

Министерство образования и науки Российской Федерации
Ярославский государственный университет им. П. Г. Демидова
Кафедра радиофизики

А. С. Захаров

Метод случайного доступа

Часть 1

Методические указания

Рекомендовано

*Научно-методическим советом университета для студентов,
обучающихся по направлениям Радиотехника,
Радиофизика, Радиофизика и электроника*

Ярославль
ЯрГУ
2012

УДК 004.072.4(072)

ББК 3973.2я73

338

Рекомендовано

*Редакционно-издательским советом университета
в качестве учебного издания. План 2012 года*

Рецензент

кафедра радиофизики ЯрГУ

338 **Метод случайного доступа. Ч. 1** : методические указания / А. С. Захаров ; Яросл. гос. ун-т им. П. Г. Демидова. – Ярославль: ЯрГУ, 2012. – 36 с.

В методических указаниях рассмотрены принципы организации случайного доступа к общей среде передачи в информационно-вычислительных сетях. Представлены описания методов АЛОНА, дискретная АЛОНА, метода множественного доступа с контролем несущей и обнаружением коллизий (CSMA/CD– carrier sense multiple access with collision detection). Рассмотрены несколько вариантов стратегии повторной передачи при разрешении коллизий. Выполнены оценки производительности сетей, построенных на основе этих методов, проведён их сравнительный анализ. С учетом особенностей построения беспроводных сетей также рассмотрен метод множественного доступа с проверкой несущей и предотвращением столкновений (CSMA/CA– carrier sense multiple access with collision avoidance) и различные режимы его работы.

Предназначено для студентов, обучающихся по направлениям 210400.62 Радиотехника, 011800.62 Радиофизика, цикл БЗ; 010801.65 Радиофизика и электроника, блок СД (дисциплина «Информационно-вычислительные сети»), очной формы обучения.

УДК 004.072.4(072)

ББК 3973.2я73

© ЯрГУ, 2012

Ключевой проблемой построения локальных сетей является управление доступом пользователей к общей среде передачи. В 70-х годах XX века Норман Абрамсон вместе с коллегами из Гавайского университета разработали новый метод решения проблемы распределения общего канала между многими пользователями, получивший в дальнейшем название *ALOHA*. Несмотря на то что его авторы использовали широкополосную радиосвязь со стационарными передатчиками, основная идея данного метода применима к любой системе, в которой независимые пользователи соревнуются за право использования общего канала.

I. Метод «ALOHA»

В основе системы *ALOHA* лежит простая идея: разрешить всем пользователям передачу сообщений в сеть, как только у них появляются данные для отсылки. Конечно, в этом случае два или более сообщений могут накладываться друг на друга во времени, что приведёт к их столкновению, и столкнувшиеся кадры будут уничтожены. Поэтому должны существовать методы распознавания столкновений и информирования о них станций, участвовавших в передаче этих кадров. Среди них широко использовались три:

- столкновения распознавались центральной станцией (в сети «*ALOHA*» все кадры направлялись по радио в центральный пункт);
- путем применения положительных подтверждений о передаче в сочетании с перерывом;
- в локальных вычислительных сетях – путем прослушивания среды передачи и обнаружения столкновений каждой станцией.

В любом случае при обнаружении столкновений пострадавшие станции предпринимают попытки повторной передачи потерянного кадра, при этом они должны распределять время попыток случайным образом в соответствии с некоторым алгоритмом разрешения столкновений для избегания нового конфликта.

Определим производительность сети типа «ALOHA». Пусть за доступ к каналу состязаются N станций. Каждая из них передает в среднем λ кадров / сек. Пусть также все передаваемые кадры имеют одну и ту же фиксированную длину, соответствующую t единицам времени передачи. Будем считать, что ин-

тенсивность нагрузки ρ характеризует использование канала вновь поступающими кадрами S :

$$S = \rho = N \cdot \lambda \cdot m \quad (1)$$

учитывая, что $1/m = \mu$ представляет собой пропускную способность канала в передаваемых кадрах в секунду, тогда величина S будет характеризовать относительное использование канала или производительность, нормированную относительно μ :

$$S = \frac{N \cdot \lambda}{\mu} \quad (2)$$

Предположим, что поток поступающих на каждую станцию кадров является пуассоновским:

$$P_n(k) = \frac{e^{-n \cdot p} (n \cdot p)^k}{k!}, \quad (3)$$

где n – число опытов, $n \cdot p = \lambda$ – параметр, описывающий среднюю скорость поступления кадров, $P_n(k)$ – вероятность того, что за время $T=1$ событие (поступление кадра) осуществится K раз. Тогда вероятность того, что в промежутке времени T произойдет K поступлений кадров, будет равна:

$$P(k) = \frac{e^{-\lambda \cdot T} (\lambda \cdot T)^k}{k!}. \quad (4)$$

Полная нагрузка одного канала состоит из вновь передаваемых кадров и кадров, передаваемых повторно, которые, будем считать, также распределены по закону Пуассона. Это утверждение справедливо, если случайное время задержки повторной передачи будет сравнительно велико. Общая суммарная интенсивность кадров, передаваемых в канал, включая генерируемые вновь и передаваемые повторно $\lambda' > \lambda$. Тогда фактическая интенсивность нагрузки, или использования канала, определяется величиной $G = N \cdot \lambda' \cdot m$ или $G = \frac{N \cdot \lambda'}{\mu}$.

Рассмотрим два кадра длительностью m секунд каждый. Столкновение между ними может произойти в промежутке времени $2m$ секунд (рис. 1).

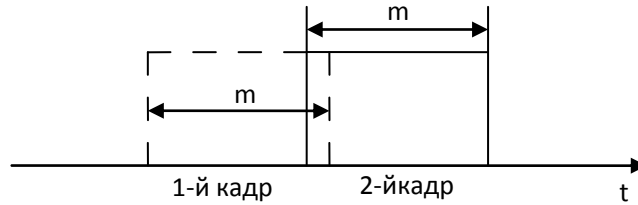


Рис. 1.

Вероятность того, что в этот промежуток не произойдет столкновения, равна вероятности того, что за это время не появится пуассоновское сообщение ($k=0$). Эта вероятность равна:

$$P(k) = e^{-\lambda T}, T = 2m \text{ – для одной станции,}$$

$$P(k) = e^{-2N\lambda'm} = e^{-2G} \text{ – для } N \text{ станций.}$$

Отношение S/G представляет долю кадров из числа передаваемых в канал, которые проходят успешно. Это число должно быть равно вероятности успеха, т. е. вероятности отсутствия столкновений. Отсюда выражение производительности для системы «ALOHA»:

$$S = G \cdot P = Ge^{-2G} \quad (5)$$

Здесь S – нормированная производительность (средняя скорость поступления новых кадров/максимальную пропускную способность сети в кадрах $1/m$), G – нормированная полная нагрузка, включающая в себя повторно передаваемые кадры и вновь поступающие в сеть за время одного кадра или, другими словами, среднее количество попыток передачи за время одного кадра.

Найдём максимальную производительность для системы «Алоха»:

$$\frac{dS}{dG} = e^{-2G} + Ge^{-2G}(-2) = e^{-2G}(1 - 2G) = 0 \Big|_{\text{при } G = \frac{1}{2}},$$

$$S_{\max} = 0.5 \cdot e^{-1} \approx 0.18 \quad (6)$$

Зависимость производительности системы «ALOHA» от поступающей нагрузки показана на рис. 2.

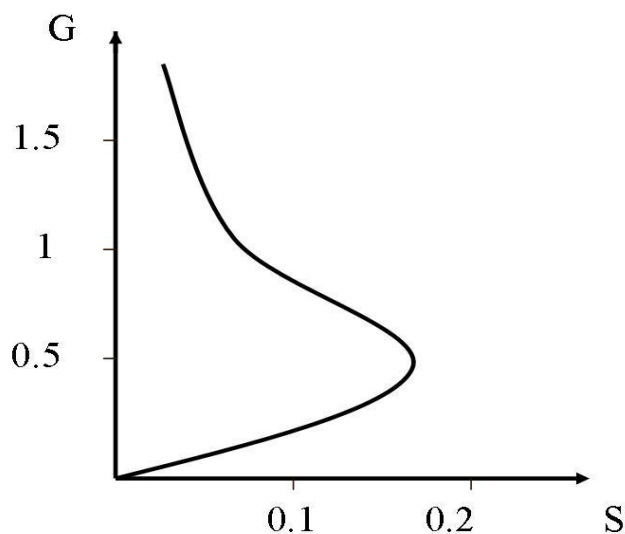


Рис. 2

При малой поступающей нагрузке S столкновения происходят редко, и $G \approx S$. При $S \rightarrow 0.18$ среднее количество попыток передачи за время одного кадра достигает $\frac{1}{2}$, что приводит к резкому росту числа столкновений. Система теряет устойчивость, количество успешно переданных кадров, а следовательно, и производительность системы S падает, а G – увеличивается до больших значений.

Лучшее, что можно достигнуть в системе «ALOHA», – это использовать канал на 18%. Однако для практических целей такой производительности часто бывает достаточно.

Оценим задержки, вносимые системой типа «ALOHA». Для этого необходимо определить стратегию повторных передач при возникновении столкновений (повторные передачи должны быть разделены во времени для уменьшения возможных повторных столкновений).

