

УДК 621.391.31

АДАПТИВНЫЙ БУФЕР КОМПЕНСАЦИИ ДЖИТТЕРА ЗАДЕРЖКИ ПРИБЫТИЯ ПАКЕТОВ НА ОСНОВЕ РОБАСТНОГО ФИЛЬТРА КАЛМАНА



[А.В. КОБРИН](#)

Харьковский национальный
университет радиоэлектроники

Abstract – In the paper author has suggested a model of the delay included jumps and emissions, it allows simulating a more real process of delays. Multimedia streaming transmission might be affected by jitter. To avoid transmissions errors there are jitter buffer used, therefore it is necessary to estimate the current jitter although it is impossible to calculate straight value online. Author has proposed to use the robust Kalman filter to estimate the delay. We have suggested the algorithm for adaptive playback buffer, which allows to smooth jumps and emissions. Filter ignores individual emissions of delay and switch to the new delay value when delay switches the trend or the level of value. The filter allows you to switch quickly, because it has window of observation allows us to take a decision about the type of distortion delay. The buffer was developed for use at boundary separating media, for example, at boundary wired and wireless media. Because there are huge quality loss.

Анотація – Запропоновано алгоритм адаптивного буфера відтворення, що дозволяє згладжувати стрибки і викиди. В основі алгоритму використовується робастний фільтр Калмана. Фільтр, завдяки своїй робастності, дозволяє ігнорувати поодинокі сплески затримки і швидко перемикатися на її нове значення при зміні тренда або рівня затримки.

Аннотация – Предложен алгоритм адаптивного буфера воспроизведения, позволяющего сглаживать скачки и выбросы. В основе алгоритма используется робастный фильтр Калмана. Фильтр, благодаря своей робастности, позволяет игнорировать единичные всплески задержки и быстро переключаться на ее новое значение при изменении тренда или уровня задержки.

Введение

Интеграция передачи данных и мультимедиа трафика реального времени в сетях с коммутацией пакетов по-прежнему является серьезной проблемой. Одной из проблем передачи речи в реальном времени по пакетным сетям является джиттер. Эту проблему можно решить на стороне отправителя, в сетевой инфраструктуре или на стороне получателя. Отправитель может адаптировать скорость передачи при обнаружении сетевых перегрузок. Сетевая инфраструктура может управлять пропускной способностью через приоритизацию трафика. На стороне получателя на основе некоторого алгоритма может компенсировать джиттер [1]. Компенсация на приемной стороне весьма перспективна в силу того, что не требует изменений в методах кодирования в речевом кодере/декодере на стороне отправителя или изменений в существующей инфраструктуре сети.

Даже в современных сетях джиттер является серьезной проблемой. На рис. 1 показано прохождение речевых пакетов от отправителя к получателю через буфер

компенсации джиттера на стороне получателя. Речевые пакеты формируются периодически (например, каждые 20 мс) в кодере отправителя и передаются в сеть, где из-за возникновения очередей и многопутевой маршрутизации, балансировки нагрузки и других причин образуется случайная задержка прибытия пакетов. На стороне отправителя буфер компенсации джиттера (он же так называемый джиттер-буфер или сглаживающий буфер) может быть использован для задержки принятых пакетов, для того чтобы минимизировать джиттер и восстановить допустимый интервал между поступающими пакетами.

