

УДК 519.224

## ПОСТРОЕНИЕ СТАТИСТИЧЕСКОЙ МОДЕЛИ РЕЧЕВОГО ТРАФИКА ПРИ МОНОЛОГЕ

Сахаров А.В.

## Введение

IP-телефония – новая технология, получившая широкое распространение во всем мире, и в настоящее время продолжает динамично развиваться. Свидетельством перспективности данной технологии может служить тот факт, что многие операторы сотовой и стационарной связи заявили о предоставлении этой услуги, а всемирно-известная компания AT&T, уже осуществляющая передачу речи по IP-сетям, объявила о долгосрочном плане перевода всего своего трафика дальней связи на платформу IP.

Создание любой более или менее разветвленной пакетной сети начинается с определения ее топологии и характера нагрузки, создаваемой каждым рабочим местом, а также требований, предъявляемых приложениями к пропускной способности, задержкам и вариации задержки пакетов при передаче. Важным этапом при построении сети является имитационное моделирование нагрузки на разных ее участках, что позволяет оптимизировать размеры буферов и, как следствие, снизить вероятность заторов в сети. Трафик в пакетных сетях хорошо изучен и достаточно точно описывается самоподобными случайными процессами с параметром Херста 0,65-0,8 [1]. Большинство клиент-серверных приложений, служба электронной почты и многие другие программы не критичны к задержкам и тем более к вариации задержки. Не имеет решающего значения, придут ли данные немного ранее или немного позднее. Поврежденные данные всегда можно запросить повторно. Здесь главное – сам факт доставки данных. Совсем иначе обстоит дело в случае с IP-телефонией. Речевой трафик не похож ни на один другой вид трафика. Приложения IP-телефонии предъявляют к сети ряд специфических требований (поскольку изначально сеть проектировалась исключительно для передачи данных), описываемых ниже. Качество передачи речи считается хорошим, согласно рекомендации G.114 союза ITU-T, если суммарная задержка при передаче сигнала в одну сторону не превышает 150 мс. Удовлетворительным считается качество при задержке до 450 мс. Пакеты с кодированной речью, задержавшиеся на время

*Предлагается новый подход к построению статистической модели трафика IP-телефонии при монологе с учетом уровня кодируемого сигнала. Введение данного параметра в модель позволяет получать на выходе трафик с разной битовой скоростью, что наиболее полно отражает реальную ситуацию. Проведено моделирование битовой скорости. Сравнение модельных и опытных данных показало достаточно высокую степень корректности полученной модели.*

более данного порогового значения, отбрасываются и считаются потерянными. Понятно, что никакой речи о повторной передаче данных здесь не может быть, поскольку это вызовет еще большие задержки. Еще одной проблемой при передаче речи поверх протокола IP является вариация задержки данных или, другими словами, джиттер. Блоки с речью генерируются кодеком через равные промежутки времени. Задержка при передаче является случайной величиной и, таким образом, пакеты приходят к месту назначения с интервалами отличными от исходных. Джиттер является причиной «дрожания» речи. С ним призваны бороться джиттер-буферы, восстанавливающие изначальные интервалы и порядок следования пакетов. Необходимая ширина канала для передачи речи довольно небольшая и, в случае подключения к сети Ethernet, колеблется в пределах 13 – 21 Кбит/с. В ряду с описанными отличиями трафик IP-телефонии не является самоподобным [3]. Все это говорит о необходимости учета при проектировании компьютерных сетей исключительных требований, предъявляемых приложениями IP-телефонии к передаче речи. Специфические особенности речевого трафика, такие как низкая битовая скорость и последствия работы детектора речевой активности (VAD – voice activity detector), отбрасывающего блоки, не содержащие речь, приводят к выводу о необходимости создания особой модели, учитывающей все характерные перечисленные свойства трафика IP-телефонии.

В ходе телефонного разговора оба абонента попеременно создают нагрузку на сеть. Таким образом, телефонный разговор можно представить как диалог, состоящий из следующих друг за другом монологов разной продолжительности (рис. 1).