Пусть интервал времени, в течение которого станции предпринимают попытки повторной передачи «битых» кадров $\Delta T = K \cdot t$, где t – время передачи кадра, K – произвольное целое число. Тогда повторная передача происходит в пределах от 1 до K таких t интервалов после обнаружения столкновения, а момент начала передачи t имеет равномерное распределение внутри этого интервала ΔT . Пусть задержка при циклическом обходе и время

обработки, требуемое для получения информации о необходимости повторной передачи, занимает целое число R интервалов времени передачи кадра m . Это соответствует простой процедуре, когда получатель информации дает положительное подтверждение каждой передаче. В этом случае станция источника кадров включалась бы для повторной передачи через R интервалов длин кадра, если к этому времени не поступит положительное подтверждение от станции-получателя об успешной передаче.

Тогда среднее время, требуемое для успешной передачи кадра, равно:

$$D = m \left[1 + R + E(R + (K + 1) / 2) \right] \quad (7)$$

где E – среднее число попыток повторной передачи на один успешно переданный кадр.

При малых ΔT возникает больше столкновений и, следовательно, больше повторных передач. При больших ΔT зависимость $E(\Delta T)$ исчезает. Если G – нормированная сумма первичных попыток и повторных передач, тогда, очевидно, $G/S = 1 + E$, учитывая (5), получим выражение для определения среднего числа попыток повторной передачи E :

$$E = e^{2G} - 1 \quad (8)$$

II. Метод дискретная «ALOHA»

В 1972 году Робертс опубликовал описание метода, который позволил существенно улучшить производительность систем «ALOHA». Его предложение заключалось в разделении времени на дискретные интервалы, соответствующие времени одного кадра. Станциям разрешалось начинать попытки передачи кадров только в начале каждого временного интервала m , равного длительности кадра (рис. 3). Одним из способов достижения синхронизации является установка специальной станции, передающей синхронизирующий сигнал в начале каждого интервала. Если повторные попытки также, как и первоначальные, описываются пуассоновским распределением, то, так как опасный интервал времени возможного столкновения стал в два раза короче, вероят-

ность успешной передачи задается функцией e^{-G} , а производительность дискретной Алохи будет:

$$S = Ge^{-G} \quad (9)$$

В этом случае максимальная производительность системы дискретная «ALOHA» будет равна: $S_{\max} = \frac{1}{e} \approx 0.368$ при $G=1$, т. е. в два раза выше, чем в системе «ALOHA».

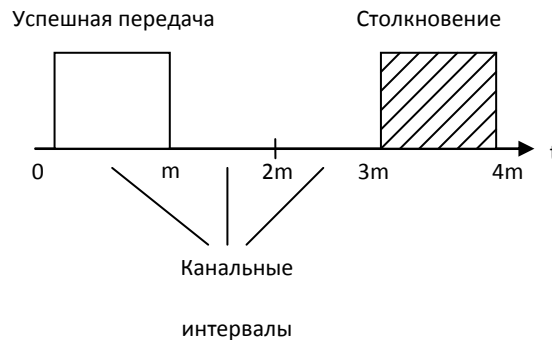


Рис. 3

Для системы дискретная «ALOHA» в оптимальной ситуации 37% интервалов будут свободными, 37% – с успешно переданными кадрами, а 26% – со столкнувшимися кадрами. При увеличении количества попыток передачи в единицу времени G , число пустых интервалов уменьшается и увеличивается количество интервалов, в которых произошли столкновения.

На рис. 4 показана зависимость производительности системы дискретная «ALOHA» от полной нагрузки на сеть. Как видно из рис. 4, при $S < S_{\max}$ возможны два значения G , т. е. для системы дискретная «ALOHA» также характерна неустойчивость. Чтобы увидеть, насколько быстро растёт количество столкновений, рассмотрим передачу тестового кадра. Вероятность того, что он избежит столкновения, равна $1 - e^{-G}$. Вероятность передачи кадра ровно за k попыток, т. е. после $k-1$ столкновения, за которыми последует успешная передача, равна

$$P_k = e^{-G} (1 - e^{-G})^{k-1} \quad (10)$$

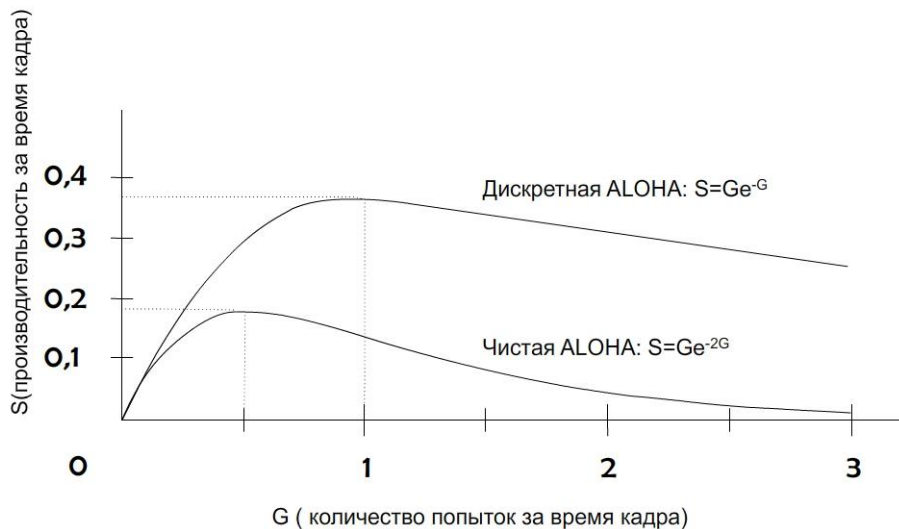


Рис. 4. Зависимость производительности систем «ALOHA» от полной нагрузки на сеть

Ожидаемое число попыток передачи для одного кадра равно:

$$E = \sum_{k=1}^{\infty} k P_k = \sum_{k=1}^{\infty} k e^{-G} (1 - e^{-G})^{k-1} = e^G \quad (11)$$

Поскольку число попыток передачи E одного кадра экспоненциально зависит от количества попыток передачи в единицу времени G , то небольшое увеличение нагрузки на систему может сильно снизить её производительность.

Для устранения трудностей, связанных с неустойчивостью схемы «ALOHA», предлагается ряд механизмов управления, например увеличение времени повторной передачи после каждого обнаружения столкновения. Эта процедура уменьшает вероятность столкновения, но увеличивает задержку в системе. Подобный метод используется в протоколе CSMA/CD (множественный доступ с контролем несущей и обнаружением коллизий).

Оценим задержки, вносимые системой дискретная «ALOHA». Стратегия повторной передачи используется такая же, как и для системы «ALOHA»: узнав, что через R интервалов после попытки передачи произошло столкновение, станция осуществляет повторную передачу в любом из K интервалов с одинаковой вероятностью. Эта ситуация проиллюстрирована на рис. 5.

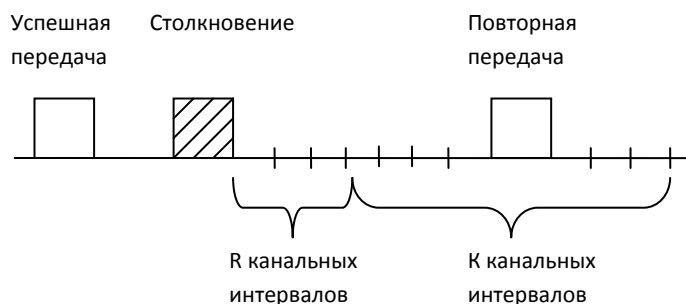


Рис. 5. Стратегия повторной передачи в системе *синхронная «ALOHA»*

В этом случае время D , требуемое для успешного завершения передачи в системе *дискретная «ALOHA»* равняется:

$$D/m = 1.5 + R + E^1 \left[R + 0.5 + (K + 1)/2 \right] \quad (12)$$

Здесь R – нормированное время цикла, необходимое для получения сигнала подтверждения успешной передачи (величина постоянная), E^1 – среднее число попыток повторной передачи. При больших K производительность $S = Ge^{-G}$, т. е. не зависит от K . Поэтому $E = G/S - 1 = e^G - 1$. Анализ показал, что при $K \geq 5$ – это уже хорошее приближение. Отличия выражения (10) от выражения (7) для системы Алоха – наличие в формуле величин 0,5 в двух местах. Эти поправки требуются для учета влияния кадров, которые появляются в системе после того, как начался интервал.

На рис. 6 приведены зависимости нормированной задержки в схемах Алоха от нагрузки.



Рис. 6. Зависимость нормированной задержки от нагрузки

Оптимальное значение величины K , минимизирующее задержку D , зависит от S – нормированной производительности и нормированной задержки R – времени циклического распространения. Хороший компромисс достигается при $K=5$. Таким образом, из рис.6 видно, при малых нагрузках система Алоха обеспечивает меньшее время доставки кадров и является более эффективной по сравнению с дискретной Алохой. С ростом нагрузки картина меняется на противоположную.

