Об одной задаче управленич сетевыми и информационными ресурсами в сетях NGN

Постоянное развитие и модернизация сетей связи, а также регулярное расширение спектра предоставляемых абонентам услуг приводит к тому, что компаниям связи приходится эксплуатировать значительное количество разнородного телекоммуникационного и ИТ оборудования, а также информационных систем. Возникает задача управления сетевыми и ИТ ресурсами и их наиболее эффективного распределения в узлах сети для оказания различных услуг разным группам абонентов с учетом многочисленных параметров (например, стоимостного критерия или параметров качества обслуживания — QoS). В докладе исследуется постановка задачи, и предлагаются подходы к ее решению.

Чукарин А.В.,
к.ф.-м.н., доцент кафедры систем телекоммуникаций РУДН,
chukarin@yandex.ru

Современные технологии предоставляют возможность внедрения широкого спектра телекоммуникационных услуг, конкретное функциональное содержание которых ограничивается только потребностями абонентов. В настоящее время количество услуг, предоставляемых в России операторами связи, исчисляется сотнями, включая как простейшие услуги по установлению местного телефонного соединения, так и комплексы интегрированных услуг, доступных подписчику независимо от того, в какой точке мира он находится. Если рассматривать услуги, предоставляемые телекоммуникационными компаниями всего мира, то перечень услуг связи будет насчитывать уже тысячи возможных наименований. Такое множество услуг, в значительной степени, связано с тем, что операторы, в зависимости от запросов пользователей, могут предоставлять им комплексные услуги, построенные на базе множества более простых услуг.

Одним из важнейших направлений развития телекоммуникаций, связанным с внедрением и предоставлением услуг, является расширение и улучшение возможностей, предназначенных для повышения качества и эффективности эксплуатации существующих разнородных ресурсов компаний связи. Такими ресурсами могут являться коммутационные станции и серверное оборудование, системное и прикладное программное обеспечение, базы данных и прочие активы телекоммуникационного предприятия. Развитие в упомянутом направлении практически невозможно без решения целого ряда задач, требующих строгих математических методов и алгоритмов, которые позволят обеспечить устойчивое и надежное управление всеми доступными ресурсами [6,7].

Рассмотрим пример, показанный на рис. 1 и иллюстрирующий упрощенную архитектуру сети, построенной в соответствии с концепцией NGN [3]. Проанализируем состав сетевых элементов, изображенных на рис. 1. Так, например, с точки зрения сигнализации основными структурными элементами сети NGN являются шлюзы сигнализации (SG — Signaling Gateway), транслирующие протоколы сигнализации между различными сетями, медиашлюзы (MG —

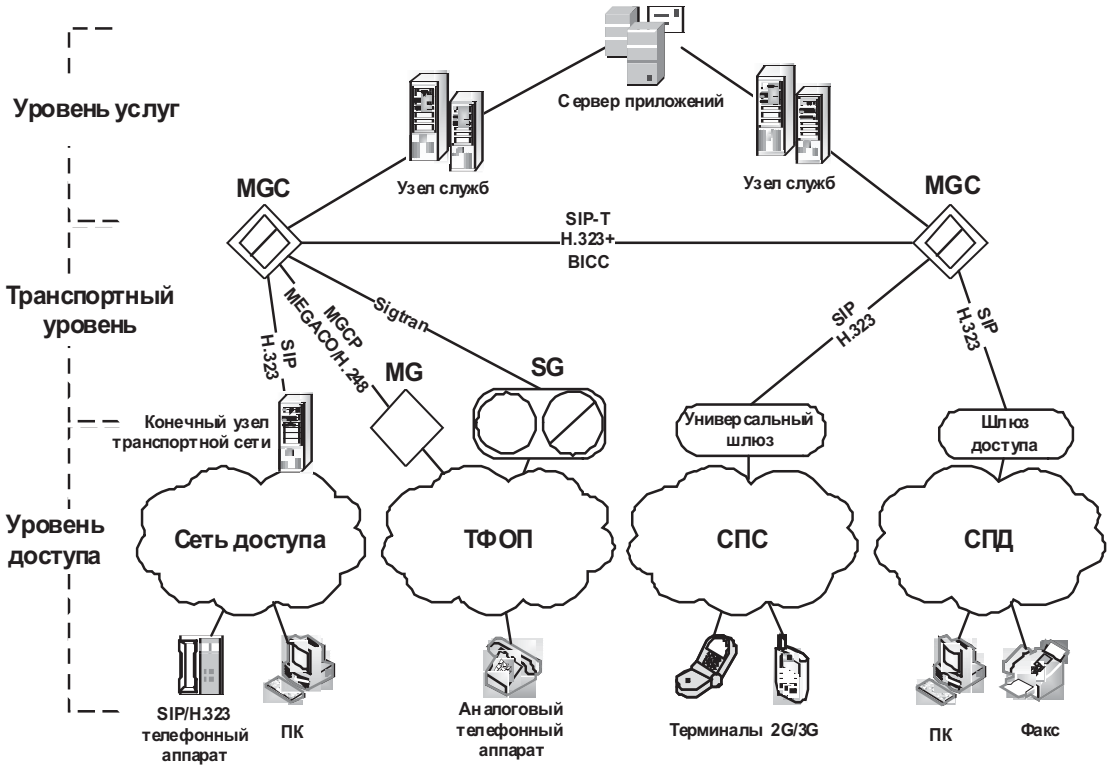


Рис. 1. Упрощенная архитектура сети NGN

Media Gateway), обеспечивающие взаимодействие между IP-сетью и сетевыми приложениями, такими, например, как сервисы ТфОП и беспроводных сетей, и контроллеры медиашлюзов (MGC — Media Gateway Controller или коммутаторы Softswitch), обеспечивающие координацию между шлюзами в соответствии с сигнальной информацией. С точки зрения поставщика услуг, основными структурными элементами будут являться узлы служб (SN — Service Node) и сервера приложений (AS — Application Server) и т.п. Таким образом, существует большое количество ракурсов, через которые можно рассматривать инфраструктуру компании связи.

Получается, что даже для упрощенной архитектуры количество элементов, подлежащих управлению весьма велико. Если дополнительно учесть, что экземпляры одного и того же типового элемента могут быть реализованы средствами различных производителей и отличаться функциональностью, то обеспечение прозрачного и эффективного управления ресурсами становится нетривиальной задачей.

Работа над рекомендациями по управлению сетями NGN ведется сегодня в крупнейших отраслевых организациях по стандартизации, включая МСЭ-Т, TMF, 3GPP (3rd Generation Partnership Project), ATIS (Alliance for Telecommunications Industry Solutions) и ETSI. Отчасти эта работа связана с гармонизацией стандартов в области управления сетями, построенными на базе различных технологий. Известен целый ряд концепций управления сетями связи (например, TMN [1] или NGOSS [2,5]), но, в основном, эти концепции либо основаны на использовании стандартных протоколов обмена управленческой информацией, либо содержат общие технические рекомендации. Интерес же представляет анализ возможности построения математической модели, пригодной для решения задачи управления сетевыми и информационными ресурсами в сетях NGN [4]. Разделение ресурсов на информационные (ИТ) и сетевые является наиболее логичным при постановке указанной задачи.

Исходя из сказанного выше, очевидно, что сеть связи представляет собой сложную распределенную систему, управление которой требует решения целого спектра разнообразных задач. В рекомендациях МСЭ-Т [1] эти задачи разделены на пять функциональных областей:

1. Управление конфигурацией:

— планирование и проектирование сети, управление установкой оборудования, ввод в эксплуатацию;

— контроль наличия и функционирования оборудования систем и сетей связи;

— обеспечение запасными частями и резервными комплектами оборудования и т. д.

2. Управление неисправностями:

— сбор и обработка сообщений о неисправностях и аварийных сигналов;

— локализация неисправности;

— устранение повреждения или неисправности;

— тестирование и восстановление работоспособности;

— проведение планово-предупредительных мероприятий и т. д.

3. Управление расчетами за услуги связи:

— определение методов измерения потребления и тарификации услуг связи;

— сбор сведений об оказанных услугах связи;

— применение тарифов, расчет сумм к оплате, выставление счетов и т. д.

4. Управление производительностью:

— отслеживание и сбор данных о функционировании сети, выявление снижения производительности;

— анализ показателей функционирования сети и принятие мер, направленных на предотвращения снижения производительности;

— повторная маршрутизация трафика, динамическое управление и т. д.

5. Управление безопасностью:

— разграничение и контроль доступа к элементам сети и системе управления;

— обеспечение конфиденциальности и целостности передаваемой информации;

— выявление и предотвращение несанкционированного пользования услугами и доступа к сетевым ресурсам;

— восстановление (программное и аппаратное) оборудования сетей и систем связи после сбоев в результате несанкционированных действий и т. д.

В англоязычных источниках для обозначения перечисленных областей задач часто используется аббревиатура FCAPS (Fault, Configuration, Accounting, Performance and Security).

Переход к сетям следующего поколения и конвергенция фиксированной и мобильной связи ставят перед разработчиками систем управления сетью новые задачи, трудность которых усугубляется расширением и усложнением цепочки создания ценности при предоставлении телекоммуникационных услуг, часто предполагающей сотрудничество нескольких операторов связи и поставщиков услуг.

Поскольку конвергентная NGN [6] состоит из совокупности взаимоувязанных гетерогенных сетей, для доступа в которые используется разнообразное терминальное оборудование, она должна обеспечивать предоставление пользователю услуг независимо от устройства и сети доступа. Различные приложения с одного терминала могут одновременно получать доступ к нескольким сетям разного типа. Таким образом, встает необходимость обеспечения на терминалах функций обнаружения и выбора сети.

Принцип обобщенной мобильности (англ. Generalized Mobility) требует от сети обеспечения пользователю возможности перехода без прерывания обслуживания не только между точками доступа одной сетевой технологии (например, Wi-Fi или GSM), но и между различными технологиями (например, 3G и 4G). К тому же, переход должен быть организован с поддержанием определенного набора параметров качества обслуживания. К функциям обеспечения мобильности относится роуминг. К NGN предъявляются требования по обеспечению национального, международного, а также так называемого вертикального роуминга, когда домашняя и гостевая сети построены на базе различных стандартов.

Мобильность пользователя и участие в предоставлении услуги нескольких телекоммуникационных компаний, состав которых (в отличие, например, от традиционных схем предоставления услуг международной телефонии) может в каждом конкретном случае меняться динамически при перемещении пользователя, усложняет функции начисления платы и биллинга. Управляющая система должна отследить и зафиксировать использование ресурсов сети на каждом этапе предоставления услуги, причем тарификация различных сегментов может быть основана на разных принципах (например, объем переданных данных, продолжительность соединения или контент). Разнообразие имеющихся и быстрое появление новых услуг связи также отнюдь не упрощает эту задачу.

Актуальнейшим вопросом управления сетями связи является обеспечение качества обслуживания. Несмотря на появление в последние годы большого количества методов управления и контроля качества обслуживания, многие операторы, особенно поставщики услуг Интернет, до сих пор предпочитают решать проблему путем

обеспечения избыточных ресурсов, что в ряде случаев оправдывается дешевизной последних. Однако ресурсы беспроводных сетей стоят дороже, и при дальнейшем распространении такого типа доступа следует ожидать более широкого внедрения различных механизмов обеспечения QoS. Обеспечение эффективной поддержки QoS является одной из основных трудностей управления сетями NGN.

Функции управления NGN в случае проводного доступа покрываются областями FCAPS. При использовании беспроводных сетей было предложено добавить дополнительные области: управление мобильностью, управление клиентами и управление терминалами.

Таким образом, возникает задача оптимального и наиболее эффективного распределения и задействования доступных сетевых и информационных ресурсов при предоставлении наиболее широкого набора услуг максимальному числу абонентов. В дальнейшем предполагается проведение исследований, направленных на математическую формализацию такой задачи, и решение ее в частном случае для услуги Triple Play (или т.н. пакета "Тройная услуга").

Литература

1. ITU-T Recommendation M.3010 (02/2000): Principles for a telecommunications management network. International Telecommuni- cation Union. — Geneva, 2000.

2. TMF GB921: Business Process Framework (eTOM) and its Addenda. Release 7.5. TMForum, 2008.

3. **Гайдамака Ю.В., Летников А.И., Пшеничников А.И., Чукарин А.В.** Системы сигнализации в сетях с коммутацией каналов и пакетов. — М.: ООО "Инсвязьиздат", МТУСИ, 2008. — 196 с.

4. **Муравьев В.В., Чукарин А.В.** К разработке модели каталога ресурсов сети NGN. Тезисы XLIII Всероссийской конференции по проблемам математики, информатики, физики и химии. Секции математики и информатики. — М.: Изд-во РУДН, 2007. — С.60.

5. **Самуйлов К.Е., Чукарин А.В., Яркина Н.В.** Бизнес-процессы и информационные технологии в управлении телекоммуникационными компаниями. — М.: Альпина Паблишерз, 2009. — 442 с.

6. **Семенов Ю.В.** Проектирование сетей связи следующего поколения. СПб.: Наука и Техника, 2005. — 240 с.

7. **Соколов Н.А.** Задачи перехода к сети связи следующего поколения. Автореф. дисс. на соиск. уч. степ. д.т.н. — СПб.: Изд-во СПбГУТ, 2006.

Методика расчета эффективности использования сетевых ресурсов в оптической пакетной транспортной сети с волновым уплотнением

Управление распределением сетевых ресурсов в процессе эксплуатации сети позволяет повысить эффективность использования уже имеющегося сетевого оборудования и линий связи. Использование эвристических алгоритмов и современных аппаратных комплексов позволяет решать задачу распределения сетевых ресурсов. При этом в процессе разработки подобных систем возникает необходимость получения независимой, эталонной оценки, позволяющей проводить анализ эффективности работы алгоритмов, реализованных в системе.

Савандюков И.М.,
Аспирант МТУСИ

Распределение ресурсов — ключевой элемент при проектировании и эксплуатации сети. Контроль за эффективностью использования ресурсов в процессе эксплуатации сети позволяет снизить капитальные затраты на строительство сети и получить дополнительные доходы с помощью линий связи и сетевых узлов, которые уже введены в эксплуатацию. Особенно это важно в оптических сетях, использующих технологию WDM, где освобождение даже одной длины волны эквивалентно освобождению полосы пропускания уровня STM-16 или даже STM-64 технологии SDH.

Не смотря на то, что расчет распределения ресурсов в пакетной сети с динамическим изменением трафика является сложной задачей, использование эвристических алгоритмов [1-3], реализуемых в современных программных комплексах позволяет решать эти задачи. В качестве примера можно рассмотреть решение задачи оптимизации распределения сетевых ресурсов в оптических пакетных транспортных сетях с оптическим уплотнением и поддержкой различных классов обслуживания. В качестве ресурсов в подобных сетях выступают длины волн, через которые организовываются световые пути между узлами-источниками и узлами-приемниками. Процесс оптимизации распределения сетевых ресурсов в нотации BPMN [4] представлен на рисунке.

При реализации приведенных процессов в программных комплексах при разработке метода распределения ресурсов возникает необходимость оценить максимально возможную эффективность использования сетевых ресурсов. Эта оценка в последствие может быть использована для сравнения результатов работы разных методов распределения ресурсов.

Одним из способов получить оценку эффективности использования сетевых ресурсов является создание аналитической модели, основанной на принципах теории массового обслуживания. Такая модель может дать оценку уровня использования ресурсов как на отдельно взятом сетевом узле, так и во всей сети в целом, а также позволит дополнительно проводить анализ параметров трафика, влияющих на качество предоставляемых услуг.

Первым шагом при оценке эффективности использования ресурсов всей сети является получение оценки эффективности использования ресурсов для каждого из сетевых узлов. Модель пакетной транспортной сети с волновым уплотнением и поддержкой классов обслуживания [5,6] предполагает наличие в сети узлов двух типов: граничных и опорных. Основное различие приведенных типов состоит в наличие у граничных узлов буферных накопителей, позволяющих им более гибко обрабатывать входные информационные потоки.

Схема процесса оптимизации распределения сетевых ресурсов оптической пакетной транспортной сети с волновым уплотнением и поддержкой классов обслуживания

Для анализа использования сетевых ресурсов узла удобно пользоваться методами теории массового обслуживания (ТМО). С точки зрения ТМО, опорный узел представляет собой систему вида $M_{N,\varepsilon_1,\ldots,\varepsilon_K}\Big|M_{\mu_1,\ldots,\mu_K}\Big|V$, где N — количество входных длин волн, V — количество выходных длин волн, K — количество классов обслуживания трафика, организованных в сети, ε_k — интенсивность поступления запросов на соединение k-й услуги, μ_k — интенсивность обслуживания запросов на соединение k-й услуги. Применяя мультисервисную модель Энгсета без мест для ожидания [7] для опорных узлов, и, определяя правила обслуживания заявок различных классов обслуживания, не трудно получить систему уравнений глобального баланса (СУГБ) для этого типа узла. Граничный узел с точки зрения ТМО представляет собой систему вида $M_{N,\varepsilon_1,\ldots,\varepsilon_K}\Big|M_{\mu_1,\ldots,\mu_K}\Big|V\Big|R_1, R_2\ldots R_K$, где N — количество входных интерфейсов, R_1, R_2...R_K — размер буферного накопителя для каждого из классов обслуживания. Использование мультисервисной модели Энгсета с местами для ожидания [7] позволяет получить СУГБ для граничных узлов.

Решение этих систем уравнений позволяет определить вероятность нахождения сетевого узла в состоянии, когда все ресурсы свободны. Для опорного узла выражение имеет вид:

$$p(\mathbf{0}) = \left(\sum_{\mathbf{n}\in \overline{S_V}} (N+\mathbf{n}-\sum_{i=1}^{K} n_i)_{\mathbf{n}} \cdot \prod_{i=1}^{K} \frac{\rho_i^{n_i}}{n_i!} + \sum_{\mathbf{n}\in S_V} (N-V+1)_{\mathbf{n}} \cdot (\prod_{i=1}^{K} \frac{\rho_i^{n_i}}{n_i!} + \sum_{l=1}^{n_k} (N-V)^l \cdot \sum_{\mathbf{n}\in S_V} \prod_{i=1}^{K} \frac{\rho_i^{n_i}}{n_i!} \right)^{-1} \quad (1)$$

где n_i — количество длин волн занятых для обслуживания заявок i-го класса обслуживания; $\mathbf{n} = (n_1,\ldots n_K)^T$ — вектор, описывающий количество находящихся в обработки заявок всех классов обслуживания;

$(N+\mathbf{n}-\sum_{i=1}^{K} n_i)_{\mathbf{n}} = \prod_{i=1}^{K} (N+n_i-\sum_{i=1}^{K} n_i)_{n_i}$; $\rho_i = \frac{\varepsilon_i}{\mu_i}$ — интенсивность нагрузки для k-ой услуги; S — пространство возможных состояний; S_V — пространство возможных состояний, при котором заняты все выходные линии; $\overline{S_V} = S \backslash S_V$.

Для граничного узла выражение имеет следующий вид:

$$p(\mathbf{0}) = \left(\sum_{\mathbf{n}\in S} (N+\mathbf{n}-\sum_{i=1}^{K} n_i)_{\mathbf{n}} \cdot \prod_{i=1}^{K} \frac{\rho_i^{n_i}}{v_i! V^{r_i}} \right)^{-1} \quad (2)$$

где r_i — количество занятых мест для ожидания в буферном накопителе, обслуживающего i-й класс трафика.

Подставляя полученные выражения (1) и (2) в соответствующие СУГБ не трудно определить вероятность нахождения любого узла сети в любом из возможных состояний. Это, в свою очередь, позволяет вычислить вероятность блокировки запросов на соединение для каждого из классов обслуживания, используя следующее выражение:

$$p_б^k = \sum_{i=1}^{k} p_{V_i} \quad (3)$$

где $p_б^k$ — вероятность блокировки запроса на содинение k-го класс обслуживания; p_{V_i} — вероятность нахождения системы в состоянии, когда запрос на соединение k-го класс обслуживания будет заблокирован.

Кроме того, использование формул (1) и (2) позволяет определить коэффициент использования ресурсов узла по формулам (4) и (5):

$$UTIL = \sum_{\mathbf{n}\in S} (m \cdot p_{S_m}), \quad (4)$$

где UTIL — среднее значение использования сетевых ресурсов; m — количество занятых ресурсов; S_m — пространство состояний системы, в котором занято ресурсов; p_{S_m} — суммарная вероятность нахождения системы в пространстве состояний S_m;

$$K = UTIL/V. \quad (5)$$

Из полученных выражений можно определить оптимальное значение коэффициента использования сетевых ресурсов. Этим значением является максимальное значение коэффициента использования сетевых ресурсов, при котором не нарушаются параметры качества обслуживания для всех классов трафика:

$$\begin{cases} K(\rho_1, \rho_2 \ldots \rho_K) \to \max; \\ p_б^1(\rho_1, \rho_2 \ldots \rho_K) < p_{бСоS1}; \\ \ldots \\ p_б^K(\rho_1, \rho_2 \ldots \rho_K) < p_{бСоSK}. \end{cases} \quad (6)$$

Получив результаты по использованию сетевых ресурсов на всех узлах сети, можно перейти к следующему этапу — оценке эффективности использования ресурсов на всей сети.

Наиболее просто оценить коэффициента использования ресурсов сети можно воспользовавшись нижеследующей формулой.

$$K_{net} = \sum_{n\subset N} UTIL_n \Big/ \sum_{l\subset L} W_l \quad (7)$$

где L — множество линий связи сети; N — множество узлов сети; $\sum_{n\subset N} UTIL_n$ — сумма всех занятых сетевых ресурсов; $\sum_{l\subset L} W_l$ — сумма всех имеющихся сетевых ресурсов.

При этом выбор оптимального значения коэффициента использования ресурсов аналогичен оценки этого параметра для отдельного сетевого узла.

Формула (6) позволяет получить верхнюю оценку эффективности использования ресурсов, поскольку не учитывает ряд факторов, влияющих на эффективность использования сетевых ресурсов. К таким факторам в первую очередь можно отнести фрагментацию сетевых ресурсов и неравномерность поступления запросов на соединение на различные граничные узлы.

Несмотря на указанные недостатки, предложенный метод обладает рядом существенных преимуществ. Во-первых, это относительная простота метода, позволяющая быстро оценить максимально допустимую нагрузку, имея в качестве исходных данных лишь информацию о количестве входных и выходных линий. Во-вторых, масштабируемость метода, который может быть с одинаковым успехом применен как для оценки эффективности использования сетевых ресурсов на отдельном узле сети, так и для оценки использования сетевых ресурсов части сети или всей сети в целом.

Литература

1. **Mohamed Kobuaa, Nicolas Puech, Maurice Gagnaire.** Traffic Engineering for Differentiated Services in Transparent WDM Networks // ITC19 / Performance Challenges for Efficient Next Generation Networks, 2005. — P.375-384.

2. **J. Phuritatkul.** Resource Allocation Algorithms for Controllable Service Differentiation in Optical Burst Switching Networks. — IEICE TRANS. COMMUN., Vol. E88-B, N4, April 2005.

3. Project P709. Planning of Full Optical Network. Deliverable 2. Basic factors influencing optical networks. - Annex C. Resource Allocation. 05.1999.

4. Business Process Modeling Notation version 1.0// http://www.bpmi.org/, 2004. — 296 p.

5. **George N. Rouskas.** Routing and Wavelength Assignment in Optical WDM Networks. // Wiley Encyclopedia of Telecommunications, (John Proakis, Editor), John Wiley & Sons, 2001.

6. **Kee Chang, Mohan Gurusamy, Yong Liu, Minh Hoang Phung.** Quality of Service in Optical Burst Switched Networks // Springer Science + Business Media, Inc., Boston, 2007. — 205 p.

7. **Башарин Г.П.** Лекции по математической теории телетрафика // Учеб. пособие. — М.: Изд-во РУДН, 2004. — 186 с.

Обзор возможных вариантов организации взаимодействия оптимизатора трафика с различными OSS-системами

Управление сетями большого размера немыслимо без использования вычислительной техники и специализированного программного обеспечения. Программно-аппаратный комплекс, развернутый у оператора связи, представляет собой набор различных систем, увязанных между собой и выполняющих определенные задачи. Добавление новой системы в этот комплекс требует серьезной проработки вопросов взаимодействия на стадии проектирования системы. В этой статье на примере внедрения оптимизатора трафика рассматриваются варианты организации взаимодействия между новой системой и различными классами систем, которые могут иметься у оператора связи.

Савандюков И.М.,
Аспирант МТУСИ

Управление сетями большого размера немыслимо без использования вычислительной техники и специализированного программного обеспечения. Кроме систем управления оборудованием, очень часто у оператора связи используются системы поддержки бизнеса и операций, так называемые OSS/BSS-системы. Эти системы позволяют проводить автоматизацию различных бизнес-процессов оператора связи, таких как выставление счетов абонентам, проведение технического учета оборудования и ресурсов сети, управление неисправностями, контроль параметров предоставляемых услуг и т.д. Классификация OSS/BSS-систем, а также перечисление функций для каждого класса систем приводится в [1, 2].

Таким образом, программно-аппаратный комплекс, используемый оператором, представляет собой большой набор различных систем, увязанных между собой и выполняющих определенные задачи. Для повышения эффекта от внедрения новой системы необходима ее интеграция в существующий программно-аппаратный комплекс оператора связи. Проведение интеграции позволяет реализовать дополнительные возможности , а также позволит использовать новую и уже имеющиеся системы как единое целое.

В качестве примера рассмотрим возможные варианты интеграции оптимизатора трафика (ОпТр) с различными системами управления и поддержки бизнеса оператора связи.

В этой статье под оптимизатором трафика будем понимать программно-аппаратный комплекс [3], основной задачей которого является оптимизация распределения ресурсов сети. Оптимизация ресурсов сети осуществляется за счет перемаршрутизации существующих каналов в зависимости от текущего состояния сети. Оптимизация распределения ресурсов может проводиться как по расписанию, определенному заранее (например, раз в неделю в выходной день), так и по событию (например, по требованию специалиста службы эксплуатации сети).

Для работы такого ОпТр необходима база данных (БД), хранящая информацию о сетевой конфигурации, а также модуль, содержащий правила проведения оптимизации распределения ресурсов.

Проведение оптимизации распределения ресурсов представляет собой набор процедур таких как: анализ состояния сети, проведение оптимизации ресурсов, формирование отчетов о проделанной работе.

Интеграция оптимизатора трафика с системами управления оборудованием

Для управления большинством типов оборудования разрабатываются специальные программные комплексы. Эти комплексы могут быть разделены на две категории:

1. Программный комплекс для управления одним экземпляром оборудования (Element Manager, EM);
2. Система для управления сетью (Network Management System, NMS).

Как правило, эти системы хранят в себе модели представления данных оборудования, которыми они управляют. Использование моделей оборудования этих систем в ОпТр позволяет уменьшить время, необходимое для внедрения оптимизатора трафика. Кроме того, если NMS имеет свою базу данных, то она может быть использована в качестве БД оптимизатора трафика. При реализации этой функции взаимодействие двух систем будет выглядеть следующим образом. Оптимизатор трафика будет формировать запрос к NMS или нужным EM, на которые система управления оборудованием будет передавать информацию, запрашиваемую оптимизатором трафика.

Кроме того, интеграция оптимизатора трафика с системой управления оборудованием может состоять в организации взаимодействия двух систем для переключений, выявленных в ходе работы оптимизатора трафика. Это позволит автоматизировать процесс оптимизации распределения ресурсов.

Однако, как правило, системы NMS и EM создаются производителем оборудования и предназначаются только для управления одним типом, реже для управления оборудованием, принадлежащего к одной продуктовой линейке. По этой причине, организация прямого взаимодействия ОпТр с системами управления оборудованием становятся сложной задачей, так как оптимизатору трафика придется работать с различными моделями оборудования. Кроме того, БД оборудования в системе управления может быть закрытой, что исключает ее использование при организации работы оптимизатора трафика.

Интеграция оптимизатора трафика с системой технического учета (NRI)

Система NRI представляет собой БД, хранящую всю необходимую информацию о сети оператора связи. Кроме информации об оборудовании, установленном на сети, эта система может хранить информацию о кабельных сооружениях, путях, абонентах, предоставляемых услугах и т.д. Модель описания объектов БД является универсальной, не зависит от типа оборудования или используемой технологии, а сама БД основана на продуктах хорошо известных, таких как Oracle или MySQL.

Интеграция оптимизатора трафика с системой учета сетевых ресурсов позволяет полностью отказаться от описания модели представления данных для оптимизатора, заимствуя ее у системы NRI.

Кроме выполнения своей основной задачи, интеграция оптимизатора трафика и системы NRI позволяет расширить функционал обоих систем. Так в результате интеграции у оператора связи появляется возможность проводить анализ изменений, вызванных добавлением или удалением тех или иных элементов сети с помощью мо-

делирования этих ситуаций. Эти возможности могут быть востребованы при составлении планов развития сети, а также при разработке схем защиты и восстановления сетевых маршрутов.

Взаимодействие двух систем может быть организовано различными способами. Во-первых, взаимодействие может быть организовано с помощью web-сервисов. Этот подход обеспечивает гибкость, позволяя, например, обращаться к данным, хранящимся в различных местах. Во-вторых, взаимодействие может быть обеспечено путем прямого обращения оптимизатора трафика к БД. При этом в БД должны быть разработаны специальные представления и процедуры, формирующие данные для работы оптимизатора трафика в заранее определенном формате.

Второй подход является менее гибким, поскольку взаимодействие систем осуществляется на уровне данных. Однако, такой подход обладает большей производительностью и скоростью работы, что немало важно при использовании оптимизатора трафика в сетях с большим количеством узлов.

Интеграция оптимизатора трафика с системами активации (SA)

Системы активации предназначены для автоматического проключения маршрутов с заданными параметрами через сеть. Для этой цели системы SA организуют подключение к оборудованию, через которое должен проходить маршрут, и производят соответствующую настройку этого оборудования. При этом, при невозможности включения маршрута, в системах SA производится обязательная проработка алгоритмов возвращения сети в исходное состояние.

Интеграция оптимизатора трафика с SA, также как и интеграция с системами управления оборудованием, позволяет полностью автоматизировать процесс оптимизации распределения ресурсов. Но в отличие от NMS, системы SA обладают большей универсальностью, что существенно упрощает организацию их взаимодействия с оптимизатором. Взаимодействие систем заключается в передаче оптимизатором трафика информации о маршруте в систему активации. Наиболее предпочтительным вариантом передачи этой информации является формат XML, позволяющий согласовать информацию о сети, хранящуюся в базах данных оптимизатора трафика и системы активации.

Максимального эффекта использования оптимизатора трафика можно добиться, если проводить его интеграцию одновременно с системами NRI и SA. Организация такого взаимодействия позволяет получить приводимые ниже преимущества, требуя при этом относительно незначительных трудозатрат на организацию такой системы:

— удается добиться полной автоматизации процесса оптимизации распределения ресурсов;

— с оптимизатора трафика снимается нагрузка, требуемая на организацию задач, непосредственно не связанных с основным назначением системы;

— все системы используют единую модель представления данных, что упрощает организацию их взаимодействия.

Интеграция оптимизатора трафика с системами управления неисправностями и контролем качества предоставления услуг (FM&PM)

Системы FM&PM с помощью получения информации от сетевого оборудования и дальнейшей ее обработки и анализа обеспечивают мониторинг состояния сети.

Интеграция систем FM&PM с оптимизатором трафика позволяет автоматически инициировать процедуру оптимизации распределения сетевых ресурсов при возникновении аварийной ситуации на сети или при ухудшении параметров качества предоставляемых услуг. В дополнении к этому системы FM&PM могут служить источниками информации о происходящих на сети изменениях.

Взаимодействие систем ОпТр и FM&PM состоит в вызове первой в случае возникновения определенной ситуации на сети. Список ситуаций, при котором происходит вызов оптимизатора трафика, должен быть определенном заранее на стадии проработки интеграции двух систем.

Интеграция оптимизатора трафик с системами управления заказами (OM)

Система управления заказами предназначена для автоматизации выполнения бизнес-процессов оператора связи. Каждый из процессов представляет собой последовательность операций, целью которых является выполнение элементарных задач, ведущих к получению конечного результата, например включению новой услуги, изменения параметров услуги, отключения услуги и т.д.

Из-за специфики OM, эта система взаимодействует со всеми системами, используемыми у оператора. Это позволяет обеспечить единый интерфейс взаимодействия оптимизатора трафика с другими системами оператора. Недостатком такого интерфейса является низкая скорость взаимодействия между различными системами. Однако, интеграция с OM позволяет включить оптимизатор трафика в те бизнес-процессы, где его применение необходимо. В конечном счете, это должно приводить к повышению эффективности использования ресурсов сети. А контроль над использованием ресурсов сети может производиться при каждом проведении изменений на сети.

Взаимодействие между оптимизатором трафик и системой OM должно быть двунаправленным. В качестве предпочтительных методов организации взаимодействия является использование языков XML (для описания представления данных) и Java или JavaScript (для описания процедур взаимодействия систем).

Подводя итоги вышесказанного можно сделать следующие выводы:

1. Интеграция оптимизатора трафика с системами управления оборудованием (EM) и сетью (NMS) является сложной, а часто и нереализуемой задачей. Более предпочтительна интеграция оптимизатора трафика с системами OSS, имеющие стандартные интерфейсы для взаимодействия и строящиеся на хорошо известных компонентах.

2. Организация взаимодействия различных OSS систем с оптимизатором трафика позволяет расширить функционал обоих систем, а также в некоторых случаях позволяет сократить время на внедрение оптимизатора трафика.

3. Наиболее интересной схемой взаимодействия является объединение оптимизатора трафика с системами NRI и SA в единое целое. Такая схема позволяет наиболее эффективно использовать функционал каждой из систем, а также избавиться от дублирования некоторых модулей систем, таких как БД, модуль анализа состояния сети, модуль отчетности и т.д.

4. Организация взаимодействия между различными системами является наиболее сложной из задач, возникающих при внедрении систем. В связи с этим вопросы возможного взаимодействия систем должны прорабатываться подробнейшим образом еще на этапе проектирования систем.

Литература

1. Enhanced telecom operations map (eTOM). The business process framework. For the information and communications services industry. Release 7.1.

2. Telecom application map. The OSS system landscape. Release 1.0.

3. **И.М. Савандюков.** Метод оптимизации ресурсов в оптических сетях с волновым уплотнением // T-Comm. Телекоммуникации и транспорт, 2009. — №4. — С.32-35.

4. ИТ сервис-менеджмент. Введение. — М.: IT Expert, 2003. — 228 с.

Численный анализ модели обслуживания сессий в сети GSM/GPRS

Построена функциональная модель процесса передачи информации по каналу сети GSM, в которой на частотный канал, состоящий из восьми физических каналов, поступают потоки GPRS-сессий. Исследована соответствующая аналитическая модель в виде многолинейной системы массового обслуживания с входящим MMPP потоком. При решении системы уравнений равновесия к матрице интенсивностей переходов применен метод LU-разложения. Получено стационарное распределение вероятностей, исследованы вероятностно-временные характеристики процесса передачи информации по каналу сети GSM.

Абаев П.О.,
ассистент кафедры систем телекоммуникаций РУДН,
pabaev@sci.pfu.edu.ru

Гайдамака Ю.В.,
доцент кафедры систем телекоммуникаций РУДН,
ygaidamaka@mail.ru

Рудикова Е.В.,
магистр кафедры систем телекоммуникаций РУДН,
eprudikova@yandex.ru

Введение

Сеть GSM, предназначенная в первую очередь для передачи речи, накладывает на использование приложений Интернет ряд технических ограничений, которые связаны, в основном, с невозможностью обеспечить требуемую для приложения скорость передачи. Для преодоления этих ограничений был разработан новый стандарт общего сервиса пакетной радиопередачи в сетях GSM технология GPRS (GPRS, General Packet Radio Service). Технология GPRS, использующая для передачи данных существующий в сети GSM радиоинтерфейс, является первым промежуточным шагом в создании систем третьего поколения (3G, 3rd Generation). Поэтому системы, поддерживающие технологию GPRS, получили название 2G+.

В работе на основе механизма распределения частот на базовой приемопередающей станции построена функциональная модель процесса передачи информации по каналу сети GSM, в которой на частотный канал, состоящий из восьми физических каналов, поступают потоки GPRS-сессий. Исследована соответствующая аналитическая модель в виде многолинейной системы массового обслуживания с входящим MMPP потоком (Markov Modulated Poisson Process). Для уменьшения объема вычислений при решении системы уравнений равновесия к матрице интенсивностей переходов применен метод LU-разложения, т.е. представления матрицы в виде произведения верхней треугольной матрицы и нижней треугольной матрицы, у которой все элементы главной диагонали — единицы. Получены стационарное распределение вероятностей и вероятностно-временные характеристики процесса передачи информации по каналу сети GSM. Для иллюстрации предложенного метода приведен пример расчета ВВХ модели.

Постановка задачи

В сети GSM используется множественный доступ с временным (TDMA, Time Division Multiple Access) и частотным (FDMA, Frequency Division Multiple Access) разделением. Как показано на рис. 1, для GSM зарезервированы две частотные полосы с разносом в 45 МГц: с 890МГц по 915 МГц для восходящего и с 935 по 960 МГц для нисходящего трафика. Каждая полоса шириной в 25 МГц разделяется на 124 несущих канала по 200КГц шириной. Частоты группируются парами, организуя дуплексный канал, за распределение несущих частот отвечает базовая приемопередающая станция.

Каждая частотная несущая содержит восемь физических каналов, соответствующих восьми временным интервалам, составляющим TDMA-кадр. В то время как в сети GSM для передачи речи используется по одному физическому каналу для передачи восходящего и нисходящего трафика, в сети GPRS для обслуживания одной сессии при передаче данных может использоваться до 8 физических каналов.

На основе механизма распределения радиочастот на базовой приемопередающей станции построим функциональную модель, в которой на частотный канал (рис. 2), состоящий из восьми физических каналов, поступают нисходящие потоки пакетов GPRS-сессий (например, данные, получаемые пользователем из сети Интернет). Восходящий трафик пакетов от мобильной станции (MS, Mobile Station) к базовой станции (BTS, Base Transceiver Station) в модели не учитывается (например, запросы пользователя на получение данных из сети Интернет), т.к. он не создает существенной нагрузки на систему.

Возможность предоставления услуги передачи данных по технологии GPRS с различными уровнями качества обеспечивается с помощью одной из стратегий управления доступом пакетов в частотный канал. Для каждой GPRS-сессии устанавливается максимально

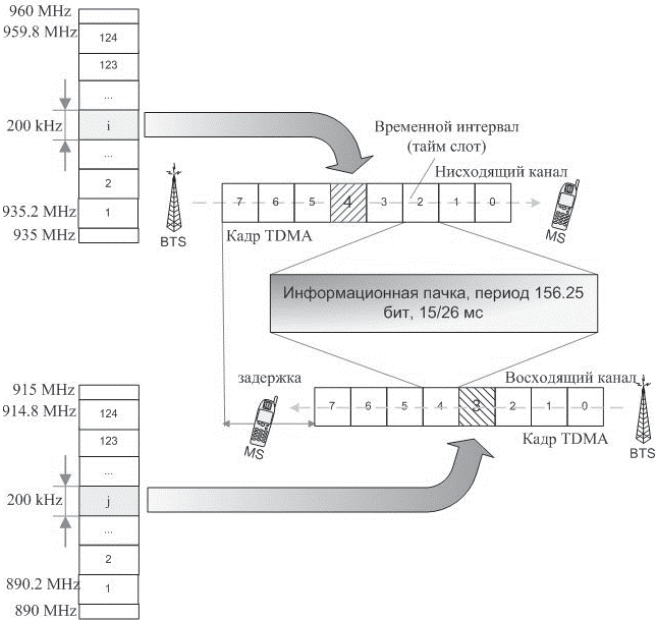


Рис. 1. Схема несущих частот GSM и кадра TDMA

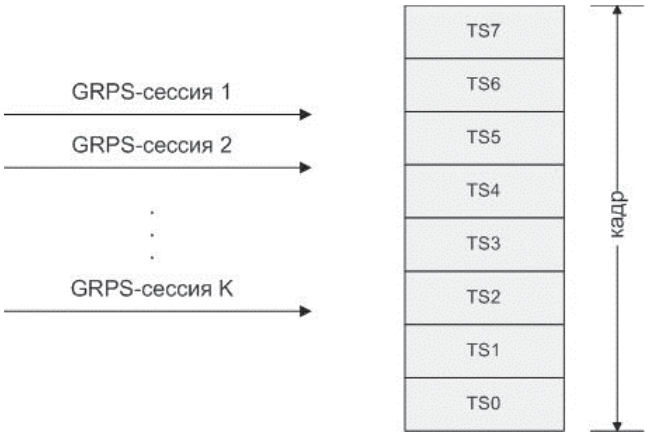


Рис. 2. Функциональная схема обслуживания трафика GPRS-сессий на частотном канале

возможное число доступных для передачи физических каналов. Если сессия исчерпала все доступные ей каналы, то вновь прибывающие пакеты данной сессии теряются.

GPRS-сессия, как показано на рис. 3, представляется последовательностью чередующихся периодов "On"/"Off" активности и паузы. В течение периода активности может быть передано некоторое количество пакетов, во время паузы пакеты не передаются, таким образом, трафик GPRS-сессий имеет пульсирующий характер.

Задачей исследования является анализ одного из основных показателей качества обслуживания трафика GPRS-сессии на участке между мобильной и базовой приемопередающими станциями вероятности потери пакета из-за отсутствия свободных каналов. С этой целью в работе построена математическая модель в виде многолинейной системы массового обслуживания (СМО) с входящим MMPP потоком (Markov Modulated Poisson Process) [2], а также предложен метод расчета характеристик модели.

Математическая модель

На рис. 4 схематично изображена СМО, моделирующая обслуживание трафика GPRS-сессий одним частотным каналом. На систему, состоящую из C приборов, от K источников (GPRS-сессий) поступает MMPP-поток заявок (GPRS-пакетов). Каждый источник может находиться в одном из двух состояний "0" (Off) или "1" (On). Длительности пребывания k-источника в состояниях "0" и "1" имеют экспоненциальное распределение с параметрами α_k и β_k соответственно, $k = \overline{1,K}$. В состоянии "0" источник не генерирует заявки. В состоянии "1" k-й источник генерирует заявки (k-заявки) с интенсивностью λ_k, $0 < \lambda_k < \infty$, $k = \overline{1,K}$. Число приборов, доступных k-заявкам, не превосходит заданной величины r_k, $0 \le r_k \le C$, $k = \overline{1,K}$. Если в момент поступления k-заявки заявками того же типа занято меньше, чем r_k приборов, k-заявка поступает на обслуживание, занимает один прибор и обслуживается на нем в течение случайного времени, имеющего экспоненциальное распределение с параметром μ, $0 < \mu < \infty$, не зависящим от типа заявки. Если в момент поступления k-заявки заявками того же типа занято r_k приборов или все приборы в системе заняты, поступившая заявка теряется. Построенную СМО будем кодировать в обозначениях Башарина-Кендалла как $MMPP(K,r)|M|C$.

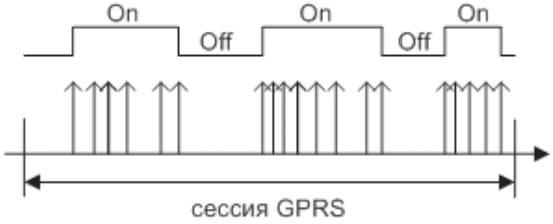


Рис. 3. Временная диаграмма трафика GPRS-сессии

Обозначим $n_k(t)$ число k-заявок в СМО, $0 < n_k(t) \le r_k(t)$, и $l_k(t) \in \{0,1\}$ состояние k-го MMPP потока (k-потока) в момент времени $t \ge 0$. Тогда составной случайный процесс $\mathbf{X}(t) = (\mathbf{n}(t), \mathbf{l}(t))$ по построению является Марковским процессом и описывает поведение исследуемой системы над множеством состояний

$$\mathcal{X} = \mathcal{N} \times \mathcal{L},\ \mathcal{N} = \left\{ \mathbf{n} : \sum_{i=1}^{K} n_i \le C,\ 0 \le n_k \le r_k,\ k = \overline{1,K} \right\},\ \mathcal{L} = \{0,1\}^K \tag{1}$$

Обозначим $p(\mathbf{n},\mathbf{l})$, $(\mathbf{n},\mathbf{l}) \in \mathcal{X}$, стационарное распределение вероятностей состояний МП $\chi(t)$ и рассмотрим $B_k \subset \mathcal{X}$ множество состояний, в которых система не может принять на обслуживание k-заявку. Будем называть B_k множеством блокировок k-заявок, и для его построения введем следующие вспомогательные множества:

$$\Pi_k^b = \{(\mathbf{n},\mathbf{l}) \in \mathcal{X} : n_k = r_k; l_k = 1\}$$

— множество состояний блокировки из-за того, что в системе уже обслуживается r_k заявок того же типа;

$$\Pi_k^C = \{(\mathbf{n},\mathbf{l}) \in \mathcal{X} : n. = C, n_k \ne r_k; l_k = 1\}$$

— множество состояний блокировок k-заявок из-за отсутствия свободных приборов.

С учетом введенных обозначений множество блокировок B_k имеют вид

$$B_k = \Pi_k^b \cup \Pi_k^C \tag{2}$$

а вероятность блокировки B_k определяется формулой

$$B_k = \sum_{(\mathbf{n},\mathbf{l}) \in \Pi_k^b} p(\mathbf{n},\mathbf{l}) + \sum_{(\mathbf{n},\mathbf{l}) \in \Pi_k^C} p(\mathbf{n},\mathbf{l}) - \sum_{(\mathbf{n},\mathbf{l}) \in \Pi_k^b \cap \Pi_k^C} p(\mathbf{n},\mathbf{l}) \tag{3}$$

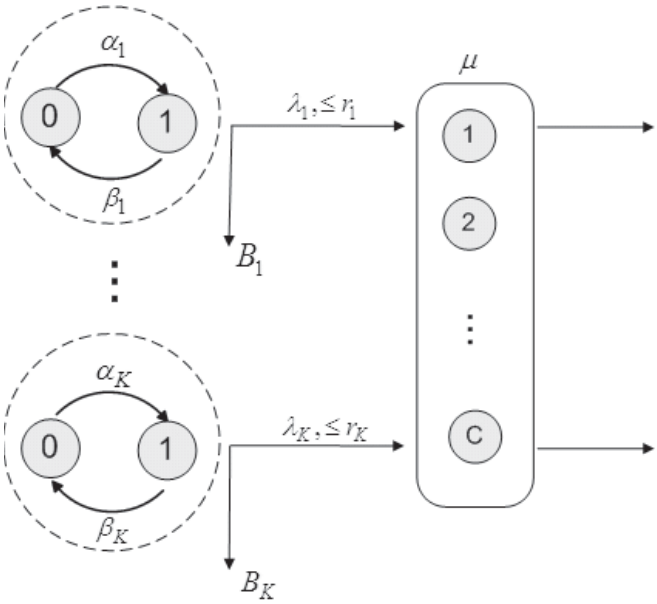


Рис. 4. СМО $MMPP(K,r)|M|C$

Задачей данной работы является разработка метода вычислений распределения вероятностей $\mathbf{p}(\mathbf{n},\mathbf{l})$, $(\mathbf{n},\mathbf{l}) \in \mathcal{X}$, и вероятностей блокировки B_k, $k = \overline{1,K}$.

Метод вычисления характеристик

Пусть $\mathbf{A}$ инфинитезимальный оператор МП $X(t)$ и $\gamma = \max\limits_{j=\overline{1,K}} r_j$ максимальное значение элементов вектора $\mathbf{n} \in N$. Обозначим $x(\gamma+1) = \overline{(x_1 x_2 \ldots x_{2K})}_{\gamma+1}$ число, составленное из элементов вектора $\mathbf{x}$ и представленное в системе счисления по основанию $(\gamma + 1)$. Очевидно, что представление числа $x(\gamma + 1)$ в десятичной системе счисления определяется формулой

$$(x(\gamma+1))_{10} = \sum_{i=1}^{2K} x_i (\gamma+1)^{2K-i} \tag{4}$$

Введем на множестве χ лексикографический порядок

$$\left(\mathbf{x}' = (\mathbf{n}',\mathbf{l}') > \mathbf{x}'' = (\mathbf{n}'',\mathbf{l}'')\right) \Leftrightarrow$$
$$\Leftrightarrow \left((\mathbf{n}' > \mathbf{n}'') \vee \left((\mathbf{n}' = \mathbf{n}'') \wedge \left(\sum_{i=1}^{2K} (x_i' - x_i'') \cdot (1+\gamma)^{2K-i} > 0 \right) \right) \right) \tag{5}$$

Нетрудно убедиться, что матрица $\mathbf{A}$ имеет блочный трехдиагональный вид, т.е.

$$\mathbf{A} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{U}_0 & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{D}_1 & \mathbf{U}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 & \mathbf{D}_2 & \mathbf{U}_2 & \mathbf{0} & \ldots \\ \ldots & \mathbf{0} & \ldots & \ldots & \ldots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{L}_{C-1} & \mathbf{D}_{C-1} & \mathbf{U}_{C-1} \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{L}_C & \mathbf{D}_C \end{pmatrix} \tag{6}$$

Заметим, что множество χ представимо в виде $\mathcal{X} = \bigcup_{i=0}^{C} \mathcal{X}_i$, где $\mathcal{X}_i = \{(\mathbf{n},\mathbf{l}) \in \mathcal{X} : n. = i\}$ и $\mathcal{X}_i \cap \mathcal{X}_j = \varnothing,\ i \neq j$. Поэтому наддиагональные блоки $\mathbf{U}_i$ описывают переходы МП $X(t)$ из состояний множества χ_i в состояние множества χ_{i+1}, $i = \overline{0, C-1}$. Аналогично поддиагональные блоки $\mathbf{L}_i$ описывают переходы из состояний множества χ_i в состояние множества χ_{i-1}, $i = \overline{1,C}$, а диагональные блоки $\mathbf{D}_i$ — переходы внутри множества χ_i, $i = \overline{0,C}$.

Обозначим $\mathbf{e}_i$ вектор длины K, i-й элемент которого равен 1, а остальные элементы равны 0. Тогда ненулевые элементы блоков матрицы $\mathbf{A}$ вычисляются по следующим формулам:

$$\mathbf{U}((\mathbf{n}',\mathbf{l}'),(\mathbf{n}'',\mathbf{l}'')) = \lambda_j, \text{ если } \mathbf{l}' = \mathbf{l}', \mathbf{n}'' = \mathbf{n}' + \mathbf{e}_j,\ 0 < n_j' < r_j, l_j' = 1 \tag{7}$$

$$\mathbf{L}((\mathbf{n}',\mathbf{l}'),(\mathbf{n}'',\mathbf{l}'')) = n_j \cdot \mu,\ \ j = \overline{1,K}, \text{если } \mathbf{l}' = \mathbf{l}'', \mathbf{n}'' = \mathbf{n}' - \mathbf{e}_j \tag{8}$$

$$\mathbf{D}((\mathbf{n}',\mathbf{l}'),(\mathbf{n}'',\mathbf{l}'')) = \begin{cases} \alpha_j, \text{если } \mathbf{n}' = \mathbf{n}', \mathbf{l}'' = \mathbf{l}' + \mathbf{e}_j,\ l_j' = 0, \\ \beta_j, \text{если } \mathbf{n}' = \mathbf{n}', \mathbf{l}'' = \mathbf{l}' - \mathbf{e}_j,\ l_j' = 1. \end{cases} \tag{9}$$

Для определения стационарных вероятностей необходимо решить систему уравнения равновесия (СУР)

$$\mathbf{p}^T \cdot \mathbf{A} = \mathbf{0}^T \tag{10}$$

Поскольку задача осложняется большой размерностью пространства состояний, а следовательно, и матрицы $\mathbf{A}$, то для поиска решения СУР воспользуемся методом LU-разложения [1], в результате которого можно получить представление матрицы в виде произведения $\mathbf{A} = \mathbf{L} \cdot \mathbf{U}$, где $\mathbf{U}$ — верхняя треугольная матрица, а $\mathbf{L}$ — нижняя треугольная матрица, у которой все элементы главной диагонали единицы. Полученные в ходе разложения матрицы $\mathbf{L}$ и $\mathbf{U}$ будут также блочными, и размеры их блоков будут совпадать с размерами соответствующих блоков матрицы $\mathbf{A}$.

$$\mathbf{A} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{U}_0 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{D}_1 & \mathbf{U}_1 & \mathbf{0} & \ldots \mathbf{0} \\ \ldots & & & & \\ \mathbf{0} & \ldots & & \mathbf{L}_C & \mathbf{D}_C \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \ldots & \mathbf{0} \\ -\mathbf{Y}_1 & \mathbf{I} & \mathbf{0} \ldots & \mathbf{0} \\ \ldots & & & \\ \mathbf{0} & \ldots & -\mathbf{Y}_C & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{W}_0 & \mathbf{V}_0 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_1 & \mathbf{V}_1 & \mathbf{0} \ldots & \mathbf{0} \\ \ldots & & & & \\ \mathbf{0} & \ldots & & & \mathbf{W}_C \end{pmatrix} \tag{11}$$

Представим вектор стационарных вероятностей в виде $\mathbf{p}^T = (\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_C)$, причем число элементов подвектора $\mathbf{p}_i$, $i = \overline{0,C}$, равно числу строк (столбцов) блока $\mathbf{D}_i$, $i = \overline{0,C}$. Согласно [1] для неразложимой матрицы $\mathbf{A}$ справедливо утверждение: существует единственное LU-разложение $\mathbf{A} = (\mathbf{L} - \mathbf{Y})\mathbf{W}$, где $\mathbf{Y}$ — нижняя треугольная блочная матрица с блоками $\mathbf{Y}_{ij}$, имеющая нулевые диагональные блоки, $\mathbf{W}$ — верхняя треугольная блочная матрица с блоками $\mathbf{W}_{ij}$, матрица $\mathbf{W}_{CC}$ — нижний блок матрицы $\mathbf{W}$ является неразложимой консервативной матрицей, а вектор $\mathbf{p}^T = (\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_C)$ является решением СУР (10) тогда и только тогда, когда

$$\mathbf{p}_C^T \cdot \mathbf{W}_{CC} = \mathbf{0}^T \tag{12}$$

$$\mathbf{p}_k^T = \sum_{i=k+1}^{C} \mathbf{p}_i^T \mathbf{Y}_{ik},\ k = C-1, \ldots, 0 \tag{13}$$

Из формул (12) и (13) следует, что вектор $\mathbf{p}^T = (\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_C)$ является решением СУР (10), если с учетом обозначений $\mathbf{Y}_{ii} \equiv \mathbf{Y}_i \underline{\mathbf{W}_{ii} \equiv \mathbf{W}_i}$ справедливы соотношения: $\mathbf{p}_C^T \cdot \mathbf{W}_C = \mathbf{0}^T$ и $\mathbf{p}_i^T = \mathbf{p}_{i+1}^T \cdot \mathbf{Y}_{i+1}$, $i = 0, C-1$, которые позволяют свести решение системы (10) к решению системы меньшего порядка.

Для успешной работы алгоритма LU-разложения элементы главной диагонали матрицы $\mathbf{A}$ не должны обращаться в нуль в процессе вычислений. Сложности могут возникать и в случае, когда эти элементы не равны нулю, но близки к нему. В прикладных задачах вместо LU-разложения чаще используют LUP-разложение $\mathbf{P} \cdot \mathbf{A} = \mathbf{L} \cdot \mathbf{U}$, где $\mathbf{P}$ — матрица перестановок, в каждом столбце которой и в каждой строке по одному элементу, равному единице, а остальные элементы равны нулю, определитель матрицы $\mathbf{P}$ по модулю равен единице, $\mathbf{p}^T$, $\mathbf{p}^{-1}$ произведение двух матриц перестановки также является матрицей перестановки.

