

Об одном способе передачи данных реального времени в сетях Wi-Fi¹

Д.В. Банков^{*,**}, Е.М. Хоров^{*,**}, А.И. Ляхов^{*}, М.Л. Сандал^{*}

** Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия*

*** Национальный исследовательский университет «Высшая школа экономики», Москва, Россия*

Поступила в редколлегию 30.11.2018

Аннотация—В развитии технологий беспроводной связи наблюдается повышенный интерес к сценариям, в которых накладываются строгие ограничения на надёжность и на время доставки данных. Данные сценарии включают в себя приложения реального времени, такие как автоматизация производства, удалённое управление, трансляция видеопотоков и виртуальная реальность. Предъявляемые требования к качеству обслуживания крайне сложно удовлетворить в рамках распространённых на данный момент технологий передачи данных. В частности, в сетях Wi-Fi в данный момент невозможно гарантировать низкую задержку из-за особенностей метода доступа к каналу. В данной работе исследуется способ обеспечивать низкую задержку и высокую надёжность связи в сетях Wi-Fi, основанный на использовании дополнительного радиоинтерфейса. Данный подход исследуется при помощи математической модели гетерогенной сети, в которой устройства генерируют приоритетные и неприоритетные пакеты данных. Результаты исследования показывают, что при соблюдении определённых ограничений на интенсивность приоритетного трафика, данный подход позволяет удовлетворить требования приложений реального времени. При этом пропускная способность, доступная для передачи неприоритетных данных уменьшается незначительно.

КЛЮЧЕВЫЕ СЛОВА: беспроводная сеть, сетевые протоколы, надёжность, низкая задержка, приложения реального времени.

1. ВВЕДЕНИЕ

Приложения реального времени (англ. Real-Time Applications, RTA), такие как автоматизация производства, удалённое управление, трансляция видеопотоков, виртуальная и дополненная реальность, приобретают всё большую популярность в качестве потенциального сценария использования сетей Wi-Fi нового поколения. Данная тенденция подтверждается, в частности, тем, что рабочая группа комитета по стандартизации IEEE 802, посвящённая разработке стандартов сетей Wi-Fi, создала тематическую группу RTA для того, чтобы определить наиболее значимые приложения реального времени, определить характерные для них требования на качество обслуживания, например, на задержку при передаче и на вероятность доставки данных, а также для того, чтобы предложить решения, которые позволят поддерживать передачу данных в реальном времени в сетях Wi-Fi [1]. Приложения реального времени представляют собой сложную задачу для сетей Wi-Fi в силу высоких требований к задержке (не более 1 мс в сценариях управления производством) и к надёжности передачи данных (вероятность потери пакета не должна превышать 10^{-9} для аварийных систем).

¹ Исследование выполнено в МИЭМ НИУ ВШЭ за счет гранта Российского научного фонда (проект №18-19-00580)

Похожая тенденция имеется и в сотовых сетях, в которых сверхнадежная связь с низкой задержкой (англ. Ultra-Reliable Low-Latency Communications, URLLC) рассматривается в качестве одной из важнейших составляющих беспроводных сетей пятого поколения. В частности, в URLLC предъявляются ограничения на вероятность потери пакета порядка 10^{-5} и низкой задержкой (менее 1 мс) [2]. Данные ограничения послужили триггером для инноваций в разработке сотовых сетей, например, для разработки нового радиointерфейса (англ. New Radio air interface) для сотовой сети, что существенно снизило гранулярность передач, поскольку текущая гранулярность планирования LTE в 1 мс слишком велика, чтобы удовлетворить нормам URLLC на задержку [3, 4].

В сетях Wi-Fi удовлетворение требований к качеству обслуживания RTA затруднено тем, что станции Wi-Fi используют нелицензируемый диапазон частот для передачи данных. При передаче данных в нелицензируемом диапазоне частот устройства должны перед передачей прослушивать канал и, в случае, если в данный момент ведётся передача, должны дождаться момента, когда канал освободится, а затем использовать метод случайного доступа. При этом передача чужого кадра и соответствующее время ожидания в сетях Wi-Fi может занять значительное время (до 5 мс). Следовательно, при существующих стандартах Wi-Fi удовлетворить требованиям RTA практически невозможно.

В данной работе исследуется подход к обслуживанию RTA в сети Wi-Fi, предложенный в ноябре 2017 года на пленарном заседании IEEE 802 [5]. Данный подход основан на использовании вторичного радиointерфейса, используемого для передачи служебных сигналов о занятости канала. Эффективность предложенного подхода оценивается при помощи разработанной математической модели передачи данных в гетерогенной сети с двумя типами станций: передающими пакеты RTA и передающими обычные пакеты. Эта модель позволяет вычислить такие важные показатели как задержка пакетов RTA, вероятность того, что пакет RTA не будет доставлен в пределах заданного ограничения, и пропускная способность обычных станций.

Дальнейшее изложение построено следующим образом. В Разделе 2 описан метод доступа к каналу, используемый в сетях Wi-Fi. Далее в Разделе 3 описана модификация метода доступа, позволяющая предоставлять приоритетный доступ к каналу для потоков RTA. В Разделе 4 приведён обзор литературы, содержащей решения для обслуживания трафика с низкой задержкой и высокой надёжностью. Раздел 5 содержит описание сценариев моделирования и формальную постановку задачи. В Разделе 6 строится математическая модель процесса передачи данных. В Разделе 7 приведены численные результаты. Раздел 8 содержит заключение.

2. ДОСТУП К КАНАЛУ В СЕТЯХ WI-FI

Основной метод доступа к каналу в сетях Wi-Fi, EDCA (англ. Enhanced Distributed Channel Access), является разновидностью метода множественного доступа с контролем несущей и избеганием коллизий с двоичной экспоненциальной отсрочкой. Приведём его краткое описание.

У каждой станции сети Wi-Fi есть очередь пакетов на передачу, куда поступают новые пакеты. В случае, когда пакет прибывает в пустую очередь, станция проверяет, занят ли канал в данное время. Если канал свободен, станция сразу передаёт новый пакет. В противном случае, станция взводит счетчик отсрочки, который выбирается равновероятно из интервала $[0, W_r - 1]$, где r — число неудавшихся попыток передачи текущего пакета (изначально равный 0), а W_r — конкурентное окно, которое определяется следующим образом:

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

где W_{min} и W_{max} означают нижнюю и верхнюю границы величины конкурентного окна соответственно.

