации, например, алгоритмов MIRA (Minimum Interference Routing Algorithm) и LMIR (Ligth Minimum Interference Routing) [68, 73, 83]. Критерий (8) нацеливает на сбалансированное использование сетевых ресурсов, отвечая требованиям концепции ТЕ, и является наиболее распространенным при рассмотрении задач многопутевой маршрутизации в рамках потоковых моделей [75 – 77]. А обеспечение сбалансированной загрузки сети, как известно, является ключевым фактором повышения качества обслуживания в ТКС.

С целью реализации дополнительных возможностей по дальнейшему повышению качества обслуживания в работе [84] предлагается развитие линейной пото-

ковой модели в направлении совместного решения задачи маршрутизации и управления доступом (профилирования входящего трафика). Для этого вектор управляющих переменных расширяется путем введения величины φ^k , которая моделирует интенсивность k -го трафика, получившего отказ в обслуживании сетью:

$$\vec{x} = \begin{bmatrix} x_{ij}^k \\ \vdots \\ \varphi^k \end{bmatrix}, (i, j) \in \mathcal{E}, k \in \mathcal{K}. \quad (10)$$

В ходе расчета вектора (10) в качестве критерия оптимальности предлагается использовать минимум линейной или линейно-квадратичной целевой функции соответственно:

$$\sum_{(i,j) \in \mathcal{E}} \left(\sum_{k \in \mathcal{K}} w_{ij}^k x_{ij}^k + \sum_{k \in \mathcal{K}} w^k \varphi^k \right) \rightarrow \mathbf{min} \quad (11) \quad \text{или} \quad \left[\frac{1}{2} \vec{x}^t H \vec{x} + W^t \vec{x} \right] \rightarrow \mathbf{min}, \quad (12)$$

которые характеризуют относительные затраты на управление трафиком на этапе доступа к сети и внутри ТКС, где H и W – стоимостные диагональная матрица и вектор соответственно.

При переходе к нелинейным критериальным функциям представляется возможным более полный и сложный учет требований к качеству обслуживания, а также других, например, экономических факторов [83]. Как показал анализ, наиболее распространенными среди нелинейных целевых функций являются следующие [69, 72, 83]:

$$\sum_{(i,j) \in \mathcal{E}} x_{ij} f(x_{ij}) \rightarrow \mathbf{min}, \quad \sum_{k \in \mathcal{K}} x_{ij}^k = x_{ij}, \quad (13)$$

$$\sum_{k \in \mathcal{K}} f_{utility}^k \left(\sum_{p \in \mathcal{P}_k} x_p^k \right) \rightarrow \mathbf{max}, \quad (14)$$

$$\sum_{(i,j) \in \mathcal{E}} f_{penalty}^k (K_{ij}^{ucn}) \rightarrow \mathbf{min}, \quad (15)$$

где $\sum_{k \in \mathcal{K}} f_{utility}^k \left(\sum_{p \in \mathcal{P}_k} x_p^k \right)$ – функция полезности (возможно прибыли) для k -го обслуживаемого трафика; $f_{penalty}^k (K_{ij}^{ucn})$ – функция штрафов за использование тракта (i, j) .

Критерии (13) – (15) представлены в общем виде, входящие в их состав функции f , $f_{utility}^k$, $f_{penalty}^k$ в зависимости от постановки задачи принимают тот или иной вид. Например, функция полезности (14) при управлении ТСП-трафиком в зависимости от версии ТСП формализуются как [77]:

$$f_{utility}(x) = \mathbf{log}(x) \quad (TCP Vegas),$$

$$f_{utility}(x) = \mathbf{arctan}(x) \quad (TCP Tahoe),$$

$$f_{utility}(x) = (1 - \delta)^{-1} x^{1-\delta}, \quad \delta = 2 \text{ (TCP Reno)}.$$

Целевая функция в рамках известной модели Галлагера [30, 86] является новизной (13) при

$$f(x_{ij}) = \frac{1}{c_{ij} - x_{ij}} + \tau_{ij}, \quad (16)$$

где τ_{ij} – постоянная задержка, связанная с обработкой и распространением в тракте передачи (i, j) .

Выражение (16) получено при рассмотрении ТКС как сети массового обслуживания, где каждый тракт передачи описывается моделью $M/M/1$. При этом предполагается пуассоновский характер поступающего в сеть трафика; независимость длин пакетов и их распределение по показательному закону; сохранение независимости между интервалами поступления и длинами пакетов при объединении нескольких потоков; абсолютная надежность и помехоустойчивость всех трактов передачи сети; неограниченность буферной емкости на узлах сети [30, 86]. Тогда с учетом (16) формализованная в виде (13) задача при наличии соответствующих ограничений (3) – (5) представляет задачу управления трафиком, минимизирующую среднее время задержки в сети. Основанный на использовании частных производных целевой функции алгоритм решения этой задачи (алгоритм Галлагера), как показано в работе [63], медленно сходится и может быть применен для стационарного и квазистационарного трафика.

В ряде работ модель Галлагера получила развитие в различных направлениях: в [87] с целью повышения сходимости алгоритма предлагается использовать вторые производные целевой функции, в работе [88] модель адаптирована под требования инжиниринга трафика (Traffic Engineering), в [89] модель развита на случай k -путевой маршрутизации, в [90] с целью устранения зависимости от глобальных констант состояния сети и требований к статичности передаваемого трафика алгоритм Галлагера рассматривается как один из этапов в рамках эвристического, комбинированного алгоритма в сочетании с комбинаторными методами кратчайших путей.

