

УДК 621.391 : 004.725.5

© 2024 г. А.В. Ритерман, Д.В. Банков, А.И. Ляхов, Е.М. Хоров

ОБ ЭФФЕКТИВНОСТИ МЕТОДА ДОСТУПА К КАНАЛУ С ВЫТЕСНЕНИЕМ В СЕТЯХ Wi-Fi 8¹

Для обеспечения надежной доставки пакетов с низкой задержкой, требуемой приложениями реального времени (англ.: real-time applications, RTAs), разрабатывается метод доступа к каналу с вытеснением, который будет определен в дополнении к стандарту Wi-Fi 8. В статье проведено исследование эффективности данного метода доступа в сети с одной RTA-станцией. Для этого разрабатывается аналитическая модель сети Wi-Fi 8, использующей метод доступа к каналу с вытеснением, с помощью которой находятся параметры метода доступа к каналу, при которых обеспечивается низкая задержка и высокая надежность доставки RTA-трафика, а эффективность использования канала станциями, передающими менее приоритетные кадры, максимальна.

Ключевые слова: приложения реального времени, доступ к каналу с вытеснением, Wi-Fi 8, IEEE 802.11bn.

DOI: 10.31857/S0555292324040041, **EDN:** QPNGPH

§ 1. Введение

Поддержка приложений реального времени (англ.: real time applications, RTA), таких как приложения виртуальной и дополненной реальности (VR/AR) и различные промышленные приложения, является важной задачей для современных сетей Wi-Fi. Обычно RTA-трафик требует обеспечивать низкую задержку доставки пакетов (порядка 1–10 мс) с высокой надежностью (вероятность своевременной доставки пакетов порядка 99,999%) [1]. На данный момент сети Wi-Fi не всегда могут удовлетворить требования к качеству обслуживания RTA-трафика. Один из источников такой проблемы состоит в невозможности передачи каких-либо данных в то время, когда канал занят другой передачей, которая может длиться несколько миллисекунд.

Для решения данной проблемы в новом дополнении к стандарту сетей Wi-Fi, IEEE 802.11bn [2, 3], предлагается добавить механизм вытеснения (англ.: preemption). При его использовании станции Wi-Fi, получив доступ к каналу, разделяют передачу неприоритетных данных на фрагменты, между которыми предусмотрены интервалы времени, когда другие станции могут захватить канал для передачи приоритетных данных.

Данный механизм новый для сетей Wi-Fi, и в литературе практически нет исследований его эффективности. На данный момент существует исследование, в котором предложены и описаны три схемы приоритетного доступа к каналу, адаптированные для механизма вытеснения [4]. В первых двух схемах RTA-станции взводят

¹ Исследование выполнено в ИППИ РАН за счет гранта Российского научного фонда № 24-19-00816, <https://rscf.ru/project/24-19-00816/>.

случайные счетчики отсрочки, прежде чем передать данные с помощью механизма вытеснения. Отличие этих схем состоит в том, когда именно уменьшаются значения этих счетчиков. В первой схеме счетчик уменьшается при детектировании кадра, позволяющего использовать механизм вытеснения. Во второй схеме счетчик уменьшается на границах слотов, аналогично тому, как это происходит в стандартном методе доступа EDCA (англ.: enhanced distributed channel access). В третьей схеме для каждой RTA-станции вводится вероятность получения доступа к каналу в момент времени, когда можно начать передачу данных с использованием механизма вытеснения. В работе [4] нет данных о том, позволяют ли рассмотренные подходы удовлетворить требования к качеству обслуживания RTA-трафика.

В статье рассматривается сценарий, в котором в сети есть точка доступа, передающая низкоприоритетные данные подключенным к ней станциям, и RTA-станция, передающая срочные кадры точке доступа. В статье исследуются зависимости распределения задержки доставки срочных кадров и эффективности использования канала от параметров механизма вытеснения и интенсивности RTA-трафика. Под задержкой подразумевается время, прошедшее с момента генерации RTA-кадра до получения подтверждения о его доставке, а под эффективностью использования канала – доля времени, в течение которого точка доступа передает полезные данные, прямо пропорциональная пропускной способности сети. Исследуется, какие факторы сильнее всего влияют на задержку доставки RTA-кадров. Разработана аналитическая модель метода доступа с вытеснением для случая одной станции, передающей RTA-кадры в сети.

Дальнейшее изложение построено следующим образом. В § 2 содержится дополнительная информация о механизме вытеснения. В § 3 приводятся сценарий и постановка задачи. В § 4 описывается математическая модель сети Wi-Fi с механизмом вытеснения. В § 5 приведены полученные численные результаты. В § 6 представлено заключение статьи.

§ 2. Объект исследования

В современных сетях Wi-Fi основным методом доступа к каналу является механизм случайного доступа EDCA, краткое и упрощенное описание которого приведено ниже.

Перед получением доступа к каналу станции Wi-Fi должны прослушать канал и определить, свободен ли он. Если новый кадр на передачу поступает в пустую очередь и канал свободен, то станция сразу начинает передачу, в противном случае она взводит счетчик отсрочки. Счетчик отсрочки инициализируется целым числом, которое равновероятно выбирается из полуинтервала $[0, W_r)$. Параметр W_r называется конкурентным окном, его значение зависит от количества неудачных попыток передачи текущего кадра r следующим образом:

$$W_r = \begin{cases} W_{\min}, & r = 0, \\ \min(2W_{r-1}, W_{\max}), & r > 0, \end{cases} \quad (1)$$

где $W_{\min}, W_{\max}$ – границы конкурентного окна.