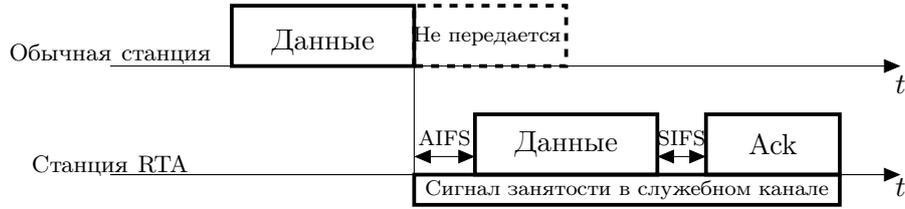


Рис. 1. Приоритетный доступ для пакетов RTA

Пока канал занят, станция замораживает свой счетчик отсрочки. В случае же, если канал свободен в течение заданного временного интервала AIFS (англ. Arbitration Inter-Frame Space, конфликтный межкадровый интервал), станция вновь активирует свой счетчик и уменьшает его на 1 каждый временной слот σ . Как только счетчик достигает нуля, станция делает попытку передать свой пакет. Получение этого пакета должно быть подтверждено адресатом посредством служебного кадра АСК по истечении интервала SIFS (англ. Short Inter-Frame Space, короткий межкадровый интервал) после получения. При своевременном получении подтверждения станция удаляет текущий пакет из очереди, сбрасывает в нуль счетчик r и переходит к следующему пакету. Если же отправляющая станция в течение ACKTimeout не получает подтверждения, она считает отправленный пакет потерянным, увеличивает свой счетчик r и пробует снова.

EDCA различает четыре типа категорий доступа (англ. Access Categories, ACs), каждая из которых имеет собственную очередь и свой счетчик отсрочки. При этом, W_{min} , W_{max} и AIFS для каждой очереди тоже различаются, что позволяет EDCA приоритезировать очереди. Следует отметить, что данный способ приоритезации трафика означает, что трафик заданной категории доступа получает в среднем больше ресурсов и имеет в среднем более низкую задержку, но это вовсе не означает, что более приоритетный трафик имеет исключительные права на доступ к каналу и всегда будет вытеснять менее приоритетный трафик.

3. ПРИОРИТЕТНАЯ СХЕМА ДОСТУПА

Идея предлагаемого метода состоит в добавлении новой категории доступа для трафика RTA и использовать добавочный радиointерфейс для передачи предупредительных служебных кадров. Для этой цели подойдет так называемый пробуждающий радиointерфейс (англ. Wake-Up Radio, WUR), работающий в узкой частотной полосе вне основного канала и обеспечивающий медленную, но надежную связь.

Когда очередь RTA не пуста, станция отправляет сигнал занятости канала (англ. busy tone) в служебном канале. По получении этого сигнала, все станции, передающие обычные (не RTA) пакеты тут же освобождают канал, чтобы дать RTA пакетам возможность быть переданными (см. рис.1.1). Далее станция RTA ждет AIFS и, если основной канал свободен, осуществляет попытку передачи. В противном случае станция соревнуется за канал с остальными станциями RTA в соответствии с правилами EDCA. Важно заметить, что первая передача осуществляется без взвода счетчика отсрочки, что аналогично передаче EDCA в случае, когда пакет приходит в пустую очередь. Ввиду того, что такая передача осуществляется асинхронно, она происходит с минимально возможной задержкой и с наименьшей вероятностью коллизии (по сравнению с передачей, сделанной по истечении счетчика отсрочки).

Пока служебный канал занят, обычные станции считают обычный канал занятым, то есть, их счетчики отсрочки заморожены. Если же обычная станция получает сигнал занятости канала, находясь в состоянии передачи, она считает, что для передаваемого ей кадра произошла коллизия.

4. ОБЗОР ЛИТЕРАТУРЫ

Задача обеспечения высокой надежности и малой задержки уже была рассмотрена в контексте технологии Ethernet, что привело к появлению решения в виде стандартов 802.1 TSN (англ. Time Sensitive Networking — сети, чувствительные к задержке) [6, 7]. Эти стандарты описывают методы обеспечения высокой надежности посредством повторения пакетов, контроля пути и резервирования потока, в то время как низкая задержка достигается путем временного планирования и механизма пресечения кадров. Проблема планирования решается на уровне сетевого коммутатора, который должен обслуживать пакеты из разных потоков в таком порядке, который гарантирует им определённую «наихудшую» задержку, что может быть интерпретировано как задача оптимизации [8]. Однако решения, предложенные для Ethernet, не решают проблемы, возникающие в беспроводных сетях, так как современные сети Ethernet состоят в основном из соединений типа «точка-точка» и в менее подвержены коллизиям кадров.

Для повышения надёжности передачи данных в сетях Ethernet также был предложен протокол IEC 62439-3 Parallel Redundancy Protocol (PRP) [9], основанный на дублировании пакетов и их одновременной отправке при помощи двух независимых интерфейсов. Данный протокол был адаптирован для сетей Wi-Fi [10], в которых в качестве двух интерфейсов использовались беспроводные интерфейсы, использующие различные частотные каналы. Такое решение повышает вероятность доставки данных, поскольку данные успешно доставляются до получателя в том случае, если пакет был передан хотя бы по одному интерфейсу. Однако данный подход к повышению надёжности имеет и свои недостатки, связанные с тем, что необходимо обеспечить взаимодействие между беспроводными интерфейсами одного устройства для того, чтобы избежать чрезмерного расходования канальных ресурсов в случае, например, когда один интерфейс пытается доставить кадр с данными тогда, когда этот кадр уже был успешно доставлен по другому интерфейсу. Данная проблема была решена путём использования «перекрёстных подтверждений» [11], которые позволяют одному интерфейсу уведомить другой о том, что заданный кадр был успешно доставлен и может быть удалён из очереди второго интерфейса. Следует отметить, что использование данного подхода хоть и позволяет повысить надёжность передачи данных, но не решает проблему приоритетного доступа к каналу, так как в случае, если каналы, используемые обоими интерфейсами, заняты передачей менее приоритетных данных, устройство должно ждать их освобождения, что может привести к тому, что задержка при передаче данных превысит заданное ограничение.

В работе [12] рассматривается проблема сосуществования устаревших станций и станций, использующих новые возможности физического уровня. Рассматривается сценарий, в котором сеть состоит из станций, поддерживающих новый метод доступа к каналу, введённый в дополнение к стандарту IEEE 802.11ax, и устаревших станций. Все они передают в общем канале и соревнуются друг с другом и с точкой доступа за доступ к каналу. Станции, поддерживающие стандарт IEEE 802.11ax имеют преимущество перед устаревшими станциями, поскольку помимо EDCA могут получать доступ к каналу по расписанию, составляемому точкой доступа. Авторы предлагают обеспечить честность между станциями, путем использования различных наборов параметров EDCA для разных типов станций. При этом для исследования эффективности предлагаемого решения строится математическая модель процесса передачи станциями разного типа. Данный подход к моделированию используется в текущей работе для того, чтобы описать процесс передачи обычных станций.