В целом описанные выше потоковые модели ТКС в отличие от комбинаторных методов, во-первых, изначально ориентированы на решение многопродуктовых и/или многополюсных задач, во-вторых, предполагают одновременный расчет множества путей в сети и обеспечивают сбалансированное распределение трафика по этому множеству. Получаемые с их помощью решения способствуют росту сетевой производительности, а также косвенно повышают качество обслуживания пользовательских трафиков. Однако представленные модели ввиду отсутствия в их рамках показателей QoS отдельных потоков (индивидуальных или агрегированных) повышают качество обслуживания в среднем, например, в рамках модели Галлагера, и не позволяют достичь распределения трафика, гарантирующего требуемое QoS. Представленные в [91 – 92] расширения потоковых моделей за счет введения в явном виде ограничений на QoS-показатели обслуживаемых трафиков носят статический характер и не охватывают случай динамического управления трафиком, что требует пере-

хода к более информативным и одновременно более сложным динамическим потоковым моделям.

Для математического описания динамических свойств ТКС могут использоваться различные подходы, среди которых основное место занимают методы, описывающие процесс функционирования ТКС как управляемый случайный марковский процесс и представляемые в пространстве переменных состояния, как правило, в виде линейных или реже нелинейных дифференциально-разностных уравнений, обобщенных в векторно-матричной форме [64].

Например, динамику информационного обмена в ТКС можно представить в виде следующей системы неавтономных уравнений загрузки буферов очередей на маршрутизаторах ТКС [64, 93]:

$$x_{i,j}(k+1) = x_{i,j}(k) - \sum_{\substack{l=1, \\ l \neq i}}^M b_{i,l}(k) \cdot u_{i,l}^j(k) + \sum_{\substack{m=1, \\ m \neq i,j}}^M b_{m,i}(k) \cdot u_{m,i}^j(k) + y_{i,j}(k), \quad (17)$$

где $b_{m,i}(k) = c_{m,i}(k) \cdot \Delta t$; $y_{i,j}(k) = \zeta_{i,j}(k) \cdot \Delta t$, ($k=0,1,2,\dots$; $\Delta t = t_{k+1} - t_k$); $x_{i,j}(k)$ – объем данных, находящихся на маршрутизаторе V_i и предназначенных для передачи маршрутизатору V_j в момент времени t_k , трактуемый в дальнейшем как переменная состояния; $i = \overline{1, M}$; $c_{i,j}(k)$ – пропускная способность тракта передачи (i, j) в момент времени t_k ; $u_{i,l}^j(k)$ – доля пропускной способности тракта передачи (i, j) , выделенная пользователю трафику с адресатом V_j в момент времени t_k и трактуемая в дальнейшем как управляющая переменная, характеризующая распределение канальных ресурсов; $\zeta_{i,j}(k)$ – средняя интенсивность поступления данных на маршрутизатор V_i в момент времени t_k с адресатом V_j от пользователей сети; Δt – период (таймер) перерасчета управляющих переменных.

На переменные состояния, а также переменные управления сетевыми ресурсами с целью предотвращения перегрузки элементов ТКС должен быть наложен ряд ограничений [64]:

$$0 \leq x_{i,j}(k) \leq x_{i,j}^{\max}, \quad 0 \leq u_{i,l}^j(k) \leq 1, \quad \sum_{n=1}^N u_{i,l}^n(k) \leq 1, \quad (18)$$

где $x_{i,j}^{\max}$ – максимально допустимая длина очереди на i -м узле для трафика с адресатом V_j .

В работе [93] с целью сбалансированного использования сетевых ресурсов (как канальных, так и буферных) предлагается модифицировать ограничения (18), которые принимают вид

$$0 \leq x_{i,j}(k) \leq \alpha_1(k) \cdot x_{i,j}^{\max}, \quad (19)$$

$$0 \leq u_{i,l}^j(k); \quad \sum_{n=1}^M u_{i,l}^n(k) \leq \alpha_2(k) \leq 1, \quad (20)$$

где $\alpha_1(k)$ – верхний динамически управляемый предел использования буферного пространства маршрутизаторов ТКС; $\alpha_2(k)$ – верхний предел использования пропускной способности трактов передачи телекоммуникационной сети, то есть $\alpha_1(k)$ и $\alpha_2(k)$ являются управляемыми переменными наряду с $u_{i,l}^j(k)$.

Сама же задача маршрутизации, рассматриваемая как слагаемое общей проблемы управления трафиком, формулируется как задача оптимальной балансировки канальных и буферных ресурсов с использованием следующего критерия:

$$\sum_{k=1}^a (r_\alpha(k)\alpha_1(k) + r_\beta(k)\alpha_2(k)) \rightarrow \min, \quad (21)$$

где r_α и r_β – весовые коэффициенты, которые характеризуют условную стоимость балансировки по буферным и канальным ресурсам сети; a – интервал рассмотрения динамики системы (интервал прогнозирования).

Заметим, что в общем случае в рамках динамической модели ТКС (17) – (18) в качестве критерия оптимальной маршрутизации может выступать также аналог любого из выражений (6) – (8), (13) – (15).

