

УДК 621.39

doi:10.31799/1684-8853-2019-4-69-81

Эффективные стратегии обслуживания пользователей в многопользовательской системе MIMO

А. В. Горбунова^а, канд. физ.-мат. наук, старший преподаватель, orcid.org/0000-0001-7073-7022

Е. Г. Медведева^{а, б}, ассистент, orcid.org/0000-0001-9104-708X, medvedeva-eg@rudn.ru

Ю. В. Гайдамака^{а, б}, доктор физ.-мат. наук, профессор, orcid.org/0000-0003-2655-4805

В. С. Шоргин^б, канд. техн. наук, старший научный сотрудник., orcid.org/0000-0002-1090-544X

К. Е. Самуйлов^{а, б}, доктор техн. наук, профессор, orcid.org/0000-0002-6368-9680, samuylov-ke@rudn.ru

^аРоссийский университет дружбы народов, Миклухо-Маклая ул., 6, Москва, 117198, РФ

^бИнститут проблем информатики Федерального исследовательского центра «Информатика и управление» РАН, Вавилова ул., 44-2, Москва, 119333, РФ

Постановка проблемы: одно из основных преимуществ технологии многоантенных систем, функционирующих в многопользовательском режиме, заключается в возможности передачи информации нескольким пользователям одновременно благодаря формированию нескольких параллельных потоков данных, что позволяет значительно повысить пропускную способность системы. Тем не менее спектральная эффективность может быть дополнительно увеличена за счет применения различных динамических стратегий обслуживания пользователей, учитывающих текущее состояние системы. Для оценки эффективности таких стратегий необходимы адекватные математические модели. **Цель:** разработка математических моделей многопользовательской системы MIMO, позволяющих оценить один из важных показателей качества восприятия пользователем — среднее время получения услуги, в данном случае — среднее время загрузки запрошенного пользователем файла. **Результаты:** представлены математические модели системы MIMO в виде систем массового обслуживания в дискретном и непрерывном времени для трех известных стратегий обслуживания пользователей. Предложены в замкнутой форме формулы для оценки среднего времени загрузки файла пользователем, позволяющие сделать вывод о наилучшей стратегии обслуживания без обязательности проведения численного эксперимента и, кроме того, имеющие меньшую вычислительную сложность по сравнению с уже существующими оценками, полученными с помощью математических моделей в непрерывном времени. Еще одним преимуществом построенных моделей является очевидная перспектива проведения дальнейших исследований по изучению влияния различных функций распределений длины файла на среднее время загрузки файла пользователем. Сделанные выводы об эффективности каждой из стратегий согласуются с более ранними работами, что подтверждается численным примером. **Практическая значимость:** разработанные модели многопользовательской системы MIMO могут быть использованы при выборе параметров функционирования реальной многопользовательской беспроводной системы MIMO.

Ключевые слова — многоантенные системы, многопользовательская беспроводная система MIMO, математическая модель системы MIMO, пропускная способность системы MIMO, время загрузки файла.

Для цитирования: Горбунова А. В., Медведева Е. Г., Гайдамака Ю. В., Шоргин В. С., Самуйлов К. Е. Эффективные стратегии обслуживания пользователей в многопользовательской системе MIMO. *Информационно-управляющие системы*, 2019, № 4, с. 69–81. doi:10.31799/1684-8853-2019-4-69-81

For citation: Gorbunova A. V., Medvedeva E. G., Gaidamaka Yu. V., Shorgin V. S., Samouylov K. E. Effective user service strategies in a multi-user MIMO system. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2019, no. 4, pp. 69–81 (In Russian). doi:10.31799/1684-8853-2019-4-69-81

Введение

В настоящее время остается актуальной задача увеличения пропускной способности современных систем сотовой связи, поскольку число пользователей постоянно растет и увеличиваются объемы передаваемой информации. В отличие от экстенсивного подхода увеличения числа базовых станций (БС), что не всегда выгодно с экономической и, возможно, технической точек зрения, одним из эффективных подходов является применение технологии многоантенной передачи (Multiple Input Multiple Output — MIMO) [1]. Основное отличие данной технологии от классической заключается в использовании нескольких антенн как на передающей, так и на приемной стороне [2]. Более подробное описание особенностей функциониро-

вания этой системы будет приведено ниже. Тем не менее, несмотря на повышение скорости передачи данных, обеспечиваемой применением самой технологии MIMO, можно добиться лучшей пропускной способности пользователей в таких системах за счет использования эффективных стратегий обслуживания [3–6]. В частности, в статье [6] было предложено несколько вариантов стратегий планировщика распределения радиоресурсов как в однопользовательском, так и в многопользовательском режимах. Авторами была построена математическая модель системы MIMO в виде системы массового обслуживания (СМО) в непрерывном времени в условиях каждой из предлагаемых стратегий, сделаны выводы и даны рекомендации относительно результативности той или иной стратегии на основе численных экспериментов.