5. ПОСТАНОВКА ЗАДАЧИ

Предложенная схема позволяет значительно уменьшить задержку доставки пакетов RTA за счет уменьшения пропускной способности для обычных станций. В то же время, станции RTA все еще должны соревноваться за доступ к каналу между собой, а потому важно опреде-

лить интенсивность трафика RTA, при которой требования на задержку и долю потерянных пакетов (англ. Packet Loss Ratio, PLR) для станций RTA будут выполнены. Поэтому целью данной работы является оценка эффективности предложенной схемы доступа с точки зрения задержки доставки RTA пакетов и доступной полосы пропускания для обычных станций, а также поиск границы применимости данной схемы.

Для достижения данной цели рассматриваются два сценария: приоритизированный и равноправный. В обоих случаях рассматривается сеть, состоящая из M станций RTA и N обычных станций. При этом считается, что обычные станции работают в режиме насыщения, то есть у них всегда есть пакеты в очереди. Станции RTA, напротив, после успешной передачи пакета получают следующий через временной промежуток, подчиняющийся экспоненциальному распределению с параметром λ . Таким образом, вероятность того, что новый пакет придет до того, как пройдет время t , равна:

$$F(t) = 1 - \exp(-\lambda t). \quad (1)$$

Считается, что все станции находятся в зоне радиовидимости друг друга (т.е. нет проблемы скрытой станции).

В приоритизированном сценарии устройства используют предложенную схему, т.е. станции RTA имеют приоритет при доступе к каналу. В равноправном сценарии обычные станции не прослушивают служебный канал и, следовательно, не реагируют на служебные сигналы занятости канала. Таким образом, они соревнуются за доступ к каналу со станциями RTA.

Для обоих сценариев ставится задача разработки математической модели процесса передачи станций в сети, позволяющей при заданной интенсивности трафика RTA и при заданном количестве станций обоих типов найти задержку при передаче кадров RTA и пропускную способность, доступную для обычных станций. При помощи построенной модели требуется определить максимальную нагрузку на сеть, при которой задержка не превышает заданный порог 1 мс. Также требуется сравнить производительность сети в обоих сценариях и таким образом оценить эффективность предложенного подхода.

6. АНАЛИТИЧЕСКАЯ МОДЕЛЬ

Чтобы численно описать предложенный метод, ниже приводится математическая модель работы сети для обоих вышеупомянутых сценариев. Для того, чтобы описать поведение станций и найти задержку для станций RTA и полосу пропускания обычных станций, используется подход к моделированию сетей Wi-Fi, похожий на представленный в [13].

6.1. Приоритизированный сценарий, станции RTA.

Станции RTA имеют исключительные права доступа к каналу, вследствие чего их поведение не зависит от обычных станций. Выберем любую станцию RTA и будем наблюдать ее состояние в моменты начала передачи пакетов и в моменты отсчета её счетчика отсрочки. Интервал между такими событиями будем называть слотом. Когда очередь станции пуста, станция по определению пребывает в свободном (idle) состоянии. Когда станция получает пакет на передачу и при этом канал не занят, происходит асинхронная передача, что приводит к переходу станции в состояние ST (определенное таким образом). Если же канал оказывается занятым, станция взводит счетчик отсрочки и начинает его обратный отсчет. Во время этого отсчета ее состояние описывается парой чисел (b, s) , где b — текущее значение счётчика отсрочки, а s — текущая стадия отсрочки.

Чтобы описать вероятности перехода из одного состояния в другое, введём обозначения τ_s и τ_n — вероятности того, что в случайно выбранный слот станция совершает соответственно

асинхронную и синхронную передачу. Предполагается, что эти вероятности равны для всех станций и не меняются во времени. Таким образом, вероятность τ того, что станция RTA передает в случайно выбранный слот равна: $\tau = \tau_s + \tau_n$.

С точки зрения рассматриваемой станции, слоты могут быть пустыми, успешными или коллизионными. Пустым слот становится в случае, если ни одна из станций в его начале не начинает передачу, и вероятность этого

$$p_e = (1 - \tau)^{M-1}. \quad (2)$$

Слот успешен, если в его начале только одна станция начинает передачу. Таким образом, его вероятность:

$$p_s = (M - 1)\tau_n(1 - \tau_n)^{M-2} + (M - 1)\tau_s. \quad (3)$$

При этом делается предположение, что все асинхронные передачи успешны.

Вероятность коллизионного слота считается как

$$p_c = 1 - p_e - p_s. \quad (4)$$

Длительности пустого, успешного и коллизионного слота обозначены как σ , T_s^u и T_c^u , соответственно.

Переход станции из свободного состояния (idle) к отсчету счетчика происходит в случае, когда во время успешного или коллизионного слота ей приходит пакет. Вероятность этого события:

$$\tilde{p}_{idl} = p_s F(T_s^u) + p_c F(T_c^u) \quad (5)$$

Похожим образом можно определить вероятность перехода станции из состояния idle в состояние ST:

$$p_{st} = p_e F(\sigma). \quad (6)$$

Пусть теперь станция находится в состоянии ($s = i, b = k$). Когда b превышает 1, станция уменьшает значение счетчика. Когда b становится нулем, станция делает попытку передачи, которая заканчивается коллизией с вероятностью

$$p = 1 - (1 - \tau_n)^{M-1}, \quad (7)$$

где M — число станций RTA. Значит, с вероятностью $1 - p$ станция переходит в состояние idle, а с вероятностью p выбирает новую отсрочку, равномерно распределенную среди W_{s+1} возможных вариантов.

Пусть m — максимально возможная стадия счетчика, т.е. такая стадия, что $W_{max} = W_{min} \times 2^m$. Если, имея $b = m$, станция совершает неуспешную попытку передачи, b не увеличивается, а остается равным m . Если обозначить $Pr(state2|state1)$ — вероятность перехода из состояния $state1$ в состояние $state2$, то получатся следующие вероятности перехода марковской цепи:

$$\left\{ \begin{array}{l} Pr(i, k|i, k + 1) = 1; k \in [0, \dots, W_i - 2], i \in [0, \dots, m] \\ Pr(i, k|i - 1, 0) = p/W_i; k \in [0, \dots, W_i - 1], i \in [1, \dots, m] \\ Pr(m, k|m, 0) = p/W_m; k \in [0, \dots, W_m - 1] \\ Pr(idle|i, 0) = (1 - p); i \in [0, \dots, m] \\ Pr(0, k|Idle) = \tilde{p}_{idl}/W_0; k \in [0, \dots, W_0 - 1] \\ Pr(ST|Idle) = p_{st} \\ Pr(Idle|ST) = 1 \\ Pr(Idle|Idle) = 1 - \tilde{p}_{idl} - p_{st} \end{array} \right. \quad (8)$$