Пока канал свободен, счетчик отсрочки уменьшается на единицу после каждого пустого временного слота длительностью σ . Когда канал занят, счетчик отсрочки замораживается. Как только канал оказывается свободным в течение межкадрового интервала AIFS (англ.: arbitration inter-frame space), станция возобновляет отсчет и уменьшает значение счетчика отсрочки.

Когда значение счетчика отсрочки достигает нуля, станция совершает попытку передачи кадра данных. Попытка считается успешной, если станция-отправитель получит кадр подтверждения АСК от станции-получателя через короткий межкад-

ровый интервал SIFS (англ.: short inter-frame spacing). Если станция-отправитель в течение интервала времени AckTimeout после окончания передачи не детектирует начало подтверждения, попытка считается неудачной. В случае неудачи станция увеличивает счетчик r и продолжает попытку передачи кадра данных. При успешной передаче счетчик r обнуляется, и станция переходит к передаче следующих данных при их наличии.

Станция получает доступ к каналу на время TXOP (англ.: transmission opportunity). При этом станция может передать как несколько кадров, так и один. Однако время, в течение которого станция занимает канал, включая кадр подтверждения, не должно превышать предельное значение длительности TXOP (TXOP limit).

Механизм случайного доступа EDCA определяет четыре категории доступа (англ.: access categories, ACs), для каждой из которых установлены свои значения $W_{\min}$, $W_{\max}$, AIFS и TXOP limit. Такое разделение на категории можно использовать для того, чтобы станции, передающие RTA-кадры, могли получать доступ к каналу раньше, чем станции, передающие неприоритетные данные. В частности, чтобы минимизировать вероятность коллизии между станциями, передающими RTA-кадры, и станциями, передающими кадры другого типа, можно установить параметры так, чтобы значение $AIFS + W_{\max}$ для RTA-кадров было меньше значения AIFS для остальных кадров. В то же время остается проблема, связанная с тем, что пока канал занят передачей одной станции, другая не может передавать данные и должна дожидаться освобождения канала, что может привести к длительным задержкам. С одной стороны, можно ограничивать максимальную длительность передачи для неприоритетного трафика так, чтобы учесть максимальную задержку на ожидание конца передачи [5], однако такой подход приведет к снижению эффективности использования канала при передаче неприоритетного трафика, связанному с накладными расходами на получение доступа к каналу и на передачу служебных данных. Этого можно избежать, используя механизм вытеснения, который не требует ограничения максимальной длительности передачи для неприоритетного трафика, и в то же время позволяет станциям, передающим RTA-кадры, получать доступ к каналу, не дожидаясь окончания длительной передачи.

Дополнение к стандарту IEEE 802.11bn еще только разрабатывается, и подробного описания механизма вытеснения в нем нет. Поэтому приведем упрощенное описание данного механизма, сохранив наиболее важные его свойства. Будем рассматривать передачу неприоритетных данных от точки доступа, предполагая, что в ее начале происходит обмен кадрами RTS/CTS (англ.: ready to send/clear to send), чтобы избежать длительной коллизии. Для этого точка доступа начинает передачу кадром RTS, в ответ на которую станция-получатель отправляет CTS, и после получения CTS точка доступа начинает передачу данных. Данные разделяются на несколько коротких фрагментов (длиной T), между которыми предусмотрены промежутки времени длительностью XIFS, которые в данной статье равны PIFS (англ.: PCF inter-frame space, $PIFS = SIFS + \sigma$), когда канал свободен. При отсутствии приоритетного трафика точка доступа совершает передачу до конца TXOP так, чтобы в конце TXOP получить кадр с блочным подтверждением (англ.: Block ACK, BACK), подтверждающим доставку каждого фрагмента. Причем если в TXOP не укладывается целое количество фрагментов длиной T и BACK, то TXOP соответствующим образом заканчивается позже (при этом длительность TXOP L изначально устанавливается такой, чтобы при “удлинении” TXOP не превысил TXOP limit). При наличии RTA-кадров появляется возможность передать их с использованием механизма вытеснения. Для этого станция через SIFS после окончания фрагмента передает RTA-кадр, на который точка доступа передает подтверждение. Через PIFS после передачи подтверждения точка доступа продолжает передачу фрагментов. Если произошло вытеснение, передачу следующего RTA-кадра станция может выполнить только после очередного фрагмента от точки доступа. Другими словами,

RTA-станция не может с помощью механизма вытеснения передать два RTA-кадра подряд. В случае коллизии при вытеснении точка доступа теряет ТХОР, и RTA-станции, чьи передачи попали в коллизию, разрешают ее методом EDCA.

Точка доступа объявляет станциям о доступе с использованием механизма вытеснения и об интервалах времени, когда можно сделать вытеснение, в заголовке первого фрагмента. Для снижения накладных расходов последующие фрагменты имеют более короткие заголовки, которые лишь содержат небольшую преамбулу, необходимую для возобновления синхронизации приемника. При этом в случае успешного вытеснения следующий фрагмент снова содержит полный заголовок, поскольку принимающая станция за время передачи RTA-данных теряет синхронизацию.

Не очевидно, как следует выбирать значение параметра механизма вытеснения T . Например, с одной стороны, значение T следует устанавливать как можно большим, чтобы минимизировать накладные расходы, связанные с межкадровыми интервалами и заголовками. С другой стороны, значение T определяет максимальное время, которое RTA-станции нужно ждать до получения возможности вытеснения. Возникает вопрос, как выбирать параметр механизма вытеснения T , удовлетворяющий требованиям к качеству обслуживания RTA-трафика, при которых эффективность использования канала при передаче неприоритетных данных была бы максимальной.