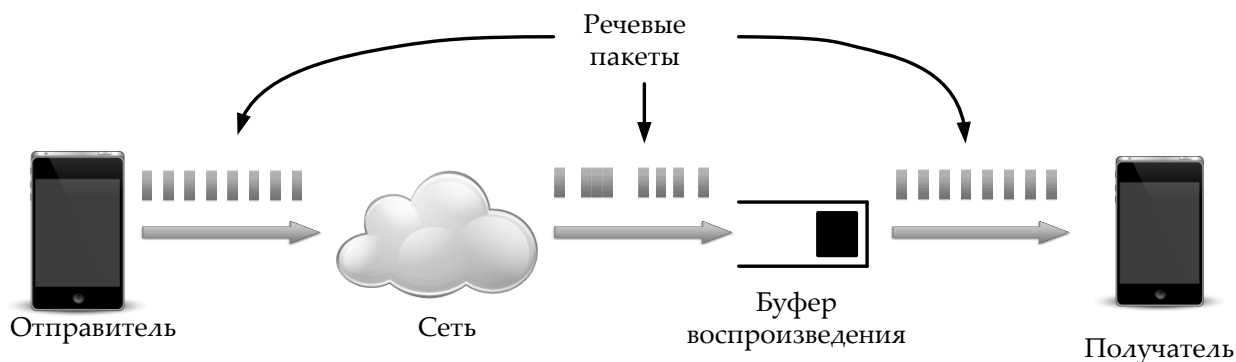


Рис. 1. Компенсация джиттера голосовых пакетов

Буфер компенсации джиттера на стороне получателя удерживает каждый принятый пакет на величину времени буфера, в котором компенсируется дисперсия задержки сети без чрезмерной задержки воспроизведения. Если межпакетная задержка будет превышать время буфера, буфер будет опустошаться, и декодеру не будет хватать пакетов, чтобы воспроизводить речь. Это приводит к неравномерности воспроизведения речи. Ясно, что чем дольше задержка, тем больше пакетов прибывает до их запланированного времени воспроизведения и тем лучше будет компенсация джиттера. В то же время длительные задержки нежелательны, так как они ухудшают интерактивность человеческого общения. Отметим, что человеческое ухо терпимо относится к максимальным задержкам от 150 до 400 миллисекунд [2]. Различные схемы кодирования также могут иметь различные допуски к потерям. Как следствие, алгоритм компенсации джиттера должен быть в состоянии найти компромисс между задержкой воспроизведения и потерями пакетов.

Цель статьи – это разработка алгоритма компенсации джиттера, позволяющего дифференцированно обрабатывать различные типы джиттера. Выделяют три вида джиттера, которые могут быть вызваны различными причинами:

1. Выброс задержки – это когда один пакет в потоке оказывается задержанным на значительно больший интервал времени по отношению к другим. Это наблюдается в тех случаях, когда происходят сетевые перегрузки, передача высокоприоритетного служебного трафика, изменение маршрута и др.

2. Постоянный джиттер – это передача с примерно постоянным изменением задержки пакетов.

3. Скачок задержки возникает из-за всплеска пакетной активности. Это явление, как правило, связано с перегрузками линии доступа или изменением маршрута.

I. Обзор существующих работ в области адаптивных алгоритмов компенсации джиттера

Перед тем как рассмотреть существующие алгоритмы компенсации джиттера, рассмотрим временные параметры (рис. 2), связанные с i -м пакетом в аудио-поток:

t_i – время, когда i -й пакет был сгенерирован на стороне отправителя;

a_i – время, когда i -й пакет был принят на стороне получателя;

p_i – время, когда i -й пакет был воспроизведен на стороне получателя;

D_{prop} – задержка распространения от отправителя к получателю (допускается, что она является константой на протяжении всего речевого соединения);

v_i – задержка в очередях между отправителем и получателем;

b_i – размер буфера компенсации джиттера на стороне получателя для i -ого пакета, $b_i = p_i - a_i$;

d_i – задержка воспроизведения для i -го пакета, $d_i = p_i - t_i$;

n_i – общая задержка, введенная сетью для i -го пакета, $n_i = a_i - t_i$.