Пусть $q_{i,k}$ — стационарное распределение цепи. Заметив, что

$$q_{i,0} = pq_{i-1,0} = p^2q_{i-2,0} = \dots = p^i q_{0,0}, i \in [1, m-1] \quad (9)$$

$$q_{m,0} = p^m q_{0,0} / (1-p) \quad (10)$$

и что остальные вероятности могут бы выражены следующим образом:

$$q_{i,k} = \frac{W_i - k}{W_i} \begin{cases} p^i q_{0,0}; 0 \leq i \leq m-1, \\ \frac{p^i}{1-p} q_{0,0}; i = m, \end{cases} \quad (11)$$

$$q_{Idle} = \frac{q_{0,0}}{\tilde{p}_{idl}}, \quad (12)$$

$$q_{ST} = q_{Idle} p_{st}. \quad (13)$$

можно, используя условие нормировки

$$q_{ST} + q_{Idle} + \sum_{i=0}^m \sum_{k=0}^{W_i-1} q_{i,k} = 1, \quad (14)$$

найти $q_{0,0}$:

$$q_{0,0} = \left(\frac{1 + p_{st}}{\tilde{p}_{idl}} + \frac{1}{2} \left[W_0 \left(\sum_{i=0}^{m-1} (2p)^i + \frac{(2p)^m}{1-p} \right) + \frac{1}{1-p} \right] \right)^{-1} \quad (15)$$

Зная эту величину, можно найти вероятности передачи:

$$\tau_n = \sum_{i=0}^m q_{i,0} = \frac{q_{0,0}}{1-p}, \quad (16)$$

$$\tau_s = q_{Idle} \tilde{p}_{st} = \frac{q_{0,0} p_{st}}{\tilde{p}_{idl}}, \quad (17)$$

$$\tau = \tau_s + \tau_n = \tau_n \left(1 + \frac{(1-p) p_{st}}{\tilde{p}_{idl}} \right). \quad (18)$$

Получившуюся систему можно решить численными методами.

Найдём теперь среднюю задержку D при передаче кадров RTA. Она состоит из двух частей:

$$D = D_1 + D_2. \quad (19)$$

Здесь D_1 — средняя задержка для асинхронной передачи (равная минимальной задержке) и может быть приближенно подсчитана как $D_1 = p_{libre} T_s$, где p_{libre} — вероятность того, что в случайный момент времени канал пуст, т.е.

$$p_{libre} = \frac{p_e \sigma}{p_e \sigma + p_c T_c^u + p_s T_s^u}. \quad (20)$$

Вторая часть $D_2 = (1 - p_{libre}) \times D'$, где D' — средняя задержка синхронной передачи. Определив среднюю длину виртуального слота, в котором рассматриваемая станция не пытается провести передачу, как

$$\hat{\sigma} = p_e \sigma + p_s T_s^u + p_c T_c^u. \quad (21)$$

можно заметить, что D' состоит, во-первых, из времени, которое станции придется ждать освобождения канала, а во-вторых, из коллизий, в которые станция попадает при неуспешных

попытках передачи. Вероятность того, что станция попадет в i коллизий равна p^i и каждая коллизия длится T_c^u . При этом попытка номер i в среднем добавляет к задержке $\frac{\hat{\sigma}(W_i-1)}{2}$ — время, в течение которого станция отсчитывает счетчик. В конце концов, с вероятностью $1-p$ станция передает успешно за время T_{succ}^u . Суммарная средняя задержка:

$$\begin{aligned} D' &= \frac{T_s^u}{1 - e^{-\lambda T_s^u}} - \frac{1}{\lambda} + \sum_{i=0}^{\infty} p^i \left(\frac{\hat{\sigma}(W_i - 1)}{2} + (1-p) T_s^u \right) + \sum_{i=1}^{\infty} p^i T_c^u = \\ &= \frac{T_s^u}{1 - e^{-\lambda T_s^u}} - \frac{1}{\lambda} + T_s^u + \frac{T_c^u p - \frac{\hat{\sigma}}{2}}{1-p} + \frac{W_0 \hat{\sigma}}{2} \left[\frac{1 - (2p)^m}{1-2p} + \frac{(2p)^m}{1-p} \right] \end{aligned} \quad (22)$$

Найдя среднюю задержку, приступим к поиску функции распределения задержки для станций URLLC.

С вероятностью p_{libre} передача станции будет асинхронной, а значит, будет длиться T_s^u . С вероятностью $1 - p_{libre}$ пакет будет сгенерирован тогда, когда канал занят другой передачей. Станция дожидается конца слота и начнет отсчитывать счетчик отсрочки. Функция распределения этого времени ожидания может быть найдена как

$$F_{wait}(t) = \frac{\Pr(T_s^u - x < t | x < T_s^u)}{\Pr(x < T_s^u)} = \begin{cases} 0, & t < 0, \\ \frac{e^{\lambda(t-T_s^u)} - e^{-\lambda T_s^u}}{1 - e^{-\lambda T_s^u}}, & 0 \leq t \leq T_s^u, \\ 1, & t > T_s^u. \end{cases} \quad (23)$$

Пусть $d_i(e, s, c)$ — вероятность того, что в течение i -ой передачи станция простаивает e пустых, s успешных и c коллизионных слотов. Тогда

$$d_i(e, s, c) = \frac{1}{W_i} \frac{(e + s + c)!}{e!s!c!} p_e^e p_s^s p_c^c, \quad (24)$$

если $0 \leq e + s + c < W_i - 1$ и 0 в противном случае. Далее, пусть $w_i(e, s, c)$ - вероятность того, что во время попыток передач с 0 по i станция ожидает e , s и c соответственно пустых, успешных и коллизионных слотов. В таком случае

$$\begin{aligned} w_0(e, s, c) &= d_0(e, s, c), \\ w_i(e, s, c) &= \sum_{e_i, s_i, c_i} d_i(e_i, s_i, c_i) \times w_{i-1}(e - e_i, s - s_i, c - c_i), i > 0, \end{aligned} \quad (25)$$

а функция распределения задержки $F_D(t)$ записывается как

$$F_D(t) = \begin{cases} 0, & t < T_s^u, \\ p_{libre} + (1 - p_{libre}) \times F_{D'}(t - T_s^u), & t > T_s^u, \end{cases} \quad (26)$$

где $F_{D'}(t)$ — функция распределения времени, проведенного в состоянии синхронной передачи:

$$F_{D'}(t) = \sum_{i=0}^{\infty} p^i (1-p) \sum_{e, s, c} w_i(e, s, c) \times F_{wait}(t - \sigma e - s T_s^u - c T_c^u). \quad (27)$$

Вероятность потери пакета PLR выражается через распределение задержки следующим образом:

$$PLR = 1 - F_D(D_{max}), \quad (28)$$

где D_{max} — максимально допустимая задержка при передаче пакета, например, 1 мс.

6.2. Приоритезированный сценарий, обычные станции

Аналогично случаю со станциями RTA, рассмотрим случайную обычную станцию. Обозначим за τ_r вероятность того, что в начале слота обычная станция делает попытку передачи.