III. Метод множественного доступа с проверкой несущей и обнаружением столкновений (CSMA/CD– carrier sense multiple access with collision detection)

Метод основан на методе чистой Алохи и позволяет улучшить ее характеристики (т. е. производительность $> 0,368$ от пропускной способности канала). Идея метода заключалась в том, что все станции перед началом передачи прослушивают линию. Станция, желающая передать сообщение, выходит на связь только после обнаружения свободного состояния канала. Эта процедура называется проверкой несущей. Однако столкновения все же могут возникнуть, поскольку станции физически разнесены

одна от другой и две или несколько станций одновременно (с точностью до задержки распространения сигнала) могут обнаружить свободное состояние канала и начать передавать, что и вызовет столкновение. Если станции обнаружат столкновение кадров, они передают всем остальным станциям специальный сигнал о помехе и отменяют свои передачи.

Возможность проверки несущей позволяет увеличить производительность канала по сравнению с чистой Алохой, а обнаружение столкновения с прекращением передачи вместо его завершения дает еще большее увеличение производительности.

Известно несколько методов *CSMA*, они отличаются друг от друга организацией управления передачи, если канал оказался занятым:

- в схеме с *p*-настойчивостью: станция, обнаружившая занятый канал, осуществляет передачу после того, как канал станет свободным, с вероятностью *p*. С вероятностью $(1-p)$ передача откладывается на промежуток времени τ распространения сигнала;

- в схеме 1-настойчивостью станция осуществляет попытку передачи, как только канал станет свободным;

- при ненастойчивой схеме станция переносит передачу на другое время в соответствии с предписанным распределением задержек передачи, проверяя в это время несущую.

Все эти схемы рассчитаны на то, что станция в состоянии обнаружить конец передачи вскоре после его завершения, т. е. они требуют, чтобы задержка распространения τ между наиболее удалёнными станциями в сети была мала по сравнению с длительностью кадра *m*: $a = \tau/m \ll 1$. Если $a \rightarrow 1$ или $a > 1$, то работа схемы *CSMA* переходит в схему «*ALOHA*». Таким образом, эти схемы применимы для локальных сетей или в крупных сетях с небольшими скоростями передачи, для которых длительность кадра *m* велика. Протокол *CSMA*, работающий по правилу 1-настойчивости с добавлением возможности определения коллизий *CD*, принят в качестве протокола в схеме Ethernet.

После обнаружения столкновения передача прерывается, попытка повторной передачи предпринимается через случайный промежуток времени, который удваивается каждый раз после обнаружения нового столкновения до некоторой максимальной величины, при которой станция прекращает дальней-

шие попытки передачи и извещает вышестоящие уровни о нарушении связи. Процесс увеличения задержки повторной передачи уменьшает вероятность столкновения и возможную нестабильность работы системы.

Для оценки производительности метода CSMA рассмотрим сеть, построенную по шинной топологии (рис. 7).

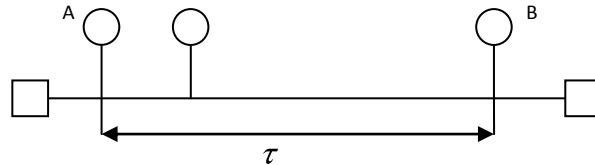


Рис. 7.

Здесь **A** и **B** – наиболее удалённые друг от друга станции в сегменте сети, τ – время распространения электромагнитного сигнала между ними.

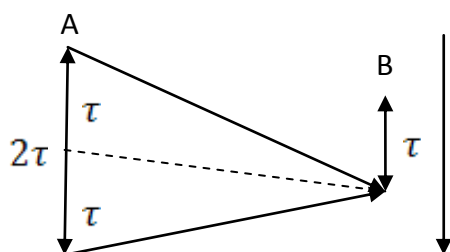
Определим среднее время, требуемое для успешного запуска кадра в шину (рис. 8). Обратная величина его – будет максимальной производительностью. Время для успешного завершения передачи кадра – назовем *виртуальным временем передачи* t_v .



Рис. 8

t_v имеет 3 составляющих: **m** – требуется для передачи сообщения (предполагаем фиксированным), τ – требуется для проверки завершения передачи, а кратное 2τ – требуется для разрешения столкновений, если они обнаружатся. Для обнаружения конца пе-

редачи требуется время τ , равное времени распространения сигнала от станции **A** до станции **B**. Только после получения последнего бита сообщения станция **B** убеждается в окончании передачи, что приводит к фактическому увеличению времени передачи кадра на величину τ и снижению максимальной производительности.



Оценим максимальную пропускную способность метода множественного доступа с контролем несущей и обнаружением коллизий. При столкновении между сигналами от станций **A** и **B**, обнаружение его зай-

мет в худшем случае время 2τ , после чего передача будет немедленно прекращена. Станция **A** начинает передачу в некоторый произвольный момент времени. До поступлением сообщения от станции **A** на станцию **B** (для этого требуется время $\leq \tau$), последняя проверяет канал и, если находит его свободным, начинает свою передачу. Возникающее в этом случае столкновение обнаруживается станцией **A** через время τ (время прохождения сообщения от станции **B** до станции **A**). Общее время обнаружения столкновения составит 2τ . Если произошло столкновение, предположим, что для его разрешения требуется $2\tau J$ -единиц времени (J – среднее число повторных передач после ОС, оно сопоставимо с параметром $E = G/S - 1$ в схеме Алоха). В этом случае выражение для виртуального времени передачи имеет вид:

$$t_v = m + \tau + 2\tau J = m [1 + a(1 + zJ)], \quad a = \tau/m(13)$$

Найдем величину J . Она зависит от стратегии повторной передачи. Предположим, что длительность интервала столкновения описывается геометрическим распределением единиц 2τ с параметром v . Интервал, равный одной единице (2τ), имеет вероятность $p = v$. 2 единицам ($2 \times 2\tau$) – $p = v(1 - v)$, 3 единицам ($3 \times 2\tau$) – $p = v(1 - v)^2$ и т. д. Таким образом, v является вероятностью успеха в конце интервала, а $(1 - v)$ – вероятностью столкновения. Тогда среднее число повторных передач J определяется выражением:

$$J = \sum_{k=1}^{\infty} kv(1 - v)^{k-1} = 1/v(14)$$

Найдем v . Пусть в возможных передачах участвуют n станций ($n \gg 1$). Пусть вероятность того, что одна станция намеревается передавать в промежутке времени 2τ , равна p . Тогда вероятность того, что передает точно одна станция и эта передача успешна:

$$v = np(1 - p)^{n-1} \quad (15)$$

$p = 1/n$ – максимизирует это выражение, т. к.

$$v' = n \left((1-p)^{n-1} + p(n-1)(1-p)^{n-2}(-1) \right) = \quad (16)$$

$$n(1-p)^{n-2}(1-pn) = 0 \quad , p = \frac{1}{n} ,$$

при этом условии вероятность успешной передачи кадра наибольшая. Используя $p = 1/n$ и учитывая, что $n \gg 1$, в пределе получим:

$$v_{max} = \left(1 - \frac{1}{n}\right)^{n-1} \cdot e^{-1}, \quad n \rightarrow \infty \quad (17)$$

В этом случае:

$$t_v = m [1 + a(1 + 2e)], \quad a \equiv \tau/m \quad (18)$$

Эта модель напоминает синхронную Алоху и фактически приводит к такой же производительности $1/e$. Максимальная производительность λ_{max} , выраженная в числе сообщений за единицу времени, равна $1/t_v$, где t_v – виртуальное время передачи, соответствует максимальному числу эквивалентных сообщений в секунду, которые могут быть переданы по общему каналу последовательно одно за другим. В этом случае:

$$\bar{\lambda}m < \frac{1}{1 + a(1 + 2e)} = \frac{1}{1 + 6,44a}, \quad a \equiv \tau/m \quad (19)$$

Здесь $\bar{\lambda}$ – среднее число сообщений, передаваемых по общему каналу за единицу времени всеми пользователями. Напомним,

что $\bar{\lambda}t \equiv \rho$, т. е. определяет интенсивность нагрузки на сеть, которую создают вновь поступающие пакеты.

Таким образом, согласно выражению (19), с учётом сделанных предположений можно сделать вывод, что при соотношении времени распространения сигнала между наиболее удалёнными станциями в сети к длине сообщения, равной 0,1, максимальная производительность сети, реализующей данный метод доступа к среде, может достигать 0,6 скорости передачи в канале. Для сравнения в случае схемы АЛОНА максимальная производительность составляет $\rho = 0,18$, а для дискретной АЛОНА – $\rho = 0,36$.

На рис. 9 представлены зависимости производительности различных протоколов случайного доступа от нагрузки на сеть и нескольких схем организации повторной передачи.