Для получения LUP-разложения используют модифицированный метод Гаусса исключения неизвестных, причем на каждой итерации происходит операция перестановки местами двух строк соответствующей матрицы, необходимая для получения ненулевых ведущих элементов.

Пример расчета

Рассмотрим пример построения матрицы $\mathbf{A}$ для случая $C = 3$, $K = 2$, $\mathbf{r} = (2,3)$. Выпишем пространство состояний χ согласно формуле (1) и первые несколько блоков матрицы $\mathbf{A}$

$$\mathcal{X} = \{(0,0), (0,1), (1,0), (1,1), (0,2), (2,0), (0,3), (1,2), (2,1)\} \times \{0,1\}^2 \tag{14}$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{U}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{D}_1 & \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 & \mathbf{D}_2 & \mathbf{U}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_3 & \mathbf{D}_3 \end{pmatrix},\ \mathbf{U}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & \lambda_1 \end{pmatrix},$$

$$\mathbf{L}_1 = \begin{pmatrix} \mu_2 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 \\ 0 & 0 & \mu_2 & 0 \\ 0 & 0 & 0 & \mu_2 \\ \mu_1 & 0 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 \\ 0 & 0 & \mu_1 & 0 \\ 0 & 0 & 0 & \mu_1 \end{pmatrix}$$

$$\mathbf{D}_0 = \begin{pmatrix} -(\alpha_1+\alpha_2) & \alpha_2 & \alpha_1 & 0 \\ \beta_2 & -(\beta_2+\alpha_1+\lambda_2) & 0 & \alpha_1 \\ \beta_1 & 0 & -(\beta_1+\alpha_2+\lambda_1) & \alpha_2 \\ 0 & \beta_1 & \beta_2 & -(\beta_1+\beta_2+\lambda_2+\lambda_1) \end{pmatrix}$$

Отметим, что на каждом шаге происходит увеличение размера блока: если блок $\mathbf{D}_0$ имеет размер 4 x 4, то блок $\mathbf{D}_3$ уже 12 x 12. Построим вспомогательные множества и множества блокировок

$$\Pi_1^b = \{(2,0,1,0),(2,0,1,1),(2,1,1,0),(2,1,1,1)\},\ \Pi_2^b = \{(0,3,0,1),(0,3,1,1)\},$$
$$\Pi_1^c = \{(0,3,1,0),(0,3,1,1),(1,2,1,0),(1,2,1,1)\},\ \Pi_2^c = \{(2,1,0,1),(2,1,1,1)\},$$
$$B_1 = \{(2,0,1,0),(2,0,1,1),(2,1,1,0),(2,1,1,1),(0,3,1,0),(0,3,1,1),(1,2,1,0),(1,2,1,1)\},$$
$$B_2 = \{(0,3,0,1),(0,3,1,1),(2,1,0,1),(2,1,1,1)\}$$

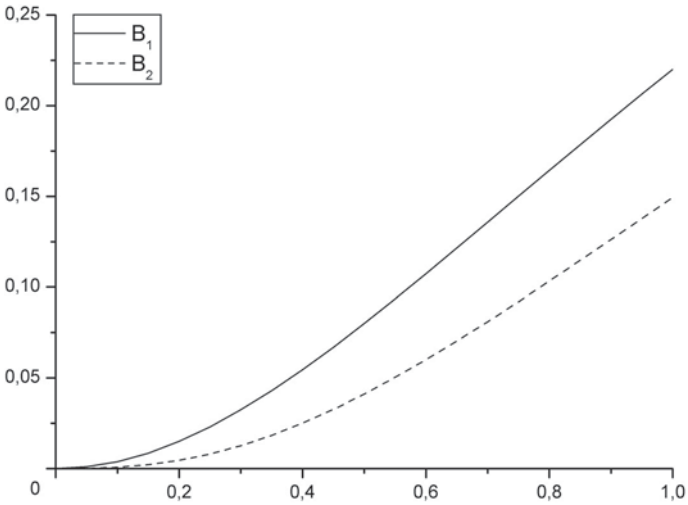


Рис. 5. Зависимость вероятностей блокировки k-заявок B_k, k = 1,2, от нагрузки $\rho_1 = \rho_2 = \rho$ ($C = 3,\ K = 2,\ \mathbf{r} = (2,3)$)

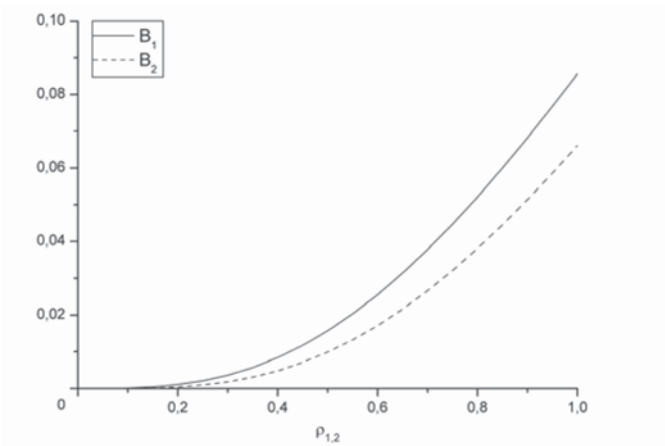


Рис. 6. Зависимость вероятностей блокировки k-заявок B_k, k = 1,2, от нагрузки $\rho_1 = \rho_2 = \rho$ ($C = 4,\ K = 2,\ \mathbf{r} = (3,4)$)

Выполним расчеты в предположении, что GPRS-пакеты передаются по схеме CS2 со средней скоростью v = 13,4 кбит/с [3]. Выберем согласно [4] среднюю длину пакета l = 480 байт, среднюю продолжительность периода "On" 2 секунды, среднюю продолжительность периода "Off" 20 с. Среднее время передачи по каналу можно рассчитать как $\mu = \frac{v}{l}\ c^{-1}$.

На рис. 5 и 6 приведены графики зависимости вероятностей блокировок k-заявок от нагрузки $\rho_k = \lambda_k/\mu$, $0 \le \rho_k < 1$, для случаев C = 3, K = 2, $\mathbf{r}$ = (2,3) и C = 4, K = 2, $\mathbf{r}$ = (3,4) . Заметим, что при увеличении числа приборов на единицу вероятность блокировки уменьшилась в два раза.

Как видно из графиков, кривая вероятностей блокировки 2 заявок лежит выше соответствующей кривой для 1-заявок. Это объясняется тем, что для обоих примеров значения компонент вектора $\mathbf{r}$ выбраны так, чтобы обеспечить приоритетное обслуживание заявок второго типа. Таким образом, пользователь второй GPRS-сессии получает услугу передачи данных по технологии GPRS с более высоким уровнем качества по сравнению с пользователем первой GPRS-сессии. Для пользователя второй GPRS-сессии это может означать, например, более высокое качество изображения в случае загрузки видеоинформации или более высокую скорость получения данных из сети Интернет, что приведет к меньшему времени ожидания до полной загрузки запрошенных данных.

Заключение

В работе построена модель процесса передачи информации по технологии GPRS и проведен численный анализ одного из показателей качества обслуживания протокола при передаче данных на участке сети сотовой подвижной связи между мобильной и базовой станциями. В настоящее время остается актуальной задача выбора оптимальной стратегии управления доступом для разделения радиоресурсов базовой приемопередающей станции между пользователями. Эта задача еще усложнится, если при ее решении учитывать голосовой трафик.

В заключение авторы выражают глубокую благодарность заведующему кафедрой систем телекоммуникаций РУДН д.т.н., проф. К.Е. Самуйлову за постановку задачи и полезные советы в процессе исследований.

Литература

1. **Наумов В.А.** Численные методы анализа марковских систем. Учебное пособие. — М.: Изд-во РУДН, 1985. — 36 с.

2. **W. Fischer, K. S.** Meier-Hellstern The Markov-Modulated Poisson Process (MMPP) cookbook. //Performance Evaluation, 18(2), 1993. — P. 149-171.

3. Digital cellular telecommunications system (Phase2+); General Packet Radio Service (GPRS); Service description; Stage 2 //ETSI, GSM recommendation 03.60, v. 7.4.1. 1998.

4. Universal Mobile Telecommunication System (UMTS); Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS //ETSI, Technical Report TR 101 112 v3.2.0. 1998.

5. **C. Lindemann, A.** Thummler Performance analysis of the general packet radio service //Computer Networks #41, 2003. — P. 1-17.

Построение и анализ модели установления соединения по протоколу SIP в сети связи следующего поколения

Целью исследования является анализ показателей качества обслуживания соединения по протоколу SIP в сети связи следующего поколения. В работе предложена математическая модель протокола в виде экспоненциальной неоднородной сети массового обслуживания. Получены формулы для расчета стационарных вероятностей и основных вероятностно-временных характеристик процесса установления соединения.

Абаев П.О.,
ассистент кафедры систем телекоммуникаций РУДН,
pabaev@sci.pfu.edu.ru

Хатунцев А.Б.,
аспирант кафедры автоматической электросвязи МТУСИ,
antkh@mail.ru

Введение

Отличительной особенностью сигнализации в сетях связи следующего поколения (NGN) является отсутствие необходимости в выделении канальных ресурсов для передачи сигнального трафика, который обслуживается на базе общих транспортных механизмов, реализованных в сетях передачи данных общего пользования оператора связи на базе протокола IP. Понятие "сеть сигнализации в NGN" обозначает совокупность всех соединений сигнализации, обеспечивающих передачу сигнального трафика. Под сигнальным трафиком понимается трафик, создаваемый сообщениями протоколов, обеспечивающих процедуры установления, поддержания, модификации и разъединения соединений. Соединение сигнализации состоит из узлов, в которых реализованы сигнальные протоколы, и звеньев сети, соединяющих эти узлы.

Операторы связи позиционируют сети NGN как надстройку над существующими телефонными сетями общего пользования (ТфОП) и IP-сетями [1-4]. Поэтому при модернизации сетей они вынуждены концентрировать нагрузку от сетей ТфОП в шлюзах, которыми управляют гибкие коммутаторы (Softswitch, SX). Данный подход называется многоуровневым, где шлюз находится на уровне доступа, а Softswitch на уровне управления.

На рис. 1 показана общая схема сети, условно разделенная на три подсети: сеть доступа ТфОП, транспортная сеть и сеть доступа IP/SIP. Сигнальное взаимодействие сети ТфОП с сетью с коммутацией пакетов осуществляется посредством протокола MEGACO и технологии Sigtran. В работе рассматриваются соединения сигнализации в предположении, что при установлении соединений в сети применяются три базовые технологии: ОКС 7, Sigtran и протокол установления сеансов связи — SIP. Поэтому соединения сигнализации содержат следующие три компонента: соединение ОКС 7, соединение Sigtran и соединение по протоколу SIP. Другие компоненты, например соединения по протоколу MEGACO, рассматривать не будем, что не нарушит общности разрабатываемых методов анализа показателей качества обслуживания соединений между агентами пользователей сети ТфОП и сети SIP. В качестве искомого показателя выберем время установления соединения из конца в конец. С этой целью в разделе 1 построена математическая модель установления

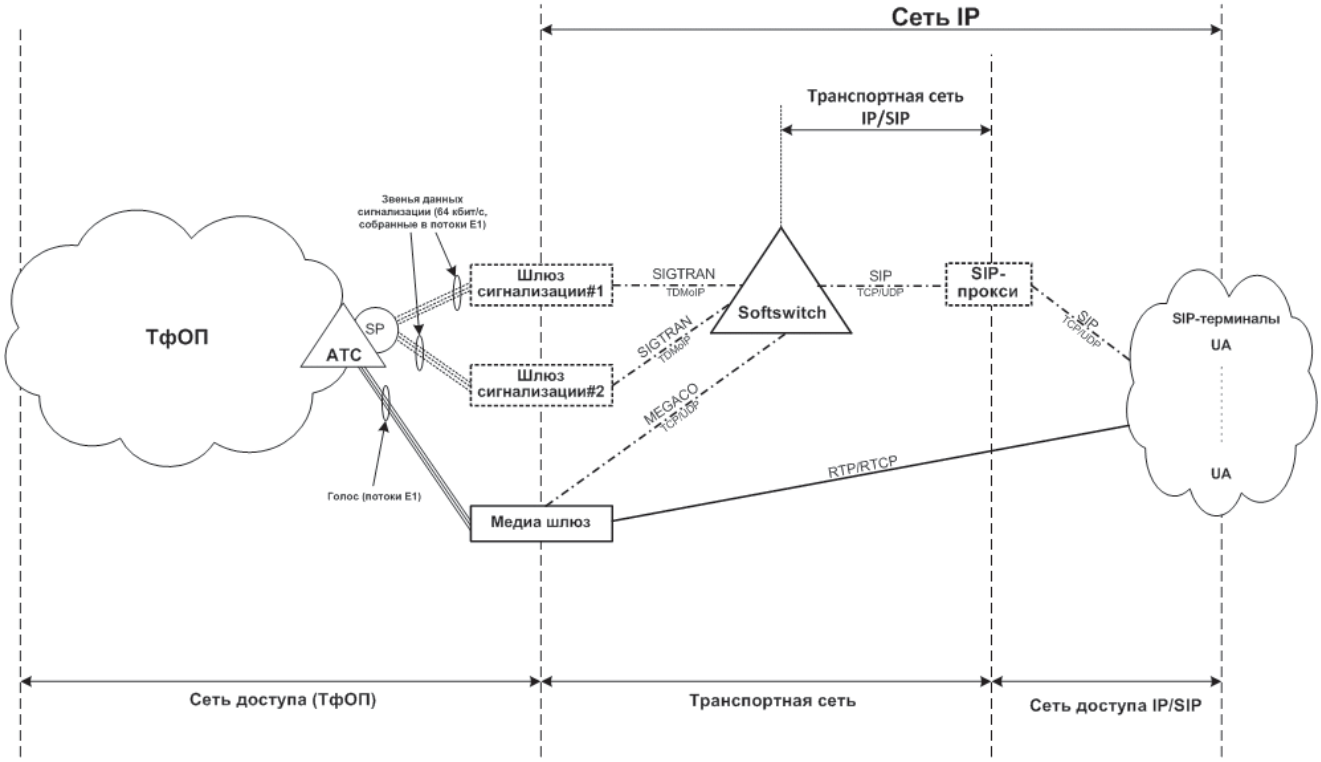


Рис. 1. Схема сети

соединения в виде открытой неоднородной экспоненциальной сети массового обслуживания (СеМО) [1, 5], в разделе 2 предложен метод расчета вероятностно-временных характеристик модели [5, 6].

1. Построение математической модели установления соединения

На рис. 2 показана диаграмма процедуры установления соединения по протоколу SIP в сети NGN между агентом пользователя сети ТфОП, инициирующим соединение, и агентом пользователя (UA) сети SIP. В работе ограничимся рассмотрением только этого случая. На основании диаграммы построим функциональную схему установления соединения, изображенную на рис. 3, в предположении, что отсутствует повторная передача сообщений. Поясним переход к функциональной схеме. Сигнальный шлюз (SGW) представлен в виде трех отдельных узлов: узел "ЦПУ", узел "ОКС-7" и узел "Sigtran". Узел "ЦПУ" моделирует работу центрального процессора шлюза, два оставшихся узла моделируют задержку до начала попадания сообщения в сеть ОКС-7 и IP. Узлы "Softswitch" и "SIP-прокси" моделируют работу одноименных узлов сети. Узел "Транспортная сеть IP/SIP" моделируют задержку передачи сообщения на участке Softswitch и SIP-прокси, Узел "Сеть доступа IP/SIP" задержку передачи сообщения на участке SIP-прокси и UA.

Установление соединения состоит из следующих этапов:

1) В центральный процессор сигнального шлюза поступает заявка IAM, что соответствует инициации нового вызова.

2) После обслуживания в центральном процессоре сигнального шлюза заявка меняет класс на IAM [Sigtran] и направляется на IP-интерфейс сигнального шлюза, где моделируется задержка передачи еепо сети до узла Softswitch.

3) Заявка IAM [Sigtran] направляется в узел Softswitch, где меняет класс на INVITE.

4) Заявка INVITE направляется в многолинейный узел "Транспортная сеть IP/SIP".

5) После обслуживания в узле "Транспортная сеть IP/SIP" заявка INVITE поступает на узел "SIP-прокси" и перед поступлением меняет класс на 100.

6) Заявка 100 обслуживается в узле "SIP-прокси" (сериализация ответа 100 Trying) и перед попаданием в узел повторно меняет класс на INVITE (передача запроса 100 по сети не моделируется, т.к. не влияет на время установления соединения).

7) Заявка INVITE передается на многолинейный узел "Сеть доступа IP/SIP" и перед поступлением меняет класс на 180.

8) Заявка 180 отправляется обратно в узел "Сеть доступа IP/SIP" (сериализация ответа 180 Ringing). Перед попаданием в узел "Сеть доступа IP/SIP" заявка 180 повторно меняет класс на 200.

9) Одновременно с этим в узле "SIP-прокси" инициируется заявка 180 и обслуживается в многолинейном узле "Транспортная сеть IP/SIP".

10) Заявка 180 возвращается в узел "SIP-прокси", где завершается ееобслуживание.

11) После этого в Softswitch инициируется заявка ACM [Sigtran], которая поступает на IP-интерфейс сигнального шлюза.

12) Заявка ACM [Sigtran] обрабатывается в ЦПУ сигнального шлюза и меняет свой класс на ACM.

13) Заявка ACM обслуживается E1-интерфейсом сигнального шлюза.

14) Заявка ACM покидает сигнальный шлюз.

15) Заявка 200 (см. п.8) заканчивает обслуживание в узле "Сеть доступа IP/SIP" (передача запроса 200 по сети) и отправляется на узел "SIP-прокси".

16) Заявка 200 обслуживается в многолинейном узле "Транспортная сеть IP/SIP".

17) Заявка 200 попадает на узел Softswitch.

18) Заявка 200 обслуживается в процессоре Softswitch, меняет свой класс на ACK и направляется на узел "Транспортная сеть IP/SIP".

19) Заявка ACK обслуживается в многолинейном узле "Транспортная сеть IP/SIP" и покидает сеть IP/SIP, соединение со стороны КП подготовлено.

20) Одновременно с отправкой сообщения ACK Softswitch генерирует заявку ANM [Sigtran], которая поступает на IP-интерфейс сигнального шлюза.

21) Заявка ANM [Sigtran] обрабатывается в ЦПУ сигнального шлюза и меняет свой класс на ANM.

22) Заявка ANM обслуживается E1- интерфейсом сигнального шлюза.

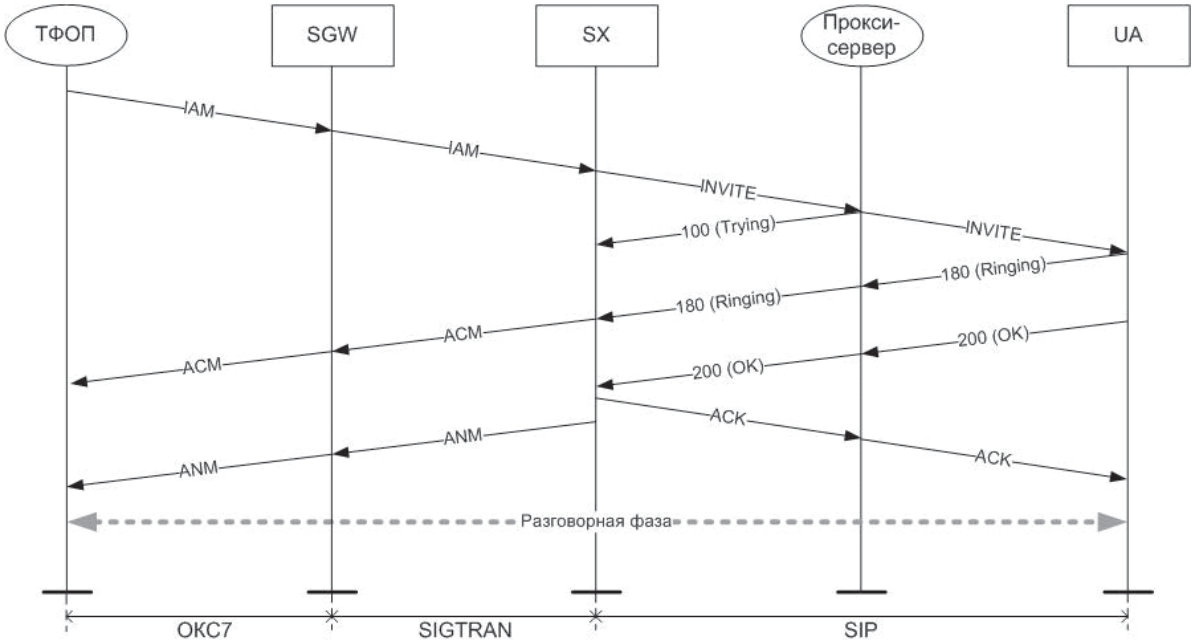


Рис. 2. Диаграмма установления соединения

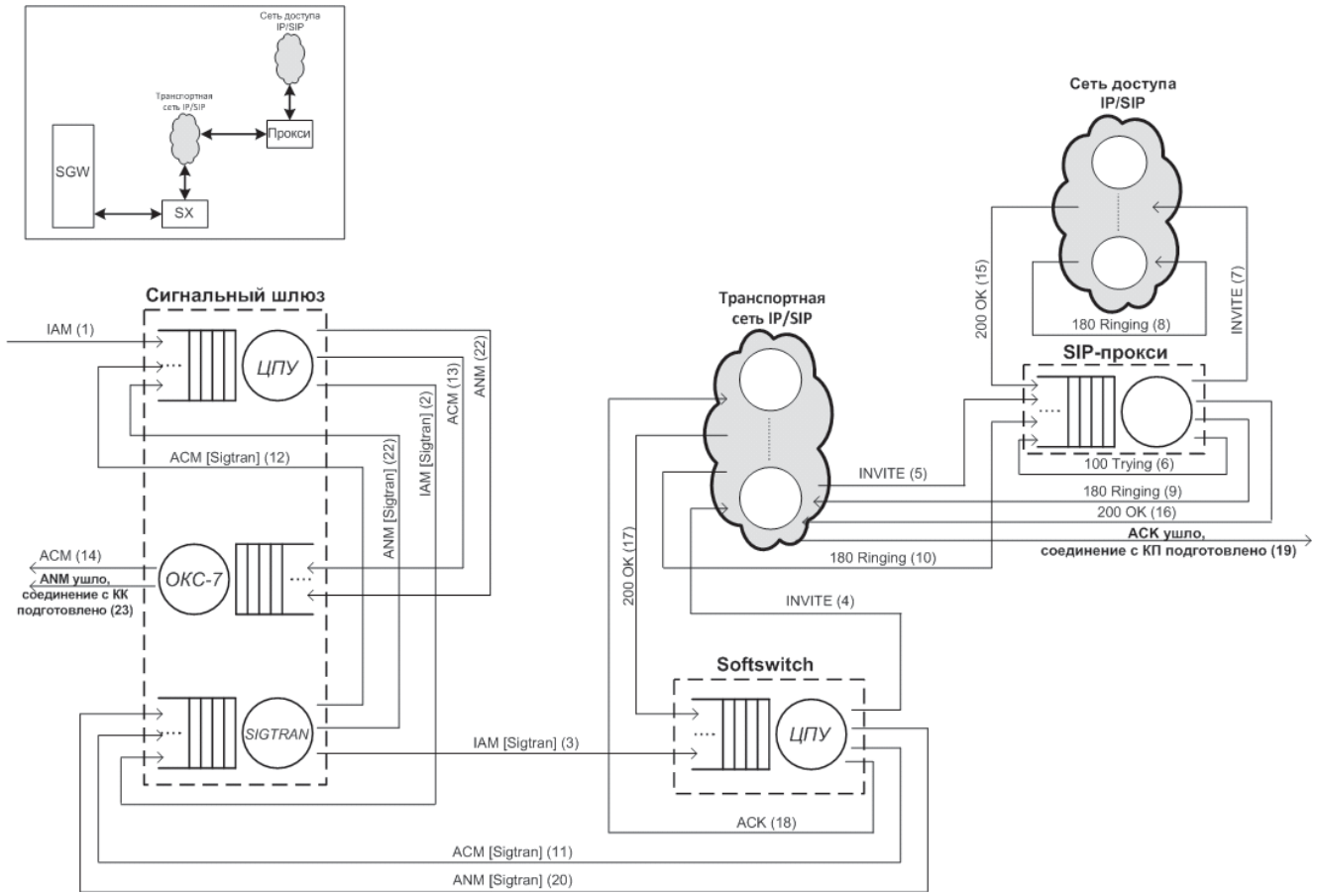


Рис. 3. Функциональная схема установления соединения

23) Заявка ANM покидает сигнальный шлюз, соединение со стороны КК подготовлено.

24) Успешное окончание фазы установления соединения

Построим математическую модель установления соединения в предположении, что события, соответствующие этапам (8-10, 15-17), а также этапам (11-14, 20-23) происходят одновременно. Тогда модель может быть описана в терминах экспоненциальной неоднородной сети массового обслуживания (CeMO), схема которой показана на рис. 4.

В этих предположениях не учитываются ответные сообщения из сети SIP в сеть ОКС 7, а именно ACM и ANM, т.е. не учитывается создаваемая ими дополнительная нагрузка на Softswitch. С целью компенсировать в модели неучтенную нагрузку, можно направить дополнительный поток заявок (сообщения ACM и ANM) на узел 4 (Softswitch), которые будут обслуживаться по цепочке узел 4 -> узел 3 -> узел 1-> узел 2 и по окончанию обслуживания на узле 2 покидать систему.

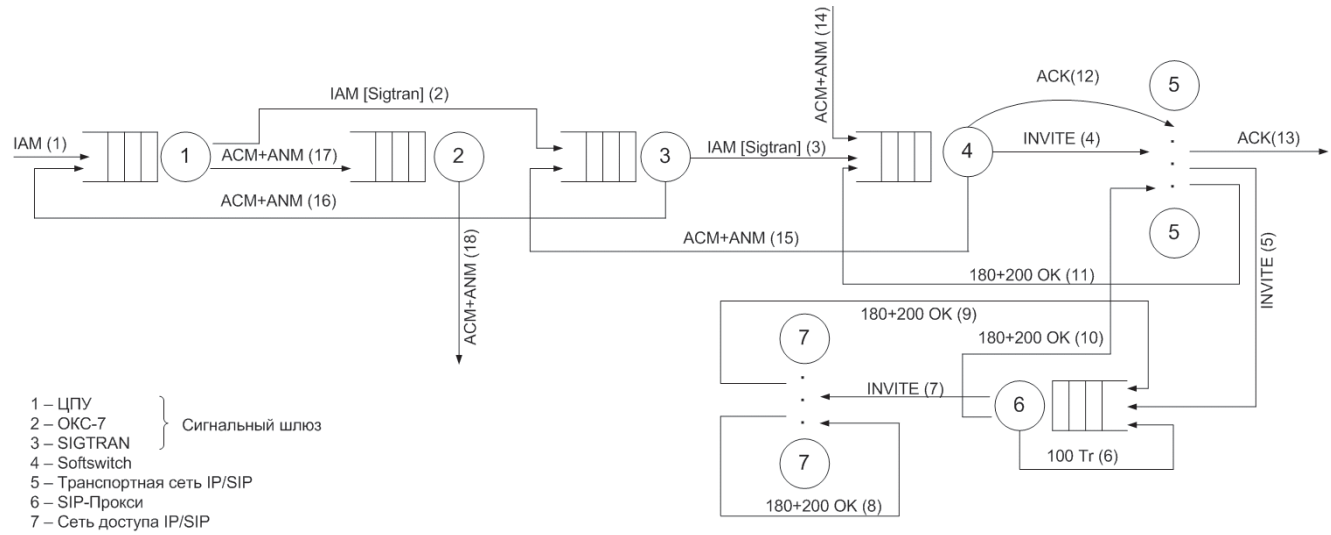


Рис. 4. Модель установления соединения в NGN

Таблица 1

Соответствие классов заявок типам сообщений

Класс заявки	Сообщение
$r_1 \in \{1\}$	IAM
$r_2 \in \{2,3\}$	IAM [Sigtran]
$r_3 \in \{4,5,7\}$	INVITE
$r_4 \in \{6\}$	100 Trying
$r_5 \in \{8,9,10,11\}$	180 Ringing, 200 OK
$r_6 \in \{12,13\}$	ACK
$r_7 \in \{14,15,16,17,18\}$	ACM, ANM

Теперь ясно, что модель установления соединения по протоколу SIP в сети NGN может быть представлена в виде BCMP-сети [5], состоящей из узлов $M = \{1,2,...,M\}$, $M = 7$ и в которой циркулируют заявки $R = \{1,2,...,R\}$, $R = 7$ классов. В табл. 1 показано соответствие классов заявок типам сообщений.

Будем называть (i,r)-заявкой заявку класса r, находящуюся в узле i и обозначим $L' = \{(i,r): i \in M,\ r \in R\}$ — множество всех допустимых типов заявок. Маршрутизацию заявок в сети описывает стохастическая матрица $\mathbf{И}' = (\theta_{ir,js}), i, j \in M,\ r, s \in R$. Для описания выхода заявок из сети введем дополнительный узел 8. Будем считать, что при переходе в узел 8 заявка меняет класс на $r_0 = 0$. Тогда $L'' = \{(8,0): i \in M,\ r \in R\}$ — дополнительное множество заявок. Переходы между множеством $L = L' \cup L''$ заявок описывает расширенная стохастическая матрица Θ, представленная в табл. 2.

2. Метод расчета вероятностно-временных характеристик

Пусть интенсивности пуассоновских потоков, поступающих на узел 1 (сообщения IAM) и на узел 4 (сообщения ACM и ANM), совпадают и равны λ. Решая уравнение равновесия для интенсивности потоков в узлах сети

$$\begin{cases} \lambda_{ir} = \sum\limits_{(j,s)\in L} \lambda_{js}\theta_{js,ir},\ (i,r) \in L \setminus \{(1,1),(4,14)\}, \\ \lambda_{1,1} = \sum\limits_{(j,s)\in L} \lambda_{js}\theta_{js,1,1} + \lambda, \\ \lambda_{4,14} = \sum\limits_{(j,s)\in L} \lambda_{js}\theta_{js,4,14} + \lambda, \end{cases} \tag{1}$$

получаем, что $\lambda_{ir} = \lambda,\ (i,r) \in L$.

Выражение для стационарных вероятностей для случая экспоненциальной сети без разделения сообщений на различные классы [5] имеет вид

$$P(\mathbf{n}) = \prod_{i=1}^{M} P_i(n_i),$$
$$P_i(n_i) = \begin{cases} (1-\rho_i)\rho_i^{n_i}, i \neq 5,7, \\ e^{\rho_i}\dfrac{\rho_i^{n_i}}{n_i!}, i = 5,7, \end{cases} \tag{2}$$

где $\rho_i = \sum\limits_{r\in R} \dfrac{\lambda_{ir}}{\mu_{ir}}$, $n = (n_1, n_2, ..., n_M)$, $n_i = (n_{i1}, n_{i2}, ..., n_{iR})$,

n_{ir} — число (i,r)-заявок, $n_i = \sum\limits_{r\in R} n_{ir}$.

Опуская для краткости промежуточные выкладки, получаем формулу для расчета среднего времени D пребывания заявки в системе

$$D = d_1 + d_3 + 2d_4 + 3d_5 + 3d_6 + 2d_7 \tag{3}$$

Время пребывания заявки в узлах можно рассчитать по формуле

$$d_i = \begin{cases} \mu_i^{-1},\ i = 5,7 \\ \dfrac{1}{\mu_i(1-\rho_i)},\ i \neq 5,7, \end{cases} \tag{4}$$

где $\rho_i = \sum\limits_{r\in R} \dfrac{\lambda_{ir}}{\mu_i}$.

Таблица 2

Матрица вероятностей переходов

$\theta_{ir,js}$	Значение индексов $(i,r),(j,s)$
1	(1, 1), (3, 2); (1, 16), (2,17);
	(2, 17), (8, 0);
	(3, 2), (4, 3); (3, 15), (1, 16);
	(4, 3), (5, 4); (4, 11), (5, 12); (4, 14), (3, 15);
	(5, 4), (6, 5); (5, 10), (4, 11); (5, 12) (8, 0);
	(6, 5), (6, 6); (6, 6), (7, 7); (6, 9), (5, 10);
	(7, 7), (7, 8); (7, 8), (6, 9);
0	в остальных случаях

Заключение

В работе выполнен анализ показателей качества обслуживания соединения по протоколу SIP в NGN в предположении экспоненциальной сети без учета класса заявок. В дальнейшем планируется получить расчетные формулы для экспоненциальной сети с учетом интенсивностей обслуживания в узле для различных классов сообщений, приближенные формулы для анализа вероятностно-временных характеристик модели для случая детерминированного обслуживания в узлах, а также разработать имитационную модель для анализа адекватности приближенных формул.

Литература

1. **Летников А.И., Пшеничников А.П., Гайдамака Ю.В., Чукарин А.В.** Системы сигнализации сетей коммутации каналов и коммутации пакетов: Уч. пособие для вузов. — М.: Изд-во МТУСИ, 2008. — 195 с.

2. **Гольдштейн А.Б., Гольдштейн Б.С.** Softswitch. — СПб.: БХВ, 2006. — 368 с.

3. ITU-T Recommendation Y.2011: General principles and general reference model for next generation networks // ITU-T, October 2004.

4. ITU-T Recommendation Y.2261: PSTN/ISDN evolution to Next Generation Networks // ITU-T, September 2006.

5. **Башарин Г. П., Толмачев А. Л.** Теория сетей массового обслуживания и ее приложения к анализу информационно-вычислительных систем // Итоги науки и техники. Теория вероятностей. Математическая статистика. Теоретическая кибернетика. — Т. 21, 1983. —120 с.

6. **Самуйлов К.Е., Лузгачев М.В., Плаксина О.Н.** Разработка вероятностной модели для анализа показателей качества протокола инициирования сеансов связи // Вестник РУДН, Серия Математика. Информатика. Физика. — Т.1. №3-4, 2007. — С.16-26.

Построение анализ марковской модели установления соединения по протоколу SIP с учетом повторных передач

Представлены результаты разработки модели установления соединения по протоколу SIP для случая, когда на транспортном уровне, для передачи сообщений применяется протокол TCP. На основании анализа процесса установления сессии по протоколу SIP поверх протоколов TCP и RTP/RTCP построена математической модель и получены формулы для расчета среднего времени установления соединения с учетом ретрансляции сообщений из-за различного рода сбоев и отказов при их передаче по IP-сети.

Нсангу М.М.,
аспирант кафедры систем телекоммуникаций РУДН,
nsangoumouchili@yandex.ru

Самуйлов К.Е.,
зав. кафедрой систем телекоммуникаций РУДН,
ksam@sci.pfu.edu.ru

Чукарин А.В.
доцент кафедры систем телекоммуникаций РУДН,
chukarin@yandex.ru

Введение

Протокол SIP (Session Initiation Protocol) — один из основных протоколов сетей связи следующего поколения (Next Generation Networks, NGN) - позволяет в полной мере осуществлять необходимые функции управления вызовами. Являясь протоколом прикладного уровня, протокол SIP поддерживает управление сеансами между абонентами NGN и обеспечивает базу для предоставления услуг связи [1,8].

Для организации услуг VoIP (Voice over IP) протокол SIP может использоваться с протоколами транспортного уровня TCP (Transmission Control Protocol) и RTP/RTCP (Real-Time Protocol/RTP Control Protocol) [10]. Протокол TCP предназначен для надежной передачи данных, а протокол RTP рекомендован IETF для использования с SIP для передачи мультимедийной информации. Протокол RTCP работает совместно с протоколом RTP и позволяет контролировать его работу с целью обеспечения качества передачи сообщений. Несмотря на широкое распространение, для протокола SIP в настоящее время не стандартизованы методики, позволяющие проводить полный анализ параметров его производительности [9].

В докладе, представлены результаты разработки модели установления соединения по протоколу SIP [1] для случая, когда на транспортном уровне, для передачи сообщений применяется протокол TCP. Анализ процесса установления сессии по протоколу SIP поверх транспортных протоколов TCP и RTP/RTCP приводит к построению его математической модели, найдены формулы для расчета вероятностно-временных характеристик (ВВХ) и выведена формула вычисления среднего времени установления соединения на уровнях протоколов SIP, TCP и RTP/RTCP с учетом ретрансляции сообщений из-за различного рода сбоев и отказов при их передаче по IP-сети. На их основе, представлены результаты численного анализа.

1. Анализ установления соединения по протоколу SIP

Схема сигнального обмена, необходимого для установления сессии по протоколу SIP, показана на рис. 1. Механизм SIP, использующийся для обмена сообщениями между шлюзами и прокси-серверами, обеспечивает высокую надежность передачи [1]. Поэтому, с точки зрения анализа возможных потерь и ошибок, наибольший интерес представляет обмен сообщениями между агентами пользователя (User Agent, UA).

В случае возникновения различных сбоев запускается механизм повторной передачи сообщения, показанный на рис. 2. При этом предлагается что механизм передачи сообщений 100-Trying и 180-Ringing является надежным [3]. Ретрансляция предусмотрена для сообщений Invite, 200-OK и ACK.

Повторная передача сообщений на уровне протокола SIP происходит с использованием таймеров [4], контролирующих процесс ретрансляции. Так, в момент передачи сообщения Invite запускаются таймер T_1 и таймер T_{INV}, ограничивающий время ожидания ответа (сообщения типа 100-Trying и 180-Ringing). При срабатывании таймера T_1 происходит повторная передача запроса Invite протоколом SIP на транспортный уровень, и таймер запускается заново, но уже с удвоенным значением. Этот процесс продолжается после каждой повторной передачи до истечения таймера T_{INV}. Аналогично

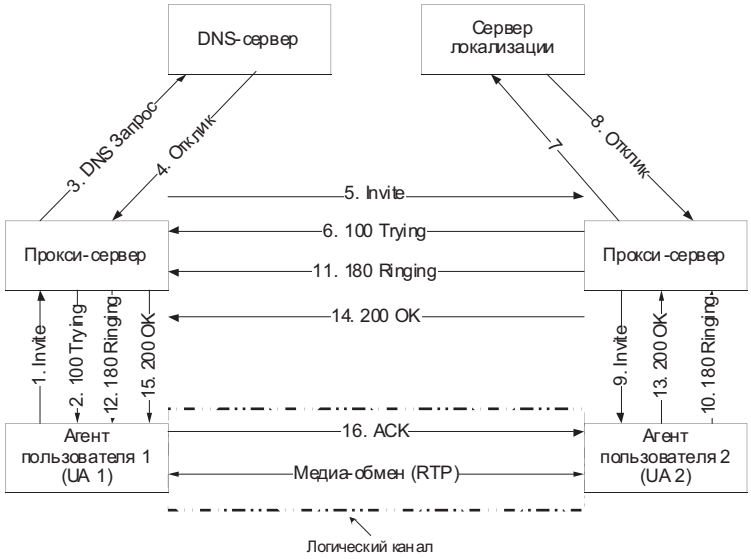


Рис. 1. Схема обмена сигнальными сообщениями

процесс повторной передачи организован и для сообщений 200-OK и ACK.

Таким образом, среднее время D_{SIP} передачи всех сообщений SIP может быть вычислено по формуле:

$$D_{SIP} = d_1 + d_2 + d_3, \quad (1)$$

где, как показано на рис. 2, величины d_1, d_2, d_3 являются средними значениями временных задержек, возникающих с учетом ретрансляций при передаче сообщений Invite, 200-OK, ACK соответственно. Тогда среднее время установления сессии D рассчитывается как сумма

$$D = D_{SIP} + D_{TCP} + D_{RTP}, \quad (2)$$

где, согласно [3], $D_{TCP} = \frac{3}{2}RTT$ — средняя задержка, возникающая в ожидании подтверждения успешной передачи сообщений SIP, D_{RTP} — средняя задержка, возникающая при установлении сессии по протоколу RTP с момента передачи SIP-терминалом сообщения ACK до момента приема первого сообщения CNAME и приклеенного к нему RTCP-пакета [5], RTT (Round Trip Time) — средняя величина сетевой задержки в петле связи.

2. Математическая модель и вероятностно-временные характеристики

Пусть величина N, число повторных передач сообщений Invite, 200-OK, ACK, по стандарту $N=7$ [4]. Обозначим g вероятность повторной передачи SIP-сообщения в прямом направлении (от UA1 к UA2), и f — вероятность повторной передачи сообщения в обратном направлении.

Тогда процесс установления сеанса описывается однородной ЦМ, диаграмма переходов которой на множестве состояний показана на рис. 3. Вероятность $P(f, g)$ успешного установления сессии по протоколу SIP с учетом повторных передач может быть вычислена по формуле:

$$P(f,g) = \alpha\left(\frac{1-g^{N+1}}{1-g}\right)^2\left(\frac{1-f^{N+1}}{1-f}\right) \quad (3)$$

где множитель $\alpha = (1-g)^2(1-f)$ равен вероятности успешной передачи сообщений Invite, 200-OK и ACK.

$$\overline{P}(f,g) = 1 - \left(1-f^{N+1}\right)\left(1-g^{N+1}\right)^2 \quad (4)$$

Если считать, что $f = g$ и начальные значения таймеров $T_1 = T_2 = T_3 = T$, среднее значение D_{SIP} задержки установления сессии с учетом повторных передач сообщений Invite, 200-OK и ACK по протоколу SIP может быть найдено по формуле

$$D_{SIP} = 3T\left(\frac{(1-g)}{(1-g^{N+1})}\frac{\left(1-(2g)^{N+1}\right)}{(1-2g)} - 1\right) \quad (5)$$

Пусть τ — интервал времени между передачей фреймов RTCP-пакета, g' — вероятность потери фрейма, k — число фреймов в RTCP-пакете, $q = 1-(1-g')^k$ — вероятность потери RTCP-пакета. Согласно [4], время передачи сообщения CNAME и RTCP-пакета по протоколу RTP может быть найдено по формуле:

$$D_{RTP} = \left[\frac{2-(1-g')^k}{(1-g')^k}\right]\frac{T_{RTCP}}{2} + \frac{RTT}{2} + (k-1)\tau \quad (6)$$

Здесь $\frac{T_{RTCP}}{2}$ — среднее время передачи RTCP-пакета; $\frac{RTT}{2}$ — время успешной передачи сообщения CNAME по IP-сети. Если предполагать, что $g' = f = g$ и $k = 1$. Тогда среднее время установления соединения по протоколу SIP поверх протоколов TCP и RTP/RTCP с учетом повторных передач равно:

$$D = 2RTT + \left[\frac{1+g}{1-g}\right]\frac{T_{RTCP}}{2} + 3T\left(\left(\left(\frac{1-g}{1-g^{N+1}}\right)\left(\frac{1-(2g)^{N+1}}{1-2g}\right)\right) - 1\right) \quad (7)$$

Пусть $g' = f = g \le 10^{-1}$, $T = 0{,}5$ с, $T_{RTCP} = 5$ с, $RTT = 0{,}425$ с, с число фреймов в RTCP-пакете $k = 1$. На рис. 4 построен график зависимости вероятности потери вызова от вероятности повторной передачи сигнального сообщения.

На рис. 5 показана зависимость времени ретрансляции SIP-сообщений от максимального числа повторных передач.

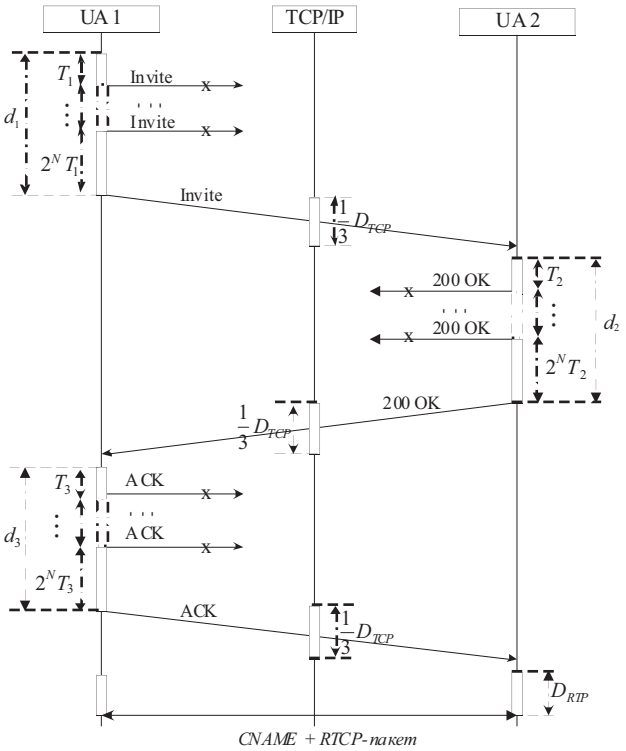


Рис. 2. Диаграмма установления соединения с учетом повторных передач

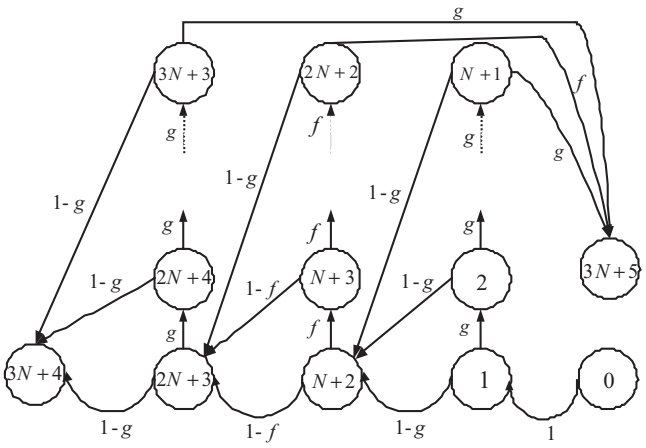


Рис. 3. Диаграмма переходов состояний модели

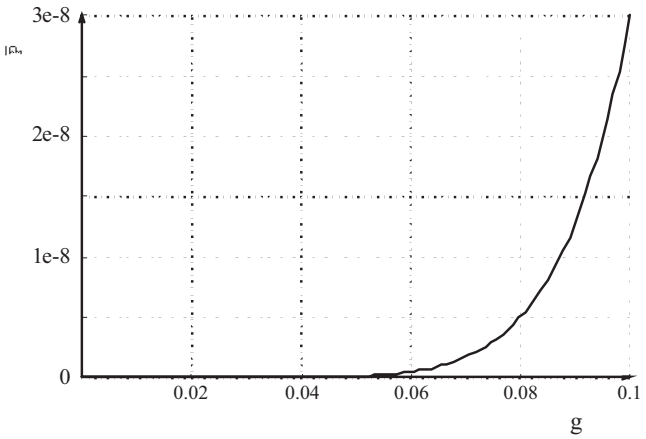


Рис. 4. Вероятность потери вызова

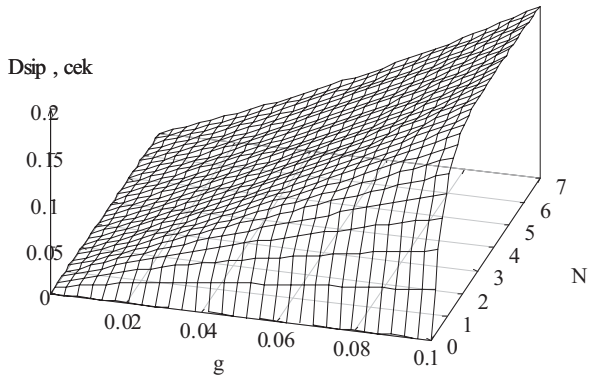


Рис. 5. Среднее время ретрансляции SIP-сообщений

Заключение

На основе результатов анализа представленной в докладе, математической модели процесса установления сессии SIP с учетом повторных передач, проверена адекватность предложенных формул, проведен численный эксперимент, результаты которого не противоречат требованиям международных стандартов к нормативным значениям задержек установления сессии по протоколу SIP [6,7]. Однако, применение надежных транспортных протоколов далеко не достаточно для гарантирования надежности передачи. Поэтому, в ряде случаев необходимо оптимизировать время установления SIP-соединения. В этой связи, существует ряд методов, в том числе, сжатие SIP-сообщений с помощью SigComp (Signaling Compression), адаптивные таймеры повторных передач и стратегии ретрансляций на канальном уровне. Последний метод наиболее эффективен и использует стеки протоколов SIP/TCP/RLP или SIP/UDP/RLP. Протокол RLP (Radio Link Protocol) передает сообщение NACK (Negative Acknowledgement)? которое потребует каналу связи ретрансляции фрейма в случае, если он был ошибочно передан. Процесс передачи NACK может осуществляться двумя способами в зависимости от того, сколько ретрансляций требует протокол RLP при каждой повторной передаче [11]. Анализ этих и других особенностей транспортировки SIP-сообщений является предметом дальнейших исследований.

Литература

1. **Гольдштейн Б.С., Зарубин А. А., Саморезов В.В.** Протокол SIP: справочник. — СПБ.: БХВ-Санкт-Петербург, 2005. — 456 с.

2. **Башарин Г.П.** Графы и цепи Маркова. Учебное пособие. — М.: Издательство УДН, 1986. — 32 с.

3. **Самуйлов К.Е., Лузгачев М.В., Плаксина О.Н.** Разработка вероятностной модели для анализа показателей качества протокола инициирования сеансов связи // Вестник РУДН "Прикладная и компьютерная математика". - №3-4, 2007. — С.53-63.

4. **De Marco G., Iacovoni G., Barolli L.** A Technique to Analyze Session Initiation Protocol Traffic. // Proceedings of the 11th International Conference on Parallel and Distributed Systems 11th. — Volume 2, 2005. — P.595-599.

5. **Das S.K., Lee E., Basul K., Kakani N., Sanjoy K.** Performance optimization of VoIP calls over wireless links using H.323 protocol. // IEEE Transactions on Computers, vol.52. — №6, 2003. — P.742-752.

6. ITU-T Recommendation Y.1541, Network Performance Objectives for IP-Based Services, 2002.

7. ITU-T Recommendation E.721, Network grade of service parameters and target values for circuit-switched services in the evolving ISDN, 1991.

8. **Летников А.И., Пшеничников А.П., Гайдамака Ю.В., Чукарин А.В.** Системы сигнализации в сетях с коммутацией каналов и пакетов. — М.: МТУСИ, 2008. —195 с.

9. **Серебренникова Н.В., Хатунцев А.Б.** Эволюция нормирования параметров качества протокола SIP // Вестник связи. — №5, 2009. — С.10-14.

10. **Schulzinne H., Casner S., Frederick R., Jacobson V.** RTP: A Transport Protocol for Real-Time Applications // RFC 1889. — January, 1996.

11. **Hanane Fathi, Shyam S. Chakraborty, Ramjee Prasad.** Voice over IP in Wireless Heterogeneous Networks: Signalling, Mobility, Security. — Springer, 2009. — 226 p.

Оценка длительности задержки телефонных сообщений от числа приоритетов в очереди и распределения времени обслуживания трафика нижних приоритетов

Проведен анализ влияния телефонного трафика на трафик нижних приоритетов и влияние трафика нижних приоритетов на характеристики обслуживания телефонного трафика. Рассматривается алгоритм приоритетной очереди (PQ-Priority Queuing) на выходном канале маршрутизатора в сетях IP. В очереди PQ телефонному трафику предоставляется наиболее приоритетная очередь, в результате чего низкоприоритетный трафик получает большую задержку. Показывается, что влияние трафика нижнего приоритета не может существенно повлиять на задержку пакетов телефонного трафика, так как полученная зависимость времени задержки пакетов высшего приоритета имеет практически линейную зависимость от загрузки каналов трафиком нижних приоритетов.

Михайлов С.К.,
Аспирант ФГУП ЦНИИС,
msk29@yandex.ru

Сергеева Т.П.,
Главный научный сотрудник НЛ-2222 ФГУП ЦНИИС, к.т.н.,
tsergeeva@zniis.ru

При обслуживании телефонного трафика в сетях NGN используется модель DiffServ с организацией приоритетного обслуживания потоков данных, отнесенных к разным классам. В докладе рассматривается алгоритм приоритетной очереди PQ (Priority Queuing), моделирующий очередность обслуживания пакетов разного класса в выходной очереди маршрутизатора. При очередности PQ обычно формируется несколько очередей, в каждой из которых хранятся пакеты с определенным приоритетом. Обслуживание последовательности очередей PQ производится в порядке уменьшения приоритета, т.е. переход к очереди более низкого приоритета производится после передачи всех пакетов (заявок) в очереди высшего приоритета. При выборе заявки на передачу используется режим "относительного приоритета", т.е. передача заявки высшего приоритета производится без прерывания передаваемого пакета низшего приоритета. В очереди PQ телефонному трафику предоставляется наивысший приоритет. Заметим, что в этом случае характеристики обслуживания телефонного трафика не зависят от количества приоритетов нижнего уровня, важен лишь параметр загрузки канала потоками низшего приоритета. Но это справедливо только для случая одинаковых параметров у потоков нижнего уровня.

Введем следующие обозначения:

ρ — коэффициент использования сервера;

λ — скорость поступления, то есть среднее количество поступающих в секунду запросов;

Ts_n — среднее время обслуживания каждого запроса приоритета n, в это время не входит время ожидания в очереди;

Tr_n — среднее время, которое запрос приоритета n проводит в системе;

ω_n — среднее количество запросов приоритета n, ожидающих обслуживания.

Для системы M/M/p/1 с двумя приоритетами получены выражения для среднего времени ожидания и средней очереди [1].

$$Tr_1 = Ts_1 + \frac{\rho_1 * Ts_1 + \rho_2 * Ts_2}{1-\rho_1} \quad (2)$$

$$Tr_2 = Ts_2 + \frac{Tq_1 - Ts_1}{1-\rho} \quad (3)$$

$$\omega_1 = \frac{\rho_1 * (\rho_1 * Ts_1 + \rho_2 * Ts_2)}{Ts_1 * (1-\rho_1)} \quad (4)$$

$$\omega_2 = \omega_1 * \frac{\lambda_2}{\lambda_1 * (1-\rho)} \quad (5)$$

$$Tq_1 = Tr_1 * \frac{Ts_2}{Ts_1} \quad (6)$$

Как известно, потоки телефонного трафика могут описываться Пуассоновским распределением. Если длительность между моментами поступления телефонных вызовов ($Tв_i$) распределены по экспоненциальному закону [1], то при пакетировании с помощью кодека G.711 телефонных сообщений пакеты следуют друг за другом с постоянным интервалом, равным времени накопления пакета в буфере шлюза. В маршрутизаторе пакеты от разных сообщений объединяются в один поток пакетов телефонии, который имеет высший приоритет. Для суммарного потока пакетов телефонии от многих источников можем считать, что в соответствии с предельной теоремой [1] суммарный поток пакетов также распределен по Пуассоновскому закону, т.к. представляет собой сумму большого числа независимых потоков с ограниченной дисперсией.

Потоки второго приоритета представляют собой потоки передачи данных трафика сети Интернет, трафика сигнализации, сообщений протоколов TCP, FTP, Telnet, видео-трафика, которые наиболее точно описываются моделями самоподобного потока [2]. Характеристики самоподобного потока демонстрируют большую степень изменчивости данных, т.е. для расчета его характеристик невозможно использовать модель M/M/1, так как значения дисперсии превосходят величину математического ожидания. При исследовании влияния потоков нижнего приоритета на трафик телефонии с высшим приоритетом в модели M/M/1 будет использовать более высокие, чем расчетные значения ρ_2, имитируя условия пакетирования самоподобного трафика низших приоритетов. Коэффициент увеличения значения ρ_2 по отношению к расчетному принимается равным 1,5, 2, 3.