§ 3. Сценарий и постановка задачи

Будем рассматривать сеть Wi-Fi, состоящую из точки доступа, передающей насыщенные потоки с неприоритетными данными, станций, принимающих данные от точки доступа, и одной станции, генерирующей ненасыщенный RTA-трафик (называемой RTA-станцией). У точки доступа всегда есть данные на передачу, в то время как у RTA-станции после успешной передачи новые данные появляются через временной промежуток, распределенный экспоненциально с параметром λ [6]. Все станции находятся в зоне уверенного приема кадров друг от друга. Будем считать, что в случае одновременной передачи кадров от разных станций происходит коллизия, в результате которой ни один кадр не доставляется успешно. При этом используется достаточно надежная сигнально-кодовая конструкция [7], такая что можно пренебречь влиянием случайного шума на передачи, и в отсутствие коллизий передача данных всегда успешна. При этом для RTA-трафика требуется, чтобы Q^* -квантиль задержки доставки пакетов не превышал ограничения D^* .

Устройства используют механизм вытеснения. Значения AIFS и конкурентных окон для трафика от точки доступа и RTA-станций установлены так, чтобы после освобождения канала RTA-станция всегда получала доступ к каналу раньше точки доступа.

Определим эффективность использования канала как долю времени, в течение которого точка доступа передает части кадров с полезными данными. Эта величина прямо пропорциональна пропускной способности сети. В статье ставится следующая задача: найти длительность фрагмента T , обеспечивающую требуемую задержку и надежность доставки RTA-кадров, при которых эффективность использования канала точкой доступа S максимальна. Данную задачу можно сформулировать как следующую задачу оптимизации:

$$\begin{aligned} \max_T S(T) \\ \text{при условии } P(D(T) \leq D^*) \geq Q^*, \end{aligned} \quad (2)$$

где D – задержка доставки RTA-кадров, D^* – ограничение на эту задержку, Q^* – минимально допустимая вероятность своевременной доставки RTA-кадров.

§ 4. Математическая модель

Для решения поставленной задачи разработана аналитическая модель, описание которой приводится ниже. Данная модель построена на основе аналитических моделей, представленных в [5, 8].

Пусть интенсивность трафика от RTA-станции намного меньше интенсивности трафика от точки доступа. Тогда можно рассматривать процесс передачи RTA-станции как некоторое возмущение поверх стационарного процесса передачи точки доступа.

В описанной ниже математической модели параметры $W_{\min}^{\text{AP}}$, $W_{\max}^{\text{AP}}$ и AIFS_{AP} обозначают границы конкурентного окна и AIFS для точки доступа, а параметры $W_{\min}^{\text{RTA}}$, $W_{\max}^{\text{RTA}}$ и AIFS_{RTA} – для RTA-станции.

4.1. Передача точки доступа. Опишем процесс передачи точки доступа в отсутствие RTA-станции. Время в канале, аналогично [9], делится на виртуальные слоты – интервалы времени между последовательным уменьшением счетчика отсрочки. Процесс передачи точки доступа можно разделить на интервалы (далее – *интервалы обслуживания*), состоящие из отсчета отсрочки и передачи кадра. Рассмотрим один такой интервал обслуживания.

Точка доступа перед передачей отсчитывает отсрочку, среднее значение которой составляет

$$\bar{b} = \frac{W_{\min}^{\text{AP}} - 1}{2}.$$

После этого точка доступа передает в течение L_{ext} – длительности ТХОР, удлиненной так, чтобы передать целое число фрагментов и ВАСК:

$$L_{\text{ext}} = T_{\text{first}} + k \cdot T_{\text{mid}} + T_{\text{last}}, \quad (3)$$

где T_{first} – длительность интервала (далее называемого *первым интервалом*) времени между началом передачи от точки доступа и первой возможностью прерывания, T_{mid} – длительность интервала (далее – *средний интервал*) времени между моментами, когда возможно прерывание, T_{last} – длительность интервала (далее – *последний интервал*) времени между концом последнего из средних интервалов и концом ТХОР, k – максимальное количество интервалов длиной T_{mid} , которые могут уложиться в L_{ext} так, чтобы в начале успел произойти обмен кадрами RTS/CTS, а в конце точка доступа получила ВАСК. Длительности интервалов равняются

$$\begin{aligned} T_{\text{first}} &= T_{\text{RTS}} + \text{SIFS} + T_{\text{CTS}} + \text{SIFS} + T + \text{SIFS}, \\ T_{\text{mid}} &= \sigma + T + \text{SIFS}, \\ T_{\text{last}} &= \sigma + T + \text{SIFS} + \text{ВАСК}, \end{aligned}$$

где T_{RTS} и T_{CTS} – длительности кадров RTS и CTS соответственно. Значение k находится из системы

$$\begin{cases} k = \left\lceil \frac{L - T_{\text{first}} - T_{\text{last}}}{T_{\text{mid}}} \right\rceil, & L - T_{\text{first}} - T_{\text{last}} > 0, \\ k = 0, & L - T_{\text{first}} - T_{\text{last}} \leq 0. \end{cases} \quad (4)$$

После передачи точка доступа ждет AIFS_{AP} . Таким образом, средняя длительность интервала обслуживания равна

$$L_{\text{period}} = \bar{b} \cdot \sigma + L_{\text{ext}} + \text{AIFS}_{\text{AP}}. \quad (5)$$

4.2. Распределение задержки. Рассмотрим передачу RTA-кадра. Распределение задержки RTA-кадра зависит от состояния канала в момент его генерации. Можно

выделить следующие возможные события: генерация кадра в пустом слоте, в первом интервале, в одном из средних и в последнем. Вероятности этих событий обозначим через p_{idle} , p_{first} , p_{mid} , p_{last} соответственно. Функции распределения задержки, соответствующие этим событиям, обозначим через $F_{\text{idle}}(t)$, $F_{\text{first}}(t)$, $F_{\text{mid}}(t)$, $F_{\text{last}}(t)$.