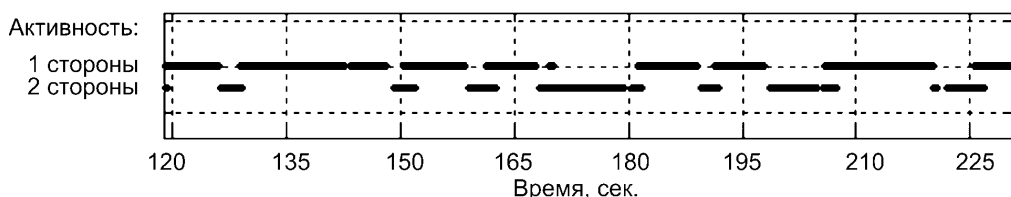


Рис. 1. Диаграмма активности участников телефонного разговора

## Способ построения статистической модели речевого трафика

В данной статье предлагается способ построения статистической модели трафика при монологе как «on/off» последовательности. Такая последовательность состоит из следующих друг за другом интервалов с активной речью и пауз. Паузы являются результатом работы детектора речевой активности, входящего в состав вокодера. Блоки, не содержащие речь, не попадают в выходной поток (рис. 2в.) и, таким образом, уменьшается необходимая ширина канала для передачи без существенного ухудшения качества передаваемой речи.



Рис. 2. Этапы обработки речи на стороне источника (а) оцифрованный поток, (б) после разбиения на блоки, (в) поток, прошедший VAD т.е. без пауз, (г) сжатый поток, (д) поток RTP пакетов (по одному блоку на пакет)

Из выше сказанного следует, что построение статистической модели трафика сводится к определению продолжительностей интервалов с паузами и речью. Таким образом, необходимо подобрать функции плотностей распределения вероятностей и их параметры, которые наиболее точно описывают вероятность появления интервала с речью или паузами той или иной продолжительности.

Экспериментальными данными послужили 106 файлов, собранные в течение монолога. Речь кодировалась кодеком G.723.1 с битовой скоростью 6.4 Кбит/с., хорошо себя зарекомендовавшим и принятым производителями средств пакетной телефонии. Каждый файл содержит пакеты с кодированной речью. Продолжительность сбора информации каждого файла составила от 3 до 5 минут. Перехват пакетов производился на том же компьютере, на котором они и генерировались, что позволило снизить «дрожание» в следовании пакетов до порядка 1 мс [4].

Продолжительности интервалов с речью и паузами определялись на основе маркеров и временных меток, являющихся частью заголовка RTP (Real Time Transport Protocol – транспортный протокол реального масштаба времени). Он описан в RFC 1889. Маркером помечается первый из пакетов, ответственных за передачу участка с речью. Ключевым параметром в определении продолжительностей явилось значение временной метки, поскольку маркер иногда выставляется и среди подряд идущих блоков, что может быть объяснено наличием речи и паузы в блоке предшествующем маркеру. Подобное случалось редко и в расчет не бралось. Значение временной метки для последовательно идущих блоков отличается на количество отсчетов закодированных в одном блоке. Для кодека G.723.1 это значение равно 240. Величина временной метки увеличивается и в том случае, когда блок отбрасывается, как не содержащий речь [2]. Таким образом, разница в значениях меток более чем на 240 (т.е. 480, 720 и т.д.)

говорит о наличии паузы. Время первых и последних пакетов с речью фиксировалось, что позволило восстановить последовательность продолжительностей участков с речью и пауз.

В первую очередь необходимо определиться с набором распределений, среди которых будет выбрано одно, наиболее точно соответствующее экспериментальным данным. Было выбрано 6 распределений: гамма, Вейбулла, логнормальное, Пирсона 5 и 6 типа, бета распределения (рис. 4).

