



МАТЕМАТИЧЕСКАЯ МОДЕЛЬ АДАПТИВНОГО УПРАВЛЕНИЯ ТРАФИКОМ В ПОДВИЖНЫХ СЕТЯХ С ПЕРЕМЕННЫМ КАЧЕСТВОМ СИГНАЛА

А. А. Есин

АО «МирВайФай, г. Москва

✉ anton.esin@imm.am

Аннотация. Рассматривается задача математического моделирования процесса передачи трафика с применением методов агрегации в подвижных сетях в условиях высокой мобильности абонентов и пространственно неоднородного покрытия, включая зоны деградации сигнала (так называемые мёртвые зоны). Для повышения надёжности и устойчивости передачи данных используется агрегация трафика на уровне канала. Для исследования эффективности алгоритмов агрегации и их адаптации к параметрам работы сети и скорости перемещения абонентов предложена марковская модель канала связи. Модель основана на периодическом аффинном движении мобильного устройства между базовыми станциями, равномерно расположенными на прямой. В рамках модели вводятся понятия зоны устойчивого покрытия и переходной зоны, определяемые расстоянием до ближайшей и следующей базовой станции. Состояние канала описывается марковской цепью с состояниями, соответствующими дискретизации уровня сигнала: стабильное соединение, ухудшенное соединение и разрыв; переходы между состояниями моделируются марковским процессом с постоянными интенсивностями, параметры которого определяются на основе данных реальных сетей. Рассмотрено также обобщение на случай временной зависимости состояний канала, приводящее к полумарковской модели. Для обоих случаев получены аналитические выражения для стационарных вероятностей состояний и сформулированы условия устойчивости системы, гарантирующие ограниченность очередей. Предлагается также модель для адаптивного управления пропускной способностью сети, оптимизирующая параметры передачи трафика в зависимости от текущего состояния канала, очереди запросов и скорости абонента. Эффективность предложенной модели подтверждается численным моделированием, демонстрирующим стабильную работу сети в широком диапазоне скоростей и параметров покрытия. Разработанная модель может применяться для оценки надёжности и оптимизации сетевых протоколов в условиях высокой мобильности абонентов, включая высокоскоростной железнодорожный транспорт, автомобильные сети и мобильные платформы.

Ключевые слова: мобильные сети, марковская модель, полумарковский процесс, устойчивость системы, преобразование Лапласа.

ВВЕДЕНИЕ

Современные системы связи, обеспечивающие устойчивую передачу данных от мобильных объектов, таких как высокоскоростные поезда, беспилотные транспортные средства, дроны и автомобили, активно используют инфраструктуру сотовых сетей общего пользования. Однако вне зон плотной застройки качество связи существенно снижается вследствие разреженности расположе-

ния базовых станций (БС), неполного покрытия и нестабильности радиоканала. Это проявляется в ослаблении сигнала, разрывах соединений и значительных задержках передачи данных.

При этом требования к качеству обслуживания (англ. *Quality of Service*, QoS) для абонентов в высокомобильной среде не уступают запросам пользователей в условиях стабильного покрытия: требуется устойчивая передача объёмного мультимедийного трафика, поддержка потокового видео и

передача данных в реальном времени с подвижных платформ. Такие сценарии особенно актуальны, например, при видеонаблюдении, телеметрии или автоматизированном управлении подвижными объектами.

Агрегация трафика является эффективным способом решения данных проблем [1, 2]. Существующие методы агрегации трафика (например, Link Aggregation Control Protocol (LACP), Multilink PPP) применяются в архитектурах LTE и LTE Advanced (LTE-A). В настоящей работе основное внимание уделяется сценариям, характерным для LTE и LTE-A, где агрегация реализуется на уровне протокола Packet Data Convergence Protocol (PDCP) и через механизмы Carrier Aggregation (CA). Модель может быть также адаптирована для сценариев 5G, однако специфика NR-архитектуры (New Radio) в данной работе не рассматривается подробно. Отметим, что данные методы демонстрируют высокую эффективность в сетях с устойчивым или умеренным покрытием, однако эффективность существенно падает при резких изменениях качества сигнала и высокой скорости движения абонентов, которая требует частого переключения между БС [2].

Анализ типовых сценариев функционирования сетей в условиях высокой мобильности абонентов позволяет выделить несколько фундаментальных проблем:

- Неполное покрытие сети. Вне городских территорий наблюдаются значительные мёртвые зоны, в которых уровень сигнала недостаточен для устойчивого соединения, что снижает общую пропускную способность сети и увеличивает потери пакетов.

- Динамика сигнала. Качество канала связи существенно зависит от траектории и скорости движения абонента. Доплеровский сдвиг, многолучевое распространение и неоднородность покрытия приводят к нестабильности канала и частым разрывам соединений.

- Неоптимальное поведение при потере связи. В случае разрыва соединения или ослабления сигнала инициируются многократные попытки повторной передачи пакетов, что приводит к перегрузке как канала, так и сетевой инфраструктуры, снижает общую эффективность передачи и увеличивает энергопотребление.

Существующие методы агрегации и управления мобильным трафиком недостаточно эффективно

справляются с перечисленными проблемами. Основные их недостатки таковы:

- Игнорирование временной корреляции. Большинство моделей предполагает независимые переходы между состояниями сети, что не учитывает инерционность оборудования и физические закономерности распространения сигнала.

- Отсутствие адаптации к топологии сети и скорости движения абонентов. В ряде подходов предполагается стационарная или квазистационарная среда, что делает их неприменимыми при высокой скорости движения и динамически изменяющемся качестве канала.

- Высокая латентность при переключении. Алгоритмы переключения между базовыми станциями могут приводить к длительным паузам в передаче трафика, особенно в условиях нестабильного сигнала.

Очевидным решением является улучшение существующих методов агрегации. Однако для оптимизации работы сети и поиска наилучшего решения требуется разработка соответствующих математических моделей. При этом существующие модели, как правило, либо являются упрощёнными, либо не учитывают ряд ключевых параметров реальных систем, таких как изменяющееся положение серверов и абонентов в пространстве [3, 4], временные корреляции в изменении состояния сети, динамическое перераспределение нагрузки, адаптивная буферизация данных; либо же учитывают пространственную геометрию элементов сети, но не учитывают сложную динамику очередей [5–7]; либо же анализируют исключительно проблемы связности компонентов сети, используя теорию графов [8] или методы теории перколяции [9–12], добавляя изменчивость состояний узлов во времени [13] и адаптируя классические методы статистической физики к задачам анализа беспроводных сетей [14, 15].

Цель данной работы – разработка новой модели управления трафиком в мобильных сетях, учитывающей стохастическую изменчивость пропускной способности каналов, временную корреляцию состояний сети и адаптивное распределение нагрузки для повышения устойчивости соединений и снижения задержек. В данной статье представлена стохастическая модель передачи агрегированного трафика, в которой характеристики канала описываются конечной марковской цепью, полученной путём дискретизации непрерывного сигнального



уровня [16]. Марковская модель оправдана ввиду экспоненциального характера статистики переходов между состояниями в реальных измерениях (например, при анализе сценариев переключения между базовыми станциями в LTE) [17].

В отличие от существующих моделей, предполагающих либо стационарность состояния канала, либо независимость переходов между состояниями, в данной работе предложена стохастическая модель, включающая временную корреляцию и пространственную структуру покрытия с учётом движения абонента между базовыми станциями по аффинной траектории. Основной вклад работы заключается:

- в построении марковской и полумарковской моделей канала связи на основе дискретизации уровня сигнала,
- в аналитическом выводе стационарных распределений состояний и условиях устойчивости очередей,
- в разработке адаптивной модели управления трафиком, оптимизирующей пропускную способность с учётом текущего состояния канала, скорости движения и длины очереди.

Работа имеет следующую структуру. В § 1 излагается процесс построения модели канала и представлена математическая модель адаптивного управления трафиком, выводятся основные уравнения модели, описывается марковский процесс переходов между состояниями и анализируются условия устойчивости системы. Формулировки и доказательства основных результатов даны в § 2. Здесь приводятся явные аналитические формулы для стационарного распределения, формулируется условие устойчивости модели. В § 3 представлены результаты численного моделирования предложенной модели в различных сценариях работы сети и осуществляется ее сравнение с существующими методами агрегации трафика. В заключении формулируются основные выводы по результатам исследования, оценивается эффективность предложенной модели и приводятся направления дальнейших исследований.

Настоящая работа ориентирована на архитектуры LTE и LTE-A, в которых агрегация трафика осуществляется на уровне PDCP и через механизмы Carrier Aggregation (CA). Рассматриваемая модель применима к сценариям с высокоскоростным перемещением (например, железнодорожного транспорта), характерным для данных технологий. Хотя в 5G NR применяются более сложные схемы

агрегации (включая Dual Connectivity и Service Data Adaptation Protocol (SDAP)), они выходят за рамки данной статьи и могут быть предметом последующего обобщения. Основное внимание сосредоточено на формализованном описании стохастической структуры радиоканала и адаптивного управления пропускной способностью в рамках LTE/LTE-A.

1. ПОСТРОЕНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Мобильную сеть будем рассматривать как совокупность базовых станций и мобильных абонентов, перемещающихся по маршрутам с различным уровнем покрытия сигнала. При этом под мобильным абонентом будем понимать единый агрегатор или сервер, формирующий объединённый трафик от множества устройств на борту высокоскоростного поезда или автомобиля. Практическая реализация такой схемы получила распространение в современных сетях 4G/5G (примером могут служить мультирелейные системы, в рамках которых поезд или автобус имеет единую точку доступа к сотовой сети) [1]. Это позволяет агрегировать трафик множества устройств в один канал и упростить процесс переключения между базовыми станциями, в то время как отдельные устройства внутри вагона подключаются к этому «серверу» [1, 18, 19]. В контексте LTE- и LTE-A-архитектур агрегатор реализуется в виде абонентского терминала или мультирелейной платформы, объединяющей трафик с борта транспортного средства и взаимодействующей с несколькими компонентными сетями (CA).