Для иллюстрации проведем расчеты. Пусть величина телефонного пакета $L1$ = 160 байт (кодек G.711), величина пакета трафика второго приоритета $L2$ = 1500 байт (максимальный пакет сети Ethernet). Для канала с пропускной способностью 10 Мбит/с,

$$Ts_1 = \frac{L1*8}{V_{кан}} = 0{,}128 мс$$

$$Ts_2 = \frac{L2*8}{V_{кан}} = 1{,}2 мс$$

В одном телефонном сообщении длительностью $t_{тел}$ = 3 мин, число пакетов определим как:

$$N_{пак} = \frac{t_{тел} * 60}{L1 * 125 * 10^{-6}} = 7200 \ пак / сообщ$$

где $125*10^{-6}$ — время поступления одного байта в телефонном канале со скоростью 64 кбит/с.

При суммарной телефонной нагрузке Y = 70 Эрл, интенсивность нагрузки вызовов

$$\lambda c = \frac{Y}{t_{тел} * 60} = 0{,}39 \quad выз / сек$$

интенсивность поступления телефонных пакетов

$$\lambda_1 = \lambda c * N_{пак} = 2800 \quad пак / сек$$

Тогда значение $\rho_1 = \lambda_1 * Ts_1 = 0{,}36$.

Определим по формуле (2) зависимость среднего времени ожидания при изменении нагрузки трафика нижних приоритетов $\rho_2 = 0{,}1 - 0{,}6$ при значениях коэффициента использования для трафика 1-го приоритета $\rho_1 = 0{,}36$.

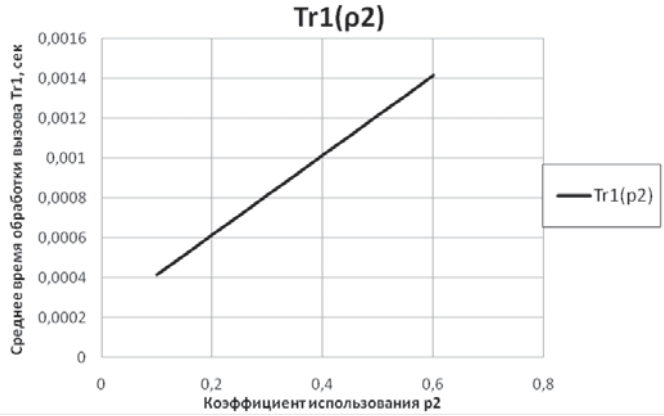


Рис. 1

Из графика на рис.1 видно, что среднее время ожидания меняется от 0,4 мс до 1,4 мс. Из этого можно заключить, что на время ожидания первого приоритета загрузка канала трафиком нижнего приоритета влияет не очень значительно, т.е. загрузка возросла в 6 раз при том, что время ожидания выросло в 3,5 раза.

В отличие от этого, зависимость времени ожидания пакетов телефонного трафика от загрузки канала собственным трафиком высшего приоритета, приведенная на рис. 2, демонстрирует очень быстрый рост времени ожидания. Загрузка увеличилась в 4 раза, а время ожидания возросло в 6 раз. Особенно важным является очень быстрый (экспоненциальный) рост времени ожидания в области ρ = 0,8 – 0,9.

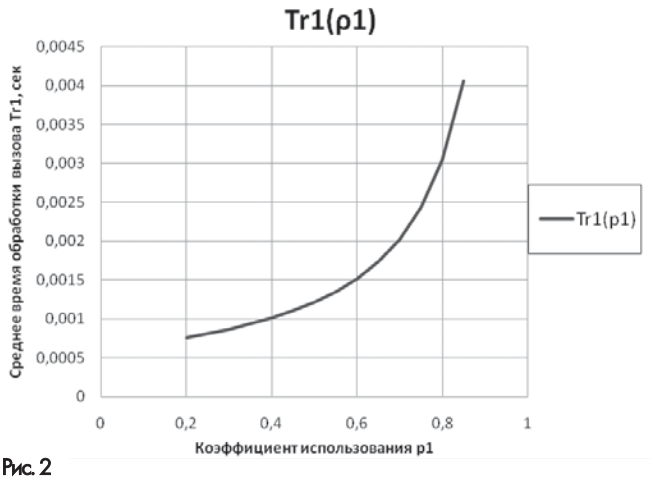


Рис. 2

Сравнение этих зависимостей позволяет сделать вывод, что колебания интенсивности трафика, или эффект самоподобия, не окажут значительного влияния на время задержки обслуживания трафика высшего приоритета в приоритетных очередях. Особенно четко этот эффект отмечается при малых размерах пакетов трафика нижнего приоритета.

Анализ выражения (2) показывает, что входящий в состав этой формулы член $\rho_2 * Ts_2$ равен $\lambda_2 * Ts_2^2$, поэтому уменьшение значения Ts_2 существенно повлияет на значение Tr_1. При длине пакета второго приоритета 300 байт значение $\rho_2 = \lambda_2 * Ts_2$ уменьшится в 5 раз. При этом среднее время ожидания в очереди высшего приоритета увеличится в 2 раза при увеличении загрузки очередью нижнего приоритета в 6 раз.

Очевидно, что время ожидания в очереди первого приоритета зависит от соотношения длины пакетов в очередях нижних приоритетов. В случае одинаковой средней длины пакетов, время ожидания в очереди высшего приоритета не зависит от количества приоритетов в очереди. Вообще, слабая (линейная) зависимость позволяет сделать вывод, что на время ожидания телефонного трафика характеристики трафика нижних приоритетов не оказывают значительного влияния, и для телефонного трафика высшего приоритета модель M/M/p/1 с приоритетной очередью незначительно отличается от простой модели M/M/1 с очередью FIFO с расчетом суммарного значения $\rho = \Sigma\rho_i$.

Литература

1. **Лившиц Б.С., Пшеничников А. П., Харкевич А.Д.** Теория телетрафика, учебник для вузов. 2-е изд., М.: Связь, 1979.

2. **Столлингс В.** Современные компьютерные сети, издание 2-е, изд. Питер, 2003.

Надежность сигнализации в МГ/МН-сети телефонной связи на базе оборудования Softswitch

Производятся численные расчеты характеристик надежности при передаче сигнальной информации в сетях междугородной и международной (МГМН) телефонной связи, построенной на базе оборудования Softswitch. Показывается, что схема с двусвязным подключением шлюзов "dual homing" обеспечивает требуемую надежность сигнализации, тогда как в сети с двухплоскостным дизайном, требующим удвоения объемов оборудования, характеристики надежности чрезмерно избыточны, вследствие чего затраты являются неоправданными.

Сергеева Т.П.,
Главный научный сотрудник НЛ-2222 ФГУП ЦНИИС, к.т.н.,
tsergeeva@zniis.ru

Михайлов С.К.,
Аспирант ФГУП ЦНИИС,
msk29@yandex.ru

В настоящее время в сетях различных операторов производится перевод трафика существующих (а также строящихся) сетей междугородной и международной телефонной связи в сети NGN. При этом основным способом перехода от телефонных сетей с коммутацией каналов (сетей TDM) к сетям с коммутацией пакетов (сетям NGN) является решение на базе оборудования Softswitch. При таком переводе трафика используется функция согласования протоколов сигнализации, обеспечивающая взаимодействие сетей коммутации каналов с IP-сетями. Обычно она выполняется входящим в состав оборудования Softswitch контроллером MGC (media gateway controller), поэтому далее мы будем употреблять термин MGC наряду с Softswitch. Схема типового решения на базе Softswitch для сетей МГМН телефонной связи (далее МГМН связи) приведена на рис. 1, где к сигнальным шлюзам SG подключаются каналы ОКС №7, а к медиашлюзам MG подключаются разговорные каналы (тракты) от цифровых телефонных станций.

В сети МГМН связи в транзитном междугородном узле (ТМгУС) и международном узле (ТМнУС) каждого Федерального округа устанавливаются маршрутизатор и контроллер шлюзов MGC, который, опираясь на протоколы сигнализации ОКС №7 (SIGTRAN), MGCP и Megaco/H.248, управляет шлюзами SG и MG, расположенными в регионах. Для передачи сигнализации между контроллерами MGC различных округов используются сигнальные сообщения протокола SIP-T, передаваемые по сети IP/MPLS между граничными маршрутизаторами.

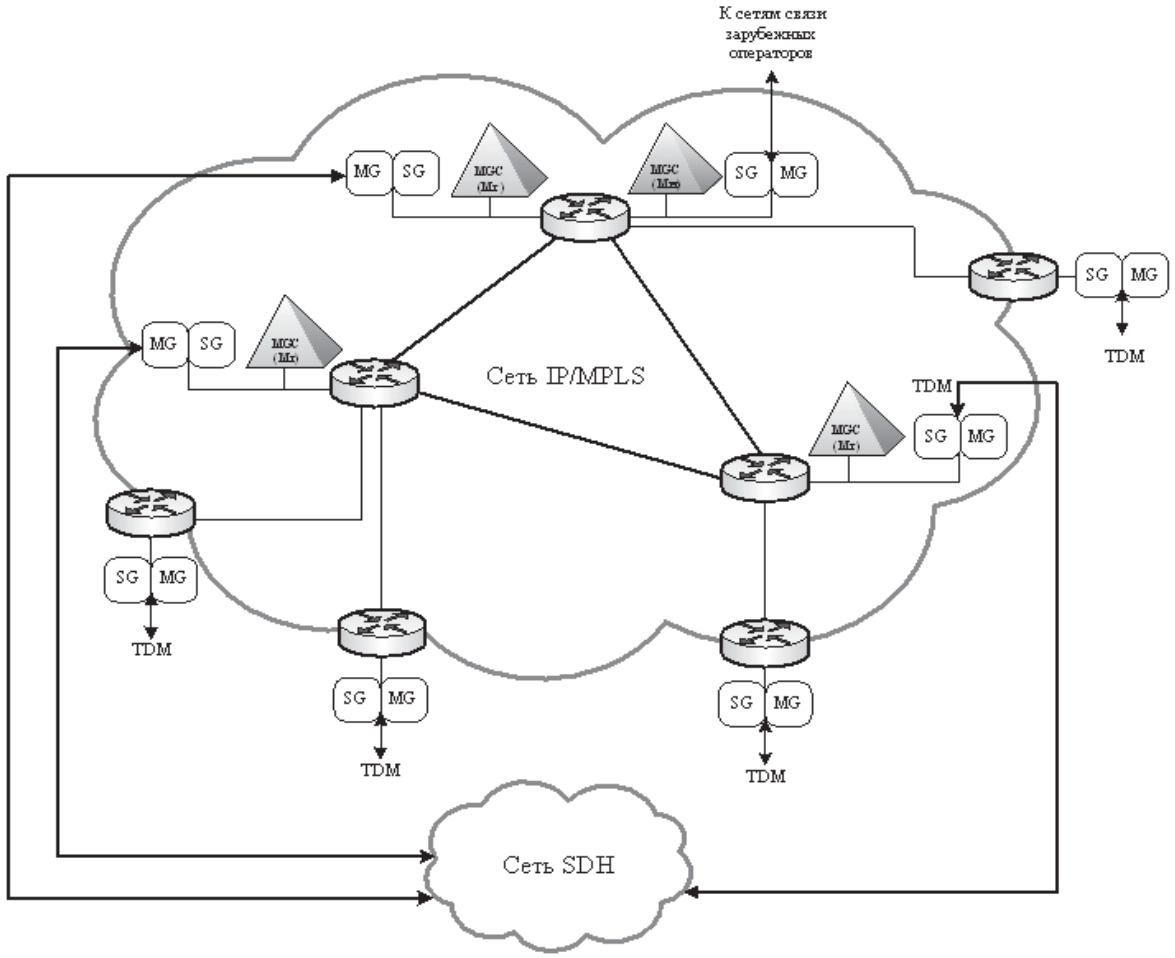


Рис. 1. Схема сети МГМН телефонной связи на базе пакетной коммутации с применением аппаратуры Softswitch

Повышению надежности системы сигнализации в сети МГМН должно быть уделено особое внимание, так как повреждение какого-либо оборудования в пути прохождения сигнальной информации приводит к прекращению связи в отдельной зоне или во всей территории Федерального округа, обслуживаемой данным MGC.

В настоящее время различные фирмы-производители оборудования, проектные организации и операторы сетей ведут дискуссии относительно способов повышения надежности пакетных сетей и, в частности, способов повышения надежности систем сигнализации. Поэтому представляет интерес определить численные характеристики надежности при различных способах резервирования.

Типовая модель для расчета надежности самого длинного пути прохождения сигнализации в сети NGN включает в себя участки (см. рис. 1):

Шлюз SG/MG — Линия от регионального узла до маршрутизатора 1 — Маршрутизатор 1- MGC1 — Линия между маршрутизаторами — маршрутизатор 2 — MGC2 — Линия от маршрутизатора до регионального узла — Шлюз SG/MG.

Реальный путь прохождения сигнальной информации несколько отличается от указанного. Так, поступающие от шлюзов к маршрутизаторам пакеты сигнализации передаются на Softswitch, и после обработки в Softswitch опять направляются в маршрутизатор, т.е. в пути сигнального сообщения возникают петли. Эти петли не должны учитываться при расчете надежности, так как каждый элемент в пути должен включаться в расчет только один раз.

Несмотря на то, что в настоящее время нормы по надежности сигнализации в сетях NGN еще не разработаны, можно ориентироваться на соответствующие нормы по надежности сигнализации по ОКС №7 в сетях TDM, так как очевидным является требование независимости качества услуг от используемой технологии. На соединение сигнализации ОКС №7 устанавливалась норма 30 мин. неготовности в год [2]. Для ее выполнения для всех каналов ОКС №7 необходимо иметь не менее двух путей, независимых по звеньям. Действительно, на рис. 2 изображена схема прохождения сигнализации ОКС №7 для вызова, проходящего от станции 1 до станции 7 через два транзитных узла 3 и 5 (треугольниками изображены транзитные узлы, а кружками — региональные телефонные узлы)

Основной канал ОКС №7 на участке "линия от регионального узла до транзитного узла" проходит по двум путям, где основной путь — по прямой связи между этими узлами, а резервный путь - через соседний транзитный узел, где организуется узел сигнализации STP. Примем значения коэффициентов готовности для всех однотипных участков длиной ***a*** (от оконечного узла до транзитного узла) равными K_a и для участков длиной ***b*** (между маршрутизаторами) равными K_b, а коэффициент готовности *n*-го узла равным K_n. Тогда на основании известных правил расчета последовательно-параллельных схем общий коэффициент готовности $K_{1\text{-}7}$ между узлами 1 и 7 может быть вычислен из выражения:

$$K_{1-7} = K_1 K_3 K_5 K_7 \left[1-(1-K_a)(1-K_2 K_a K_b)\right]\times \\ \times\left[1-(1-K_b)(1-K_4 K_b^2)\right]\left[1-(1-K_a)(1-K_6 K_a K_b)\right] \quad (1)$$

Известно [1], что для оптической линии протяженностью 13900 км нормативное значение коэффициента готовности $K_L = 0{,}985$, поэтому для *a* = км вычислим значение коэффициента готовности как $K_a = 0{,}985270/13900 = 0{,}9997$. Аналогичным образом вычисляем значение $K_b = 0{,}9983$ для *b* = 1500 км. Значение коэффициента готовности всех узлов примем равным 0,999995 (в соответствии с требованиями по надежности цифровых телефонных станций [2]). Тогда вычисленный из выражения (1) коэффициент готовности между оконечными точками 1-7 будет иметь значение $K_{1\text{-}7} = 0{,}99995$. Эта величина соответствует неготовности канала ОКС №7 30 мин. за 14 месяцев, т.е. при принятом для сетей ОКС №7 резервировании получаем результат даже лучший, чем требуемая норма неготовности 30 мин. в год.

Отметим, что если в сетях передачи данных достаточной считается готовность 99,9%, то указанное значение не отвечает требованиям телефонных сетей, которым требуется знаменитые "пять девяток" (99,999%) для готовности коммутационных узлов и 0,99995 для соединений сигнализации и телефонных соединений. При отказе от двух последних знаков после запятой, т.е. при уменьшении готовности от 99,995% до 99,9% среднее время простоя увеличивается от 30 секунд до 43 минут в месяц. Считается, что для телефонных соединений в сетях NGN необходимо получать показатели надежности не худшие, чем в телефонных сетях TDM.

По данным различных компаний-производителей оборудования Softswitch коэффициент готовности $Kr_{S\text{-}W}$ может достигать величины 0,99995 при максимальном внутреннем резервировании оборудования. Такую же величину имеет коэффициент готовности маршрутизатора $Kr_M = 0{,}99995$.

Для оценки надежности соединений в пакетных сетях рассмотрим, например, самую простую, без резервирования, схему прохождения сигнальной информации, состоящую из цепочки: телефонная станция — сигнальный шлюз — линия до маршрутизатора — маршрутизатор — Softswitch, для которой приняты следующие значения коэффициентов готовности: для телефонной станции — $Kr_{TC} = 0{,}99999$, для линии длиной *a* =270 км от станции до узла — $Kr_{Лa} = 0{,}9997$, для маршрутизатора и Softswitch значения коэффициента готовности $Kr_M = Kr_{S\text{-}W} = 0{,}99995$, значение коэффициента готовности шлюза $Kr_Ш = 0{,}99999$, так как его надежность считается не ниже надежности цифровой телефонной станции или мультиплексора SDH. Общий коэффициент готовности для данного сигнального соединения KrC, вычисляется по формуле:

$$Кг_C = Кг_{TC} \cdot Кг_Ш \cdot Кг_{Лa} \cdot Кг_M \cdot Кг_{S-W}$$

При подстановке указанных значений получаем $Kr_C = 0{,}99958$. Такой коэффициент готовности соответствует перерыву связи на 30 мин. каждый 1 месяц и 20 дней. Очевидно, что такие показатели надежности не удовлетворят практику, поэтому требуется использовать сетевое резервирование. Это предполагает, что более низкая по сравнению с телефонными станциями надежность оборудования

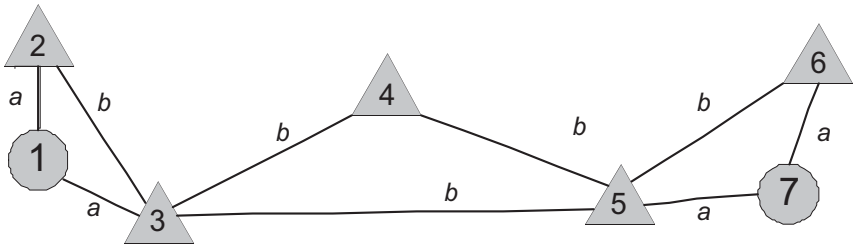


Рис. 2. Схема резервирования в системе сигнализации ОКС №7

Softswitch и маршрутизаторов должна быть компенсирована либо дублированием самого этого оборудования, либо повышением надежности передачи на линиях связи с использованием механизмов сетевого резервирования без дублирования узлов (Softswitch и маршрутизаторов).

Ряд компаний-производителей оборудования маршрутизаторов и Softswitch разработали типовые структуры резервирования в сетях NGN, которые построены по схеме так называемого двухплоскостного дизайна, где предусмотрено дублирование всех устройств, входящих в маршрут сигнального и голосового трафика. При очевидной очень высокой надежности таких схем затраты на построение сети с двухплоскостным дизайном возрастают более чем вдвое. В большинстве случаев такое удвоение оборудования является неоправданным, так как существуют более экономичные способы резервирования, которые обеспечат достаточную надежность сетей NGN. Например, в телефонных сетях коммутационные станции не дублируются, тем не менее, коэффициент готовности для междугородного телефонного соединения, проходящего через один транзитный узел с обходом через два транзитных узла составляет 0,999989, что соответствует неготовности 6 мин. в год. Такая надежность устраивает всех пользователей.

Один из основных способов повышения надежности системы сигнализации заключается в переключении или перемаршрутизации пакетного трафика от шлюзов SG/MG на резервный маршрут, если в пути сигнального сообщения обнаружен отказ любого оборудования, т.е. линии, маршрутизатора или MGC. При отказе должно обеспечиваться автоматическое переключение трафика на резервный маршрут с минимальной задержкой в обслуживании абонентов сети. Распространенным вариантом такого резервирования является использование механизма "двусвязного подключения" (dual homing) шлюзов, который в настоящее время реализован в большинстве типов оборудования шлюзов SG/MG. При этом весь сигнальный трафик с выходных портов шлюзов одновременно присутствует на двух MGC (например, своем и соседнем). Таким образом, MGC разбиваются на пары, взаимно поддерживающие возможность обработки двойного сигнального трафика. Если предполагать равным сигнальный трафик в зонах ответственности обоих MGC, то при отсутствии отказов каждый MGC обрабатывает только собственный сигнальный трафик и его производительность используется только наполовину. При отказе какого-либо MGC у соседнего MGC активизируется порты, по которым поступает трафик от шлюзов, находящихся в зоне ответственности поврежденного MGC.

Выполнение операции переключения в сети IP/MPLS производится за время порядка 50 мс, что обеспечивается специальным протоколом BFD (Bidirectinal Failure Detection), который определяет повреждение на направлении прохождения любого тракта с коммутацией по меткам и путем изменений в соответствующих таблицах продвижения переводит трафик на резервный путь. При этом удваивается загрузка исправного MGC, который, помимо собственного, обрабатывает также сигнальный трафик соседнего MGC. В настоящее время некоторые производители сообщают о создании нового оборудования контроллеров MGC с механизмом кластерного подключения шлюзов. Действительно, с учетом того, что сигнализация между шлюзами и контроллерами проходит по сети IP/MPLS с возможностью соединения по принципу "каждый с каждым" шлюзов, контроллеров, а также шлюзов с контроллерами, имеется техническая возможность управления любым шлюзом от любого контроллера. Однако при этом каждый контроллер должен иметь данные не только о своей зоне ответственности, но также данные обо всех управляемых зонах. Единственной проблемой в этом случае может являться производительность контроллеров, необходимая при подключении резервных зон. Но если учесть, что одновременное повреждение двух MGC, вероятность которого определяем как $(1-0,99995)^2 = 2,5\times10^{-9}$, является событием чрезвычайно редким, то производительность контроллеров можно рассчитывать только на случай повреждения одной резервируемой зоны.

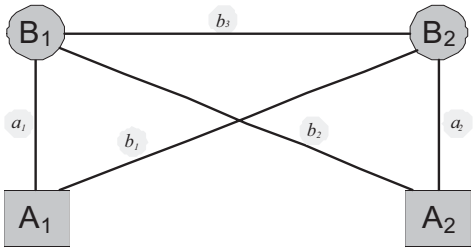


Рис. 3. Модель взаимного резервирования сигнальной информации

Сравним по надежности две схемы резервирования — а) с использованием подключения шлюзов к двум MGC (dual homing) без резервирования оборудования узлов и б) с резервированием оборудования шлюзов, маршрутизаторов и MGC (двухплоскостной дизайн).

На рис. 3 представлена модель взаимного резервирования путей сигнального трафика с помощью механизма dual homing с подключением шлюзов к двум соседним MGC, где имеется по три пути для сигнальной информации. На практике обычно применяется такое взаимное резервирование подключения шлюзов к MGC, но даже если используется резервирование с кластерным подключением шлюзов, то это не нарушает требования о наличии трех путей для сигнализации, при которых производится резервирование обоих участков линий от шлюзов до MGC.

На рис. 3 узлы A1 и A2 представляют собой объединение телефонной станции и шлюза, узлы B1 и B2 — объединение маршрутизатора и контроллера MGC. По классификации теории вероятностных графов двухполюсная сеть, изображенная на рис. 3, является неразложимой, т.е. ее коэффициент готовности не может быть определен путем разложения на последовательно-параллельные фрагменты. Для ее расчета могут быть использованы методы получения верхней и нижней оценок, например, методом Литвака-Ушакова [3]. Также существует способ получения точного результата путем перебора всех вариантов состояния схемы с выбором вариантов успешного прохождения сигнала и расчета их вероятности. В данной модели, содержащей 9 элементов, число рассматриваемых состояний равно 512. Эти состояния создают полный набор событий, сумма вероятностей который равна "1" по формуле полной вероятности. Вручную такое количество событий обработать достаточно сложно. В ФГУП ЦНИИС разработана программа, позволяющая получать точные выражения для коэффициентов готовности для неразложимых двухполюсных сетей с размерностью не более 20 узлов и линий, получаемых также после автоматического сжатия последовательно включенных элементов сети. Данная программа состоит из следующих блоков:

— построение множества всех путей между полюсами;

— отбор путей с заданными ограничениями (ранг пути, количество входящих в путь узлов разного типа), т.е. определение векторов успешного соединения полюсов;

— определение пересечения найденных векторов с множеством состояний рассматриваемой схемы;

— составление формулы для расчета Кг двухполюсной сети;

— получение численных результатов расчетов по составленной формуле на основе введенных коэффициентов готовности элементов сети.

Вычисление векторов путей в двухполюсной сети между узлами A1 и A2 согласно рис. 3 показывает, что имеется три пути, т.е. три вектора, в которых все входящие в него элементы должны иметь значение "1" (т.е. находиться в исправном состоянии). При пересечении этих векторов (с равными "1" фиксированными значениями элементов) с множеством состояний остальные элементы, дополняющие каждый вектор до 9 элементов, могут иметь произвольное значение, т.е. являются свободными. Значения векторов путей следующие:

$$A1 \cdot A2 \cdot B1 \cdot a_1 \cdot b_2; \quad A1 \cdot A2 \cdot B1 \cdot B2 \cdot b_3 \cdot a_2 \cdot a_1;$$
$$A1 \cdot A2 \cdot B2 \cdot a_2 \cdot b_1.$$

При фиксированном значении 5-ти элементов первого вектора количество состояний в пересечении этого вектора с множеством всех состояний составляет 16 (так как в каждый вектор не входят 4 элемента из 9-ти, которые могут принимать любые значения, т.е. всего иметь 2^4 значений), пересечение третьего вектора также имеет 16 состояний, а пересечение второго вектора, состоящего из семи элементов, имеет четыре свободных состояния. Таким образом, имеется всего 39 состояний из 512, которые приводят к благоприятному исходу прохождения сигнализации между полюсами A1 и A2 двухполюсной сети. Рассчитанный при прежних параметрах по описанной выше программе коэффициент готовности двухполюсной сети равен 0,999986. Этот коэффициент готовности соответствует времени неготовности 30 мин в 4 года, т.е. надежность достаточна и существенно выше требуемых показателей. Таким образом, структура сети с резервным подключением шлюзов к двум комплектам оборудования Softswitch обеспечивает достаточный уровень надежности при передаче сигнальных сообщений.

Рассмотрим второй вариант. Схема сети с двухплоскостным дизайном приведена на рис. 4.

При упрощенном описании модели (без изображенных пунктиром линий, которые в путях сигнализации не используются) количество элементов в этой схеме составляет 36, а количество основных путей — 16. При описании данной схемы невозможно укрупнить элементы и состояния, так как распараллеливание в данной схеме производится на всех уровнях, начиная с первых элементов. Схему с таким числом состояний не удается рассчитать автоматически в связи с ограничениями вышеупомянутой программы по размерности.

В [3] указывается, что наиболее простыми и точными являются оценки Литвака-Ушакова (ОЛУ). Для получения пригодной для практики нижней ОЛУ используется сеть, представляющая собой параллельное соединение непересекающихся (т.е. не имеющих общих ребер) простых (без циклов) путей, в каждом из которых все ребра данного пути соединяются последовательно. Если рассматривать модель, изображенную на рис. 4, без пунктирных линий, то в нижнюю оценку входят два параллельно соединенных пути, для которых коэффициент готовности составляет 0,9999973, соответствующий неготовности 30 мин за 21 год и 5 мес. При задействовании также изображенных пунктиром линий получаем 4 параллельных пути с общим коэффициентом готовности 0,999999999993. С учетом того, что нижняя ОЛУ, по определению, является достаточно слабой, реальный коэффициент готовности схемы, изображенной на рис.4, будет еще выше.

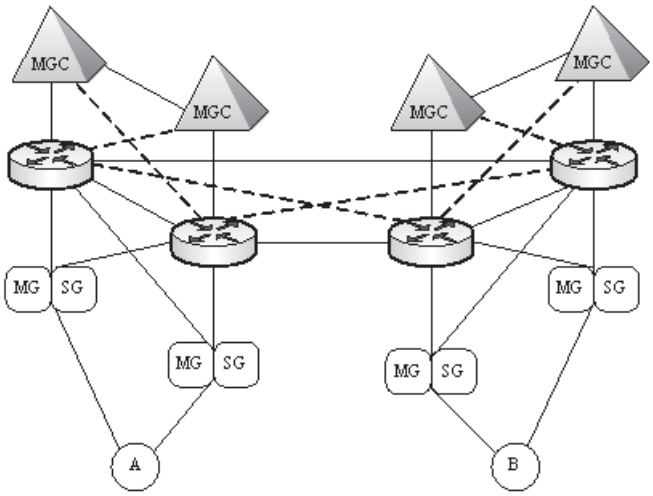


Рис. 4. Модель двухплоскостного дизайна на фрагменте сети МГМН телефонной связи

Из этих результатов следует вывод, что надежность схемы с двухплоскостным дизайном является чрезмерно избыточной, вследствие чего дополнительные затраты на удваивание оборудования узлов и линий при двухплоскостном дизайне в сетях МГМН на базе оборудования Softswitch являются неоправданными.

Литература

1. Правила проектирования, строительства и эксплуатации волоконно-оптических линий передачи по ВЛ напряжения более 110 кВ. Минэнерго России. — М.,1999.

2. Основные положения развития Взаимоувязанной сети связи Российской Федерации на перспективу до 2005 года, руководящий документ, книга 3, Основные положения развития телефонной сети связи общего пользования. — Москва, 1996.

3. **Нетес В.А.** Об оценке вероятности связности двухполюсных сетей // Электросвязь, 2001. — № 11. — С.39-41.

О проблеме перегрузок в мультимедийных IP-ориентированных сетях

Будущее сетей операторов связи невозможно представить без подсистемы IMS (IP Multimedia Subsystem). Активный переход к полностью IP-ориентированным сетям для предоставления мультимедийных услуг, включая голосовые услуги, ставит новые вопросы о возможных перегрузках в сетях связи и необходимости анализа возникновения перегрузок и способов их предупреждения. Сигнальный трафик в IP-ориентированных сетях порождаемый протоколом SIP (Session Initiation Protocol) при установлении, модификации и разъединении сеансов связи, существенно отличается от сигнального трафика ОКС 7 в традиционных сетях связи. В докладе показана необходимость исследования перегрузок в сетях, построенных на базе подсистемы IMS.

Реймер Д.А.,
Директор Департамента интегрированных бизнес решений, ЛАНИТ,
reimer@lanit.ru

Введение

Бурный рост интереса к IP-ориентированным сетям и активный переход операторов связи к предоставлению мультимедийных услуг на базе подсистемы IMS (IP Multimedia Subsystem) открывает перед операторами широкие возможности по конвергенции технологий и способов предоставления услуг. Использование пакетной коммутации и протокола IP для передачи и данных и голоса позволяет существенно расширить спектр услуг с одной стороны, но в то же время ставит новые задачи по обеспечению требуемого качества предоставляемых услуг. Предоставление услуг с требуемым уровнем качества в мультимедийных IP-ориентированных возможно при отсутствии блокировок в сети или выполнении автоматических процедур с использованием механизмов управления перегрузками.

Мультимедийные IP-ориентированые сети

Идея построения мобильной сети на базе IP технологий прорабатывалась консорциумом 3GPP (3rd Generation Partnership Project) [1] для сетей мобильной связи. Первая спецификация 3GPP появилась в 1999 году и носила название 3GPP Release 99 [2]. В ней впервые были введены классы качества обслуживания QoS (Quality of Service) и намечен революционный переход от стандартизации услуг к стандартизации возможностей предоставления услуг (Services Capabilities). После этого в 3GPP начали разрабатывать концепцию мобильной сети, полностью базирующейся на протоколе IP, но в силу трудоемкости задачи, ее решение было разбито на Release 4[2] и Release 5[2]. Архитектура IMS, полностью базирующаяся на IP впервые была представлена в Release 5. Предполагалось, что она должна стать независимой от технологии доступа и полностью базирующейся на IP сетевой архитектурой. В Release 6[2] в архитектуру IMS были введены изменения, а также появилась поддержка Wireless LAN сетей. Благодаря работе группы TISPAN, Release 7[2] 3GPP добавил поддержку фиксированных сетей и IMS стала основой и для сетей фиксированной связи следующего поколения (NGN, Next Generation Networks).

Основными функциональными элементами IMS являются модуль, реализующий функцию управления сеансами связи CSCF (Call Session Control Function) и сервер домашних абонентов HSS (Home Subscriber Server), фактически являющийся базой пользовательских данных.

В состав модуля CSCF включены три функции — функция обслуживания (S-CSCF, Serving CSCF), функция запроса (Interrogating CSCF, I-CSCF) и функция прокси-сервера (P-CSCF, Proxy CSCF). Функция S-CSCF предназначена для регистрации абонентов в базе данных HSS, а также осуществляет контроль и управление сеансами связи. Функция I-CSCF реализует доступ в домашнюю сеть и, в случае необходимости, может выполнять функции сетевого экрана, скрывая внутреннюю топологию сети. Функция P-CSCF обеспечивает первую точку контакта для терминала IMS, а также может выполнять шифрование и сжатие сообщений. Важно, что в общем случае при управлении сеансами связи функциональные узлы IMS (далее, для краткости — узлы IMS) обмениваются сигнальными сообщениями и создают сигнальный трафик, который должен обслуживаться с заданными параметрами качества, которые могут определяться требованиями стандартов, а также соглашениями об уровне обслуживания (SLA, Service Level Agreement). Узлы CSCF обмениваются между собой сообщениями протокола SIP, а для обращения к серверу HSS используется протокол Diameter. Терминалы абонентов (UE, User Equipment) соединяются посредством узлов P-CSCF тех сетей, в которых они на данный момент находятся и через эти функции взаимодействуют с другими узлами IMS, находящимися в их домашних сетях. При регистрации абонента в некоторой сети, функция P-CSCF этой сети определяет домашнюю сеть абонента и соответствующую функцию S-CSCF, в которой он обслуживается.

Особенности механизма управления перегрузками сигнального трафика

Важным звеном в подсистеме IMS является протокол SIP (Session Initiation Protocol) [3], выполняющий роль основного протокола сигнализации, который используется для установления, модификации и разъединения сеансов связи, аналогично ОКС 7 для сети GSM и для интеллектуальной сети связи.

Протокол SIP позволяет устанавливать сессии, передавая согласованные между инициаторами соединения параметры и совместимые типы медиаданных. SIP использует так называемые прокси-серверы для более эффективной маршрутизации запросов текущего местонахождения пользователя, его аутентификации и авторизации и других различных услуг, а также осуществляет функции маршрутизации, регистрации пользователей и др. Существует множество приложений, требующих установления и управления соединениями, где сессия представляет собой обмен данными между некоторым количеством участников. Внедрение подобных приложений осложняется поведением пользователей, которые могут перемещаться между оконечными устройствами, иметь несколько адресных имен, одновременно использовать различные мультимедиа сервисы. Некоторые протоколы могут осуществлять передачу разных типов данных в режиме реального времени, таких, как голос, видео, текстовые сообщения. SIP работает в соответствии с концепцией таких протоко-

лов, устанавливая оконечные точки (называемые пользовательскими агентами UA) для установления местонахождения друг друга, а также для согласования параметров соединения. Для локализации возможных участников соединения, SIP поддерживает создание инфраструктуры сетевых хостов (прокси-серверов), на которые пользовательские агенты UA отправляют сообщения о регистрации, запросы на установление сессий. Таким образом, SIP является гибким многофункциональным инструментом создания, управления и завершения сессий, работающим независимо от транспортных протоколов более низких уровней, а также независимо от типа устанавливаемого соединения.

Несмотря на популярность и увеличение количества внедрений в промышленной эксплуатации механизм управления перегрузками в протоколе SIP малоизучен. Хотя протокол SIP и обладает базовым механизмом управления перегрузками, но на практике этот алгоритм оказался малоэффективным и часто не способным предотвратить перегрузку SIP серверов. Проблема становится еще более актуальной в настоящее время, т.к. внедрения становятся более масштабными и охватывают все большее число пользователей.

Перегрузка на SIP серверах возникает, если интенсивность поступления сообщений превышает возможности по обработке сообщений. При повышении нагрузки производительность сервера существенно падает. Кроме этого SIP сервер генерирует внутренние сообщения, например, для повторной передачи сообщения, если не был получен ответ на предыдущее сообщение в требуемое время. Это приводит к тому, что перегрузка сервера может возникнуть и в случае, когда интенсивность поступления сообщений меньше, чем скорость обработки сообщений. Таким образом, перегрузка на одном сервере может повлечь за собой перегрузки на соседних серверах и привести к перегрузке всей сети.

Рассмотрим механизм возникновения блокировок более подробно. Для обеспечения надежных соединений SIP-сообщения должны передаваться поверх протокола TCP, однако, иногда используется ненадежное UDP-соединение в зависимости от технических характеристик оконечного оборудования. В этом случае для того, чтобы компенсировать ненадежность передачи SIP-сообщений, SIP-запросы передаются вновь, если не был получен адекватный ответ в течение заданного промежутка времени. Хотя пересылки повышают надежность, они вызывают перегрузки сигнальной сети SIP.

На Рис.1 представлена типичная сетевая структура протокола SIP. Перед установлением сессии между вызывающим (пользователь-A (user A) на Рис.1) и вызываемым (пользователь-B (user B) на Рис.1) их пользовательские агенты (UA) обмениваются информацией, необходимой для установления сессии посредством SIP-сигнализации, которая осуществляется посредством отправки запросов и ответов через прокси-серверы SIP, причем пути доставки таких запросов независимы от пути установления сессии. На Рис.1 обмен сигнальными SIP-сообщениями происходит между соседними компонентами сети 1, 2, 3. Присваивая адрес SIP URI, каждый прокси-сервер SIP выполняет маршрутизацию запросов и ответов протокола SIP.

На рис.2 представлена типичная временная диаграмма установления SIP-сессии. Пользователь-A запрашивает пользователя-B, используя его уникальный идентификатор SIP URI, посылая ему первоначально сообщения типа INVITE. Узел, получивший такое сообщение, возвращает предварительный ответ 100 Trying, подтверждающий получение INVITE и содержащий параметры для продолжения установления сессии, и в случае успеха пользователь-B возвращает 180 Ringing. Когда пользователь-B отвечает, отправляя сообщение типа 200 OK, и пользователь-A получает его, он немедленно отправляет обратно подтверждение установления сессии ACK. Как только это произошло, оба пользователя обмениваются различными мультимедиа пакетами, и по окончании вызова сторона-инициатор отправляет сообщение типа BYE для завершения сессии.

Базовый механизм управления перегрузками в SIP протоколе использует код ошибки 503 (Service Unavailable). В случае если SIP сервер не может обработать запрос из-за временной перегрузки, то он отклоняет запрос с кодом ошибки 503. Перегруженный сервер может добавить Retry-After заголовок в SIP сообщение об ошибке, указав время в секундах, в течение которого сервер не хочет получать каких-либо сообщений от отправителя. Отправитель, прекращает отправку сообщений на этот период времени и возобновляет после его окончания.

Кроме этого, протокол SIP предусматривает два типа процедур повторной отправки сообщений [4]: сообщения INVITE и остальные (200 OK, BYE, ACK), причем время отправки T1 определено в документе RFC 3261[3]. Пересылка сообщений INVITE со стороны клиента происходит с первоначальным интервалом в T1 секунд (значение T1 по умолчанию равно 500ms), который затем удваивается после каж-

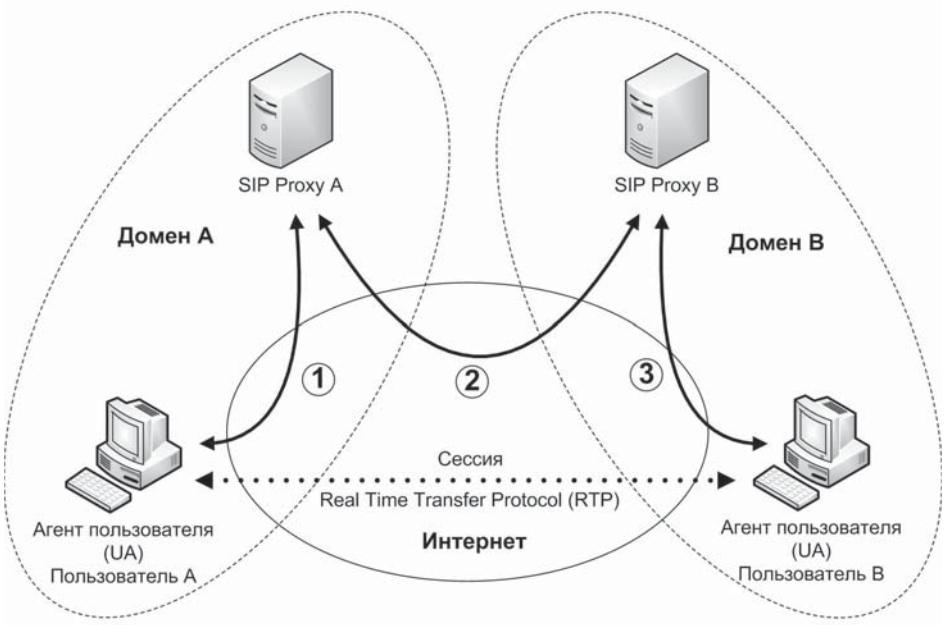


Рис.1. Сетевая структура протокола SIP

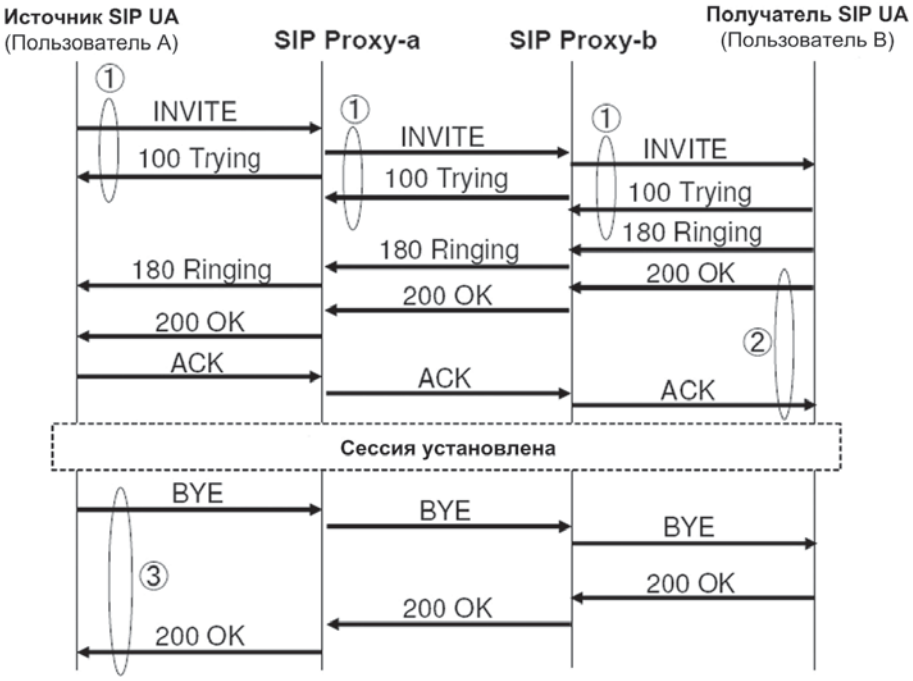


Рис. 2. Временная диаграмма обмена SIP-сообщениями

дой передачи пакетов, и прекращается в случае, когда получен требуемый ответ или когда с момента первой отправки прошло более 64?T1, то есть 32 секунды. Повторная пересылка сообщений, отличных от INVITE (200 OK и BYE), происходит иначе. В RFC 3261 задан интервал T2: запросы вновь отправляются через время T1, которое удваивается для каждого следующего пересылаемого пакета, пока не достигает значения T2 (по умолчанию равно 4 секундам). В соответствии со стандартом RFC 3261, сообщения типа INVITE пересылаются 7 раз, а 200 OK и BYE — 11 раз в общем итоге, что может негативно влиять на производительность сигнальной сети в целом.

Примеры простейших моделей управления перегрузками

Рассмотрим структуру сигнальной SIP-сети на рис. 3. Источниками являются пользовательские SIP-агенты (UA), показанные квадратиками в левой части рис. 3, соединенными с роутерами, обозначенными кружками. Прокси-серверы SIP, обозначенные квадратиками серого цвета, также соединены с роутерами. Области, окруженные пунктирными линиями, составляют домен, причем каждый 0 ~ n-1 домен содержит *m* пользовательских агентов UA и один прокси-сервер SIP. Домен *n* включает в себя *n* x *m* получателей-пользовательских агентов SIP (обозначенных в правой части рис. 3 небольшими квадратами) и один прокси-сервер SIP. Роутеры 0 ~ *n*-1 соединены с роутером *n*, а пути маршрутизации сессий (обозначены на рис. 3 пунктирными стрелками) независимы от путей рассылки SIP-сообщений. Все *n* x *m* пар источников и приемников UA могут устанавливать сессии, и если каждый приемник находится в домене n, то все SIP-сообщения направляются на прокси-сервер SIP, что в таком случае может привести к образованию пробки на сервере.

Управление перегрузками может быть как локальным, применяемым на конкретном SIP-сервере, так и распределенным, включающих несколько SIP-серверов.

Первый алгоритм локального управления — это простой алгоритм bang-bang control (BBC)[4]. В рамках аппарата теории массового обслуживания система поступления SIP-сообщений на сервер — это система с очередью заданной емкости *B*, с дисциплиной обслуживания FIFO, причем время обслуживания таких сообщений на сервере зависит от их типа. Обычно сообщения типа INVITE требуют больше времени, чем остальные, поскольку они предполагают также предварительный запрос SIP URI в базе данных [3,4]. Введем два пороговых значения наполнения очереди *h* и *l* с целью обнаружения перегрузок: если в очереди находится более *h* заявок, то прокси-сервер SIP переходит в состояние перегрузки, впоследствии, если число заявок в буфере становится меньше *l*, сервер считает перегрузку ликвидированной. Таким образом, сервер может находиться в двух состояниях: рабочем "normal" и состоянии перегрузки "congestion", и переходы из одного в другого можно описать следующими неравенствами. Если x — количество заявок в очереди, то в стационарном режиме данной СМО сервер переходит в состояние "congestion" при $x > h$ и возвращается об-

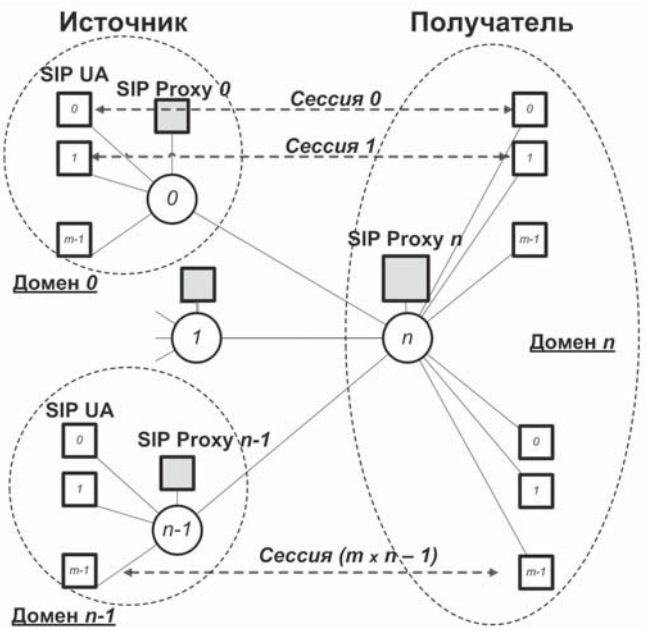


Рис. 3. Модель сигнальной сети на базе протокола SIP

ратно при $x < l$. Когда сервер находится в состоянии перегрузки, сигнальная сеть SIP регулирует поступление новых вызовов, и если в обычном случае прокси-сервер SIP отправляет "100" в ответ на запрос "INVITE", то в данном — возвращает "503", что означает, что сервис недоступен. Согласно RFC 3261, когда источник SIP UA (клиент) получает ответ с кодом "300-699", он должен в течение времени *D* (равного 32 секундам по умолчанию) оставаться в состоянии, когда он не может посылать сообщения "INVITE". Регулируя поступление новых вызовов изменением значения *D*, можно уменьшить или временно ликвидировать нагрузку на сеть.

Второй алгоритм локального управления называется occupancy (OCC)[5], в котором входящие INVITE запросы принимаются с вероятность *f* (или отклоняются с вероятностью 1-*f*). Алгоритм OOC базируется на том, что нагрузка (занятость) на сервер равна ρ. Целью алгоритма является динамическое регулирование значения *f* для поддержания коэффициента нагрузки на уровне или ниже полученного целевого значения ρ_{targ}. Нагрузка ρ периодически обновляется каждые τ секунд.

Каждые *t-th* моменты времени коэффициент нагрузки ρ_t обновляется и сравнивается с ρ_{targ}. Основная идея алгоритма состоит в увеличении значения *f*, если $\rho < \rho_{targ}$, и уменьшении в противном случае.

Разумное управление ретрансляциями ответов может позволить предотвратить отправку лишних копий потерянных сообщений и таким образом увеличить поступающую нагрузку.

Вместо просмотра сообщений, локальное управление перегрузками позволяет сдерживать ретрансляции при отказе в обслуживании. Кроме того, когда сервер может отклонить индивидуально каждый запрос, он может легко выбрать и группу запросов для отклонения, тем самым "сглаживать" контроль обслуживания поступающих запросов. При этом локальное управление перегрузками не требует объединения серверов.

В свою очередь распределенные механизмы управления перегрузками позволяют загруженному серверу избавиться от необходимости ретрансляции запросов на другие сервера по маршруту.

При распределенном управлении перегрузками SIP-сервера задействованы в обеспечении обратной связи (для контроля перегрузок) с серверами, которые располагаются дальше в сети в направлении пересылки сообщений. Обратная связь для контроля перегрузки может осуществляться, например, за счет передачи информации в заголовках SIP-ответов [6]. Цикл обратной связи для контроля перегрузки может быть применен как между парами соседних SIP серверов на всем маршруте (Hop-by-hop), так и централизованно, обеспечивая сквозное управление на всем маршруте (End-to-end).

Неэффективность работы механизма управления перегрузками в SIP протоколе исследована в работе [7], в которой было проведено имитационное моделирование на нескольких топологиях сети с использованием представленных алгоритмов BBC (Bang-Bang Control) и OCC (Occupancy Algorithm) в различных моделях контроля перегрузками, в результате которого показано, что использование механизма ретрансляции различными SIP-таймерами приводит как к перегрузкам серверов, так и может вызвать лавинную перегрузку во всей сети.

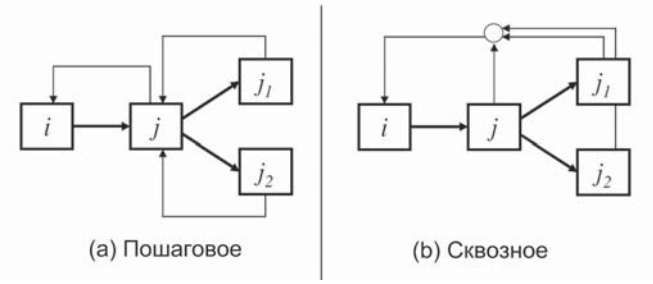


Рис. 4. Модели распределенного управления перегрузками

Необходимо дальнейшее более детальное изучение моделей сигнального SIP-трафика в масштабе всей сети. Распределенные алгоритмы управления являются более эффективными для предотвращения перегрузки всей сети, но и существенно более сложными и трудоемкими по сравнению с алгоритмами локального управления.

Заключение

Увеличение масштабов внедрения мультимедийных IP-ориентированных сетей и бурный рост числа пользователей ставит новые вопросы в исследовании механизмов сигнализации на базе протокола SIP. Существующие механизмы управления перегрузками в протоколе SIP являются неэффективными и требуют дополнительной проработки и, возможно, развития и доработок в протоколе. Исследование моделей поведения пользователей и серверов в сети IMS и более глубокий анализ работы механизмов сигнализации в крупных IMS сетях позволит повысить стабильность работы сети и обеспечить требуемый уровень качества услуг для пользователей.

Литература

1. 3rd Generation Partnership Project, http://www.3gpp.org.

2. 3GPP Release 99, 4-7, http://www.3gpp.org/releases.

3. **J. Rosenberg.** "SIP: Session Initiation Protocol", RFC 3261, http://www.ietf.org/rfc/rfc3261.txt, 269 pp.

4. **Masataka Ohta.** "Overload Control in a SIP Signaling Network", PWASET VOLUME 12 march 2006 ISSN 1307-6884, 205-210 pp.

5. **B. L. Cyr, J. S. Kaufman and P. T. Lee.** "Load balancing and overload control in a distributed processing telecommunication systems," United States Patent No. 4,974,256, 1990.

6. **V. Hilt, I. Widjaja and H. Schulzrinne.** "Session Initiation Protocol (SIP) Overload Control," IETF, Internet-Draft, draft-hilt-sipping-overload, 2008, 21 pp.

7. **V. Hilt and Indra Widjaja.** "Controlling Overload in Networks of SIP Servers",IEEE, 2008, 83-93 pp.

Методика получения нижних оценок коэффициента готовности тракта в системе взаимосвязанных кольцевых структур SDH

Расчет коэффициента готовности для тракта, проходящего по системе взаимосвязанных кольцевых структур, в общем случае затруднен в связи с необходимостью учёта зависимых элементов. В статье предлагается метод нижней оценки коэффициента готовности для тракта, проходящего по сети SDH, использующей кольцевые схемы резервирования. Погрешность метода по сравнению с точным значением коэффициента готовности не превышает 10^{-8}.

Баркова И.В.,
Начальник НЛ-2222 ФГУП ЦНИИС, к.т.н.,
barkova@zniis.ru

Сергеева Т.П.,
Главный научный сотрудник НЛ-2222 ФГУП ЦНИИС, к.т.н.,
tsergeeva@zniis.ru

Тетекин Н.Н.,
Старший научный сотрудник НЛ-2222 ФГУП ЦНИИС, к.т.н.,
tetekin@zniis.ru

Введение

В качестве основного показателя надежности сети обычно [1,2] используется вероятность связности двухполюсной сети или вероятность получения соединения между двумя точками сети (двухточечная надежность). В общем случае вероятность связности двухполюсной сети не полностью соответствует двухточечной надежности. Расчет вероятности связности учитывает все возможные пути между полюсами (узлами). Однако на практике затруднительно задействовать для резервирования все возможные на сети пути между двумя узлами. Резервирование обычно производится по какой-то конкретной схеме.

В данной статье в качестве такой схемы резервирования рассматривается резервирование с использованием системы самовосстанавливающихся кольцевых структур SNCP (SubNetwork Connection Protection — резервирование соединений подсети) и MS SPRing (Multiplex Section Shared Protection Ring — кольцо с совместным резервированием мультиплексных секций). В качестве числовой характеристики двухточечной надежности будем использовать коэффициент готовности (Кг), определяемый, например, в [1-3].

Выражения для расчета Кг изолированного кольца известны — их можно найти, например, в [1]. Однако для системы связанных колец выражений для расчета Кг не имеется, вследствие этого для каждого случая объединения колец нужно получать собственное выражение, что для большого количества и разнообразия возможных вариантов соединений колец в сети становится невозможным. Поэтому на практике в больших сетях используется метод расчета коэффициента готовности тракта путем получения нижних и верхних оценок для вероятности связности двухполюсной сети [2].