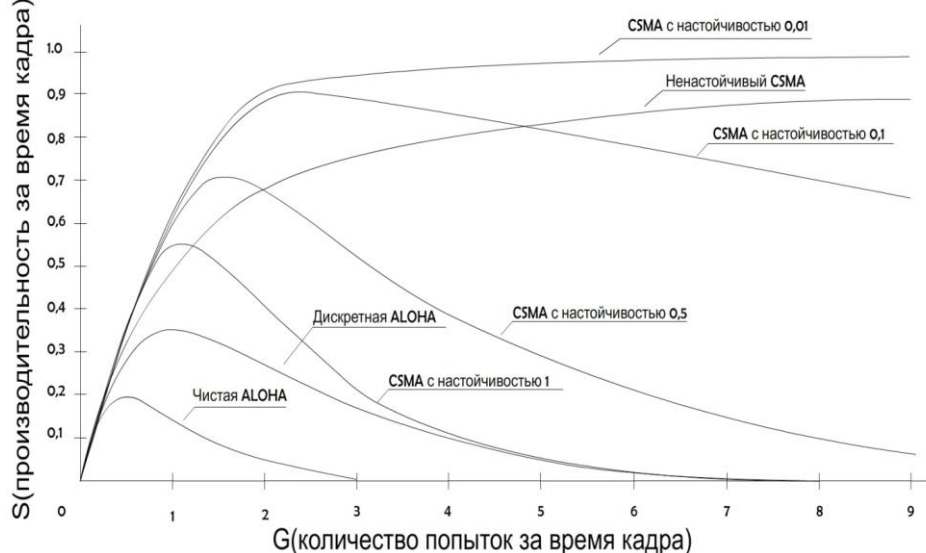


Рис. 9. Сравнение эффективности использования канала в зависимости от нагрузки на сеть

При увеличении длины передаваемых сообщений, согласно выражению (16), эффективность использования канала возрастёт, но при этом ухудшится доступность сети для станций. Зависимость эффективности использования канала в сети с CSMA/CD при максимальной нагрузке, скорости передачи 10 Мбит/с и 64-байтовыми интервалами разрешения столкновений от числа станций в сети для различных длин сообщений представлены на рис. 10.

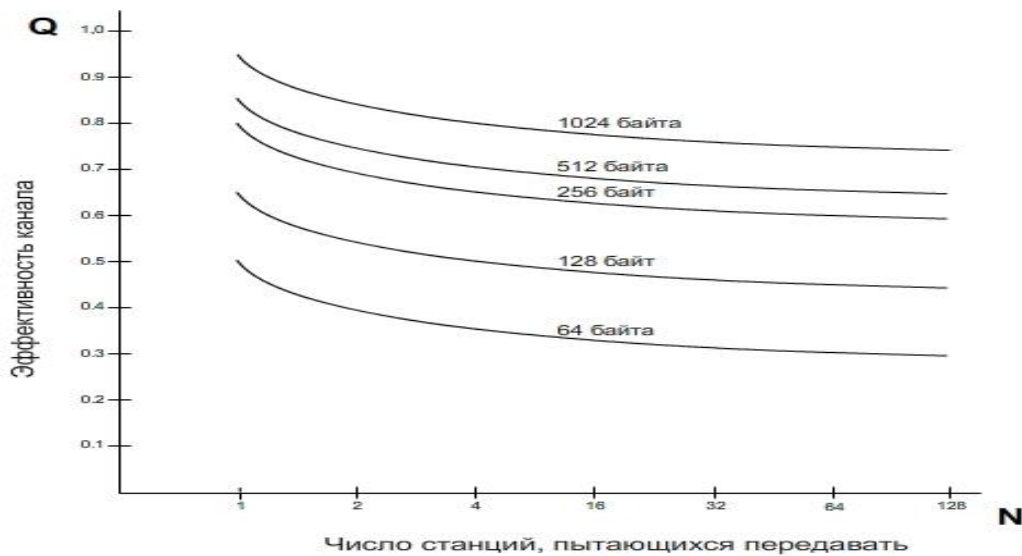


Рис.10.Эффективность использования канала в сети сCSMA/CD при скорости10Мбит/с и 64-байтовыми интервалами разрешения столкновений

Как видно из рис. 9, при возрастании общей нагрузки в сети, связанной с увеличением количества станций либо их активности, произойдёт насыщение сети, что приведёт к резкому увеличению времени доставки сообщений. Одним из возможных решений является увеличение скорости передачи данных, например, со 100 Мбит/с до 1000 Мбит/с или 10 Гбит/с. Такой подход требует замены сетевого оборудования. Однако существует не столь радикальное решение, такое как увеличение линейной скорости передачи данных в сети, а можно ограничиться изменением её структуры, т. е. отказаться от шинной топологии и перейти к звездообразной, а в современных больших сетях – к древовидной (рис. 11).Такой подход легко реализуется в сетях, где в качестве среды передачи используется витая пара или волоконно-оптический кабель.

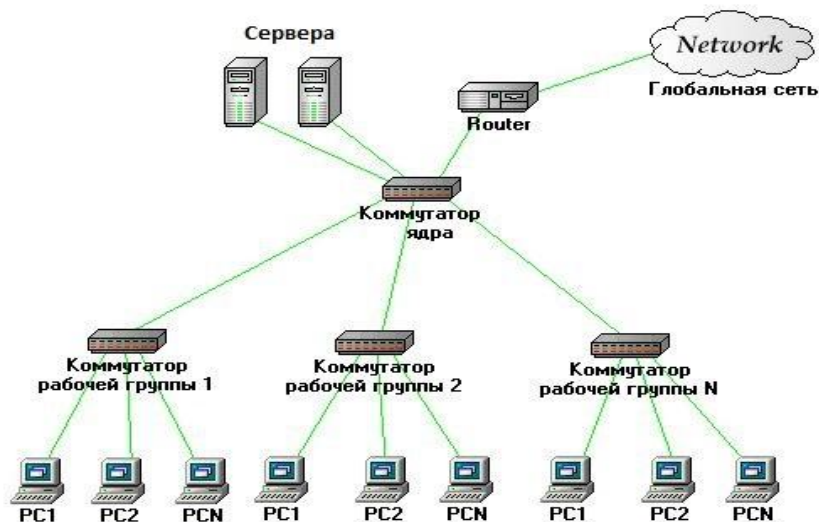


Рис. 11. Современная топология построения локальной сети

Сердцем такой системы является коммутатор (switch), включающий в себя высокоскоростную плату (коммутационную матрицу) и от 4 до 96 контроллеров линий (портов). Структурная схема его представлена на рис. 12. Коммутатор представляет собой многопортовый мост и работает на втором канальном уровне модели OSI. Главное назначение коммутатора – обеспечение разгрузки сети посредством локализации трафика в пределах отдельных сегментов. При построении коммутаторов используется архитектура без блокирования, которая позволяет установить множественные связи между разными парами портов одновременно, причём передаваемая станция микадры не теряются в процессе коммутации. При этом трафик между взаимодействующими сетевыми устройствами оказывается локализованным. Локализация осуществляется с помощью адресных таблиц, устанавливающих связь каждого порта с адресами сетевых устройств, относящихся к сегменту этого порта. Таблица заполняется в процессе анализа коммутатором адресов станций, подключенным к его портам, при передаче ими кадров.

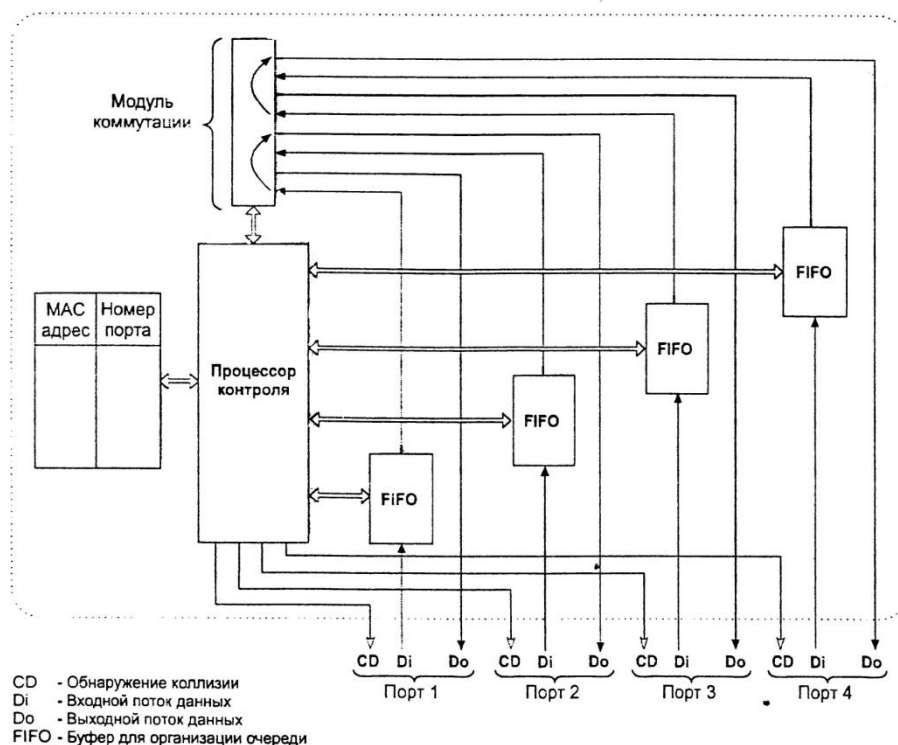


Рис. 12. Структурная схема коммутатора

Кадр передаётся через коммутатор локально в соответствующий порт только тогда, когда адрес станции назначения, указанный в поле адреса, уже содержится в адресной таблице этого порта. В случае отсутствия адреса станции назначения в таблице кадр рассылается во все остальные сегменты. Если коммутатор обнаруживает, что MAC-адрес станции назначения у входящего кадра находится в таблице MAC-адресов, приписанной за портом, в который поступил данный кадр, то этот кадр сбрасывается – его непосредственно получит станция назначения, находящаяся в этом же сегменте. Если входящий кадр является широковещательным (broadcast), т. е. если все биты поля MAC-адреса получателя в кадре задаются равными 1, то такой кадр будет направлен коммутатором во все остальные порты. Архитектура и типичная реализация контроллера линии (порта), а также выполняемые им функции приведены на рис. 13.