Будем предполагать, что покрытие сети неплотное, т. е. БС располагаются на таком удалении друг от друга, чтобы имели место разрывы в покрытии, а скорость движения абонентов достаточно высокая, чтобы для передачи необходимого объема трафика нужно было задействовать последовательно несколько БС.

1.1. Предположения модели

Топологию сети естественно моделировать в виде ориентированного графа $G = (V, E)$ [3, 5, 8], где V – множество узлов сети (например, базовых станций); $E \subseteq V \times V$ – множество ориентированных рёбер, представляющих возможные маршруты передачи данных между узлами.

Сделаем несколько предположений относительно рассматриваемой модели, которые существенно упрощают анализ, но при этом не ограничивают общность подхода и основные качественные выводы.

• **Граф базовых станций является линейным:** будем предполагать, что БС расположены вдоль прямой линии в точках с координатами

$$x_j = j \Delta x, \quad j \in \mathbb{Z}_+,$$

где Δx – фиксированное расстояние между соседними БС. Другими словами, предполагается, что БС расположены равномерно в узлах одномерной решётки $\mathbb{Z}_+$. Такое предположение отражает типичную ситуацию покрытия вдоль транспортных магистралей или железных дорог.

Замечание. Следует подчеркнуть, что рассматриваемая линейная расстановка БС с равномерным шагом вдоль прямой является идеализированным допущением, характерным в первую очередь для магистральных сценариев (железнодорожных, автомагистралей и т. д.). В реальных системах могут встречаться вариации во взаимном расположении БС, участки трассы с кривизной, зоны неоднородного рельефа (туннели, мосты и пр.). Тем не менее, линейная модель даёт удобный базовый каркас для анализа основных закономерностей, особенно при движении по относительно прямым участкам пути. Это упрощение не влияет на общую логику методов и выводы, но при необходимости может быть уточнено.

• **Движение агрегатора равномерно:** агрегатор перемещается вдоль указанной прямой линии с постоянной скоростью:

$$v(t) = v_0, \quad v_0 > 0.$$

Это предположение оправдано при моделировании стандартных сценариев (например, движения высокоскоростных поездов или автомобилей на автомагистралях), в которых изменения скорости относительно малы и агрегатор движется стабильно.

Замечание. В работе предполагается постоянная скорость агрегатора $v(t) = v_0$ (или нормированная единичная скорость). Такой выбор оправдан стационарностью условий высокоскоростного транспорта на длинных магистралях. При этом реальные сценарии могут включать участки разгона (торможения), а также колебания скорости. Если такие эффекты существенны, модель можно расширить. В данном же исследовании сосредоточимся на базовом сценарии с постоянной или почти постоянной скоростью, что обеспечивает перио-

дичность движения и упрощает дальнейшие аналитические выкладки без потери общности.

• **Влияние БС локализовано:** качество соединения в каждой точке траектории агрегатора определяется только двумя ближайшими БС – текущей (позади агрегатора) и следующей (перед агрегатором). Такое упрощение обосновано быстротой затухания сигнала с расстоянием и позволяет пренебречь влиянием дальних БС без существенной потери точности.

Таким образом, при условии равномерного расположения базовых станций и равномерного движения агрегатора процесс перемещения агрегатора можно моделировать как периодическое движение точки по полуинтервалу $[0, 1)$. Можно считать также, что при достижении агрегатором положения 1 происходит смена текущей и следующей базовых станций, и агрегатор снова начинает движение из точки 0. Такая интерпретация существенно упрощает математический анализ и численное моделирование динамики сети.

С учётом этих предположений состояние агрегатора в любой момент времени может быть полностью параметризовано его положением на аффинном полуинтервале $[0, 1)$, а исследование модели можно ограничить периодом прохождения данного расстояния.

Таким образом, при условии равномерного расположения базовых станций и равномерного движения агрегатора его движение можно строго описать следующим образом.

Обозначим через $T = \frac{\Delta x}{v_0}$ время прохождения

агрегатором расстояния Δx между двумя соседними БС.

Тогда положение агрегатора в любой момент времени t моделируется с помощью периодического аффинного параметра $s(t) \in [0, 1)$, задаваемого формулой

$$s(t) = \frac{t \bmod T}{T},$$

где $s(t) = 0$ соответствует положению агрегатора ровно в точке текущей базовой станции, а $s(t) = 1$ – достижению следующей БС.

При достижении агрегатором положения $s(t) = 1$ он мгновенно начинает движение на новом отрезке, где следующая базовая станция становится текущей, а новая соседняя базовая станция – следующей.

Таким образом, движение агрегатора по сети представляет собой периодический процесс, кото-



рый в рамках модели можно представить как повторяющееся движение точки по полуинтервалу $[0, 1]$:

$$x(t) = s(t).$$

Важным частным случаем, который позволяет максимально упростить модель и последующий анализ, является выбор единичной скорости агрегатора $v_0 = 1$ и единичного расстояния между базовыми станциями $\Delta x = 1$.

В этом случае время прохождения между соседними БС становится единичным ($T = 1$), и положение агрегатора $x(t)$ на полуинтервале $[0, 1]$ задаётся простой периодической зависимостью

$$x(t) = t \bmod 1.$$

Таким образом, агрегатор перемещается по полуинтервалу $[0, 1]$ с постоянной единичной скоростью и, достигая точки 1, мгновенно переходит к следующему полуинтервалу, начиная вновь с позиции 0. Графически эта функция представляет собой периодическую пилообразную зависимость с единичным периодом.

Таким образом, положение агрегатора на полуинтервале $[0, 1]$ полностью определяет его удалённость от двух соседних базовых станций.

Ввиду периодичности модели для анализа свойств системы достаточно рассмотреть её поведение на одном фиксированном периоде $t \in [0, 1]$. Результаты на остальных временных интервалах будут аналогичны вследствие указанной периодичности.

1.2. Динамика расстояний до базовых станций и качества сигнала с учётом топологии сети

Расстояния от агрегатора до текущей (позади него) и следующей (впереди него) базовых станций при положении $x(t) \in [0, 1]$ задаются следующими формулами:

$$d_{\text{current}}(t) = x(t), d_{\text{next}}(t) = 1 - x(t), 0 \leq x(t) \leq 0,5;$$

$$d_{\text{current}}(t) = 1 - x(t), d_{\text{next}}(t) = x(t), 0,5 < x(t) \leq 1.$$

Из этих формул явно следует, что:

- в начале периода ($x(t) = 0$) агрегатор находится непосредственно на текущей базовой станции (расстояние до неё равно нулю), а до следующей станции расстояние максимально (равно единице);
- в середине периода ($x(t) = 0,5$) агрегатор находится ровно посередине между двумя соседними базовыми станциями, и расстояния до них равны 0,5;

- в конце периода ($x(t) = 1$) агрегатор достигает следующей базовой станции, после чего происходит мгновенный переход на новый отрезок, где следующая базовая станция становится текущей, и положение агрегатора снова становится равным нулю.

Таким образом, за один период агрегатор сначала удаляется от текущей базовой станции (при движении от 0 до 0,5), а затем приближается к следующей станции (при движении от 0,5 до 1). В результате расстояние до ближайшей базовой станции за один период сначала монотонно возрастает от 0 до 0,5, а затем монотонно убывает обратно до нуля.

Дополнительно предположим, что каждая базовая станция характеризуется конечным радиусом действия R , в пределах которого агрегатор получает устойчивый сигнал (состояние S_1). Кроме того, за пределами радиуса R существует небольшая переходная зона шириной δ , в которой сигнал ослаблен (состояние S_2).

Таким образом, для каждой базовой станции выделяются три зоны покрытия:

- зона устойчивого сигнала (состояние S_1): $0 \leq d \leq R$,
- переходная зона ослабленного сигнала (состояние S_2): $R < d \leq R + \delta$,
- мёртвая зона (состояние S_3): $d > R + \delta$, где d – текущее расстояние до базовой станции

$$d = d(t) = \min \{d_{\text{current}}(t), d_{\text{next}}(t)\}.$$

Предполагается, что выполнено следующее естественное условие:

$$2(R + \delta) < 1,$$

т. е. суммарный диаметр зон покрытия одной базовой станции (зона устойчивого сигнала плюс переходная зона с обеих сторон) существенно меньше расстояния между соседними базовыми станциями. Это гарантирует наличие мёртвых зон между зонами покрытия соседних БС, где сигнал полностью отсутствует.

При движении агрегатора по полуинтервалу $[0, 1]$ агрегатор последовательно проходит:

- через зону устойчивого сигнала текущей станции,
- через переходную зону ослабленного сигнала текущей станции,
- через мёртвую зону,
- через переходную зону ослабленного сигнала следующей станции,
- через зону устойчивого сигнала следующей станции.

В общем случае влияние скорости движения и расстояния до ближайшей базовой станции на качество соединения моделируется специальной функцией потерь, объединяющей линейные и нелинейные эффекты [20]:

$$L(d(t), v(t)) = L_0 + a d(t)^\gamma + k_1 v(t) + k_2 \log(1 + v(t)) + k_3 v(t)^\alpha,$$

где L_0 – базовые потери при идеальном случае (нулевое расстояние и нулевая скорость); $a d(t)^\gamma$ – компонент затухания сигнала, зависящий от расстояния $d(t)$ до базовой станции (степенная модель распространения сигнала, обычно $\gamma \in [2, 4]$; $k_1 v(t)$ – линейный компонент, связанный с доплеровским сдвигом частоты; $k_2 \log(1 + v(t))$ – логарифмическое затухание, вызванное многолучевыми замираниями (*fading*); $k_3 v(t)^\alpha$ – нелинейные эффекты, возникающие при переключении между базовыми станциями и высоких скоростях ($\alpha > 1$).

