УДК 621.396.931

ОЦЕНКА СРЕДНЕГО УДЕЛЬНОГО МЕСЯЧНОГО ТРАФИКА ПРИ НЕДОСТАТОЧНОЙ СТАТИСТИКЕ

Ю.А. Савостицкий, гл. научный сотрудник Института автоматизированных систем (ИАС), к.т.н.

Введение. Формируемые оператором условия присоединения абонентов к своей номерной емкости базируются на документе Минсвязи России РД45.120–2000 и содержат ограничение на допустимую удельную нагрузку абонентской линии. Нагрузка на соединительную линию должна определяться не по максимуму допустимой месячной наработки абонента M в объемах трафика исходящих вызовов (в минутах в месяц), а по его среднему значению m. Для 80-ти регионов на сайте ОАО «Связьинвест» регулярно публикуются статистические данные по количеству абонентов по трафикам (квартальному и годовому). По некоторым регионам (в частности, по Москве и Татарстану) эти данные практически недоступны.

На практике телефонная активность разных абонентов существенно различается. Так, в московской сети ЗАО «Комкор-ТВ» 20% абонентов создают 40% внутреннего трафика, в МГТС 15–20% абонентов – 50% трафика.

В статье предлагается инженерный метод оценки среднего трафика (число минут занятия канала) по известным значениям максимально допустимого объема трафика и проценту трафика всех абонентов, приходящегося на долю наиболее активных абонентов. Предлагаемый метод основан на аппроксимации месячного трафика абонентов плотностью усеченного нормального распределения.

Аппроксимация месячного трафика. Профиль месячного трафика абонентов можно представить кривой плотности усеченного в $x = 0$ нормального распределения, где по оси абсцисс на отрезке $0 < x < \max x$ приведен условный показатель численности абонентов, а по оси ординат — условный показатель $f^*(x)$ объема месячного трафика абонента (рис. 1). При заданном общем числе N абонентов и допустимом месячном трафике M условные показатели приводятся к численности n абонентов и удельному месячному трафику t в натуральном выражении:

$$n = N x / \max x;\ t = M f^* / \max f^*, \qquad (1)$$

где параметры x и f^* (включая $\max x$ и $\max f^*$) определяются профилем $f^*(x)$ месячного трафика пользователя.

В качестве минимального набора исходных данных для выбора профиля $f^*(x)$ могут использоваться:

- $\min t / M$ – отношение минимального месячного трафика (ниже которого пользователь не учитывается в числе N) к максимально допустимому;
- процент наиболее активных абонентов, создающих 50% трафика.

В качестве $\min t$ представляется разумным принять время, составляющее менее половины средней продолжительности одного вызова. Если принять $M = 1000$ мин. (т. е. допустимая среднесуточная удельная нагрузка 0,023 Эрл, а нагрузка в ЧНН 0,097 Эрл), то при трехминутной средней продолжительности вызова в число абонентов не включаются пассивные пользователи, чей месячный трафик составляет менее $\min t / M = 0{,}005$, т. е. менее 0,5% от максимально допустимого. Процент наиболее активных абонентов (50% трафика) принимается равным 20%.

Профиль месячного трафика пользователя аппроксимируется кривой плотности усеченного нормального распределения $f^*(x)$ – областью положительных значений x. При отбрасывании области определения $x < 0$ степень усечения исходного нормального распределения $f(x)$ составляет 0,5, а интегральная функция распределения $\Phi^*(x)$ – усеченного в точке $x = 0$ распределения $\Phi(x)$ – есть $\Phi^*(x) = 2\Phi(x) - 1$ [1]. Максимальное значение $f^*(x)$ в точке $x = 0$:

$$\max f^*(x) = f^*(0) = \sqrt{2/\pi} = 0{,}79788;$$

$$f^*(x) = 0{,}79788 \exp(-x^2/2). \qquad (2)$$

Математическое ожидание E и дисперсия D случайной величины X с плотностью распределения $f^*(x)$ составляют:

$$E(X) = \sqrt{2/\pi}\int_0^\infty x\exp(-x^2/2)dx =$$

$$= 2\sqrt{2/\pi}\int_0^\infty y\exp(-y^2)dy = \sqrt{2/\pi} = 0{,}79788;$$

$$D(X) = \sqrt{2/\pi}\int_0^\infty (x - E')^2\exp(-x^2/2)dx =$$

$$= 2/\sqrt{\pi}\Gamma(1{,}5) - E^2 = 0{,}34083. \qquad (3)$$

По «правилу трех сигм» при нормальном распределении вероятность выхода случайной величины за пределы диапазона $E + 3D$ составляет 0,0014. На этом основании 100% абонентов соответствует аргументу функции распределения $x = 3{,}2$ $\Phi^*(3{,}2) = 0{,}9986$ и $f^*(3{,}2) = 0{,}0048$, т. е. в число N не включаются абоненты с месячным трафиком менее 6 мин. Таким образом, 20% абонентов N соответствуют значению $x = 0{,}64$. Условный показатель общего трафика всех включенных в рассмотрение абонентов