Как отмечалось выше, метод множественного доступа с контролем несущей и обнаружением столкновений лежит в основе популярной в настоящее время технологии построения локальных сетей Ethernet и определяет базовый функциональный

профиль 802.3.Общепринято считать, что Ethernet был изобретён 22 мая 1973 года, когда Роберт Меткалф (Robert Metcalfe) составил докладную записку для главы PARC о потенциале технологии Ethernet. Но законное право на технологию Меткалф получил через несколько лет. В 1976 году он и его ассистент Дэвид Боггс (David Boggs) издали брошюру под названием «Ethernet: Distributed Packet-Switching For Local Computer Networks». Впервые стандарт был опубликован 30 сентября 1980 года. Он начал соперничество с двумя крупными запатентованными технологиями: Token Ring и ARCNET, и вскоре стал самым распространённым стандартом для локальных сетей. Причиной такого успеха Ethernet были высокая скорость передачи и недорогое оборудование.

Базовый функциональный профиль 802.3 имеет двухуровневую архитектуру (3-уровневую, если рассматривать физическую среду как отдельный уровень), включает в себя канальный и физический уровни. **Канальный уровень** отвечает за прием пакетов от подключенных к сети устройств, формирует кадры данных для передачи, распознаёт пакеты, предназначенные для конкретной станции, обнаруживает (но не исправляет) ошибки путём анализа контрольной последовательности кадра, управляет связью, реализуя алгоритм случайного доступа CSMA/CD к шине. **Физический уровень** выполняет кодирование и декодирование приходящего с канального уровня блока данных, создание и удаление преамбул кадров, преобразование из двоичного в манчестерский код и обратно, передачу и прием, согласование физических характеристик сигналов со средой передачи, прослушивание шины, обнаружение конфликтов. **Физическая среда** обеспечивает соединения между станциями. Раньше в качестве среды передачи использовались «толстый» и «тонкий» коаксиальные кабели, многомодовый волоконно-оптический кабель, витая пара 3 и 5 категорий, повторители. Современные локальные сети строятся на основе витой пары 5Е и 6 категорий и оптического кабеля, при этом всё чаще одномодового, вследствие снижения стоимости элементной базы оптической линии связи.

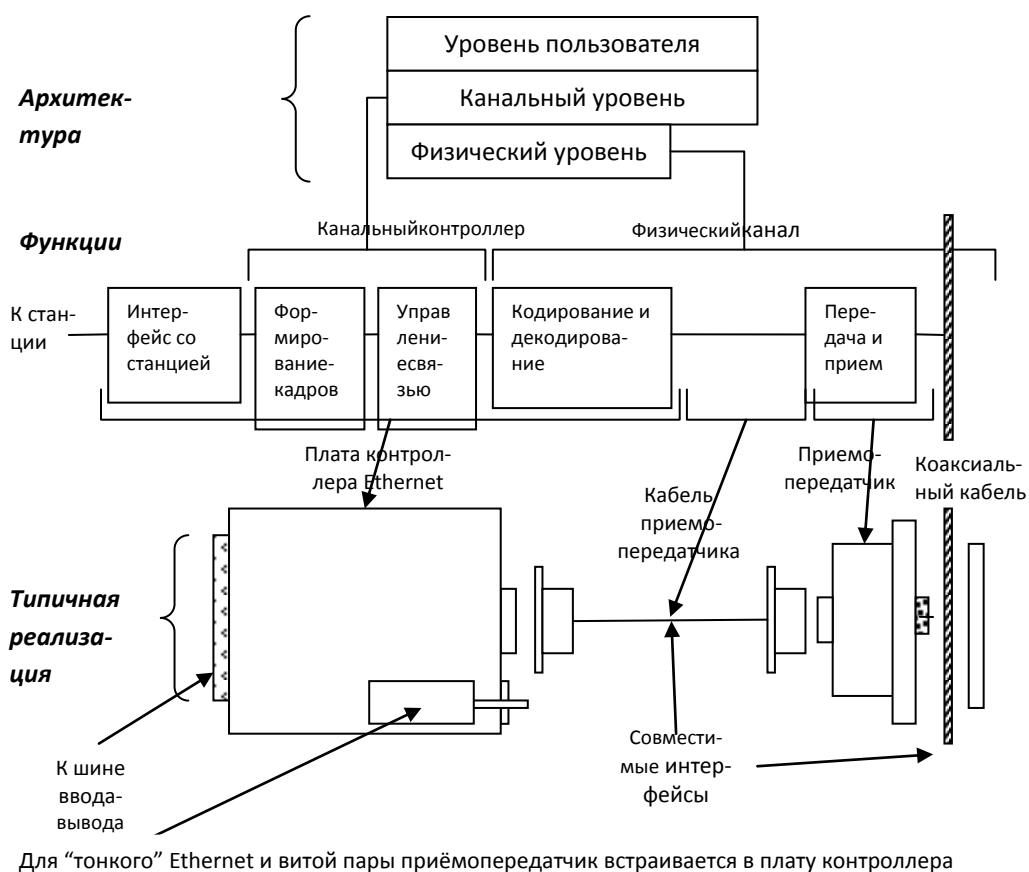


Рис. 13. Архитектура, функции и типичная сетевого контроллера

Уровни базового функционального профиля 802.3 соотносятся с уровнями эталонной модели ISO (IEEE) следующим образом (рис. 14).

Уровни	Функции уровней в модели ISO (IEEE)	Функции уровней в стеке Ethernet
2	Управление логическим каналом(LLC)	Формирование кадров
	Управление доступом к среде (MAC)	Управление связью
1	Передача физических сигналов (PLS)	Кодирование и декодирование
	Подключение к физической среде (PMA)	Передача и приём

Рис. 14.Соответствие функций уровней модели ISO и Ethernet

Для реализации рассмотренных выше функций, стандарт 802.3 предусматривает формирование кадров, имеющих следующую структуру (рис. 15).

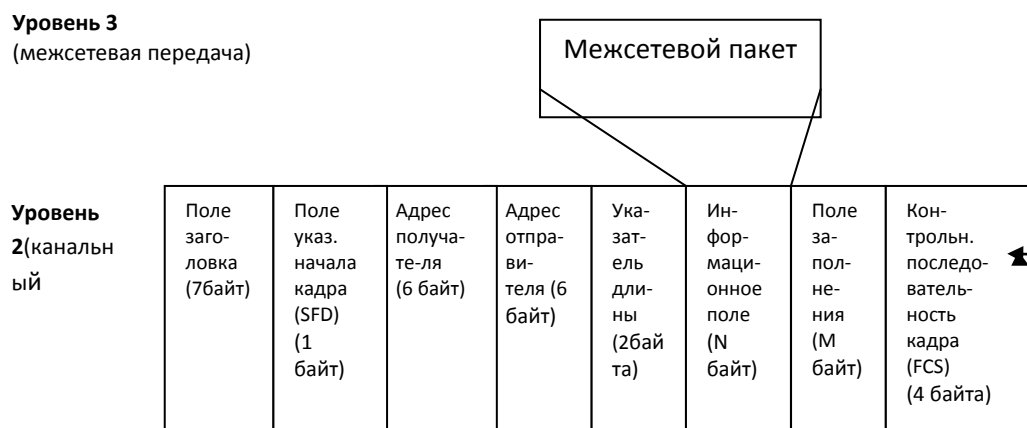


Рис. 15. Структура кадра сети Ethernet

Стандарт 802.3 определяет восемь полей заголовка:

- *Поле преамбулы (поле заголовка)* состоит из семи байтов синхронизирующих данных. Каждый байт содержит одну и ту же последовательность битов – 10101010. При манчестерском кодировании эта комбинация представляется в физической среде периодическим волновым сигналом. Преамбула используется для того, чтобы дать время и возможность схемам приемопередатчиков (transceiver) прийти в устойчивый синхронизм с принимаемыми тактовыми сигналами.

- *Начальный ограничитель (поле указателя начала, SFD – start delimiter frame)* кадра состоит из одного байта с набором битов 10101011. Появление этой комбинации указывает, что далее последуют информационные поля кадра.

- *Адрес получателя* – может быть длиной 6 байтов (MAC-адрес получателя). Первый бит адреса получателя – это признак того, является адрес индивидуальным или групповым: если 0, то адрес указывает на определенную станцию, если 1, то это групповой адрес нескольких (возможно всех) станций сети. При широковещательной адресации все биты поля адреса устанавливаются в 1. Общепринятым является использование 6-байтовых адресов.