Таким образом, представленная модель одновременно учитывает влияние обоих ключевых факторов – удалённости от базовой станции и скорости движения агрегатора – что делает её физически адекватной и практически полезной для дальнейшего анализа.

Линейная модель $L(v) = L_0 + a d(t)^\gamma + k_1 v(t)$ сохраняется как частный случай ($k_2 = k_3 = 0$), обеспечивая аналитическую простоту для базового анализа [20], и в рамках данной работы будет рассматриваться именно данная модель.

1.3. Процесс переходов между дискретными уровнями сигнала в общем случае

В рамках предложенной модели агрегатор последовательно проходит три зоны с различным качеством сигнала (устойчивый сигнал, ослабленный сигнал и мёртвая зона). В реальных сетях переходы между состояниями соединения (S_1, S_2, S_3) происходят не мгновенно, а характеризуются некоторой задержкой и инерционностью. Основными причинами таких эффектов являются [20]:

- Постоянство внешних условий: параметры канала связи остаются относительно стабильными на протяжении некоторого времени, поэтому переходы между состояниями происходят постепенно, а не резко.

- Инерционность оборудования: переключение между базовыми станциями требует дополнительного времени на согласование и обработку сигналов.

- Физическая траектория движения: при движении по определённой траектории переходы между состояниями связаны с топологией сети и регулярностью расположения базовых станций.

Эти эффекты создают временную корреляцию между состояниями канала, что означает зависимость вероятностей переходов не только от текущего состояния, но и от времени пребывания в нём. Строгое математическое описание таких ситуаций приводит к необходимости рассмотрения более общего класса процессов, называемых полумарковскими (англ. *Semi-Markov Processes*, SMP) [21, 22].

Рассмотрим случайный процесс $X(t)$, моделирующий изменение уровня сигнала в мобильной сети. Этот процесс представляет собой полумарковскую модель с тремя состояниями: S_1 (устойчивый сигнал), S_2 (ослабленный сигнал), S_3 (мёртвая зона). Моменты переходов образуют возрастающую последовательность $\{\tau_n\}$ случайных величин, при этом $X(t)$ сохраняет текущее состояние между переходами [23].

Полумарковская модель уровней сигнала задаётся следующей системой интегро-дифференциальных уравнений для вероятностей состояний $p_i(t) = P\{X(t) = S_i\}$, $i = 1, 2, 3$:

$$\begin{aligned} \frac{dp_i(t)}{dt} = & \sum_{j \neq i} \int_0^t p_j(\tau) \alpha_{ji}(t - \tau) d\tau - \\ & - \sum_{k \neq i} \int_0^t p_i(\tau) \alpha_{ik}(t - \tau) d\tau, \end{aligned} \quad (1)$$

где интенсивность перехода задаётся функцией

$$\begin{aligned} \alpha_{ij}(t) = & \lambda_{ij}(t) \exp\left(-\int_0^t \lambda_i(u) du\right), \\ \lambda_i(t) = & \sum_{k \neq i} \lambda_{ik}(t), \quad i = 1, 2, 3, \end{aligned}$$

Будем полагать, что заданы фиксированные начальные условия $p_1(0) = 1, p_2(0) = p_3(0) = 0$.

Для решения системы (1) удобно применять преобразование Лапласа [24]. При этом в пространстве Лапласа получаем систему алгебраических уравнений



$$\begin{aligned} & s\tilde{p}_i(s) - p_i(0) = \\ & = \sum_{j \neq i} \tilde{p}_j(s) \tilde{\alpha}_{ji}(s) - \sum_{k \neq i} \tilde{p}_i(s) \tilde{\alpha}_{ik}(s). \end{aligned}$$

В общем случае времена пребывания могут иметь произвольные распределения $F_i(t)$: $T_i \sim F_i(t)$. Тогда система уравнений (1) становится достаточно сложной для аналитического решения. В такой ситуации можно применять численные схемы по временной сетке $\{t_n\}$, например:

$$\begin{aligned} p_i(t_{n+1}) = p_i(t_n) + \Delta t \left[\sum_{j \neq i} \sum_{m=0}^n p_j(t_m) \alpha_{ji}(t_{n+1} - t_m) - \right. \\ \left. - \sum_{k \neq i} \sum_{m=0}^n p_i(t_m) \alpha_{ik}(t_{n+1} - t_m) \right]. \end{aligned}$$

При более сложных распределениях (гамма-, Вейбулла и т. п.) такой прямой численный подход позволяет эффективно получать решения.

На практике часто используются следующие распределения времён пребывания в состояниях: гамма- (Эрланга), Вейбулла и равномерное распределения. Ниже приведены распределения времён пребывания и соответствующие выражения для интенсивностей переходов $\alpha_{ij}(t) = \lambda_{ij}(t) \times \exp\left(-\int_0^t \lambda_i(u) du\right)$, а также их преобразования Лапласа для данных распределений $\tilde{\alpha}_{ij}(s) = \mathcal{L}\{\alpha_{ij}(t)\}$:

• Гамма-распределение (распределение Эрланга):

Функция интенсивности, интенсивность перехода в соседнее состояние и преобразование Лапласа соответственно имеют следующий вид:

$$\lambda_i(t) = \frac{\lambda_i^k t^{k-1} e^{-\lambda_i t}}{(k-1)! - \sum_{m=0}^{k-1} \frac{(\lambda_i t)^m}{m!}},$$

$$\alpha_{ij}(t) = p_{ij} \lambda_i(t), \quad \tilde{\alpha}_{ij}(s) = p_{ij} \frac{\lambda_i^k}{(s + \lambda_i)^k}.$$

• Распределение Вейбулла:

$$\lambda_i(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1}, \quad \alpha_{ij}(t) = p_{ij} \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\beta\right],$$

$$\tilde{\alpha}_{ij}(s) = p_{ij} \frac{\beta}{\alpha^\beta} \int_0^\infty t^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\beta - st\right] dt.$$

• Равномерное распределение на отрезке $[a, b]$:

$$\lambda_i(t) = \begin{cases} 0, & t < a, \\ \frac{1}{b-t}, & a \leq t < b, \\ \infty, & t = b, \end{cases}$$

$$\alpha_{ij}(t) = \begin{cases} 0, & t < a, \\ \frac{p_{ij}}{b-a}, & a \leq t \leq b, \\ 0, & t > b, \end{cases}$$

$$\tilde{\alpha}_{ij}(s) = p_{ij} \frac{e^{-as} - e^{-bs}}{(b-a)s}.$$

Подстановка вышеописанных параметров в систему уравнений (1) позволяет получить явные выражения для вероятностных характеристик модели при различных распределениях времени пребывания в состояниях, а также вывести важнейшие метрики качества функционирования системы.

Однако аналитическое исследование таких моделей в общем случае приводит к громоздким выражениям, затрудняющим их интерпретацию и применение. Поэтому в дальнейшем ограничимся рассмотрением марковского случая, предполагающего экспоненциальное распределение времени пребывания в каждом состоянии:

$$\mathbb{P}(T_i > t) = e^{-\lambda_i t},$$

что соответствует отсутствию памяти и существенно упрощает анализ. Такое допущение означает, что временные корреляции и инерционность каналов считаются усреднёнными и аппроксимируются постоянными интенсивностями переходов между состояниями.

Это упрощение позволяет получить строго аналитические и легко интерпретируемые результаты, которые и будут изложены в § 2 как основные аналитические следствия модели.

2. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

2.1. Механизм оптимального управления

Для адаптивной корректировки параметров передачи трафика (формально) введём функцию $u(t)$, зависящую от текущего состояния сети:

$$u(t) = \begin{cases} \mu_{\text{norm}}, & \text{если } X(t) = S_1, \\ \mu_{\text{slow}} g(\tau_2, \theta), & \text{если } X(t) = S_2, \\ 0, & \text{если } X(t) = S_3, \end{cases}$$

где $g(\tau_2, \theta)$ – адаптивная функция, зависящая от времени пребывания в состоянии S_2 и параметров θ , подлежащих оптимизации.

Для оптимизации работы системы формулируется целевая функция

$$J(u) = \int_0^T [\omega_1 L_{\text{pack}}(u(t)) + \omega_2 D(u(t))] dt, \quad (2)$$

где $L_{\text{pack}}(u(t))$ – функция, отражающая потери пакетов в зависимости от выбранной стратегии $u(t)$; $D(u(t))$ – функция задержек передачи данных; ω_1 и ω_2 – коэффициенты важности соответствующих критериев; T – рассматриваемый временной интервал.

Оптимальная стратегия $u^*(t)$ достигается путём минимизации целевой функции (2). В качестве метода оптимизации можно использовать, например, градиентный спуск или методы динамического программирования [25, 26]. Для градиентного спуска по параметрам θ корректировка осуществляется по хорошо известному алгоритму [25, 27]

$$\theta_{n+1} = \theta_n - \eta \nabla_{\theta} J(u),$$

где η – параметр алгоритма, а $\nabla_{\theta} J(u)$ – градиент целевой функции по параметрам θ . Подобные подходы широко применяются в задачах калибровки систем связи и управления потоками в сетях.

Кроме управления скоростью передачи, функция $u(t)$ может влиять и на переходные интенсивности между состояниями сети. Например, можно модифицировать интенсивности так, чтобы уменьшить вероятность разрыва связи или ускорить восстановление:

$$\lambda_{ij}(v, u(t)) = \lambda_{ij}^0 h(u(t)) e^{-k_v v(t)}, \quad (3)$$

где λ_{ij}^0 – базовые интенсивности; функция $h(u(t))$ описывает корректирующий эффект управления, а $e^{-k_v v(t)}$ отражает влияние скорости движения $v(t)$ агрегатора (при необходимости). Подобные поправки позволяют моделировать физические и протокольные механизмы, при которых, например, более агрессивное использование ресурса (большая мощность передачи) увеличивает значение $h(u(t))$ и тем самым снижает вероятность перехода в мёртвую зону.