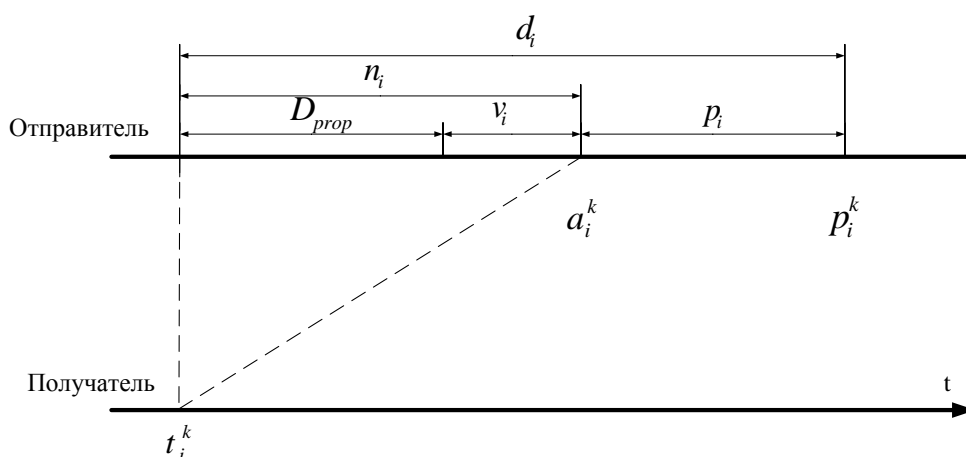


Рис. 2. Временные параметры, ассоциированные с i -м пакетом

Алгоритмы, позволяющие выполнить компенсацию джиттера в режиме реального времени путем буферизации пакетов на стороне получателя, могут быть классифицированы как фиксированные или адаптивные [2]. Хотя фиксированный алгоритм, который использует фиксированное время буфера, проще в реализации, чем другие алгоритмы, он может привести к неудовлетворительному качеству речи [2]. Адаптивные алгоритмы позволяют изменять задержку воспроизведения, используя небольшие правки, которые незаметны в восстановленной речи. Адаптивные алгоритмы различают по способу компенсации джиттера на алгоритмы, выполняющие компенсацию в промежутках между речевыми потоками, и на алгоритмы, выполняющие компенсацию во время речевого потока [2]. В первом случае (рис. 3) произ-

водится корректировка времени воспроизведения первого пакета речевого потока, а все остальные пакеты воспроизводятся через фиксированный интервал после первого пакета. Пакеты, полученные позже запланированного времени их воспроизведения, отбрасываются, и кодер может либо повторить последний принятый пакет, вставить паузу либо воспроизвести другие экстраполированные звуки.

В схеме с компенсацией джиттера во время речевого потока время воспроизведения всех последующих пакетов растягивается всякий раз, когда пакет чрезмерно задерживается в сети. Этот способ сводит к минимуму потери пакетов, но влияет на длину исходного речевого потока, что приводит к проблемам с разборчивостью восстановленной речи [2].

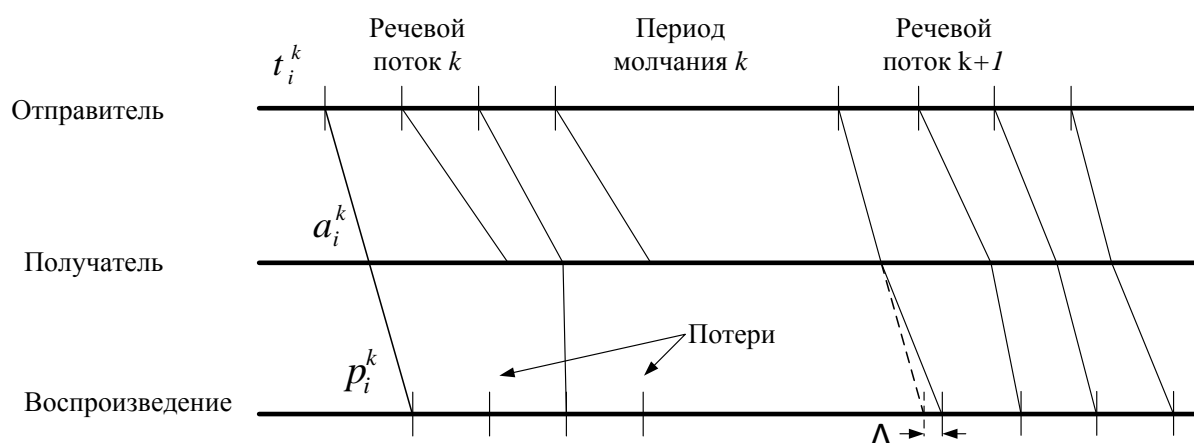


Рис. 3. Схема регулировки задержки воспроизведения в паузах между речевыми потоками