Таким образом, подход, основанный на представлении ТКС в пространстве состояний, обеспечивает динамическое управление трафиком (динамическую маршрутизацию), учет доступных сетевых ресурсов (как канальных, так и буферных) и сбалансированное их использование. Однако сохраняет ранее отмеченный недостаток: ввиду отсутствия в модели (17) – (21) в явном виде показателей качества обслуживания (скорости передачи, средних задержек, уровня потерь в трактах передачи сети) данный подход не позволяет реализовать идею QoS-маршрутизации в полной мере. С целью обеспечения гарантированного качества обслуживания математическая постановка задачи маршрутизации должна содержать QoS-показатели обслуживаемых потоков трафика, которые могут быть введены в целевую функцию или выступать в роли дополнительных ограничений в рамках какой-либо потоковой модели. В первом случае речь идет о достижении наилучшего качества обслуживания (по какому-либо одному отдельно взятому или комплексному показателю), во втором – о выполнении заданных QoS-требований согласно договору об обслуживании SLA. Второй подход представляется более целесообразным с точки зрения эффективности функционирования сети в целом и может быть осуществлен, например, за счет введения в потоковую модель ТКС тензорных ограничений следующего вида [94 – 97]:

$$\lambda^{\langle mp\bar{b} \rangle} \leq \left(G_{\pi\eta}^{\langle 4,1 \rangle} - G_{\pi\eta}^{\langle 4,2 \rangle} \left[G_{\pi\eta}^{\langle 4,4 \rangle} \right]^{-1} G_{\pi\eta}^{\langle 4,3 \rangle} \right) \tau_{\langle mp\bar{b} \rangle}, \quad (22)$$

$$\lambda^{\langle mp\bar{b} \rangle} \geq \left(X_{\pi\eta}^{\langle 4,1 \rangle} - X_{\pi\eta}^{\langle 4,2 \rangle} \left[X_{\pi\eta}^{\langle 4,4 \rangle} \right]^{-1} X_{\pi\eta}^{\langle 4,3 \rangle} \right) p_{\langle mp\bar{b} \rangle}, \quad (23)$$

$$\text{где } \left\| \begin{array}{c|c} G_{\pi\eta}^{(4,1)} & G_{\pi\eta}^{(4,2)} \\ \hline \text{---} & \text{---} \\ \hline G_{\pi\eta}^{(4,3)} & G_{\pi\eta}^{(4,4)} \end{array} \right\| = G_{\pi\eta}^{(4)}, \quad \left\| \begin{array}{c|c} G_{\pi\eta}^{(1)} & G_{\pi\eta}^{(2)} \\ \hline \text{---} & \text{---} \\ \hline G_{\pi\eta}^{(3)} & G_{\pi\eta}^{(4)} \end{array} \right\| = G_{\pi\eta}, \quad \left\| \begin{array}{c|c} X_{\pi\eta}^{(4,1)} & X_{\pi\eta}^{(4,2)} \\ \hline \text{---} & \text{---} \\ \hline X_{\pi\eta}^{(4,3)} & X_{\pi\eta}^{(4,4)} \end{array} \right\| = X_{\pi\eta}^{(4)},$$

$$\left\| \begin{array}{c|c} X_{\pi\eta}^{(1)} & X_{\pi\eta}^{(2)} \\ \hline \text{---} & \text{---} \\ \hline X_{\pi\eta}^{(3)} & X_{\pi\eta}^{(4)} \end{array} \right\| = X_{\pi\eta}, \text{ причем } G_{\pi\eta}^{(4,1)} \text{ и } X_{\pi\eta}^{(4,1)} - \text{первые элементы матриц } G_{\pi\eta}^{(4)} \text{ и } X_{\pi\eta}^{(4)}$$

соответственно; $G_{\pi\eta}^{(4)}$ и $X_{\pi\eta}^{(4)}$ – квадратные подматрицы размера $\phi \times \phi$ матриц $G_{\pi\eta}$ и $X_{\pi\eta}$ соответственно, $\phi = M - 1$; $G_{\pi\eta}$ и $X_{\pi\eta}$ – проекции метрических тензоров \mathbf{G} и \mathbf{X} временной и вероятностной тензорной моделей ТКС соответственно в системе координат контуров и узловых пар; $\lambda_{\langle mp\bar{b} \rangle}^{(mp\bar{b})}$, $\tau_{\langle mp\bar{b} \rangle}$ – требуемые значения скоростных и временных показателей качества обслуживания: требуемая интенсивность трафика и допустимая средняя задержка (IPTD), $p_{\langle mp\bar{b} \rangle} = \log_a(1 - P_{\text{пот}}^{\langle mp\bar{b} \rangle})$, $P_{\text{пот}}^{\langle mp\bar{b} \rangle}$ – допустимая величина потерь при передаче трафика между заданной парой адресатов (IPLR).