В рамках данной работы используется экспоненциальная форма управляющей функции (3) как

базовая. Она хорошо отражает интуитивно ожидаемое экспоненциальное ухудшение параметров канала при длительном нахождении в ослабленном состоянии [28], выбор экспоненты оправдан аналитической простотой, применимостью в известных моделях каналов с затуханием возможностью калибровки параметра α в рамках процедуры оптимизации [29].

Однако возможны и другие виды функций (логистическая, линейная и др.), в зависимости от физических или протокольных ограничений. Подобный выбор функции $g(\tau_2, \theta)$ может быть обоснован с помощью данных или в рамках аппроксимации механизмов, реализуемых, например, в адаптивной модуляции современных протоколов связи (см. LTE, 5G [30]).

В настоящей работе целевая функция (2) рассматривается в общем виде, однако для численного моделирования можно конкретизировать её следующим образом:

$$L_{\text{pack}} = \frac{1}{T} \int_0^T \mathbb{I}_{\text{loss}}(t) dt,$$

где $L_{\text{pack}}(u(t))$ – доля потерянных пакетов, $\mathbb{I}_{\text{loss}}(t) = 1$, если пакет в момент t был отброшен, и $\mathbb{I}_{\text{loss}}(t) = 0$ в противном случае.

Усреднённая задержка передачи пакетов $D(u(t))$ может быть оценена, например, по следующей формуле:

$$D(u(t)) = \frac{1}{N} \sum_{i=1}^N W_i,$$

где W_i – суммарное время пребывания i -го пакета в системе.

Таким образом, введение управления $u(t)$ и целевой функции даёт возможность комплексного подхода к оптимизации в сетях мобильной связи [31]: как регулирования скорости передачи (или других параметров канала), так и управления вероятностью разрывов и восстановлений путём изменения интенсивностей λ_{ij} . На практике выбор конкретной формы функции $u(t)$ и метода оптимизации определяется требованиями к качеству (снижению задержек, потерь) и доступными аппаратными (протокольными) возможностями [32].

Кроме того, подобное управление может быть реализовано посредством дополнительного модуля (платы) с программным управлением, в котором, например, применяются нейронные сети для прогнозирования состояния канала и динамической



настройки параметров. Тем самым удаётся вовремя реагировать на ухудшения связи и заранее адаптировать скорость передачи и вероятность разрывов в реальном времени.

В зависимости от специфических требований к обслуживанию (например, ограничения на задержку, порог вероятности потерь или ограничения на энергоёмкость передачи) предложенная целевая функция (2) может быть уточнена путём добавления дополнительных критериев или штрафов за превышение норм качества. Таким образом, при решении конкретных инженерных задач форма $L_{\text{pack}}(\cdot)$ и $D(\cdot)$ может быть выбрана с учётом практической значимости конкретных метрик качества работы сети (пакетные потери, время простоя и пр.). В настоящей работе функция $J(u)$ остаётся в общем виде; в разделе, посвящённом частным предположениям и упрощениям, продемонстрируем, как именно происходит выбор целевой функции для специальных случаев.

2.2. Основные свойства системы

Теперь рассмотрим основные характеристики системы. Для простоты и прозрачности дальнейшего анализа предположим, что интенсивности переходов между состояниями связи постоянны в пределах каждой зоны и зависят только от текущего состояния, в котором находится агрегатор. Это означает, что переходы между уровнями качества сигнала можно описывать с помощью фиксированных параметров, отражающих физические характеристики среды и динамику перемещения.

Таким образом, модель использует конечное число постоянных интенсивностей переходов, каждая из которых характеризует типичный режим перехода между зонами покрытия базовой станции. При предположении об экспоненциальном распределении времени пребывания в каждом состоянии модель сводится к непрерывному марковскому процессу с конечным числом состояний, задаваемому системой уравнений Колмогорова и, таким образом, вместо рассмотрения системы (1) можно перейти к рассмотрению инфинитезимальной матрицы переходов

$$Q = \begin{pmatrix} -\lambda_{12} & \lambda_{12} & 0 \\ \lambda_{21} & -(\lambda_{21} + \lambda_{23}) & \lambda_{23} \\ 0 & \lambda_{32} & -\lambda_{32} \end{pmatrix},$$

а вектор вероятностей состояний $p(t) = (p_1(t), p_2(t), p_3(t))$ удовлетворяет системе уравнений Колмогорова

$$\frac{dp(t)}{dt} = p(t)Q.$$

Здесь коэффициенты λ_{ij} имеют следующую физическую интерпретацию:

- λ_{12} – интенсивность перехода от устойчивого сигнала S_1 к ослабленному S_2 , соответствующая движению из зоны $d \leq R$ в зону $R < d \leq R + \delta$;
- λ_{21} – интенсивность восстановления сигнала от состояния S_2 к S_1 при обратном переходе в зону хорошего покрытия;
- λ_{23} – интенсивность потери сигнала (переход от состояния S_2 к S_3), что соответствует попаданию в мёртвую зону;
- λ_{32} – интенсивность восстановления сигнала при возвращении из мёртвой зоны S_3 в переходную зону S_2 .

Замечание. В описываемом подходе структура сети явно не входит в модель. Однако она может быть косвенно учтена путём параметризации матрицы Q на основе матрицы смежности графа покрытия или других структурных характеристик сети. Такой подход открывает возможность для обобщения модели на случай пространственно неоднородных сетей или реальных телекоммуникационных инфраструктур.

Теорема 1 (о существовании кусочно-гладкой управляющей функции). Для канала, находящегося в одном из трёх состояний: S_1 (устойчивое состояние), S_2 (ослабленное состояние) и S_3 (разрыв связи), существует кусочно-гладкая управляющая функция общего вида $u(t)$, зависящая только от текущего состояния канала $X(t)$ такая, что

$$u(t) = \begin{cases} u_1, & X(t) = S_1, \\ u_2, & X(t) = S_2, \\ u_3, & X(t) = S_3, \end{cases}$$

где u_i – параметры управления в состоянии S_i ($i=1, 2, 3$). Интенсивности обслуживания $\mu_i(u(t))$ в каждом состоянии канала и переходов $\lambda_{ij}(u(t))$ между состояниями при этом корректно определены и управляемы на каждом интервале постоянства состояния.

Данная функция управления определяет интенсивность обслуживания $\mu_i(u(t))$ в каждом состоянии канала, а также может влиять на интенсивно-

сти переходов $\lambda_{ij}(u(t))$ между состояниями (например, путём управляемого ускорения восстановления канала или замедления его ухудшения). Иными словами, когда канал находится в состоянии S_i , система применяет управление u_i , которое задаёт текущую скорость обслуживания $\mu_i(u_i)$ и может изменять переходные интенсивности $\lambda_{ij}(u_i)$ в другие состояния.

Доказательство (обоснование выбора формы $u(t)$). Управление, зависящее только от текущего состояния $X(t)$, естественно в контексте адаптивных систем связи. Такая кусочно-заданная функция $u(t)$ позволяет мгновенно реагировать на изменения качества канала: при переходе из устойчивого состояния в ослабленное или при разрыве связи управление скачкообразно меняет свое значение на новое, соответствующее ухудшившимся условиям. Это обеспечивает адаптивное регулирование обслуживания – например, при ухудшении канала (S_2) можно уменьшить скорость обслуживания или включить более надёжный режим передачи, а при разрыве (S_3) – инициировать процедуры восстановления связи (что эквивалентно увеличению интенсивности перехода обратно в состояние S_1 или S_2). В устойчивом состоянии S_1 управление может возвращаться в режим максимальной пропускной способности. Такая форма функции $u(t)$ гарантирует, что система динамически подстраивается под текущее состояние канала, улучшая устойчивость обслуживания при деградации канала и сокращая длительность простоев при разрывах. ♦

Теорема 2 (критерий устойчивости системы с очередью). Рассмотрим марковский процесс $X(t)$, принимающий три состояния $\{S_1, S_2, S_3\}$, где состояния S_1 и S_2 имеют интенсивности обслуживания μ_1 и μ_2 соответственно, а в состоянии S_3 передача невозможна ($\mu_3 = 0$). Пусть λ_{in} – интенсивность входного (пуассоновского) потока заявок. Обозначим через π_1, π_2, π_3 стационарные вероятности пребывания канала в состояниях S_1, S_2, S_3 (если они существуют).

Система очередей (с неограниченным буфером) и переменным качеством сигнала устойчива тогда и только тогда, когда

$$\lambda_{in} < \pi_1 \mu_1 + \pi_2 \mu_2. \quad (4)$$

Условие (4) означает, что средняя входная нагрузка должна быть меньше средней пропускной способности канала, равной $\pi_1 \mu_1 + \pi_2 \mu_2$. При

нарушении условия (4) очередь (при неограниченной ёмкости) возрастает неограниченно, см. работы [3, 24].

Доказательство. В модели $M/M/1$ условие устойчивости $\lambda_{in} < \mu$ есть необходимое и достаточное [2]. Для случая с переменной μ , зависящей от состояния канала, средняя скорость обслуживания равна $\sum_{i=1}^3 \pi_i \mu_i$. Так как $\mu_3 = 0$, вклад состояния S_3 в пропуск-

ную способность нулевой. Следовательно, критерий переходит в неравенство (4). При выполнении условия (4) существует единственное стационарное распределение размера очереди; иначе стационарное распределение отсутствует.