$$\Phi^*(x) = \int_0^{3{,}2} f^*(x)dx = 1. \qquad (4)$$

Погрешность значения $\Phi^*(x) = 1$ из-за неучета интервала $3{,}2 < x < \infty$ составляет 0,14%, а 20% абонентов N соответствуют значению $\Phi^*(0{,}62) = 0{,}48$, т. е. 20% абонентов создают около

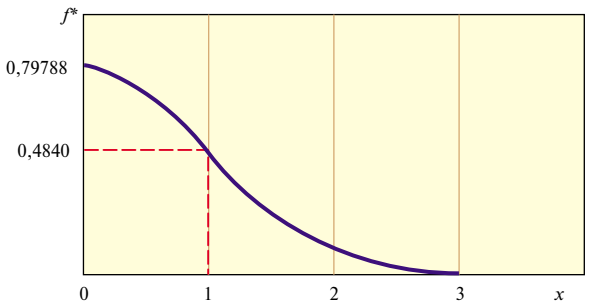

Рис. 1

50% трафика. В этом смысле аппроксимация профиля месячного трафика пользователя кривой $f^*(x)$ приемлема (рис.1).

Обоснование принятого профиля. При выборе вероятностных распределений на основе неполной информации обычно определяющим становится принцип максимума энтропии приемлемого распределения. Это позволяет избежать субъективного завышения «степени упорядоченности» распределения, не обоснованного доступными исходными данными. В нашем случае они ограничены следующей информацией:

- численность абонентов неотрицательна;
- 20% абонентов создают 50% трафика;
- весь трафик – 100%.

В интуитивно выбранном распределении, приведенном выше, эти данные привели к следующим обобщенным характеристикам распределения: $x > 0$, $E = \sqrt{2/\pi} = 0{,}79788$, $D = 0{,}34083$. Эти характеристики при поиске неизвестного распределения могут быть формализованы в виде условий:

$$\int_0^\infty f^*(x)dx = 1; \qquad (5)$$

$$\int_0^\infty f(x)dx = E = 0{,}79788; \qquad (6)$$

$$\int_0^\infty (x-E)^2 f(x)dx = D = 0{,}34083. \qquad (7)$$

Максимизируемая энтропия представляется в виде:

$$Э(X) = -\int_0^\infty f(x)\log f(x)dx = -(1/\ln 2)\int_0^\infty f(x)\ln f(x)dx. \qquad (8)$$

Нахождение функции $f(x)$, реализующей максимум определенного интеграла (8) и подчиняющейся дополнительным условиям (уравнениям связи) (5)–(7), является типичной изопериметрической задачей вариационного исчисления, к решению которой при достаточно гладких подынтегральных функциях в уравнениях связи применим метод множителей Лагранжа, т. е. искомая функция $f(x)$ может быть получена как решение системы уравнений Эйлера [2] . В нашем случае, когда выражения (5)–(8) не содержат $f(x)$, уравнение Эйлера имеет вид:

$$dF / df(x) = 0, \qquad (9)$$

где
$$F = dЭ/df + \lambda_1 \psi_1 + \lambda_2 \psi_2 + \lambda_3 \psi_3; \qquad (10)$$

$$dЭ/df = -f(x)\log f(x) = -(1/\ln 2) f(x)\ln f(x), \qquad (11)$$

где ψ_1, ψ_2, ψ_3 – подынтегральные функции левых частей уравнений связи (5)–(7); λ_1, λ_2, λ_3 – постоянные множители Лагранжа, определяемые из решения уравнения (9).