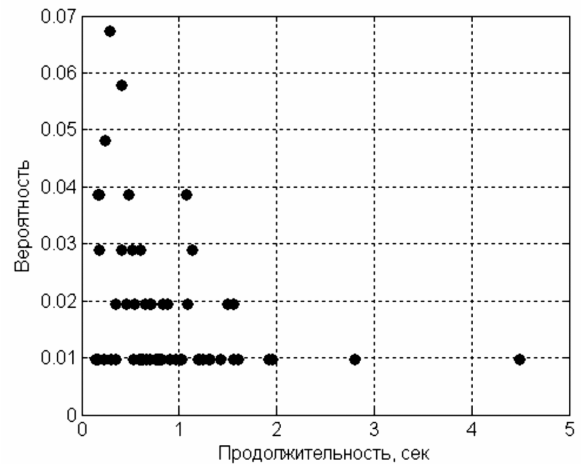


Рис. 3. Гистограмма плотности распределения вероятности продолжительности пауз

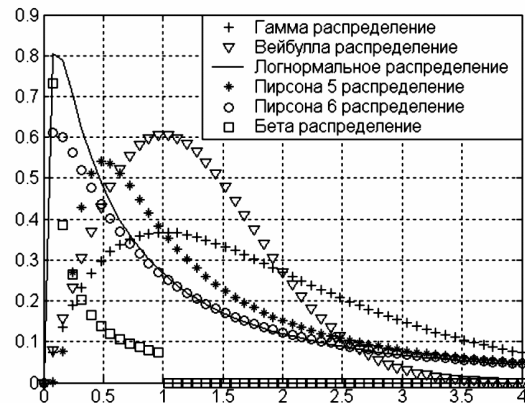


Рис. 4. Функции плотностей распределения вероятностей

Выбор производился на основе соответствия графиков функций плотностей распределения вероятностей и гистограмм плотностей распределения вероятностей экспериментальных данных (рис. 3).

Выбор наиболее подходящего распределения производился на основе критерия согласия Колмогорова-Смирнова. В начале, с помощью метода максимального правдоподобия, для каждого распределения на основе опытных данных подбирались его параметры. В этом методе предполагается независимость экспериментальных наблюдений  $x_1, x_2, \dots, x_n$ . В доказательство истинности данного предположения был подсчитан коэффициент корреляции, значение которого по абсолютной величине в среднем не превысило 0,23. Поскольку использование этого коэффициента не всегда оправдано, была проведена еще одна проверка. Неформальной оценкой независи-

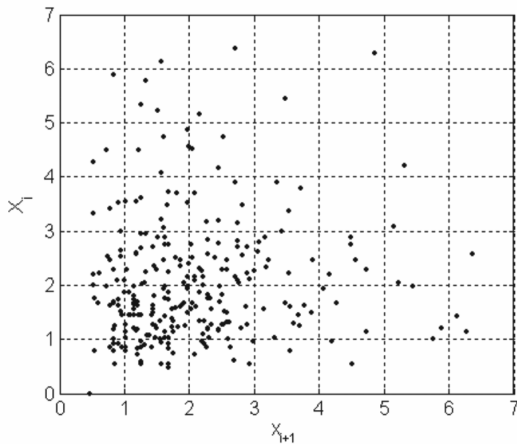


Рис. 5. Диаграмма разброса опытных данных

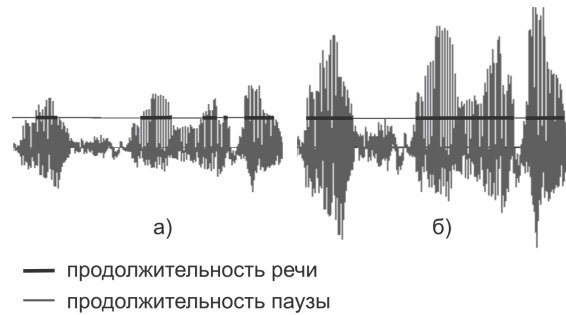


Рис. 6. Влияние амплитуды сигнала на продолжительность интервалов с речью и паузами а) сигнал с амплитудой A1 б) тот же сигнал с амплитудой A2, (A2 > A1).

Таблица 1

Результаты тестирования по критерию Колмогорова-Смирнова

Распределения:	Гамма	Вейбулла	Логнормальное	Пирсона 5 тип	Пирсона 6 тип	Бета
Речь	76	67	77	71	76	58
Пауза	101	101	106	96	106	85
Общий балл	177	168	183	167	182	143

мости данных, упорядоченных по времени их сбора, является диаграмма разброса (рис. 5). В случае, если величины  $x_1, x_2, \dots, x_n$  являются зависимыми, то точки  $(x_i, x_{i+1})$  стремятся расположиться вдоль линии с положительным или отрицательным уклоном в первом квадранте. В настоящем случае точки разбросаны хаотично, что подтверждает независимость данных.

Далее по критерию согласия Колмогорова-Смирнова с уровнем значимости  $\alpha = 0,05$  проводилась проверка на соответствие распределений опытным данным. В случае положительного результата распределение «получало» один балл, иначе – ноль баллов. Результаты тестирования по всем 106 файлам представлены в таблице 1.

В итоге получается, что логнормальное распределение наиболее точно описывает как распределение продолжительностей интервалов с речью, так и с паузами.