Итоговая функция распределения задержки имеет вид

$$F(t) = p_{\text{idle}} \cdot F_{\text{idle}}(t) + p_{\text{first}} \cdot F_{\text{first}}(t) + p_{\text{mid}} \cdot F_{\text{mid}}(t) + p_{\text{last}} \cdot F_{\text{last}}(t). \quad (6)$$

Рассмотрим каждый из четырех случаев подробнее.

Генерация в пустом слоте. Согласно сценарию значение AIFS для RTA-станции меньше, чем для точки доступа, поэтому после освобождения канала RTA-станция начинает отсчет отсрочки на

$$\Delta_{\text{AIFS}} = \frac{\text{AIFS}_{\text{AP}} - \text{AIFS}_{\text{RTA}}}{\sigma}$$

пустых слотов раньше, чем точка доступа. Учитывая это, вероятность p_{idle} может быть найдена как доля времени, когда канал свободен:

$$p_{\text{idle}} = \frac{(\bar{b} + \Delta_{\text{AIFS}}) \cdot \sigma}{L_{\text{period}}}. \quad (7)$$

Рассмотрим произвольный пустой слот с точки зрения RTA-станции. Вероятность того, что точка доступа начнет передавать на границе рассматриваемого слота, равна среднему количеству передач за интервал обслуживания, деленному на среднее количество слотов в интервале обслуживания:

$$\tau = \frac{1}{\bar{b} + \Delta_{\text{AIFS}} + 1}. \quad (8)$$

При генерации RTA-кадра во время пустого слота RTA-станция начинает передачу кадра на границе слота, и возможно несколько случаев:

- Если точка доступа не передает в следующем слоте (с вероятностью $1 - \tau$), происходит успешная передача RTA-станции. Соответствующую функцию распределения задержки обозначим через $F_{\text{idle}}^s(t)$;
- Если точка доступа передает в следующем слоте (с вероятностью τ), происходит коллизия между кадрами RTA-станции и точки доступа. Соответствующую функцию распределения задержки обозначим через $F_{\text{idle}}^c(t)$.

Результирующая функция распределения задержки $F_{\text{idle}}(t)$ имеет вид

$$F_{\text{idle}}(t) = (1 - \tau) \cdot F_{\text{idle}}^s(t) + \tau \cdot F_{\text{idle}}^c(t). \quad (9)$$

Найдем входящие в нее вероятности и функции распределения задержек.

Успешная передача RTA-станции. Задержка передачи кадра в случае успешной передачи складывается из времени ожидания от момента генерации кадра до начала следующего слота и последующей передачи кадра. Чтобы найти ее функцию распределения, рассмотрим задачу в общем случае, когда RTA-кадр генерируется в течение интервала времени длительностью g , а после окончания интервала передается за время s . Найдем вероятность того, что задержка RTA-кадра, которую обозначим через D , будет не больше чем t . Пусть x – время от начала интервала g до момента генерации RTA-кадра, причем x – экспоненциально распределенная на отрезке $[0, g]$ случайная величина.

Искомая вероятность равна

$$\mathbf{P}(D < t | x < g) = \mathbf{P}(g - x + s < t | x < g) = \frac{\mathbf{P}(g + s - t < x < g)}{\mathbf{P}(x < g)}.$$

Введем функцию $\Phi(t, g, s)$ для упрощения выкладок:

$$\Phi(t, g, s) = \frac{\mathbf{P}(g + s - t < x < g)}{\mathbf{P}(x < g)} = \begin{cases} 0, & t < s, \\ \frac{e^{-\lambda(g+s-t)} - e^{-\lambda g}}{1 - e^{-\lambda g}}, & s \leq t < g + s, \\ 1, & t \geq g + s. \end{cases} \quad (10)$$

Используя (10), функцию распределения задержки $F_{\text{idle}}^s(t)$ можно записать как

$$F_{\text{idle}}^s(t) = \Phi(t, \sigma, T_r), \quad (11)$$

где

$$T_r = \text{DATA} + \text{SIFS} + \text{ACK}$$

– время успешной передачи RTA-станции (DATA – длительность кадра данных RTA-станции, ACK – длительность подтверждения).

Коллизия между RTA-станцией и точкой доступа. С вероятностью τ при попытке передачи RTA-кадра возникает коллизия с точкой доступа. При коллизии с точкой доступа RTA-станция взведет счетчик отсрочки b , отсчет которого начнется после интервала времени, равного длительности коллизии

$$T_c = \max\{\text{RTS}, \text{DATA}\} + \text{ACKTIMEOUT} + \text{AIFS}_{\text{RTA}},$$

где ACKTIMEOUT – время, в течение которого станция ожидает получения подтверждения. Затем через время $b \cdot \sigma$ RTA-станция совершит повторную попытку передачи. Функция распределения задержки при конкретном значении b будет равна

$$\Phi(t, \sigma, T_c + b \cdot \sigma + T_r).$$

Согласно сценарию у RTA-станции установлены приоритетные значения AIFS и конкурентных окон (таблица 1), поэтому повторная коллизия с точкой доступа невозможна. Учитывая то, что после коллизии счетчик отсрочки выбирается равномерно из полуинтервала $[0, W_1^{\text{RTA}}]$, где

$$W_1^{\text{RTA}} = \min\{2W_{\min}^{\text{RTA}}, W_{\max}^{\text{RTA}}\},$$

получим функцию распределения задержки $F_{\text{idle}}^c(t)$ в случае коллизии RTA-станции с точкой доступа:

$$F_{\text{idle}}^c(t) = \frac{1}{W_1^{\text{RTA}}} \sum_{b=0}^{W_1^{\text{RTA}}-1} \Phi(t, \sigma, T_c + b \cdot \sigma + T_r). \quad (12)$$