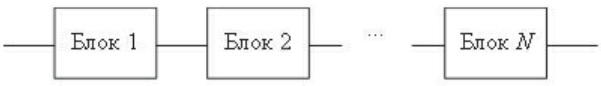
Существующие методики расчета надежности сетей базируются на понятии приводимых и неприводимых сетей. Приводимыми сетями являются все сети с последовательно-параллельным соединением фрагментов сети. Для приводимых сетей модель расчетов основана на том представлении, что всегда имеется минимальная совокупность включенных последовательно функциональных блоков, так что полная система работоспособна только в том случае, когда исправны все блоки этой совокупности. Поэтому для оценки надежности приводимой сети ее наиболее простая обобщенная модель представляется в виде цепочки последовательно соединенных функциональных блоков (рис. 1а).

Каждый из этих функциональных блоков может представлять собой узел или линию сети или может иметь собственную схему резервирования, например, схему параллельного или кольцевого резервирования. Отметим, что изолированное кольцо является приводимой сетью и может быть сведено к параллельному соединению двух цепочек из звеньев, соединяющих полюса по двум сторонам кольца, тогда как схема из соединенных с помощью общих ребер двух и более кольцевых структур является неприводимой сетью.

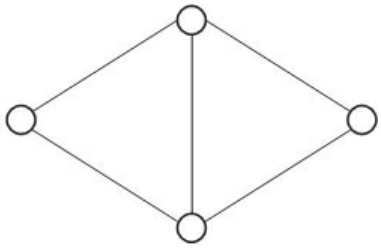
Для приводимых сетей расчет коэффициента готовности производится для каждого блока, далее для объединения функциональных блоков. Для неприводимой сети расчет Кг не может быть выполнен так просто. Наиболее простым примером неприводимой сети является мостиковая схема, изображенная на рис. 1б.

Точный расчет надежности неприводимых сетей является достаточно сложным и трудоемким, поэтому обычно используется способ оценки снизу и сверху надежности неприводимых сетей. Примеры таких оценок можно найти, например, в [2].

В данной статье для получения нижней оценки Кг тракта, проходящего по системе взаимосвязанных кольцевых структур предлагается использовать метод так называемой спец-оценки. Суть метода заключается в том, что для получения Кг заданного тракта получаем структуру, которую назовем "действующая структура", когда из се-



а) Схема последовательного соединения функциональных блоков



б) Мостиковая схема

Рис. 1. Примеры схем сетей: а) приводимая, б) неприводимая

ти вырезаются только те кольцевые структуры, через которые проходит данный тракт. На основе действующей структуры строится новая "расчетная структура", в которой общие ребра в кольцевых структурах дублируются, а сами кольца связываются через узлы перехода рассматриваемого тракта из одного кольца в другое. Эта структура отображает тот факт, что физически общие ребра соответствуют двум параллельным участкам (волокнам или оптическим каналам). По сути, предлагаемый метод представляет собой расчет Кг для последовательности колец, связанных не с помощью общих ребер, а с помощью общих узлов, которыми являются узлы перехода тракта из одного кольца в другое.

Математическая формулировка метода спец-оценки

Рассмотрим тракт, проходящей по системе взаимосвязанных кольцевых структур SDH.

Для формулирования метода получения Кг введем ряд обозначений:

B — путь прохождения рассматриваемого тракта; $B = \bigcup_{i \in R} B_i$

R — множество всех кольцевых структур, по которым проходит путь рассматриваемого тракта;

B_i — часть пути тракта, проходящая по i-му кольцу, т.е. ;

K_i — множество участков i-го кольца в пути B,

N_i — множество всех участков i-го кольца ($N_i > K_i$),

$N_i \backslash K_i$ — множество всех участков i-го кольца, за исключением тех, по которым проходит основной путь тракта.

Тогда оценка Кг тракта, проходящего по системе взаимосвязанных кольцевых структур может быть получена по формуле (1):

$$Кг_{спец} = \prod_{i \in R} Кг_{B_i} \quad (1)$$

где

$Кг_{Bi}$ — Кг части пути B_i, вычисляемый по формуле (2) для MS SPRing и (3) для SNCP.

$$Кг_{B_i} = \prod_{j \in K_i} Кг_{уч_j} + \sum_{j \in K_i} \left(\left(1 - Кг_{уч_j}\right) \prod_{l \in N_i, l \neq j} Кг_{уч_l} \right) \quad (2)$$

$$Кг_{B_i} = \prod_{j \in K_i} Кг_{уч_j} + \left(1 - \prod_{j \in K_i} Кг_{уч_j}\right) \cdot \prod_{l \in N_i \backslash K_i} Кг_{уч_l} \quad (3)$$

где $Кг_{уч_j}$, $j \in K_i$ — Кг участков i-го кольца, по которым проходит (основной) путь тракта; $Кг_{уч_l}$, $l \in N_i$ — Кг участков i-го кольца, по которым не проходит основной путь тракта.

Здесь под "участком кольца" понимается мультиплексная секция (МС), включающая, вообще говоря, C участков кабеля (C_j для j-й МС) и E элементов оборудования (E_j для j-й МС) (мультиплексоров, регенераторов). Кг j-го участка кольца вычисляется по формуле (4):

$$Кг_{уч_j} = \prod_{e=1}^{E_j} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \quad (4)$$

Кг участка кабеля ($Кг_{каб}$) вычисляется по формуле (5):

$$Кг_{каб} = (Кг_{1км})^k \quad (5)$$

где k — протяженность участка кабеля в км; $Кг_{1км}$ — Кг 1 км кабеля данного типа.

Таким образом, учитывая формулы (1)-(5) окончательно получаем следующее выражение для оценки Кг тракта, проходящего по системе взаимосвязанных кольцевых структур с защитой MS SPRing (6) или SNCP (7):

$$Кг_{спец(MSSPRing)} = \prod_{i \in R} \left\{ \prod_{j \in K_i} \left(\prod_{e=1}^{E_j} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) + \sum_{j \in K_i} \left[1 - \left(\prod_{e=1}^{E_j - 1} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) \right] \times \prod_{l \in N_i, l \neq j} \left(\prod_{e=1}^{E_j - 1} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) \right\}, \quad (6)$$

$$Кг_{спец(SNCP)} = \prod_{i \in R} \left\{ \prod_{j \in K_i} \left(\prod_{e=1}^{E_j} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) + \left(1 - \prod_{i \in R} \left\{ \prod_{j \in K_i} \left(\prod_{e=1}^{E_j - 1} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) \right\} \times \prod_{l \in N_i \backslash K_i} \left(\prod_{e=1}^{E_j - 1} Кг_{обор_e} \cdot \prod_{c=1}^{C_j} Кг_{каб_c} \right) \right\}, \quad (7)$$

где (E_j-1) означает, что при расчете для j-го участка не учитывается Кг оборудования, установленного на одном конце j-й МС, так как в противном случае в общей формуле это оборудование учитывалось бы дважды (в силу замкнутости кольца).

Кг участка кабеля вычисляется по формуле (8):

$$Кг_{каб} = (Кг_{1км})^k \quad (8)$$

где k — протяженность участка кабеля в километрах; $Кг_{1км}$ — Кг 1 км кабеля.

Пример оценки Кг по предлагаемому методу

Рассмотрим тракт, проходящий из A в D через 3 кольца MS SPRing (рис. 2).

Пусть тракт A-D проходит по участкам A-B, B-C, C-D. При этом участок A-B проходит по кольцу 1, включающему 3 участка A-B, B-F и A-F. Участок B-C входит в кольцо 2, состоящее из 4 участков: B-C, C-E, E-F и B-F. Участок C-D входит в кольцо 3, состоящее из 3 участков C-E, E-D и C-D. Пусть все кольца организованы как кольца MS SPRing. Пусть протяженность всех участков составляет 1 км. Кг оборудования положим равным 1.

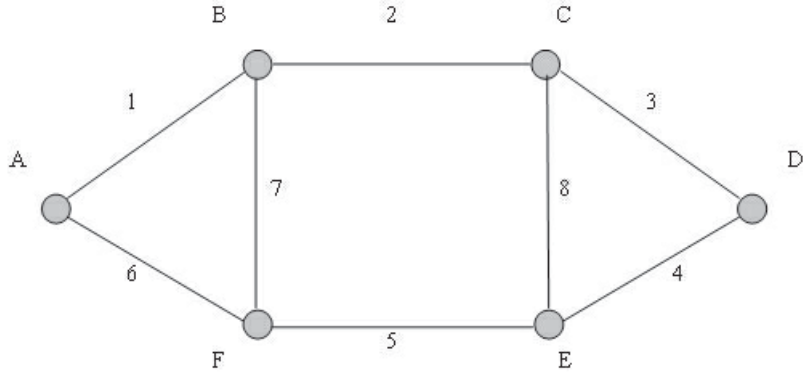


Рис. 2. Схема взаимодействия трех кольцевых структур

Точное значение Кг между узлами *A* и *D* может быть вычислено путем анализа всех вероятностей работоспособных состояний и здесь не приводится в связи с большим размером такой таблицы (число возможных состояний $2^8 = 256$). Формулу для спец-оценки (9) запишем, пронумеровав участки сети в соответствии с рис. 2:

$$Кг_{A-D(спец)} = [Кг_A Кг_1 Кг_B + (1 - Кг_1 Кг_B) Кг_7 Кг_F Кг_6 Кг_A] \cdot$$
$$\cdot [Кг_C Кг_2 Кг_B + (1 - Кг_C Кг_2) Кг_8 Кг_E Кг_5 Кг_F Кг_7 Кг_B] \quad (9)$$
$$\cdot [Кг_C Кг_3 Кг_D + (1 - Кг_3 Кг_D) Кг_4 Кг_E Кг_8 Кг_C]$$

где $Кг_i$ — Кг *i*-го участка.

Результаты расчетов Кг схемы, представленной на рис. 2, сведены в таблицу. Анализ таблицы показывет, что с увеличением значения Кг участка точность спец-оценки повышается. Для Кг участка кабеля протяженностью 1 км, равного 0,99999864 (0,14 обрывов кабеля на 100 км в год при времени восстановления 8,5 часов), соответствующего данным измерений на линиях дальней связи при подземной укладке, погрешность метода спец-оценки Кг по отношению к точному значению не превышает 10^{-10}.

Доказательство справедливости использования спец-оценки в качестве нижней оценки Кг тракта, проходящего по системе взаимосвязанных кольцевых структур

Докажем, что разделение общих ребер в связанных кольцевых структурах на два отдельных ребра, приводит к уменьшению надежности получаемой структуры.

Рассмотрим наиболее простой случай, приведенный на рис. 3, когда два кольца соединенные по общему участку 5 (рис. 3а), в соответствии с методом спец-оценки, заменяются на два кольца, связанные через узел B (рис. 3б), где тракт *A-C* переходит из первого кольца во второе.

Для схемы, изображенной на рис. 3а, Кг может быть рассчитан по формуле:

$$Кг_{точн} = Кг_1 Кг_2 + (1 - Кг_1) \cdot Кг_2 Кг_3 Кг_5 + (1 - Кг_2) \cdot Кг_1 Кг_4 Кг_5 +$$
$$+ (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5 \quad (10)$$

Кг схемы, изображенной на рис. 3б рассчитывается по формуле:

$$Кг_{спец} = (Кг_1 + (1 - Кг_1) \cdot Кг_3 Кг_5) \cdot (Кг_2 + (1 - Кг_2) \cdot Кг_4 Кг_5) \quad (11)$$

Определим разность между Кг для схемы 3б (спец-оценка) и схемы 3а (точное значение).

Результаты расчета Кг системы из трех колец

Табл. 1. Результаты расчета Кг системы из трех колец

Тип оценки *Кг*	*Кг* при $Кг_{участка}$=0,99	*Кг* при $Кг_{участка}$=0,999	*Кг* при $Кг_{участка}$=0,9999
Точное значение	0,999307069	0,999993007	0,99999993
Спец-оценка кольцевых структур	0,999305148	0,999993005	0,99999993

$$Кг_{спец} - Кг_{точн} = (Кг_1 + (1 - Кг_1) \cdot Кг_3 Кг_5) \cdot (Кг_2 + (1 - Кг_2) \cdot Кг_4 Кг_5) -$$
$$- (Кг_1 Кг_2 + (1 - Кг_1) \cdot Кг_2 Кг_3 Кг_5 + (1 - Кг_2) \cdot Кг_1 Кг_4 Кг_5 +$$
$$+ (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5) =$$
$$= (Кг_1 Кг_2 + (1 - Кг_1) \cdot Кг_2 Кг_3 Кг_5 + (1 - Кг_2) \cdot Кг_1 Кг_4 Кг_5 +$$
$$+ (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5^2) - (Кг_1 Кг_2 + (1 - Кг_1) \cdot Кг_2 Кг_3 Кг_5 +$$
$$+ (1 - Кг_2) \cdot Кг_1 Кг_4 Кг_5 + (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5) =$$
$$= (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5^2 - (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5 =$$
$$= (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5 \cdot (Кг_5 - 1). \quad (12)$$

Оценим знак выражения (12) во всем диапазоне значении Кг, т.е. при условии, что

$$Кг_i \in [0;1] \quad \forall i \quad (13)$$

Для получения оценки требуется оценить знак каждого из сомножителей:

$(1 - Кг_1) \geq 0$, так как $Кг_1 \leq 1$;
$(1 - Кг_2) \geq 0$, так как $Кг_2 \leq 1$;
$Кг_3 \geq 0$, $Кг_4 \geq 0$, $Кг_5 \geq 0$ по условию (13).
$(Кг_5 - 1) \leq 0$, так как $Кг_2 \leq 1$.

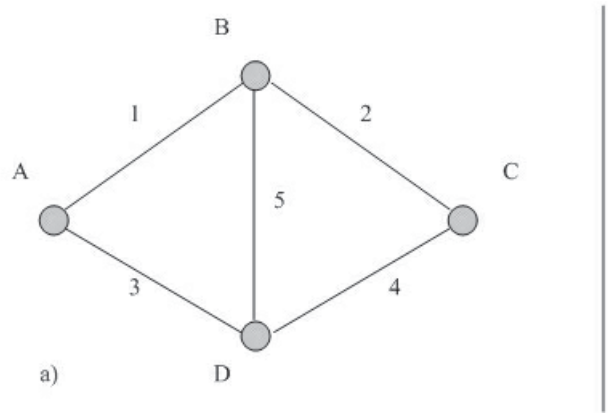
Произведение пяти неотрицательных и одного неположительного множителя дает неположительный результат, т.е.

$$Кг_{спец} - Кг_{точн} \leq 0 \quad (14)$$

Рассмотрим, насколько далеко спец-оценка коэффициента готовности (11) отличается от точной оценки (10). Для этого вычислим отношение модуля величины ошибки (12) к точному значению (10), что даст нам величину относительной погрешности:

$$\Delta = \frac{(1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5 \cdot (Кг_5 - 1)}{Кг_1 Кг_2 + +(1 - Кг_1) \cdot Кг_2 Кг_3 Кг_5 + (1 - Кг_2) \cdot Кг_1 Кг_4 Кг_5 + (1 - Кг_1) \cdot (1 - Кг_2) \cdot Кг_3 Кг_4 Кг_5} \quad (15)$$

В знаменателе дроби (15) фигурирует сумма четырех положительных (точнее — неотрицательных) слагаемых, их которых наи-



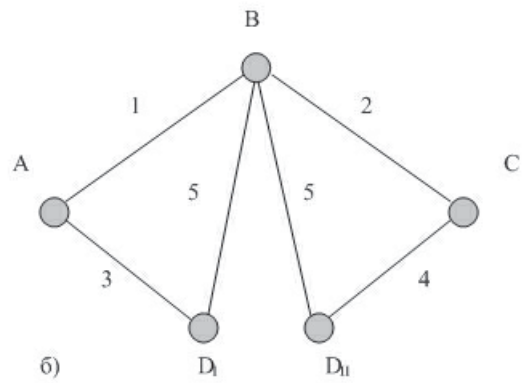


Рис. 3. Схема сети

большим является первое, а три остальных — малы в сравнении с ним. Они имеют порядок 1 – $Кг_i$, i = 1,2 или меньше относительно первого слагаемого. Отбрасывая их, мы уменьшим знаменатель дроби (15), т.е. уменьшим относительную погрешность. Таким образом,

$$\Delta \le \frac{(1-Кг_1)\cdot(1-Кг_2)\cdot Кг_3 Кг_4 Кг_5\cdot(Кг_5-1)}{Кг_1\cdot Кг_2} \qquad (16)$$

В реальных условиях коэффициенты готовности участков 1-5 всегда составляют величины порядка 0,998-0,999 и более, так что разности 1 – $Кг_i$, i = 1,2,5 имеют порядок $1\div2\cdot10^{-3}$ или меньше. Поэтому относительная погрешность спец-оценки имеет порядок не хуже 10^{-8}. Поскольку сами значения Кг близки к единице, то и абсолютная погрешность спец-оценки не превышает 10^{-8}.

Заключение

Представленный метод спец-оценки коэффициента готовности тракта, проходящего по системе взаимосвязанных кольцевых структур SDH, дает результат, погрешность которого не превышает 10^{-8}, что приемлемо при практических расчетах надежности. Вместе с тем метод пригоден для алгоритмизации и может быть использован в соответствующем программном обеспечении. Описанный метод используется при расчётах надежности в программных комплексах КаНаТ и NetLines, разработанных ФГУП ЦНИИС.

Литература

1. **Баркова И.В., Сергеева Т.П.** Математические модели оценки надежности кольцевых структур в сетях SDH // Электросвязь, 2001. — №11. — С.36-38.

2. **Нетес В.А.** Об оценке вероятности связности двухполюсных сетей // Электросвязь, 2001. — №11. — С.39-41.

3. ГОСТ Р 53111-2008. Устойчивость функционирования сети связи общего пользования. Требования и методы проверки.

Автоматизированынй расчет коэффициента готовности времени задержки для трактов сети SDH с помощью программного комплекса КаНаТ

В настоящее время операторы связи большое внимание уделяют обеспечению качества и надежности предоставляемых ресурсов и услуг. Ключевыми качественными показателями функционирования сети связи физического уровня являются коэффициент готовности и время задержки прохождения сигнала. В статье описывается алгоритм расчёта качественных показателей и надежности для тракта, проходящего по отказоустойчивым структурам ВОЛС, при помощи разработанного ФГУП ЦНИИС программного комплекса КаНаТ. Программный комплекс позволяет рассчитывать время задержки прохождения сигнала из конца в конец, а также с высокой точностью (погрешность порядка 10-8) получать оценки коэффициента готовности для тракта, проходящего по сети, использующей различные, в том числе кольцевые, схемы резервирования.

Баркова И.В.,
Начальник НЛ-222 ФГУП ЦНИИС, к.т.н.,
barkova@zniis.ru

Слободской Н.В.,
software@slobodsky.ru

Введение

В настоящее время обязательной практикой при предоставлении клиентам ресурсов сети стало заключение договора SLA. Ключевыми элементами такого договора являются пункты, касающиеся качества и надежности предоставляемых ресурсов [1]. Таким образом, уже на этапе заключения договора оператор должен знать, какие именно параметры качества и надежности он может предоставить для заказываемого соединения. Использование для этой цели специальных компьютерных программ позволяет существенно ускорить расчёты и повысить их точность.

Разработанный ФГУП ЦНИИС программный комплекс КаНаТ позволяет для заданного тракта получить значение коэффициента готовности и времени задержки пакета. Программный комплекс КаНаТ может быть использован в качестве инструмента поддержки SLA, а также для нужд эксплуатационных служб оператора. Использование программы, конвертирующей информацию из базы данных оператора путем обмена с внешними файлами, позволяет производить расчёты для реальных трактов, а также значительно облегчает ввод данных. Модульность структуры комплекса позволяет приспособить его для нужд любого оператора путем замены конвертирующей программы.

Описание программного комплекса

Программный комплекс КаНаТ предназначен для определения коэффициента готовности (Кг) и качественных показателей (времени задержки) трактов, проходящих по отказоустойчивым структурам ВОЛС. Название комплекса образовано от словосочетания "КАчество и НАдежность Трактов", отражающего его назначение.

Программный комплекс КаНаТ позволяет проводить расчеты для двух задач: расчёт коэффициента готовности и расчёт времени задержки. Поскольку обе задачи опираются на одни и те же базовые характеристики сети (структуру первичной сети оператора), оба расчёта сведены в один комплекс. Выбор того или иного вида расчётов производится самим пользователем из меню программы. При необходимости, после корректировки данных возможно проведение повторных расчетов тех же (или других) трактов. По желанию пользователя в любой момент возможен переход к другому типу расчетов с сохранением результатов, полученных ранее. По окончании работы с программным комплексом информацию о сети можно сохранить в файле на жестком диске (или другом носителе информации — по желанию пользователя).

Алгоритм работы с программным комплексом КаНаТ, представленный на рис. 1, включает три основных этапа:

- этап ввода и корректировки информации, на котором производится подготовка данных о первичной сети и трактах, для которых требуется произвести расчет, а также настройка параметров сети, необходимых для выбранного расчета;
- этап расчета, на котором производится настройка параметров расчета и собственно расчет требуемых характеристик трактов;
- информационный этап, на котором пользователь, работающий с программным обеспечением, имеет возможность анализировать информацию о результатах проведенных расчетов.

Этап ввода и корректировки данных

Программный комплекс КаНаТ предусматривает два основных варианта ввода данных:

1) Заполнение данных средствами редактирования, встроенными в программный комплекс КаНаТ.

2) Импорт из базы данных оператора с помощью специальной конвертирующей программы, встраиваемой в программный комплекс КаНаТ.

Этап расчета

Программный комплекс КаНаТ позволяет производить расчёты двух параметров:

— коэффициента готовности (Кг) для тракта с учетом схем резервирования, используемых для заданного тракта;

— времени задержки на распространение сигнала из конца в конец для заданного тракта.

Программный комплекс КаНаТ дает возможность пользователю производить многовариантные расчеты тракта, меняя различные параметры, такие как Кг на участках сети, схемы защиты для участка трассы тракта, время задержки на оборудовании и др.

Информационный этап

Результаты расчетов формируются в виде файлов в формате гипертекста (HTML — Hyper Text Markup Language), что позволяет анализировать их с помощью табличных процессоров, обладающих развитыми возможностями обработки информации, например, Microsoft Excel. Кроме того, результаты расчетов отображаются в окне результатов и сохраняются в файлах данных.

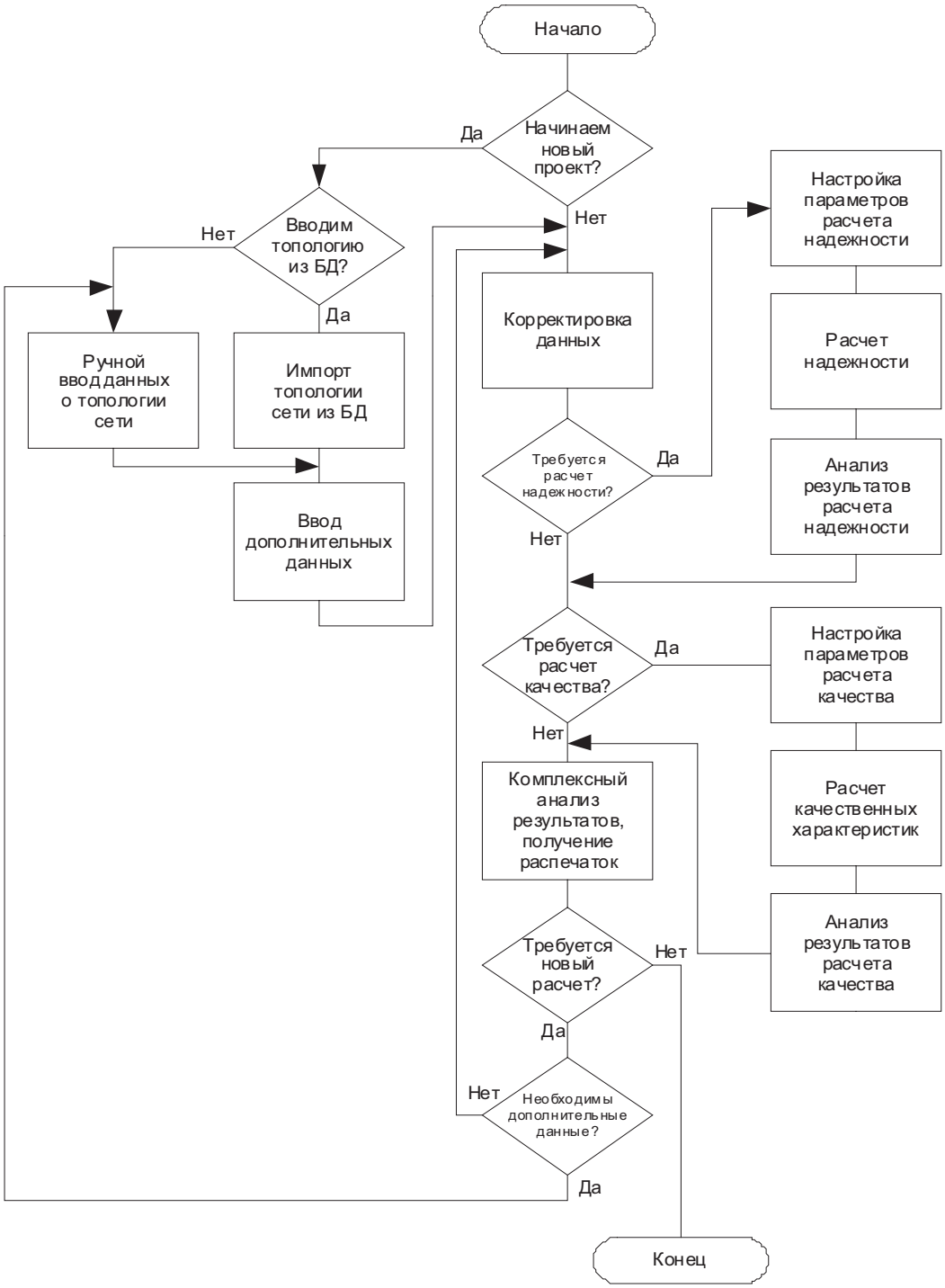


Рис. 1. Блок-схема выполнения расчетов коэффициента готовности и качественных характеристик сети связи с использованием программного комплекса КаНаТ

Методика расчета качества и надежности, используемая в программном комплексе КаНаТ

Исходные данные

Исходными данными для расчёта времени задержки являются:

— схема прохождения тракта с указанием типов оборудования и длин участков кабеля, по которым проходит тракт;

— время на обработку тракта в оборудовании всех типов;

— время на прождение сигнала по кабелям всех типов протяженностью 1 км.

Исходными данными для расчёта Кг являются:

— схема прохождения тракта с указанием типа оборудования (мультиплексор, регенератор) и характеристик участков кабеля (тип прокладки, длина и др.), по которым проходит тракт;

— перечень самовосстанавливающихся структур SDH, имеющихся на сети, по которой проходит тракт (для тракта, заразервированного как SNCP 1+1 указываются две схемы прохождения тракта);

— Кг оборудования всех типов (возможно устанавливать особые характеристики для оборудования разных производителей);

— Кг 1 км кабеля каждого типа.

Методика расчета

Методика расчета времени задержки

Методика вычисления времени задержки сводится к последовательному суммированию задержек в оборудовании и кабеле на всем пути прохождения тракта.

Методика расчета надежности

Для расчёта надежности в программном комплексе KaHaT используется алгоритм, в основе которого лежит метод так называемой спец-оценки, позволяющий получить нижнюю оценку Кг тракта, проходящего по системе отказоустойчивых кольцевых структур SDH (MS SPRing и SNCP). По сути, метод спец-оценки представляет собой расчет Кг для последовательности колец, связанных не с помощью общих ребер, а с помощью общих узлов, которыми являются узлы перехода рассматриваемого тракта из одного кольца в другое. Подробно метод спец-оценки описан в статье "Методика получения нижних оценок коэффициента готовности тракта, проходящего по системе взаимосвязанных кольцевых структур SDH" (авторы И.В. Баркова, Т.П. Сергеева и Н.Н.Тетекин), представленной в этом сборнике. Погрешность метода спец-оценки не превышает 10^{-8}.

Программный комплекс позволяет рассчитывать Кг трактов, использующих следующие варианты резервирования:

— резервирование тракта по схеме 1+1 (SNCP 1+1);

— резервирование участков трассы тракта (мультиплексных секций — МС) по схеме MSP 1:1;

— резервирование МС в кольцах MS SPRing;

— резервирование МС в кольцах SNCP;

— для разных участков трассы тракта используются разные схемы резервирования (MSP, MS SPRing, SNCP или без резервирования).

Для расчета Кг тракта трасса тракта разбивается на сегменты, характеризующиеся использованием одной схемы резервирования. Кг тракта рассчитывается как произведение Кг сегментов трассы этого тракта. Формулы для расчёта Кг различных схем резервирования можно найти, например, в [2].

Расчет Кг в программном комплексе KaHaT производится с учётом большого числа факторов: используемых схем резервирования, типов оборудования и кабелей, по которым проходит трасса тракта.

Программный комплекс KaHaT встроенный справочником по измеренным значениям Кг, а также встроенным калькулятор для пересчёта Кг на время безотказной работы в соответствии с формулой, определенной, например, в [3].

Результаты расчета

Расчет времени задержки

Необходимость оценки времени распространения информации объясняется тем, что в трактах транспортной сети передаются сообщения, которые не должны быть задержаны в сети от абонента до абонента на время, большее некоторого заданного граничного значения, зависящего от класса обслуживания данного сообщения. Расчёты показали, что наибольшее влияние на задержку оказывает протяженность трассы тракта, т.к. задержка в оборудовании существенно меньше задержки в кабеле.

Расчет надежности

Для тракта, трасса которого проходит по системе взаимосвязанных кольцевых схем резервирования, программный комплекс KaHaT дает нижнюю оценку Кг. При этом отклонение от точного значения не превышает величину 10-8. Для тракта, зарезервированного по схеме 1+1 программный комплекс позволяет получить точное значение Кг. Также точное значение получается, если трасса тракта не проходит через кольцевые структуры, взаимодействующие друг с другом через общие участки.

Программный комплекс KaHaT позволяет задать особые значени времени задержки и Кг для оборудования и кабелей разных производителей, разных типов оборудования (например, мультиплексор SDH, регенератор DWDM и др.), разных типов кабелей (здесь тип характеризуется особым Кг и/или временем задержки), различных вариантов прокладки кабеля.

Пример расчета

Пусть необходимо рассчитать качественные характеристики (время задержки и Кг) для тракта A-F, проходящего по сети, схема которой представлена на рис. 2.

Перечень исходных данных:

— производители оборудования и кабелей;

— типы оборудования:

- мультиплексор SDH: задержка = 7 мкс, Кг = 0,9999994,
- регенератор SDH: задержка = 50 мкс, Кг = 0,9999949.

— тип прокладки кабелей — в грунте: на 1 км кабеля задержка = 5 мкс, Кг = 0,9999987;

— узлы сети (*A-J*);

— устройства, установленные в узлах (в узле D — регенератор SDH, в остальных узлах — мультиплексоры SDH);

— участки кабеля (*A-B, B-C, C-D, D-E, E-F, A-G, G-H, C-H, E-I, I-J, J-F*);

— мультиплексные секции (*A-B, B-C, C-E, E-F, A-G, G-H, C-H, E-I, I-J, J-F*);

— кольцевые схемы резервирования: *A-B-C-H-G* типа MS SPRing и *E-I-J-F* типа SNCP;

— тракт *A-F* и его (основная) трасса, записанная в мультиплексных секциях: *A-B-C-E-F*.

Будем считать, что длина каждого участка кабеля составляет 10 км, все кабели произведены одним производителем, имеют один

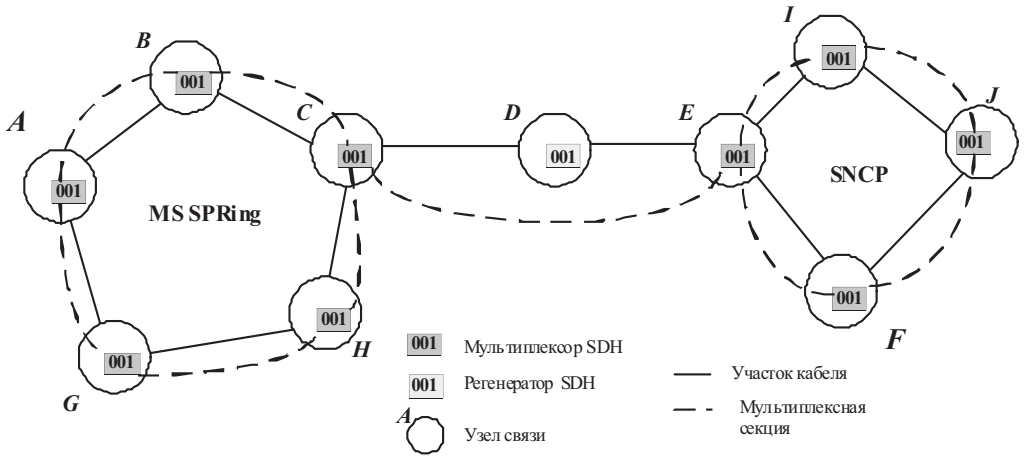


Рис. 2. Схема сети

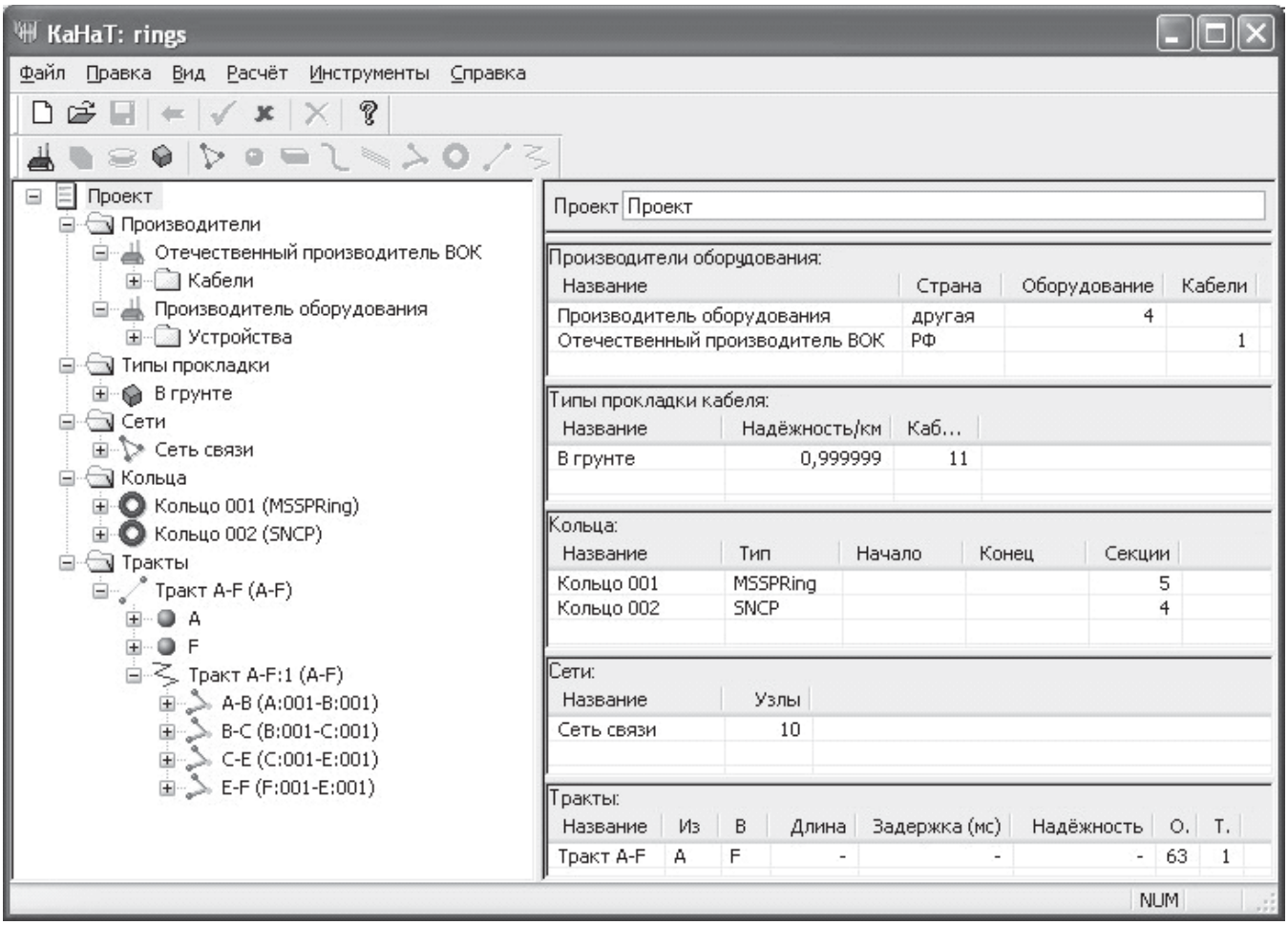


Рис. 3. Окно программного комплекса KaHaT с данными о тракте

тип проладки, т.е. обладают одинаковыми параметрами качества и надежности. Считаем, что всеоборудование произведено одним производителем, параметры качества различаются в зависимости от типа — мультиплексор или регенератор.

Окно программы после ввода всех исходных данных представлено на рис. 3.

Расчет времени задержки

Время задержки для рассчитываемого тракта протяженностью 50 км составлет 0,321 мс.

Следует отметить, что расчёт времени задержки реального тракта протяженностью 11531,2 км показал, что время задержки составляет 58,278 мс. Найденное значение соответствует нормам для 0-го класса [4] и совпадает с результатами измерений для трактов сходной протяженности.

Расчет Кг

Трассу рассчитываемого тракта можно разделить на три сегмента, каждый из которых резервируется по собственной схеме. Мультиплексные секции *A-B* и *B-C* резервируются в кольце MS SPRing; мультиплексная секция *C-E*, проходящая транзитом через узел *D* (через регенератор SDH), не резервируется; мультиплексная секция *E-F* резервируется в кольце SNCP. Расчет показал, что Кг тракта *A-F* составляет 0,999968, что при времени восстановления 8,5 часов соответствует 30 годам безотказной работы.

Заключение

Программный комплекс KaHaT представляет собой удобный инструмент для расчёта ключевых качественных функционирования сети связи физического уровня — времени задержки и коэффициента готовности. Для расчётов параметров реальных трактов, с целью сокращения времени на ввод данных, необходимо использование специальной программы для переноса необходимой для расчёта информации из базы данных сети оператора в программный комплекс KaHaT. Такая программа должна разрабатываться под конкретную базу данных оператора связи и может быть без труда подключена к программному комплексу, благодаря модульности его структуры.

Программный комплекс KaHaT, снабженный программным модулем для конвертирования информации о реальной сети из базы данных оператора, будет полезен различным службам операторов связи, в том числе службе эксплуатации и формирования сети, и позволит с большей достоверностью гарантировать качество предоставляемых арендаторам ресурсов.

Литература

1. ITU-T Recommendation E.860, Framework of a service level agreement — 2002.

2. Баркова И.В., Сергеева Т.П., Математические модели оценки надежности кольцевых структур в сетях SDH // Электросвязь, 2001. — № 11. — С.36-38.

3. ГОСТ Р 53111-2008. Устойчивость функционирования сети связи общего пользования. Требования и методы проверки.

4. ITU-T Recommendation Y.1541, Network performance objectives for IP-based services — 2002.

Анализ модели Call-центра с двумя классами сообщений и тремя группами операторов

Анализируется модель Call-центра, предназначенного для обслуживания двух классов абонентов тремя различными по квалификации группами операторов; исследованы вероятностные характеристики модели. Предложена возможность "переобучения" персонала для обслуживания вызовов любого класса с целью уменьшения вероятности потерь до заданной величины.

Зарипова Э.Р.,
РУДН,
Леонтьева Е.А.,

Введение

Многие организации (экстренные службы, Интернет-магазины, справочно-информационные службы, службы технической поддержки и т.п.), деятельность которых связана с необходимостью принимать и обрабатывать большое количество вызовов, используют в своей работе Call-центры. Клиенты могут дозвониться в Call-центр компании и услышать ответ автоинформатора или оператора, отвечающего на звонки клиента. Разработке моделей Call-центров посвящены работы Б.С. Гольдштейна и его учеников [1], А.В. Рослякова и С.В. Ваняшина [2], а также зарубежных авторов G.Kolle [3], R. Stolletz [4] и др.

Для усовершенствования работы с клиентами компания предусматривает наличие операторов разной квалификации. Входящий поток звонков клиентов разбивается по заранее определенному критерию и направляется к одному из операторов требуемой квалификации. В работе исследована математическая модель Call-центра, основанная на подходе [4], получены вероятностные характеристики модели, проведены расчеты этих характеристик для исходных данных, близких к реальным.

Анализ математической модели Call-центра

Для исследования выбрана модель Call-центра с количеством обслуживающих операторов до нескольких десятков, обслуживающего входящие голосовые вызовы. Поступающие в Call-центры вызовы различаются по длительности обслуживания и требуют соответствующей квалификации обслуживающего персонала. В работе предлагается разделить поток вызовов по заранее заданным критериям на два разных потока и, анализируя содержание запросов, направлять вызов к своему оператору. Абоненты первого класса имеют относительный приоритет. Вызовы обслуживают три группы операторов. Каждая группа операторов характеризуется различной квалификацией. В первую группу входят c_1 операторов, обслуживающих только абонентов первого класса (1-специалисты), во вторую — c_2 операторов, обслуживающих только абонентов второго класса (2-специалисты), и третью группу образуют операторов широкого профиля, способных обслуживать абонентов как первого, так и второго классов. В системе предусмотрены две очереди для каждого класса абонентов. Будем предполагать, что максимальное число каналов для обслуживания абонентов первого (второго) класса, находящихся в системе, т.е. на ожидании в 1-очереди (2-очереди) или на обслуживании у оператора, ограничено числом $K_1(K_2)$. ($K_i > c_i + c_{12}$, $i = \overline{1,2}$). Все абоненты в рассматриваемой модели являются "терпеливыми", т.е. не покидают систему до тех пор, пока не будут обслужены оператором [2,3,4].

Вызов, поступивший в Call-центр, немедленно поступает на обслуживание к своему специалисту. Если все операторы для данного класса вызовов заняты, и есть свободные операторы в третьей группе, то вызов направляется на обслуживание к оператору широкого профиля. Если же заняты все операторы, то вызов присоединяется к очереди, соответствующей его классу. Оператор из третьей группы, освободившись, принимает на обслуживание а) вызовы первого класса, если 1-очередь не пуста; б) вызовы второго класса, если 2-очередь не пуста и нет ожидающих заявок первого класса.

В терминах ТМО, вызовы первого (второго) класса соответствуют 1-заявкам (2-заявкам), операторы, обслуживающие 1-абонентов (2-абонентов) — 1-приборам (2-приборам), специалисты широкого профиля — (1,2)-приборам. Предположим, что в систему поступают пуассоновские потоки 1- и 2-заявок с интенсивностями λ_1 и λ_2 соответственно. Время обслуживания 1-заявки (2-заявки) одним из 1-приборов (2-приборов) распределено по экспоненциальному закону с параметром $\mu_1(\mu_1)$. Времена обслуживания 1- и 2-заявок одним из (1,2)-приборов имеют экспоненциальные распределения с параметрами μ'_1, μ'_1 соответственно. Состояния рассматриваемой системы должны содержать информацию о числе заявок каждого типа, ожидающих в очереди, числе занятых приборов, обслуживающих заявки каждого типа, количестве (1,2)-приборов, обрабатывающих 1- и 2-заявки.

Определим четырехмерный Марковский процесс (МП), описывающий состояние системы в момент времени $t \geq 0$:

$$\{X_1(t), X_2(t), Y_1(t), Y_2(t)\},$$

где $X_1(t)$ — общее число 1-заявок в системе в момент времени (ожидающих в 1-очереди и находящихся на обслуживании на приборах); $X_2(t)$ — общее число 2-заявок в системе в момент времени (ожидающих в 2-очереди и находящихся на обслуживании на приборах); $Y_1(t)$ — число (1,2)-приборов, обслуживающих 1-заявки в момент времени t; $Y_2(t)$ — число (1,2)-приборов, обслуживающих 2-заявки в момент времени t.

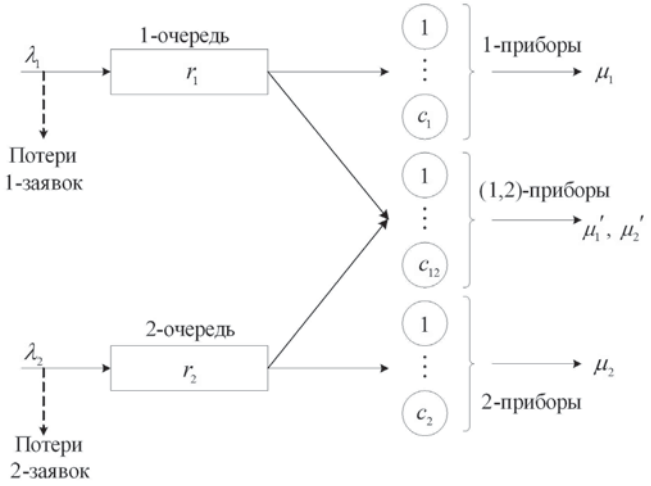


Рис. 1. Схема математической модели Call-центра

Пространство состояний S разбивается на четыре подпространства:

$$S_1 = \{(i,j,k,l) \in S : r_1 > 0 \wedge r_2 > 0\},$$
$$S_2 = \{(i,j,k,l) \in S : r_1 > 0 \wedge r_2 = 0\},$$
$$S_3 = \{(i,j,k,l) \in S : r_1 = 0 \wedge r_2 > 0\},$$
$$S_4 = \{(i,j,k,l) \in S : r_1 = 0 \wedge r_2 = 0\},$$
$$S = \bigcup_{m=1}^{4} S_m,\ S_m \cap S_n = \varnothing,\ m,n = \overline{1,4},\ m \neq n$$

Пусть $p(i,j,k,l) = \lim_{t\to\infty}\{X(t) = (i,j,k,l)\}$ определяет стационарное распределение вероятностей МП $\{X(t),\ t \geq 0\}$. Стационарные вероятности $\{p(i,j,k,l),\ (i,j,k,l) \in S\}$ находим из системы уравнения равновесия (СУР).

Для иллюстрации ниже приведены уравнения равновесия для состояний из подпространства S_1. В подпространство S_1 входят состояния $(i,j,k,l) \in S$, когда в системе находятся ожидающие заявки обоих типов. В этом случае все приборы заняты. Число занятых 1-приборов — c_1, число занятых 2-приборов — c_2. Из всех (1,2)-приборов k приборов заняты обслуживанием 1-заявок, а остальные $l = c_{12} - k$ обслуживают 2-заявки. Таким образом, состояниям (i, j, k, l) из подпространства S_1 соответствуют следующие условия:

$$c_1 + k < i \leq K_1,\ c_2 + l < j \leq K_2,\ 0 \leq k \leq c_{12},\ l = c_{12} - k. \qquad (1)$$

Уравнения равновесия для состояний (i, j, k, l) из подпространства S_1 имеют вид:

$$\begin{aligned}
&[\lambda_1 I(K_1 - i) + \lambda_2 I(K_2 - j) + c_1\mu_1 + k\mu_1' + c_2\mu_2 + l\mu_2']p(i,j,k,l) = \\
&= \lambda_1 p(i-1,j,k,l) + \lambda_2 p(i,j-1,k,l) + \\
&+[(c_1\mu_1 + k\mu_1')I(K_1 - i)]p(i+1,j,k,l) + \\
&+[(c_2\mu_2)I(K_2 - j)]p(i,j+1,k,l) + \\
&+[(l+1)\mu_2' I(k) I(K_2 - j)]p(i,j+1,k-1,l+1), \qquad (2)\\
&(i,j,k,l) \in S_1,
\end{aligned}$$

где $I(x)$ — функция Хевисайда.

Аналогично можно получить уравнения равновесия для подпространств S_2, S_3, S_4.

Основными вероятностными характеристиками являются вероятности потери 1- и 2-заявок. Потеря 1-заявок (т.е. абонент слышит сигнал "занято") происходит в состояниях (K_1, j, k, l), когда общее число 1-заявок в системе достигает своего максимального значения K_1. Такие состояния являются элементами из подпространства S_1 (с ожидающими заявками обоих типов) и подпространства S_2 (с ожидающими 1-заявками). Таким образом, вероятность потери 1-заявок вычисляется следующим образом:

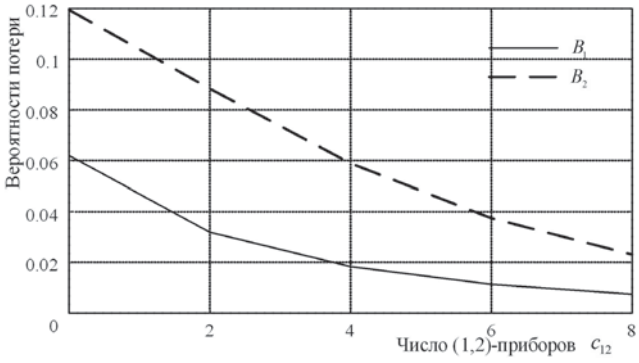


Рис. 2. Зависимость вероятностей потери B_1, B_2 от числа (1,2)-приборов

$$B_1 = \sum_{\substack{(i,j,k,l)\in S_1,\\ i=K_1}} p(i,j,k,l) + \sum_{\substack{(i,j,k,l)\in S_2,\\ i=K_1}} p(i,j,k,l). \qquad (3)$$

Поступающая в систему 2-заявка теряется, если общее число 2-заявок в системе достигает максимального значения K_2. Это справедливо для элементов из подпространства S_1 (с ожидающими 1- и 2-заявками) и подпространства S_3 (с ожидающими 2-заявками и отсутствием ожидающих 1-заявок). Для вычисления вероятности потери 2-заявок используем формулу:

$$B_2 = \sum_{\substack{(i,j,k,l)\in S_1,\\ j=K_2}} p(i,j,k,l) + \sum_{\substack{(i,j,k,l)\in S_3,\\ j=K_2}} p(i,j,k,l). \qquad (4)$$

Пример численного анализа

В Call-центрах необходимо достаточное количество операторов, способных обслуживать вызовы разных классов. Руководитель Call-центра может тратить некоторые ресурсы для переквалификации специалистов узкого профиля (обслуживающих только один тип вызовов) в специалистов широкого профиля. Будем считать, что выполнено условие $c_1 + c_2 + c_{12} = \text{const}$. Величина c_{12} варьируется в интервале [0;8] с шагом 2, при этом одновременно число 1- и 2-приборов уменьшается на 1 на каждом шаге. Предположим, что (1,2)-приборы тратят меньше времени на обработку вызова. На рисунке 2 представлена зависимость вероятности потери 1- и 2-заявок (B_1 и B_2) от числа (1,2)-приборов. Полученные результаты позволяют оптимизировать работу Call-центра и, учитывая экономические и другие параметры, оценить количество операторов узкого профиля, которых необходимо переквалифицировать в специалистов широкого профиля. При построении графика использовались следующие данные: $K_1 = 15$, $K_2 = 13$, $\lambda_1 = 6$, $\lambda_2 = 4$, $\mu_1 = 0{,}8$, $\mu_2 = 0{,}75$, $\mu_1' = \mu_2' = 1$, на первом шаге $c_1 = 8$, $c_2 = 5$.

График показывает, что переквалификация 1-операторов и 2-операторов в специалистов широкого профиля позволяет уменьшить вероятности потери 1- и 2-заявок. Так, например, при отсутствии в системе (1,2)-операторов вероятность потери 1-заявок составляет более 6%, а вероятность потери 2-заявок — 12%. При переквалификации 4 операторов, обслуживающих 1-заявки и 4 операторов, обслуживающих 2-заявки, в специалистов широкого профиля, 1-заявки практически не теряются (потери менее 1%), а вероятности потери 2-заявок снижаются до 2%.

В рамках дальнейших исследований планируется провести анализ других характеристик модели Call-центра, таких как длины очередей, вероятности немедленного обслуживания заявок обоих типов и коэффициент занятости операторов трех групп. Работа будет продолжена введением "нетерпеливых" абонентов, будет построена имитационная модель для обслуживания вызовов в Call-центре с большим количеством классов клиентов.

Литература

1. **Гольдштейн Б.С., Зарубин А.А., Поташов А.И.** Центры обработки вызовов для органов внутренних дел // Учебное пособие. — СПб.: СПбГУТ, 2005. — 40 с.

2. **Росляков А.В., Ваняшин С.В.** Математические модели центров обслуживания вызовов. — М.:ИРИАС, 2006. — 336 с.

3. **Koole G., Pot A.** An overview of Routing and Staffing Algorithms in Multi-Skill Customer Contact Centers // Submitted Version, 2006. — 1-42 p.

4. **Stolletz R.** Performance analysis and optimization of inbound call centers // Lecture Notes in Economics and Mathematical Systems, Chapter 5, Springer, 2003. — 219 p.

Архитектура QoS для конвергентных сетей и особенности ее применения

Предлагается архитектура QoS для конвергентных сетей. Рассмотрены особенности ее применения для сетей, построенных на основе платформ IMS и оборудовании SoftSwitch. Обсуждаются варианты реализации архитектур QoS в конвергентных сетях, рассмотрены их достоинства и недостатки. На основе проведенного анализа обозначены актуальные проблемы, разработаны рекомендации по повышению эффективности применения архитектур QoS для конвергентных сетей.

Ефимушкин В.А.,
директор по науке ФГУП ЦНИИС, к.ф.-м.н., доцент,
ef@zniis.ru

Углов И.В.,
ст. технический специалист ОАО "МТС",
ivuglov@mts.ru

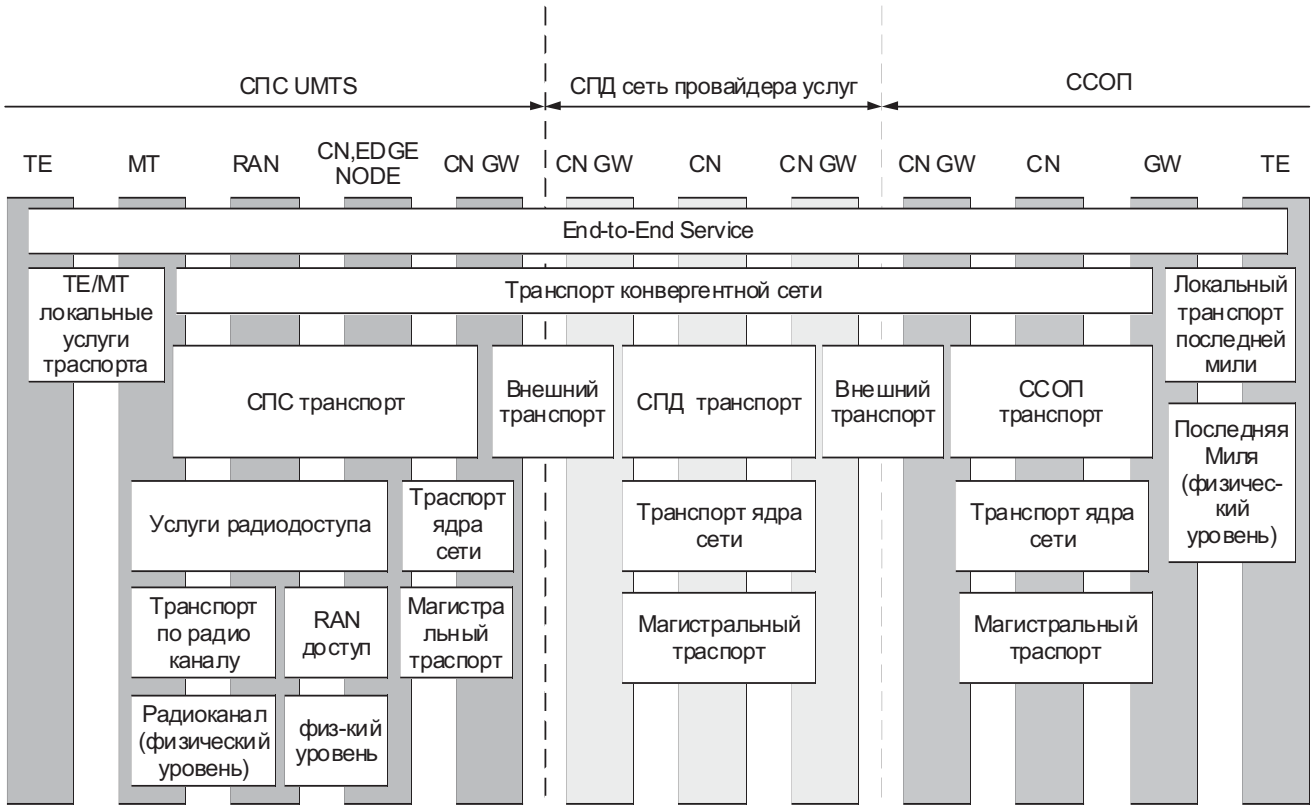
Введение

Прошедший 2009 г. можно охарактеризовать как один из переломных моментов в истории развития телекоммуникационных сетей в России. В этом году разрешился вопрос о повсеместном использованием частот для сетей третьего поколения. Разрешение на использование данных частот способствует развитию мобильных и конвергентных сетей в России, о чем можно судить по росту абонентской базы мобильных сетей в 2009 г. и прогнозам аналитиков на следующий год [1, 2].