Опишем метод, введенный в [1], который является базовым линейным рекурсивным фильтром. Чтобы определить время воспроизведения для i -ого пакета, мы рассмотрим два случая, в зависимости от того является i -й пакет первым в речевом потоке или нет. Если i -й пакет является первым в речевом потоке, то его время воспроизведения рассчитывается как:

$$p_i = t_i + \hat{d}_i + \gamma \cdot \hat{v}_i,$$

где \hat{d} и \hat{v}_i – это оценка среднего и отклонения сквозной задержки в течение речевого потока; γ – это константа, используемая для установки времени воспроизведения так, чтобы только небольшая часть поступающих пакетов была потеряна [1]. Эта константа равна 4 во всех экспериментах, выполняемых в [1]. В [2] это значение варьируют от 1 до 20, чтобы добиться различного процента потерь. Чтобы пересчитать среднюю задержку и ее отклонения, используются следующие уравнения:

$$\hat{d}_i^k = \alpha \cdot \hat{d}_{i-1}^k + (1 - \alpha) \cdot d_i^k,$$

$$\hat{v}_i^k = \alpha \cdot \hat{v}_{i-1}^k + (1 - \alpha) \cdot |\hat{d}_i^k - d_i^k|.$$

Эти уравнения представляют собой линейные рекурсивные фильтры, где коэффициент α называется шаговой постоянной, $\alpha \leq 1$, и обеспечивает устойчивость процедуры. В работе [1] значение α выбрано равным 0,98002.

Очевидно, что эти уравнения являются преобразованными уравнениями стохастической аппроксимации типа Роббинса-Монро. Можно показать, что получаемые оценки \hat{d}_i^k и \hat{v}_i^k , являются оптимальными для оценивания случайной величины, для которых имеет место уравнение состояния $x_k = x_{k-1}$. В нашем случае сетевая задержка представляет случайный процесс, для которой оптимальной является процедура Калмана-Бьюси из-за оптимизационной постановки задачи.

II. Синтез алгоритма буфера компенсации джиттера на основе фильтра Калмана

Достаточно конструктивной моделью случайного динамического процесса является формирующий фильтр, описываемый уравнением состояния.

$$dx(t)/dt = F(t)x(t) + G(t)\zeta(t).$$

Учитывая то, что рассматривается информационный обмен в цифровой форме в виде пакетов, уравнение состояния, описывающее случайные изменения задержки на каждом k -ом шаге дискретизации, представляется в виде:

$$x(k+1) = F(k+1, k)x(k) + G(k)\zeta(k), \quad (1)$$

где $F(k+1, k)$ – матрица перехода; $G(k)$ – порождающий коэффициент; $\zeta(k)$ – порождающая последовательность, выбранная из гауссового белого шума (ГБШ), со спектральной плотностью мощности N_ζ .

Процесс измерения задержки для упрощения будем считать линейным. Уравнение наблюдения в линейном приближении представляется в виде:

$$y(k) = Hx(k) + v(k), \quad (2)$$

где $v(k)$ – фазовый шум ошибки измерения, являющийся белым Гауссовым шумом со спектральной плотностью мощности N_v , некоррелированный с процессом $\zeta(k)$.

Но на практике, модель описанная уравнением (1), является идеализированной, в реальных же ситуациях процесс задержки претерпевает различные случайные отклонения от нормального закона распределения, обусловленные наличием инерционных элементов, таких как буферы, маршрутизаторы и др.