Во время проведения анализа опытных данных было отмечено, что среднее число интервалов с речью и с паузами  $\bar{N}_{on/off}$  составило 63 интервала на файл, при этом каждый файл содержит различное количество интервалов с речью и паузами  $N_{on/off}$ , варьирующееся от 2 до 160, что нельзя объяснить разницей только в продолжительности сбора информации в файл. Причем с уменьшением  $N_{on/off}$  средняя продолжительность интервалов с речью  $\bar{T}_{on}$  и паузами  $\bar{T}_{off}$  увеличивается и уменьшается соответственно. Это говорит о том, что чувствительность микрофона (несмотря на активированные функции «настройка чувствительности во время вызова» и «автоматическое обнаружение тишины») или громкость говорящего менялись от монолога к монологу, что влияло на уровень кодируемого сигнала S (рис. 6).

Таким образом, получается, что создаваемая нагрузка на сеть и параметры распределения случайных величин  $T_{on}$  и  $T_{off}$  зависят от уровня S, который можно определить следующим образом:

$$S = \bar{N}_{on} / \bar{N}_{off}.$$

Поскольку случайная величина, распределенная по выбранному по результатам тестирования логнормальному распределению  $LN(\mu, \sigma)$ , принимает значения из области  $[0, \infty)$ , то к параметрам  $\mu$  и  $\sigma$  данного распределения следует добавить максимальные и минимальные возможные значения случайной величины. Обозначим  $T_{on}^{max}(S), T_{off}^{max}(S), T_{on}^{min}(S), T_{off}^{min}(S)$  - максимальные и минимальные продолжительности интервалов с речью и паузами соответственно. Из выше сказанного следует:

$$T_{on} \sim LN(\mu_{on}(S), \sigma_{on}(S)), \quad T_{on} \in [T_{on}^{min}(S), T_{on}^{max}(S)];$$

$$T_{off} \sim LN(\mu_{off}(S), \sigma_{off}(S)), \quad T_{off} \in [T_{off}^{min}(S), T_{off}^{max}(S)],$$

где функции

$$\mu_{on}(S), \sigma_{on}(S), T_{on}^{min}(S), T_{on}^{max}(S),$$

$$\mu_{off}(S), \sigma_{off}(S), T_{off}^{min}(S), T_{off}^{max}(S)$$

были получены в результате аппроксимации экспериментальных данных в смысле наименьших квадратов полиномами до третьей степени включительно (рис. 7):

$$T_{on}^{max}(S) = \begin{cases} 0.05 \times S^3 - 0.87 \times S^2 + 5.96 \times S - 1.73, & S \in [0.72, 15.29] \\ -0.0002 \times S^2 + 0.45 \times S + 42.99, & S \in [15.29, 910.3], \end{cases}$$

$$T_{on}^{min}(S) = -3e - 7 \times S^3 + 0.0002 \times S^2 - 0.002 \times S + 0.54,$$

$$T_{off}^{max}(S) = 6/\sqrt{S},$$

$$T_{off}^{min}(S) = 2e - 10 \times S^3 - 4e - 7 \times S^2 + 0.0002 \times S + 0.09,$$

$$\mu_{on}(S) = \begin{cases} 0.0005 \times S^3 - 0.02 \times S^2 + 0.29 \times S - 0.37, & S \in [0.72, 25.8] \\ 1e - 9 \times S^3 - 2e - 5 \times S^2 + 0.01 \times S + 1.63, & S \in [25.8, 910.3], \end{cases}$$

$$\sigma_{on}(S) = \begin{cases} 0.002 \times S^3 - 0.04 \times S^2 + 0.20 \times S + 0.32, & S \in [0.72, 10.69] \\ 2e - 9 \times S^3 - 2e - 6 \times S^2 + 0.0006 \times S + 1.12, & S \in [10.69, 910.3], \end{cases}$$

$$\mu_{off}(S) = \begin{cases} -0.005 \times S^3 + 0.09 \times S^2 - 0.57 \times S + 0.09, & S \in [0.72, 10.29] \\ -3e - 9 \times S^3 + 5e - 6 \times S^2 - 0.003 \times S - 1.19, & S \in [10.29, 910.3], \end{cases}$$

$$\sigma_{off}(S) = \begin{cases} -0.0002 \times S^3 + 0.007 \times S^2 - 0.08 \times S + 1.03, & S \in [0.72, 27.56] \\ -1e - 9 \times S^3 + 2e - 6 \times S^2 - 0.001 \times S + 0.65, & S \in [27.56, 910.3]. \end{cases}$$

По причине громоздкости коэффициенты были округлены до первых значащих цифр.