Стационарные вероятности π_1, π_2, π_3 определяются из уравнений Колмогорова (1). Стационарный вектор $\pi = (\pi_1, \pi_2, \pi_3)$ удовлетворяет системе

$$\begin{cases} -\lambda_{12}\pi_1 + \lambda_{21}\pi_2 = 0, \\ \lambda_{12}\pi_1 - (\lambda_{21} + \lambda_{23})\pi_2 + \lambda_{32}\pi_3 = 0, \\ \pi_1 + \pi_2 + \pi_3 = 1. \end{cases}$$

Решая систему, получаем:

$$\pi_2 = \frac{\lambda_{12}}{\lambda_{21}} \pi_1, \quad \pi_3 = \frac{\lambda_{23}}{\lambda_{32}} \pi_2 = \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} \pi_1.$$

Подставляя в условие нормировки:

$$\begin{aligned} \pi_1 + \pi_2 + \pi_3 &= \pi_1 \left(1 + \frac{\lambda_{12}}{\lambda_{21}} + \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} \right) = 1 \Rightarrow \\ \Rightarrow \pi_1 &= \left(1 + \frac{\lambda_{12}}{\lambda_{21}} + \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} \right)^{-1}. \end{aligned}$$

Подстановка π_1, π_2, π_3 в неравенство (4) даёт явное численное условие

$$\lambda_{in} \leq \left(\frac{\lambda_{21}\lambda_{32}}{\Delta} \right) \mu_1 + \left(\frac{\lambda_{12}\lambda_{32}}{\Delta} \right) \mu_2,$$

где

$$\Delta = \lambda_{21}\lambda_{32} + \lambda_{12}\lambda_{32} + \lambda_{12}\lambda_{23}. \quad \blacklozenge$$

Следствие (условие устойчивости системы с управлением). В стационарном режиме условие устойчивости системы с учётом оптимального управления записывается как

$$\lambda_{in} < \pi_1(u^*) \mu_{norm} + \pi_2(u^*) \mu_{slow},$$

где стационарные вероятности $\pi_i(u^*)$ зависят от оптимальной стратегии $u^*(t)$.



2.3. Операционные характеристики системы в стационарном режиме

Хотя в предыдущих разделах изложение было сосредоточено на описании канала (состояний S_1, S_2, S_3) и его переходных интенсивностях, в реальной системе с очередью полное состояние в момент времени t задаётся парой «состояние канала – длина очереди».

Умея оценивать эволюцию канала в терминах состояний $\{S_1, S_2, S_3\}$ и управляющей функции $u(t)$, можно оценить качество работы системы с учётом трафика в реальных системах. Для этого объединим марковскую модель канала с моделью очереди типа $M/M/1$ (либо $M/M/1/N$ при конечном буфере), где скорость обслуживания зависит от текущего состояния канала и, при необходимости, от управления $u(t)$.

Такой подход основан на классических результатах теории массового обслуживания [3, 4], применённых к ситуации, когда параметры обслуживания (и даже переходы канала из одного состояния в другое) могут меняться в зависимости от стратегии управления. Как как будет показано далее, в стационарном режиме ключевые вероятностные характеристики – вероятность потерь, среднее число заявок в системе, пропускная способность – выражаются через стандартные формулы Эрланга и Литтла, но с учётом средних (эффективных) скоростей, задаваемых состояниями канала.

Рассмотрим систему с очередью, которую обслуживает *управляемый* марковский канал с тремя состояниями $\{S_1, S_2, S_3\}$, где управление $u(t)$ может влиять как на интенсивность обслуживания $\mu_i(u)$ в состояниях S_i , так и на интенсивности переходов $\lambda_{ij}(u)$. Предположим следующее:

- Очередь имеет конечную ёмкость: максимум N пакетов в ожидании (значит, всего $N+1$ пакетов вместе с обслуживаемым).

- В стационарном режиме канал проводит долю времени $\pi_i(u^*)$ в состоянии S_i (при оптимальном или фиксированном управлении u^*).

- Если $\mu_3(u) = 0$ (пребывание в «мёртвой» зоне), то $\bar{\mu} = \pi_1(u^*)\mu_1(u^*) + \pi_2(u^*)\mu_2(u^*)$ есть *эффективная* (средняя) скорость обслуживания.

- Интенсивность входного потока заявок – λ_{in} .

Тогда с учётом предположений можно найти операционные характеристики работы системы.

- *Коэффициент загрузки.* Определим относительную нагрузку (загрузку):

$$\rho = \frac{\lambda_{in}}{\bar{\mu}} = \frac{\lambda_{in}}{\pi_1(u^*)\mu_1(u^*) + \pi_2(u^*)\mu_2(u^*)}.$$

Если $\rho < 1$ при данном u^* , система может работать без «коллапса» в неограниченном буфере; если $\rho \geq 1$, очередь растёт или происходят большие потери.

- *Вероятность потерь.* Для конечного N вероятность потери пакета при переполнении очереди задаётся (аналогично формуле Эрланга) через

$$P_{loss} = P\{\text{очередь} = N\} = \frac{\rho^{N+1}}{\sum_{k=0}^{N+1} \rho^k},$$

где ρ определяется через $\pi_i(u^*)$ и $\mu_i(u^*)$. При $\rho < 1$ и $N \rightarrow \infty$ получаем $P_{loss} \rightarrow 0$, а при $\rho > 1$ имеем $P_{loss} \rightarrow 1$.

- *Среднее число пакетов и задержка.* Обозначим $L = E[n]$ математическое ожидание числа пакетов в очереди и на обслуживании. При $\rho \neq 1$ справедлива формула

$$L = \frac{\rho}{1-\rho} - \frac{(N+2)\rho^{N+2}(1-\rho)}{1-\rho^{N+2}}.$$

Среднее число пакетов в очереди L_q получается вычитанием средней занятой позиции обслуживания; аналогично для среднего времени пребывания (W) и ожидания (W_q) применяем закон Литтла:

$$L = \lambda_{eff} W, \quad L_q = \lambda_{eff} W_q,$$

где $\lambda_{eff} = \lambda_{in}(1 - P_{loss})$ – интенсивность фактического потока поступающих и обслуживаемых пакетов.

- *Средняя пропускная способность.* В стационарном режиме система пропускает поток с интенсивностью

$$\lambda_{out} = \lambda_{eff} = \lambda_{in}(1 - P_{loss}).$$

Если $\rho < 1$ и $N \rightarrow \infty$, то $\lambda_{out} = \lambda_{in}$ (нет потерь). При $\rho > 1$ очередь переполнена и $\lambda_{out} \approx \bar{\mu}$.

2.4. Следствия и замечания

- *Граничные случаи (сводимость к классическим моделям).* Если канал практически всегда

находится в хорошем состоянии S_1 (т. е. $\pi_1 \approx 1$, $\pi_2 \approx \pi_3 \approx 0$, $\mu_1 = \mu$), то вся система сводится к классическому $M/M/1$. Тогда критерий устойчивости принимает вид $\lambda_{in} < \mu$, а формулы для вероятности потерь, среднего числа пакетов и времени ожидания упрощаются до хорошо известных результатов (формула Эрланга, формулы Литтла и т. д.). Напротив, если значительную долю времени агрегатор пребывает в мёртвой зоне S_3 (π_3 велика, $\mu_3 \approx 0$), то эффективная скорость $\bar{\mu}$ резко падает, и система ведёт себя как «медленный сервер» с высокой нагрузкой ρ .

• **Пример влияния пребывания в мёртвой зоне на потери.** Пусть $\mu_1 = 1$, $\mu_2 = 0,5$, $\mu_3 = 0$, а доли времени в состояниях таковы: $\pi_1 = 0,8$, $\pi_2 = 0,15$, $\pi_3 = 0,05$. Тогда $\bar{\mu} = 0,8 \cdot 1 + 0,15 \cdot 0,5 = 0,875$. При входном потоке $\lambda_{in} = 0,7$ имеем $\rho \approx 0,8 < 1$, и система работает без существенных потерь. Однако при $\lambda_{in} = 1,0$ ($\rho \approx 1,14 > 1$) очередь начинает неограниченно расти (при большом буфере) либо достигает максимума N (при конечном буфере); возникает ощутимая вероятность потерь ($P_{loss} \approx 0,2$), а выходная пропускная способность насыщается на уровне $\lambda_{out} \approx \bar{\mu} = 0,875$.

• **Физический смысл управления переходами и скоростью обслуживания.** Если в системе можно управлять интенсивностями $\lambda_{ij}(u(t))$ (ускоряя восстановление из состояния S_3 или замедляя уход из состояния S_1) и (или) изменять $\mu_i(u(t))$ (выбирая более агрессивные или безопасные режимы передачи), то фактически можно контролировать долю времени π_3 (пребывание в мёртвой зоне) и эффективную скорость $\bar{\mu}$. Тем самым можно существенно понизить нагрузку ρ , уменьшить вероятность потерь и среднюю задержку. Однако на практике повышение μ_i или сокращение времени пребывания в состоянии S_3 может требовать дополнительных ресурсов (мощность, резервные каналы и т. д.), поэтому приходится искать компромисс между затратами и выигрышем в качестве обслуживания.

• **Область применения и перспективы.** Сделанные допущения (экспоненциальные распределения, марковские или полумарковские переходы) упрощают анализ и дают элегантные формулы, но могут быть лишь приближением к реальным каналам с коррелированным трафиком и сложной динамикой распространения сигнала в неоднородной среде. Тем не менее, основной вывод (необходимость выполнения условия $\lambda_{in} < \bar{\mu}$ для устойчиво-

сти) и принципы оптимизации управления $u(t)$ сохраняют силу. Предложенная модель позволяет оценить предельную производительность сети, а также адаптировать стратегию обслуживания и переходов в современных мобильных сетях, где важны как пропускная способность сети, так и энергопотребление, время восстановления связи и пр.

• Связь с $M/G/1$ -системами с прерываниями.