- *Адрес отправителя* – 6-байтовое поле, содержащее адрес станции отправителя. Первый бит всегда имеет значение 0.
- *Двухбайтовое поле длины* определяет длину поля данных в кадре.
- *Поле данных* может содержать от 0 до 1500 байт. Но если длина поля меньше 46 байт, то используется следующее поле – поле заполнения, чтобы дополнить кадр до минимально допустимой длины.
- *Поле заполнения* состоит из такого количества байтов заполнителей, которое обеспечивает определенную минимальную длину поля данных (46 байт). Это обеспечивает корректную работу механизма обнаружения коллизий. Если длина поля данных достаточна, то поле заполнения в кадре не появляется.
- *Поле контрольной последовательности кадра (FCS– frame check sequence)* – 4 байта, содержащие значение, которое вычисляется по определенному алгоритму (полиному CRC-32). После получения кадра рабочая станция выполняет собственное вычисление контрольной суммы для этого кадра, сравнивает полученное значение со значением поля контрольной суммы и таким образом определяет, не искажен ли полученный кадр.

Обратим внимание, что максимальная длина кадра равна 1526 байтам. Это важно для выбора размера буферной памяти сетевого оборудования и для оценки общей загруженности сети. Дальнейшее увеличение длины кадра негативно скажется на доступности сети для станций.

Рассмотрим процедуры передачи и приема кадров в сети, работающей со скоростью 10 Мбит/с и использующей в качестве среды передачи коаксиальный кабель. Это позволит нам проиллюстрировать реализацию метода CSMA/CD в сети Ethernet.

Передача кадра. Алгоритм работы станции при передаче представлен на рис. 15. Канальный уровень оповещается физическим о наличии сигналов в кабеле. Передача не начинается до их исчезновения, и даже после этого канальный уровень ожидает еще не менее 9,6 мкс для того, чтобы канал освободился.

Если канальный уровень оповещается о возникновении конфликта во время передачи, то последняя прерывается посылкой сигнала JAM, который содержит произвольные данные и гаран-

тирует надежное обнаружение конфликта всеми станциями. Длина сигнала JAM – 32 бита.

После конфликта все станции выполняют процедуру «усеченный двоичный экспоненциальный алгоритм отсрочки», которая состоит в следующем: в каждой станции генерируется случайное число, которое определяет длительность задержки до следующей попытки передачи. Время задержки кратно $51,2 \text{ мкс} >$ наихудшего времени круговой задержки. В результате любая станция, «выбросившая» наименьшее случайное число, получает возможность завершить передачу без конфликта. Если наименьшее случайное число выбросили 2 или более станций, процедура повторяется с увеличенным диапазоном возможных чисел:

- в течение первых 10 попыток диапазон генерирования случайных чисел экспоненциально возрастает от $0 \div 1$ до $0 \div 1023$ (то есть $\tau_{max} = 65 \text{ мс}$), а для 5 последующих попыток остается на том же уровне;

- если 16-я попытка неудачна, канальный уровень отказывается от передачи и извещает об этом верхний уровень.

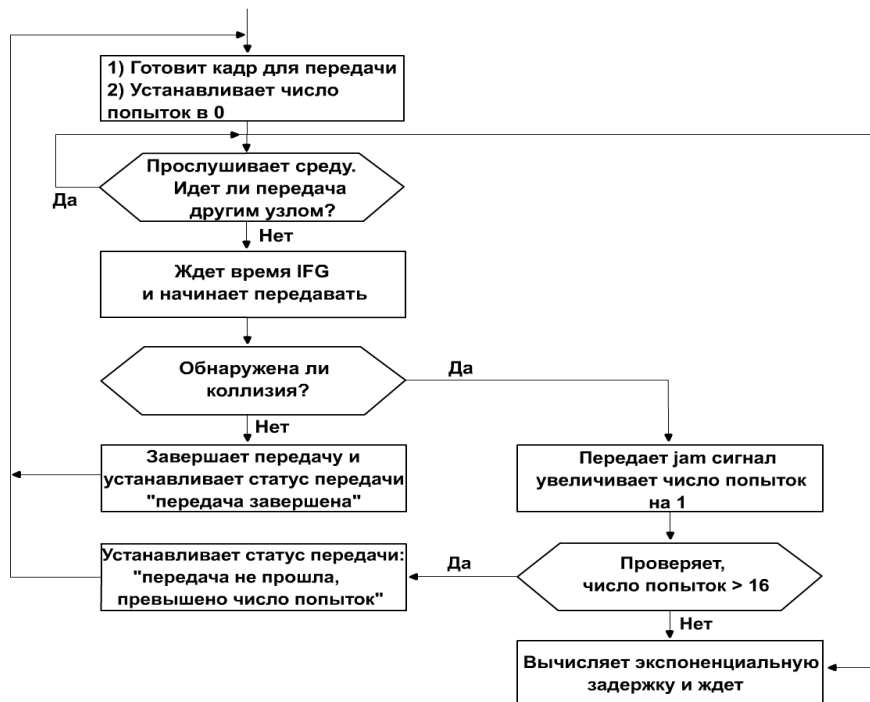


Рис. 16. Структурная схема алгоритма работы станции при передаче

Прием кадра (рис. 17). При обнаружении сигнала в линии, принимающая станция на физическом уровне первым делом синхронизируется по преамбуле и затем преобразует манчестерский код в бинарную форму. Далее на канальном уровне оставшиеся биты преамбулы отбрасываются. Канальный уровень отвергает также кадры короче 64 байт или кадры, длительность которых не кратна байту. Такая ситуация может возникнуть только в результате конфликта.

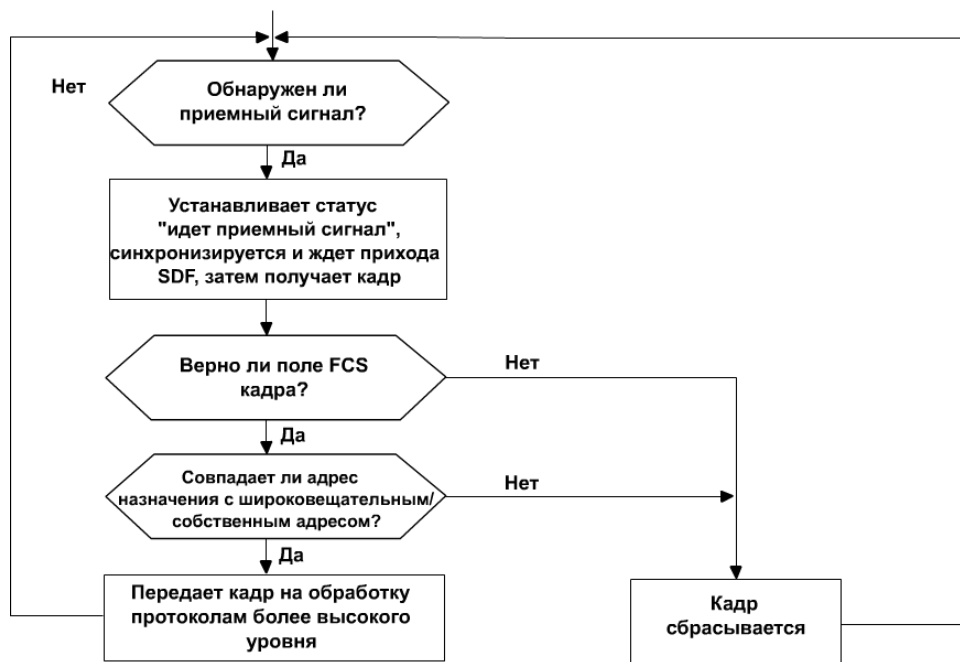


Рис.17. Структурная схема алгоритма работы станции при приёме

Затем станция читает адрес назначения и сравнивает его со своим собственным. Если адреса совпадают, то поля кадра, за исключением полей указателя начала и контрольной последовательности, помещаются в буфер и вычисляется контрольная сумма, которая сравнивается с полем контрольной последовательности кадра. Если они равны, то содержимое буфера передаётся протоколу более высокого уровня. В противном случае кадр отбрасывается. Возникновение коллизии при приёме кадра обнаруживается либо по изменению электрического потенциала в линии (если используется коаксиальный кабель), либо по факту приёма дефектного кадра (неверная контрольная сумма), если используется витая пара или оптическое волокно. В обоих случаях принятая информация отбрасывается.

IV. Метод множественного доступа с проверкой несущей и предотвращением столкновений (CSMA/CA- carrier sense multiple access with collision avoidance)

В проводных сетях Ethernet, построенных на основе метода CSMA/CD, станция ожидает, когда в канале будут отсутствовать сигналы, и тогда начинает передачу. Если шумовой всплеск не приходит обратно в течение времени, необходимого на пересылку 64 байт, то можно утверждать, что передаваемый кадр доставлен корректно. Напомним, что необходимым условием корректного применения данного метода является возможность прослушивания канала всеми станциями в сети. При построении беспроводных сетей ввиду их специфики такой подход не работает. Причиной этого является проблема скрытой станции, которая возникает вследствие особенностей распространения радиоволн. Поскольку не все станции могут слышать друг друга, передача, идущая в одной части соты, например станция **A** передаёт сообщение станции **B**, может быть не принята станцией **C**, находящейся в другой его части. При этом последняя сделает ложный вывод о том, что канал свободен и она имеет право также начать передачу для станции **B**. Кроме этого, в беспроводных сетях возможны ситуации, когда станция **B** хочет отправить данные для станции **A** и прослушивает канал. Услышав, что в нём уже осуществляется какая-то передача, например станцией **C**, станция **B** делает ложный вывод о том, что передача в данный момент для станции **A** невозможна, хотя станция **C** передавала данные другой станции, в то время как станция **A** находилась вне зоны радиовидимости для станции **C**, производившей передачу, то есть была для неё недоступна. В этом случае сигнал от станции **C** не мог повлиять на работу станции **A** и создать помехи при приёме. Данная ситуация показана на рис. 18.