Очевидно, указанные выбросы и скачки можно представить как в уравнении состояния (1), так и в уравнении наблюдения (2). Учитывая то, что для оценки задержки будем использовать рекуррентный фильтр Калмана (ФК), оставим уравнение состояния (1) без изменения и преобразуем уравнение наблюдения (2). При наличии кратковременных выбросов компонент помех $v(k)$ соответственно преобразовывается с учетом вероятности появления выброса r_v , приобретая вид [3]:

$$v(k) = (1 - r_v)L(v(k) + r_v L(v_v(k))), \quad (3)$$

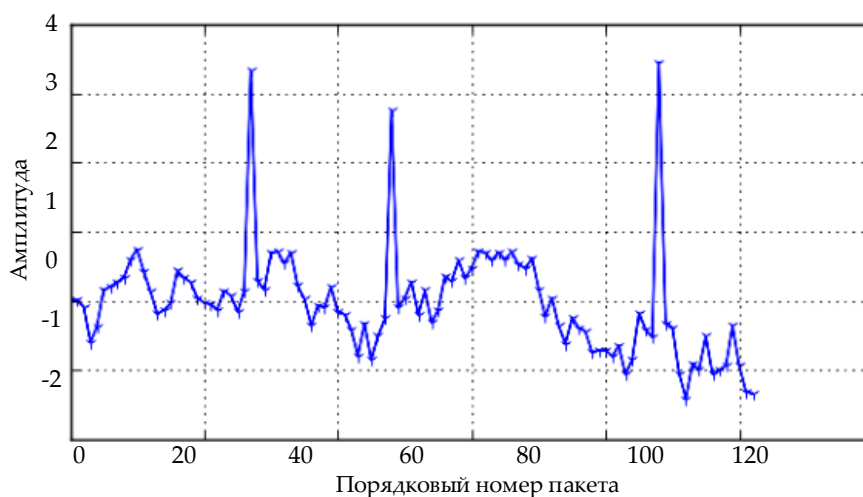
где $L(v_v(k))$ представляет собой случайный процесс выброса, а $0 \leq r_v \leq 1$ – вероятность появления выброса относительно стационарного случайного процесса $v(k)$.

Иная ситуация при появлении скачка, который влияет на уравнение наблюдения в целом, после скачка в уравнении наблюдения появляется смещение, изменяющее несколько последующих состояний, это уравнение представлено в виде:

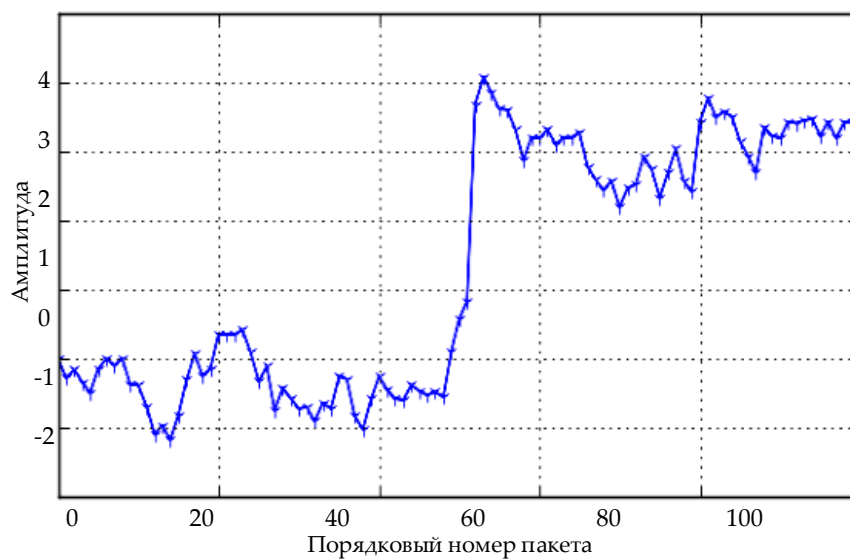
$$y(k) = (1 - r_s)L(y(k) + r_s L(y_s(k))), \quad (4)$$

где $L(y_s(k))$ представляет собой уравнение наблюдения случайного процесса скачка, а $0 \leq r_s \leq 1$ представляет собой вероятность появления скачка.