В настоящей работе исследуется система ММО в условиях тех же стратегий, однако предложены математические модели в виде СМО в дискретном времени, что по своей сути ближе к цифровым системам, кроме того, в замкнутой форме получены формулы для оценки среднего времени загрузки файла пользователем, которые позволяют сделать выводы об эффективности каждой из стратегий, не прибегая к численным экспериментам.

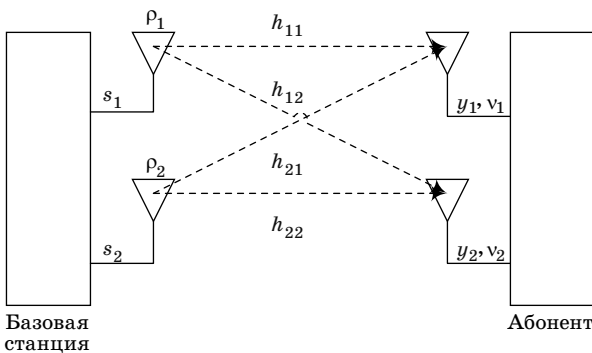
Параметры и показатели эффективности системы ММО

Рассмотрим многопользовательскую систему ММО, в которой несколько абонентских станций пользователей с M_R приемными антеннами каждая одновременно работают в определенной полосе частот с одной БС с M_T передающими антеннами. Пользователи системы ММО, находящиеся в зоне покрытия БС, подключаются к ней для того, чтобы получить файл конечного объема, после чего отключаются от системы.

Рассматривается модель с передачей данных по нисходящему каналу (downlink), т. е. от БС к пользователю (рис. 1), при этом важной характеристикой радиоканала является мощность сигнала на приемной стороне, которая зависит от местоположения пользователя. Мощность сигнала, принятого i -м пользователем, определяется следующим выражением:

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s} + \mathbf{v}_i,$$

где \mathbf{y}_i — вектор сигналов, принятых i -м пользователем, размерности $M_R \times 1$; \mathbf{s} — передаваемый вектор комплексных информационных символов размерности $M_T \times 1$; \mathbf{H}_i — прямоугольная комплексная матрица каналов связи для i -го пользователя размерности $M_R \times M_T$; \mathbf{v}_i — комплексный



■ **Рис. 1.** Схема передачи сигнала в системе ММО, $M_T = M_R = 2$, с одним пользователем

■ **Fig. 1.** MIMO system for single-user transmission, $M_T = M_R = 2$

случайный гауссов вектор размерности $M_R \times 1$ независимых одинаково распределенных шумов одинаковой мощности N , дБм, на каждой ветви разнесения. При этом каждая компонента вектора \mathbf{v}_i имеет комплексное нормальное распределение с нулевым математическим ожиданием и дисперсией σ^2 , равной $N_0/2$, т. е. матрица ковариации шумов равна $(N_0/2)\mathbf{I}_{M_R}$, где \mathbf{I}_{M_R} — единичная матрица размерности $M_R \times M_R$, а N_0 — спектральная плотность мощности шума. Также предполагается, что мощность сигнала, передаваемого БС, ограничена величиной P_T [дБм].

Метрикой качества передачи данных в радиоканале является отношение средней мощности сигнала к средней мощности шума, или среднее отношение сигнал/шум (Signal to Noise Ratio — SNR), т. е. $SNR = P_T/N$. Обычно отношение сигнал/шум измеряется в децибелах [дБ], поэтому для получения данной характеристики используется формула

$$SNR_{[дБ]} = P_{T[дБм]} - N_{[дБм]}.$$

Одним из основных показателей эффективности систем связи, в том числе системы с технологией ММО, является пропускная способность системы, т. е. максимально возможная (или достижимая) скорость безошибочной передачи данных (achievable rate) [2]. Прежде чем переходить непосредственно к анализу системы, приведем некоторые формулы, известные из теории информации, которые будут использоваться в дальнейшем. В предположении о том, что матрица \mathbf{H} каналов связи (далее — канальная матрица) известна на приемной стороне, в случае однопользовательской системы ММО пропускная способность, оптимизирующая распределение мощности каналов связи, вычисляется с помощью формулы [2, 7, 8]