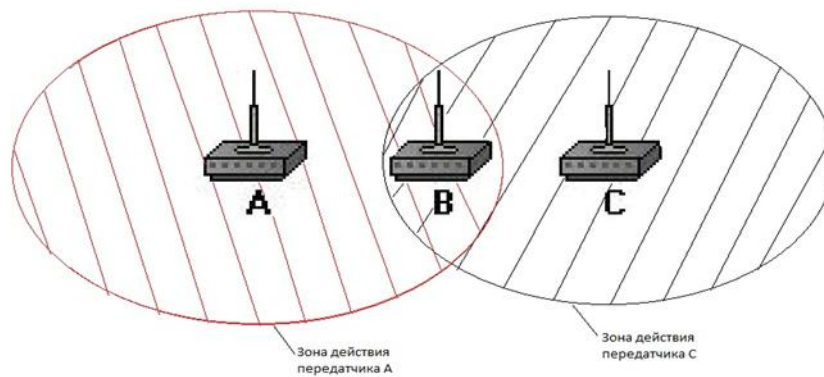


Рис. 18. Проблемы скрытой и засвеченной станций

Таким образом, в данном случае станция **В** откладывает передачу для станции **А** на некоторое время, хотя могла бы её успешно произвести. Ситуация усугубляется ещё и тем, что большинство радиосистем являются полудуплексными, то есть не могут одновременно и на одной и той же частоте производить передачу и контролировать канал. Поэтому применение метода CSMA/CD в беспроводных сетях является невозможным.

В связи с этими обстоятельствами для беспроводных сетей метод CSMA/CD был усовершенствован. Были разработаны два режима работы. Первый называется **DCF (Distributed Coordination Function – распределённая координация)** и не имеет никаких средств централизованного управления, в этом смысле напоминая Ethernet. Второй режим – **PCF (Point Coordination Function – сосредоточенная координация)** – подразумевает, что базовая станция берёт на себя функцию управления активностью всех станций данной соты. Данные доработки реализованы в базовом функциональном профиле, описанном в стандарте 802.11.

В режиме DCF 802.11 использует протокол, называемый CSMA/CA (carrier sense multiple access with collision avoidance). Здесь ведётся прослушивание как физического, так и виртуального канала. Протокол CSMA/CA может работать в двух режимах. В первом – станция перед передачей прослушивает канал и, если он свободен, начинается пересылка данных. Во время пересылки канал не прослушивается, и станция передаёт кадр целиком, при-

чём он может быть разрушен на входе приёмника из-за интерференции сигналов. Если канал занят, отправитель дожидается его освобождения и затем начинает передачу. Если возникает коллизия, станции, не поделившие между собой канал, выжидают в течение случайных интервалов времени (используется двоичный экспоненциальный откат такой же, как в Ethernet) и затем снова пытаются отправить кадр.

Другой режим CSMA/CA основан на протоколе MACAW и использует контроль виртуального канала, как показано на рис. 19. Например, станция **A** хочет передать данные станции **B**. Станция **C** также находится в зоне действия **A**, т. е. принимает от неё сигналы, а также возможно в зоне действия **B**. Станция **D** расположена в зоне действия **B**, но не видит станцию **A**.



Рис.19.Прослушивания виртуального канала в протоколе MACAW

Протокол начинает работать тогда, когда станции **A** необходимо передать сообщение **B**. В этом случае **A** посылает станции **B** кадр RTS, запрашивая разрешение на передачу. Если **B** может принять данные, она отсылает обратно положительное подтверждение, кадр CTS. После приёма CTS **A** запускает таймер ACK и начинает передачу данных. В случае корректного приёма **B** генерирует кадр ACK, сообщая станции **A** о конце передачи. Если интервал времени таймера на станции **A** истекает прежде, чем получен ACK от **B**, весь алгоритм работы протокола повторяется сначала.

Рассмотрим этот же процесс с точки зрения станций **C** и **D**. **C** находится в зоне действия **A**, поэтому она также принимает кадр RTS и понимает, что скоро по каналу будут передаваться какие-то данные, при этом откладывает свою передачу данных до окончания активности соседних станций. Исходя из информации, содержащейся в RTS, станция **C** может предположить, сколько времени займёт передача последовательности, включая конечный ACK. В течение этого промежутка **C** считает, что её виртуальный канал занят и необходимо отказаться от передачи. Индикацией такого состояния является последовательность NAV (Network Allocation Vector – вектор выделенной сети). Станция **D** не слышит RTS, посылаемый **A**, зато слышит CTS, посланный станцией **B**, и также выставляет NAV. Особенностью NAV является то, что они не передаются, а являются напоминанием станциям, что необходимо отложить передачу на определённый интервал времени.

Особенностью беспроводных каналов является то, что они гораздо более сильно подвержены влиянию различных помех по сравнению с проводными. Результатом этого является уменьшение вероятности успешной передачи кадра пропорционально увеличению его длины. Если вероятность ошибки в одном бите равна P , то вероятность того, что n -битный кадр будет принят корректно, равна $(1 - P)^n$. Например, для $P = 10^{-4}$ вероятность успешной передачи кадра длиной 12144 бит составляет менее 30%. Если $P = 10^{-6}$, то и в этом случае более 1% кадров будут испорчены. Чем длиннее кадр, тем меньше у него вероятность дойти до получателя неповреждённым, и его надо передавать повторно.

Для решения проблемы зашумлённых каналов в беспроводных сетях применяется разбиение кадров на небольшие фрагменты, каждый из которых содержит свою контрольную сумму. Фрагменты нумеруются, и их успешная передача подтверждается индивидуально с использованием протокола с ожиданием (то есть отправитель не может передать фрагмент с номером $k+1$, пока не получит подтверждение о доставке фрагмента с номером k). Захватив канал с помощью диалога, состоящего из RTS и CTS, отправитель может передать несколько фрагментов подряд, как показано на рис. 20.



Рис. 20.

Фрагментация повышает производительность путём повторной пересылки отдельных коротких фрагментов, в которых произошла ошибка, а не кадров целиком. Таким образом, в режиме DCF (режиме распределённой координации) отсутствует централизованный контроль, и станции самостоятельно борются за эфирное время примерно так же, как в Ethernet.

Кроме режима DCF, беспроводные сети могут работать и в режиме PCF. Здесь базовая станция опрашивает все подчиненные ей станции, выявляя те из них, которые требуют предоставить им канал. Порядок передач полностью и централизованно координируется базовой станцией, поэтому коллизии в режиме PCF исключены. Стандарт лишь предписывает осуществлять такую координацию, но не дает конкретных указаний, касающихся частоты, порядка опросов или наличия либо отсутствия каких-либо приоритетов у отдельных станций. Механизм основан на том, что базовая станция широковещательным способом периодически (10–100 раз в секунду) передает сигнальный кадр. В нем содержатся такие системные параметры, как последовательности смены частот и периоды работы на частотах (для систем, реализующих метод частотных скачков для расширения спектра, – FHSS), данные для синхронизации и т. д. Сигнальный кадр также является приглашением для новых станций, которые желают войти в список опрашиваемых станций. Попав в этот список, станция получает гарантированную долю пропускной способности (при оп-

ределенных параметрах скорости), то есть ей гарантируется качество обслуживания.

Режимы PCF и DCF могут сосуществовать даже внутри одной соты сети. Это делается путем точного определения межкадрового интервала. После отправки кадра необходимо какое-то время простаивать, прежде чем какая-либо станция получит разрешение послать кадр. Всего определено четыре интервала, каждый из которых имеет собственное предназначение. Они изображены на рис. 21.

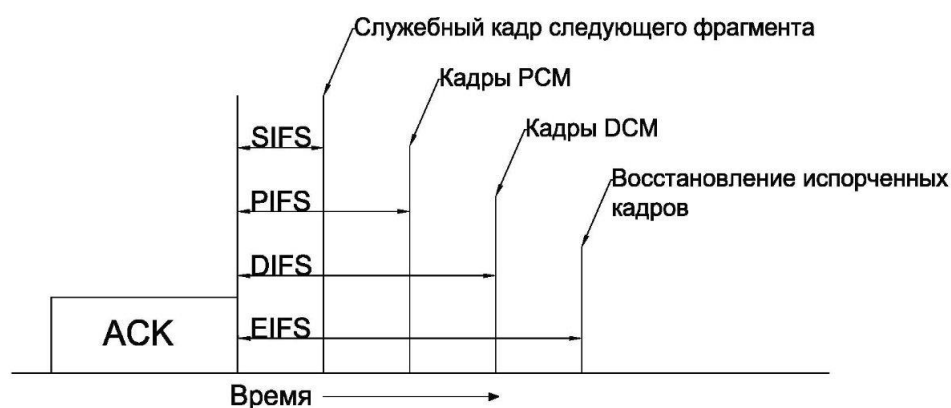


Рис. 21. Межкадровые интервалы в стандарте 802.11

Самый короткий интервал – это SIFS (Short Inter Frame Interval – короткий межкадровый интервал). Он используется для того, чтобы одна из сторон, ведущих диалог с помощью управляющих кадров, могла получить шанс начать первой. Здесь может быть CTS, посылаемый приемником в ответ на запрос RTS; ACK, посылаемый им же после окончания приема фрагмента или целого кадра; очередная часть пакета фрагментов, посылаемая отправителем (то есть он не посылает RTS после каждого фрагмента).