Генерация в первом и средних интервалах Вероятности генерации RTA-кадра в первом интервале и в одном из средних интервалов могут быть приближенно найдены как

$$p_{\text{first}} = \frac{T_{\text{first}}}{L_{\text{period}}}, \quad (13)$$

$$p_{\text{mid}} = \frac{k \cdot T_{\text{mid}}}{L_{\text{period}}}. \quad (14)$$

Кадр, сгенерированный в одном из таких интервалов, будет успешно передан с использованием механизма вытеснения при первой такой возможности. Задержка при успешной передаче складывается из времени ожидания, отсчитанного от момента генерации кадра до конца интервала, т.е. до возможности сделать вытеснение, и длительности T_r . Длительности T_{first} и T_{mid} различны, поэтому функция распределения задержки при генерации RТА-кадра в первом интервале находится как

$$F_{\text{first}}(t) = \Phi(t, T_{\text{first}}, T_r), \quad (15)$$

а при генерации RТА-кадра в одном из средних интервалов – как

$$F_{\text{mid}}(t) = \Phi(t, T_{\text{mid}}, T_r) \quad (16)$$

соответственно.

Генерация в последнем интервале. Вероятность генерации RТА-кадра в последнем интервале равна

$$p_{\text{last}} = \frac{T_{\text{last}}}{L_{\text{period}}}. \quad (17)$$

При генерации кадра в последнем слоте RТА-станция взводит счетчик отсрочки, который равновероятно выбирается из полуинтервала $[0, W_{\text{min}}^{\text{RТА}})$, ждет окончания TXOP, после чего через время $\text{AIFS}_{\text{RТА}}$ начинает отсчет отсрочки, который достигнет значения 0 спустя время $b \cdot \sigma$, и затем успешно передает RТА-кадр.

Задержка передачи RТА-кадра будет складываться из времени, прошедшего от момента генерации кадра до конца TXOP, $\text{AIFS}_{\text{RТА}}$, времени отсчета отсрочки и T_r . Согласно сценарию у RТА-станции установлены приоритетные значения AIFS и конкурентных окон, поэтому коллизия с точкой доступа невозможна. По аналогии с формулой (12) функция распределения задержки при конкретном значении счетчика отсрочки будет равна

$$\Phi(t, T_{\text{last}} + \text{AIFS}_{\text{RТА}}, b \cdot \sigma + T_r).$$

Учитывая, что счетчик отсрочки выбирается равновероятно из полуинтервала $[0, W_{\text{min}}^{\text{RТА}})$, получим функцию распределения задержки $F_{\text{last}}(t)$:

$$F_{\text{last}}(t) = \frac{1}{W_{\text{min}}^{\text{RТА}}} \sum_{b=0}^{W_{\text{min}}^{\text{RТА}}-1} \Phi(t, T_{\text{last}} + \text{AIFS}_{\text{RТА}}, b \cdot \sigma + T_r), \quad (18)$$

4.3. Эффективность использования канала. Оценим эффективность использования канала точкой доступа.

Эффективность использования канала при отсутствии передачи RТА-кадра за интервал обслуживания обозначим через S_0 . За это время (L_{period}) точка доступа передаст $k + 2$ фрагмента длиной T : по одному в первом и последнем интервалах, k фрагментов в средних интервалах. Учтем, что каждый фрагмент имеет заголовок, поэтому эффективность использования канала будет равна

$$S_0 = \frac{T \cdot (k + 2) - t_0(k)}{L_{\text{period}}}, \quad (19)$$

где $t_0(k)$ – суммарная длина всех заголовков в этом случае.

Учтем наличие передач RТА-кадров. Найдем среднее количество сгенерированных за интервал обслуживания RТА-кадров ($\bar{n}$) как отношение длительности ин-

тервала к сумме средней задержки передачи RTA-кадра ($\bar{D}$) и среднего времени ожидания перед генерацией нового RTA-кадра:

$$\bar{n} = \frac{L_{\text{period}}}{\bar{D} + \frac{1}{\lambda}}. \quad (20)$$

Величину $\bar{D}$ можно найти, используя распределение задержки $F(t)$, найденное в (6):

$$\bar{D} = \int_0^{\infty} x dF(x). \quad (21)$$

Если RTA-кадр сгенерирован во время передачи точки доступа длительностью L_{ext} , то он будет передан с использованием механизма вытеснения. В противном случае, при генерации во время интервала длительностью

$$\text{AIFS}_{\text{AP}} + \bar{b} \cdot \sigma$$

(т.е. $L_{\text{period}} - L_{\text{ext}}$), когда канал свободен, кадр будет передан без использования механизма вытеснения. Для достижения высокой пропускной способности точке доступа целесообразно устанавливать как можно большую длительность передачи. При этом будет выполняться

$$L_{\text{ext}} \gg \text{AIFS}_{\text{AP}} + \bar{b} \cdot \sigma,$$

и большинство RTA-кадров будут передаваться с использованием механизма вытеснения, поэтому оценим эффективность использования канала следующим образом:

$$S \approx S_0 \cdot \left(1 - \frac{\bar{n} \cdot (T_r + \text{PIFS})}{L_{\text{period}}}\right) = S_0 \cdot \left(1 - \frac{T_r + \text{PIFS}}{\frac{1}{\lambda} + \bar{D}}\right), \quad (22)$$

где эффективность S_0 в случае передачи точки доступа без RTA-станции умножается на масштабирующий множитель, равный доле времени, в течение которого точка доступа передает свои кадры. Эта доля времени оценивается с учетом того, что за интервал обслуживания L_{period} RTA-станция генерирует в среднем $\bar{n}$ кадров, для передачи каждого из которых отнимает у точки доступа $T_r + \text{PIFS}$ канального времени.