Решение уравнения Эйлера

$$\ln f(x) = [\,\lambda_1 (x - E)^2 + \lambda_2 x + \lambda_3\,]\ln 2 - 1,\ 0 < x < \infty. \qquad (12)$$

Для удовлетворения ограничений (5)—(7) подбираем константы:

$$\lambda_1 = -0{,}7214,\ \lambda_2 = -1{,}1512,\ \lambda_3 = 0{,}5747 \qquad (13)$$

и получаем

$$f(x) = \sqrt{2/\pi}\exp(-x^2/2),\ 0 < x < \infty. \qquad (14)$$

Таким образом, при $E = 0{,}79788$ и $D = 0{,}34083$, что соответствует $\Phi^*(0{,}62) \approx 0{,}5$ (т. е. 20% абонентов создают половину объема трафика). Усеченное в $x = 0$ нормальное распределение является наименее упорядоченным в смысле максимума энтропии и не содержит субъективизма, не обоснованного исходными данными. Так что выбор профиля (2) можно считать верным.

Оценка средней удельной наработки. Если рассмотреть обратную к $f^*(x)$ функцию $x(f^*)$, то произведение малого промежутка Δf^* в окрестности любого f^* на среднее по этому промежутку значение x, т. е. $x\,\Delta f^*(x)$, представляет собой долю суммарного месячного трафика всех рассматриваемых абонентов, приходящуюся на малую окрестность $x(f^* + \Delta f^*/2) < x < x(f^* - \Delta f^*/2)$. Суммирование всех $x\,\Delta f^*(x)$ по f^*, т. е. интеграл

$$\int_0^{0{,}79788} x df^* = 1 \qquad (15)$$

(та же величина, что и площадь Φ^* под кривой $f^*(x)$ на рис. 1) есть значение интегральной функции распределения случайной величины T месячного трафика в точке $f^* = \max f$. При неограниченном уменьшении Δf^* значение x характеризует плотность, с которой распределяются значения случайной величины T в точке $f^*(x)$. Таким образом, $x(f^*)$ представляет собой функцию плотности распределения T (рис. 2).

Среднее значение месячного трафика одного абонента:

$$E(f^*) = \int_0^{0{,}79788} f^*(x)df^* \qquad (16)$$

удается оценить лишь приближенно, аппроксимируя функцию $x(f^*)$ участками кривых разного порядка, упрощающими интегрирование:

$$x = 3{,}49 - 58 f^*, \quad f^* < 0{,}007;$$

$$x = 3{,}43 - 50 f^*, \quad 0{,}007 < f^* < 0{,}01;$$

$$x = 3{,}18 - 22{,}5 f^*, \quad 0{,}01 < f^* < 0{,}025;$$

$$x = 2{,}91 - 11{,}2 f^*, \quad 0{,}025 < f^* < 0{,}05;$$

$$x = 2{,}666 - 6{,}3 f^*, \quad 0{,}05 < f^* < 0{,}1; \qquad (17)$$

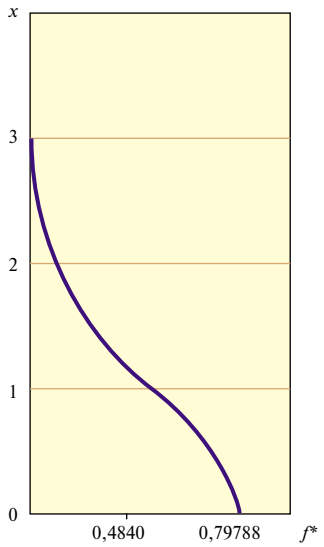

Рис. 2

$$x = 3{,}24 - 3f^{*\,0{,}4}, \quad 0{,}1<f^*<0{,}24;$$

$$x = 2{,}06 - 2{,}19f^*, \quad 0{,}24<f^*<0{,}57;$$

$$x = 2{,}069 - 2{,}19f^*, \quad 0{,}57<f^*<0{,}675; \qquad (17)$$

$$x = k\,(0{,}79788 - f^*)^{0{,}5}, \quad 0{,}675<f^*<0{,}79788,$$

где $k = 1{,}64$ при $f^* < 0{,}718$; $k = 1{,}61$ при $0{,}718 < f^* < 0{,}775$; $k = 1{,}59$ при $0{,}775 < f^*$.

Ошибка Δx аппроксимации (17) составляет 0,1–1,1% всюду кроме малых областей $x < 0{,}02$ ($0{,}79775 < f^* < 0{,}79788$) и $x > 3{,}345$ ($f^* < 0{,}003$), в которых неточность аппроксимации влияет на точность оценки интеграла (16) меньше, чем ошибки Δx в области $0{,}02 < x < 3{,}345$.

Подстановка (17) в (16) после вычислений приводит к $E(f^*) = 0{,}25$, т. е. среднее значение $f^*(x)$ составляет 0,3 M. При $M = 1000$ мин. месячный трафик среднего пользователя равен 300 мин.