Используемые в (22) и (23) проекции тензоров \mathbf{G} и \mathbf{X} в пространстве контуров и узловых пар $G_{\pi\eta}$ и $X_{\pi\eta}$ могут быть получены путем координатного преобразования при переходе в данную систему координат из более простой системы координат ветвей, представляющую собой совокупность отдельно взятых трактов передачи сети:

$$X_{\pi\eta} = A^t X_v A, \quad (24)$$

$$G_{\pi\eta} = A^t G_v A, \quad (25)$$

где A – матрица контравариантного координатного преобразования при переходе от системы координат контуров и узловых пар к системе координат ветвей (определяется структурными свойствами сети); G_v и X_v – проекции метрических тензоров \mathbf{G} и \mathbf{X} системе координат ветвей данной сети, представляющие собой диагональные $n \times n$ матрицы, n – количество трактов передачи.

Выражение для расчета элементов главной диагонали матриц G_v и X_v зависит от принятой модели трафика и процессов его обслуживания. Например, при рассмотрении тракта передачи как системы массового обслуживания $M/M/1$ [94–95] получаем

$$g_{ii}^{(v)} = \lambda_{(v)}^i (c_i - \lambda_{(v)}^i), \quad i = \overline{1, n}, \quad (26)$$

где $\lambda_{(v)}^i$ – интенсивность трафика, протекающего в i -м тракте передачи сети; c_i – пакетная пропускная способность i -го тракта передачи.

В случае реализации на интерфейсах сети механизма отбрасывания пакетов по переполнению (Tail Drop) имеем [96 – 97]

$$x_{ii}^{(v)} = \left[\frac{1}{\lambda_{(v)}^i} \log_a \left(1 - \frac{\left(1 - \frac{\lambda_{\Sigma(v)}^i}{c_i} \right) \left(\frac{\lambda_{\Sigma(v)}^i}{c_i} \right)^N}{1 - \left(\frac{\lambda_{\Sigma(v)}^i}{c_i} \right)^{N+1}} \right) \right]^{-1}. \quad (27)$$

Преимуществом использования тензорных ограничений вида (22) – (23) является их ориентация на многопутевой способ маршрутизации с балансировкой трафика по всем доступным к использованию маршрутам, учет потокового характера трафика, возможность выполнения заданных требований по каждому из показателей качества обслуживания *в отдельности*, что в более полной мере отвечает формуле концепции QoS-маршрутизации. Своеобразной платой за это является усложнение модели маршрутизации и процесса получения решения, что, в свою очередь, требует ряда мер по повышению их масштабируемости.

III. Выводы. Основные направления развития средств QoS-маршрутизации

На основании проведенного как качественного, так количественного анализа различных подходов к решению задач QoS-маршрутизации в качестве основных направлений их дальнейшего развития следует отметить следующие. В первую очередь, пересмотр моделей ТКС и смещение акцентов с графокомбинаторных на потоковые модели, желательно динамического характера, что позволит в полной мере реализовать преимущества динамического управления трафиком, основываясь на текущем состоянии сети, например, текущей загруженности канальных и буферных ресурсов. В свою очередь это потребует пересмотра метрик маршрутов: отказа от метрик, основанных на номинальных или статически конфигурируемых параметрах трактов передачи и маршрутизаторов сети, и перехода к динамическим метрикам, отражающим качество обслуживания, достигаемое в текущий момент времени для передаваемого через данный интерфейс трафика. Возможно, ввиду высокой дифференциации трафика в современной мультисервисной ТКС, будет целесообразным введение различных метрик, а точнее правил их формирования для различных типов трафика.

С целью предоставления сетью телекоммуникационных услуг гарантированного качества и принятия решения о маршрутизации того или иного трафика на основании запрашиваемого качества обслуживания модель ТКС должна быть расширена за счет введения ряда QoS-ограничений, выполнение которых в ходе решения оптимизационной задачи будет гарантировать выполнение заданных QoS-требований. В общем случае эти ограничения могут вводиться для каждого QoS-показателя в отдельности, а могут, как например (22) – (23), связываться в рамках единого аналитического выражения несколько показателей качества.

Достигаемое качество обслуживания пользователей, а также производительность сети в целом во многом зависят от степени сбалансированного использования сетевых ресурсов. В этом плане одним из эффективных средств повышения качества обслужи-

вания видится балансировка трафика не только с точки зрения равномерной загруженности сетевых ресурсов (канальных и буферных), а балансировка с точки зрения QoS, целью которой является равенство QoS-показателей, достигаемых вдоль различных путей передачи трафика при реализации многопутевой стратегии маршрутизации.

Еще одним немаловажным направлением развития средств QoS-маршрутизации является повышение согласованности в решении задач маршрутизации с другими задачами управления трафиком, например, управления доступом, очередями и резервирования ресурсов, путем совместного их решения в рамках единой математической модели ТКС.

Список литературы:

1. *Chao H. J., Xiaolei Guo* Quality of service control in high-speed networks. – Wiley-IEEE, 2002. – 432 p.
2. *Peng B. Kemp A.H., Boussakta S.* QoS routing with bandwidth and hop-count consideration – a performance perspective // *Journal of Communications*. – 2006. – No. 1 (2). – P. 1 – 11.
3. *Medhi Deepankar, Ramasamy Karthikeyan* Network routing: algorithms, protocols, and architectures. – Morgan Kaufmann Publishers, 2007. – 788 p.
4. *Pi6ro M., Medhi D.* Routing, Flow, and Capacity Design in Communication and Computer Networks. – Morgan Kaufmann, 2004. – 765 p.
5. *Руденко И.* Tsunami Computing, Маршрутизаторы CISCO для IP-сетей. – М.: Кудиц-Образ, 2003. – 656 с.
6. *Димарцио Д.Ф.* Маршрутизаторы Cisco. Пособие для самостоятельного изучения. – СПб.: Символ-Плюс, 2003. – 512 с.
7. *Остерлох Х.* Маршрутизация в IP-сетях. Принципы, протоколы, настройка. – С.Пб.: ВHV-С.Пб., 2002. – 512 с.
8. Руководство по технологиям объединенных сетей; [пер. с англ.]. – [4-е изд.]. – М.: Изд. дом «Вильямс», 2005. – 1040 с.
9. *Doyle J., Carroll J.* CCIE Professional Development Routing TCP/IP. – Vol. I. – Cisco Press, 2005. – 936 p.
10. *Jiayue He, Rexford J.* Toward internet-wide multipath routing // *IEEE Network*. – 2008. – Vol. 22, Issue 2. – P. 16 – 21.
11. A scalable intra-domain resource management architecture for DiffServ networks / H.A. Mantar, I.T. Okumus, J. Hwang [and other] // *Journal of High Speed Networks*. – 2006. – Vol. 15. – P. 185–205.
12. A Scalable And Efficient Inter-Domain QoS Routing Architecture For Diffserv Networks / H.A. Mantar, J. Hwang, S.J. Chapin [and other] // *IFIP/IEEE Eighth International Symposium on Integrated Network Management*, 24 – 28 March 2003: proceedings of the symposium. – P. 463 – 467.
13. *Orda A., Sprintson A.* Precomputation schemes for QoS routing // *IEEE/ACM Transactions on Networking*. – 2003. – Vol. 11, Issue 4. – P. 578 – 591.
14. *Cui Y., Xu K., Wu J.* Precomputation for multi-constrained QoS routing in high-speed networks // *IEEE INFOCOM 2003: proceedings of the conference*, Vol. 1. – P. 1305 – 1315.
15. *Orda A., Sprintson A.* QoS Routing: The Precomputation Perspective // *IEEE INFOCOM 2000: proceedings of the conference*, Vol. 3. – P. 283 – 291.

16. *Pelsser C., Bonaventure O.* Path Selection Techniques to Establish Constrained Interdomain MPLS LSPs // 5th International IFIP-TC6 Networking Conference, 15 – 19 May 2006: proceedings of the conference. – P. 209 – 220.
17. *Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks / A. Sprintson, M. Yannuzzi, A. Orda [and other]* // IEEE INFOCOM 2007, May 2007: proceedings of the conference. – P. 1820 – 1828.
18. *Dasgupta S., de Oliveira J.C., Vasseur J.-P.* Path-Computation-Element- Based Architecture for Interdomain MPLS/GMPLS Traffic Engineering: Overview and Performance // IEEE Network. – 2007. – Vol. 21, Issue 4. – P. 38 – 45.
19. Final specification of protocols and algorithms for interdomain SLS management and traffic engineering for QoS-based IP service delivery. Report IST-2001-37961 D1.3. / P. Morand, M. Boucadair, T. Coadic [and other]. – MESCAL Consortium, 2001.
20. A management and control architecture for providing IP differentiated services in MPLS-based networks / P. Trimintzios, I. Andrikopoulos, G. Pavlou [and other] // IEEE Communications Magazine. – 2001. – Vol. 39, Issue 5. – P. 80 – 88.
21. *Shaleeza Sohail, Sanjay Jha* The survey of bandwidth broker. Technical reports UNSW-CSE-0206. – University of New South Wales, School of Computer Science and Engineering, 2002.
22. A scalable model for interbandwidth broker resource reservation and provisioning / H.A. Mantar, J. Hwang, I.T.Okumus [and other] // IEEE Journal on Selected Areas in Communications. – 2004. – Vol. 22, Issue 10. – P. 2019 – 2034.
23. RATES: a server for MPLS traffic engineering / P. Aukia, M. Kodialam, P.V.N. Koppol [and other] // IEEE Network. – 2000. – Vol. 14, Issue 2. – P. 34 – 41.
24. Design and implementation of a Routing Control Platform / M. Caesar, D. Caldwell, N. Feamster [and other] // Networked Systems Design and Implementation (NSDI), May 2005: proceedings of the conference. – P. 54 – 67.
25. *Дикер Пулдуш Г.* Сети АТМ корпорации Cisco. – М.:Изд-во «Вильямс», 2004. –880 с.
26. QoS Extension to BGP / L. Xiao, K.-S.Lui, J.Wang [and other] // 10th IEEE International Conference on Network Protocols, 2002: proceedings of the conference. – P. 100 – 109.
27. Inter-Domain QoS Routing on Diffserv Networks: A Region Based Approach / I.T.Okumus, H.A. Mantar, J. Hwang [and other] // Journal of Computer Communication. – 2005. – Vol. 28. – P 174 – 188.
28. *Key P., Massoulié L., Towsley D.* Combined Multipath Routing and Congestion Control: a Robust Internet Architecture. TechReport MSR-TR-2005-111. – Microsoft Research, 2005.
29. *Key P., Massoulié L., Towsley D.* Path Selection and Multipath Congestion Control // IEEE INFOCOM 2007: International Conference on Computer Communications, 6 – 12 May 2007: proceedings of the conference. – P. 143 – 151.
30. *Бертсекас Д., Галлагер Р.* Сети передачи данных. – М.: Мир, 1989. – 544 с.
31. *Ahuja R. K., Magnanti T. L., Orlin J. B.* Network. Flows: Theory, Algorithms, and Applications. – Prentice Hall, 1993. – 864 p.
32. Research challenges in QoS routing / X. Masip-Bruin, M. Yannuzzi, J. Domingo-Pascual [and other] // Computer Communications. – 2006. – Vol. 29, Issue 5. – P. 563 – 581.
33. *Younis O., Fahmy S.* Constraint-Based Routing in the Internet: Basic Principles and Recent Research // IEEE Communications Surveys and Tutorials. – Vol. 5, Issue 1. – P. 2 – 13.
34. *Chen S., Nahrstedt K.* An overview of quality-of-service routing for the next generation high-speed networks: problems and solutions // IEEE Network Magazine. – 1998. – Vol. 12. – P. 64 – 79.