Рассматриваемая в данной работе модель канала с тремя состояниями (как обобщение системы с $n > 3$ состояниями [33–35]) с изменяющейся скоростью обслуживания и периодами полной недоступности (разрывами) концептуально близка к $M/G/1$ -системам с прерываниями (см., например, работы [36, 37]). Однако, в отличие от типовых моделей $M/G/1$ с прерываниями, учитывается дополнительная управляющая функция $u(t)$, влияющая на интенсивности переходов и скорость обслуживания. Это позволяет подстраивать параметры под текущие условия мобильного канала, что расширяет классические подходы и открывает возможность адаптивной оптимизации.

• Таким образом, сходство с $M/G/1$ с прерываниями присутствует, и полученные в настоящей работе результаты (в частности, критерий $\lambda_{in} < \sum_i \pi_i \mu_i$) повторяют логику классических моделей. Однако из-за динамического управления, физической интерпретации состояний (радиоканал) и целевой функции $J(u)$ представленная модель приобретает более гибкую структуру, открывая новые возможности адаптации в инженерных приложениях.

3. ЧИСЛЕННЫЙ АНАЛИЗ УСТОЙЧИВОСТИ И ПРОИЗВОДИТЕЛЬНОСТИ АДАПТИВНОЙ МОДЕЛИ

В данном параграфе представлены результаты имитационного моделирования, подтверждающие теоретические выводы, изложенные выше. Рассматривается стохастический канал с тремя состояниями $\{S_1, S_2, S_3\}$ и входным потоком заявок интенсивности λ_{in} . Пропускная способность системы зависит от текущего состояния канала и управляющей функции $u(t)$.

Моделирование выполнено методом дискретных событий на интервале T_{max} . Исследуются ключевые характеристики: устойчивость системы, длина очереди, вероятность потерь, а также влияние параметров сети и стратегии управления на



качество работы сети. Ниже приведены параметры эксперимента, метрики и визуализация поведения системы в различных режимах.

3.1. Параметры модели

Поток заявок моделируется пуассоновским процессом с интенсивностью λ_{in} . Заявки обрабатываются по принципу FIFO (*first in, first out*) с буфером конечной ёмкости.

Канал может находиться в трёх состояниях:

- S_1 : устойчивая связь, пропускная способность μ_{norm} ;
- S_2 : ослабленный сигнал, пропускная способность μ_{slow} (или $\mu_{slow}e^{-\alpha\tau_2}$);
- S_3 : разрыв связи, обслуживание недоступно ($\mu = 0$).

Переходы между состояниями описываются либо фиксированной матрицей интенсивностей (марковский случай), либо функциями $\alpha_{ij}(t)$ в полумарковской модели. Обслуживание выполняется, если канал находится в состоянии S_1 или S_2 и очередь не пуста. В состоянии S_3 приём заявок продолжается, но обслуживание приостанавливается.

Основные метрики качества работы сети: динамика длины очереди и влияние интенсивности λ_{in} и скорости движения на устойчивость.

3.2. Результаты моделирования

Приведённые ниже графики демонстрируют поведение модели в двух режимах:

- устойчивый: $\lambda_{in} < \pi_1\mu_{norm} + \pi_2\mu_{slow}$;
- неустойчивый: $\lambda_{in} > \pi_1\mu_{norm} + \pi_2\mu_{slow}$.

В устойчивом режиме (рис. 1) очередь (размер буфера агрегатора) стабилизируется и не растёт неограниченно. Однако при перегрузке (рис. 2) имеется положительный снос.

На рис. 3 видно, что в устойчивом режиме даже при разрывах связи восстанавливается достаточно быстро. При перегрузке даже кратковременные перебои в обслуживании вызывают лавинообразный рост буфера агрегатора.

На рис. 4 показано влияние экспоненциального затухания в состоянии S_2 : при высокой нагрузке увеличивается время пребывания в состоянии S_3 , что резко ухудшает устойчивость.

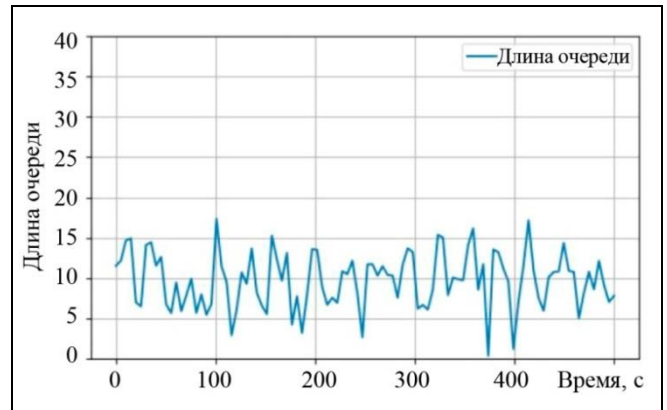


Рис. 1. Динамика длины очереди в устойчивом режиме

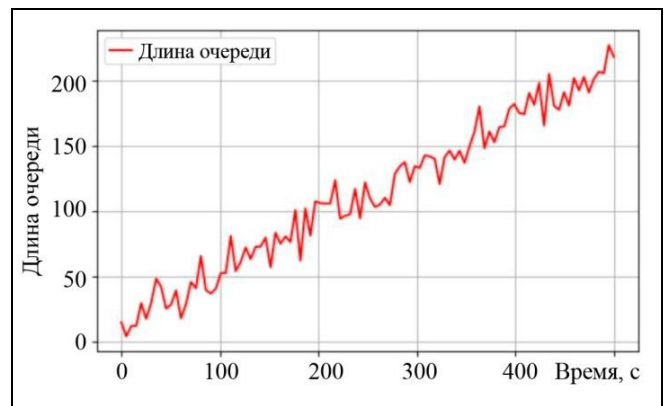


Рис. 2. Рост очереди в неустойчивом режиме

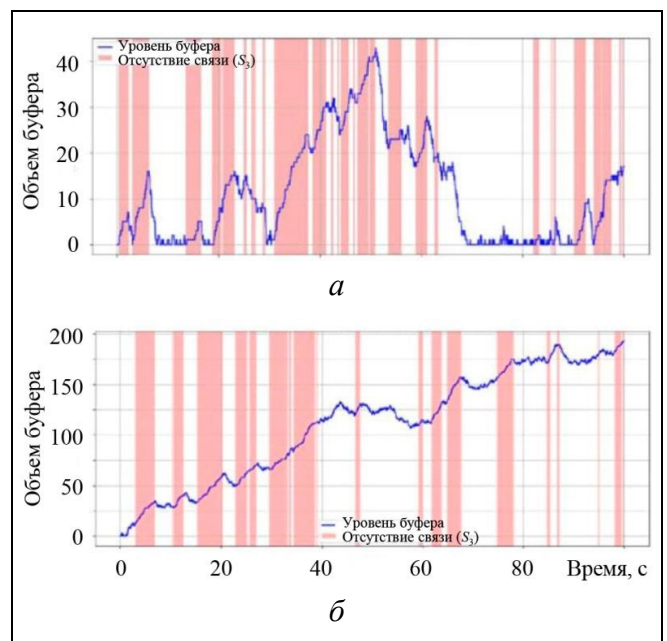


Рис. 3. Очередь: а – в устойчивом режиме ($\lambda = 3,0$ ед./с, $p = 30$ м/с), б – в неустойчивом режиме ($\lambda = 7,0$ ед./с, $p = 30$ м/с). Красные зоны – периоды разрыва связи