Основываясь на прогнозах и анализе фактических данных о развитии телекоммуникационных технологий, таких как WiMAX и LTE (Long Term Evolution) можно предположить, еще большую актуализацию задачи о предоставлении услуг с гарантированным качеством. Особую остроту она приобретает в случае оказания конвергентных услуг, где на стыках разнотехнологичных сетей возникают вопросы о соглашениях об уровне качества обслуживания (Service Level Agreement, SLA) и взаимодействии различных технологий построения сети и доставки сигнала. Базовая архитектура обеспечения качества услуг (Quality of Service, QoS) в мобильных сетях представлена в технической спецификации ETSI TS 123 107 V8.0.0 (2009-01) [3]. В статье предлагается расширение этой архитектуры QoS до конвергентных сетей.

Архитектура QoS для конвергентных сетей

Построим архитектурную модель QoS (рис.1) для конвергентной сети, состоящей из трех разнотехнологичных сетей: сети сотовой подвижной связи (ССПС), сети передачи данный (СПД), сети связи общего пользования (ССОП), и рассмотрим ее особенности. Услуги на сети рассматриваются на участке предоставления от одного терминала пользователя (Terminal Equipment, TE) до другого TE (End-to-end Service). Такого рода услуги могут иметь определенный



Архитектурная модель QoS конвергентной сети, состоящей из трех разнотехнологичных сетей: СПС (UMTS), СПД (SIP сеть провайдера услуг, не имеющая прямых абонентов), ССОП

уровень QoS, предоставляемый на сети пользователям.

Для реализации уровней QoS к уровню транспорта и его услугам предъявляются четкие требования на всем пути предоставления услуги. Уровень транспорта должен удовлетворять всем требованиям для данного QoS, которые могут различаться для различных типов трафика (сигнального, пользовательского и служебного).

Архитектурная модель QoS имеет многоуровневую структуру, где качество передачи трафика на одном логическом уровне обеспечивается подмножеством логических транспортных уровней, расположенных на уровень ниже.

Проходя путь от одного TE к другому, трафик проходит через разнотехнологичные сети, характеризующиеся различными протоколами передачи сигнализации, данных и различной средой передачи. Оборудование TE, подключенное к сети СПС, взаимодействует с ней посредством мобильной точки терминирования трафика (Mobile Termanation, MT). Другой терминал подключен через сеть ССОП к шлюзу, роль которого в данном случае может выполнять автоматическая телефонная станция либо программный коммутатор класса 5.

При предоставлении End-to-End услуг на конвергентной сети, уровень приложений использует ресурсы транспортных сетей, входящих в конвергентную. Таким образом, необходимо добиться обеспечения требуемых характеристик к транспортному уровню на всем маршруте прохождения вызова.

Уровень End-to-End услуг опирается на уровень TE/MT локального транспорта, транспорта конвергентной сети и локальный транспорт последней мили.

Транспорт конвергентной сети можно разделить на блоки транспортных услуг на стыках сетей СПС-СПД и СПД-ССОП, а также на UMTS, СПД, ССОП транспортные услуги. Можно заключить, что UMTS транспорт предоставляет UMTS QoS в рамках СПС сети, СПД транспорт — в рамках СПД, и ССОП транспорт обеспечивает QoS на ССОП, но качество обслуживания на конвергентной сети определяется минимальным показателем QoS на любом ее участке, включая межсетевые стыки. Договоры об уровнях обслуживания между разнотехнологичными сетями должны быть определены контрактами о качестве предоставляемых услуг (Service Level Agreement, SLA).

Транспорт UMTS можно разделить на две составляющие — транспортные услуги радио доступа (Radio Access Bearer Service, RA BS) и транспортные услуги ядра сети (Core Network Bearer Service, CN BS). Оба класса услуг отражают попытки оптимизировать и разделить службы UMTS транспорта, с учетом топологии и особенностей СПС, таких как мобильность и профили мобильных абонентов профили.

Транспортные услуги радио доступа предоставляют защищенный транспорт для сигнального и пользовательского трафика между MT и пограничным узлом ядра сети UMTS (Core Network Edge Node, CN EN), с QoS, заданным UMTS транспортным уровнем или же со значениями QoS, принятыми по умолчанию для сигнального трафика. Эти услуги основаны на характеристиках радио интерфейса и служат для обеспечения возможности перемещения MT.

Услуга CN BS ядра сети СПС, СПД, ССОП связывает узлы CN EN и шлюзовое оборудование ядра сети (Core Network Gateway, CN GW), CN GW и CN GW, CN GW с шлюзовым оборудованием (Gateway, GW) соответственно, для каждой из сетей. Оборудование CN GW служит для обеспечения межсетевого взаимодействия, а GW — для организации подключения конечных абонентов в сети ССОП. Роль транспортных блоков услуг CN BS — эффективный контроль и использование магистральной сети в порядке предоставления услуг верхнему уровню. Пакетное ядро сетей СПС, СПД, ССОП может поддерживать различные магистральные технологии передачи данных.

Рассматривая следующий за CN BS уровень архитектурной модели QoS, можно заметить, что в сетях СПС он делиться на три блока:

— услуги транспорта радиоканала (Radio Bearer Service, RBS);

— услуги сети радиодоступа (RAN Access — Bearer Service, RAN A — BS);

— магистральный транспорт (Backbone).

Услуги RBS определяют все аспекты транспорта через радиоинтерфейс. Услуги транспорта предоставляются UTRAN FDD/TDD или GERAN.

Блок услуг RAN A-BS вместе с услугами физического уровня (Physical Bearer Service, PBS) предоставляют транспорт между радиосетью доступа и ядром сети СПС. Для пакетного трафика с различными QoS данный блок услуг предоставляет транспорт с требуемыми характеристиками.

Блок услуг магистральной сети используется главным образом блоком CN BS. Услуги магистральной сети включают услуги уровней 1 и 2 OSI, выбираются в соответствии с предпочтениями оператора для наиболее полного соответствия возможным требованиям QoS, поступающим с уровня CN BS.

Заключение

В данной статье рассмотрена архитектурная модель QoS для сетей UMTS на основе ETSI TS 123 107 [3]. На основе анализа технической спецификации была предложена обобщенная модель QoS для конвергентной сети. Рассмотрены особенности логических уровней архитектурной модели и их функциональность. Проанализированы ключевые аспекты построения архитектуры QoS в конвергентных сетях.

Литература

1. Развитие сетей 3G в России // J'son & Partners. Информационный бюллетень, ноябрь 2009.

2. Сотовая связь в России // J'son & Partners. Информационный бюллетень, май 2009.

3. Technical Specification ETSI TS 123 107. Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Quality of Service (QoS) concept and architecture // 2009-01.

Построение адресного пространства в сенсорных сетях. Проблемы стандартных подходов и методы их решения

Проводится сравнительный анализ алгоритмов построения адресного пространства в беспроводной сенсорной сети. Рассматриваются механизм распределенного присваивания адресов (Distributed Address Assigning Mechanism, DAAM) и схема адресации с распределенным заимствованием (Distributed Borrowing Addressing Scheme, DIBA). Алгоритм DAAM, стандартизованный для сенсорных сетей IEEE 801.25.4/ZigBee, имеет ряд недостатков, связанных с таким параметром сети, как глубина дерева, приводящих к снижению эффективности процедуры получения адреса новым устройством сети или даже к отказу в получении адреса. В отличие от DAAM, алгоритм DIBA предлагает решение данной проблемы, вводя новую схему распределения адресов в сети и модернизированный алгоритм маршрутизации, и обеспечивает большее покрытие, в отличие от DAAM, но обладает своими узкими местами. По итогам анализа выделяются сильные и слабые стороны каждого из алгоритмов. Предлагаются решения проблем рассматриваемых алгоритмов.

Ефимушкин В.А.,
Директор по науке ФГУП ЦНИИС,
к.ф.-м.н., доцент,
ef@zniis.ru

Шуваев Я.В.,
Специалист по инфраструктуре "Группа компаний ЛАНИТ",
Аспирант ФГУП ЦНИИС,
shuvaev@lanit.ru

Введение

ZigBee — стандарт беспроводной связи для сетей с минимизированными показателями по передаче информации и расходу энергии, основанный на IEEE 802.15.4 MAC/PHY [1] Этот стандарт используется как основной при построении беспроводных сенсорных сетей (Wireless Sensor Networks, WSN). В стандарте ZigBee [2] определены два алгоритма построения адресно-го пространства сети: стохастического присваивания адресов SAAM (Stochastic Address As-signment Mechanism) и распределенного присваивания адресов DAAM (Distributed Address As-signment Mechanism). В работе обоих алгоритмов используется схема, при которой каждый узел, становясь родительским, самостоятельно назначает адреса своим дочерним узлам без взаимодействия с какими-либо явно выделенным центральным узлом сети. Такое поведение характерно для мобильных ad-hoc сетей [3-5].

Суть работы алгоритма SAAM заключается в присвоении новому узлу случайного адре-са. Из-за этой особенности алгоритма, чтобы избежать дублирования адресов необходимо каж-дый раз при добавлении нового узла проверять, не был ли назначен данный адрес ранее, и не является ли он действующим в данный момент [6]. Для реализации такой проверки необходимо использование специальных протоколов маршрутизации по запросу (on-demand, reactive) [7-9] или управляемых таблицей (table-driven, proactive) [10]. Все протоколы маршрутизации указанных типов требуют передачи дополнительной служебной информации и хранения больших таблиц маршрутизации, что нежелательно для сенсорной сети в виду ограничений на величину энергопотребления и объемы передаваемых данных.

1. Алгоритм DAAM

Алгоритм DAAM представляет сеть в виде дерева, а маршрутизация пакетов осуществ-ляется только на основе имеющейся адресной информации без обращений к дополнительным таблицам маршрутизации и предварительной прокладки маршрута. Гарантия уникальности присваиваемых адресов обеспечивается благодаря использованию функции *CSkip* (Child Skip) [11], которая определяет уникальное для каждого уровня дерева число. Вместе с *CSkip* в проце-дуре распределения адресов в алгоритме DAAM для узла m используются такие величины, как максимальная глубина дерева L_m, максимальное число C_m дочерних узлов не маршрутизаторов и максимальное число R_m дочерних узлов, которые могут быть маршрутизаторами.

Однако, когда речь заходит о построении самоорганизующейся сети [12, 13], важным показателем становится площадь покрытия, при этом учитывая фиксированные начальные па-раметры, достаточно сложно добиться высокого покрытия при использовании алгоритма DAAM.

В качестве примера рассмотрим сеть, имеющую характеристики $C_m = 3$, $R_m = 3$, $L_m = 4$, состоящую из трех узлов, расположенных вокруг координатора сети ZigBee — узла, находяще-гося в вершине дерева. Когда в сети появляется новый узел, между ним и уже существующими узлами устанавливаются взаимоотношения, причем их характер изначально не определен. В нашем примере все три узла могут стать дочерними узлами координатора, то есть смежными узлами координатора в терминах теории графов, а могут образовать линейную цепь, увеличив глубину дерева.

Если будет иметь место последнее, лимит параметра L_m будет быстро достигнут, а об-ласть покрытия сетью будет иметь вытянутую форму по сравнению с первым случаем. Рис. 1 иллюстрирует проблему распределения адресов при использовании алгоритма DAAM. Новый узел *i* пытается присоединиться к сети, ассоциировав себя с узлом m, то есть, запросив у него адрес. Однако узел m не может присвоить адрес новому узлу сети, так как, имея три дочерних узла ($C_m = 3$), он уже исчерпал свой лимит по подключению новых узлов.

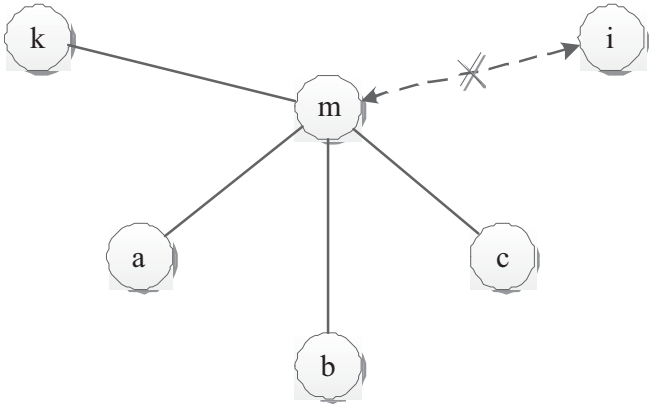


Рис. 1. Проблема присваивания адреса новому узлу в алгоритме DAAM

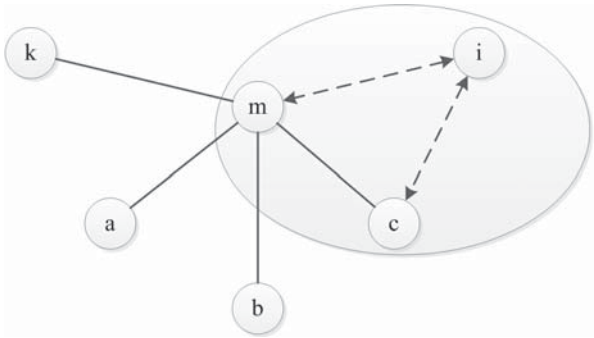


Рис. 2. Новый узел входит в сеть

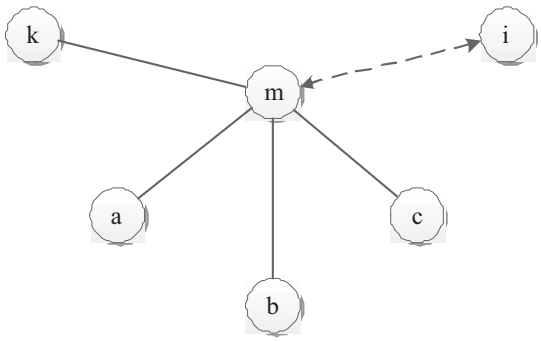


Рис. 3. Поиск подходящего узла

В [14] предложено три решения этой проблемы. Первое решение заключается в исполь-зовании центрального реестра адресов (Centralized Stateful Address Configuration, CSAC), с ис-пользованием которого происходит управление всеми адресами сети центральным координатором. Второе решение состоит в использовании гибридного алгоритма, при котором, распределение адресов с использованием метода CSAC происходит только в том случае, когда узел не может получить адрес при помощи стандартного метода DAAM. Третий подход к решению проблемы предлагает использовать блоки адресов определенного размера, которые делегируются маршрутизатору с разрешением дальнейшего управления.

Несмотря на то, что данные сценарии тем или иным образом решают проблему распре-деления адресов новым узлам сети, все они требуют наличия центрального узла, который либо частично, либо полностью участвует в процедуре адресации. Это делает сеть зависимой от дан-ного узла и снижает уровень отказоустойчивости сети.

Существует еще один метод решения данной задачи, описанный в [15]. Метод линейного адресного присваивания (Linear address assignment, LAA) также подразумевает использование центрального узла для осуществления адресации, однако предполагает использование дополнительной таблицы маршрутизации, рассылки служебных сообщений при расчете маршрута и внесения изменений в формат заголовка передаваемых пакетов. Этот подход, в отличие от предыдущих, не эксплуатирует алгоритм DAAM, а полностью заменяет его.

Все подходы, основанные на централизованной архитектуре, после достижения сетью определенного критического размера, с большой долей вероятности будут приводить к возник-новению длительных задержек между моментом отправки запроса на присоединение до момента получения адреса новым узлом. Такое поведение связано с тем, что за данный период времени кандидату на роль родительского узла необходимо осуществить несколько дополнительных процедур: найти и установить связь с центральным узлом, отправить ему запрос, получить и обработать ответ, - на все эти действия требуется время, особенно на процесс передачи запроса к и от координатора сети, который может находиться достаточно далеко. Накапливаясь, задержки могут привести к срабатыванию таймаута, установленного для процедуры получения адреса в стандарте IEEE 802.15.4, что повлечет за собой остановку процедуры и узел не получит адрес.

Как мы видим, алгоритм DAAM обеспечивает гарантированное построение адресного пространства беспроводной сенсорной сети. Однако, для решения задачи увеличения площади покрытия сети данный алгоритм перестает отвечать заданным требованиям.

2. Алгоритм DIBA

Адресация

Алгоритм, на котором остановимся подробнее, — схема адресации с распределенным за-имствованием (DIBA). Он использует сильные стороны DAAM и позволяет строить сети с большей площадью покрытия за счет увеличения числа узлов, имеющих возможность получить адрес. Кроме того, алгоритм DIBA позволяет гарантировать уникальность распределяемых ад-ресов. Здесь для осуществления маршрутизации совместно со схемой адресации DIBA предла-гается использовать доработанный алгоритм с простой таблицей маршрутизации.

Когда новый узел хочет присоединиться к сети, работающей в стандарте ZigBee, ему необходимо выбрать родительский узел среди доступных в зоне видимости узлов. Предположим, что новый узел *i* находится в зоне действия узлов *m* и *c*, показанной на рис. 2 серой областью, тогда автоматически узлы *m* и *c* становятся кандидатами на роль родительского узла для *i*. В этом случае узел *i* будет присоединен к тому узлу, у которого имеется наибольшее число свободных мест для дочерних узлов. Все узлы, имеющие возможность присоединять дочерние узлы — маршрутизаторы, рассылают широковещательные сигнальные сообщения (beacon), которые содержат информацию о числе оставшихся свободных адресов (Available address count, AAC). Получая сообщения данного типа, новые узлы проводят первичный отбор на роль родительского узла.

Родительский узел, получив запрос на присоединение нового узла, проверяет, если ли у него свободный адрес (Available address, AA). Если свободный адрес имеется, то он выделяется новому узлу. Если же число свободных адресов равно нулю или узел находится на уровне L_m (последнем) дерева, он инициирует процесс заимствования адреса с использованием механизма DIBA. На рис. 3-8 изображена схема работы механизма DIBA.

В нашем примере если узел *m*, получивший запрос на присоединение от нового узла *i*, как показано на рис.3 и 4, имеет ограничение $C_m = 3$ на число дочерних узлов, то узел *i* не получит адрес, в виду того, что узел *m* уже имеет три дочерних узла: *a, b* и *c*.

Для решения этой проблемы, узел m должен разослать широковещательное сообщение AB_REQ с запросом на заимствование адреса, рис. 5. Данное сообщение отсылается родительскому узлу узла *m* и его дочерним узлам. Родительский и дочерние узлы вместе образуют множество соседей узла *m*.

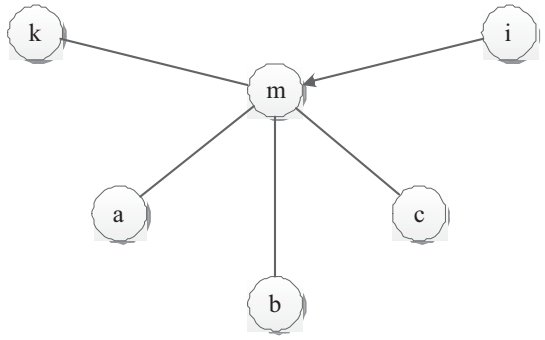


Рис. 4. Отправка запроса на получение адреса A_REQ

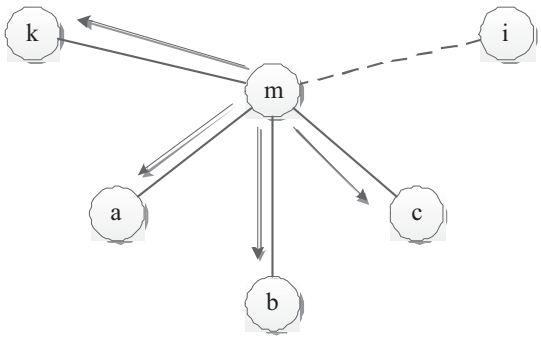


Рис. 5. Отправка запроса на заимствование адреса AB_REQ

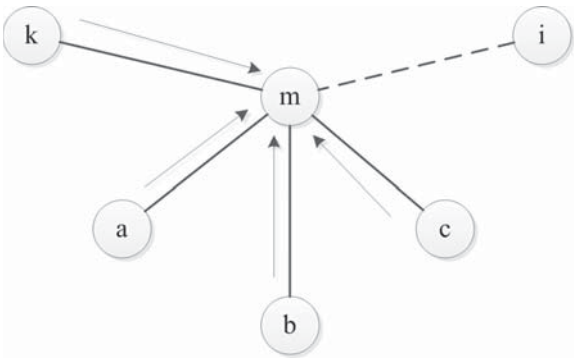


Рис. 6. Получение ответа на запрос AB_RSP

Соседние узлы *k, a, b* и *c*, получив *AB_REQ*, отправляют ответное сообщение *AB_RSP* узлу *m*, которое содержит параметры *AA* и *AAC*, рис. 6.

Если число *AAC* свободных адресов равно нулю, ответное сообщение *AB_RSP* не отсы-лается. Узел m при отправке *AB_REQ* выставляет таймаут, в течение которого, узел m проверяет полученные от соседей сообщения *AB_RSP*. Выбор узла и адреса среди соседних узлу m узлов будет происходить по следующему алгоритму:

• Первым выбирается узел с самым большим значением параметра *AAC*, т.е. узел с наименьшим числом дочерних узлов на текущий момент.

• Если значения параметра *AAC* у соседних узлов совпадают, выбирается узел, с наи-большим значением адреса *AA*.

Если узел m решит заимствовать адрес у узла c, используя описываемый алгоритм, узел m ответит на сообщение *AB_RSP* сообщением с подтверждением *AB_ACK*, как это показано на рис. 7.

Получив подтверждение, узел с заносит адрес, который он отдал в аренду, в свою табли-цу маршрутизации. Ответным сообщением на запрос о присоединении узел m делает узел i сво-им дочерним узлом, присвоив ему заимствованный у узла c адрес, рис.8. Таким образом, новый узел i становится присоединенным к сети.

В соответствии со стандартом IEEE 802.15.4 узел отправляет команду на присоединение координатору сети через промежуток времени, равный macResponseWaitTime символов после получения подтверждения команды на присоединение. Максимальное значение времени ожидания ответа, macResponseWaitTime, равно 64 aBaseSuperframeDuration, что составляет 983 мс при работе сети на частоте 2,45 Ггц. В случае использования DIBA, после того, как родитель получает запрос ассоциации, он рассылает сообщение *AB_REQ* и ожидает сообщений формата *AB_RSP*. После того, как срабатывает таймер, родитель выбирает *AA* и отсылает ответное сообщение новому узлу. Таким образом, задержка между командой запроса ассоциации и командой ответа на запрос ассоциации, равна времени, которое требуется для обмена сообщениями *AB_REQ и AB_RSP*.

В соответствии со стандартом IEEE 802.15.4, атрибут macAckWaitDuration равен макси-мальному числу символов, в течение которых необходимо ожидать получения фрейма с под-тверждением на переданный фрейм — это неизменяемый параметр имеющий значение 864 нс. Так как скорость передачи данных в сети стандарта IEEE 802.15.4 равна 250 кбит/с при частоте 2,45 Ггц, задержка при передаче *AB_REQ* и *AB_RSP*, суммарный объем которых меньше 40 байт, находится в пределах 1 мс для каждого сообщения. Таким образом, при отсутствии потерь, узлу требуется меньше 3 мс для заимствования адреса у одного из соседних узлов. Даже в случае потерь пакетов при передаче, задержка между командой запроса на ассоциацию и командой ответного запроса достаточно мала по сравнению с максимальным значением парамет-ра macResponceWaitTime.

В DIBA узел рассылает широковещательное сообщение *AB_REQ* только своему родительскому узлу и детям, находящимся от него в пределах одного скачка (хопа). Если бы алгоритм позволял узлу рассылать сообщение *AB_REQ* соседям, находящимся на расстоянии более одного хопа, вероятность заимствования адреса существенно увеличилась. Однако такой сценарий очевидно приводит к увеличению задержки. Кроме того, сообщениям *AB_RSP* пришлось бы прокладывать маршрут внутри сети, что привело бы к усложнению таблицы маршрутизации. Таким образом, заимствование адреса у потомков более, чем первого порядка, приводит к усложнению организации сети и не желательно для беспроводных сенсорных сетей.

На рис. 9 представлен код алгоритма заимствования адреса на языке Java, выполняемого в узле — кандидате на родительский узел.

Узлы, получившие сообщения *AB_REQ*, пытаются найти пригодные AA для сдачи в аренду с использованием следующей формулы:

$$AA = My_addr = (R_m - num_BA) \cdot CSkip(d) + 1,$$

где *num_BA* < *Rm* • *CSkip*(*d*) — параметр метода DAAM, описанный в спецификации ZigBee, *d* — уровень (глубина) дерева, на котором расположен данный узел, *num_BA* — число заимствованных адресов. Таким образом, *AA* — наибольший адрес из всех адресов, не присвоенных дочерним узлам или отданных в аренду другим узлам.

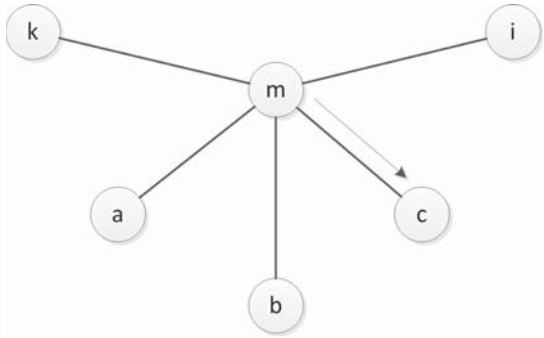


Рис. 7. Отправка подтверждения о заимствовании адреса AB_ACK

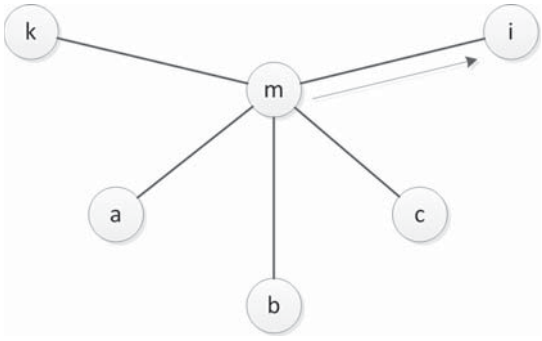


Рис. 8. Ответ на запрос о получении адреса A_RSP

```
if (my_depth < Lm && num_child < Cm) {
  assign_available_address();
  num_child++;
}
else{
  AB_REQ.broadcast();
  Timer T1 = new Timer(1);
  while(new Date().getTime() < T1.getTime()){
    if(AB_REQ.hasArrived()){
    select_address_to_borrow_using_priority_rule();
    break;
  }
}
if (T1.expired()){
  select_address_to_borrow_using_priority_rule();
}
AB_ACK.transmit_to_lender();
AA.assign_to_new_node();
routingTable.update();
```

Рис. 9. Алгоритм работы родительского узла

Если новому узлу присваивается доступный для заимствования адрес (Available address for lending, *AAL*) другого узла, то узел блокирует столько адресного пространства, сколько за-нимает адресный блок размера [*AAL, AAL + CSkip*(*d*) –1], где *d* — глубина дерева, на котором находится узел, отдающий адресный блок. Данный адресный блок вычисляется на основе мето-да присваивания адресов DAAM. Таким образом, заимствование адреса *AAL* у узла является заимствованием объема равного размеру адресного блока [*AAL, AAL + CSkip*(*d*) –1]. Часть заим-ствованного адресного блока не может быть заимствована снова.

Код алгоритма на языке Java, выполняющегося на узлах, получающих сообщение *AB_REQ*, представлен на рис.10.

```
if (num_child + num_BA < Cm && my_depth < Lm){
  AA.select();
  AB_RSP.transmitInclAA();
  Timer T2 = new Timer(2);
  while(new Date().getTime() < T2.getTime()){
   if(AB_ACK.isReceived()){
   routingTable.update();
   num_BA++;
  }
}
```

Рис.10. Алгоритм обработки запроса на заимствование адреса *AB_REQ*

Если родительский узел, который позаимствовал адрес для нового узла сети, покидает сеть и узел, у которого он его позаимствовал, является его дочерним узлом, то этот дочерний узел должен попытаться ассоциироваться с другим узлом сети и получить новый адрес, так как в сети должна поддерживаться структура дерева. Однако, если узел, у которого был заимствован адрес, является родительским узлом для данного узла, то данный родительский узел должен быть уведомлен о том, что заемщик покидает сеть, тем самым возвратив под контроль отданный адрес. Если сеть покидает узел, который отдал свой адрес в аренду, при этом узел, которому он отдал свой адрес, является его потомком, то потомок, получивший адрес не должен возвращать арендованный адрес и может оставить его себе.

Алгоритм маршрутизации

Пример сети ZigBee/IEEE 802.15.4, которая использует алгоритм DIBA, представлен на рис.11. Сеть имеет древовидную топологию. В момент, когда новый узел пытается ассоцииро-ваться с узлом 81, у которого уже есть 3 дочерних узла ($C_m = 3$), он не может получить адрес и поэтому узел 81 заимствует адрес 104, у узла 95, у которого данный адрес зарезервирован для своих потомков. В результате данной операции новый узел получает адрес 104 и становится дочерним узлом узла 81. В свою очередь, узел 104 принимает два новых узла в качестве уже своих дочерних узлов и присваивает им адреса 105 и 106.

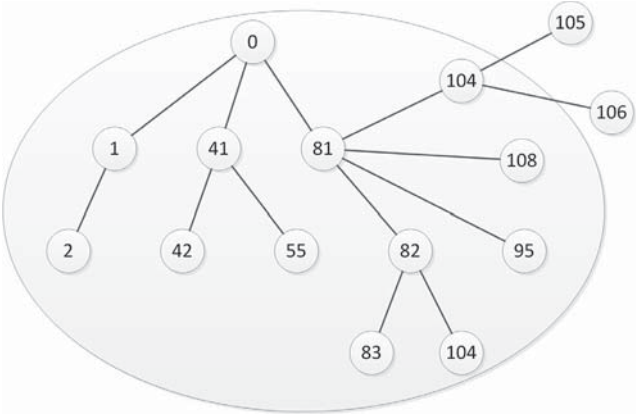


Рис. 11. Пример беспроводной сенсорной сети, использующей алгоритм DIBA

DIBA использует маршрутизацию по дереву (tree routing), основанную на DAAM, что дает ему преимущество перед другими методами в виде простого алгоритма маршрутизации.

Для демонстрации различий в схемах маршрутизации DIBA и DAAM предположим, что узел 104 является узлом получателем, а узел отправитель находится внутри сети, покрытие ко-торой показано на рис.11 серой областью. При использовании метода маршрутизации по дереву DAAM, пакет данных передается к узлу 95 через узел 81, так как адрес 104 изначально предназначался для потомка узла 95. Однако пакет данных не может быть перенаправлен в пункт назначения, так как узел 95 отдал свой адрес 104 узлу 81, как показано на рис.11. Для правильной маршрутизации пакетов в алгоритме DIBA каждый узел поддерживает таблицу маршрутизации, в которой хранится информация о заимствованных адресах, узлы арендаторы и узлы заемщики (табл.1 и 2). Получив пакет с данными, узел 81 перенаправляет этот пакет узлу 104, а не узлу 95, в соответствии с таблицей маршрутизации.

Таблица 1

Таблица маршрутизации узла 81

	Адрес	Арендодатель	Арендатор
Родитель	0	-	-
Дочерний узел 1	82	-	-
Дочерний узел 2	95	-	-
Дочерний узел 3	108	-	-
Дочерний узел 4 (адрес отдан)	104	95	-

Таблица 2

Таблица маршрутизации узла 95

	Адрес	Арендодатель	Арендатор
Родитель	81	-	-
Дочерний узел №1	104	-	81

Таблица маршрутизации, как показано в табл.1 и 2, содержит 16-битные короткие адре-са, адреса узлов-арендодателей и адреса узлов, заимствовавших адреса. Арендодателем называ-ется узел, который отдает адрес, арендатором — узел, получающий адрес. Размер таблицы мар-шрутизации может быть вычислен. Когда узел заимствует адрес, он добавляет заимствованный адрес и адрес узла арендодателя в свою таблицу маршрутизации. Учитывая, что размер адреса составляет 2 байта, таблица маршрутизации увеличивается на 4 байта каждый раз, когда узел заимствует адрес. Так как, согласно алгоритму DIBA, адрес может быть заимствован только у родителя или у дочернего узла, дополнительно в таблице маршрутизации требуется простран-ство в размере, определяемом неравенством:

$$(C_m + 1) \cdot 2\,Байта < Размер < (C_m \cdot 3 + 2) \cdot 2\,Байта$$

В случае сети с параметрами $C_m = 3$, $R_m = 3$, $L_m = 3$ получаем

8 *Байт* $\leq$ *Размер* $\leq$ 22 *Байта*.

Таким образом, для таблицы маршрутизации требуется небольшое увеличение памяти.

Каждый узел маршрутизирует пакеты на основе алгоритма, представленного на рис.12, основываясь на следующей информации об адресе. Получив пакет данных, узел проверяет, не является ли адрес узла получателя его собственным или принадлежит блоку адресов его дочер-них узлов. Если адрес получателя пакета принадлежит блоку адресов его дочерних узлов, узел проверяет в своей таблице маршрутизации, не был ли этот адрес из блока адресов отдан в арен-ду другому узлу. Если адрес был отдан в аренду, то пакет согласно таблице маршрутизации пе-ренаправляется заемщику. Если адрес не был заимствован, то пакет перенаправляется соответствующему дочернему узлу. Если адрес получателя не присутствует среди адресов адресного блока дочерних узлов, пакет перенаправляется родительскому узлу в соответствии с алгорит-мом маршрутизации DAAM.

```
if (packet.getDest_address() == my_address){
  packet.get();
}
else{
  if(inMyChildBlock(packet.getDest_address())){
    if(myChildsBlockisLent()){
      packet.forward(borrower);
    }
    else{
      packet.forward(child);
    }
  }
  else {
    packet.forward(parent);
  }
}
```

Рис. 12. Алгоритм маршрутизации DIBA

3. Предложения по усовершенствованию алгоритмов

Несмотря на все достоинства алгоритма DIBA, возможно возникновение ситуации, когда кандидат на роль родительского узла не сможет найти свободный адрес у своих соседних узлов. В этом случае новый узел не получит адрес и не попадет в сеть. Предлагается рассмотреть несколько вариантов оптимизации алгоритма для предотвращения данной ситуации.

Авторы алгоритма DIBA в [16] вводят ограничение на глубину поиска свободного адреса в один хоп от узла кандидата на роль родительского узла. Данное ограничение объясняется тем, что при увеличении глубины поиска увеличивается задержка передачи сообщений между узлами, что может привести к превышению таймаута ожидания ответа на запрос присоединения на новом узле. Однако, время поиска адреса у соседних узлов в радиусе одного хопа не превышает 3 мс для алгоритма DIBA при определенном в спецификации протокола ZigBee таймауте в 983 мс. На наш взгляд зона поиска адреса для заимствования может быть увеличена до двух-трех хопов. Это позволило бы значительно расширить пул адресов, доступных для заимствования и при этом, не превысить таймаут ожидания ответа на новом узле. Глубина поиска должна определяться в зависимости от начальных параметров сети C_m, R_m, L_m.

Второе предложение по оптимизации алгоритма построения адресного пространства беспроводной сенсорной сети касается ситуации, когда в результате поиска адреса у соседних узлов в радиусе, определенном в алгоритме, то есть, учитывая первое предложение по оптими-зации, не найдется свободных адресов. В таком случае предлагается для присвоения адреса но-вому узлу использовать алгоритм SAAM, выделяя новому узлу случайный адрес из зоны пула доступных адресов, который с наименьшей вероятностью будет занят.

Вычислить такой адрес можно будет по следующему алгоритму. Зная характеристики C_m, R_m, L_m сети, вычисляем максимально возможное число устройств в сети, а соответственно и адресов. Учитывая, что структура сети представляет собой дерево, корнем которого является координатор сети с адресом 0, а присвоение адресов всем новым узлам идет в порядке возрас-тания, то можно с большой долей вероятности предположить, что адреса из пула возможных адресов сети с наибольшим значением будут свободны с большей вероятностью. Таким обра-зом, если для нового узла не находится свободного адреса, предлагается использовать алгоритм SAAM для присвоения этому узлу случайного адреса из конца пула возможных адресов сети.

Заключение

В статье проведен сравнительный анализ алгоритмов построения адресного пространства в беспроводных сенсорных сетях. Были рассмотрены алгоритмы DAAM и DIBA. По итогам анализа алгоритмов были выявлены сильные и слабые стороны каждого. Были предложены усовершенствования, призванные устранить недостатки алгоритма DIBA, такие как расширение области поиска адреса для заимствования, а также использование алгоритма SAAM для присваивания адреса в случаях, когда адрес для нового узла сети нельзя найти, используя основной алгоритм.

Литература

1. IEEE TG 15.4b, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifi-cations for Low-Rate Wireless Personal Area Networks (WPANs) // IEEE Standard for Infor-mation Technology, Part 15.4, 2006.

2. ZigBee Alliance, ZigBee Specification Document, no. 0534r17, Jan. 2008.

3. **Perkins C.E.** IP Address Autoconfiguration of Ad Hoc Networks / C.E. Perkins // IETF Inter-net Draft / draft-ietf-manet-autoconf-01.txt, Nov. 2001.

4. **Nesargi S.** MANETconf Configuration of Hosts in a Mobile Ad Hoc Network / S. Nesargi, R. Prakash // IEEE / INFOCOM, 1059-1068, 2002.

5. **Zhou H.** Prophet Address Allocation for Large Scale MANETs / H. Zhou // IEEE INFOCOM, 1304-1311, 2003.

6. **Yao Z.** Dynamic Address Allocation for Management and Control in Wireless Sensor Net-works / Z. Yao, F. Dressler // Proc. HICSS, 292b, 2007.

7. **Perkins C.E.** Ad-hoc On-Demand Distance Vector Routing / C.E. Perkins, E.M. Rouyer // Proc. IEEE WMCSA, 90-100, 1999.

8. **Akkaya K. A** Survey on Routing Protocols for Wireless Sensor Networks / K. Akkaya, E.M. Rouyer // Ad hoc Networks, vol. 3, no. 3, 325-349, May 2005.

9. **Al-Karaki J.N.** Routing Techniques in Wireless Sensor Networks: A Survey / J.N. Al-Karaki, A.E. Kamal // IEEE Wireless Commun., vol. 11, no. 6, 6-28, Dec. 2004.

10. **Perkins C.E.** Highly Dynamic Destination Sequence-Vector Routing (DSDV) for Mobile Computers / C.E. Perkins, P. Bhagwat // Proc. ACM SIG-COMM, 234-244, 1994.

11. ZigBee Alliance, ZigBee Specification Document, no. 0534r17, Jan. 2008. Part 3.6.1.6.

12. **Cerpa A.** ASCENT: Adaptive Self-Configuring Sensor Networks Topologies / A. Cerpa, D. Estrin // IEEE Trans. Mobile Computing,vol. 3, no. 3, 272-285, July 2004.

13. **Sohrabi K.** Protocols for Self-Organization of a Wireless Sensor Network / K. Sohrabi // IEEE Pers. Commun., vol. 7, no. 5, 16-27, Oct. 2000.

14. **Yen L.H.** Flexible Address Configurations for Tree-Based ZigBee/IEEE 802.15.4 Wireless Networks / L.H. Yen, W.T. Tsai // Proc. AINA, 395-402, 2008.

15. **Jeon H.I.** Efficient, Real-Time Short Address Allocations for USN Devices Using LAA (Last Address Assigned) Algorithm / H.I. Jeon, Y. Kim // Proc. ICACT, 689-693, 2007.

16. **Park S.** Distributed Borrowing Addressing Scheme for ZigBee/IEEE 802.15.4 Wireless Sensor Networks / S. Park, J.H. Ryu, S. Joo, H.S. Kim // ETRI, vol. 31, no. 5, 527-528, Oct. 2009.

Вопросы использования коротких и длинных номеров, особенности нумерации в IP-сетях

Несмотря на то, что в настоящее время национальные планы нумерации большинства стран уже сформированы, появляются новые требования и возможности, которые вносят дополнения в форматы номеров и правила их применения. Традиционные телефонные номера в соответствии с рек. Е.164 МСЭ-Т остаются востребованными как для новых приложений M2M, появившихся в сетях с коммутацией каналов, так и для нумерации абонентов IP-терминалов в сетях с пакетной коммутацией. Цифровизация местных сетей связи и появление Интернет изменили подход к использованию коротких номеров.

Каледин В.В.,
зам.Директора по науке ФГУП ЦНИИС,
e-mail:kaledin@zniis.ru

Козадаева Л.А.,
ведущий научный сотрудник ФГУП ЦНИИС,
kozadaeva@zniis.ru

Использование коротких номеров

Еще в 1998 г. CEPT определил, что короткие номера — это номера длиной от 3 до 5, максимально — до 6 цифр, которые нужны для вызова социально значимых служб федерального и местного уровня на сети фиксированной связи. Короткие номера от 4 до 6 цифр предназначались для развития конкуренции операторов, предоставляющих услуги в рамках местной сети связи. Такие номера состоят из 3-значного кода доступа и 1-3-значного кода идентификации оператора или номера услуги. Трехзначные номера, как самые удобные, предназначались для служб федерального значения. Для коротких номеров, согласованных (гармонизированных) для действия в Европе, рекомендовался диапазон цифры "1".

Анализ международного опыта по применению рекомендаций CEPT показывает, что в настоящее время подход к использованию коротких номеров изменился, что связано с изменением характера потребностей. В целом номера делятся на три основных типа:

• тип А — номера для федеральных служб, которые действуют повсеместно и являются обязательными для всех операторов сетей фиксированной и подвижной связи. К этому типу относится диапазон гармонизированных в Европе номеров "11X(X(X))", в котором уже определены номера "112, 118xx и 116x(x)x))" для общеевропейской службы спасения, справочных служб и срочной психологической помощи;

• тип Б — номера для служб/услуг, действующих на ограниченной территории и/или в сетях нескольких операторов. Такие номера согласуются операторами и провайдрами услуг на условиях саморегулирования внутри каждой страны и имеют длину 4-6 цифр или обычные номера местных сетей;

• тип В — внутрисетевые номера (network specific numbers), которые используются только в рамках конкретной сети и назначаются самим оператором.

Факторами, которые влияют на целесообразность использования таких номеров являются: "прожорливость" коротких номеров в отношении номеров Е.164 (в зависимости от типа плана нумерации), сложившаяся практика использования обычных местных номеров для вызова местных услуг/служб, человеческий фактор, который заключается в неспособности человека запомнить много коротких номеров, различные потребности регионов в разных службах, сложность их назначения, альтернативные возможности Интернета и возможности цифровых станций.

При открытом плане нумерации короткие номера являются "очень прожорливым" ресурсом, т.к. они "съедают" до 10% ресурса номеров Е.164 в каждом коде АВС. Например, в РФ при 7-значном номере абонента 100 коротких номеров "1XX" "съедают" по 1 миллиону номеров Е.164 в каждом коде АВС, т.е. в целом 89 миллионов (по количеству субъектов РФ). При этом в качестве номеров типа Б и В используются обычные местные номера. Поэтому эффективность использования коротких номеров низкая, а стоимость одного трехзначного номера в РФ равна стоимости блока в 10 тысяч номеров Е.164.

При закрытом плане нумерации с префиксом Пн короткие номера не занимают ресурса номеров Е.164, поэтому в качестве первой цифры короткого номера можно использовать все цифры, кроме префикса выхода в национальную сеть: "Пн= 0". В этом случае ресурс коротких номеров становится очень значительным, т.к. можно использовать номера в диапазоне "1xx — 9xx". При закрытом плане нумерации с префиксом Пн, который в настоящее время действует в РФ на сетях подвижной связи, короткие номера типа Б и В используются активно.

Также следует отметить изменение в обработке и маршрутизации экстренных и других вызовов, осуществляемых по коротким номерам. На аналоговой технике такие вызовы устанавливались по специально выделенным линиям. Преимущества цифровых станций дали возможность маршрутизировать такие вызовы в общем потоке телефонного трафика, обеспечивая при этом для экстренных вызовов такие функции как приоритетность вызова и установление обратной связи. Для единообразной маршрутизации короткие номера пересчитываются в обычные номера Е.164, что означает выделение каждому короткому номеру в дополнение номер Е.164.

Отдельные короткие номера остаются необходимыми только для служб высокой социальной значимости, таких как экстренные службы, и их должно быть немного. В настоящее время сеть Интернет становится все более мощным конкурентом справочным службам, организуемым на сети фиксированной связи, поэтому нет необходимости вводить большое количество коротких номеров вместо ныне действующих обычных номеров. Использование коротких номеров в качестве внутрисетевых целесообразно только при закрытом плане нумерации с префиксом Пн.

Использование длинных номеров — нумерация для M2M

В настоящее время значительно возрастает количество и разнообразие приложений, основанных на межмашинном обмене (M2M — Machine-to-Machine, Mobile-to-Machine). Примерами таких приложений являются: проведение измерений на расстоянии, управление торговыми аппаратами. Для этих приложений телефонных номеров пока не требовалось, но в ближайшем будущем в странах Ев-

росоюза намечается запуск системы eCall, для которой потребуются телефонные номера, поскольку будут задействованы сети подвижной связи GSM/GPRS или UMTS. Предполагается, что вскоре все новые автомобили в Евросоюзе будут оснащаться мобильными модулями eCall, способными обеспечить вызов по номеру 112 (общеевропейская служба спасения) или в формате SMS передать информацию о месте происшествия, ускоряя тем самым поступление помощи пострадавшим.

Использование сетей подвижной связи для реализации приложений M2M ставит вопрос о необходимости использования ресурсов нумерации для идентификации объектов/устройств, обеспечивающих эти приложения в соответствии с рекомендациями E.164 и E.212 МСЭ-Т, по которым традиционно осуществляется идентификация абонентов, терминалов и операторов сетей. При реализации этих рекомендаций устройства M2M будут иметь доступ к сетям подвижной связи и действовать в рамках правил, принятых для обычных подвижных терминалов, в том числе на них могут распространяться принятые положения роуминга.

Выделение значительного ресурса номеров E.164 в традиционном формате национального номера (NDC+SN) для многочисленных объектов разных приложений может вызвать большое давление на национальный план нумерации страны. Поэтому для сохранения стабильности национального плана нумерации, некоторые страны (Нидерланды, Норвегия), в которых приложения M2M уже активно развиваются, предлагают использовать специальный диапазон номеров (NDC=ABC=DEF), который обладал бы двумя важными характеристиками: возможностью использования длинных номеров и повторное использованием одних и тех же номеров.

Под длинными номерами следует понимать номера в национальном формате, длина которых превышает длину стандартного национального номера, но не выходит за границу длины в 15 цифр международного номера в соответствии с рек. E.164. Можно сказать, что длинный номер имеет максимально возможную длину национального номера, которую может обеспечить рек. E.164. При однозначном коде страны эта длина составляет 14 цифр, при двухзначном -13 цифр, а при трехзначном — 12 цифр. Такой подход обеспечивает внутри одного трехзначного кода пункта назначении (NDC) емкость абонентских номеров (SN) от 1 миллиарда номеров (при трехзначном коде страны CC и 9-значном SN) до 100 миллиардов номеров (при однозначном коде страны и 11-значном SN в РФ и США).

Вопрос повторного использования номеров не так однозначен, как вопрос использования длинных номеров. С одной стороны, текущее использование приложений M2M не требует от используемых номеров, чтобы они были уникальны повсеместно, поскольку они используются только внутри сети одного оператора (аналогия с закрытой группой пользователей), поэтому один и тот же диапазон номеров E.164 может использоваться в сетях разных операторов M2M. С другой стороны, нельзя исключить, что в будущем потребуется возможность традиционного использования номеров E.164, т.к. может возникнуть необходимость доступа к устройствам M2M из внедомашней сети оператора M2M. В этом случае, возможно, предусмотреть другой новый диапазон длинных номеров E.164, который бы уже обеспечивал повсеместную уникальность номеров E.164. Также следует учитывать возможности реорганизации компаний, обеспечивающих M2M, т.е. их разделения или объединения, что может потребовать пересмотра выделенного совпадающего ресурса нумерации.

Для доступа устройств M2M к сетям подвижной связи необходимо инсталлировать SIM-карты, сформированные в соответствии с рек. E.212. Эта необходимость создает большую зависимость операторов M2M от операторов сетей подвижной связи, т.к. оператор M2M не имеет собственного оборудования СПС и должен заключить контракт на обслуживание с оператором СПС. Если же оператор M2M захочет поменять своего оператора мобильной связи, то ему надо заменить IMSI, т.е. заменить все SIM-карты на всех устройствах M2M, что довольно дорого Имея свой собственный код MNC, оператор M2M будет находиться как бы в роуминге по отношению ко всем операторам СПС и может заключать контракты на роуминговое обслуживание с разными операторами на более выгодных для себя условиях, что можно считать важным аргументом в пользу назначения собственного кода MNC для операторов M2M.

Особенности нумерации в IP-сетях

У абонентов IP-сетей, по сравнению с абонентами традиционных сетей фиксированной связи, появляются новые типы терминалов (IP-терминалы), которые обеспечивают им как привычные речевые, так и новые широкополосные услуги связи. Специфика установления соединений в IP-сетях требует выполнения процедуры регистрации IP-терминала для определения местоположения пользователя в пакетной сети. Благодаря этой процедуре IP-терминалы получили возможность реализации кочевания (nomadicity) - способности пользователя при необходимости менять свою точку сетевого доступа. При этом сеанс предоставления услуги пользователю полностью останавливается, а потом начинается сначала, т.е. при этом не обеспечиваются непрерывность связи, хендовер.

Для обеспечения взаимодействия абонентов традиционных телефонных сетей с абонентами сетей с пакетной коммутацией необходимо для идентификации IP-терминалов использовать действующую нумерацию в соответствии с рек. E.164. При выборе диапазона номеров, который необходимо назначить для абонентов IP-терминалов, должна учитываться возможность "кочевания".

В отчете CEPT "Numbering for VoIP services" (2004г.), в котором обстоятельно исследовались вопросы нумерации для речевых услуг по IP протоколу, отмечалось, что для таких услуг целесообразно использовать номера или из диапазонов географических зон нумерации, или же открыть новый диапазон номеров. Оба диапазона имеют свои преимущества и недостатки. Считается, что номера из географических диапазонов — лучше поддерживают конкуренцию, т.к. они привычны для традиционных абонентов данной местности и не вызывают вопросов о тарифах при установлении соединений. Однако поскольку IP-терминалы обладают свойством "кочевания", то более удобным является новый диапазон номеров, поскольку определение географического местоположения таких абонентов при их обращении к экстренным службам потребует особой обработки

Для того чтобы узнать современное состояние назначения нумерации для IP-терминалов, которое определено Администрациями разных стран, в мае 2009г. был направлен запрос членам группы по нумерации и адресации CEPT, членом которой является ЦНИИС. Были заданы следующие два вопроса:

— какого типа номера могут быть использованы для речевых услуг, предоставляемых с IP-терминала: географические, негеографические, отдельные?

— поддерживается ли возможность кочевания, включая доступ к экстренным службам?

В таблице 1 представлено 10 ответов, полученных из 7 стран (информация других источников отмечены *).

Анализируя представленный международный опыт назначения номеров для IP-терминалов можно сделать следующие выводы:

1. В основном назначаются географические номера, хотя есть страны, в которых назначаются номера из негеографических диа-

Таблица 1

Диапазоны номеров для IP-терминалов, возможность поддержки кочевания и доступ к экстренным службам

Страна	Диапазон номеров и возможность поддержки кочевания при установлении обычных телефонных соединений	Доступ к экстренным службам при возможности кочевания
Австрия	Все типы номеров могут быть использованы. Кочевание поддерживается.	Если VoIP считается общедоступной услугой, то доступ к экстренным службам должен поддерживаться независимо от кочевания. В настоящее время нет законных требований, чтобы маршрутизировать вызов к нужным экстренным службам, если нет технической возможности.
Англия	Возможны номера из любых диапазонов, но наиболее общим являются географические номера. Кочевание поддерживается.	Те операторы, которые обеспечивают МГ вызовы, должны обеспечить доступ к экстренным службам.
Германия	Географические и негеографические номера. Кочевание поддерживается.	Не поддерживают, т.к. нет соответствующих рекомендаций. Операторы обязаны информировать таких абонентов об отсутствии доступа к экстренным службам.
Голландия	Географические, отдельные и некоторые типы негеографических номеров. Кочевание поддерживается.	Некоторые операторы поддерживают
Португалия	Географические номера для абонентов, которые не будут пользоваться возможностью «кочевания» и отдельный диапазон «30» для тех абонентов, которые будут использовать возможность «кочевания.	Провайдеры, предоставляющие услуги «кочевания» должны поддерживать маршрутизацию по номеру 112.
Швейцария	Географические номера. Кочевание поддерживается.	Предоставляется по мере технической возможности оператора в каждой местности. В контракте с абонентом указываются ограничения на конкретные территории, из которых доступа нет. При появлении более подходящих технологических возможностей сети доступ будет расширен.
Швеция	Географические и негеографические номера. Кочевание поддерживается	Доступ к экстренным службам от пользователей при их кочевании далеко от дома не поддерживается.
Финляндия [Ficora] (*)	Географические номера, корпоративные номера Кочевание поддерживается.	Предоставляется по мере технической возможности оператора в каждой местности и с разной степенью точности определения местоположения абонента.
США [Dutch] (*)	Географические номера	Доступа к экстренным службам нет, т.к. VoIP считается услугой передачи данных
Япония [Dutch] (*)	Отдельный диапазон негеографических номеров (050-abcd-xxxx)	Информации нет

пазонов, в т.ч. и из новых. В некоторых странах новый диапазон используется только для IP-терминалов с возможностью кочевания, а без кочевания используются географические номера. В других странах новый диапазон используется в целом для IP-терминалов.

2. Все страны поддерживают возможность кочевания при установлении вызова к обычным абонентам.