35. An overview of constraint-based path selection algorithms for QoS routing / F. Kuipers, P. Van Mieghem, T. Korkmaz [and other] // IEEE Communications Magazine. – 2002. – Vol. 40, Issue 12. – P. 50 – 55.
36. Polynomial Time Approximation Algorithms for Multi-Constrained QoS Routing / Guoliang Xue, Weiyi Zhang, Jian Tang [and other] // IEEE/ACM Transactions on Networking. – 2008. – Vol. 16, Issue 3. – P. 656 – 669.
37. *Azizul R. Mohd Shariff, Mike E. Woodward* A Delay Constrained Minimum Hop Distributed Routing Algorithm using Adaptive Path Prediction // Journal of Networks. – 2007. – Vol. 2, Issue 3. – P. 46 – 57.
38. *Prakash P.S., Selvan S.* An Efficient and Optimized Multi Constrained Path Computation for Real Time Interactive Applications in Packet Switched Networks // International Journal of Computer Science. – 2008. – Vol. 3 (4). – P. 293 – 300.
39. Finding a path subject to many additive QoS constraints / Guoliang Xue, Arunabha Sen, Weiyi Zhang [and other] // IEEE/ACM Transactions on Networking. – 2007. – Vol. 15, Issue 1. – P. 201 – 211.
40. *Xiao Y., Thulasiraman K., Xue G.* QoS routing in communication networks: approximation algorithms based on the primal simplex method of linear programming // IEEE Transactions on Computers. – 2006. – Vol. 55, Issue 7. – P. 815 – 829.
41. *Costa L. H., Fdida S., Duarte O. C.* Developing scalable protocols for three-metric QoS routing // Computer Networks: The International Journal of Computer and Telecommunications Networking. – 2002. – Vol. 39, Issue 6. – P. 713 – 727.
42. Highly responsive and efficient QoS routing using pre- and on-demand computations along with a new normal measure / Yanxing Zheng, Turgay Korkmaz, Wenhua Dou [and other] // Computer Networks: The International Journal of Computer and Telecommunications Networking. – 2006. – Vol. 50, Issue 18. – P. 3743 – 3762.
43. *Pragyansmita P., Raghavan S. V.* Survey of QoS routing // 15th international conference on Computer communication, 2002: proceedings of the conference. – P. 50 – 75.
44. *Andrew L.H., Kusuma A.N.* Generalized analysis of a QoS-aware routing algorithm // IEEE GLOBECOM, 1998: proceedings of the conference. Vol. 1. – P. 1 – 6.
45. *Van Mieghem P., Kuipers F.A.* Concepts of exact QoS routing algorithms // IEEE/ACM Transactions on Networking. – 2004. – Vol. 12, Issue 5. – P. 851 – 864.
46. *Neve H. De., Van Mieghem P.* TAMCRA: A Tunable Accuracy multiple constraints routing Algorithm // Computer Communications. – 2000. – Vol. 23. – P. 667 – 679. .
47. *Korkmaz T., Krunz M.* Multi-constrained Optimal Path Selection // INFOCOM 2001: proceedings of the conference, Vol. 2. – P. 834 – 843.
48. *Chen S., Nahrstedt K.* On finding multi-constrained path // IEEE International Conference on Communication, June 1998: proceedings of the conference, Vol. 2. – P. 874 – 879.
49. *Gouda M.G., Schneider M.* Maximizable routing metrics // Sixth International Conference on Network Protocols, 13 – 16 Oct 1998 : proceedings of the conference. – P. 71 – 78.
50. A Survey on Routing Metrics. TIK Report 262/ [Rainer Baumann, Simon Heimlicher, Mario Strasser, Andreas Weibel]. – Computer Engineering and Networks Laboratory, ETH-Zentrum, Switzerland, 2007. – 56 p.
51. *Whang Z., Crowcroft J.* Quality-of-service routing for supporting multimedia applications // IEEE Journal on Selected Areas in Communications. – 1996. – Vol. 14. – P. 1228 – 1234.
52. *Liang Guo, Matta I.* Search space reduction in QoS routing // 19th IEEE International Conference on Distributed Computing Systems, 1999: proceedings of the conference. – P. 142 – 149.