Вероятность того, что передача является неуспешной равна:

$$p_r = 1 - (1 - \tau_r)^{N-1} e^{-M\lambda T_s^r}, \quad (29)$$

где T_s^r — длительность успешной передачи для обычных станций. Здесь учтен тот факт, что, передача обычной станции становится неудачной (коллизией), если хотя бы одна станция RTA в это время передает пакет.

Поскольку обычные станции работают в режиме насыщения, τ_r может быть найдена более простым путем, а именно как среднее число попыток передачи на один кадр, деленное на среднее время, необходимое для успешной передачи кадра:

$$\begin{aligned} \tau_r &= \frac{\sum_{i=0}^{\infty} p_r^i}{\sum_{i=0}^m (\frac{W_i-1}{2} + 1) p_r^i + \sum_{i=m+1}^{\infty} (\frac{W_m-1}{2} + 1) p_r^i} = \\ &= \frac{2(1-2p_r)}{(1-2p_r)(W_0+1) + p_r W_0 (1-(2p_r)^m)}. \end{aligned} \quad (30)$$

Решив систему из уравнений 29 и 30, можно получить τ_r .

Чтобы найти пропускную способность обычных станций нужно учесть, что на их передачу в значительной степени влияют станции RTA: даже если слот обычной станции начался как пустой или успешный, он может быть прерван сигналом занятости канала и таким образом стать коллизией (с точки зрения рассматриваемой обычной станции).

Пусть p_{empty} — вероятность того, что ни одна из обычных станций не начинает передачу в выбранный слот. Тогда

$$p_{empty} = (1 - \tau_r)^N. \quad (31)$$

Как было упомянуто выше, пустой слот может быть прерван сигналом занятости канала, в результате чего канал будет занят до тех пор, пока станции RTA не передадут свои кадры. Средняя длина такого слота равна

$$\langle \sigma \rangle = \sigma (1 - F(\sigma))^M + D \left(1 - (1 - F(\sigma))^M \right), \quad (32)$$

где D — задержка станций RTA, о которой говорилось ранее.

Пусть p_{suc} — вероятность того, что в случайно выбранный слот лишь одна обычная станций совершает попытку передачи:

$$p_{suc} = N\tau_r (1 - \tau_r)^{N-1}. \quad (33)$$

Он закончится успешно лишь в том случае, если ни одна из станций RTA не начнет передачу до его окончания. Вероятность такого исхода:

$$p_{S,S} = p_{suc} (1 - F(T_s^r))^M. \quad (34)$$

Длина такого слота — T_s^r .

Если в течение слота приходит пакет RTA, потенциально успешный слот завершается коллизией. Вероятность этого равна

$$p_{S,C} = p_{suc} \left(1 - (1 - F(T_s^r))^M \right). \quad (35)$$

Средняя длительность такого слота может быть приближена выражением

$$\langle D_{S,C} \rangle = \frac{\int_0^{T_s^r} M\lambda t e^{-M\lambda t} dt}{1 - e^{-M\lambda T_s^r}} + D = \frac{1}{M\lambda} - \frac{T_s^r e^{-M\lambda T_s^r}}{1 - e^{-M\lambda T_s^r}} + D, \quad (36)$$

где первое слагаемое — среднее время, прошедшее до того, как сигнал занятости канала прервал передачу кадра.

Коллизии можно разделить на два типа. Обыкновенные коллизии (между пакетами обычных станций) делятся T_c^r и случаются с вероятностью

$$p_{C,R} = (1 - p_{empty} - p_{suc}) (1 - F(T_c^r))^M. \quad (37)$$

Как и успешные передачи, обыкновенные коллизии могут прерваны станциями RTA, что случается с вероятностью

$$p_{C,BT} = (1 - p_{empty} - p_{suc}) \left(1 - (1 - F(T_c^r))^M\right). \quad (38)$$

Средняя длина такой коллизии:

$$\langle D_{C,BT} \rangle = \frac{1}{M\lambda} - \frac{T_c^r e^{-M\lambda T_c^r}}{1 - e^{-M\lambda T_c^r}} + D, \quad (39)$$

что аналогично уравнению (36), только вместо T_s^r используется T_c^r . Таким образом, средняя длина коллизионного слота дается выражением

$$\langle T_{coll} \rangle = p_{S,C} \langle D_{S,C} \rangle + p_{C,BT} \langle D_{C,BT} \rangle + p_{C,R} T_c^r, \quad (40)$$

в то время как пропускная способность приближается формулой

$$S = \frac{p_{S,S} L}{\langle \sigma \rangle p_{empty} + T_s^r p_{S,S} + \langle T_{coll} \rangle}, \quad (41)$$

где L — размер кадра. Далее будет показано, что это приближение справедливо, когда плотность потока пакетов RTA низка.

6.3. Равноправный сценарий

В этом сценарии обычные станции не реагируют на сигналы занятости канала, что приводит к равноправной конкуренции за канал между ними и станциями RTA.

Как и раньше, p и τ описывают станции RTA, а p_r τ_r — обычные. Если выбрать случайную обычную станцию, то вероятность того, что случайный слот будет пустым с ее точки зрения при условии, что сама станция не передает, будет

$$p_{er} = (1 - \tau_r)^{N-1} (1 - \tau)^M. \quad (42)$$

Аналогичная вероятность с точки зрения случайной станции RTA:

$$p_{eu} = (1 - \tau_r)^N (1 - \tau)^{M-1}. \quad (43)$$

Аналогичным образом описываются остальные вероятности, описывающие ситуацию с точки зрения выбранной станции (индекс s для успешных слотов, c — для коллизий)

$$p_{su} = (M-1)\tau_n(1-\tau_n)^{M-2}(1-\tau_r)^N + (M-1)\tau_s + N\tau_r(1-\tau_r)^{N-1}(1-\tau_n)^{M-1}, \quad (44)$$

$$p_{sr} = (N-1)\tau_r(1-\tau_r)^{N-2}(1-\tau_n)^M + M\tau_n(1-\tau_n)^{M-1}(1-\tau_r)^{N-1} + M\tau_s, \quad (45)$$

$$p_{cu} = 1 - p_{su} - p_{eu}, \quad (46)$$

$$p_{cr} = 1 - p_{sr} - p_{er}. \quad (47)$$

Коллизии делятся на две группы: с участием обычных станций и без них. С точки зрения станции РТА, вероятность коллизии без участия обычных станций:

$$p_{cnu} = (1-\tau_r)^N(1-(1-\tau)^{M-1} - (M-1)\tau_n(1-\tau_n)^{M-2} - (M-1)\tau_s), \quad (48)$$

а вероятность коллизии с участием как минимум одной обычной станции:

$$p_{cru} = 1 - (1-\tau_r)^N - N\tau_r(1-\tau_r)^{N-1}(1-\tau)^{M-1}. \quad (49)$$

Полная вероятность коллизии вычисляется как $p_{cu} = p_{cnu} + p_{cru}$. Похожим образом можно записать $p_{su} = p_{snu} + p_{sru}$, где p_{snu} и p_{sru} — вероятности успешной передачи для РТА и обычных станций соответственно. В этих обозначениях нужно переопределить \tilde{p}_{idl} и p_{st} :

$$p_{st} = p_{eu}F(\sigma), \quad (50)$$

$$\tilde{p}_{idl} = p_{snu}F(T_s^u) + p_{sru}F(T_s^r) + p_{cnu}F(T_c^u) + p_{cru}F(T_c^r), \quad (51)$$

а также вероятность того, что синхронная передача станции РТА станет коллизией:

$$p = 1 - (1-\tau_r)^N(1-\tau_n)^{M-1}, \quad (52)$$

Та же вероятность для обычных станций:

$$p_r = 1 - (1-\tau_r)^{N-1}(1-\tau_n)^M. \quad (53)$$

Таким образом получается та же цепь Маркова для станций РТА, что была описана системой (8).