3. Выход на экстренные службы с IP-терминалов, обладающих свойством кочевания,, крайне ограничен, т.е. реализуется по мере появления технических возможностей. В контракте с абонентами указываются ограничения на территории, с которых такой доступ возможен.

IP-адресация, преобразование нумерации и адресации

Организация взаимодействия абонентов пакетной сети и сети с коммутацией каналов не ограничивается только выбором наиболее подходящего диапазона номеров. Крайне важным является вопрос правильного преобразования номера E.164 в соответствующие идентификаторы URI, действующие внутри протоколов пакетной сети. В настоящее время для IP-телефонии наиболее общепринятым является протокол SIP. Вопросы идентификации абонентов IP сетей рассматриваются в документах RFC 3261 и RFC 3966, разрабо-

танных IETF. Согласно RFC 3261 для идентификации ресурсов связи в целом используется унифицированный идентификатор ресурса "Sip URI", схема которого аналогична схеме идентификатора электронной почты.

Схема "sip:номертелефона@домен" поддерживает форматы международных (глобальных) номеров, которые определены в рекомендации Е.164 МСЭ-Т, а также форматы номеров, которые не соответствуют этой рекомендации, но составляют важную часть национального плана нумерации страны или какой-либо компании. К таким номерам относятся: короткие номера для вызова экстренных, справочных служб; номера абонентов УПАТС и др. Номера такого типа называются номера не-Е.164 или локальные номера.

Глобальный номер — это международный номер Е.164, который состоит из кода страны (CC), кода пункта назначения (NDC), абонентского номера (SN), т.е. имеет вид:

CC NDC SN

Глобальные номера уникальны и однозначны во всем мире.

Локальный номер — это все остальные номера, которые невозможно представить в виде международного номера Е.164. Локальные номера уникальны только в рамках какой-либо конкретной географической зоны, или части телефонной сети или страны.

В схеме Sip URI глобальный номер обязательно должен иметь знак "+" перед кодом страны и желательно параметр "user=phone". В общем виде схема "Sip URI" для глобального и локального номера имеет следующий вид:

sip: +номерЕ.164@gateway; user=phone (глобальный номер)

sip: номерне Е.164@gateway; user=phone (локальный номер)

Следует подчеркнуть важность наличия знака "+" для идентификации глобального номера Е.164, т.к. параметр "user=phone" является опциональным.

В подтверждение необходимости следовать положениям RFC, можно привести мнение двух стран, от которых ЦНИИС получил ответы на вопрос о необходимости использования знака "+" для глобального номера:

Ирландия — если оператор предполагает действовать только внутри своей сети, то для него нет необходимости принимать во внимание международный набор и он свободен делать то, что выберет. Однако если оператор планирует быть доступным внутри своей сети по международным номерам или передавать телефонные номера своих собственных пользователей для внешних контактов, то оператор должен быть информирован о включении знака: "+".

Англия — в разных странах подходы разные, т.к. одни регуляторы считают, что работа внутри сети — это дело самого оператора, а другие регуляторы требуют следования международным стандартам. Однако, для "NGN взаимодействия" необходимо иметь ясные стандарты и следовать им".

Представление телефонного номера через схему "Sip URI" не является единственно возможным в пакетных сетях. В документе RFC 3966 представлена схема "Tel URI", которую можно использовать во всех протоколах, использующих механизм URI. В части универсальности схема "Tel URI" приближается к схеме "URN", которая также может быть использована в разных протоколах. В процессе установления соединения схема "Tel URI" транслируется в другие схемы URI для телефонных номеров, которые соответствуют конкретному протоколу. Для протокола SIP схема "Tel URI" преобразуется в схему "Sip URI".

Согласно RFC 3966 идентификатор "Tel URI" представляет так же, как и "Sip URI", только два типа телефонных номеров: глобальный и локальный, различие между которыми также происходит по наличию знака "+". Глобальный номер имеет знак "+", а локальный номер имеет параметр "phone-context", который определяет рамки правильности применения данного номера. Этот параметр м.б. представлен либо через номер Е.164 или через доменное имя. Внутри конкретной сети использование локальных номеров не ограничивается. В общем виде схема "Tel URI" имеет следующий вид:

tel: +CC NDC SN; глобальный номер Е.164

tel: 1234; phone-context=+E.164: локальный номер

tel: 1234; phone-context=DNS: локальный номер

Если схема "Sip URI" помимо телефонного номера может идентифицировать разнообразные ресурсы связи, в т.ч. и страницы сайтов Интернет, то схема "Tel URI" предназначена для идентификации только телефонных номеров абонентов или службы/услуги. Разработка специального идентификатора для использования телефонного номера в пакетных сетях свидетельствует о том, что, несмотря на появление в IP-сетях таких новых типов идентификации, как адресации и именования, система телефонных номеров Е.164 будет действовать еще очень долго. Это обусловлено тем, что телефонные номера повсеместно внедрены давно и цифры привычны и понятны в широком контексте языковых культур.

Следует отметить, что номера экстренных служб не попадают под схему URI для телефонного номера, т.е. ни под схему Sip URI, ни под схему Tel URI . В настоящее время для идентификации экстренных служб в пакетных сетях IETF разработал RFC 5031 "Универсальные имена ресурсов (URN) для экстренных и других широко известных служб". Предлагается проводить идентификацию через схему "URN", которая включает в себя 10 схем для экстренных служб и 4 схемы для служб консультации. Эти схемы имеют следующий вид:

Urn:servis:sos	служба 112)
Urn:servis:sos.police	(полиции)
Urn:servis:sos.ambulance	(скорая помощь)
Urn:servis:sos.fire	(пожарной службы)
Urn:servis:sos.gas	(службы газовой сети)
Urn:servis:counseling	(общая консультация)
Urn:servis:counseling.children	(помощь детям) и др.

Схема "URN" была выбрана, поскольку она может быть реализована во многих протоколах, в которых реализован механизм URI.

Примеры схем идентификации телефонного номера через схемы SipURI и TelURI представлены в табл. 2.

При передаче номерной информации из сетей с коммутацией каналов (TDM) в сети с коммутацией пакетов (IP сети) происходит преобразование номеров Е.164 и не Е.164, представленных в соответствующих параметрах протокола ISUP, в идентификаторы "Sip URI" протокола SIP. Параметр "nature of address" протокола ISUP обеспечивает правильное преобразование международных и национальных номеров в "SipURI" глобальных номеров, а других номеров — в "Sip URI" локальных номеров.

Таблица 2

Примеры схем идентификации через схемы SipURI и TelURI

Телефонный номер	Sip URI	Tel URI
+7 495 3062311 (глобальный номер аб.ЦНИИС)	Sip:+74953062311@zniis.ru; user=phone	Tel:+74953062311
3062311 (локальный номер аб.ЦНИИС)	Sip:3062311@zniis.ru	Tel:3062311;phone-context=zniis.ru
112	Urn:servis:sos	Urn:servis:sos

При передаче номерной информации из сетей с пакетной коммутацией в сети с коммутацией каналов происходит обратное преобразование идентификаторов "Sip URI" протокола SIP в международный, национальный или местный номер протокола ISUP. Преобразование выполняется на основании анализа наличия знака "+". Если "+" есть и код страны совпадает с кодом страны РФ, т.е. Кс = 7, то номер представляется в формате "national", иначе - "international". Если "+" нет и "Sip URI" не содержит параметр "user=phone", то такие вызовы передаются в формате "локальный номер".

Заключение

1) Короткие номера остаются необходимыми только для служб высокой социальной значимости и повсеместного использования, таких как экстренные службы, справочные службы, которых немного. Доступ к службам местного значения (вызов такси, муниципальные службы и др.) целесообразно, как и в настоящее время осуществлять по обычным местным номерам. Тип плана нумерации является определяющим фактором при создании ресурса коротких номеров и стоимости одного номера. При открытом плане нумерации этот ресурс значительно меньше, а стоимость номера дороже, чем при закрытом плане нумерации с префиксом Пн, поэтому при закрытом плане нумерации внутрисетевые номера могут быть короткими, а при открытом плане целесообразно использовать обычные местные номера. Интернет является мощным конкурентом справочных служб, поэтому механизма использования номера "118x(x(x)) для доступа к конкурентным справочным службам достаточно. На цифровых станциях короткие номера пересчитываются в обычные номера E.164, т.е. каждому короткому номеру соответствует еще в дополнение номер E.164.

2) Для реализации приложений M2M через сети подвижной связи можно использовать номера E.164, представленные с максимально возможной длиной национального номера, так называемые - "длинные номера". Для организации гибкого взаимодействия операторов M2M с операторами сетей подвижной связи целесообразно рассмотреть возможность назначения операторам M2M собственных кодов MNC, по рек.212.

3) Для согласованного взаимодействия абонентов IP-сетей с абонентами сетей с коммутацией каналов необходимо использовать номера E.164. При определении диапазона номеров следует учитывать способность "кочевания" IP-терминалов, которая будет играть важную роль для определения местоположения абонента при выходе к экстренным службам. Международный опыт показывает, что одним из вариантов назначения подходящего диапазона номеров является следующее решение: если абонент не планирует использовать "кочевание", то ему назначается географический номер той местной сети, к которой он принадлежит, в противном случае абоненту назначается номер из нового негеографического диапазона номеров, который специально создается для "кочующих" абонентов.

Влияние электромагнитной совместимости ЛЭП и линий связи на качество телекоммуникационных услуг

Описывается влияние магнитного поля высоковольтной линии электропередачи на линии связи в режиме короткого замыкания. Рассматривается вопрос обеспечения качества передачи информации по кабельным линиям связи при данном виде влияния и некоторые меры защиты от него. Так же производится оценка качества передачи информации в системах ИКМ-120 и ИКМ-480.

Коротовских Д.В.,
Инженер НЛ-2222 ФГУП ЦНИИС,
Korotovskikh-d@yandex.ru

Введение

В настоящее время особенно актуальными являются вопросы, касающиеся надежности работы линий связи, а так же вопросы качества передачи информации по этим линиям. Развитие электроэнергетики требует строительства мощных электростанций с высоковольтными воздушными линиями электропередачи (ВЛ), и, как правило, трассы магистральных линий связи совпадают с направлениями ВЛ. При этом электромагнитные поля около линий электропередач создают наведенные напряжения и токи в цепях и каналах связи. Различают длительные (при нормальной эксплуатации ВЛ) и кратковременные (при аварийных режимах работы ВЛ) посторонние напряжения и токи, которые могут оказывать мешающее действие на передачу сигналов, снижая ее качество.

Влияние магнитного поля ВЛ в режиме КЗ на линии связи

В данной статье рассматриваются кратковременные опасные напряжения и токи, которые имеют место при однофазных коротких замыканиях ВЛ (с заземленной нейтралью). Время действия опасных напряжений, индуцированных в линиях связи при их сближении с ВЛ, определяется скоростью срабатывания автоматов релейной защиты. Рассмотрим схему, представленную на рис. 1.

На основании рисунка 1 можно сделать вывод, что рассматриваемый нами аварийный режим возникает в трех случаях:

1) при заземлении фазового провода трехфазных ВЛ с заземленной нейтралью;

2) при заземления одного или двух фазовых проводов симметричных ВЛ с изолированной нейтралью;

3) при заземлении фазового провода несимметричных ВЛ переменного тока.

Последний вариант (3) имеет место быть только для воздушных и кабельных цепей полуавтоматической блокировки с блок-механизмами, второй (2) - для воздушных ЛС всех видов, а первый (1) - для всех воздушных и кабельных ЛС [2]. Поэтому можно исключить второй (2) и третий (3) варианты, так как рассматриваются только подземные кабельные линии связи, а первый (1) вариант разобрать более детально.

Трехфазная ВЛ с заземленной нейтралью является симметричной линией передачи. Напряжения и токи на всех проводах этих линий равны по величине, но сдвинуты по фазе на 120°, и при нормальном режиме работы ВЛ (при равенстве нагрузок фаз) сумма напряжений относительно земли и сумма токов фазовых проводов практически равны нулю. Другими словами, электрическое влияние, обусловленное неуравновешенным напряжением фазовых прово-

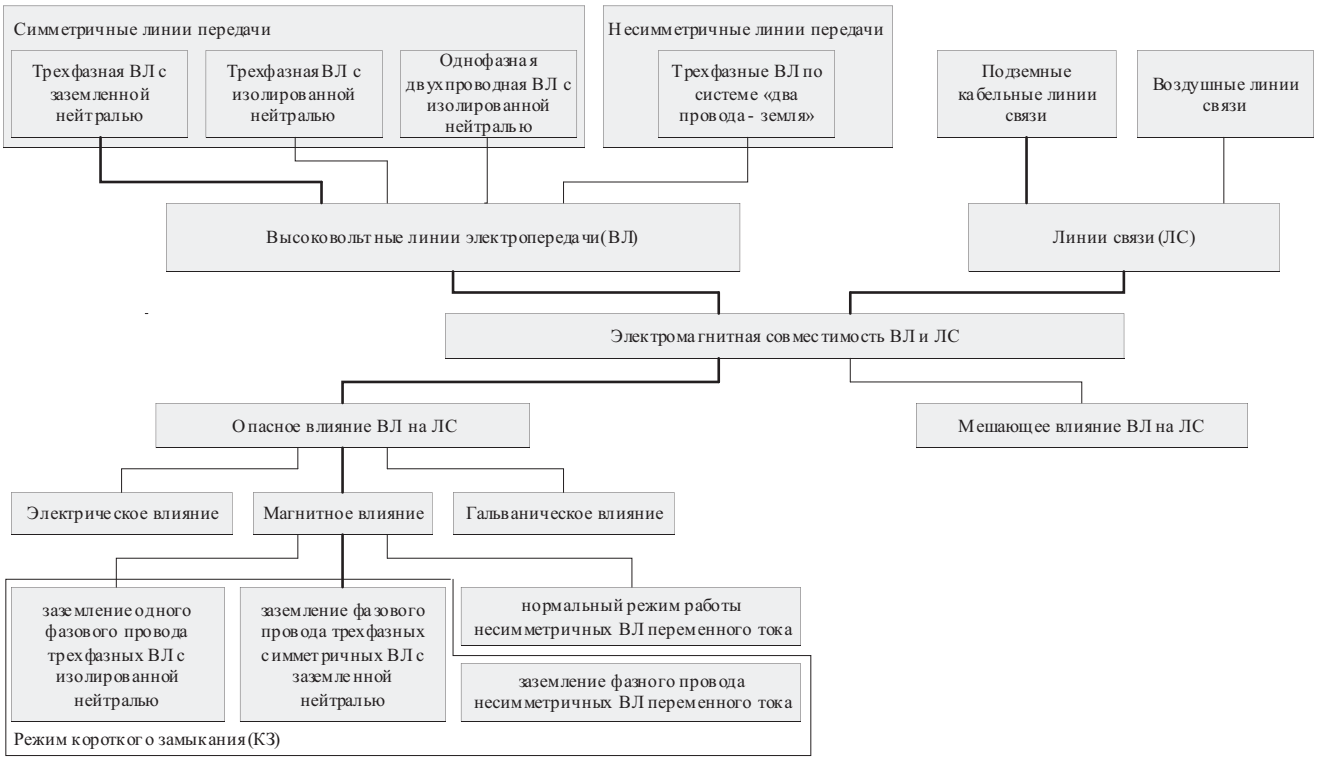


Рис. 1. Электромагнитная совместимость ВЛ и линий связи

Причины повреждения ВЛ

Причины повреждений ВЛ	Частость отказов, %
Перекрытие и разрушение изоляторов	57,0
Перекрытие с фазы на опору	3,8
Перекрытие с фазы на проезжающие механизмы	3,3
Обрыв грозозащитного троса и контакт с фазовым проводом	3,1
Падение фазы на землю	2,4
Набросы на фазу	1,8
Перекрытия на деревья	1,6
Прочие	27,0

дов ВЛ по отношению к земле, отсутствует. Так же отсутствует и магнитное влияние, связанное с неуравновешенным током в проводах ВЛ. Отсюда можно сделать вывод, что при нормальном режиме работы симметричные ВЛ не опасны.

Другое дело — возникновение короткого замыкания, возможные причины которого указаны в таблице.

При обрыве (заземлении) одного из фазных проводов трехфазной ВЛ с заземленной нейтралью в неисправном проводе возникает ток короткого замыкания, достигающий больших значений, при которых переменное магнитное поле ВЛ индуктирует в линии связи значительную продольную э.д.с. (рис. 2). В этом случае может произойти повреждение кабельных цепей и подключенной к этим цепям аппаратуры.

Из выше сказанного можно сделать вывод, что подземные кабели связи, которые покрыты металлической оболочкой, при прокладке в земле подвержены только магнитным влияниям (режим короткого замыкания) и практически защищены от электрических влияний.

Нормы качества передачи информации и надежности линий связи

Для того чтобы обеспечить качественную передачу информации и защитить линию связи и включенную в нее аппаратуру от опасного воздействия внешних электромагнитных полей применяют различные типы мер защиты, приведенные на рисунке 3. Их совместное рассмотрение позволяет более глубоко понять механизм влияния ВЛ на линии связи.

Заметим, что разделение ситуаций влияния на опасные и мешающие несколько условно. Например, опасное влияние, при котором происходит срабатывание разрядника в линейном регенераторе, защищающее оборудование от разрушения, становится по сути мешающим, так как в период закорачивания проводов цепи на "землю" качество передачи информации резко снижается.

Рассмотрим совместное использование мер защиты, применяемых как на линиях электропередачи, так и в аппаратуре связи (рисунок 3). При обрыве фазного провода на землю быстродействующий автоматический выключатель отключает поврежденный участок ВЛ. Современные устройства высокочастотной защиты на линиях электропередачи с заземленной нейтралью позволяют производить отключение ВЛ в течение 0,3-0,4 с. На этом промежутке времени происходит постоянное срабатывание защитных разрядников, установленных в аппаратуре связи, что существенно снижает качество передачи.

Что касается качества функционирования ОЦК или цифрового тракта, то оно оценивается коэффициентом ошибок ($K_{ош}$) или коэффициентом пораженных секунд (SES).

В соответствии с международными нормами снижение качества ниже определенного порога считается отказом. При этом ОЦК, цифровой тракт или соединение переходят из состояния готовности в состояние неготовности [6].

Рекомендации МСЭ-Т G.821, G.826 и G.827 и рекомендация МСЭ-Р F.557 [7, 8, 9, 10] определяют критерии перехода между этими состояниями следующим образом: период времени неготовности начинается с интервала времени, содержащего 10 последовательных SES. Эти 10 секунд рассматриваются как часть времени

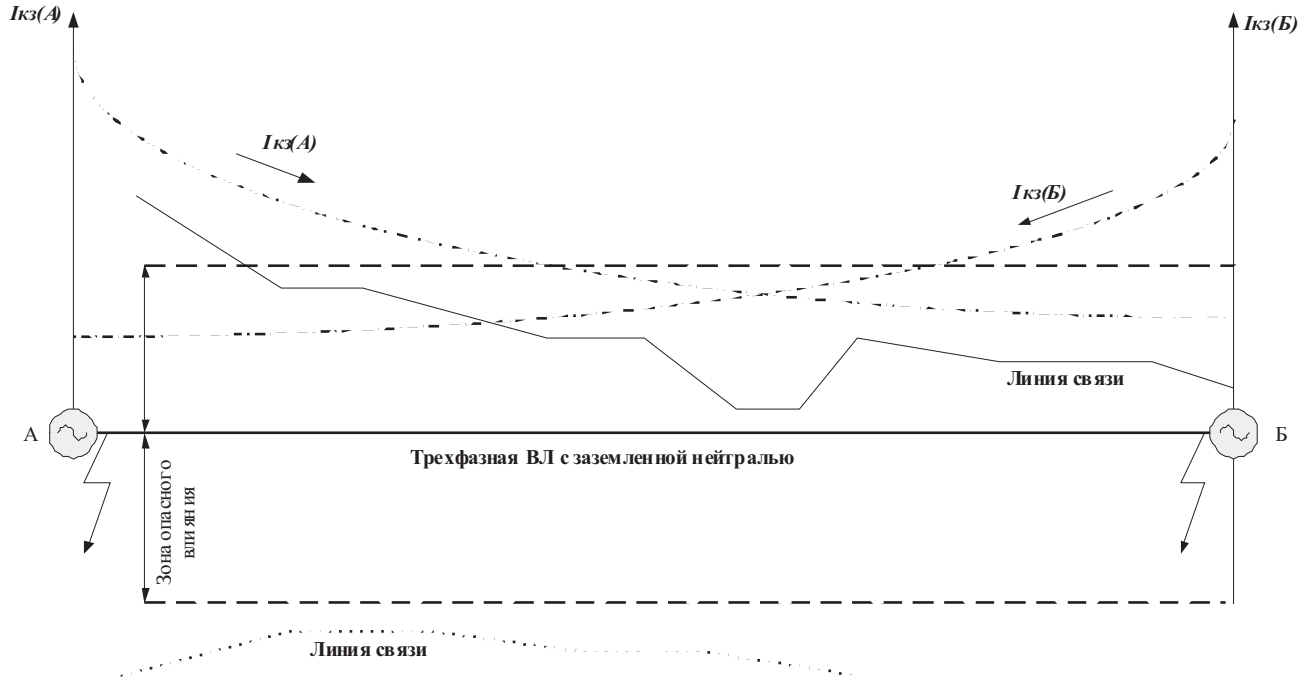


Рис. 2. Сближение ВЛ с линией связи

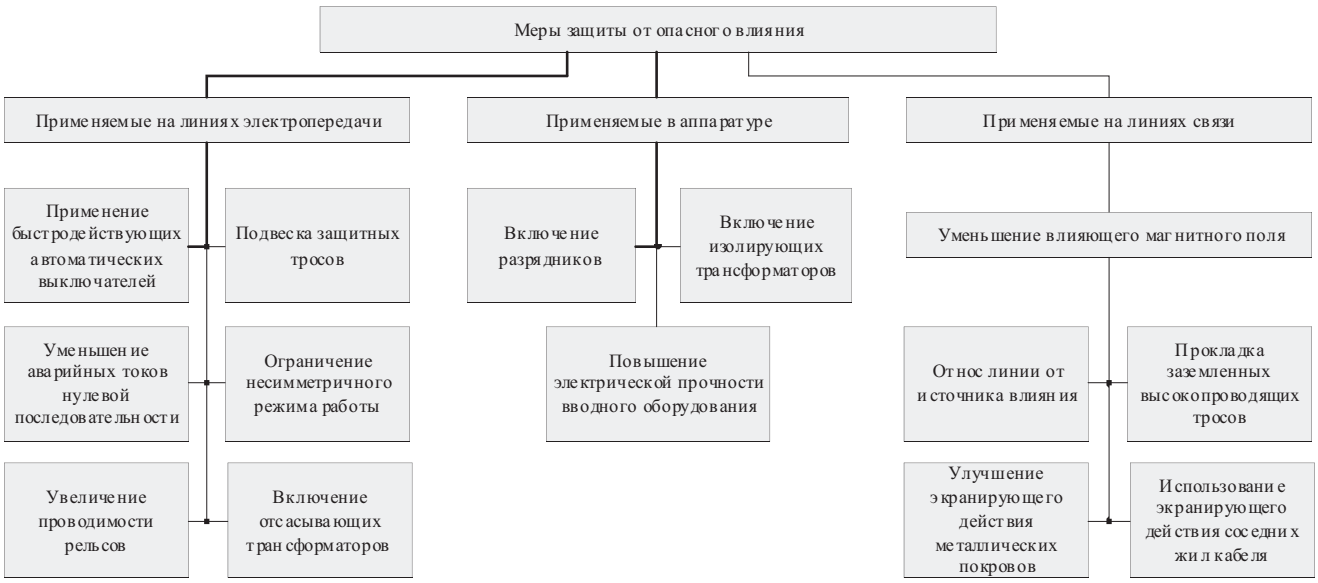


Рис. 3. Меры защиты от опасного влияния линий высокого напряжения

неготовности. Новый период готовности начинается с 10 последовательных секунд, не пораженных ошибками (не — SES).

Напомним, что по определению SES — это односекундные периоды, в которых $K_{ош} \geq 10^{-3}$. Таким образом, значение $K_{ош} = 10^{-3}$ в односекундный период является границей между показателем качества и показателем готовности. Если $K_{ош}$ меньше 10^{-3}, то объект функционирует (возможно, с пониженным качеством); если $K_{ош}$ больше 10^{-3}, то объект находится в состоянии неготовности и, возможно, необходимо предпринимать специальные меры для устранения этого отказа.

Оценка качества передачи информации при влиянии ВЛ на линии связи

Учитывая все вышесказанное, проведем для примера оценку качества передачи информации в системе ИКМ-120. ИКМ-120 — это вторичная цифровая система передачи с импульсно-кодовой модуляцией 120 каналов тональной частоты (ОЦК). Скорость передачи в данной системе составляет 8448 кбит/с. С момента КЗ до момента срабатывания устройств противоаварийной автоматики проходит 0,3 — 0,4 с. На этом промежутке времени защитные разрядники постоянно срабатывают и система ИКМ-120 оказывается неспособной обеспечить требуемое качество передачи информации. За это время она потеряет от 2,5 до 3,4 Мбит информации.

Аналогичная ситуация происходит и в системе ИКМ-480, которая является третичной цифровой системой передачи со скоростью 34368 кбит/с. Потери в данной системе при КЗ составят от 10,3 до 13,7 Мбит.

Литература

1. **М.И. Михайлов, Л.Д. Разумов, С.А. Соколов.** Защита сооружений связи от опасных и мешающих влияний. — М.: "Связь", 1978. — 288 с.

2. Отчет ЦНИИС "Анализ достижений и нормативных документов по вопросу опасного влияния ВЛ на проводные линии связи", 2007. — 34 с.

3. Отчет ЦНИИС "Теоретическая модель опасного влияния магнитного поля ВЛ в режиме однофазного КЗ. Алгоритм расчета опасных влияний ВЛ на проводные линии связи", 2007. — 40 с.

4. **А.А. Воронков.** Методическая разработка к лабораторной работе "Испытание защитных устройств электросвязи" - 2001. - 19 с.

5. Directives concerning the protection of telecommunication lines against harmful effects from electric power and electrified railway lines, Volume II Calculating induced voltages and currents in practical cases, CCITT, Geneva 1989. — 606 с.

6. **Минкин В.М.** Нормирование качественных показателей в свете последних решений МСЭ // Технический материал НИИР, 1998.

7. Rec. ITU-T G.821 Error performance of an international digital connection operating at a bit rate below the primary rate and forming part of an integrated services digital network.

7. Rec. ITU-T G.826 Error performance parameters and objectives for international, constant bit rate digital paths at or above the primary rate.

9. Rec. ITU-T G.827 Availability parameters and objectives for path elements of international constant bit-rate digital paths at above the primary rate.

10. Rec. ITU-T F.557.

Совершенствование принципов организации баз данных компании — оператора связи

Рассматривается подход к решению задачи операторов связи по оптимизации деятельности в условиях насыщения рынка и повышения конкуренции за счет оптимизации принципов хранения и обработки информации, используемой при оказании услуг связи. Представлена информация об используемых операторами связи базах данных, а также базах данных, внедряемых совместно с системами поддержки операционной деятельности, и характерные проблемы, возникающие при этом.

Темкина Т.А.,
Ведущий инженер НЦ "Гармонизация услуг связи"
ФГУП ЦНИИС,
ttletterbox@mail.ru

В настоящее время практически во всех развитых странах рынок услуг связи является высококонкурентным. В частности, рынок услуг фиксированной связи находится в стадии насыщения, и его участники — в основном, традиционные операторы, вынуждены конкурировать с компаниями, оказывающими альтернативные, более дешевые услуги SIP-телефонии.

Рынок услуг подвижной связи (СПС) изначально характеризовался более высоким уровнем конкуренции из-за отсутствия на нем монополистов и практически одновременном появлении нескольких участников. За последние же несколько лет данный рынок в развитых странах перешел в стадию насыщения, и уровень проникновения данных услуг в некоторых странах превысил 130%, что еще более обострило конкурентную борьбу.

Рынок услуг широкополосного доступа (ШПД) в настоящее время интенсивно развивается и характеризуется высоким уровнем конкуренции. На нем работают как традиционные операторы связи, так и альтернативные операторы и поставщики услуг доступа. Для традиционных операторов, оказывающих услуги фиксированной телефонной связи, услуги ШПД являются перспективой развития и средством удержания абонентов. Для альтернативных операторов, занимающих обычно значительно меньшую долю рынка, на первом плане также стоит задача укрепления своих позиций, а для поставщиков услуг доступа, для которых услуги ШПД являются основным источников дохода, привлечение новых абонентов является первостепенной задачей.

В сложившихся условиях для привлечения новых абонентов, а при насыщении рынка — для удержания существующих, операторы связи, в целях сохранения конкурентоспособной позиции, выводят на рынок новые предложения, такие как:

— пакеты услуг;

— новые, дополнительные услуги.

Важными условиями успеха при выводе новых предложений на конкурентном рынке являются:

— обеспечение минимальных операционных расходов на введение новых предложений;

— обеспечение минимальных сроков введения новых предложений.

Выполнение данных условий позволит оператору связи быстро реагировать на действия конкурентов и предпринимать ответные меры, тем самым удерживая позицию на рынке, и при этом избежать существенного падения прибыли из-за затрат на реализацию новых предложений, часть из которых может предлагаться бесплатно в маркетинговых целях, и, соответственно, не приносить дохода.

Однако задача по выполнению данных условий встала перед операторами связи относительно недавно, в связи с чем используемые ими принципы организации операционной деятельности по оказанию услуг связи не позволяют ее решить.

Некоторые операторы связи сейчас используют лишь отдельные, не связанные между собой системы для автоматизации отдельных функций управления, а у некоторых такие системы вообще не используются. В первом случае для внедрения новой услуги необходимо отдельно настраивать используемые системы, прописывая в них информацию о новой услуге и порядке ее оказания, а во втором — вручную настраивать используемое при оказании услуги оборудование. В обоих случаях операционные расходы оператора связи на внедрение новой услуги неоправданно велики, как и объемы работ по ее внедрению.

Таким образом, на сегодняшний день одной из важных задач для операторов связи для обеспечения успешной деятельности в условиях жесткой конкуренции является реорганизация операционной деятельности по оказанию услуг связи в направлении внедрения, объединения и унификации систем управления, используемых в процессе оказания услуг связи.

Вопросом реорганизации бизнес-процессов телекоммуникационных компаний с 2000 года в рамках программы NGOSS (New Generation Operations Systems and Software) занимается международный форум TMF (TeleManagement Forum). В рамках данной программы им были разработаны следующие документы:

— eTOM (enhanced Telecom Operations Map) - расширенная карта бизнес-процессов телекоммуникационных компаний;

— SID (Shared Information and Data) - общая модель информации и данных;

— TAM (Telecom Application Map) - карта приложений телекоммуникационных компаний;

— TNA (Technology Neutral Architecture) - технологически нейтральная архитектура (программных решений).

Первые два документа уже утверждены как стандарты ITU, и разработанные в рамках NGOSS документы часто используются операторами связи как основа при реорганизации и оптимизации деятельности.

В соответствии с eTOM бизнес-процессы телекоммуникационной компании делятся на три вида:

— разработка стратегии развития, инфраструктуры сети и предлагаемых продуктов;

— операционные процессы;

— процессы управления предприятием.

Соответственно на три вида делятся и системы поддержки деятельности телекоммуникационной компании:

— BSS (Business Support System) — система поддержки бизнеса;

— OSS (Operations Support System) — система поддержки операционной) деятельности;

— ESS (Enterprise Support System) - система поддержки управления предприятием.

Бизнес-процессы, связанные с оказание услуг связи, относятся к операционным процессам и поддерживаются системами OSS. К ним относятся следующие:

— продажа и подключение услуги;

— оказание услуги;

— расчет суммы для оплаты и выставление счета;

— решение проблемы клиента;

— решение проблемы, выявленной путем мониторинга средств и сетей связи;

— приостановление/возобновление оказания услуги;

— прекращение оказания услуги по инициативе клиента.

Для автоматизации данных бизнес-процессов, и соответственно, их оптимальной организации, в TAM предусмотрены следующие системы OSS:

— автоматизированная система расчетов;

— система управления взаимоотношениями с клиентами;

— система учета ресурсов (инвентаризация);

— система управления заявками

— система регистрации проблем с ресурсами и услугами и управления их устранением;

— система активации услуг;

— система управления устранением неисправностей;

— система управления параметрами работы и др.

В каждой из перечисленных систем предусмотрена работа с определенным набором данных, а значит, имеется база данных и СУБД для хранения, обработки и анализа информации о состоянии соответствующего объекта, или поддерживаются функции по удаленному доступу к базе данных, содержащей необходимую информацию. При этом требуется обеспечить взаимодействие между системами OSS, поскольку данные и деятельность одной могут зависеть от изменения данных в другой системе.

Эти базы данных предназначены для оптимизации деятельности оператора при оказании услуг, в том числе оптимизации хранения всех данных оператора, относящихся к бизнес-процессам оказания услуги, которые до внедрения систем OSS хранились отдельно, поразному и частично на бумажных носителях.

Например, практически у всех операторов автоматизированы хранение и обработка тех данных, которые постоянно обновляются и нуждаются в отслеживании для обеспечения прибыльности работы компании. Такими данными являются данные о состоянии счета абонента, особенно при авансовой схеме платежей и списании средств со счета по факту использования услуги (например, услуги подвижной связи), а для операторов СПС — данные о местоположении абонента, в зависимости от которого изменяется тарификация вызовов.

Остальные данные, относящиеся, например, к системам инвентаризации и др. системам поддержки операционной деятельности, до сих пор хранятся на бумажных носителях (вплоть до чертежей сетей связи) и в изолированных и не систематизированных и не унифицированных базах данных.

При переводе всех данных в автоматизированные системы OSS необходимо обеспечить их безопасность, защиту от потери, а также возможность оперативного доступа со стороны сотрудников. При внедрении нового принципа обработки данных (автоматизация) при реализации бизнес-процессов также появляется требование по предотвращению перегрузок в сети, соединяющей системы OSS, и перегрузок самих систем, СУБД в них и баз данных оператора.

Реализация этих требований позволит поддерживать деятельность оператора по оказанию услуг. В случае же возникновения сбоев, перегрузок, и других проблем, приводящих к временной блокировке доступа к данным или к полной их потере, оператор не сможет обеспечить качество услуг, а при большом масштабе "аварии" — даже оказывать услуги.

Учитывая вышесказанное, при внедрении систем OSS необходимо комплексно и детально продумывать принципы работы баз данных, с которыми взаимодействуют системы OSS, от их числа и моделей данных до физической реализации и распределения по серверам, для достижения изначально стоящих целей по минимизации операционных расходов, без возникновения проблем при оказании услуг. При этом необходимо учитывать все планы по внедрению систем поддержки в компании, чтобы затем не понадобилось адаптировать следующую внедряемую систему с уже установленным их комплексом, т.е. осуществлять повторную работу.

Однако до сих пор ни один производитель систем OSS не предложил операторам комплексное решение, проработав эти вопросы. Что касается операторов связи, то для них эта задача еще более сложна. В настоящее время внедрение систем OSS производится у операторов связи одним из следующих способов:

— автоматизация только отдельных, целевых бизнес-процессов;

— покупка готового решения OSS и адаптация под него бизнес-процессов;

— анализ деятельности и разработка структуры систем OSS.

Достоинства и недостатки данных способов приведены в таблице.

При любом из трех способов внедрения систем OSS оператор сам или производитель оборудования разрабатывает модель данных, причем чаще всего для отдельной системы OSS. Отметим, что ее разработка с учетом особенностей деятельности оператора и на основе имеющихся у него данных является очень трудоемкой. Что касается вопросов организации и совместного и взаимозависимого функционирования баз данных, связанных с комплексной системой поддержки операционной деятельности, то они пока остаются нерешенными. Подлежат разработке оптимальные методы построения таких БД, порядок их выбора для конкретного случая и средства оценки эффективности их работы.

Сравнение подходов к внедрению систем OSS

Подход к внедрению систем OSS		Достоинства	Недостатки
Автоматизация отдельных бизнес-процессов	Собственная разработка	– Процессы оператора реализованы в OSS – Быстрая реакция на потребности Оператора	– Необходимость поддержки собственными силами – Немасштабируемые решения
	Специализированное ПО	– Поддержка со стороны производителя – Масштабируемость – Возможность интеграции	– Подготовка ТЗ на разработку
Покупка готового решения OSS		– Поддержка со стороны производителя – Поддержка стандартов	– Адаптация БП оператора к системам OSS – Избыточность / недостаточность функций
Самостоятельная разработка структуры систем OSS		– Процессы оператора реализованы в OSS – Последовательность внедрения соответствует потребностям	– Значительные затраты на интеграцию систем

Литература

1. **Самуйлов К.Е.** Бизнес-процессы и информационные технологии в управлении телекоммуникационными компаниями / Самуйлов К.Е., Чукарин А.В., Яркина Н.В. — М.: Альпина Бизнес Букс, 2009. — 442 с.

2. Управление сетями электросвязи в России: сборник научных материалов конференции [Электронный ресурс] / Экспо-Телеком. - М., 2009. — 1 электронный оптический диск (CD-ROM).

Тенденции развития интерфейсов терминалов сетей связи общего пользования

Приведена классификация терминальных устройств сетей общего пользования и их интерфейсов. Рассмотрены проводные и беспроводные аппаратные интерфейсы, а также тенденции их развития в долгосрочной перспективе.

Кривоногов В.М.,
Аспирант ФГУП ЦНИИС,
billyg@inbox.ru

Определения.

Терминал или пользовательское оборудование: технические средства для передачи и (или) приема сигналов электросвязи по линиям связи, подключенные к абонентским линиям и находящиеся в пользовании абонентов или предназначенные для таких целей.

Сеть связи — технологическая система, включающая в себя средства и линии связи и предназначенная для электросвязи или почтовой связи.

Сеть связи общего пользования предназначена для возмездного оказания услуг электросвязи любому пользователю услугами связи на территории Российской Федерации и включает в себя сети электросвязи, определяемые географически в пределах обслуживаемой территории и ресурса нумерации и не определяемые географически в пределах территории Российской Федерации и ресурса нумерации, а также сети связи, определяемые по технологии реализации оказания услуг связи.

Сеть связи общего пользования представляет собой комплекс взаимодействующих сетей электросвязи, в том числе сети связи для распространения программ телевизионного вещания и радиовещания.

Интерфе?йс (от англ. interface — поверхность раздела, перегородка) - совокупность средств и методов взаимодействия между элементами системы.

В зависимости от контекста, понятие применимо как к отдельному элементу (интерфейс элемента), так и к связкам элементов (интерфейс сопряжения элементов).

Этот термин используется практически во всех областях науки и техники. Его значение относится к любому сопряжению взаимодействующих сущностей. Под интерфейсом понимают не только устройства, но и правила (протокол) взаимодействия этих устройств.

В контексте отдельного элемента интерфейс элемента противоположен реализации элемента (внутреннему устройству и функционированию). Пользователю элемента незачем знать, как реализован используемый элемент, чтобы управлять им, но используемый элемент должен предоставить интерфейс управления. Например, водителю вовсе не обязательно знать, как устроен двигатель, чтобы управлять автомобилем, достаточно пользоваться интерфейсом автомобиля (рулем и педалями).

Классификация терминалов сетей связи общего пользования.

По степени мобильности: носимые; возимые; стационарные.

По использованию среды передачи: проводные; беспроводные.

По месту использования: наземные; спутниковые.

По дальности действия: до 10 м; до 300 м; до 1000м; более 1000 м.

По диапазону частот: НЧ; ВЧ; СВЧ; оптический диапазон.

По способу представления информации: аналоговые и цифровые.

По типу сети доступа: с коммутацией каналов; с коммутацией пакетов; смешанные.

По типу связи: симплексные; полудуплексные; дуплексные.

По типу передаваемой информации: для приема или приемопередачи голоса; для приема или приемопередачи видеоизображения и звука; для передачи данных; терминал для приемопередачи голоса, видеоизображения и звука и данных.

Интерфейсы терминалов сетей общего пользования. Классификация интерфейсов.

Основной задачей интерфейса терминала является обработка входного сигнала и передача его на устройство, обрабатывающее информацию.

Классификация интерфейсов.

По способу реализации:

— аппаратный — неотъемлемая часть терминала, предназначенная для взаимодействия с другими терминалами или сетью связи, или для взаимодействия других аппаратных составляющих терминала;

— программно-аппаратный — внешнее устройство, взаимодействующее с терминалом по средством аппаратного интерфейса, предназначенное обеспечить взаимодействие терминала с другими терминалами или сетью связи;

— программный — интерфейс взаимодействия программного обеспечения или взаимодействия терминала с пользователем.

По использованию среды передачи: проводные; беспроводные.

По диапазону частот: НЧ; ВЧ; СВЧ; оптический диапазон.

По типу связи: симплексные; полудуплексные; дуплексные.

По способу представления принимаемой (или) передаваемой информации: аналоговые; цифровые.

Аппаратные интерфейсы.

Проанализировав структурные и принципиальные схемы различных терминальных устройств можно отметить следующие:

Аппаратную составляющую терминала сети связи общего пользования можно условно разделить на 2 части:

— интерфейс взаимодействия с сетью связи или аппаратный интерфейс;

— устройство обрабатывающее информационный сигнал.

Следовательно, от аппаратного интерфейса зависит то, с какой сетью связи сможет взаимодействовать терминал, и с какой скоростью принимать и обрабатывать информацию.

В современных терминальных устройствах применяются высокоскоростные сигнальные процессоры в качестве блока цифровой обработки сигнала (далее ЦОС) аппаратного интерфейса, а устройством, обрабатывающим информацию, является процессор или микроконтроллер (в зависимости от скорости и сложности выполняемых задач).

Тенденции развития интерфейсов терминалов сетей общего пользования.

Основными направлениями развития сетей связи общего пользования можно считать:

1. увеличение скорости передачи данных, что невозможно без увеличения быстродействия терминальных устройств (увеличение производительности центрального процессора и увеличение быстродействия сигнальных процессоров в блоках ЦОС).

Осенью 2009 г. был утвержден стандарт 802.11 n, поддерживающий скорость передачи данных до 600 Мбит/с. Рабочая группа 802.16 готовит к 2010 г. утверждение релиза 2.0 стандарта 802.16.e с анонсированной скоростью передачи данных до 100

Мбит/с. Так же готовиться к утверждению стандарт 802.16.m к 2011 г.; скорость передачи данных составит до 300 Мбит/с.

2. Отказ от проводных сетей доступа. Повсеместное внедрение беспроводных локальных вычислительных сетей (далее ЛВС). Например, сети стандарта 802.11 и 802.16. Даже интерфейсы периферийных устройств переходят на беспроводную среду передачи. Пример: Создана рабочая группа 802.15.3 (Wireless USB — беспроводная универсальная последовательная шина).

3. Возможность работы терминальных устройств с различными сетями связи. Последнее возможно несколькими способами. Вот некоторые из них:

Использование дополнительных программно-аппаратных средств для обеспечения доступа терминальных устройств к сервисам сети.

Когда только начали появляться персональные компьютеры (далее ПК), для объединения их в локально вычислительную сеть потребовалось использование модемов (периферийный программно-аппаратное средство). Позднее компания Intel интегрировала в наборы логики материнских плат сетевую карту. В набор логики для мобильных ПК компания Intel интегрирует кроме сетевой платы модем для взаимодействия с ТФОП. При появлении стандарта 802.11 устройства взаимодействия с этими сетями существовали только в виде периферийных программно — аппаратных средств. С 2005 года интерфейс 802.11 интегрируется в некоторые модели терминалов сотовых сетей. Для ПК и мобильных ПК поддержка работы с сетями 802.11 a/b/g и 802.16.e так и не была включёна в набор логики материнской платы. Компания Lenovo добавила модуль 3G сетей в нетбук IdeaPad S10-2-3G (построен на базе платформы Intel Atom, аппаратный интерфейс 3G не интегрирован в набор логики).

Обеспечение поддержки доступа к сервисам сети средствами терминальных устройств, изменение или внесение изменений в интерфейс взаимодействия терминального устройства с другими устройствами или сетевыми компонентами.

Компания MediaTek анонсировала три новые микросхемы для работы с беспроводными сетями стандарта IEEE 802.16e WiMAX. Они полностью соответствуют всем требованиям спецификаций WiMAX Forum Wave2.

Использование дополнительных программно-аппаратных средств для обеспечения доступа терминальных устройств к сервисам сети.

В связи с началом повсеместного распространения корпоративных, домашних и общественных беспроводных сетей, работающих в стандартах Wi-Fi, WiMAX и 3GPP (3rd Generation Partnership Project), возникает необходимость поддержки мобильными платформами связи в разных частотных диапазонах и стандартах. Постоянная и повсеместная связь скоро станет неотъемлемой частью жизни конечных пользователей, которым необходимы коммуникаторы и другие мобильные устройства, чтобы предоставить им лучший в своем классе уровень сервиса. Для реализации этой концепции отрасль должна решить серьезные технические проблемы, не самая последняя из которых — разработка технологий интеллектуального роуминга, которые позволят организовать "прозрачное" переключение как между точками доступа одной беспроводной сети, так и между разными беспроводными сетями.

Для непрерывного доступа необходима мобильность и поддержка переключения как в гомогенных, так и гетерогенных средах передачи. Гетерогенность заключается в возможности переключения между различными сетями — WLAN, WiMAX и сотовыми сетями. Гомогенность — это возможность "прозрачного" переключения между разными точками подключения одной сети (точками доступа беспроводной локальной сети или базовыми станциями WiMAX).

В случае переключения в гомогенной среде беспроводной локальной сети мобильные платформы в первую очередь должны автоматически определять структуру беспроводной среды, в которой находится пользователь, и выбирать точку доступа с наилучшим качеством связи. Во-вторых, необходимо обеспечить доступ к ресурсам, чтобы предоставить гарантированное качество обслуживания, а также выполнить аутентификацию — или до, или в процессе переключения.

IEEE 802.11k и 802.11r — основные отраслевые стандарты, которые в настоящее время находятся в стадии разработки. Они должны обеспечить базовый набор сервисов (Basic Service Set, BSS) при "прозрачных" переключениях между точками доступа беспроводной локальной сети. Стандарт IEEE 802.11k позволяет найти наилучшую из возможных точку доступа, а стандарт IEEE 802.11r определяет механизмы обеспечения безопасности и быстрого переключения между точками доступа в рамках одного и того же расширенного набора сервисов (Extended Service Set, ESS).

Стандарт IEEE 802.21 предназначен для того, чтобы обеспечить "прозрачное" переключение пользователей между разными сетями. Он определяет функцию переключения, независимого от носителя (media independent handover, MIH), для уровней 2 и 3 модели протоколов передачи данных OSI (Open Systems Interconnection, взаимодействие открытых систем), что должно обеспечить мобильную связь в гетерогенных сетях. Стандарт IEEE 802.21 предусматривает механизмы, позволяющие различным клиентским устройствам динамически переключаться между различными типами сетей — Wi-Fi, WiMAX и сотовыми сетями, а также выбирать оптимальный тип подключения, что может существенно повысить возможности мобильной связи. Для того чтобы обеспечить повсеместное распространение различных беспроводных сетей для передачи голоса, видео и цифровых данных, корпорация Intel возглавила группу IEEE по разработке этих стандартов и сейчас занимается проектированием компонентов, которые должны предоставить непрерывный мобильный доступ к беспроводным сетям для мобильных платформ Intel.

Вывод: На сегодняшний день тема исследования и разработки интерфейсов терминалов весьма актуальна, так как определяет возможности работы терминальных устройств с активно развивающимися сетями связи общего пользования.

Литература

1. ISO/IEC 27002 Information technology — Security techniques — Code of practice for information security management — Информационные технологии. Технологии безопасности. Практические правила менеджмента информационной безопасности.

2. http://tssonline.ru/main.php Технологии и средства связи — телекоммуникации, широкополосный доступ, телефония, беспроводные сети, абонентский доступ, кабели связи.

3. http://www.compnets.narod.ru/content.html- Учебный курс: "Компьютерные сети".

4. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 3-е изд. / В. Г. Олифер, Н. А. Олифер. — СПб.: Питер, 2008. — 958 с..

5. Сети UMTS. Архитектура, мобильность, сервисы. / Кааранен Х., Ахтианен А., Лаитинен Л., Найян С., Ниеми В. — М.: Техносфера, 2007. — 464 с.

6. Конвергенция сетей связи в российских условиях. / А.А. Иванов, В.А. Соколов, Д.С. Терентьев, С.М. Ярлыкова. Журнал "Технологии и средства связи" №5-2006.

7. NGN принципы построения и организации /Под ред. Ю.Н. Чернышова. — М.: Эко- Трендз, 2008. — 400 с.

8. Системы и сети радиодоступа. / Григорьев В. А., Лагутенко О.И., Распаев Ю. А. — М.: Эко-Трендз, 2005. — 384 с.

9. Платформы для непрерывного мобильного доступа к беспроводным сетям. / Кейпил Суд (Kapil Sood), Эмили Х. Кью (Emily H. Qi), Вайвик Г. Гапта (Vivek G. Gupta) — Журнал Technology@Intel.

10. http://wimax.livebusiness.ru/tags/802.16E.

11. Федеральный закон о связи. 7 июля 2003 г. — № 126-ФЗ.

Задачи оптимизации оптоволоконных сетей связи на участке абонентского доступа

Исследуются вопросы оптимизации оптоволоконных сетей связи, построенных на основе технологий Ethernet и пассивной оптической сети (Passive Optical Network, PON). Приводятся особенности наиболее широко применяемых топологий построения сетей абонентского доступа. Рассматриваются положительные и отрицательные стороны данных технологий с точки зрения поставщика услуг связи. Обсуждается вопрос о необходимости оптимизации сети.

Егоршев Д.Е.,
Системный администратор
ООО "Меридиан Экспресс Интернэшнл",
unexperienced@rambler.ru

Введение

В настоящее время проектируется и вводится в эксплуатацию все возрастающее количество сетей с использованием оптоволоконного кабеля. Большую роль играет оптоволокно и в широкополосных сетях доступа. И следует отметить, что сеть доступа, это наименее медленно обновляющийся участок сети связи. Так как инвестиции в инфраструктуру сети доступа (кабель и его прокладку) являются фундаментальными, то при проектировании сети необходимо учитывать, не станет ли использование той или иной технологии доступа узким местом в будущем [1].

Задачи оптимизации оптоволоконной сети заключаются в выборе такой структуры сети и показателей ее узлов, при которых выбранная характеристика максимальна. Главными целями для операторов сети могут быть: поддержание постоянной скорости передачи данных, уменьшение задержек и джиттера при передаче трафика реального времени, увеличение отказоустойчивости сети на выбранных участках и уменьшение стоимости развертывания и обслуживания сетевой инфраструктуры.

Выбирая между проводными технологиями передачи сигнала до абонента, оператор сети связи европейского региона будет строить сеть доступа на базе одной из двух наиболее перспективных: Ethernet или PON.

1. Проблемы выбора технологии построения сети доступа

Так как строительство сети доступа — это процесс капитальный, трудо- и ресурсоемкий, то подходить к нему следует основательно. Принято считать, что главные затраты при развертывании оптоволоконной сети до дома (Fiber to the home, FTTH) приходятся на строительство инфраструктуры, т.е. на работы, связанные с прокладкой кабеля и строительством помещений для размещения оборудования [2]. Стоимость самого кабеля и размещенного на линии оборудования будет намного ниже.

Замена вышедшего из строя или устаревшего активного оборудования не обойдется слишком дорого оператору связи, однако внесение изменений в инфраструктуру после завершения прокладки кабеля потребует больших затрат на строительные работы. Статистика [2] показывает, что потребности пользователей в пропускной способности канала растут в геометрической прогрессии. В настоящее время многие поставщики услуг доступа в Интернет предлагают для частных абонентов доступ со скоростью 100 Мбит/с, а за рубежом предлагают доступ со скоростью 1 Гбит/с.

Длительный срок службы и большие затраты на построение оптоволоконной сети предъявляют высокие требования к правильному выбору топологии и технологии.

Для сетей на базе технологии Ethernet в основном применяют топологии "звезда" и "кольцо", в то время как сети на базе технологии PON всегда строятся по топологии "древо".

Сети на базе Ethernet вытеснили такие технологии как Token Ring за счет своего быстрого развития и снижения стоимости активного оборудования в расчете на одного абонента.

Услуги передачи данных по сети Ethernet и Ethernet-коммутация стали приносить доход на рынке корпоративных сетей и привели к снижению цен, появлению законченных продуктов и ускорению освоения новых решений [2]. Даже начальные проекты обеспечили прекрасную устойчивость к различного рода повреждениям кабеля и оказались весьма рентабельными. К недостаткам кольцевой топологии можно отнести разделение полосы пропускания внутри каждого кольца доступа, и снижение пропускной способности при подключении нового абонента, что в свою очередь порождает трудности масштабирования архитектуры. Архитектура "звезда" предполагает наличие выделенных оптоволоконных линий от каждого оконечного устройства абонента к коммутатору.

Сравнительный анализ технологий APON, EPON, GPON и 10G-EPON

Характеристики	APON (BPON)	EPON	GPON	10G-EPON
Дата принятия стандарта	октябрь 1998	июль 2004	октябрь 2003	Сентябрь 2009
Стандарт	ITU-T G.981.x	IEEE 802.3ah	ITU-T G.984.x	IEEE 802.3av
Скорость передачи, прямой/обратный поток, Мбит/с	155/155 622/155 622/622	1000/1000	1244/155,622,1244 2488/622,1244,2488	10.3125/1.25 10.3125/10.3125
Базовый протокол	ATM	Ethernet	SDH	Ethernet
Максимальное число абонентских узлов на одно волокно	32	16	64 (128)	1024

При построении сети PON используются устройства разветвители (Splitters), с помощью которых оптоволоконная линия разделяется между абонентами с неким коэффициентом. Максимальное количество абонентов, которых можно подключить к 1 линии зависит от конкретной технологии PON и достигает 1000 в случае стандарта IEEE 802.3av [5]. Архитектура FTTH на базе EPON построена на использовании протокола Ethernet (см. таблицу).

2. Основные преимущества архитектуры на базе PON

Перечислим основные преимущества архитектуры на базе PON.

1) Экономия оптоволоконного кабеля.

Пассивные оптические сети требуют прокладки существенно меньшего количества жил от устройства Optical Line Termination (OLT) до устройств Optical Network Termination (ONT) устанавливаемых у пользователя, благодаря использованию разветвителей. Также возможна воздушная прокладка оптического кабеля, с вводимыми ограничениями на количество жил.

При отсутствии существующей инфраструктуры или развертывании сети в новых районах экономия на оптоволоконном кабеле невелика, поскольку увеличение затрат на использование кабеля с большим количеством жил не может сравниться со стоимостью рытья траншей и необходимостью использования чужой инфраструктуры, например, канализационных коллекторов [2].

2) Топология сети PON

Технология PON очень хорошо подходит для построения сети в коттеджных поселках для подачи услуги Triple Play. Конечно, возможно тянуть в каждый дом отдельное волокно, но в данном случае выгоднее строить сеть по технологии PON. Также, PON может быть применим при модернизации существующих сетей кабельного телевидения. В большинстве своем при их строительстве использовался оптический кабель с малым количеством оптических жил (от 4 до 8). Кроме того, топология кабельного телевидения совпадает с топологией дерева PON. Следовательно, оператор кабельного телевидения, выбрав PON, может быстро развернуть услуги широкополосного доступа в Интернет и начать предоставлять услуги Triple Play, т.к. телевидение в инфраструктуре PON может передаваться по отдельной длине волны [1].

3) Введения новых услуг без нужды обновлять оборудование на линии.