53. Peng B., Kemp A.H., Boussakta S. Impact of network conditions on QoS routing algorithms // 3rd IEEE Consumer Communications and Networking Conference, CCNC 2006, 8 – 10 Jan. 2006: proceedings of the conference, Vol. 1. – P. 25 – 29.
54. Guerin R. A., Orda A., Williams D. QoS Routing Mechanisms and OSPF Extensions // IEEE GLOBECOM, 1997: proceedings of the conference. – P. 1903 – 1908.
55. Ma Q., Steenkiste P. Quality-of-service routing for traffic with performance guarantees // IFIP International Workshop on Quality of Service, 1997: proceedings of the conference. – P. 115–126.
56. Shigang Chen, Meongchul Song, S. Sahni Two Techniques for Fast Computation of Constrained Shortest Paths // IEEE/ACM Transactions on Networking. – 2008. – Vol. 16, Issue 1. –P. 105 – 115.
57. Baradaran M., Yaghmaee M. H. A Constraint Based Routing Algorithm for Multimedia Networking // IAENG International Journal of Computer Science. – Vol. 33, Issue 2. – IJCS_33_2_2.
58. Alkahtani A., Woodward M., Al-Begain K. An Overview of QoS Routing in Communication Networks // 4th conference PGNet, 16 – 17 June 2003: proceedings of the conference. – P. 236–244.
59. Gang Liu, Ramakrishnan K.G. A*Prune: an algorithm for finding K-shortest paths subject to multiple constraints // Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2001, 22 26 apr. 2001: proceedings of the conference, Vol. 2. – P. 743 – 749.
60. Naumov V. Aggregate multipath QoS routing in the Internet // International Workshop Next Generation Network Technologies, 2002: proceedings of the conference. – P. 25 – 30.
61. Lee G. M. A survey of multipath routing for traffic engineering // Lecture Notes in Computer Science (LNCS). – Springer-Verlag, 2005. – Vol. 3391. – P. 264 – 274.
62. Обзор и сравнительный анализ основных моделей и алгоритмов многопутевой маршрутизации в мультисервисных телекоммуникационных сетях / В.В. Поповский, А.В. Лемешко, Л.И. Мельникова [и др.] // Прикладная радиоэлектроника. – 2005. – Т. 4, № 4. – С. 372 – 382.
63. Vutukury S., Garcia-Luna-Aceves J.J. MPATH: a loop-free multipath routing algorithm // Elsevier Journal of Microprocessors and Microsystems. – 2001. – No. 24 (6). – P. 319 – 327.
64. Поповский В.В., Лемешко А.В., Евсева О.Ю. Динамическое управление ресурсами ТКС: математические модели в пространстве состояний // Наукові записки УНДІЗ. – 2009. – № 1 (9). – С. 3 – 26.
65. Seok Yo., Lee Yo., Choi Ya Dynamic constrained multipath routing for MPLS networks // IEEE International Conference on Computer Communications and Networks, 2001: proceedings of the conference, Vol. 2., Issue 1. – P. 348 – 353.
66. Capone A., Fratta L., Martignon F. Dynamic Online QoS Routing Schemes: Performance and Bounds // Computer Networks. – 2006. – Vol. 50 (7). –P. 966 – 981.
67. Mitchell J.E., Farwell K., Ramsden D. Interior Point Methods for Large-Scale Linear Programming // Handbook of Optimization in Telecommunications; Edited by Mauricio G.C. Resende, Panos M. Pardalos. – Springer, 2006. – P. 3 – 25.
68. Ying-Xiao Xu, Gen-Du Zhang Models and algorithms of QoS-based routing with MPLS traffic engineering // 5th IEEE International Conference on High Speed Networks and Multimedia Communications, 2002: proceedings of the conference. – P. 128 – 132.
69. Migdalas A. Nonlinear Programming in Telecommunications // Handbook of Optimization in Telecommunications; Edited by Mauricio G.C. Resende, Panos M. Pardalos. – Springer, 2006. – P. 27 – 66.
70. Multipath protocol for delay-sensitive traffic / Umar Javed, Martin Suchara, Jiayue He [and other] // First international conference on COMMunication Systems And NETworks, 2009: proceedings of the conference. – P. 438 – 445.