Учитывая сделанные переопределения, уравнения для τ из предыдущего раздела, до сих пор справедливы. Однако, чтобы найти его численное значение, придется решить систему уравнений (16), (30), (52) и (53), так как вероятности попасть в коллизию зависят как от τ_n , так и от τ_r .

Зная эти вероятности, можно найти среднюю задержку для станций РТА и пропускную способность для обычных методами, аналогичными тем, что были применены в приоритезированном сценарии. Так, пропускная способность будет определена формулой

$$S = \frac{p_{sru}L}{\langle\sigma\rangle p_{empty} + T_c^u p_{su} + T_s^r p_{sr} + T_c^r p_{cr} + T_c^u p_{cu}} = p_{sru}L/\langle\tilde{\sigma}\rangle, \quad (54)$$

где $p_{empty} = (1-\tau)^M(1-\tau_r)^N$, $\langle\tilde{\sigma}\rangle$ — средняя длина слота, а средняя задержка станций РТА находится как

$$\langle D \rangle = p_{libre}T_s^u + (1-p_{libre})[\langle\tilde{\sigma}\rangle\frac{W_0-1}{2} + (1-p)T_s^u + p(T_c^u + \langle\tilde{\sigma}\rangle\frac{W_1-1}{2} + (1-p)T_s^u + p(\dots)] \quad (55)$$

где $p_{libre} = p_{empty}\sigma/\langle\tilde{\sigma}\rangle$. Последняя формула имеет простую вероятностную интерпретацию: либо пакет передается сразу, асинхронно, либо начинается отсчет счетчика в диапазоне $[0, W_0-1]$. Далее с вероятностью $(1-p)$ пакет успешно передается, а с вероятностью p случается коллизия, взводится счетчик уже в большем диапазоне, и т.д.

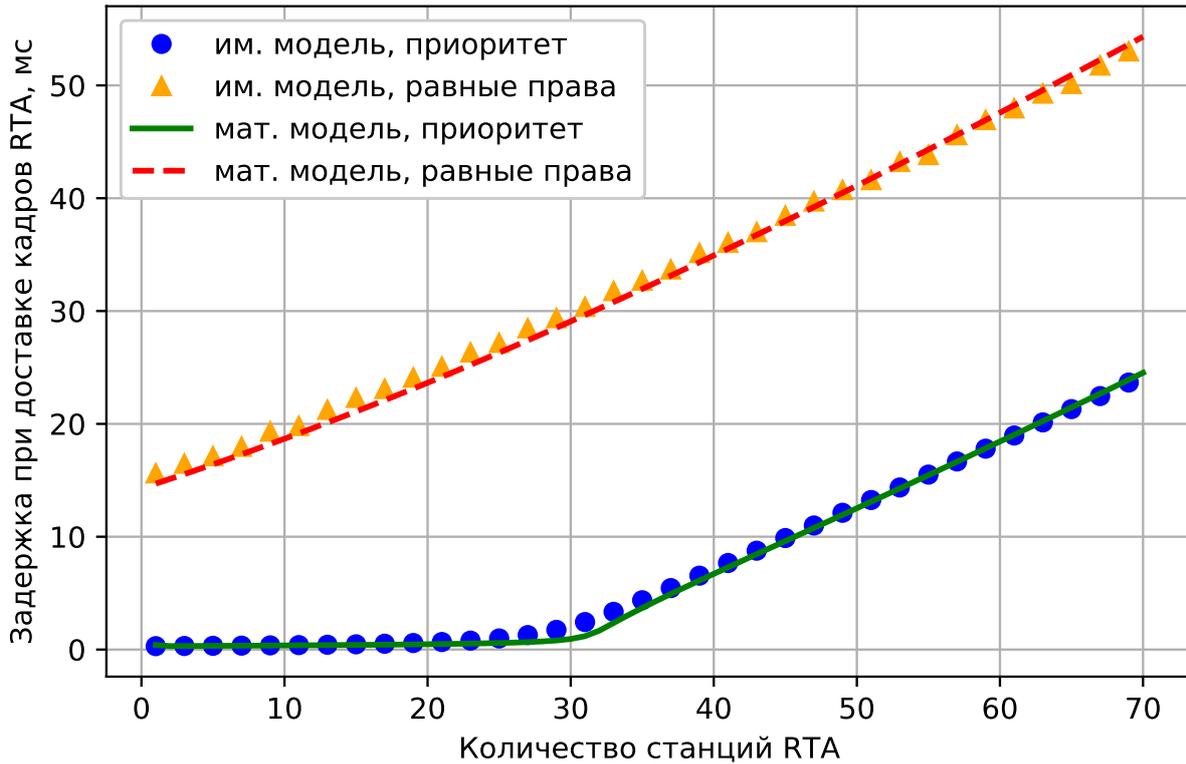


Рис. 2. Зависимость задержки пакетов RTA от числа станций RTA. Интенсивность поступления пакетов: $\lambda = 100c^{-1}$.

7. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Далее приведены результаты имитационного моделирования и их сравнение с результатами, полученными при помощи аналитической модели.

При имитационном моделировании станции передавали данные в канале шириной 20 МГц на сигнально-кодовой конструкции MCS 5, обеспечивающей скорость передачи данных 25 Мбит в секунду. Станции RTA передавали кадры размером 200 байт, а обычные станции — размером 1 Кбайт. Станции находились на таком расстоянии от точки доступа, что заданная сигнально-кодовая конструкция обеспечивала надёжную передачу данных, то есть передачи могли быть неуспешными только из-за коллизий кадров от разных станций.

На рис. 2 представлена зависимость средней задержки пакетов RTA от числа станций RTA для обоих сценариев.

Как видно, наделение станций RTA исключительными правами доступа к каналу значительно (более, чем на порядок) уменьшает время доставки при условии, что интенсивность потока пакетов RTA относительно низка. Так же можно наблюдать резкий рост задержки в равноправном сценарии, начиная с некоторого количества станций. Это можно объяснить тем, что при определенном количестве станций RTA, начинается постоянная конкуренция за канал даже между ними, поскольку оказывается, что всегда есть станция, желающая передавать. В равноправном сценарии такого скачка нет, так как на протяжении всей кривой станции RTA передают, преимущественно соревнуясь за канал с обычными станциями, то есть станции *сразу* находятся в условиях конкуренции.