После интервала SIFS ответить может всегда только одна станция. Если она упускает свой шанс и время PIFS (Priority Inter Frame Spacing – межкадровый интервал PCF) истекает, то базовая станция может послать сигнальный кадр или кадр опроса. Этот механизм позволяет станции, посылающей кадр данных или последовательность фрагментов, закончить свою передачу без какого-либо вмешательства со стороны соседей, но дает и базовой станции возможность после окончания передачи станцией захватить канал, не борясь за него с другими желающими.

Если базовой станции нечего сказать и интервал DIFS (DCF Inter Frame Spac-Hig – межкадровый интервал DCF) истекает, то любая станция может попытаться захватить канал. Применяются при этом обычные правила борьбы, включая двоичный экспоненциальный откат в случае коллизии.

Последний временной интервал называется EIFS (Extended Inter Frame Spacing – расширенный межкадровый интервал). Он используется только той станцией, которая только что получила испорченный или неопознанный кадр и хочет сообщить об этом факте. Почему наиболее низкий приоритет отдан именно этому событию? Дело в том, что принявшая такой кадр станция может сразу не определить, что происходит, и ей нужно выждать некоторое время (интервал EIFS) до начала передачи сигнала коллизии, чтобы не прервать идущий в это время диалог между другими станциями. Для реализации этих режимов в структуре кадра стандарта 802.11, в отличие от кадра стандарта 802.3 (рис. 15), наряду со стандартными полями, такими как адреса отправителя, получателя, контрольной суммы и др., в заголовке кадров стандарта 802.11 появляются дополнительные поля.

Стандарт 802.11 определяет три класса кадров, передаваемых по каналу – информационные, служебные и управляющие. Все они имеют заголовки с множеством полей, используемых подуровнем MAC. Кроме того, в кадрах появились поля, используемые физическим уровнем, но они в основном относятся к методам модуляции, поэтому здесь они не рассмотрены. Формат информационного кадра стандарта 802.11 показан на рис. 22.

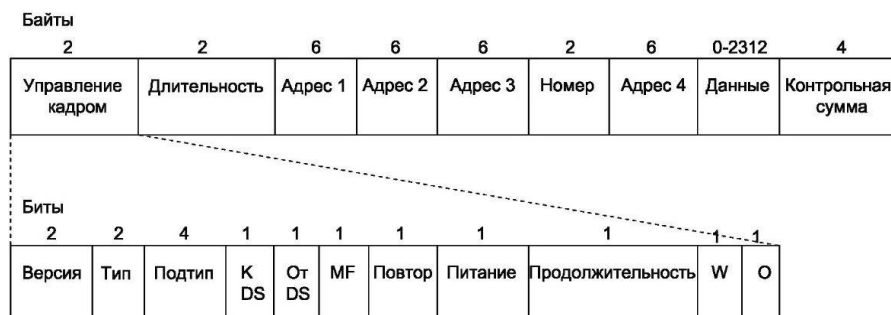


Рис. 22. Информационный кадр стандарта 802.11

Вначале идет поле **Управление кадром** (Frame Control). Оно содержит 11 вложенных полей. Первое из них – **Версия протокола**, именно оно позволяет двум протоколам работать одновременно в одной ячейке сети. Затем следуют поля **Тип** (информационный, служебный или управляющий) и **Подтип** (например, RTS или CTS). Биты **К DS** и **От DS** говорят о направлении движения кадра – к межсотовой системе распределения (например, Ethernet) или от нее. Бит **MF** говорит о том, что далее следует еще один фрагмент. Бит **Повтор** маркирует повторно посылаемый фрагмент. Бит **Управление питанием** используется базовой станцией для переключения станции в режим пониженного потребления или выхода из этого режима. Бит **Продолжительность** говорит о том, что вообще-то у отправителя имеются еще кадры для пересылки. Бит **W** является индикатором использования шифрования в теле кадра по алгоритму WEP (Wired Equivalent Protocol – протокол обеспечения конфиденциальности). Наконец, бит **O** говорит принимающей станции о том, что кадры с этим битом должны обрабатываться строго по порядку.

Второе основное поле информационного кадра – это поле **Длительность**. В нем задается время, которое будет потрачено на передачу кадра и получения подтверждения её успешного завершения. Это поле можно найти и в служебных кадрах, и именно в соответствии с ним станции выставляют признаки NAV. Заголовок кадра содержит уже четыре адреса в формате, соответствующем стандарту IEEE 802. Наряду с адресами отправителя и получателя, так как кадры могут входить в ячейку или покидать ее через базовую станцию, дополнительно в структуру кадра включены два адреса, которые как раз и хранят адреса исходной и целевой ячеек при передаче между ними сообщений.

Поле **Номер** позволяет нумеровать фрагменты. Из 16 доступных бит 12 идентифицируют кадр, а 4 – фрагмент. Поле **Данные** содержит передаваемую по каналу информацию, его длина может достигать 2312 байт. В конце, как обычно, расположено поле **Контрольная сумма**.

Управляющие кадры имеют формат, сходный с форматом информационных кадров, за одним исключением нескольких моментов: в управляющем кадре отсутствуют поля базовых станций, поскольку таким кадрам незачем выходить за пределы соты.

Служебные кадры гораздо короче: в них содержится один или два адреса, отсутствуют поля *Данные* и *Номер*. Ключевой здесь является информация, содержащаяся в поле *Подтип*. Значениями его обычно являются RTS, CTS или ACK.

Заключение

Как показало время, технологии информационных сетей, в основе которых лежат различные варианты метода случайного доступа, одержали уверенную победу над остальными. Основой её явились: высокая производительность, минимальное время доступа к сетевым ресурсам, относительная простота реализации на аппаратном уровне и, как следствие, невысокая стоимость сетевых устройств. Дальнейшее развитие элементной базы и технологий её производства, сетевых протоколов (в том числе реализации режима Q&S – Quality of Service и др.) позволили в сетях, построенных на основе метода случайного доступа, передавать в реальном времени все виды трафика, в том числе и чувствительные к временным задержкам, такие как речь и видео. Это даёт прекрасную возможность широко использовать технологии случайного доступа и в локальных, и при построении городских сетей, в том числе и беспроводных.

Рекомендуемая литература

1. Олифер, В. Г. Компьютерные сети: Принципы, технологии, протоколы: учебник для вузов / В. Г. Олифер, Н. А. Олифер. – 3-е изд. – СПб.: Питер, 2009. – 958 с.
2. Таненбаум, Э. Компьютерные сети / Э. Таненбаум. – 4-е изд. – СПб.: Питер, 2009. – 992 с.
3. Корпорация CiscoSystems, Inc. Программа сетевой академии CiscoCCNA 1и 2: Вспомогательное руководство. – 3-е изд., с испр.: пер. с англ. – М.: Вильямс, 2005. – 1168 с.
4. Корпорация Cisco Systems, Inc. Программа сетевой академии Cisco CCNA 3и 4: Вспомогательное руководство: пер. с англ. – М.: Вильямс, 2007. – 994 с.

5. Шварц, М. Сети связи: протоколы, моделирование и анализ: в 2 ч./ М. Шварц. – Ч. 1. – М.: Наука; Гл. ред. физ.-мат. лит., 1992. – 336 с.

6. Шварц, М. Сети связи: протоколы, моделирование и анализ: в 2 ч./ М. Шварц. – Ч. 2. – М.: Наука; Гл. ред. физ.-мат. лит., 1992. – 368 с.

7. Вишневский, В. М. Широкополосные беспроводные сети передачи информации/ В. М. Вишневский и др. М.: Техносфера, 2005 – 592 с.

8. Григорьев, В. А. Сети и системы радиодоступа/ В. А. Григорьев О. И. Лагутенко, Ю. А. Распаев. – М: Эко-Трендз, 2005. – 384 с.

9. Закер, К. Компьютерные сети. Модернизация. Поиск неисправностей / К. Закер. –СПб.: БХВ – Петербург, 2005. – 987 с.

Оглавление

I. Метод «ALOHA».....	3
II. Метод дискретная «ALOHA»	7
III. Метод множественного доступа с проверкой несущей и обнаружением столкновений (CSMA/CD– carriersensmultipleaccesswithcollisiondetection)	11
IV. Метод множественного доступа с проверкой несущей и предотвращением столкновений (CSMA/CA– carriersensmultipleaccesswithcollisionavoidance)	26
Заключение.....	34
Рекомендуемая литература	34

Учебное издание

Захаров Александр Сергеевич

Метод случайного доступа

Часть 1

Методические указания

Редактор, корректор М. В. Никулина
Правка, верстка М. В. Никулина

Подписано в печать 04.12.2012. Формат 60×84¹/₁₆.
Усл. печ. л. 2,32. Уч.-изд. л. 2,0. Тираж 20 экз. Заказ

Оригинал-макет подготовлен
в редакционно-издательском отделе ЯрГУ.

Ярославский государственный университет
им. П. Г. Демидова.
150000, Ярославль, ул. Советская, 14.