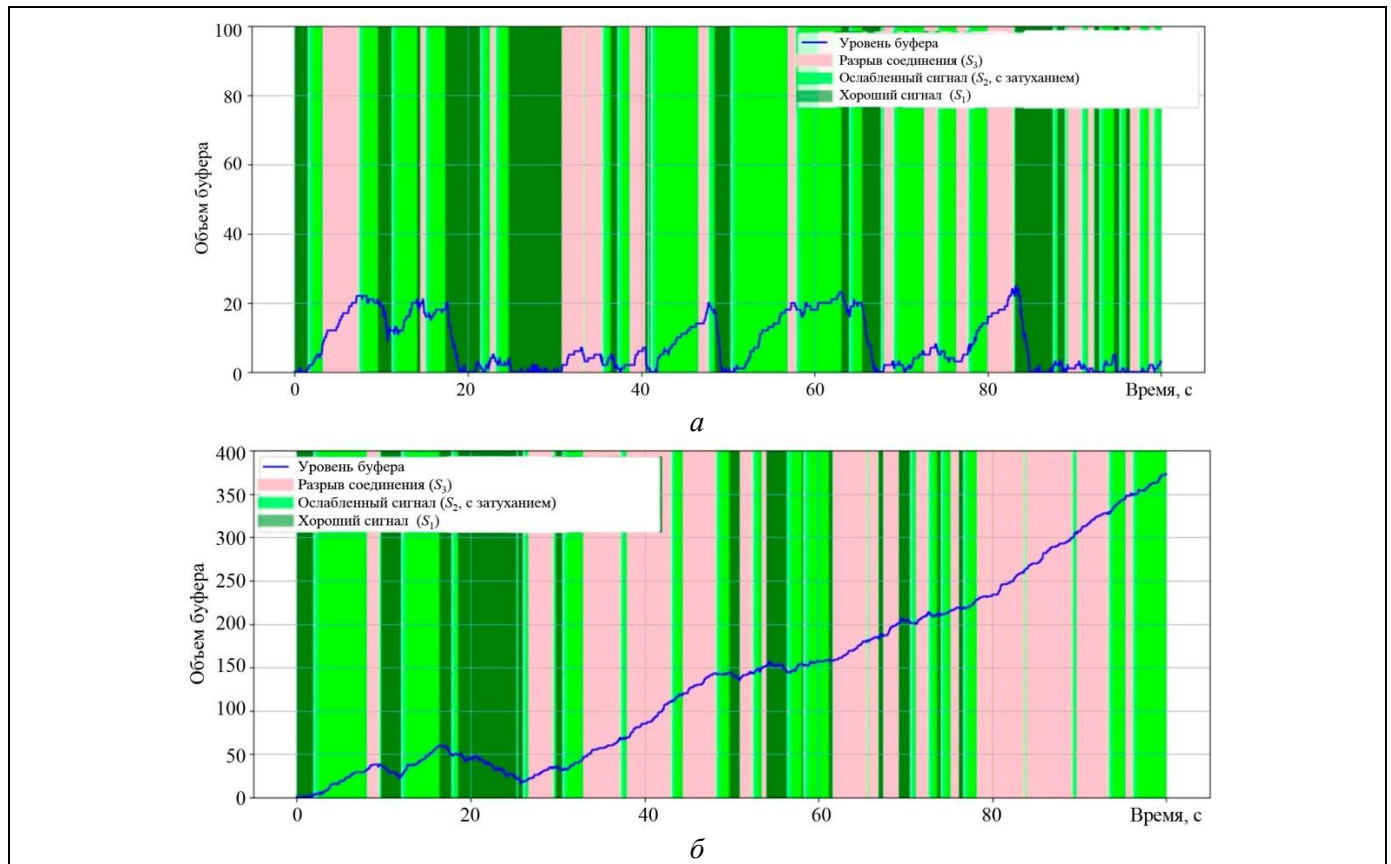


Рис. 4. Затухающая пропускная способность в состоянии S_2 : a – устойчивый режим ($\lambda = 3,0$), b – неустойчивый режим ($\lambda = 7,0$)

Рис. 5 демонстрирует преимущество адаптивного управления: вблизи критической нагрузки оно эффективно стабилизирует длину очереди.

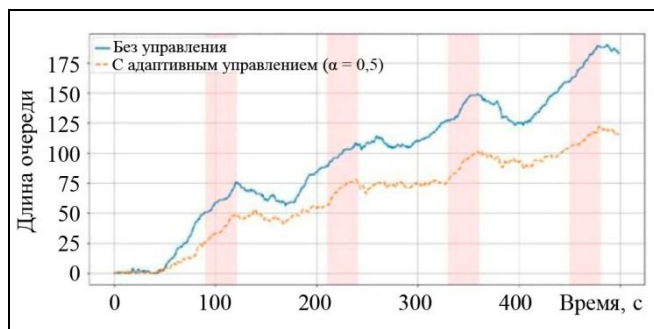


Рис. 5. Сравнение: динамика длины очереди при адаптивном управлении и при его отсутствии

На рис. 6 показана граница устойчивости. Адаптивная модель демонстрирует смещение порога перегрузки, позволяя системе функционировать ближе к критическим значениям.

Рис. 7 показывает влияние полумарковских эффектов: при гамма-распределении и распределении Вейбулла система становится более инерционной, что повышает чувствительность к перегрузкам.

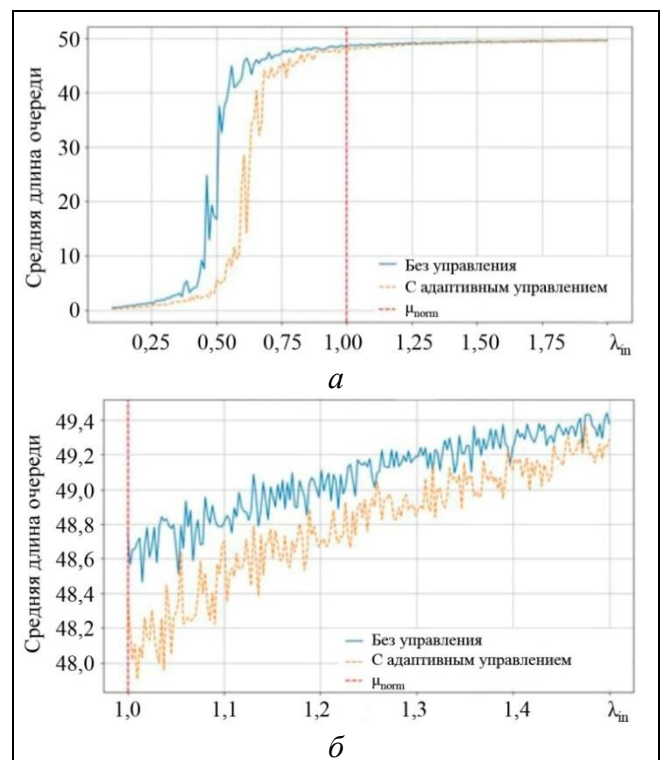


Рис. 6. Фазовый переход: средняя длина очереди в зависимости от значения λ_{in}

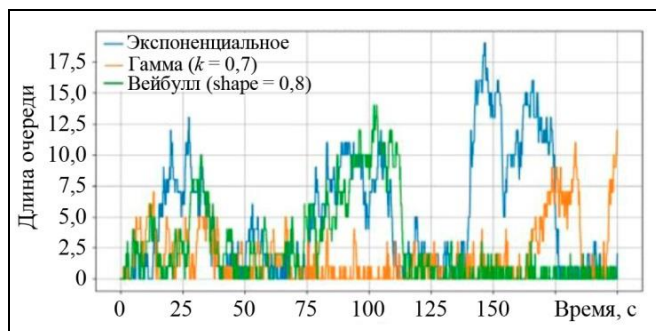


Рис. 7. Зависимость динамики длины очереди от закона распределения времени обслуживания

ЗАКЛЮЧЕНИЕ И ПЕРСПЕКТИВЫ

Предложенная модель описывает систему с адаптивным управлением в условиях стохастически меняющихся характеристик канала, находящегося в одном из трёх состояний: стабильной связи, ослабленного сигнала и полного разрыва связи. Было показано, что включение управляющей функции $u(t)$, влияющей как на интенсивность обслуживания, так и на переходы между состояниями, позволяет динамически адаптировать поведение системы к текущим условиям. В отличие от классических моделей, в которых параметры фиксированы, в предлагаемом подходе они становятся функциями стратегии, что даёт возможность прямо учитывать влияние управления на производительность сети.

Проведённый численный анализ подтверждает, что при $\lambda_{in} < \bar{\mu}$ система устойчива: очередь остаётся ограниченной. При перегрузке ($\lambda_{in} > \bar{\mu}$) происходит лавинообразный рост длины очереди. Однако адаптивное управление замедляет этот рост и фактически сдвигает границу устойчивости. Были также продемонстрированы эффекты временной корреляции: при переходе к полумарковским моделям система становится инерционной, но общие критерии устойчивости сохраняются. Это подчёркивает универсальность предложенной конструкции и её применимость как в классических, так и в более сложных телекоммуникационных сценариях.

Практическая реализация модели возможна в виде программируемого блока, встроенного в устройство агрегации. Такой модуль может использовать нейросетевые схемы для оценки текущего состояния канала и управления параметрами передачи. Кроме того, модель может быть обобщена в сторону марковского управления: когда решения зависят не только от текущего состояния, но и от времени пребывания в нём. Это естественным образом реализуется через системы с памятью

или логики, способные учитывать историю состояний. Также возможна реализация промежуточных состояний связи, задаваемых средствами многозначной логики, что позволяет описывать не только резкие, но и плавные деградации канала.

Таким образом, модель объединяет стохастическую динамику, управление и очереди в единую систему и может быть использована как в теоретических исследованиях устойчивости подобных систем, так и при разработке новых стеков протоколов передачи данных, особенно в контексте мобильных сетей нового поколения. Перспективы дальнейших исследований включают многопользовательские сценарии, исследование стратегий управления на данных и построение энергоэффективных алгоритмов для устройств с ограниченными вычислительными ресурсами на базе интеллектуального управления. В качестве самостоятельной теоретической задачи отметим исследование фазового перехода, открытым является вопрос устранения экспоненциального характера роста очереди и замедление адаптивного управления.

Автор выражает глубокую признательность анонимным рецензентам за ценные замечания и конструктивные предложения, которые значительно улучшили качество данной работы.

ЛИТЕРАТУРА

1. Патент RU 2694025 C1. Система агрегации сетевых данных в компьютерных сетях : 2019102447 : заявл. 29.02.2019 : опубл. 08.07.2019, бюл. № 19 / Марченков А.А., Есин А.А. [Patent RU 2694025 C1. Network Data Aggregation System in Computer Networks : 2019102447 : appl. 29.02.2019 : reg. 08.07.2019, bul. no. 19 / Marchenkov, A.A., Esin, A.A.]
2. Chang, C.S. Performance Guarantees in Communication Networks; Telecommunication Networks and Computer Systems. – London: Springer-Verlag, 2000. – 392 p.
3. Kleinrock, L. Queueing Systems: Volume 1. – New York: Wiley-Interscience, 1975. – 431 p.
4. Gross, D., Shortle, J., Thompson, J., Harris, C. Fundamentals of Queueing Theory; Wiley Series in Probability and Statistics. – New York: Wiley, 2011. – 527 p.
5. Baccelli, F., Blaszczyszyn, B. Stochastic Geometry and Wireless Networks: Volume I Theory // Foundations and Trends® in Networking. – 2010. – Vol. 3, no. 3–4. – P. 249–449. – DOI: <https://doi.org/10.1561/13000000006>
6. Haenggi, M., Andrews, J.G., Baccelli, F., et al. Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks // IEEE Journal on Selected Areas in Communications. – 2009. – Vol. 27, no. 7. – P. 1029–1046. – DOI: <https://doi.org/10.1109/JSAC.2009.090902>
7. Baccelli, F., Blaszczyszyn, B. On a Coverage Process Ranging from the Boolean Model to the Poisson–Voronoi Tessellation with Applications to Wireless Communications // Advances in Applied Probability. – 2001. – Vol. 33, no. 2. – P. 293–323. – DOI: <https://doi.org/10.1017/S0001867800010806>