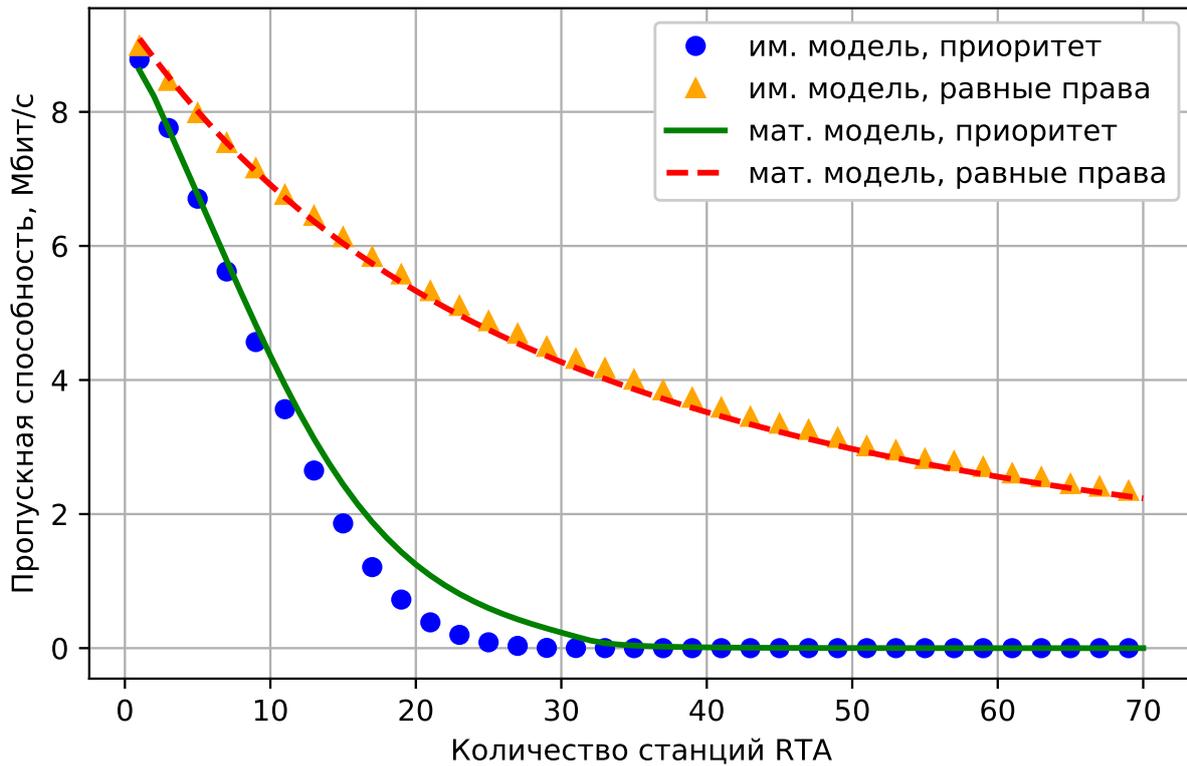


Рис. 3. Зависимость пропускной способности обычных станций от числа станций RTA.

На рис. 3 представлена зависимость пропускной способности обычных станций от числа станций RTA. В приоритизированном сценарии на определенном этапе станции RTA занимают своими передачами все время, из-за чего пропускная способность обычных станций обращается в нуль. В равноправном сценарии такого не происходит, пропускная способность плавно убывает с добавлением новых претендентов на канал.

На рис. 4 представлена зависимость вероятности потери пакета RTA от интенсивности срочного трафика для случая, когда в сети присутствует 50 станций RTA, и пакет считается потерянным, если он не доставлен в течение 1 мс с момента появления. В первую очередь следует отметить, что в рассматриваемом сценарии в случае, когда не используется схема приоритетного доступа к каналу, практически все пакеты RTA не могут быть доставлены в пределах 1 мс, что связано с тем, что станциям RTA почти всегда приходится ждать, пока обычные станции освободят канал, и время ожидания может быть очень велико. В случае, когда у станций RTA есть приоритет при доступе к каналу, при низкой интенсивности трафика даже 50 станций могут передавать свои пакеты в пределах ограничения на задержку в 1 мс. При дальнейшем увеличении нагрузки, станции RTA начинают конкурировать за доступ к каналу, что приводит к задержкам, связанным с ожиданием освобождения канала, коллизиям при передаче кадров и возможным повторным передачам.

Анализируя получившиеся результаты, можно заключить, что при малой интенсивности трафика RTA предложенный метод очень эффективен: незначительно понижая пропускную способность обычных станций, он позволяет станциям RTA передавать данные гораздо быстрее. Тем не менее, данный метод имеет свои ограничения, связанные с тем, что при росте интенсивности трафика RTA возрастает конкуренция за доступ к каналу между станциями RTA,