Это возможно благодаря использованию другой длины волны для передачи данных для этой услуги и полному отделению всех потоков данных друг от друга.

При развертывании архитектуры PON поставщики услуг связи сталкиваются с несколькими проблемами:

— Полоса пропускания одного канала сети PON распределяется между всеми абонентами, подключенными к одному "дереву". Чем больше абонентов подключено в одном "дереве", тем меньше затраты на кабель, но меньше и пропускная способность в расчете на одного абонента.

— Проблема защиты трафика. В отличие от обычной Ethernet-коммутации, в PON используется общая среда передачи, поэтому требуется шифрование нисходящего потока данных. Для этого используется алгоритм шифрования Advanced Encryption Standard (AES) [7] с 256-разрядным ключом для нисходящего потока. Однако использование функций шифрования влияет на производительность устройств и их стоимость.

— При каждом разветвлении оптического кабеля в соотношении 1:2 уровень сигнала падает примерно на 3,4 дБ [2]. Таким образом, накладывается ограничение на максимальное количество ветвлений одной линии PON.

Пассивные оптические разветвители не имеют возможности передавать служебную информацию о неполадках в сети, что затрудняет поиск и выявление неисправности на линии.

Таким образом, видно, что у технологии PON, которая считается весьма перспективной технологией абонентского доступа, присутствует также немалое количество недостатков. Часть из них, в частности нехватка пропускной способности канала, исправляются в новых версиях рекомендаций Сектора стандартизации Международного союза электросвязи и IEEE, однако позже опять напоминают о себе. Также, стандарты и спецификации выходят чаще, чем оператор сети связи может позволить себе полную модернизацию кабельной инфраструктуры и оборудования на линии. И даже проектирование сети с большим запасом "прочности" отнюдь не гарантирует готовность сети к предоставлению новых услуг все увеличивающемуся количеству абонентов.

В этом случае поставщикам услуг связи пригодится возможность анализировать состояние и динамику изменений уже запущенных в эксплуатацию участков сети и, на основании данных о среде, в которой предстоит строить новый участок, производить процесс оптимизации сети.

Заключение

С учетом вышеприведенных факторов, задачи оптимизации оптоволоконной сети будут актуальны для обеих технологий. При этом, параметры оптимизации для Ethernet и PON могут отличаться. Для сети на базе технологии PON поставщики услуг связи смогут использовать оптимизацию по наименьшему значению затуханию сигнала, в зависимости от мощности установленного оборудования и расположения разветвителей. Или оптимизацию по наименьшей стоимости инфраструктуры сети, в зависимости от количества используемых оптоволоконных линий до OLT, расположения оконечного устройства и разветвителей и потребностей каждого абонента в пропускной способности канала. Для сети на базе технологии Ethernet большое значение будет играть оптимизация по наименьшему показателю задержки и джиттера при передаче трафика реального времени, такого как IP-телефония, видео- и аудиоконференции, а также оптимизация по стоимости, с учетом установленного коммутационного оборудования, топологии сети, требуемой пропускной способности до абонента и типов предоставляемых услуг.

Для поставщиков услуг связи проведение подобных мероприятий становится особенно актуальным после введения сети в эксплуатацию и, соответственно, отсутствия возможности вносить в нее крупные изменения без больших издержек на проведение строительных работ.

Литература

1. **Шилов Д.** Ethernet или PON / Шилов, Д. — Кабельщик, 2009. — №3.

2. **Гасымов И.** Архитектура оптических сетей доступа FTTH (Fiber-to-the-Home) / Гасымов И. — Официальный документ компании Cisco, 2007.

3. **Иванов В.** Сейчас наступил оптимальный момент для разворачивания сетей EPON / Ивано В. — Connect! Мир Связи, 2006. — №11.

4. **Сергеев А.** PON: многообещающая "пассивность" / Сергеев А. — Connect! Мир Связи, 2004. — №8.

5. 10Gb/s Ethernet Passive Optical Network standard // IEEE standard 802.3av / 2009.

6. Ethernet in the First Mile (EFM) // IEEE standard 802.3ah / 2004.

7. **Daemen J.** AES Proposal: Rijndael / J. Daemen, V. Rijmen // NIST Rijndael Specification / 2003.

8. **Бродский Н.Ю.** Тенденции и перспективы распространения технологий широкополосного доступа в России / Бродский, Н.Ю. — Вестник связи, 2008. — №9.

Моделирование сети следующего поколения при оценке качества инфокоммуникационных услуг

Предлагается модель функционирования сети связи следующего поколения. Модель наглядно представляет аспекты, влияющие на характеристики качества инфокоммуникационных услуг на базе открытых интерфейсов. Получены математические функции этих характеристик на уровне средних значений. Модель позволяет также исследовать характеристики на уровне распределений.

Шалагинов В.А.,
старший инженер Технопарка ФГУП ЦНИИС,
аспирант ФГУП ЦНИИС,
shalaginov@zniis.ru

Цель расчета качественных характеристик

Целью расчета характеристик качества услуг является переход от субъективных методов оценки качества передачи речи, существующих в настоящее время к объективным оценкам, которые могут быть численно выражены. Контроль значений характеристик качества позволит обеспечить гарантированное качество обслуживания для инфокоммуникационных услуг в сетях следующего поколения.

Характеристики качества инфокоммуникационных услуг

Характеристики качества инфокоммуникационных услуг, предложенные в рекомендации МСЭ-T Y.1540 [1] предназначены для измерения параметров передачи пакетов IP. В рекомендации Y.1541 [2] предложены численные значения этих параметров. В зависимости от численных значений характеристик определены классы качества обслуживания.

Для инфокоммуникационных услуг целесообразно использовать многомерный вектор. Определим *n*-мерный вектор $\overleftarrow{a}, n < \infty, (a_1 1, a_1 2, ..., a_1 n)$. Координаты вектора $\overleftarrow{a}(a_1, a_2, ..., a_n)$ — являются характеристиками качества инфокоммуникационных услуг.

В качестве примера рассмотрим значения векторов многомерного критерия десятимерного вектора $\overleftarrow{a}, n = 10, (a_1 1, a_1 2, ..., a_1 10)$, где

a_1 — требуемая ШПП, Мб/с;

a_2 — задержка передачи пакета, мс;

a_3 — вариация задержки при передаче пакета, мс;

a_4 — коэффициент потерь пакетов;

a_5 — доля пакетов содержащих ошибку;

a_6 — коэффициент разупорядоченности пакетов

a_7 — интервал времени, от отправки запроса до начала предоставления услуги (включая время на передачу сигнальной информации для предоставления услуги), мс;

a_8 — интервал времени, в процессе предоставления услуги для изменения характеристик услуги (включая время на передачу сигнальной информации для вступления в силу изменений), мс;

a_9 — тип передачи: в реальном времени/не в реальном времени.

a_{10} — вероятность отказа при запросе и в процессе предоставления услуги.

Рассмотрим более подробно характеристики $a_2,...,a_5$, приведенные в рекомендации МСЭ-T Y.1541 и имеющие привязку к существующим классам качества обслуживания. Согласно рекомендации каждый класс качества обслуживания образован с помощью комбинации значений определенных характеристик.

Задержка передачи пакета — это время между двумя событиями, одним из которых является момент отправки пакета в сетевой сегмент от источника (t_1), вторым — момент прибытия пакета к узлу назначения (t_2). При необходимости фрагментация пакета для передачи в следующую сеть, для которой размер пакета является слишком большим, время на фрагментацию также входит в величину t_2. Под фрагментацией понимается разбиение слишком длинного для конкретной сети пакета на более короткие пакеты, с созданием соответствующих служебных полей для последующей сборки фрагментов в исходное сообщение. При этом всегда выполняется соотношение $t_2 > t_1$. Задержка передачи пакета может быть измерена и для пакетов, которые содержат ошибку.

Вариация задержки при передаче пакета вычисляется для двух узлов сети. Вариация задержки определяется как разброс задержки очередного пакета по отношению к предыдущему. Величина вариации задержки пакета равна разности $\upsilon_{i,i-1} = x_i - x_{i-1}$, где x_i — задержки при передаче *i*-го и (*i*– 1)-го пакетов соответственно.

Коэффициент потерь пакетов (k_1) — это отношение общего числа потерянных пакетов к общему числу переданных пакетов в измеряемой совокупности. В число переданных пакетов входят три типа пакетов:

— пакеты, которые будут переданы без ошибок (их число обозначим p_s);

— пакеты, переданные с ошибкой (их число обозначим p_m);

— пакеты, которые будут потеряны (их число обозначим p_1).

Таким образом,

$$k_1 = \frac{p_1}{p_s + p_m + p_1}.$$

Доля пакетов, содержащих ошибку (k_m) — это отношение числа пакетов с ошибками к числу всех переданных пакетов, в измеряемой совокупности.

$$k_m = \frac{p_m}{p_s + p_m}.$$

Коэффициент разупорядоченности пакетов — это отношение числа успешно переданных, но разупорядоченных пакетов, к общему числу успешно переданных пакетов. Разупорядоченность пакетов может быть обнаружена в случае, когда номер одного из пакетов в последовательности полученных пакетов превосходит номер ожидаемого пакета. Разупорядоченным пакетом в таком случае считается пакет, номер которого равен ожидаемому, но данный пакет пришел позже пакета, с превосходящим номером. Интервал между разупорядоченными пакетами может определяться как временем, так и количеством байт пакета полученного раньше пакета, который считается разупорядоченным [3].

Модель сети следующего поколения при предоставлении инфокоммуникационных услуг

Пример модели сети следующего поколения при предоставлении инфокоммуникационных услуг приведен на рис. 1. На схеме

указаны уровни сети следующего поколения, задействованные в процессе предоставления инфокоммуникационных услуг. На каждом из уровней приведены элементы оказывающие влияние на качество инфокоммуникационных услуг. Показаны интерфейсы передачи данных 1 и 10 Gigabit Ethernet (GE).

Пример расчета средних значений задержки

Задержка передачи пакета данных является важным параметром для инфокоммуникационных услуг реального времени. В современной литературе используется термины "эластичный трафик" и "потоковый трафик" [4], применительно к трафику услуг некритичных к задержкам и критичных к задержкам (услуги реального времени) соответственно.

Задержка передачи пакета данных, возникающая при предоставлении инфокоммуникационной услуги, будет представлять собой сумму:

$$T = T_A + T_T + T_R,$$

где T_A, T_T, T_R — задержка в сети доступа, транспортной сети и на обработку услуги.

Рассмотрим подробнее эти три типа задержки и вычислим примерное числовое значение задержки при передаче пакета с данными.

В случае постоянного времени обслуживания и предоставления услуг одним сервером для каждой услуги известна формула [5]:

$$T_R = \frac{T_S(2-\rho)}{2(1-\rho)},$$

где T_S — среднее время обслуживания одного запроса (для веб-сервера составляет порядка 20 мс), ρ — коэффициент использования сервера.

Задержка в транспортной сети T_H при каждом переходе представляет собой сумму задержек сериализации, распространения и коммутации [6]:

$$T_H = T_S + T_D + T_K,$$

Задержка сериализации (T_1S) — это время, которое требуется устройству для передачи данных при заданной ШПП

$$T_S = A/R,$$

где A — количество бит, R — скорость передачи.

Задержка распространения (T_D) — время, которое требуется данным для достижения принимающего устройства на другом конце канала. Это время зависит от расстояния и типа среды передачи. При передаче данных на большие расстояния задержка распространения из конца в конец составляет порядка нескольких десятков миллисекунд:

$$T_D = D/R,$$

где D — длина сетевого пути.

Задержка коммутации (T_K) — время, которое требуется устройству, получившему пакет, для его передачи следующему устройству. В современных устройствах значение не превышает 10 нс.

Задержка в транспортной сети вычисляется по формуле:

$$T_T = T_H \cdot N,$$

где N — количество переходов между узлами сети.

Задержка в сети доступа (T_1A) в значительной степени зависит от применяемой технологии доступа. Исследования показали [7], что задержка в реальной сети при использовании технологии DSL составляет в среднем 65 мс.

Результаты измерений задержки на фрагменте сети Интернет

Рассмотрим графики распределений, полученные в ходе измерений задержки передачи пакетов. Задержка измерялась при прохождении пакетов по пути, состоящем из семи узлов, количество измерений равно 100.

Схема испытаний представлена на рис. 2.

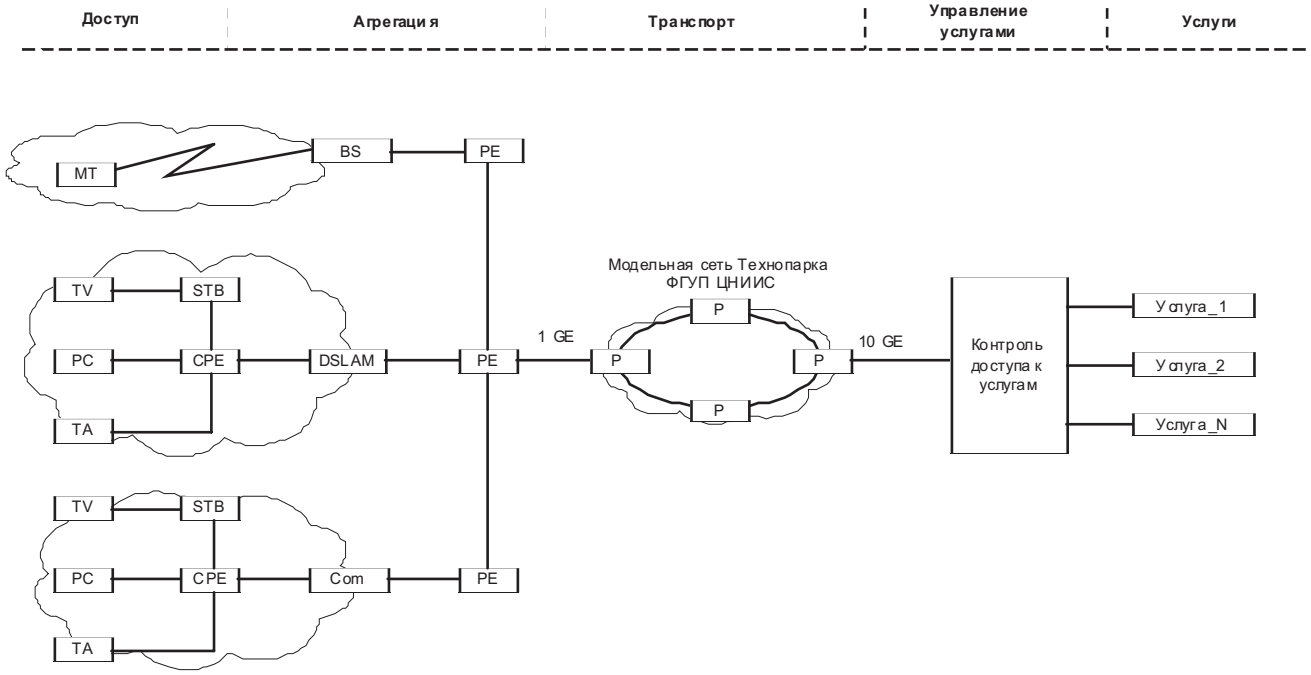


Рис. 1. Модель сети следующего поколения при предоставлении инфокоммуникационных услуг
Обозначения: MT — мобильный терминал; TA — телефонный аппарат; BS — базовая станция; PC — персональный компьютер; TV — дисплей (например, телевизор); Com — коммутатор; STB — приставка пользователя; P, PE — маршрутизатор уровня P и PE; CPE — оборудование в помещении пользователя; DSLAM — мультиплексор доступа цифровой абонентской линии

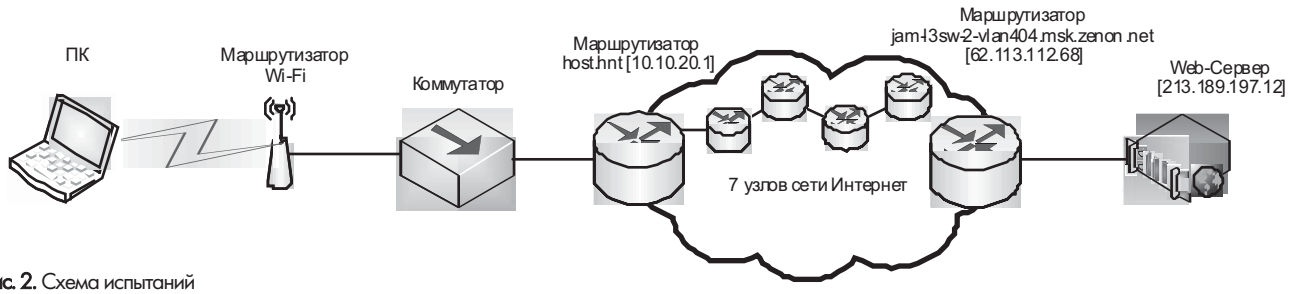


Рис. 2. Схема испытаний

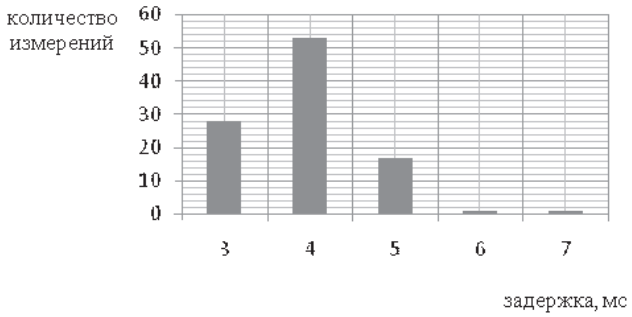


Рис. 3. Измерение задержки для кадров размером 64 байта

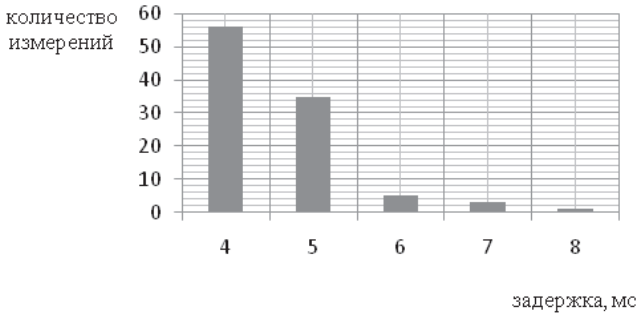


Рис. 4. Измерение задержки для кадров размером 512 байт

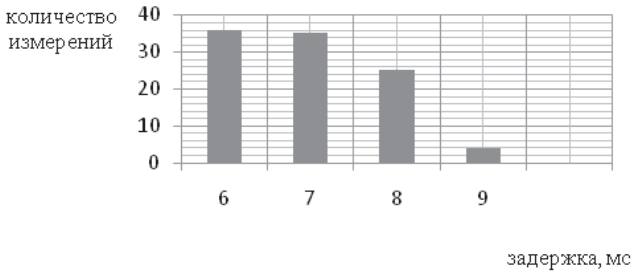


Рис. 5. Измерение задержки для кадров размером 1518 байт

Получены результаты для размеров кадров 64 байт, 512 байт и 1518 байт, приведенных в документах RFC 2544 [8] и RFC 5180 [9]. Гистограммы измерений приведены на рис. 3, 4 и 5.

Из гистограмм видно, что задержке соответствует усеченное нормальное распределение. С увеличением размера кадра происходит смещение графика вправо по оси x, средняя задержка передачи пакета увеличивается. В настоящее время на модельной сети Технопарка ФГУП ЦНИИС готовится проведение измерений других характеристик качества по схеме приведенной на рис. 1.

Выводы и рекомендации

Измерение характеристик качества становиться важной задачей в сетях следующего поколения для инфокоммуникационных услуг реального времени, передающих потоки данных мультимедиа. Для успешного внедрения таких услуг и предоставления гарантированного качества необходимо вычисление характеристик качества из конца в конец. Испытания на модельной сети позволит определить законы распределения характеристик качества. Предварительные измерения показали, что задержке передачи пакетов соответвует усеченное нормальное распределение.

Литература

1. Internet protocol aspects — Quality of service and network performance, Network performance objectives for IP-based services, // Рекомендация МСЭ-Т Y.1541, 2006.

2. Internet protocol aspects — Quality of service and network performance, Internet protocol data communication service — IP packet transfer and availability performance parameters // Рекомендация МСЭ-Т Y.1540, 2007.

3. Packet Reordering Metrics // IETF RFC 4737, 2006.

4. **Наумов В.А., Самуйлов К.Е., Яркина Н.В.** Теория телетрафика мультисервисных сетей // М.: РУДН, 2007.

5. **Столлингс В.** Современные компьютерные сети // СПб.: Питер, 2003.

6. **Шринивас В.** Качество обслуживания в сетях IP // М.: Издательский дом "Вильямс", 2003.

7. Delay Characteristics of High-Speed Internet Access Network // ETRI, 2003, http://www.apnoms.org/2003/slide/Technical-Session/T1-3.pdf.

8. Benchmarking Methodology for Network Interconnect Devices // IETF RFC 2544, 1999.

9. IPv6 Benchmarking Methodology for Network Interconnect Devices // IETF RFC 5180, 2008.

Проблема учета и тарификации количества информации при предоставлении инфокоммуникационных услуг

Описан принцип передачи информации в сетях следующего поколения на базе коммутации пакетов и показаны проблемы, встающие перед операторами в части учета объема данных для тарификации при предоставлении инфокоммуникационных услуг. Приведены результаты испытаний на сетях операторов связи, и предложен способ решения имеющихся проблем.

Шалагинов В.А,
старший инженер Технопарка ФГУП ЦНИИС,
аспирант ФГУП ЦНИИС,
shalaginov@zniis.ru

Введение

В настоящее время во всем мире происходит постепенная миграция телекоммуникационных услуг от существующих, на базе сетей с коммутацией каналов, к новым инфокоммуникационным услугам, на базе сетей с коммутацией пакетов. В результате такой миграции, тарификация услуг все чаще осуществляется по количеству (объему) информации, что ставит перед операторами задачу учета трафика передачи данных. В статье описан принцип передачи информации в сетях следующего поколения и показаны проблемы учёта объема данных для тарификации при предоставлении инфокоммуникационных услуг.

Способы тарификации

Учет объема услуг с целью их тарификации осуществляется операторами, в основном, тремя способами или с помощью их комбинаций:

— по времени;

— по объему принятой и/или переданной информации;

— без ограничения объема данных с ограничением по скорости передачи.

Первые два метода обычно используются в комбинации с абонентской системой оплаты. При этом с абонента взимается абонентская плата, например, за некоторый объем данных. В случае превышения оплаченного за определенный период объема, дополнительный объем оплачивается отдельно.

Все указанные методы тарификации предполагают измерение объема информации и/или продолжительности соединения с целью тарификации. Ограничение по скорости передачи предполагает учет объема передаваемых данных за малый промежуток времени, обычно за 1 секунду (например, 5 Мбит/с).

Нормативные документы

К нормативным документам, регламентирующим учёт объема переданной и принятой информации, при предоставлении услуг по передаче данных в РФ относятся:

— Федеральный Закон "О Связи" [1];

— Правила оказания услуг связи по передаче данных [2];

— Правила применения автоматизированных систем расчётов [3].

В п. 35 "Правил оказания услуг связи по передаче данных" указано: "Оплата услуг связи по передаче данных может осуществляться по абонентской или повременной системе оплаты либо по объему принятой и (или) переданной информации".

В п. 22 "Правил применения автоматизированных систем расчётов" указано:

"В случае измерения автоматизированной системой расчётов объема переданной (принятой) информации объем информации измеряется с точностью до байта".

Модель ВОС

Для предоставления инфокоммуникационных услуг для организации передачи данных в большинстве случаев используется стек протоколов TCP/IP. Архитектура стека TCP/IP построена соответствии с моделью взаимодействия открытых систем ВОС [4]. Модель ВОС использует понятия уровней, каждый протокол стека TCP/IP может быть ассоциирован с определенным уровнем (или несколькими уровнями). Назначение модели и функции уровней широко описаны в существующей литературе, поэтому отметим лишь соотношение модели ВОС и стека TCP/IP. Модель ВОС является теоретической концепцией, а модель TCP/IP является одной из практических реализаций модели ВОС. На рис. 1 приведены уровни модели ВОС и их соотношение с уровнями стека протоколов TCP/IP.

На рис. 2 приведен пример добавления заголовков протоколов различных уровней при передаче данных от пользователя к пользователю из конца в конец. Протоколы каждого уровня добавляют за-

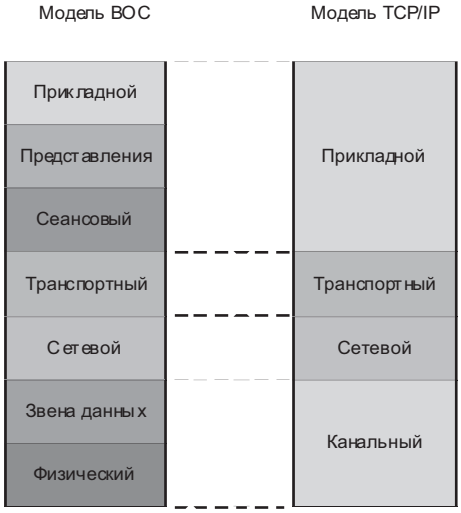


Рис. 1. Уровни модели ВОС и их соотношение с уровнями стека протоколов TCP/IP

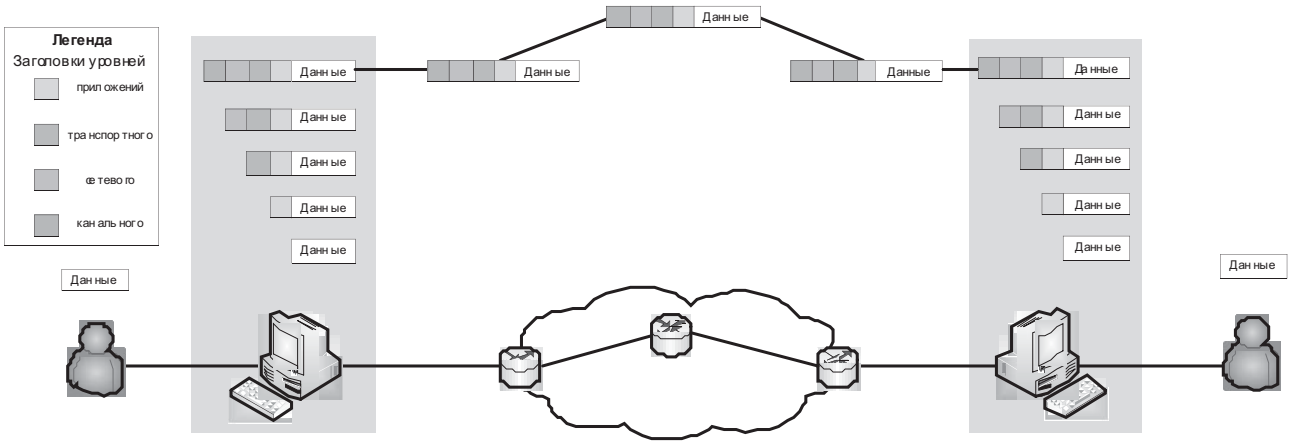


Рис. 2. Пример добавления заголовков протоколов

головки со служебной информации необходимые для передачи блоков данных.

Изучение указанных выше нормативных документов выявило отсутствие указания на способ (уровень, логику и протокол) учёта объема переданной и принятой информации, при предоставлении услуг по передаче данных. Испытания и анализ технической реализации установленных систем учета трафика на сетях операторов и оборудования передачи данных показали, что существующие различные методы учета трафика, применение которых не регламентировано документами, что может привести к неправильной тарификации инфокоммуникационных услуг в сетях следующего поколения.

Результаты измерений

Результаты измерений проведенных на оборудовании, находящимся в эксплуатации на сетях операторов связи, показывают, что учёт объема переданной и принятой информации зависит от настроек, а также от программной и аппаратной реализации оборудования. Как правило, оборудование разных производителей по-разному выполняет учёт объема информации.

Например, на рис. 3 серым цветом выделены поля пакета информации на сетевом уровне, учитываемые при использовании протокола PPPoE [4]. В статистику пиопадает объем информации, переданной на сетевом уровне.

Изменение настроек оборудования приводит к тому, что дополнительно учитывается поле Протокол PPP равное 2 или 1 байтам [6], которые добавляются к каждому передаваемому пакету информации (рис. 4).

Различия в результатах учета переданной и принятой информации в зависимости от размера данных пользователя приведены в таблице.

Разница в расчетах переданной и принятой информации превышает величину в 1 байт, допустимую согласно "Правилам применения автоматизированных систем расчётов" в Российской Федерации. Однако данная разница не противоречит "Правилам оказания услуг связи по передаче данных", в которых не указан способ учёта объема информации. Данная проблема позволяет операторам искусственно "увеличивать" объем переданных данных за счет учета служебной информации оказании услуг передачи данных.

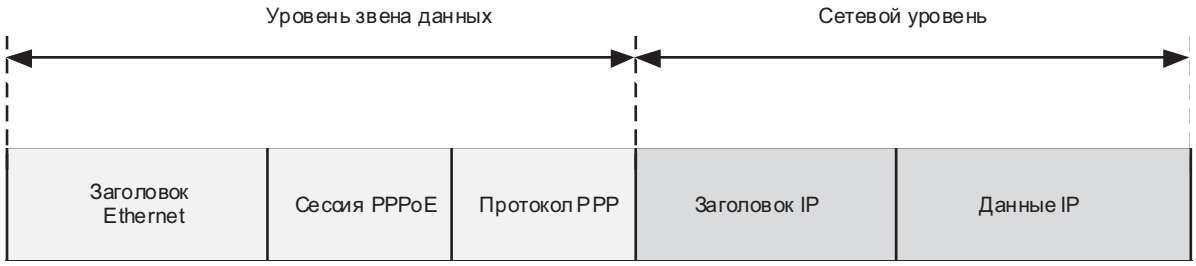


Рис. 3. Поля пакета информации на сетевом уровне, учитываемые при использовании протокола PPPoE

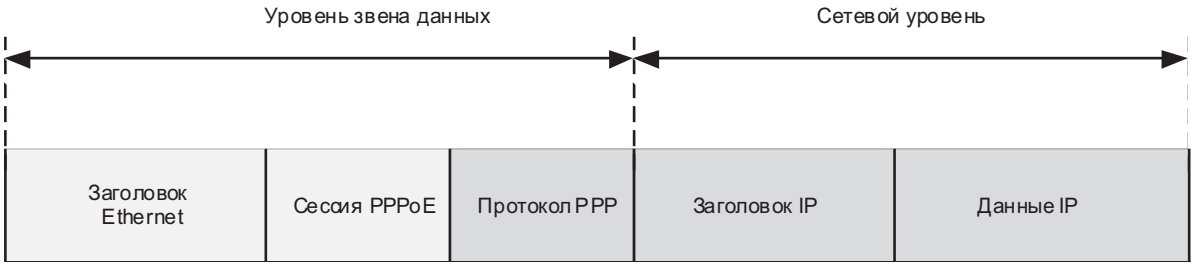


Рис. 4. Поля пакета информации на сетевом уровне, учитываемые при изменении настроек

Разница в результатах учета переданной и принятой информации

Данные пользователя	Данные учета на сетевом уровне, байт	Данные учета при изменении настроек, байт	Разница, байт
512 байт	656	706	50
10 Кбайт	11599	11725	126
100 Кбайт	115280	116126	846

Выводы

С целью защиты интересов пользователей услугами связи необходимо обеспечить единые способы учета объема переданной и принятой информации для тарификации за услуги. Выявленные разногласия в учёте информации указывают на необходимость разработки соответствующих документов регламентирующих способы учёта объема переданной и принятой информации, при предоставлении инфокоммуникационных услуг.

Литература

1. ФЗ "О Связи" 7 июля 2003 года № 126-ФЗ, с изм. 2004-2008 гг.

2. Правила оказания услуг связи по передаче данных (Утверждены постановлением Правительства Российской Федерации от 23 января 2006 г. № 32).

3. Правила применения автоматизированных систем расчётов" (Утверждены приказом Министерства информационных технологий и связи Российской Федерации от 02 июля 2007 г. № 73).

4. ГОСТ Р ИСО/МЭК 7498-1-99 Информационная технология. Взаимосвязь открытых систем. Базовая эталонная модель. Часть 1. Базовая модель // принят и введен в действие постановлением Госстандарта России от 18 марта 1999 г. № 78.

5. RFC2516 — A Method for Transmitting PPP Over Ethernet (PPPoE) // http://www.faqs.org/rfcs/rfc2516.html.

6. RFC1661 — The Point-to-Point Protocol (PPP) // http://www.faqs.org/rfcs/rfc1661.html.

Стохастическая модель минимизации длины соединительных линий телекоммуникационной сети

Рассматривается задача минимизации средней совокупной длины линий, объединяющих магистральные и абонентские узлы телекоммуникационной сети при использовании промежуточных распределительных пунктов, концентрирующих пользовательский трафик. Исследование производится на основе модели стохастической двухуровневой сети, построенной исходя из принципа минимально-достаточной совокупности узловых пунктов.

Шашаев И.Ю.,
аспирант ФГУП ЦНИИС,
i.shashaev@gmail.com

При достаточном удалении абонентских терминалов от пограничных узловых пунктов транспортного (магистрального) участка крупной, разветвленной телекоммуникационной сети целесообразно вводить промежуточный, распределительный уровень с целью сокращения совокупной длины соединительных линий на этом участке. Во многих случаях это позволяет существенно сэкономить на общей длине кабеля [1].

От рационального выбора положения распределительных пунктов (узлов) в топологической структуре сети доступа (СД) во многом зависит величина затрат оператора на построение абонентского участка сети, составляющих, как правило, наибольшую долю в общих CAPEX затратах на реализацию системы.

Исходя из сказанного, целью настоящего исследования является минимизация совокупной длины кабельных соединений сети доступа за счет оптимизации степени удаления распределительных пунктов от граничных узлов на участке между магистральной сетью и абонентскими терминалами. Оптимизации будет подлежать только параметр длины линий, так как с использованием волоконно-оптических кабелей при проектировании и строительстве линий связи, такой параметр как требуемая пропускная способность будет влиять на величину капитальных затрат в существенно меньшей степени. Это обусловлено тем, что оптические волокна обладают практически неограниченной пропускной способностью при сохранении той же стоимости линий, в отличие от медных кабелей связи [2].

Согласно работе [3] проблема решается при следующих предположениях:

- Все узлы являются абсолютно надежными. Время обработки в них пренебрежимо мало.
- Все каналы свободны от ошибок и являются абсолютно надежными. Время распространения сигнала в линиях пренебрежимо мало. Пропускные способности линий могуп принимать любые неотрицательные значения.
- В сети используется статическая маршрутизация.
- Трафик, создаваемый абонентами или поступающий в сеть из внешних источников представляет собой пуассоновский процесс со средним значением γ сообщений в секунду, длина которых распределена по экспоненциальному закону со средним значением μ бит.

Рассмотрим соответствующую оптимизационную задачу, где определение экстремума зависит только от топологических данных узлов сети.

Будем исследовать классическую звездообразную структуру СД, где имеется *M* граничных узловых пунктов магистральной сети, *N* — абонентских терминалов (в ассоциативном их понимании [4]) и вводятся *K* — промежуточных распределительных пунктов (РП).

Рассмотрим сначала некую абстрактную стохастическую задачу по минимизации усредненного значения попарного евклидового расстояния между точками двух случайных множеств, отражающих сетевые узлы двух смежных уровней.

Пусть имеются два независимых упорядоченных множества векторов $\vec{B}^{(n)}$ и $\vec{B}^{(m)}$, соответственно, *n* и *m* в 2-х мерном евклидовом векторном пространстве $\vec{A}$ над полем действительных чисел *E*, инцидентных множествам точек $B^{(n)}$ и $B^{(m)}$, соответствующих положению узлов условно нижнего- (*n*-точек) и условно верхнего — (*m*-точек) смежных сетевых уровней.

На рис.1 представлена схема сети и некоторые обозначения, используемые в дальнейшем исследовании.

Минимизируем некоторую скалярную целевую функцию $L^{(n,m)}$, отражающую суммарную длину попарных евклидовых расстояний $l_{i,j}^{(n,m)}$ между точками множеств $\vec{B}^{(n)}$ и $\vec{B}^{(m)}$ так, чтобы каждая точка $\vec{b}_i^{(n)}$ из $\vec{B}^{(n)}$ была связана только с одной точкой $\vec{b}_j^{(m)}$ из $\vec{B}^{(m)}$.

$$\mathbf{L}^{(n,m)} = \min_{\substack{n\to\infty \\ m\to\infty}} \sum_{i,j}{}^{*} \left| \vec{b}_i^{(n)} - \vec{b}_j^{(m)} \right|,$$

$$i \in (1 \div n),\ j \in (1 \div m),\ \vec{B}^{(n)} = \{\vec{b}_i^{(n)}\}_n,\ \vec{B}^{(m)} = \{\vec{b}_j^{(n)}\}_m.$$

Звездочка у знака суммы означает, что суммирование ведется не по всем комбинациям (*i, j*), а только по тем, которые соответствуют минимально необходимой связности точек из $\vec{B}^{(n)}$ с точками из $\vec{B}^{(m)}$.

Путем перехода от постановочной векторной интерпретации к матричной, осуществляя процедуру минимизации с применением методов матричного исчисления в итоге получаем итоговую формулу, отражающую минимально необходимую совокупную длину коммуникаций между двумя рассматриваемыми уровнями (*m*) и (*n*) – $L_{\min}^{(n,m)}$:

$$\boldsymbol{L}_{\min}^{(n,m)} \Rightarrow \mathbf{L}_{\min}^{(n,m)} = \alpha_B \sqrt{\frac{S_0}{m}} \cdot n, \qquad (1)$$

где α_B — масштабирующий коэффициент, S_0 — площадь двумерного пространства, в котором определены точки обоих множеств.

Одним из полезных результатов вычислений, проведенных на базе настоящей модели, является определение данных о ранге вершин R_i в точках $b_i^{(m)}$ при формировании локальных межуровневых структур, получаемых как результат сопряжения точек из $B^{(m)}$ с инцидентными им точками из $B^{(n)}$. Такие структуры в данном случае будут иметь вид множества несвязанных графов, в количестве (*m*) [5].

Оценим удельное количество вершин $R_{(r)}^{(m)}$, имеющих одинаковый ранг r, попутно оценив значение масштабирующего коэффициента α_B.

Пусть в уже упомянутом двухмерном декартовом пространстве имеется некоторая ограниченная поверхность *S* площадью S_0, на которой существует одна единственная точка $b^{(n)}$, принадлежащая $B^{(n)}$ (нижний уровень).

Предположим, кроме того, что на данной поверхности размещено по случайному закону *m* точек из множества $B^{(m)}$ (верхний уровень). Очертим вокруг точки $b^{(n)}$ некоторую окружность радиуса ρ. Тогда вероятность P_ρ того, что ни одна точка из множества $B^{(m)}$ не попадет внутрь круга, будет

$$P_c = \left(1 - \frac{\rho c^2}{S_0}\right)^m = \left(1 - \rho н^{(m)} \frac{c^2}{m}\right)^m$$

где $\nu^{(m)}$ — плотность распределения точек из $B^{(m)}$ по поверхности *S*.

Давая малое приращение радиусу ρ на величину Δρ можно определить, какова будет вероятность $P_{\Delta\rho}$ того, что в полученное кольцо площадью $2\pi\rho\Delta\rho$ попадет всего одна единственная точка из $B^{(m)}$:

$$P_{Дc} = \frac{2\rho c Дc}{S} m = 2\rho н^{(m)} Дc$$

Устремляя $\Delta\rho \to \partial\rho$ и $m \to \infty$, получаем выражение для плотности вероятностей *q*(ρ) того, что ближайшая к $b^{(n)}$ точка из $B^{(m)}$ будет от нее на расстоянии ρ:

$$q(c) = \lim_{\substack{\Delta\rho \to \partial\rho \\ m \to \infty}} \left(1 - \rho н^{(m)} c^2 \frac{1}{m}\right)^m \cdot 2\rho н^{(m)} \rho \cdot n$$

Отсюда, учитывая, что $\lim_{m \to \infty}\left(1 - \frac{x}{m}\right)^m = e^{-x}$, имеем:

$$q(c) = 2\rho н^{(m)} c e^{-\rho н^{(m)} c^2}. \qquad (2)$$

Интегрируя *q*(ρ) в пределах от 0 до ∞, приходим, но уже другим путем, к ранее приведенному выражению (1) для среднего минимального расстояния между точками $b^{(n)}$ и ближайшими к ним точками из $B^{(m)}$:

$$L_{\min}^{(n,m)} \Rightarrow \mathbf{L}_{\min}^{(n,m)} = 2\pi n \nu^{(m)} \int_0^\infty \rho e^{-\pi\nu^{(m)}\rho^2} \partial\rho = \frac{1}{2}\sqrt{\frac{S_0}{m}} n. \qquad (3)$$

Из которого видно, что α = 0,5.

Используя (2) в качестве базового, можно путем несложных преобразований найти выражения для относительного количества узлов $R_{(r)}^{(m)}$ с соответствующим рангом *r*:

$$R_{(r=0)}^{(m)} = \frac{1}{1 + \frac{\pi}{2} \cdot \frac{n}{m}},$$

$$R_{(r>1)}^{(m)} = \frac{1}{1 + \frac{\pi}{2} \cdot \mathbf{r}} \sin \pi \frac{\frac{2}{\pi \cdot \mathbf{r}} \cdot \frac{n}{m}}{1 + \frac{2}{\pi \cdot \mathbf{r}} \cdot \frac{n}{m}}.$$

Анализируя полученные результаты, нетрудно заметить, что удельное количество сетевых узлов для разных $r \geq 1$ имеет явные максимумы в области значений $\frac{n}{m} = \frac{\pi}{2} r$, что весьма важно для последующего анализа.

Возвращаясь к (1), можно предположить, что для уровня {*M, K*} имеет место:

$$\boldsymbol{L}_{\min}^{(K,M)} = \eta \alpha_B \sqrt{\frac{S_0}{M}} \cdot K, \qquad (4)$$

а для уровня {*K*, N} соответственно:

$$\boldsymbol{L}_{\min}^{(N,K)} = \eta \alpha_B \sqrt{\frac{S_0}{K}} \cdot N. \qquad (5)$$

В представленные формулы введен т.н. коэффициент искривления η, который отражает то, что реальные линии связи в условиях городской застройки имеют не прямую, а искривленную форму (для Москвы, например, η = 2,14 [4]).

Объединяя (4) и (5) можем найти общую длину коммуникаций для абонентского уровня сети.

$$\boldsymbol{L}_{\min}^{(N,K,M)} = \eta \alpha_B \sqrt{S_0} \left(\frac{K}{\sqrt{M}} + \frac{N}{\sqrt{K}}\right). \qquad (6)$$

Беря производную от $L_{\min}^{(N,K,M)}$ по *K* и приравнивая к 0 можем найти оптимальное значение $K = K_{ОПТ}$:

$$\frac{\partial \boldsymbol{L}_{\min}^{(N,K)}}{\partial K} = \eta \alpha_B \sqrt{S_0} \left(\frac{1}{\sqrt{M}} - \frac{1}{2}\frac{N}{\sqrt{K_{ОПТ}^3}}\right) = 0.$$

Отсюда,

$$\boldsymbol{K}_{ОПТ} = 0{,}8\sqrt[3]{N^2 M} \qquad (7)$$

Оценим выигрыш в длине коммуникаций на абонентском уровне за счет введения промежуточных распределительных пунктов. Исходные данные примем для одной из крупнейших московских телекоммуникационных сетей — сети "АКАДО Телеком" (КОМКОР).

На основании (7), находим:

$K_{ОПТ}$ = 2750 распределительных пунктов. (8)

Подставляя (8) в (6), получаем:

$L_{\min}^{(N,K,M)}$ = 7450 км. (9)

Используя формулу (1) для случая отсутствия распределительного уровня $(m \to M, n \to N)$ находим:

$$L_{\min}^{(N,M)} = \eta \alpha_B \sqrt{S_0} \frac{N}{\sqrt{M}} = 11200 км. \quad (10)$$

Соотнеся (10) и (9) видим, что введение распределительного уровня на абонентских участках сети приводит к экономии по длине соединительного кабеля более чем на 30%

Следует отметить, что настоящий анализ, учитывая многообразие видов телекоммуникаций на участке сети доступа, можно принять лишь ориентировочными. Однако качественный положительный эффект от внедрения распределительного уровня прослеживается весьма наглядно.

Литература

1. **Припачкин Ю.И., Тамм Ю.А.** Расчет технико-экономических характеристик мультисервисных инфокоммуникационных сетей регионального типа в условиях ограниченных априорных данных // Электросвязь, 2004. — №10.
2. **Фриман Р.** Волоконно-оптические системы связи // Техносфера, 2006. — С.266-280.
3. **Клейнрок Л.** Вычислительные системы с очередями // Мир, 1979. — 600 с.
4. **Припачкин Ю.И., Тамм Ю.А.** Математическая модель для расчета иерархических телекоммуникационных сетей // Электросвязь, 2001. — №5. — С.35-38.
5. **Тамм Ю.А.** Вычисление рангов сетевых узлов в модели инфокоммуникационных систем // Труды НИИР, 2005. — 300 с.

Применение методов многокритериальной оптимизации структурных характеристик при проектировании телекоммуникационных сетей

Работа посвящена теоретическим аспектам применения методов многокритериальной оптимизации сетевой структуры при проектировании сетей метро-уровня в условиях противоречивых требований к реализуемой системе.

Шашаев И.Ю.,
аспирант ФГУП ЦНИИС,
i.shashaev@gmail.com

Растущая важность проектов, высокая стоимость и масштабы телекоммуникационных систем предъявляют особые требования к качеству и срокам проектирования инфокоммуникационных сетей (ИКС).

Для снижения капитальных затрат оператора и обеспечения требуемых параметров связи необходимо эффективное планирование, проведение которого невозможно без использования многокритериальной оптимизации с учетом совокупности множества противоречивых требований. В частности, в процессе создания и совершенствования сети связи масштаба города решаются две неразрывно связанные между собой задачи: минимизация затрат на реализацию сети и обеспечение требуемой производительности [1].

Методология многокритериальной оптимизации при проектировании ИКС

В самом общем случае ИКС можно рассматривать как упорядоченное множество элементов, их взаимоотношений и свойств. Однозначное задание таких параметров полностью определяет систему, то есть особенности ее функционирования, структуру сети, в том числе, ее эффективность.

Рассмотрим общую методологию многокритериальной оптимизации систем, как взаимосвязанную совокупность множества допустимых проектных решений, в том числе выбора подмножества Парето-оптимальных решений и дальнейшее сужение его до единственного проектного решения.

Решение задачи выбора оптимального проектного варианта системы включает формирование множества допустимых вариантов системы, определение совокупности показателей качества, задание критерия оптимальности системы, а также выбор вариантов системы, оптимальных по заданному критерию оптимальности [2,3].

Полагается, что каждый вариант системы определяется структурой (совокупностью элементов и связей) и набором параметров. Вместе они определяют множество допустимых проектных решений. Здесь существуют противоречивые требования. С одной стороны, желательно с максимальной полнотой представить все возможные варианты системы, чтобы не пропустить потенциально лучших вариантов. С другой стороны, существуют ограничения, определяемые допустимыми затратами (времени и средств) на процесс проектирования системы.

Для формализации постановки задачи оптимального проектирования системы должно быть составлено математическое описание условий работы, структуры, показателей качества и критерия оптимальности системы в целом. Формализованная постановка задачи проектирования системы дает возможность использовать при выборе оптимальных проектных решений математические методы моделирования и многокритериальной оптимизации систем.

Процесс задания критерия оптимальности для выбора наилучшей альтернативы на множестве допустимых проектных решений системы связан с формализацией представления заказчика системы про ее оптимальность. При этом могут быть использованы два подхода: ординалистический и кардиналистический [3].

Ординалистический подход апеллирует к порядку "лучше-хуже" и основан на введении некоторых бинарных отношений на множестве допустимых альтернатив. В зависимости от структуры допустимого множества множество оптимальных решений может содержать единственный элемент, конечное либо бесконечное число элементов.

Кардиналистический подход к описанию предпочтений приписывает каждой альтернативе некоторое число, интерпретируемое как полезность (ценность) альтернативы.

С учетом существования совокупности требований к полезности системы возникают задачи оценивания и оптимизации проектных решений по совокупности показателей качества Задачи оптимизации по совокупности показателей качества называются задачами многокритериальной либо векторной оптимизации. В результате решения таких задач находится подмножество Парето-оптимальных (эффективных) проектных решений, которое в общем случае содержит не один, а несколько вариантов системы, недоминируемых по отношению строгого предпочтения.

Следует отметить, что по существу постановка многокритериальных задач оптимизации связана с заменой (аппроксимацией) представления заказчика про оптимальность системы некоторым другим понятием оптимальности, которое удается строго формализовать и свести задачу проектирования системы до конструктивной оптимизационной процедуры при векторном критерии оптимальности.

Методы выбора Парето-оптимальных решений

Парето-оптимальные решения могут быть найдены как непосредственно на множестве допустимых решений с применением введенных бинарных отношений предпочтения, так и в пространстве оценок введенных показателей качества, которое также называется критериальным пространством. Важно отметить, что для заказчика желательно по каждому показателю получить наилучшее значение. Однако на практике этот случай встречается очень редко.

Показатели качества системы могут быть трех типов: нейтральными, согласованными между собой и конкурировать между собой. В первых двух случаях оптимизация системы может осуществляться в отдельности по каждому из показателей качества. В третьем случае достигнуть потенциального значения каждого из показателей в отдельности не представляется возможным. При этом может быть достигнут лишь согласованный оптимум введенных целевых функций — оптимум по критерию Парето. Такой оптимум означает, что даль-

нейшее улучшение каждого из показателей может быть достигнуто лишь за счет ухудшения остальных показателей качества системы.

Таким образом, последовательно выполняя перебор и сравнение всех вариантов в критериальном пространстве согласно на множестве допустимых решений выделяют подмножество Парето-оптимальных вариантов системы. Остальные варианты системы являются безусловно худшими.

Нахождение Парето-оптимальных проектных решений может производится либо непосредственно, либо с использованием специальных методов, например, весового метода, метода рабочих характеристик [4,5].

При решении оптимизационных задач получается некоторая многомерная поверхность в критериальном пространстве, которая при определенных условиях совпадает с Парето-оптимальной поверхностью [3,4]. Полученная в критериальном пространстве Парето-оптимальная поверхность связывает между собой потенциально достижимые значения показателей качества (в общем случае зависимых и конкурирующих между собой). Кроме того, проводя анализ можно выяснить, как необходимо изменить значения одних показателей качества системы для улучшения других показателей, а также выяснить, как при этом следует изменить структуру и параметры системы.

Если найденное подмножество Парето-оптимальных вариантов системы оказалось узким, то на дальнейших этапах проектирования можно использовать любой из них. Как правило, подмножество Парето-оптимальных вариантов содержит много вариантов систем. Поэтому возникает задача сужения найденного подмножества Парето-оптимальных проектных решений с привлечением дополнительной информации об отношениях предпочтения заказчика системы.

Методы сужения множества Парето до единственного проектного решения

Как правило, для последующих этапов проектирования, должен быть выбран единственный вариант системы, поэтому возникает необходимость сужения подмножества Парето-оптимальных проектных решений до единственного варианта системы с привлечением дополнительной информации об отношениях предпочтения заказчика.

Такая информация появляется в результате всестороннего анализа Парето-оптимальных вариантов системы. Полученные при этом дополнительная информация о предпочтениях заказчика используется для построения некоторой скалярной целевой функции, оптимизация которой приводит к выбору единственного варианта системы.

Одним из распространенных способов сужения подмножества Парето-оптимальных решений является построение скалярной функции ценности, оптимизация которой приводит к выбору одного варианта системы из подмножества Парето.

Литература

1. **Шмалько А.В.** Цифровые сети связи: основы планирования и построения. — М.: Эко-Трендз, 2001. —282 с.

2. **Вишневский В.М.** Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 512 с.

3. **В.К. Борзенко, Л.М. Кемпнер.** Многокритериальная оптимизация. Математические аспекты. — М.: Наука, 1989. — 128 с.

4. **Bezruk V.M., Rybalko D.V.** Automation of communication systems design // IEEE "East-West Design & Test" (Symposium EWDTS'07). — Yerevan, 2007.

5. **Ж. Тель.** Введение в распределенные алгоритмы. — М.: МЦНМО, 2009. — 113с.

Особенности функционирвоания сетей беспроводного широкополосного доступа WIMAX с использованием технологии MIMO

Использование технологии MIMO позволяет уменьшить число ошибок при радиообмене данными (BER) без снижения скорости передачи в условиях множественных переотражений сигналов. При этом многоэлементные антенные устройства обеспечивают расширение зоны покрытия радиосигналами, использование нескольких путей распространения сигнала, увеличение пропускной способности линий связи. В MIMO-системах используются различные подходы к разделению сигналов, среди которых пространственно-временное, пространственно-частотное, пространственно-поляризационное кодирование, а также сверхразрешение по направлению прихода сигнала в приемник. Именно обилие подходов к разделению сигналов обусловило столь долгую разработку стандартов на использование систем MIMO в средствах связи.

Бухарев И.А.
Аспирант ФГУП ЦНИИС,
bukharev@zniis.ru

Простейшая антенна MIMO — это система из двух несимметричных вибраторов (монополей), ориентированных, например, под углом ±45° относительно вертикальной оси.

Такой угол поляризации ставит оба канала в равные условия, поскольку при горизонтально-вертикальной ориентации излучателей одна из поляризационных составляющих неизбежно получила бы большее затухание при распространении вдоль земной поверхности. Сигналы, излучаемые независимо каждым монополем, поляризованы взаимно ортогонально с достаточно высокой взаимной развязкой по кросс-поляризационной составляющей (не менее 20 дБ). Аналогичная антенна используется и на приемной стороне. Такой подход позволяет одновременно передавать сигналы с одинаковыми несущими, модулированными различным образом. Принцип поляризационного разделения обеспечивает удвоение пропускной способности линии радиосвязи по сравнению со случаем одиночного монополя (в идеальных условиях прямой видимости при идентичной ориентации приемных и передающих антенн). Таким образом, по сути любую систему с двойной поляризацией можно считать системой MIMO.

Ортогональное кодирование

Другое направление реализации принципа MIMO — использование ортогонального кодирования (как правило, по фазе) сигналов, независимо излучаемых разными антенными элементами. Характерный пример — применение MIMO в устройствах стандарта IEEE 802.11.

Один из возможных подходов к реализации ортогонального кодирования заключается в расширении спектра передаваемого сигнала методом прямой последовательности (DSSS — Direct Sequence Spread Spectrum).

Дальнейшим развитием принципа DSSS в семействе стандартов IEEE 802.11 стало использование в протоколе 802.11b метода фазовой модуляции с помощью многофазной (полифазной) комплементарной кодовой последовательности (метод CCK — Complementary Code Keying). Первым изложил идею бинарных комплементарных кодов Марсель Голей (M.J.E. Golay), более полувека назад описав их принцип и методы генерации [2]. Суть комплементарного кодирования состоит в использовании двух последовательностей *a* и *b*, каждая из *k* элементов, с автокорреляционными функциями $A(i)=\sum_{j=1}^{k-i} a_j a_{j+1}$ и $B(i)=\sum_{j=1}^{k-1} b_j b_{j+1}$,

где $i = 0...k-1$ — возможный сдвиг, для которых $A(i) + B(i) = 0$ при любом $i \neq 0$ и $A(0) + B(0) = 2k$.