§ 5. Численные результаты

В данном параграфе приводятся численные результаты исследования эффективности механизма вытеснения. Сначала в п. 5.1 показаны результаты оценки точности модели. Далее в п. 5.2 показано влияние параметра T механизма вытеснения на значение Q -квантиля задержки и эффективности использования канала. В п. 5.3 найдены оптимальные параметры механизма вытеснения.

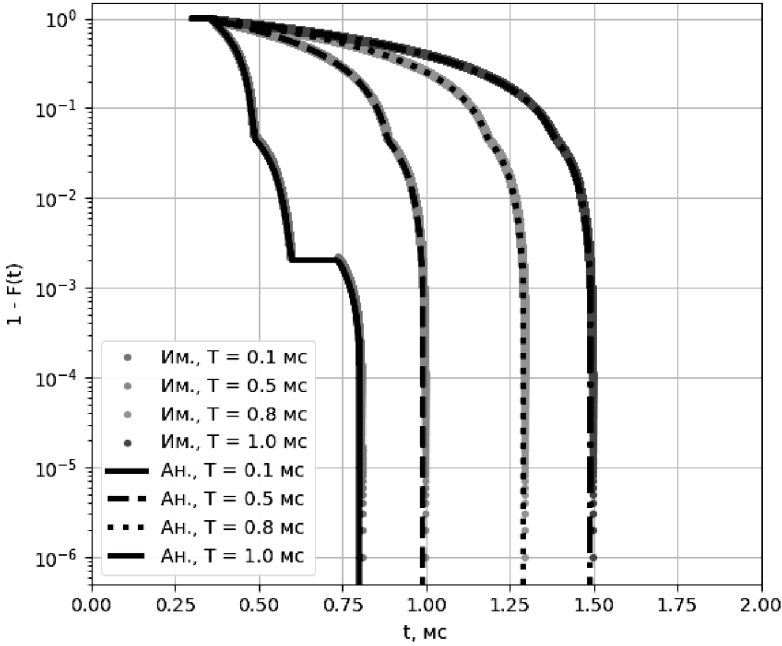
В табл. 1 приведены значения временных интервалов и параметров EDCA, используемых при получении численных результатов.

5.1. Оценка точности модели. Для оценки точности аналитической модели была разработана имитационная модель. В отличие от аналитической модели имитационная модель не использует предположения о стационарности вероятности передачи точки доступа, а подробно моделирует процесс отсчета отсрочки. Также в имитационной модели не предполагается, что процесс передачи RTA-станции не влияет на передачи точки доступа, и учитываются изменения конкурентного окна точки

Таблица 1

Значения временных интервалов и параметров EDCA

RTS = 44 мкс	PIFS = SIFS + σ	$W_{\min}^{\text{RTA}} = 4$	$L = 4$ мс
CTS = 44 мкс	AIFS _{RTA} = SIFS + 2σ	$W_{\max}^{\text{RTA}} = 8$	DATA = 300 мкс
SIFS = 16 мкс	AIFS _{AP} = SIFS + 10σ	$W_{\min}^{\text{AP}} = 16$	
$\sigma = 9$ мкс	ACKTIMEOUT = 45 мкс	$W_{\max}^{\text{AP}} = 1024$	

Рис. 1. Дополненная функция распределения задержки, $\lambda = 50 \text{ с}^{-1}$

доступа при коллизиях. На рис. 1 и 2 представлены графики зависимости дополненной функции распределения задержки от времени и эффективности использования канала от интенсивности RTA-трафика. Как видно из графиков, результаты аналитической модели (кривые “Ан.”) отличаются от результатов имитационной (кривые “Им.”) не более чем на 2% для дополненной функции распределения задержки, и не более чем на 1% для эффективности использования канала при малой интенсивности RTA-трафика. При большей интенсивности RTA-трафика значение эффективности для математической и имитационной моделей незначительно возрастает из-за роста вероятностей возникновения коллизии с точкой доступа и передачи RTA-кадра без использования механизма вытеснения, однако оно по-прежнему меньше 1%. Разработанная аналитическая модель позволяет с высокой точностью оценивать квантили задержки порядка 0,99999, поэтому ее можно использовать для выбора параметров механизма вытеснения для RTA-трафика. На графиках дополненной функции распределения задержки от времени видны особые точки, в которых меняется характер зависимости. По абсциссе они соответствуют максимальной задержке доставки RTA-кадра, сгенерированного в пустом слоте. Большие задержки возможны при передаче RTA-кадра с вытеснением или при коллизии. На рис. 1 для $T = 0,1$ мс виден горизонтальный участок, наличие которого обусловлено тем, что для такой длительности фрагмента максимальная задержка при передаче с вы-