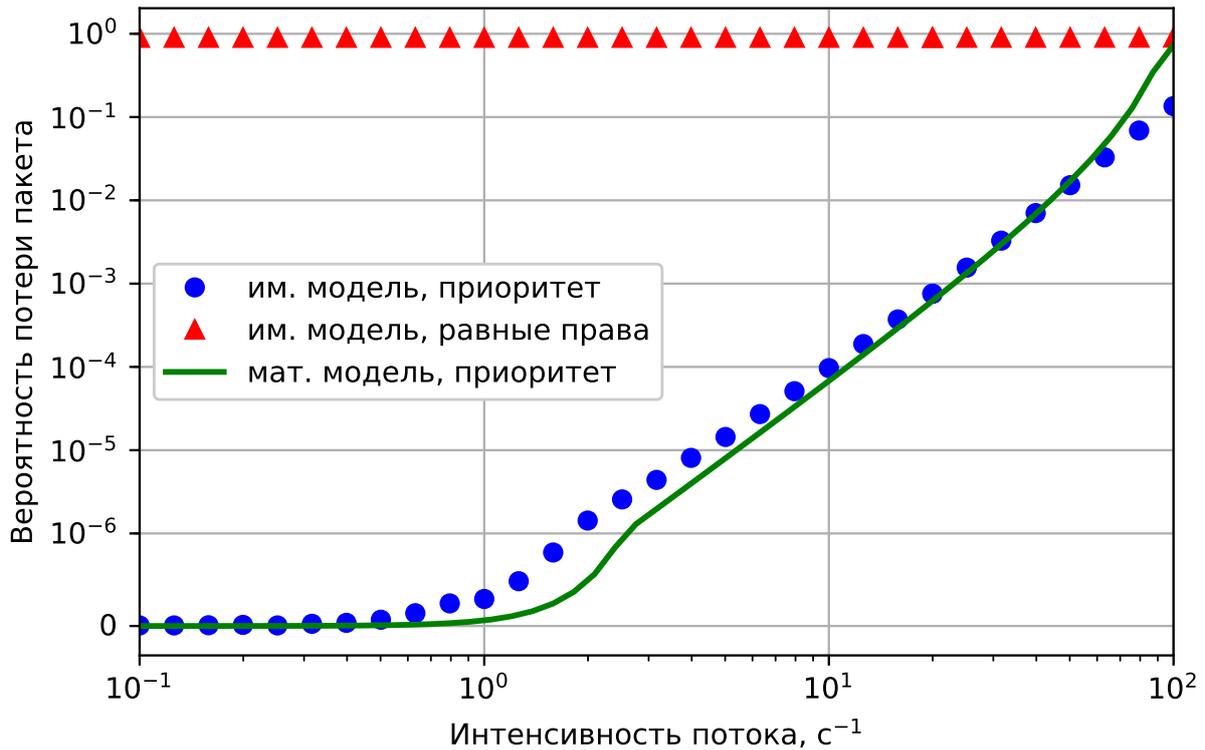


Рис. 4. Зависимость вероятности потери пакета от интенсивности трафика RTA, в сети 50 станций RTA.

из-за чего выигрыш во времени доставки пакетов RTA перестаёт быть настолько значимым, и требования на качество обслуживания срочного трафика могут не выполняться. Также следует отметить, что применение данной схемы может привести к полному прекращению передач обычных станций. Таким образом, целесообразность применения описанного метода зависит от параметров сети, а с помощью предложенной в работе математической модели можно найти границы применимости метода.

8. ЗАКЛЮЧЕНИЕ

В данной работе представлено исследование метода обслуживания приложений реального времени (RTA) в сетях Wi-Fi. Данный метод основан на использовании дополнительного надёжного и низкоскоростного радиointерфейса для того, чтобы уведомлять устройства в сети о необходимости передачи срочных кадров и таким образом организовывать приоритетный доступ к каналу для передающих устройств.

Описанная схема позволяет доставлять срочные пакеты с очень малой задержкой, что невозможно в стандартных сетях Wi-Fi вследствие того, что, согласно действующим протоколам, устройства не могут передавать данные, если канал уже занят другими станциями. Полученные результаты показывают, что данный метод позволяет значительно сократить время доставки пакетов, и при низкой интенсивности трафика RTA позволяет гарантировать сверхнизкую задержку (менее 1 мс) при высокой надёжности передачи (0.99999%), и при этом потери пропускной способности для обычного трафика являются незначительными. Предложенный метод доступа имеет ограничения, в рамках которых он может удовлетворить требования к качеству обслуживания срочного трафика. Данные ограничения зависят от параметров трафика

и количества устройств в сети и могут быть найдены при помощи представленной в работе математической модели процесса передачи данных с использованием метода приоритетного доступа.

СПИСОК ЛИТЕРАТУРЫ

1. RTA Report Discussion. — Access mode: <https://mentor.ieee.org/802.11/dcn/18/11-18-1690-00-Orta-rta-report-discussion.pptx>.
2. Framework and overall objectives of the future development of IMT for 2020 and beyond : Recommendation : M.2083 / ITU-R : 2015. — September.
3. J. Vihriälä et al. Numerology and frame structure for 5G radio access // 2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). — 2016. — Sept. — P. 1–5.
4. NR; Overall description; Stage-2 : Rep. : TS 38.300 / 3GPP : 2017. — December.
5. Wi-Fi Time Sensitive Networking. — Access mode: <https://mentor.ieee.org/802.11/dcn/17/11-17-1734-01-0wng-wtsn.pptx>.
6. Kehrer Stephan, Kleineberg Oliver, Heffernan Donal. A Comparison of Fault-Tolerance Concepts for IEEE 802.1 Time Sensitive Networks (TSN) // Emerging Technology and Factory Automation (ETFA), 2014 IEEE / IEEE. — 2014. — P. 1–8.
7. Morteza Hashemi Farzaneh, Alois Knoll. Time-Sensitive Networking: an experimental setup. IEEE — 2017.
8. Scheduling Real-Time Communication in IEEE 802.1 Qbv Time Sensitive Networks / Silviu S Craciunas, Ramon Serna Oliver, Martin Chmelík, Wilfried Steiner // Proceedings of the 24th International Conference on Real-Time Networks and Systems / ACM. — 2016. — P. 183–192.
9. Rentschler Markus, Heine Holger. The parallel redundancy protocol for industrial ip networks // Industrial Technology (ICIT), 2013 IEEE International Conference on / IEEE. — 2013. — P. 1404–1409.
10. Rentschler Markus, Laukemann Per. Performance analysis of parallel redundant WLAN. // ETFA. — 2012. — P. 1–8.
11. An enhanced MAC to increase reliability in redundant Wi-Fi networks / Gianluca Cena, Stefano Scanzio, Adriano Valenzano, Claudio Zunino // Factory Communication Systems (WFCS), 2014 10th IEEE Workshop on / IEEE. — 2014. — P. 1–10.
12. Khorov Evgeny, Loginov Vyacheslav, Lyakhov Andrey. Several EDCA Parameter Sets for Improving Channel Access in IEEE 802.11ax Networks // Wireless Communication Systems (ISWCS), 2016 International Symposium on / IEEE. — 2016. — P. 419–423.
13. Bianchi Giuseppe. Performance Analysis of the IEEE 802.11 Distributed Coordination Function // IEEE Journal on selected areas in communications. — 2000. — Vol. 18, no. 3. — P. 535–547.

An Approach to Real-Time Communications in Wi-Fi Networks¹

D.V. Bankov, E.M. Khorov, A.I. Lyakhov, M.L. Sandal

Development of the wireless communication technologies exhibits the increased interest to the scenarios which involve severe restrictions on the reliability and latency of data transmission. Such scenarios include Real-Time Applications, such as the industrial automation, remote control, video streaming and virtual reality. It is very hard to satisfy the imposed quality of service requirements with the currently widespread communication technologies. Specifically, it is currently impossible to guarantee the low delay in Wi-Fi networks due to the peculiarities of the applied channel access methods. In this work we study an approach to guarantee low latency and high reliability of communications in Wi-Fi networks, based on the usage of an additional radiointerface. This approach is studied using the developed mathematical model of a heterogeneous network, which consists of devices generating prioritized and non-prioritized data packets. The obtained results show that in case when the prioritized traffic intensity is lower than the specific limits, this approach can be used to satisfy the requirements of the Ultra-Reliable Low Latency Communications, while the throughput degradation for non-prioritized traffic is insignificant.

KEYWORDS: wireless network, network protocol, reliability, low latency, real-time applications.

¹ The research was done at MIEM NRU HSE and supported by the Russian Science Foundation (agreement No 18-19-00580)