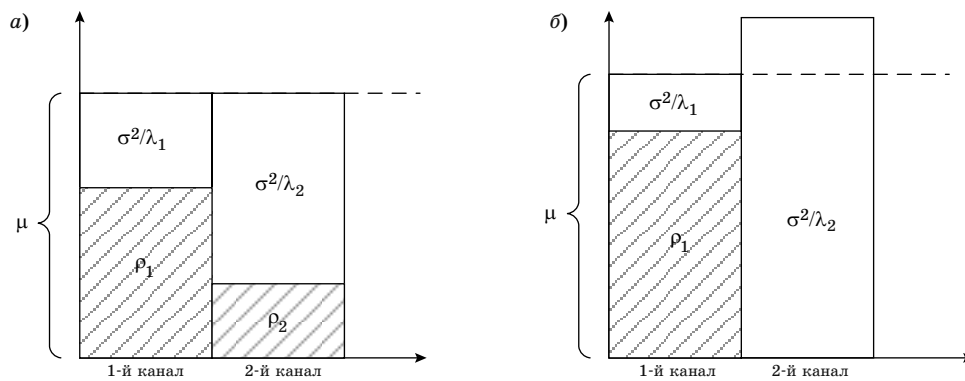
$$C(\mathbf{H}) = \sum_{k=1}^m \log_2 \left(1 + \frac{\rho_k}{\sigma^2} \lambda_k \right), \quad (1)$$

где $m = \min(M_T, M_R)$; λ_k — собственные числа матрицы \mathbf{H} ; ρ_k — мощность передачи в k -м пространственном радиоканале:

$$\rho_k = \left(\mu - \frac{\sigma^2}{\lambda_k} \right)^+.$$

Здесь $(x)^+ = \max(0, x)$; μ — так называемый «уровень заполнения водой» (water filling level), т. е. константа, которая выбирается исходя из условия нормировки (рис. 2):

$$\sum_{k=1}^m \rho_k = P_T.$$



■ **Рис. 2.** Пример применения «водоналивного алгоритма» для системы MIMO 2×2 : а — $\rho_1 + \rho_2 = P_T$; б — $\rho_1 = P_T$, $\rho_2 = 0$

■ **Fig. 2.** Example of water filling principle for MIMO system configuration 2×2

Идея «алгоритма заполнения водой», или «водоналивного алгоритма» (water filling principle) [7], заключается в неравномерном распределении мощности между передающими антеннами для того, чтобы максимизировать пропускную способность системы MIMO (рис. 2, а). При этом пропускная способность отдельных пространственных радиоканалов ρ_k , $k = 1, m$, может быть уменьшена до нуля (рис. 2, б) в зависимости от характеристик спектральной эффективности каждого подканала.

В случае многопользовательского режима передачи сигнал можно передавать нескольким пользователям одновременно. При этом, в отличие от однопользовательского режима, на скорость получения данных пользователем на приемной стороне оказывает влияние интерференция, т. е. деструктивное взаимное влияние радиосигналов, передаваемых на той же или близких частотах. Поэтому скорость получения данных на приемной стороне существенно зависит от набора пользователей, которым в многопользовательском режиме одновременно передаются данные. Таким образом, суммарная пропускная способность получается путем нахождения набора матриц ковариации передачи, которые ее максимизируют:

$$C_{SumRate}(\mathbf{H}_k) = \max_{\Sigma_k > 0, \sum_k T_r(\Sigma_k) \leq P_T} \times \sum_{k=1}^m \log_2 \left| \mathbf{I} + \sum_k \mathbf{H}_k \Sigma_k \mathbf{H}_k' \right|, k = 1, 2, \dots \quad (2)$$

Выражение $\Sigma_k > 0$ означает, что матрицы ковариации передачи положительно определены. Аргументы, которые максимизируют $C_{SumRate}$, определяют оптимальное подмножество пользователей и соответствующие матрицы ковариации передачи Σ_k [6].