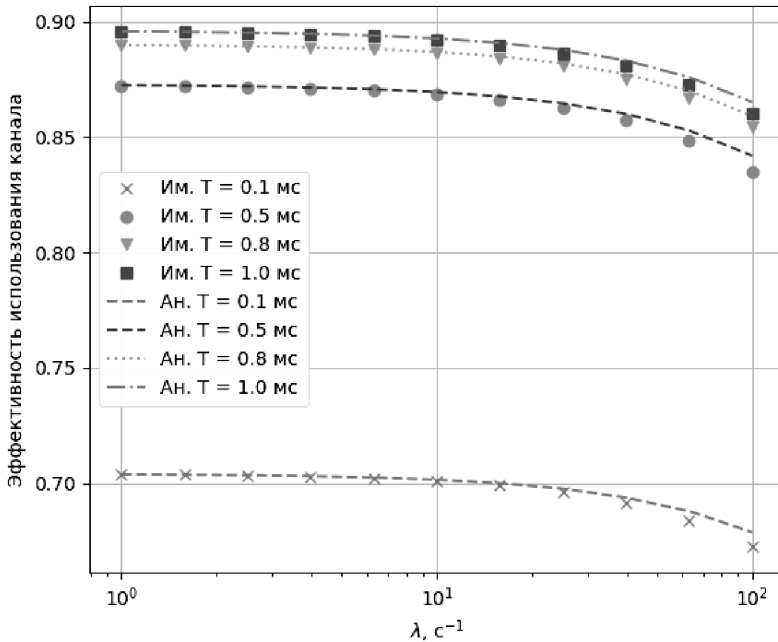


Рис. 2. Зависимость эффективности использования канала от интенсивности трафика

теснением оказывается меньше, чем минимальная задержка в случае коллизии. Таким образом, при малых значениях T квантиль задержки порядка 0,999–0,99999 определяется вероятностью коллизии RТА-станции с точкой доступа, а при больших T – вероятностью генерации RТА-кадра в первом интервале. Из зависимости эффективности использования канала от интенсивности видно, что эффективность использования канала падает с увеличением λ и растет с увеличением T .

5.2. Влияние параметров механизма вытеснения. Рассмотрим, как выбор параметров механизма вытеснения влияет на значения 0,99999-квантиля задержки и эффективности использования канала. Значение 0,99999-квантиля задержки практически не изменяется от интенсивности трафика и совпадает с показанным на рис. 1. На рис. 2 показана зависимость эффективности использования канала от интенсивности трафика. Эффективность возрастает с ростом T , что объясняется снижением накладных расходов на передачу. Таким образом, для максимизации эффективности использования канала точкой доступа и выполнения ограничения на квантиль задержки доставки RТА-кадров нужно установить T наибольшим, но так, чтобы при нем выполнялось ограничение на квантиль задержки.

5.3. Выбор параметра T . Рассмотрим, как выбирать значение параметра T . На рис. 3 представлены зависимости Q -квантиля задержки от длительности фрагмента T . Данные зависимости показывают, что при значениях T , превышающих некоторое значение T^* , квантили задержки близки к максимальной задержке при передаче с использованием механизма вытеснения RТА-кадра, сгенерированного в первом интервале, равной

$$D_{\text{first}}^{\text{max}} = RTS + \text{SIFS} + \text{CTS} + \text{SIFS} + T + \text{SIFS} + T_r + \text{AIFS}_{\text{RTA}}, \quad (23)$$

причем квантили порядка 99,9–99,99% практически совпадают с данной величиной, а квантиль 99% ниже максимальной задержки примерно на 0,05 мс.

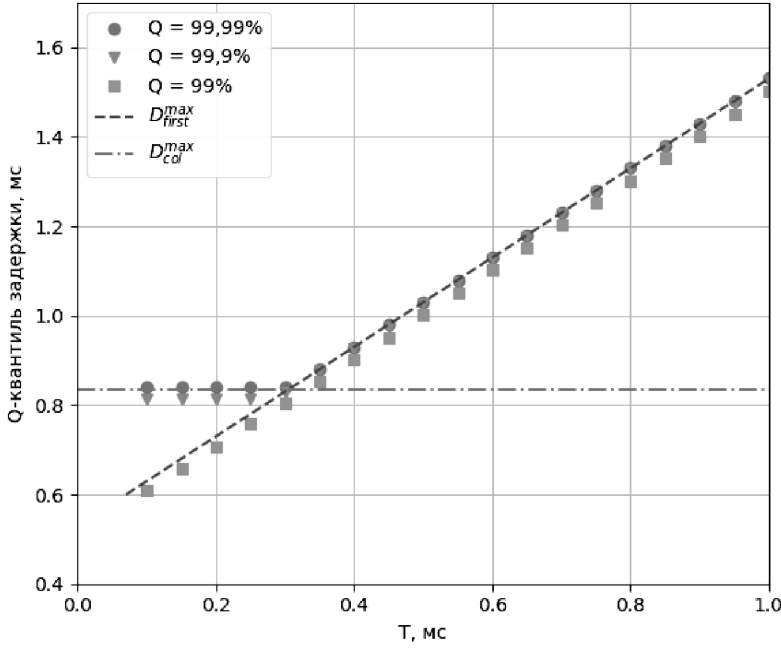


Рис. 3. Зависимость Q -квантилей от длительности фрагмента

При $T < T^*$ квантиль задержки порядка 99% также близок к $D_{\text{first}}^{\text{max}}$, а квантили задержки порядка 0,999–0,9999 определяются максимальной задержкой, соответствующей случаю коллизии кадров RTA-станции и точки доступа, равной

$$D_{\text{col}}^{\text{max}} = \sigma + T_c + (W_1^{\text{RTA}} - 1) \cdot \sigma + T_r + \text{AIFS}_{\text{RTA}}. \quad (24)$$

Длину фрагмента T^* , начиная с которого меняется зависимость, можно найти из уравнения

$$D_{\text{first}}^{\text{max}} = D_{\text{col}}^{\text{max}}.$$

В рассматриваемом случае $T^* \approx 0,3$ мс.