Данный подход может быть использован для других исходных данных (доли наиболее активных абонентов и приходящегося на эту долю процента трафика всех абонентов), нежели рассмотренные выше применительно к МГТС (например, 25% абонентов создают 45% трафика). При этом должны быть скорректированы характеристики кривой аппроксимирующего нормального распределения (параметры E и D) и ограничения (6), (7). После этого изменится результат.

При рассмотрении различных типов нагрузки и варьируемого соотношения между ее отдельными составляющими (доли передачи речи и данных) может быть использован подход, изложенный в [3]. Данная статья ориентирована исключительно на ТфОП МГТС.

Заключение. Методика использована в работах ОАО ИАС в 2006 г. при расчетах требуемых канальных емкостей для подключения телефонных абонентов к МГТС.

ЛИТЕРАТУРА

1. **Корн Г., Корн Е.** Справочник по математике для научных работников и инженеров. – М.: Наука, 1974. – 831 с.
2. **Иоффе А.Д., Тихомиров В.М.** Теория экстремальных задач. – М.: Наука, 1974. – 479 с.
3. **Савостицкий Ю.А.** Метод определения требуемой полосы магистрали для пропуска мультимедийного трафика // Электросвязь. – 2004. – №3.

Получено после доработки 24.04.07

УДК 621.396

ЭВОЛЮЦИЯ ТРЕБОВАНИЙ К СЕТЯМ ДОСТУПА

М. Звягинцев, аспирант МТУСИ

Введение. Современный подход к развитию системы электросвязи базируется на концепции сети следующего поколения (ССП), известной также как NGN — Next Generation Network. Операторы связи предполагают переходить к ССП за счет модернизации сетей телефонной связи [1, 2], в которой “узким местом” является сеть доступа (СД) [3], созданная для передачи информации в полосе пропускания канала тональной частоты. При переходе к ССП самые сложные проблемы возникают при развитии СД.

Большинство технических решений по модернизации СД основаны на максимальном использовании эксплуатируемых многопарных абонентских кабелей. Этот подход оправдан для получения максимального экономического эффекта в течение ближайших лет. Он может рассматриваться как задача нахождения локального максимума в точке, близкой к $t = 0_+$, некоторой функции $f(t)$, определяющей доходы оператора связи. Анализ ряда прогностических оценок [4] показывает, что функция $f(t)$ имеет несколько экстремумов. Используя новые средства в СД, целесообразно найти экстремумы функции $f(t)$ в границах интервала $(0, T_{max})$, выделив среди них глобальный максимум.

Выбор значения T_{max} осуществляется с учетом характера поставленной задачи. При проектировании СД величиной T_{max} можно считать “время жизни” T_l. Обычно временем жизни технического устройства считается срок службы T_0. Если $T_l \approx T_0$, то проект введения в эксплуатацию устройства очень удачен. На практике встречаются ситуации, когда $T_{lf} << T_0$. Это говорит об ошибочности проекта. Для выбора эффективного пути модернизации СД целесообразно определить соответствующие требования к моменту $T_{max} = T_0$.

Перечень требований к перспективной СД. В рекомендациях МСЭ подчеркивается, что в СД не осуществляется обработка информации пользователя и сообщений, которые передаются в системе сигнализации. Это положение позволяет ограничить перечень показателей, предъявляемых к перспективной СД.

Причина появления технологии ADSL [5] — необходимость повышения скорости обмена информацией. Тенденция роста скорости обмена информацией набирает силу [4]. Этот показатель можно выразить как функцию времени $B(t)$. Несомненно, что функция $B(t)$ существенно влияет на основные технические характеристики перспективной СД и на объем инвестиции для ее создания. Есть все основания считать, что функция $B(t)$ монотонна на отрезке времени $(0, T_{max})$.

Сети следующего поколения относятся к конвергентным сетям, предоставляющим широкий спектр услуг. Поддержка множества услуг одной сетью возможна только при обеспечении высокой надежности ССП. Мерами надежности ССП можно считать коэффициент готовности $A(t)$ и время устранения неисправностей $R(t)$. На отрезке времени $(0, T_{max})$ функция $A(t)$ будет возрастать. Характер этого процесса можно описать дискретной функцией, стремящейся к единице. Функция $R(t)$ на отрезке $(0, T_{max})$ будет убывать. Этот процесс также можно представить дискретной функцией, стремящейся к минимально возможному уровню.

Появление новых видов услуг (например, распределение программ телевидения высокой четкости) требует повышения показателей качества передачи информации. Примером таких показателей служит коэффициент ошибок по битам, которые можно выразить с помощью функции $E(t)$. Соответствующий