С помощью предложенной модели случайного процесса сгенерируем ряд задержек с выбросами рис. 4 а) и скачками рис. 4 б).



а)



б)

Рис. 4. Моделирование ряда задержек с выбросами (а) и скачками (б)

Поскольку задержка в соответствии с уравнением (2) наблюдается на фоне гауссового белого шума, а само значение задержки случайно из-за множества факторов, то можно утверждать, что в силу центральной предельной теоремы распределение случайной задержки также подчиняется нормальному закону. Знание закона распределения и использование в таких случаях минимума среднеквадратичного отклонения позволяет рассчитывать на то, что полученные оценки окажутся более точными из-за оптимизационной постановки задачи.

Уравнение оценки в виде условного среднего значения задержки с использованием ФК [4] имеет вид:

$$\hat{x}(k+1) = F\hat{x}(k) + K(k)\Delta y, \quad (5)$$

где $\Delta y = HF\hat{x}(k) - y(k)$ – невязка; $K(k)$ – коэффициент, обеспечивающий устойчивость и сходимость процедуры.

В частности, для упрощения процедуры часто выбирают K в виде константы ($K \leq 1$). Коэффициент усиления ФК $K(k)$ является функцией от апостериорной дисперсии $v(k)$, что ускоряет его сходимость:

$$K(k) = V(k)H^T N_v^{-1}.$$

На рис. 5 представлена схема сглаживающего фильтра, построенная в соответствии с (5). Ключевую роль в оценке ФК апостериорного среднего значения задержки, играет параметр F , позволяющий регулировать сглаживающие свойства фильтра.

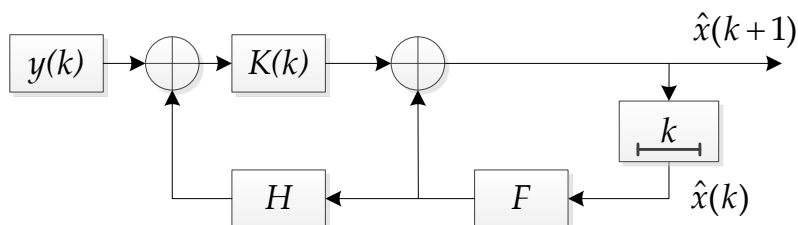


Рис. 5. Схема фильтра Калмана-Бьюси

Очевидно, что с использованием модернизированного уравнения наблюдения, ФК теряет оптимальность, а его сходимость к установившемуся состоянию может оказаться проблематичной. Существует несколько методов, позволяющих обеспечить сходимость ФК, среди которых метод кусочно-линейной аппроксимации, метод дискретно-непрерывных моделей и др. [5, 6]. Данные методы являются близкими к оптимальным, но вычислительно сложны в реализации, поэтому рассмотрим робастный фильтр Калмана (РФК), который является более простым и обеспечивает запас устойчивости в широком диапазоне входных воздействий.

Воспользуемся методологией работы [3]. Короткий выброс задержки в уравнении (3) и возвращение ее в стационарное состояние является, как правило, ложным. Тогда основной задачей фильтра является сгладить данный выброс, обеспечив теку-

щую величину этого значения неизменной по сравнению с прошлым шагом. Процедура Калмана-Бьюси при этом приобретает следующий вид:

$$\hat{x}(k+1) = F(k+1, k)\hat{x}(k) + K(k)\Delta y \min\left\{1, \frac{b}{|K(k)\Delta y|}\right\}, \quad (6)$$

где b является некоторым ограничителем изменения значения функции. Это предложение устраняет проблему неограниченного изменения оценки за один шаг в классическом фильтре Калмана, оставаясь при этом достаточно простым.