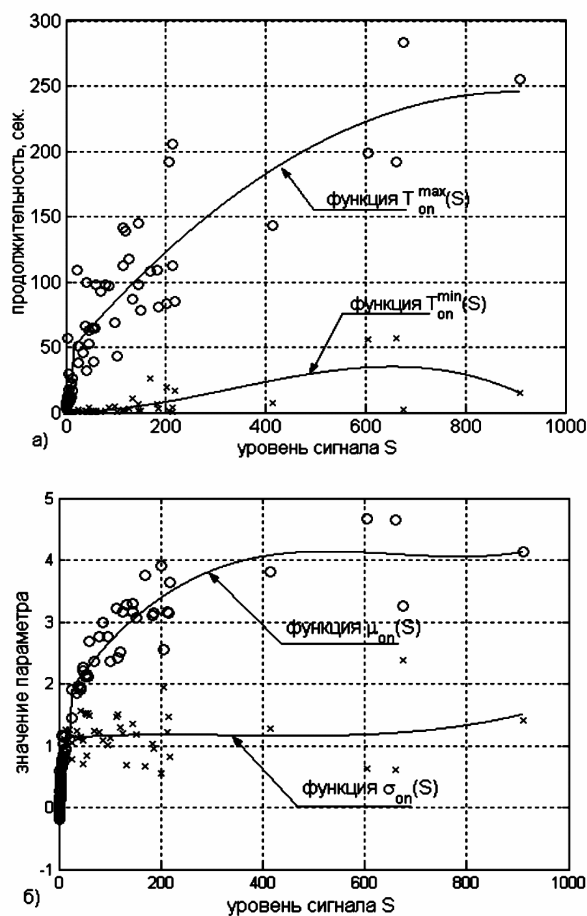


Рис. 7. Аппроксимация экспериментальных данных а) максимальное и минимальное значения интервала речи при заданном  $S$  – функции  $T_{on}^{min}(S)$  и  $T_{on}^{max}(S)$  б) параметры логнормального распределения интервала речи при заданном  $S$  – функции  $\mu_{on}(S)$  и  $\sigma_{on}(S)$

Значение переменной  $S$  изменяется в пределах от 0.72 до 910.3.  $S = 0.72$  соответствует нижнему граничному уровню слышимости и минимальному объему передаваемых данных. В противоположность этому значение  $S = 910.3$  соответствует практически полному отсутствию пауз и, как следствие, максимально возможной нагрузки на сеть. При отсутствии постороннего шума и ручной настройки чувствительности микрофона (присутствие сигнала только в присутствии речи) значение  $S$  составляет в среднем, как правило, 3.6.

В соответствии с полученными результатами в среде моделирования J-sim было проведено моделирование нагрузки, создаваемой на сеть при различных значениях уровня кодируемого сигнала  $S$ . При значении  $S = 3.6$  ошибка составила 0.2%. Максимальная ошибка 7.4% была получена при значении  $S = 0.72$ . Значительная разница в значениях ошибки объясняется крутизной кривых, полученных при аппроксимации, и малым количеством экспериментальных данных в данной области. При необходимости положение можно поправить дополнительными измерениями.

### Заключение

Построенная статистическая модель трафика с достаточно высокой степенью точности соответствует опытным данным и может быть полезна при оптимизации IP-сети, осуществляющей передачу речи. Логичным продолжением данной работы является построение статистической модели трафика при диалоге.

### Литература

1. A. Estepa, R. Estepa, and J. Vozmediano A New Approach for VoIP Traffic Characterization, IEEE communications letters, vol. 8, no. 10, october 2004
2. Leland W. E., Taqqu M. S., Willinger W., Wilson D. V. On the self-similar nature of Ethernet traffic, IEEE Transactions on networking. – Vol. 12. – 1994. - № 1. – P. 2-15.
3. RFC 1889 A Transport Protocol for Real-Time Applications H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, January 1996
4. Сахаров А. В. Анализ частотных свойств трафика. Сборник докладов Всероссийской научно-технической конференции, посвященной 60 летию Победы в Великой Отечественной войне и 110 летию изобретения радио А. С. Поповым – Нижний Новгород 2005
5. Сахаров А.В. Задержка в стеке протоколов UDP/IP/ETHERNET. Сборник трудов по работе 2-го Международного форума «Актуальные проблемы современной науки», Самара 2006