8. Raigorodskii, A.M. Random Graph Models and Their Applications // Proceedings of MIPT. – 2010. – Vol. 2. – P. 130–140.
9. Stauffer, D., Aharony, A. Introduction to Percolation Theory. – Milton Park: Taylor & Francis, 1994. – 186 p.
10. Saberi, A.A. Recent Advances in Percolation Theory and Its Applications // Physics Reports. – 2015. – Vol. 578. – P. 1–32. – DOI: <https://doi.org/10.1016/j.physrep.2015.03.003>
11. Olle, H., Yuval, P., Jeffrey, E.S. Dynamical Percolation // Annales de l'Institut Henri Poincaré (B) Probability and Statistics. – 1997. – Vol. 33. – P. 497–528.
12. Li, M., Liu, R.R., Lü, L., et al. Percolation on complex networks: Theory and application // Physics Reports. – 2021. – Vol. 907. – P. 1–68. – DOI: <https://doi.org/10.1016/j.physrep.2020.12.003>
13. Badie-Modiri, A., Rizi, A.K., Karsai, M., Kivelä, M. Directed Percolation in Temporal Networks // Phys. Rev. Res. – 2022. – Vol. 4. – art. No. L022047. – DOI: <https://doi.org/10.1103/PhysRevResearch.4.L022047>
14. ElSawy, H., Zhaikhan, A., Kishk, M.A., Alouini, M.S. A Tutorial-Cum-Survey on Percolation Theory with Applications in Large-Scale Wireless Networks // IEEE Communications Surveys & Tutorials. – 2024. – Vol. 26. – P. 428–460. – DOI: <https://doi.org/10.1109/COMST.2023.3336194>
15. Lin, H., Kishk, M.A., Alouini, M.S. Performance Analysis of Infrastructure Sharing Techniques in Cellular Networks: A Percolation Theory Approach // arXiv:2502.08023. – 2025. – DOI: <https://doi.org/10.48550/arXiv.2502.08023>
16. Zhang, Q., Kassam, S.A. Finite-State Markov Model for Rayleigh Fading Channels // IEEE Transactions on Communications. – 1999. – Vol. 47, no. 11. – P. 1688–1692. DOI: 10.1109/26.803503
17. Paranthaman, V.V., Mapp, G., Shah, P., et al. Exploring Markov Models for the Allocation of Resources for Proactive Handover in a Mobile Environment // Proceedings of 2015 IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops). – Clearwater Beach, 2015. – P. 855–861. DOI: 10.1109/LCNW.2015.7365938
18. Kalimulina, E.Y. Application of Multi-Valued Logic Models in Traffic Aggregation Problems in Mobile Networks // Proceedings of the 15th IEEE International Conference on Application of Information and Communication Technologies (AICT). – Baku, Azerbaijan, 2021. – P. 1–6. – DOI: 10.1109/AICT52784.2021.9620244
19. Kalimulina, E.Y. Lattice Structure of Some Closed Classes for Non-binary Logic and Its Applications // In: Mathematical Methods for Engineering Applications. ICMASE 2021 / ed. by F. Yilmaz, A. Queiruga-Dios, M.J. Santos Sánchez, D. Rasteiro, V. Gayoso Martínez, J. Martín Vaquero. – Springer Proceedings in Mathematics & Statistics. – Vol. 384. – Cham: Springer, 2022. – P. 25–34. – DOI: 10.1007/978-3-030-96401-6_2
20. Goldsmith, A. Wireless Communications. – Cambridge: Cambridge University Press, 2005. – 648 p.
21. Semi-Markov Models and Applications / ed. by J. Janssen, N. Limnios. – Cham: Springer Science & Business Media. – 2013.
22. Limnios, N., Oprisan, G. Semi-Markov Processes and Reliability. – Cham: Springer Science & Business Media. – 2012.
23. Cinlar, E. Introduction to Stochastic Processes. – Hoboken: Prentice-Hall, 1975.
24. Ross, S. Stochastic Processes. – New York: Wiley, 1996. – 328 p.
25. Kirk, D.E. Optimal Control Theory: An Introduction. – Hoboken: Prentice-Hall, 1970.
26. Bertsekas, D.P. Dynamic Programming and Optimal Control. – Raleigh: Athena Scientific, 2005. – 484 p.
27. Boyd, S., Vandenberghe, L. Convex Optimization. – Cambridge: Cambridge University Press, 2004.
28. Wang, T., Proakis, J.G., Masry, E., Zeidler, J.R. Performance Degradation of OFDM Systems due to Doppler Spreading // IEEE Transactions on Wireless Communications. – 2006. – Vol. 5, no. 6. – P. 1422–1432. – DOI: 10.1109/TWC.2006.1638663
29. Yin, X., Cheng, X. Propagation Channel Characterization, Parameter Estimation, and Modeling for Wireless Communications. – Hoboken: John Wiley & Sons, 2016. – 350 p.
30. Modeas, I., Kaloxylas, A., Merakos, L., Tsolkas, D. An Adaptive and Distributed Network Selection Mechanism for 5G Networks // Computer Networks. – 2021. – Vol. 189. – Art. no. 107943. – DOI: 10.1016/j.comnet.2021.107943
31. Kalimulina, E.Y. Mathematical Model for Reliability Optimization of Distributed Telecommunications Networks // Proceedings of 2011 International Conference on Computer Science and Network Technology. – Harbin, China, 2011. – P. 2847–2853. – DOI: <https://doi.org/10.1109/ICCSNT.2011.61825570>
32. Kalimulina, E.Y. A New Approach for Dependability Planning of Network Systems // International Journal of System Assurance Engineering and Management. – 2013. – Vol. 4. – P. 215–222.
33. Есин А.А. Анализ современных принципов построения систем управления на основе моделей многозначной логики // Управление большими системами. – 2020. – Вып. 88. – С. 69–98. – DOI: <https://doi.org/10.25728/ubs.2020.88.4>. [Esin, A.A. Analysis and Design Principles of Modern Control Systems Based on Multi-valued Logic Models // Large-Scale Systems Control. – 2020. – Iss. 88. – P. 69–98. – DOI: <https://doi.org/10.25728/ubs.2020.88.4>. (In Russian)]
34. Esin, A.A. Characteristics of Structurally Finite Classes of Order-Preserving Three-Valued Logic Maps // Logic Journal of the IGPL. – 2024. – Art. no. jzae128. – URL: <https://academic.oup.com/jigpal/article-lookup/doi/10.1093/jigpal/jzae128>.
35. Esin, A.A. Structural Analysis of Precomplete Classes and Closure Diagrams in Multi-valued Logic // Iranian Journal of Fuzzy Systems. – 2025. – Vol. 21, no. 6. – P. 127–145.
36. Gao, S., Liu, Z. An M/G/1 Queue with Single Working Vacation and Vacation Interruption under Bernoulli Schedule // Applied Mathematical Modelling. – 2013. – Vol. 37, no. 3. – P. 1564–1579. – DOI: 10.1016/j.apm.2012.04.045
37. Zhang, M., Hou, Z. Performance Analysis of M/G/1 Queue with Working Vacations and Vacation Interruption // Journal of Computational and Applied Mathematics. – 2010. – Vol. 234, no. 10. – P. 2977–2985. – DOI: 10.1016/j.cam.2010.04.010

Статья представлена к публикации членом редколлегии
В. М. Вишневым.

Поступила в редакцию 17.05.2025,
после доработки 19.05.2025.
Принята к публикации 20.05.2025.

Есин Антон Анатольевич – директор по развитию,
АО «МирВайФай», г. Москва, ✉ anton.esin@imm.am.
ORCID iD: <https://orcid.org/0009-0007-0680-8288>

© 2025 г. Есин А. А.



Эта статья доступна по лицензии Creative Commons
«Attribution» («Атрибуция») 4.0 Всемирная.



A MATHEMATICAL MODEL OF ADAPTIVE TRAFFIC CONTROL IN MOBILE NETWORKS WITH VARIABLE SIGNAL QUALITY

A. A. Esin

AO MirWiFi, Moscow, Russia

✉ anton.esin@imm.am

Abstract. This paper considers the mathematical modeling problem of traffic transmission in mobile networks under high user mobility and spatially heterogeneous coverage, including signal degradation (“dead”) zones. Traffic aggregation at the channel level is applied to increase the reliability and stability of data transmission. A Markov model of a communication channel is proposed to study the effectiveness of aggregation algorithms and adapt them to network operating parameters and user velocity. The model is based on a periodic affine motion of a mobile device between base stations (BSs) uniformly distributed along a straight line. Within this model, the concepts of stable coverage zones and transition zones are introduced in terms of the distances to the nearest and next BSs. The channel state is described by a Markov chain with states corresponding to signal quality sampling: stable connection, degraded connection, and disconnection. Transitions between states are governed by a continuous-time Markov process with constant rates, and the parameters of this process are determined from empirical network data. An extension of the model to incorporate the time dependence of channel states is also considered, leading to a semi-Markov framework. For both cases (Markov and semi-Markov chains), explicit expressions are derived for the stationary probabilities of states, and system stability conditions are formulated to ensure bounded traffic queues. In addition, an adaptive control model of the channel throughput is proposed; this model optimizes transmission parameters depending on the current channel state, request queue length, and user velocity. The effectiveness of the approach is demonstrated by numerical simulations: the network has stable performance across a wide range of mobility levels and coverage parameters. The model can be applied to the reliability analysis and optimization of network protocols in highly mobile environments, including high-speed railway transport, vehicle networks, and mobile platforms.

Keywords: mobile networks, Markov model, semi-Markov process, system stability, Laplace transform.

Acknowledgments. The author is grateful to the reviewers for careful reading of the manuscript and helpful remarks.