Физическая интерпретация сдвига в данных формулах — прием прямо распространяющегося сигнала и сигнала с фазовой задержкой на *i* элементов (чипов). Иными словами, если система связи, использующая комплементарное кодирование, работает в условиях многопутевого распространения сигналов, то в идеале межсимвольная интерференция, вызванная наложением сигналов с задержками распространения, должна отсутствовать, так как сумма их автокорреляционных функций равна нулю.

Поскольку в стандарте IEEE 802.11 принята дифференциальная квадратурная фазовая модуляция (DQPSK), всего может быть четыре значения фазы, сдвинутые друг относительно друга на 90°. Суть применения CCK состоит в том, что поток транслируемых данных разбивается на группы по 8 бит каждая. Группа заменяется сигналом C из 8 чипов $C = (c_1...c_8)$. Каждый чип — это сигнал, фаза которого определяется квартетом ϕ_1, ϕ_2, ϕ_3, ϕ_4 так что

$$C=(c1...c8)=[e^{j(\phi_1+\phi_2+\phi_3+\phi_4)}, e^{j(\phi_1+\phi_3+\phi_4)}, e^{j(\phi_1+\phi_2+\phi_4)}, -e^{j(\phi_1+\phi_4)},$$
$$e^{j(\phi_1+\phi_2+\phi_3)}, e^{j(\phi_1+\phi_3)}, -e^{j(\phi_1+\phi_2)}, e^{j\phi_1}$$

Применительно к квадратурной модуляции действительная и мнимая составляющие относятся к синфазному и квадратурному каналам соответственно.

В группе все 8 бит разбиваются на пары (дибиты), каждому дибиту ставится в соответствие определенный элемент фазового квартета $\phi_i\ (d_0d_1 \to \phi_i, d_2d_3 \to \phi_2, d_4d_5 \to \phi_3, d_6d_7 \to \phi_4)$. В зависимости от значений дибитов, ϕ_i оказывается равным 0°, 90°, 180° или 270°. Причем ϕ_1 определяется фазой предыдущего символа и четностью текущего символа C в потоке, поэтому модуляция и называется дифференциальной. Можно показать, что при любых значениях дибитов последовательности C оказываются комплементарными.

На приемной стороне, при условии синхронного приема, декодирующее устройство восстанавливает значения ϕ_1, ϕ_2, ϕ_3, ϕ_4, а по ним — и значения информационных бит. В простейшем случае обработка сводится к следующим вычислениям:

$$\phi_2 = \arg\{r_1r_2^* + r_3r_4^* + r_5r_6^* + r_7r_8^*\},$$
$$\phi_3 = \arg\{r_1r_3^* + r_2r_4^* + r_5r_7^* + r_6r_8^*\},$$
$$\phi_3 = \arg\{r_1r_5^* + r_2r_6^* + r_3r_7^* + r_4r_8^*\},$$
$$\phi_3 = \arg\{r_4e^{-j\phi_4} + r_6e^{-j\phi_3} + r_7e^{-j\phi_2} + r_8\},$$

где $r = [r_1,...,r_8]$ — принятая 8-битовая последовательность. Знак * означает комплексное сопряжение: если $r_i = a + jb$, то $r_i^* = a - jb$. Функция $\arg(r) = \mathrm{arctg}[Im(r)/Re(r)] = \mathrm{arctg}(b/a)$. Фактически в процессе такой обработки отношение сигнал-шум улучшается в среднем вдвое.

Поскольку CCK-последовательности комплементарны, при применении техники MIMO возможна их одновременная трансляция. Например, четные чипы (или символы) передаются по одному антенному каналу, нечетные — по другому, при этом они оказываются практически ортогональными.

Развитием принципа DSSS стала технология мультиплексирования с кодовым разделением каналов CDMA (Code Division Multiple Access). При этом на одной частоте излучаются сигналы с различными взаимно ортогональными базами (псевдослучайными кодовыми последовательностями). Причем для коррелятора, настроенного на определенную базовую последовательность, все другие сигналы выглядят как белый шум. Очевидно, что в этом случае при использовании технологии MIMO каждому антенному каналу ставится в соответствие определенный CDMA-код, и сигналы в них оказываются ортогональными.

Помимо кодового разделения каналов, все большее развитие получают системы с разделением на основе ортогональных частот (OFDM). Идея метода — вместо высокоскоростной модуляции одной несущей применять модуляцию набора поднесущих с гораздо меньшей скоростью. Весь диапазон разбивается на несколько ортогональных частот вида $s_n(t) = A \cdot \sin \cdot (2\pi[f_0 + n\Delta f])$, где $n = 0,...,N$; Δf — интервал между поднесущими.

Входной поток данных делится на группы (символы), которые используются для одновременной модуляции каждой поднесущей. К символам добавляют защитные интервалы (паузы) как эффективное средство борьбы с межсимвольной интерференцией. Чтобы метод был эффективен, число поднесущих должно быть достаточно большим — от десятков до тысяч. Так, в стандарте IEEE 802.11 a и g используют 52 поднесущие, в стандарте IEEE 802.16 — от 200 до 2048, в спецификации наземного цифрового телевизионного вещания DVB-T — 6817 поднесущих частот. Каждая из них модулируется посредством многоуровневой квадратурной модуляции. Выходной многочастотный сигнал синтезируют посредством обратного быстрого преобразования Фурье (ОБПФ).

Очевидно, что системы OFDM по природе своей приспособлены для применения технологии MIMO, причем различными способами. Техника MIMO-OFDM предусмотрена стандартами IEEE 802.16, проектом IEEE 802.11n (рис. 1), рядом других перспективных разработок в области беспроводных сетей передачи информации.

Следующим шагом развития техники CDMA и ее интеграции с OFDM стало частотно-кодовое разделение каналов MC-CDMA (multi carrier CDMA — многочастотный CDMA). Впервые эта технология была предложена в 1993 г. профессором Калифорнийского университета в Беркли Линнарецом (Jean-Paul Linnartz) на традиционном Международном симпозиуме по персональным мобильным и внутриофисным коммуникациям (PIMRC — Personal Indoor and Mobile Radio Communications) в Йокогаме [3]. Изначально предлагалось передавать каждый бит исходного сообщения на нескольких ортогональных поднесущих. На каждой из частот бит должен заменяться CDMA-кодом длины N (рис. 2). В результате на всех поднесущих сигнал оказывается защищенным от межсимвольной и межканальной интерференции дважды - частотным и кодовым разделением.

По прошествии времени, в первую очередь — в связи с поистине взрывным развитием OFDM-устройств (системы таких массовых стандартов, как IEEE 802.11 a/g, DVB, системы цифрового радиовещания и т.п.), стали появляться различные варианты реализации технологии MC-CDMA. Сегодня ее можно рассматривать как дальнейшее развитие кодированной OFDM-модуляции (COFDM), тем более что как синоним термина MC-CDMA зачастую используется понятие OFDM-CDMA.

В "классических" системах связи OFDM-CDMA используется один трансивер. В случае применения систем MIMO исходный поток данных разбивается на субпотоки по числу излучателей MIMO-передатчиков. В каждом субпотоке применяется свой CDMA-код, и далее кодированная последовательность битов используется для формирования OFDM-символов, параллельно транслируемых каждым передатчиком (рис. 3). Такой подход позволяет при одних и тех же номиналах частот разделять каналы передачи MIMO-системы за счет дополнительной ортогональной кодовой модуляции. На приемной стороне из OFDM-сигнала сначала восстанавливаются ортогональные несущие (с помощью БПФ), а затем на каждой из них сигнал декодируется путем корреляционной обработки.

Для MIMO-систем возможен и другой вариант реализации MC-CDMA, когда сначала с помощью ОБПФ синтезируется поток OFDM-символов, а затем он подвергается дополнительному CDMA-кодированию. Причем исходный поток OFDM-символов распределяется на субпотоки для каждого излучателя. Для каждого субпотока используется свой CDMA-код, кодеры располагаются перед входом ЦАП антенных излучателей. При этом сигналы разных антенн имеют одни и те же частоты, а разделение при приеме происходит с помощью CDMA-кода: в приемнике на первом этапе производится корреляционная обработка с восстановлением пакетов гармонических сигналов с непрерывной фазой (свой синусоидальный пакет для каждого излучателя MIMO), а затем восстанавливаются сигналы на каждой ортогональной поднесущей OFDM с помощью БПФ. Оба рассмотренных варианта MC-CDMA по качеству передачи информации, в принципе, равноценны.

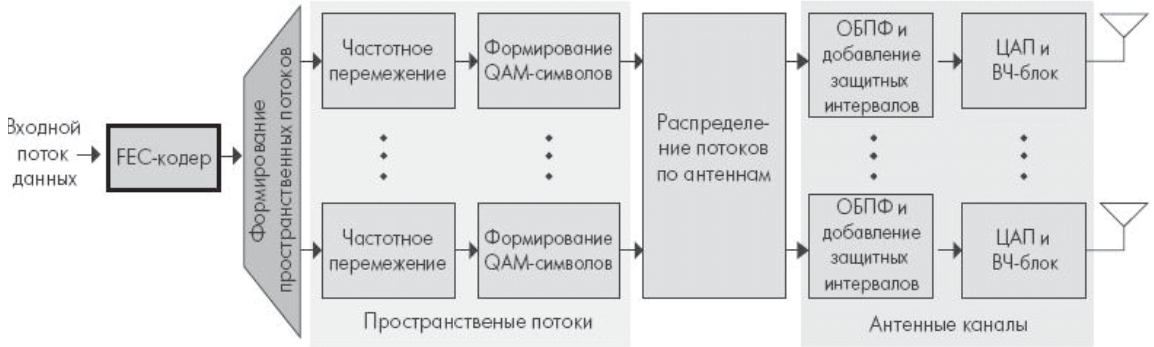


Рис. 1. Система передачи MIMO-OFDM

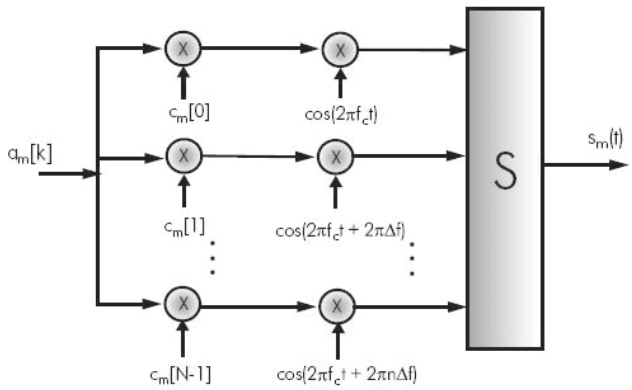


Рис. 2. Система MC-CDMA по Линнарецу

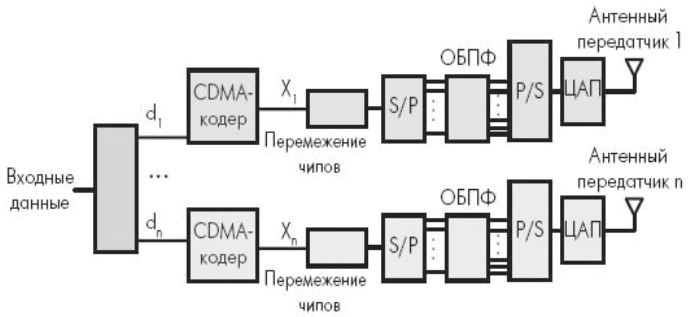


Рис. 3. Система OFDM-CDMA

Пространственно-временное кодирование

Блочные методы пространственно-временного кодирования

В 1998 г. Аламоути (Siavash Alamouti, сегодня — сотрудник компании Vivato) предложил новый подход к разделению MIMO-сигналов на приемной стороне [4]. Новая схема, названная по имени автора, относится к классу ортогонального пространственно-временного блокового кодирования (OSTBC, orthogonal space-time block codes).

Принцип кодирования по Аламоути состоит в том, что подлежащая передаче последовательность символов разбивается на пары (например, смежные четный и нечетный символы) x_i и x_{i+1}. Для передачи такого блока требуются два излучателя и два интервала передачи. В первом интервале передающая антенна 1 будет излучать сигнал символа x_i, тогда как антенна 2 — сигнал x_{i+1}. В следующем временном интервале антенна 1 передает сигнал x^*_{i+1}, а антенна 2 — сигнал x^*_i.

Физическая сущность таких манипуляций с излучаемыми сигналами становится ясной, если воспользоваться их математическим представлением:

$$x_i = a_i + jb_i,$$
$$x_{i+1} = a_{i+1} + jb_{i+1},$$
$$x_i^* = a_i - jb_i,$$
$$-x_{i+1}^* = -a_{i+1} + jb_{i+1}.$$

Несложно заметить, что если в первом временном интервале сигналы синфазны, во втором они обязательно будут в противофазе, и наоборот. Если же фазы сигналов в первом интервале ортогональны (разность 90°), то они таковыми останутся и во втором, с той лишь разницей, что опережение фазы излучения поочередно будет возникать то на первой, то на второй антеннах.

Когда передающая антенная система состоит из двух несимметричных вибраторов, результирующая диаграмма направленности (ДН) существенно зависит от соотношения фаз входных сигналов (рис. 6). В общем случае результирующая ДН отклоняется от нормали к линии, соединяющей вибраторы, в сторону элемента, на котором фаза сигнала запаздывает.

Таким образом, максимум энергии сигнальной смеси MIMO, кодированной по Аламоути, излучается каждый раз в различных направлениях, причем различных для каждой новой передаваемой пары символов в зависимости от разности фаз сигналов. Это повышает вероятность их прохождения с учетом наличия множественных переотражателей.

Для приема кодированного по Аламоути двухсимвольного сигнала достаточно одной приемной антенны и пары временных отсчетов сигнальной смеси. Таким образом, фактически можно обойтись системой MISO. При оцифровке сигнальной смеси в двух последовательных временных интервалах получим совокупность напряжений y_i и y_{i+1}:

$$y_i = h_1 x_i + h_2 x_{i+1} + n_i,$$
$$y_{i+1} = -h_1 x_{i+1}^* + h_2 x_i^* + n_{i+1},$$

где n_i, n_{i+1} — отсчеты напряжений внутреннего шума приемника, а h_1, h_2 — передаточные характеристики канала для сигналов, излученных первой и второй антеннами, соответственно. Два временных отсчета необходимы для того, чтобы число уравнений в системе равнялось числу неизвестных.

Для оценки характеристик канала передачи h_1 и h_2 при вхождении в связь транслируются заранее известные пилотные сигналы z_i и z_{i+1}. В приемнике решается приведенная выше система уравнений, где в качестве неизвестных выступают передаточные характеристики канала h_1 и h_2:

$$h_1 = \frac{y_i z_i^* - y_{i+1} z_{i+1}}{z_i^2 + z_{i+1}^2},$$

$$h_2 = \frac{y_{i+1} z_i + y_i z_{i+1}^*}{z_i^2 + z_{i+1}^2}.$$

После установления связи по известным значениям передаточных характеристик декодируются пары переданных символов [4]:

$$x_i = h_1^* y_i + h_2 y_{i+1}^*,$$
$$x_{i+1} = h_2^* y_i - h_1 y_{i+1}^*.$$

Данные соотношения являются оптимальными оценками максимального правдоподобия. Следует, однако, иметь в виду, что характер переотражений на трассе распространения сильно зависит от направления излучения, поэтому для систем MIMO известный в антенной теории принцип взаимности передающих и приемных антенн может не выполняться. Это вынуждает независимо рассчитывать характеристики каналов передачи в прямом и обратном направлениях связи. Однако на передающей стороне знать свойства трассы распространения сигналов не требуется.

16	9	6	3
5	4	15	10
11	14	1	8
2	7	12	13

→

4	1	2	3
1	4	3	2
3	2	1	4
2	3	4	1

Рис. 4. Преобразование элементов магического квадрата 4x4

MISO-схема Аламоути используется в стандарте IEEE 802.16-2004 (WiMAX). Согласно данным корпорации Intel, такое кодирование в условиях многократных переотражений позволяет в конечном счете получить выигрыш в отношении сигнал-шум около 5 дБ для модуляции BPSK и до 10 дБ — для 64-QAM.

Главное ограничение в применении рассмотренного метода кодирования — допущение о неизменности характеристик канала не только в двух последовательных временных интервалах, по которым рассчитываются передаточные характеристики, но и вплоть до момента окончания приема полезной информации. По этой причине пространственно-временное кодирование Аламоути для мобильных абонентов сопровождается падением эффективности передачи.

Вместо излучения пары сигналов в двух последовательных временных интервалах их можно передавать одновременно на двух ортогональных частотах. Такой метод кодирования называется пространственно-частотным и позволяет увеличить пропускную способность по сравнению с пространственно-временным методом. Принцип кодирования сигнала состоит в том, что пары символов x_i и x_{i+1} передаются соответственно антеннами 1 и 2 на первой частоте, а символы x^*_{i+1}, x^*_i — на второй частоте, при этом сигналы обеих частот излучаются одновременно.

Альтернативный методу Аламоути подход к пространственно-временному кодированию состоит в использовании мультиплексирования сигналов по принципу магического квадрата [6]. Этот механизм позволяет применять систему из 2M излучателей, работающих на приемную цифровую антенную решетку (ЦАР) из 2M элементов.

Магический квадрат порядка n представляет собой квадратную таблицу чисел размером *n* x *n*, заполненную целыми числами от 1 до n^2. Сумма чисел в каждой строке, в каждом столбце, а также в двух главных диагоналях должна быть одинакова.

Такая числовая схема положена и в основу рассматриваемого метода кодирования. Предложенный в публикации [6] код предполагает преобразование элементов магического квадрата 4x4: вместо исходных чисел X используются $X_{mod\,4}$ — их значения по модулю 4 ($X_{mod\,4} = 4$, если X кратно 4; $X_{mod\,4} = X$, если $X < 4$; при $X > 4$ $X_{mod\,4}$ равен остатку от деления X на 4). Для выбранной в работе [6] версии магического квадрата имеем (рис. 4).

В полученном числовом квадрате суммы элементов по всем строкам, столбцам и главным диагоналям равны 10. Числа соответствуют номерам четырех последовательных символов S_1, S_2, S_3 и S_4, подлежащих передаче. Строки соответствуют антенным излучателям, а столбцы — временным интервалам передачи.

На приемном терминале с одной антенной получают четыре последовательных отсчета сигнальной смеси y_i, y_{i+1}, y_{i+2}, y_{i+3} и решают систему:

$$y_i = h_{i,1}S_4 + h_{i,2}S_1 + h_{i,3}S_3 + h_{i,4}S_2 + n_i,$$
$$y_{i+1} = h_{i+1,1}S_1 + h_{i+1,2}S_4 + h_{i+1,3}S_2 + h_{i+1,4}S_3 + n_{i+1},$$
$$y_{i+2} = h_{i+2,1}S_2 + h_{i+2,2}S_3 + h_{i+2,3}S_1 + h_{i+2,4}S_4 + n_{i+2},$$
$$y_{i+3} = h_{i+3,1}S_3 + h_{i+3,2}S_2 + h_{i+3,3}S_2 + h_{i+3,4}S_1 + n_{i+3},$$

где n_i, n_{i+1}, n_{i+2}, n_{i+3} — отсчеты напряжений внутреннего шума приемника, а $h_{i,m}$ — известная передаточная характеристика канала для сигналов, излученных *m*-й антенной в *i*-й момент времени. Данная система уравнений относительно неизвестных S_1-S_4 решается методом наименьших квадратов.

В четырехантенном приемнике оценка амплитуд может быть произведена по одному отсчету напряжений по выходам этих антенн. Причем вместо временных интервалов можно использовать четыре ортогональные несущие, что позволяет повысить пропускную способность линии связи. Для декодирования переданного таким способом четырехчастотного пакета данных потребуется решить систему 16 уравнений.

Метод магического квадрата [6] эффективнее схемы Аламоути при быстрых замираниях сигналов на трассе распространения [7]. Если отношение сигнал-шум составляет 18 дБ, метод магического квадрата обеспечивает вероятность ошибки, на два порядка меньшую, чем метод Аламоути. В то же время при медленных замираниях картина меняется на противоположную, хотя и с меньшим разрывом по значению вероятности ошибки.

Решетчатое пространственно-временное кодирование

Пропускная способность системы в целом и ее BER также в немалой степени определяются выбранными алгоритмами декодирования. Все основные алгоритмы декодирования строятся на следующих возможных принципах: принцип максимального правдоподобия, принцип минимальной среднеквадратичной ошибки, принцип обнуления. Принцип решетчатого кодирования выражается в присвоении каждому переходу от одного символа к другому уникальной последовательности бит, формируемой на основе заранее известного полинома. Кодер STTC представляет собой совокупность M-PSK либо M-QAM модулятора и решетчатого кодера с заданным полиномом (в частности, кодера Витерби).

Пространственная селекция

Разделить сигналы MIMO-систем можно и на основе пространственной их селекции, по углам прихода на приемную антенну. Если сигналы на приемник поступают с разных угловых направлений, различающихся более чем на ширину луча ДН приемной антенны, то их можно разделить обычной пространственной селекцией. Например, в случае двух сигналов, с помощью приемной smart-антенны (ЦАР) можно сформировать два независимых луча ДН и сориентировать их в направлениях максимальной приходящей мощности. Для увеличения углового разноса трасс прохождения сигналов можно искусственно ориентировать ДН передающей антенны не в на-

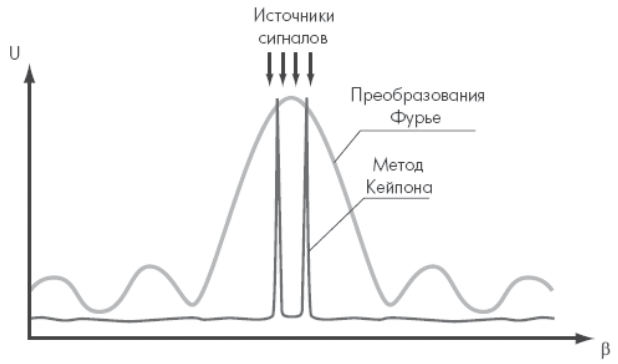


Рис. 5. Эффект сверхрэлеевского разрешения двух сигналов по методу Кейпона (в сравнении с классической обработкой посредством преобразования Фурье)

правлении приемника, а в сторону мощного переотражателя (горы, высотного здания и т.п.).

Если же различия в направлениях приема сигналов не превышают ширины главного луча результирующей приемной ДН, а остальные их параметры совпадают, сигналы передатчика MIMO могут быть разделены на основе методов углового "сверхразрешения". Поясним сущность одной из таких процедур на примере двухвибраторной антенны.

Если угловые координаты излучателей (bm) относительно нормали к приемной антенне известны, задача разделения сигналов, излученных парой вибраторов, сводится к решению системы уравнений, составленных по одному отсчету АЦП:

$$y_1 = h_1(\beta_1)x_1 + h_1(\beta_2)x_2,$$
$$y_2 = h_2(\beta_1)x_1 + h_2(\beta_2)x_2,$$

где y_1, y_2 — выходные напряжения приемных антенн; x_1, x_2 — неизвестные комплексные амплитуды излученных сигналов; $h_1(\beta_m)$, $h_2(\beta_m)$ — известные ДН приемных антенных элементов в направлениях источников излучения.

Неизвестные угловые координаты источников излучения определяются на этапе вхождения в связь при цифровом формировании ДН, для этого можно применять нелинейные математические операции — например, процедуру Кейпона (рис. 5). В результате ДН подобных приемных антенн (являющиеся виртуальными функциями) будут крайне узконаправленными и остроконечными, что позволяет повысить пространственную избирательность антенной системы. Характерно, что передавать сигналы в данном случае может антенна с широкой ДН. Это особенно важно, поскольку при нелинейной обработке принцип взаимности не выполняется и воспроизвести столь же остроконечные ДН для передающей антенны невозможно.

Следует отметить, что в режиме MIMO цифровое диаграммообразование со сверхразрешением быстро теряет свою эффективность с увеличением расстояния передачи. Например, в базовой станции (БС) сверхразрешение сигналов двух излучателей терминалов можно реализовать на расстояниях в сотни метров, а при большем числе независимых элементов в передающей антенне — и того меньше.

Иное дело — прием сигналов БС самим терминалом. Как правило, в MIMO-системах на БС может быть создан сравнительно большой разнос антенных элементов — до 10 длин волн. Это обеспечивает лучшую декорреляцию сигналов в режиме передачи на терминал. Потенциально такое решение позволяет применять разные методы для разделения MIMO-каналов в отношении входящего и исходящего трафиков.

Литература

1. **Golay, M.J.E.** Complementary series. — IRE Transactions, 1961, IT-7. — С. 82-87.

2. **N. Yee, J. P. Linnartz and G. Fettweis.** Multi-Carrier CDMA in Indoor Wireless Radio Networks. — Proceedings PIMRC'93, Yokohama, Japan, 1993. — С.109-113.

3. **Alamouti S.M.** Space-time block coding: A simple transmitter diversity technique for wireless communications. — IEEE Journal on Selected Areas in Communications, Oct. 1998, vol. 16, С.1451-1458.

4. **Biljana Badic, Markus Rupp, Hans Weinrichter.** Adaptive Channel-Matched Extended Alamouti Space-Time Code Exploiting Partial Feedback. — ETRI Journal, Volume 26, Number 5, 2004 (http://etrij.etri.re.kr).

5. **Yong Xiong, Zhiyong Bu.** Magic Squares transmission scheme for MIMO OFDM systems under fast fading channel. — In: 1st International MAGNET Workshop, Shanghai, 11-12 November 2004 (www.ist-magnet.org).

Современные методы оптимизации бизнес-процессов разработки и внедрения новых инфокоммуникационных услуг

На сегодняшний день на рынке инфокоммуникаций наблюдается появление все большего количества новых услуг, имеющих непродолжительный жизненный цикл и приносящих небольшие доходы. В то же время их доля на рынке инфокоммуникаций продолжает расти. При существующих процессах предоставления услуг такие услуги не могут стать для операторов прибыльными. Рассматриваются новые методы и технические решения, появившиеся на рынке инфокоммуникаций, которые позволяют существенно оптимизировать бизнес-процессы разработки и внедрения новых услуг, привлечь небольшие компании или даже конечных потребителей к процессу создания услуг для небольших целевых групп и, тем самым занять свое место в данном сегменте рынка.

Полканов Е.И.,
технический директор ООО "Фирма "СВЕТЕЦ",
epolkanov@svetets.ru

Ефимушкин И.В.,
ведущий специалист ООО "СВЕТЕЦ Технолоджи",
i.efimushkin@svetets.ru

Введение

Последние несколько лет рынок инфокоммуникационных услуг продолжает развиваться значительными темпами. Растет не только суммарный объем доходов от услуг, но и спектр предлагаемых услуг. Сегодня абонентам уже недостаточно традиционной телефонной связи, их требования и запросы неуклонно возрастают, становясь все более содержательными и избирательными.

При этом стоит отметить, что вместе с количеством изменились и сами услуги. В начале 21 века основную часть доходов составляло всего несколько услуг с длительным жизненным циклом, например, телефония и доступ в Интернет. В настоящее время в Интернете активно развиваются услуги сервис-провайдеров, отличительной особенностью которых является короткий жизненный цикл (ЖЦ) и, как следствие, небольшие прибыли.

В то же время, хотя прибыли от таких услуг не велики, десятки тысяч услуг предлагаются как небольшими сервис-провайдерами, так и частными лицами (рис.1), и за счет их большого числа, суммарные объемы доходов от их предоставления составляют существенную часть современного рынка инфокоммуникаций.

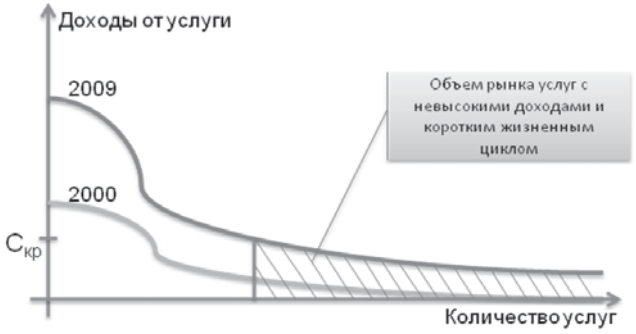


Рис.1. Объем рынка услуг с невысокими доходами

Операторские ограничения. Услуги с коротким ЖЦ

На сегодняшний день операторы связи не могут охватить весь спектр услуг. Значительные затраты на разработку и внедрение каждой услуги ограничивают операторов связи в выборе услуг, доступных для внедрения. Внедрение услуг с коротким жизненным циклом и невысокими доходами операторам связи экономически не выгодны (например, на рис.1 услуги с показателем дохода ниже Скр).

Прежде всего, это связано с

- длительными сроками внедрения и тестирования услуги;
- отсутствием эффективных возможностей интеграции с инфраструктурой оператора;
- невозможностью позиционировать услуги на узкие целевые группы.

Существенно затрудняет процесс внедрения огромная сеть связи и множество оборудования в инфраструктуре оператора, с которым необходимо провести интеграцию [1]. Кроме того, отсутствие специального персонала (особенно у операторов сетей фиксированной связи) приводит к тому, что серьезные решения внедряемые поставщиками оборудования используются с низкой эффективностью. Например, решения, рассчитанные на предоставление множества разноплановых услуг, задействуются только для предоставления нескольких услуг, предварительно настроенных специалистами поставщиков оборудования.

Устоявшееся выражение - "одна платформа — одна услуга" не может конкурировать на современном рынке инфокоммуникаций с гибкостью сервис-провайдеров, которым для предоставления услуг достаточно выхода в Интернет и внешнего адреса IP.

Рассмотрим множество сервис-провайдеров, которые делают основной упор на новизну услуг — новые возможности, новый контент, новые розыгрыши. Низкая стоимость разработки и внедрения, позволяет таким компаниям вывести услугу на рынок,получить доходы от первой волны спроса пользователей, желающих попробовать новую возможность и, уже через несколько недель, вывести на рынок новую услуг с новым контентом.

Для успешной деятельности подобных компаний необходимо решить два основных вопроса:

- монетизация услуг;
- выход на абонентскую базу.

Монетизация

Вопросы монетизации для сервис-провайдеров уже много лет решают операторы сетей подвижной связи. При этом доля операторов составляет процент от платежа абонента, который может варьироваться от 15 до 50% в зависимости от установленных договоренностей и доходов услуги.

Развитие электронных платежных систем, комиссия которых при оплате услуг сервис-провайдеров составляет 2-7%, может существенно сказаться на доходах операторов СПС. В настоящее время для сервис-провайдеров обычной практикой является предоставление множества способов оплаты собственных услуг:

- отправка сообщения SMS с кодовым словом на специальный короткий номер;
- оплата услуги средствами платежных терминалов;
- оплата услуг с помощью электронных денег.

Существует множество предложений уже готовых решений, которые предоставляют сервис-провайдерам в одном пакете различные способы оплаты услуг без необходимости заключения договорных отношений с каждым участником бизнес-процесса.

В то же время для рядового потребителя услуг наглядно проявляется разница цен при различных вариантах оплаты. С одной стороны оплата услуг с мобильного телефона удобна, поскольку у каждого абонента уже есть счет, с которого можно совершить платеж. С другой стороны разница цены на оплату той же самой услуги с помощью электронных платежных систем заставляет абонентов задуматься о создании собственного электронного кошелька.

Постепенно с развитием средств доступа в Интернет и платежных терминалов все большее количество абонентов выбирают для оплаты электронные платежные системы, цены которых, несмотря на наличие комиссий при пополнении счета и совершении платежей, все равно остаются ниже, чем стоимость оплаты услуг средствами мобильного телефона. Кроме того, использование электронных денег снимает с абонента услуг территориальные ограничения, позволяя получить доступ к возможностям зарубежного сервис-провайдера.

Абонентская база

Еще одним важным аспектом для успешного предоставления услуг в качестве сервис-провайдера является абонентская база. Успех современных услуг зачастую зависит от предварительной рекламной кампании и начального продвижения: чем больше пользователей узнали о новых возможностях, тем выше интерес со стороны других пользователей [2].

В части данного вопроса значительную роль начинают играть социальные сети, при этом не только для рекламы нового продукта. Согласно последним тенденциям социальные сети становятся площадками для услуг небольших сервис-провайдеров. То есть, сервис-провайдеры разрабатывают и внедряют свои услуги на базе абонентских и технологических ресурсов социальных сетей. Новые возможности сервис-провайдеров встраиваются в уже существующие возможности Интернет проектов и предлагаются наравне с остальными.

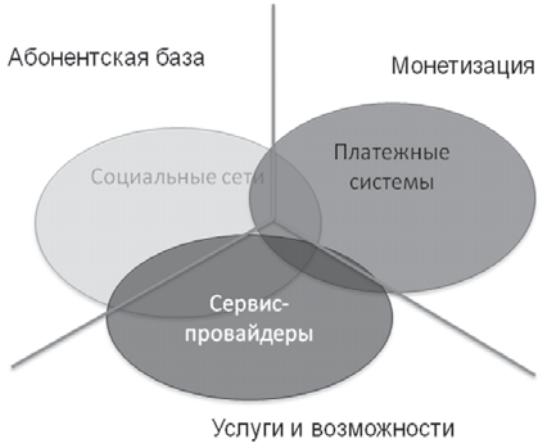


Рис.2. Необходимые ресурсы для организации успешной деятельности в качестве сервис-провайдера

Значительно способствует такому развитию наличие в социальных сетях внутренней валюты, за которую можно покупать доступ к специальным платным возможностям.

Ярким примером успешной монетизации является Российский проект "В контакте" и его зарубежный аналог "Facebook". Каждый пользователь социальной сети может пополнить свой внутренний счет множеством удобных для него способов, включая электронные платежные системы и сообщения SMS с добавочной стоимостью. Все сервис-повайдеры, интегрирующие свои возможности в данные социальные сети, как правило, предоставляют набор бесплатных возможностей для привлечения пользователей и некоторые уникальные возможности, доступные только после их оплаты внутренней валютой системы.

В качестве сервис-провайдеров с социальными сетями начинаются интегрироваться и сами платежные системы. Например, сеть платежных терминалов Qiwi, предоставляющая своим абонентам личный кабинет с электронным кошельком, недавно добавила свои возможности в социальную сеть "В контакте". Теперь пользователям социальной сети доступны средства пополнения внутреннего счета, оплаты услуг сервис-провайдеров социальной сети со счета личного кошелька Qiwi и другие возможности личного кабинета платежной системы.

Таким образом, социальные сети стали универсальными площадками для развития Интернет-сервисов за счет объединения всех необходимых ресурсов (рис.2).

Выделение отдельной структуры для разработки и внедрения новых услуг

С одной стороны у операторов связи с их сетевыми и людскими ресурсами и телефонией есть все возможности стать аналогичной площадкой для сервис-провайдеров с целью выхода на рынок услуг с небольших уровнем доходов и коротким жизненным циклом. С другой стороны традиционным операторам связи, в силу недостаточного опыта и квалифицированных кадров, затруднительно в короткий срок выйти на рынок услуг Интернета и самостоятельно подключать небольших сервис-провайдеров с их творческим и интеллектуальным потенциалом, который может быть направлен на любые, даже самые небольшие целевые группы пользователей.

В этой ситуации, как показывает практика зарубежных компаний, существует несколько подходов, активно использующихся крупнейшими зарубежными компаниями:

- глобальная реструктуризация компании с выделением одной или нескольких отдельных структур, занимающихся непосредственно взаимодействием с сервис-провайдерами (Deutsche Telekom, France Telecom);
- партнерское соглашение с крупной компанией в области инфокоммуникаций (Frane Telecom);
- покупка и развитие небольшой компании в области инфокоммуникаций (British Telekom, Verizon).

Одной из основных задач новых структур оператора связи является оптимизация бизнес-процессов разработки и внедрения новых услуг целью выхода на рынок услуг связи и Интернета с коротким жизненным циклом.

Выделение новой компании, как правило, направлено на решение следующих вопросов:

- внедрение единого решения для предоставления услуг (например, с использованием технологии OSA и платформы SDP, Service Delivery Platform) [3];

• интеграция решения с инфраструктурой оператора связи (ЦКПС, ACP, OSS/BSS и пр.);

• разработка и настройка единых внешних интерфейсов для сервис-провайдеров;

• проработка вопросов информационной безопасности;

• отработка и оптимизация бизнес-процессов взаимодействия с сервис-провайдерами (присоединение/отключение, вывод новой услуги и настройка ее тарификации, продвижение новой услуги и пр.).

При этом внедрение современного решения SDP позволяет не только автоматически решить большинство перечисленных вопросов, но и предоставить следующие возможности:

• API (Application Program Interface) - специфицированный и опубликованный набор функций, на базе которых программируются услуги;

• SCIM (Service Capabilities Interaction Management) - функции координирования взаимодействия нескольких услуг на разных серверах приложений, позволяющие организовать пакетирование услуг;

• OSA (Open Service Architecture) - набор открытых протоколов для взаимодействия с серверами приложений (J2EE, SOAP, Parlay, HTTP, VoiceXML);

• PRM (Partnership Relation Management) - функции управления отношениями с сервис-провайдерами, использующиеся для оперативного контроля действий;

• IDM (ID Management) - функции ведения единых идентификаторов, отличающихся от сетевых, что позволяет предотвратить работу сервис-провайдеров с абонентами напрямую и избежать злоупотреблений полученной информации;

• CF (Content Filtering) - функции пропуска трафика услуг для их анализа, позволяют выявлять предпочтения и интересов абонентов и предоставлять "родительские" услуги;

• SCE (Service Creation Environment) - среда разработки, отладки, тестирования услуг, позволяющая привязать сервис-провайдеров к оператору связи и снизить их затраты на инструментальные средства;

• SLEE (Service Logic Execution Environment) - среда выполнения услуг, позволяющая снизить затраты сервис-провайдеров на сервера приложений;

• CG (Charging Gateway) - функции начисления оплаты в зависимости от содержания услуги, позволяющие вести программы лояльности для абонентов и предоставлять единую точку взаимодействия с ACP для сервис-провайдеров.

Перечисленные выше функции и возможности платформ SDP прежде всего направлены на централизованное предоставление новых уникальных возможностей и снижение затрат на разработку и внедрение новых услуг, как операторами связи, так и сервис-провайдерами.

Бизнес-инкубаторы

Новые подразделения занимаются и самостоятельной разработкой и выводом новых услуг на рынок. В связи с этим наблюдается тенденция увеличения инвестиций ведущих телекоммуникационных компаний в научно-исследовательские и опытно-конструкторские работы. Кроме того, крупнейшие операторы связи (например, France Telekom, British Telecom) организуют собственные научно-исследовательские подразделения и лаборатории, среди которых можно выделить:

• исследовательские и аналитические центры;

• эксплуатационные центры;

• технопарки и модельные сети.

Целью эксплуатационных центров является поиск и отбор интересных идей новых услуг и возможностей. При этом в последнее время представители операторских компаний начинают привлекать для этого не только специалистов отрасли, но небольшие компании (например, Deutsche Telekom и British Telecom), студентов и даже обычных пользователей услуг (например, компании France Telecom, British Telecom). Для этого операторы связи создают специальные площадки — бизнес-икубаторы, на базе которых все желающие, в том числе и обычные пользователи услуг, могут попробовать реализовать свои идеи, после чего наиболее удачные услуги операторы связи разрабатывают, тестируют на модельной сети и внедряют на коммерческую сеть [4].

Технопарки и модельные сети операторы используются для отработки технологических цепочек по разработке и выводу новых услуг на телекоммуникационный рынок. Такой подход позволяет не только существенно снизить риски и непредвиденные расходы, но и подготовить квалифицированных специалистов.

Заключение

Новые методы и технические решения, появившиеся на рынке инфокоммуникаций позволяют существенно оптимизировать бизнес-процессы разработки и внедрения новых услуг, привлечь небольшие компании или даже конечных потребителей к процессу создания услуг для небольших целевых групп и, тем самым занять свое место в данном сегменте рынка. Проанализированы бизнес модели внедрения новых инфокоммуникационных услуг операторами связи и сервис провайдерами.

Литература

1. **Ефимушкин И.В., Журавлев С.В.** Концепция UGC как основа создания саморазвивающихся услуг // В сб.: Технологии информационного сообщества. Тез. докл. Московской отраслевой научно-технической конференции. — М.: Инсвязьиздат, 2007.

2. **Ефимушкин И.В.** Метод приведения услуги связи к модели саморазвивающейся услуги // В сб.: Труды Московского техн. ун-та связи и информатики / М.: ИД "Медиа Паблишер", 2008. — Т.1.

3. **Чурбанов Д.В.** Открытая сервисная архитектура OSA // Сети и системы связи, 2002. — №3.

4. **Ефимушкин И.В.** Изменение роли пользователей в процессе создания и оказания услуг // Электросвязь, 2009. — № 9.

Исследование вероятностно-временных характеристик функционирования соты WiMAX с несколькими режимами модуляции и эластичным трафиком данных

Рассматривается аналитическая модель соты WiMAX, применяющей одну из ключевых технологий, входящих в состав стандарта IEEE 802.16, адаптивную модуляцию и кодирование. Она позволяет динамически изменять режим модуляции для того, чтобы снизить потери и увеличить скорость передачи данных. В модели рассматривается два типа трафика: голос и сеансы данных, последние — эластичного типа.

Ефимушкина Т.В.,
бакалавр кафедры Систем телекоммуникаций РУДН,
tvefimushkina@gmail.com

Молчанов Д.А.,
старший научный исследователь кафедры Техники связи факультета Вычислений и электротехники Технологического университета Тампере, Финляндия
moltchan@cs.tut.fi

Кучерявый Е.А.,
профессор кафедры Техники связи факультета Вычислений и электротехники Технологического университета Тампере, Финляндия
yk@cs.tut.fi

Введение

Сети WiMAX (Worldwide Interoperability for Microwave Access) являются реализацией технологии широкополосного беспроводного доступа Broadband Wireless Access (BWA) стандарта IEEE 802.16, и реализуют самые последние достижения науки и техники в области радиосвязи, телекоммуникаций и компьютерных сетей.

В системах WiMAX, применяется широкополосный Orthogonal Frequency Division Multiplexing (OFDM) сигнал, образованный из множества разнесенных по частотному спектру узкополосных сигналов. Сигнал OFDM в сетях WiMAX фиксированного доступа стандарта IEEE 802.16-2004 имеет до 256 поднесущих, в сетях мобильного WiMAX стандарта IEEE 802.16e — до 2048 поднесущих.

В WiMAX используется технология адаптивной модуляции и кодирования AMC (Adaptive Modulation and Coding), которая заключается в автоматическом выборе модуляции в зависимости от текущего соотношения сигнал/шум в канале передачи. Это позволяет использовать спектрально эффективные режимы модуляции (64-QAM) на меньших расстояниях и достигать тем самым большей скорости передачи данных. При увеличении расстояния между базовой станцией и мобильной станцией (или при уменьшении излучаемой мощности) возможен выбор более помехоустойчивой, но менее скоростной модуляции (QPSK).

В табл.1 представлены различные режимы модуляции и кодирования (РМ), поддерживаемые системой WiMAX [1,2]. Далее рассматривается общий случай, когда в зависимости от условий в канале может быть выбран один из N РМ. Всего WiMAX поддерживает 64 различных РМ.

Описание модели

Рассмотрим функционирование соты сети подвижной связи, построенной на базе технологии WiMAX. Будем считать, что в соте поддерживается N РМ, $N < \infty$.

В соте рассматривается два типа трафика: голосовые вызовы и эластичные сеансы данных.

Будем считать, что весь частотный интервал соты поделен на M базовых частотных каналов (БК), при этом под БК будем подразумевать канал с минимальной используемой частотной полосой.

Для установления соединения с *i* РМ (*i*-РМ), каждый голосовой вызов требует БК, в то время как для сеансов данных необходимо t_i БК, $i = \overline{1,N}$.

Состояние системы обозначим, как $(\vec{v},\vec{d}) = (v_1,...,v_N,d_1,...,d_N)$, где v_i представляет число обслуживающихся голосовых вызовов в соте, а d_i — число обслуживающихся сеансов данных с *i*-РМ, $i = \overline{1,N}$. Далее везде, $i = \overline{1,N}$. Примем следующие обозначения $\vec{v}^{\mathrm{T}} = (v_1,...,v_N)$, $\vec{d}^{\mathrm{T}} = (d_1,...,d_N)$, $\vec{s} = (s_1,...,s_N)$, $\vec{t} = (t_1,...,t_N)$ и $M^{\mathrm{v}} = \vec{v}^{\mathrm{T}}\vec{s}$, $M^{\mathrm{d}} = \vec{d}^{\mathrm{T}}\vec{t}$, $M^{\mathrm{vd}} = M^{\mathrm{v}} + M^{\mathrm{d}}$.

Для наиболее эффективного использования ширины полосы пропускания (ШПП) вся свободная от голосовых вызовов полоса предоставляется сеансам данных. При этом сеансы данных, находящиеся на обслуживании занимают $M - M^{\mathrm{v}}$ БК пропорционально доле занимаемого минимального числа t_i БК: $\left\lfloor (M - M^{\mathrm{v}})\frac{t_i d_i}{M^{\mathrm{d}}} \right\rfloor$. Таким образом, каждому сеансу с *i*-РМ дополнительно к минимально выделяемым БК предоставляется необязательно целое число $n_i(\vec{v},\vec{d}) = \left\lfloor (M - M^{\mathrm{v}})\frac{t_i d_i}{M^{\mathrm{d}}} \right\rfloor \Big/ d_i$ БК, где $M - M^{\mathrm{v}}$ — это число БК, занимаемых сеансами данных.

Таблица 1

Схемы модуляции и кодирования, поддерживаемые WiMAX

РМ	Нисходящий канал	Восходящий канал
Модуляция	BPSK (опцион.), QPSK, 16QAM, 64QAM	BPSK, QPSK, 16QAM; 64QAM (опцион.)
Обязательное кодирование	Сверточное кодирование со скоростью 1/2, 2/3,3/4, 5/6	Сверточное кодирование со скоростью 1/2, 2/3,3/4, 5/6
Опциональное кодирование	Сверточное турбокодирование со скоростью 1/2, 2/3, 3/4, 5/6; повторное кодирование со скоростью 1/2, 1/3, 1/6, LDPC-код, код Рида-Соломона	Сверточное турбокодирование со скоростью 1/2, 2/3, 3/4, 5/6; повторное кодирование со скоростью 1/2, 1/3, 1/6, LDPC-код

Схема принятия поступающих вызовов

В данной модели рассматривается три типа вызовов: новый вызов, хэндовер-вызов и вызов, меняющий модуляцию. В качестве правила установления соединения, дающего приоритет как хэндовер-вызовам, так и вызовам, меняющим модуляцию, используется выделение резервных каналов. Это означает, что базовой стан-

ции назначен некоторый порог g, причем $0 < g < M$.

Число свободной ШПП для нового вызова вычисляется в соответствии с $f^{\mathrm{n}}(\vec{v},\vec{d}) = M - g - M^{\mathrm{vd}}$. Для хэндовер-вызова и вызова, меняющего модуляцию число свободной ШПП — $f^{\mathrm{n}}(\vec{v},\vec{d}) = M - g - M^{\mathrm{vd}}$, $f^{\mathrm{h}}(\vec{v},\vec{d}) = M - M^{\mathrm{vd}}$.

При поступлении в соте нового голосового вызова с i-РМ, он будет принят, если находящиеся на обслуживании сеансы данных будут иметь необходимый минимум БК для продолжения обслуживания после передачи s_i БК новому голосовому вызову: $f^{\mathrm{n}}(\vec{v},\vec{d}) \geq s_i$.

При поступлении в соте нового хэндовер-вызова с i-РМ, он будет принят, если $f^{\mathrm{h}}(\vec{v},\vec{d}) \geq s_i$.

При поступлении вызова, меняющего модуляцию с i-РМ на $(i + 1)$-РМ, он будет принят, если в системе есть необходимое для смены число БК $f^{\mathrm{h}}(\vec{v},\vec{d}) \geq \Delta s$, где $\Delta s = s_{i+1} - s_i$, причем для данного случая $i = \overline{1, N-1}$, принимая во внимание возможное освобождение БК сеансами данных, находящихся на обслуживании.

При поступлении нового, хэндовер-запроса или запроса, меняющего модуляцию с i-РМ на $(i + 1)$-РМ, на открытие сеанса данных, он будет принят, если $f^{\mathrm{n}}(\vec{v},\vec{d}) \geq t_i$, $f^{\mathrm{h}}(\vec{v},\vec{d}) \geq t_i$ и $f^{\mathrm{h}}(\vec{v},\vec{d}) \geq \Delta t$, где $\Delta t = t_{i+1} - t_i$, $i = \overline{1, N-1}$, соответственно.

В противном случае в нашей модели указанный голосовой вызов или запрос на сеанс блокируется, теряется и не оказывает влияния на потоки вызовов в соте. И наоборот, при окончании обслуживания голосового вызова или завершении сеанса, освободившиеся ресурсы будут добавлены действующим в этот момент сеансам передачи данных в соответствии с алгоритмом, описываемым выше.

Нагрузочные параметры

Примем, что поступающий поток голосовых вызовов и запросов на открытие сеансов данных — пуассоновский; а длительность обслуживания и процесс изменения модуляции имеют экспоненциальные распределения. Введем их интенсивности:

λ_i^{vn} — интенсивность поступления нового голосового вызова с i-РМ; λ_i^{vh} — интенсивность поступления голосового хэндовер-вызова с i-РМ; λ_i^{dn} — интенсивность поступления нового запроса на открытие сеанса данных с i-РМ; λ_i^{vh} — интенсивность поступления хэндовер-запроса на открытие сеанса данных с i-РМ; $\mu_i^{\mathrm{v}} = s_i\mu_i$ — интенсивность обслуживания голосового вызова с i-РМ; $\mu_i^{\mathrm{d}}(\vec{v},\vec{d}) = n_i(\vec{v},\vec{d})\mu_i$ — интенсивность обслуживания сеанса данных с i-РМ в состоянии $(\vec{v},\vec{d})$; γ_i^{v} — интенсивность изменения РМ голосового вызова с i-РМ на $|i - 2| + 1$-РМ; γ_i^{d} — интенсивность изменения РМ сеанса данных с i-РМ на $|i - 2| + 1$-РМ.

Функционирование данной системы описывается марковским процессом с пространством состояний [3]:

$$S = \left\{ (\vec{v},\vec{d}) \mid v_i = \overline{0, \left\lfloor \frac{M}{s_i} \right\rfloor}, d_i = \overline{0, \left\lfloor \frac{M}{t_i} \right\rfloor}, 0 \leq M^{\mathrm{vd}} \leq M, i = \overline{1, N} \right\}.$$

Вследствие сделанных предположений о марковском процессе, существует стационарное распределение вероятностей $p(\vec{v},\vec{d})$, которое можно получить с помощью системы уравнений равновесия и нормализующей функции $\sum_S p(\vec{v},\vec{d}) = 1$.

Вероятностно-временные характеристики

С помощью стационарного распределения вероятностей можно получить формулы для вероятностей блокировок поступающего голосового вызова и запроса на открытие сеанса данных.

Пусть π — вероятность, того что поступающий вызов или сеанс данных будет заблокирован. Тогда $\pi^B = \sum_B p(\vec{v},\vec{d})$,
где B — подпространство состояний марковского процесса, в которых вызов/сеанс данных блокируется.

Тогда, для каждого типа вызова/сеанса данных подпространство B можно представить следующим образом:

- для нового голосового вызова

$$B_i^{\mathrm{vn}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{n}}(\vec{v}+\vec{e}_i,\vec{d}) < 0 \right\}$$

- для голосового хэндовер-вызова

$$B_i^{\mathrm{vh}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{h}}(\vec{v}+\vec{e}_i,\vec{d}) < 0 \right\}$$

- для голосового вызова, меняющего модуляцию с i-РМ на с $|i - 2| + 1$-РМ;

$$B_i^{\mathrm{vm}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{h}}(\vec{v},\vec{d}) < \Delta s \right\}$$

- для нового запроса на открытие сеанса данных;

$$B_i^{\mathrm{dn}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{n}}(\vec{v},\vec{d}+\vec{e}_i) < 0 \right\}$$

- для хэндовер-запроса на открытие сеанса данных;

$$B_i^{\mathrm{dh}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{h}}(\vec{v},\vec{d}+\vec{e}_i) < 0 \right\}$$

- для сеанса данных, меняющего модуляцию с i-РМ на с $|i - 2| + 1$-РМ.

$$B_i^{\mathrm{dm}} = \left\{ (\vec{v},\vec{d}) \in S \mid f^{\mathrm{h}}(\vec{v},\vec{d}) < \Delta t \right\}$$

Обозначим среднее число голосовых вызовов и сеансов данных, как N. Тогда

$$N = \sum_S (\vec{v},\vec{d})^{\mathrm{T}} \vec{1} p(\vec{v},\vec{d}).$$

Среднее время пребывания голосовых вызовов в системе с i-РМ, как W_i^{v}, где

$$W_i^{\mathrm{v}} = \sum_S v_i p(\vec{v},\vec{d}) \Big/ \left[\lambda_i^{\mathrm{vn}} \left(1 - \pi^{B_i^{\mathrm{vn}}}\right) + \right.$$
$$\left. + \lambda_i^{\mathrm{vh}} \left(1 - \pi^{B_i^{\mathrm{vh}}}\right) + \gamma_{|i-2|+1}^{\mathrm{v}} \sum_S u^{\mathrm{h}}\left(\vec{v}+\vec{e}_i-\vec{e}_{|i-2|+1},\vec{d}\right) v_{|i-2|+1} p(\vec{v},\vec{d}) \right],$$

в то время как среднее время пребывание сеансов данных в системе с i-РМ, как W_i^{d}

$$W_i^{\mathrm{d}} = \sum_S d_i p(\vec{v},\vec{d}) \Big/ \left[\lambda_i^{\mathrm{dn}} \left(1 - \pi^{B_i^{\mathrm{dn}}}\right) + \lambda_i^{\mathrm{dh}} \left(1 - \pi^{B_i^{\mathrm{dh}}}\right) + \right.$$
$$\left. + \gamma_{|i-2|+1}^{\mathrm{d}} \sum_S u^{\mathrm{h}}\left(\vec{v},\vec{d}+\vec{e}_i-\vec{e}_{|i-2|+1}\right) d_{|i-2|+1} p(\vec{v},\vec{d}) \right].$$

Среднее число занятых каналов может быть представлено в виде

$$U^{\mathrm{e}} = \frac{1}{M} \sum_{S-S''} M^{\mathrm{v}} p(\vec{v},\vec{d}) + \sum_{S''} p(\vec{v},\vec{d})$$

где $S'' = \left\{ (\vec{v},\vec{d}) \in S \mid M^{\mathrm{d}} > 0 \right\}$

Заключение

Исследована модель с применением технологии адаптивной модуляции и кодирования в рамках стандарта IEEE 802.16. Модель учитывает два типа трафика: голоса и сеансов данных, причем сеансы данных могут использовать все имеющиеся канальные ресурсы, но при поступлении более приоритетного голосового вызова часть ресурсов освобождается для голосового вызова.

В работе получена система уравнений равновесия для данной модели и вероятностно-временные характеристики ее функционирования, включая вероятности блокировок голосовых вызовов и запросов на сеансы данных для всех видов модуляции, средние значения для числа вызовов в системе и времен пребывания голосовых вызовов и сеансов данных.

Полученные характеристики позволяют провести необходимые численные исследования функционирования модели.

Литература

1. **Вишневский В.М., Портной С.Л., Шахнович И.В.** Энциклопедия WiMAX. Путь к 4G. — М.: Техносфера, 2009. — 472 с.

2. **Kwon E., Lee J., Jung K., Ryu S.** A Performance Model for Admission Control in IEEE 802.16 // WWIC, 2005. — P.159-168.

3. **Башарин Г.П.** Лекции по математической теории телетрафика : учеб. пособие. изд. 2-е испр. и доп. — М.: РУДН, 2007. — 268 с.