Если $b \geq |K(k)\Delta y|$, то $\min\left\{1, \frac{b}{|K(k)\Delta y|}\right\} = 1$, то фильтр работает в обычном режиме как ФК (5). Если же $b < |K(k)\Delta y|$, то из этого следует, что произошел выброс и невязка умножается на понижающий коэффициент, равный $\frac{b}{|K(k)\Delta y|}$, что опять же приводит РФК к обычному виду (5).

РФК для ситуации скачка ведет себя обратно, приобретая вид:

$$\hat{x}(k+1) = F(k+1, k)\hat{x}(k) + H(k)[I - H(k)K(k)\Delta y] \min\left\{1, \frac{b}{|I - H(k)K(k)\Delta y|}\right\},$$

где b тот же аргумент, ограничивающий изменение значения функции, что и для РФК в ситуации выброса (6).

Одновременная фильтрация возможна только с некоторой задержкой $\Delta\tau$. Это необходимо для принятия решения о типе выброса. Вследствие выброса будет отмечен один большой всплеск (3), а вследствие скачка – подряд целая последовательность (4). Гибридный робастный фильтр Калмана (ГРФК) может быть реализован следующим образом [3]: РФК для ситуации скачка работает по умолчанию, и всякий раз, когда ширина выброса больше окна наблюдения $\Delta\tau$, единожды используется процедура РФК для ситуации скачка.

III. Экспериментальное сравнение алгоритмов

Сравним предложенный алгоритм с алгоритмом, рассмотренным в [1]. На рис. 6 изображен график изменения размера буфера от задержки для двух алгоритмов. Как можно увидеть, предложенный алгоритм позволяет определить буфер более точно во время скачков и выбросов задержки.

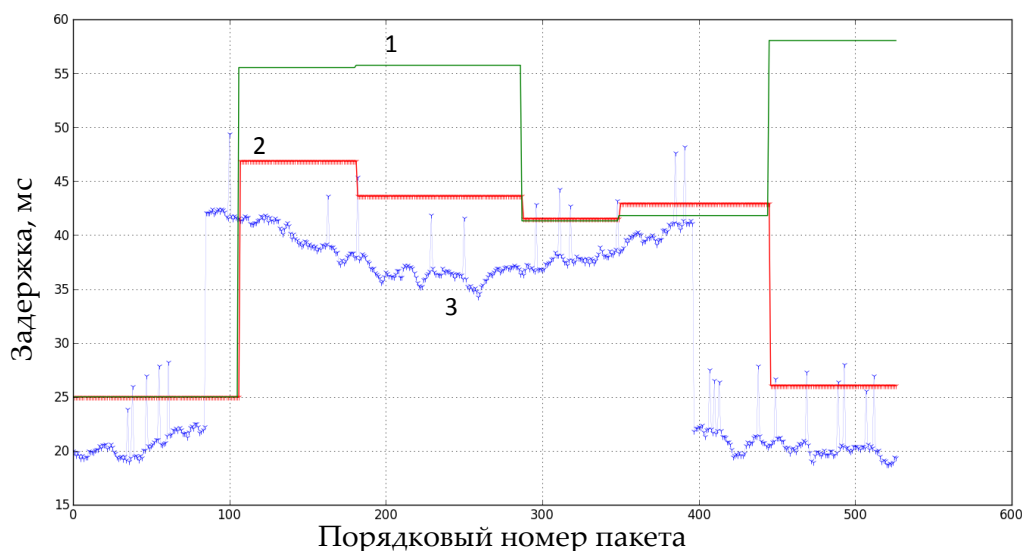


Рис. 6. Изменение размера буфера от задержки
1) адаптивный алгоритм [1],
2) алгоритм на основе фильтра Калмана,
3) процесс задержки

Чтобы оценить получаемое преимущество, воспользуемся [7] для оценки качества речи. Построим график зависимости качества переданной речи для предложенного алгоритма (пунктирная линия) и алгоритма, рассмотренного в [1] (сплошная линия) к вероятности появления выброса рис. 7 а) или скачка рис. 7 б). Размер выброса и скачка в эксперименте был определен до 20 мс.