Таким образом, при $D^* > D_{\text{col}}^{\text{max}}$ и $Q \geq 99\%$ параметр T можно выбирать, исходя из уравнения

$$D_{\text{first}}^{\text{max}}(T) = D^*.$$

В противном случае, если $D^* < D_{\text{col}}^{\text{max}}$, нужно сравнить $1 - Q$ с вероятностью коллизии передач RTA-станции и точки доступа. Если

$$p_{\text{idle}} \cdot \tau < 1 - Q,$$

параметр T можно выбрать исходя из того же уравнения. Если

$$p_{\text{idle}} \cdot \tau \geq 1 - Q,$$

то квантиль задержки близок к значению $D_{\text{col}}^{\text{max}}$, и меньшую задержку обеспечить нельзя.

§ 6. Заключение

В данной статье исследован механизм вытеснения, который планируется использовать для обслуживания приложений реального времени (RTA) в сетях Wi-Fi 8. При его использовании передачи неприоритетных данных разделяются на фрагменты, между которыми предусмотрены интервалы времени, когда можно передать приоритетные данные, даже если доступ к каналу был получен другим устройством.

В статье представлена аналитическая модель передачи данных при использовании метода доступа к каналу с вытеснением. Даны рекомендации для выбора параметров метода доступа для доставки приоритетных данных с требуемым квантилем задержки доставки RTA-кадров и максимальной эффективностью использования канала точкой доступа.

В расширенной версии данной статьи [10] также проведено сравнение метода доступа с вытеснением и стандартного метода доступа к каналу EDCA и показано, что при оптимальных параметрах обоих методов метод доступа с вытеснением обеспечивает существенно большую пропускную способность сети.

В дальнейшем планируется расширить разработанную модель для сети с неидеальным каналом и с большим количеством RTA-станций.

СПИСОК ЛИТЕРАТУРЫ

1. Карамышев А.Ю., Порай Е.Д., Хоров Е.М. Оценка емкости системы сверхнадежной связи с низкими задержками с помощью аппроксимаций для многосерверных систем массового обслуживания $G/G/s$ // Пробл. передачи информ. 2024. Т. 60. № 2. С. 36–52. <https://doi.org/10.31857/S0555292324020049>
2. Fang J., Akhmetov D., Park M., Cariou L., Stacey R. Preemption for Low Latency Application. IEEE 802.11-23/0092r0. Mar. 13, 2023. <https://mentor.ieee.org/802.11/dcn/23/11-23-0092-00-0uhr-preemption.pptx>.
3. Ruy K., Chu L., Wang H., Cao R., Zhang H. Low Latency Support in UHR. IEEE 802.11-23/0018r1. Feb. 5, 2023. <https://mentor.ieee.org/802.11/dcn/23/11-23-0018-01-0uhr-low-latency-support-in-uhr.pptx>.
4. Moon J., Kim R. Y. Preemptive Channel Access Scheme for Next Generation Wi-Fi // Proc. 2024 IEEE Int. Conf. on Big Data and Smart Computing (IEEE BigComp 2024). Bangkok, Thailand. Feb. 18–21, 2024. P. 131–135. <https://doi.org/10.1109/BigComp60711.2024.00029>
5. Bankov D., Chemrov K., Khorov E. Tuning Channel Access to Enable Real-Time Applications in Wi-Fi 7 // 12th Int. Congr. on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT 2020). Brno, Czech Republic. Oct. 5–7, 2020. P. 20–25. <https://doi.org/10.1109/ICUMT51630.2020.9222409>
6. Листциндер Б.Я., Привалов А.Ю. Обобщение формул для моментов очереди при неординарном пуассоновском потоке для очередей пакетов в системах телекоммуникаций // Пробл. передачи информ. 2023. Т. 59. № 4. С. 32–37. <https://doi.org/10.31857/S0555292323040046>
7. Угловский А.Ю., Мельников И.А., Алексеев И.А., Куреев А.А. Оценка низкого уровня ошибок с помощью выборки по значимости с равномерным распределением // Пробл. передачи информ. 2023. Т. 59. № 4. С. 3–12. <https://doi.org/10.31857/S0555292323040010>
8. Bankov D.V., Khorov E.M., Lyakhov A.I., Sandal M.L. Approach to Real-Time Communications in Wi-Fi Networks // J. Commun. Technol. Electron. 2019. V. 64. P. 880–889. <https://doi.org/10.1134/S1064226919080205>
9. Vishnevsky V., Lyakhov A. 802.11 LANs: Saturation Throughput in the Presence of Noise // Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications. Proc. 2nd Int. IFIP-TC6 Network-

ing Conf. (NETWORKING 2002). Pisa, Italy. May 19–24, 2002. Lect. Notes Comp. Sci. V. 2345. Berlin: Springer, 2002. P. 1008–1019. https://doi.org/10.1007/3-540-47906-6_82

10. *Ritterman A.V., Bankov D.V., Lyakhov A.I., Khorov E.M.* Modeling of Channel Access with Preemption in Wi-Fi 8 Networks // Probl. Inf. Transm. 2024. V. 60. № 4 (to appear). <https://doi.org/10.1134/S0032946024040045>

Ритерман Алиса Вадимовна

Банков Дмитрий Викторович

Ляхов Андрей Игоревич

Хоров Евгений Михайлович

Институт проблем передачи информации

им. А.А. Харкевича Российской академии наук, Москва

riterman@wireless.iitp.ru

bankov@iitp.ru

lyakhov@iitp.ru

khorov@wireless.iitp.ru

Поступила в редакцию

19.11.2024

После доработки

19.11.2024

Принята к публикации

10.12.2024