71. Ahuja R. K., Magnanti T. L., Orlin J. B. Network. Flows: Theory, Algorithms, and Applications. – Prentice Hall, 1993. – 864 p.
72. Towards Robust Multi-Layer Traffic Engineering: Optimization of Congestion Control and Routing / J. He, M. Bresler, M. Chiang [and other] // IEEE Journal on Selected Areas in Communications. – 2007. – Vol. 25. – P. 868 – 880.
73. Kar K., Kodialam M., Lakshman T.V. Minimum interference routing of bandwidth guaranteed tunnels with MPLS traffic engineering applications // IEEE Journal on Selected Areas in Communications. – 2000. – Vol. 18. – P. 2566 – 2579.
74. Anton Riedl, Dominic A. Schupke Routing Optimization in IP Networks Utilizing Additive and Concave Link Metrics // IEEE/ACM Transactions on Networking. – 2007. – Vol. 15, Issue 5. – P. 1136 – 1148.
75. Wang Y., Wang Z. Explicit routing algorithms for Internet Traffic Engineering // Proc. of 8th International Conference on Computer Communications and Networks. Paris, 1999. – P. 582-588.
76. Merindol P., Pansiot J.-J., Cateloin S. Improving Load Balancing with Multipath Routing // 2008 Proceedings of 17th International Conference on Computer Communications and Networks. – 2008. – P. 1-8.
77. Лемешко А.В., Вавенко Т.В. Усовершенствование потоковой модели многопутевой маршрутизации на основе балансировки нагрузки [Електронний ресурс] // Проблеми телекомунікацій. – 2012. – № 1 (6). – С. 12 – 29. – Режим доступу до журн.: http://pt.journal.kh.ua/2012/1/1/121_lemeshko_multipath.pdf.
78. Optimal Load-Balancing / I. Keslassy, C. Chang, N. McKeown [and other] // IEEE INFOCOM '05: proceedings of the conference, Vol. 4, Issue 3. – P. 1054 – 1065.
79. Fortz B., Thorup M. Internet Traffic Engineering by Optimizing OSPF weights // 19th IEEE Conf. Comp. Comm., 2000: proceedings of the conference, Vol. 2. – P. 519 – 528.
80. Requirements for Traffic Engineering Over MPLS. RFC 2702/ [D. Awduche, J. Malcolm, J. Agogbua and other]. – 1999.
81. Banner R., Orda A. Multipath Routing Algorithms for Congestion Minimization // IEEE/ACM Transactions on Networking. – 2007. – Vol. 15. – P. 413 – 424.
82. Форд Л., Фалкерсон Д. Потоки в сетях. – М.: Мир, 1966. – 276 с.
83. Figueiredo G. B., da Fonseca N.L.S., Monteiro J.A.S. A minimum interference routing algorithm with reduced computational complexity // Computer Networks: The International Journal of Computer and Telecommunications Networking. – 2006. – Vol. 50. – P. 1710 – 1732.
84. Комплексна модель маршрутизації та обмеження трафіку в телекомунікаційних мережах військового призначення / Саваневич В.Е., Лемешко О.В., Агеев Д.В., Добришкін Ю.М. // Системи озброєння та військова техніка. – 2010. – № 2 (22). – С. 78 – 84.
85. Objective functions for optimization of resilient and non-resilient IP routing / M.Hartmann, D. Hock, M. Menth [and other] // 7th International Workshop on Design of Reliable Communication Networks, 25 – 28 Oct. 2009: proceedings of the conference. – P. 289 – 296.
86. Gallager R. G. A minimum delay routing algorithm using distributed computation // IEEE Trans. on communications. – 1975. – Vol. 25, Issue 1. – P. 73 – 85.
87. Vutukury S., Garcia-Luna-Aceves J.J. A Practical Framework for Minimum-Delay Routing in Computer Networks // Journal of High Speed Networks. – 1999. – Vol. 8, Issue 4. – P. 241 – 263.
88. Vutukury S., Garcia-Luna-Aceves J.J. A traffic engineering approach based on minimum-delay routing // IEEE IC3N, 2000: proceedings of the conference. – P. 42 – 47.
89. Математические модели исследования алгоритмов маршрутизации в сетях передачи данных / М.П. Березко, В.М. Вишневикий, Е.В. Левнер [и др.] // Информационные процессы. – 2001. – Т. 1, № 2. – С. 103–125.

90. *Vutukury S., Garcia-Luna-Aceves J.J.* A Simple Approximation to Minimum Delay Routing // ACM SIGCOMM, 1999: proceedings of the conference. – P. 39 – 50.
91. *Лемешко А.В., Дробот О.А.* Модель многопутевой QoS-маршрутизации в мультисервисной телекоммуникационной сети // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2006. – Вып. 144. – С. 16 – 22.
92. *Дробот О.А.* Комплексная модель обеспечения гарантированного качества обслуживания с реализацией динамических стратегий распределения сетевых ресурсов // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2007. – Вып. 148. – С.43 – 54.
93. *Лемешко А.В., Симоненко Д.В.* Динамическая модель балансировки буферных и канальных ресурсов транспортной сети телекоммуникационной системы [Электронный ресурс] // Проблемы телекоммуникаций. – 2010. – № 2 (2). – С. 42 – 49. – Режим доступа к журн.: http://pt.journal.kh.ua/2010/2/2/102_lemeshko_dynamic.pdf.
94. *Лемешко А.В., Евсева О.Ю.* Тензорная геометризация структурно-функционального представления телекоммуникационной системы в базисе межполюсных путей и внутренних разрезов // Наукові записки УНДІЗ. – 2010. – Вип. 1 (13). – С. 14 – 26.
95. *Лемешко А.В.* Тензорная модель многопутевой маршрутизации агрегированных потоков с резервированием сетевых ресурсов, представленная в пространстве с кривизной // Праці УНДІРТ. – 2004. – Вип. 4 (40). – С. 12 – 18.
96. *Лемешко А.В., Тимочко А.И., Кравчук А.А.* Вероятностная тензорная модель телекоммуникационной системы, представленной одномерной контурной сетью // Системи обробки інформації. Зб. наук. праць. – 2004. – Вип. 12 (40). – С. 118-127.
97. *Lemeshko A. V.* Probabilistic-Temporal Model of QOS-Routing with Precomputation of Routes under the Terms of Non-Ideal Reliability of Telecommunication Network // Telecommunications and Radio Engineering. – Vol. 66, Issue 13. – 2007. – P. 1151-1166.