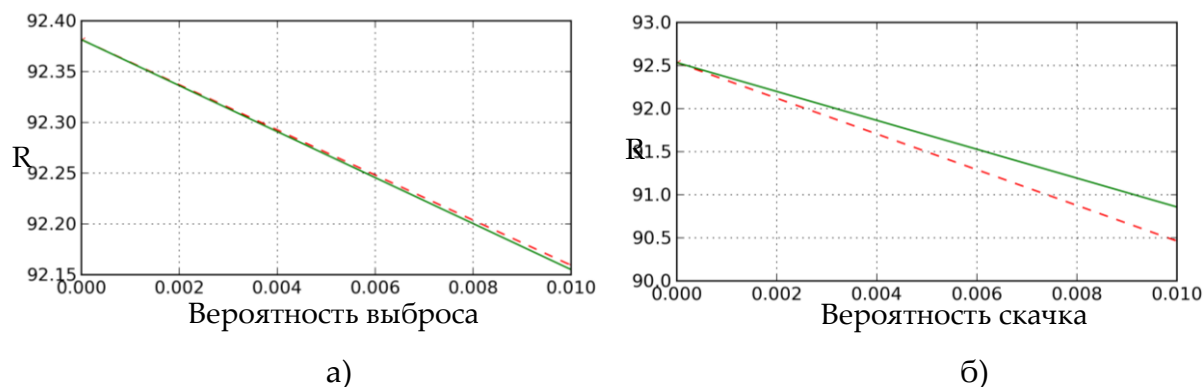


Рис. 7. Качество передачи речи а) для выброса и б) для скачка

Выводы

1. Приведен анализ причин искажения качества передачи пакетной речи, обусловленных наличием выбросов задержки получения пакетов или трендом этих задержек, и предложена модель в виде двухкомпонентной суммы уравнение (3) и (4), позволяющая учитывать эти задержки.

2. Известные алгоритмы адаптации буфера для сглаживания фазового дрожания основаны на оценках средних величин задержки и не являются оптимальными для реальных значений, представленных случайными процессами. Предложено использовать более общую математическую модель робастного фильтра Калмана-Бьюси для компенсации возмущений задержки.

3. Анализ качества работы РФКБ показывает, что по мере возрастания вероятности или амплитуды случайных возмущений задержки выигрыш в воспринимаемом качестве обслуживания предложенного алгоритма увеличивается по сравнению с процедурой, приведенной в [1], приблизительно на 10% и зависит от помеховой ситуации. В случае возмущения задержки буфер позволяет быстро переключиться на новый уровень и обеспечить удовлетворительное качество обслуживания в сложной сетевой ситуации.

Список литературы:

1. *Ramjee R.* Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks / R. Ramjee, J. Kurose, D. Towsley, H. Schulzrinne // Proceedings of IEEE INFOCOM. – 1994. – Toronto. – P. 680- 686.
2. *Moon S.* Packet Audio Playout Delay Adjustment: Performance Bounds and Algorithms / J. Kurose, D. Towsley // Multimedia Systems. – 1998. – Vol. 6. – P. 17-28.
3. *Ruckdeschel P.* Optimally robust kalman filtering // Berichte des Fraunhofer ITWM. – 2010. – No. 185. – 53 p.
4. *Поповский В.В.* Математические основы управления и адаптации в телекоммуникационных системах / В.В. Поповский, В.Ф. Олейник. – Х.: СМИТ, 2011. – 362 с.
5. *Чернавский А.В.* Кусочно линейная аппроксимация вложений многообразий в коразмерностях, больших двух. // Матем. сб. – 1970. – С. 499-500.
6. *Расина И.В.* Дискретно-непрерывные модели и оптимизация управляемых процессов. // Программные системы: теория и приложения. – 2011. – С. 49-72.
7. G.107. ITU-T. Recommendation G.107: The E-model: a computational model for use in transmission planning [Text] / ITU-T. Geneva – 2003. – P. 26.