Важным показателем качества восприятия пользователя (QoE) является среднее время получения услуги, в случае рассмотренной многопользовательской системы MIMO — среднее время загрузки запрошенного пользователем файла конечного объема. Для описанной технической системы время T загрузки файла соответствует интервалу от момента подключения пользователя к системе до момента отправки запроса на получение файла на БС до момента поступления последнего бита запрошенного файла в оборудование пользователя и отключения пользователя от системы. В статье этот показатель исследуется для трех стратегий обслуживания пользователей в системе MIMO: в однопользовательском (стратегии 1 и 2) и в многопользовательском (стратегия 3) режимах передачи данных, — предложенных в работе [6]. Работа планировщика распределения радиоресурсов на БС при применении каждой стратегии описана в следующих разделах. Построены математические модели в дискретном и непрерывном времени для сравнения эффективности исследованных стратегий.

Моделирование системы MIMO в дискретном времени

Для моделирования функционирования системы MIMO в дискретном времени разделим временную ось на интервалы (такты) постоянной длины Δt и сделаем следующие предположения.

1. Все запросы пользователей на загрузку файла, отправленные в течение текущего такта, поступают на БС одновременно в начале следующего такта.

2. Пользователи меняют свое местоположение в момент сразу после начала такта. В течение такта пользователи не перемещаются, так что для каждого пользователя скорость получения данных, зависящая от соответствующей его место-

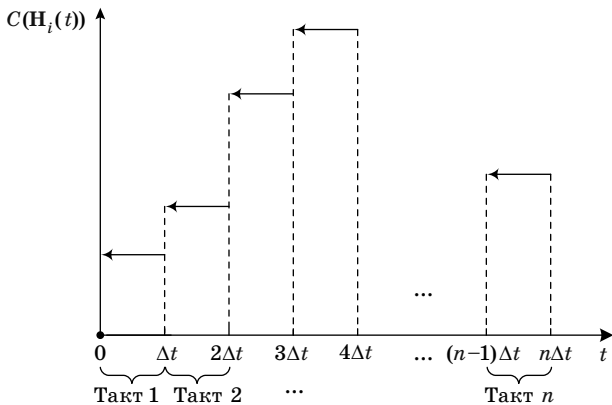
положению канальной матрицы \mathbf{H} , определяется в начале такта и не меняется до окончания такта.

3. Пользователи, завершившие загрузку файла в течение текущего такта, покидают систему одновременно в момент непосредственно перед началом следующего такта.

В этих предположениях определенные выше величины на каждом такте будут принимать различные, но неизменные в течение такта значения. При этом будем считать, что значения величин $y_i(t)$, $\mathbf{H}_i(t)$, $s(t)$, $v_i(t)$, $C(\mathbf{H}_i(t))$ для i -го пользователя и значение величины $C_{SumRate}(\mathbf{H}_k(t), k = 1, 2, \dots)$ для подмножества пользователей в момент $t \in ((n-1)\Delta t, n\Delta t]$ на такте с номером n — это значения, принимаемые непосредственно после начала i -го такта, например (рис. 3):

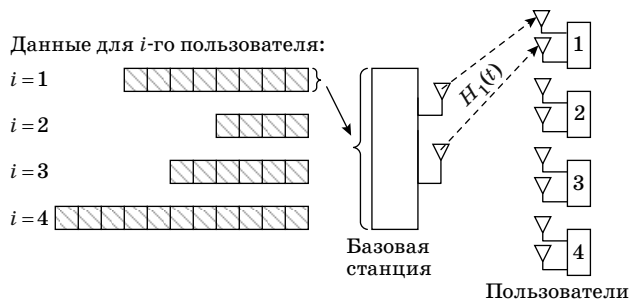
$$C(\mathbf{H}_i(t)) = C(\mathbf{H}_i((n-1)\Delta t + 0)), \\ t \in ((n-1)\Delta t, n\Delta t], n = 1, 2, \dots$$

На рис. 4 представлена схема модели многопользовательской системы MIMO с $M_T = 2$ передающими антеннами, обслуживающей четырех



■ Рис. 3. Пропускная способность канала i -го пользователя в системе MIMO в дискретном времени

■ Fig. 3. The capacity of the i -user in the MIMO system with a discrete time



■ Рис. 4. Модель многопользовательской системы MIMO 2×2 , обслуживающей четырех пользователей

■ Fig. 4. The model of multi-user 2×2 MIMO system serving four users

пользователей, каждый из которых имеет $M_R = 2$ приемные антенны. Здесь же схематически изображены файлы случайной длины, находящиеся в очереди на передачу на БС.

Предположим, что при наличии подключенных к системе MIMO пользователей БС непрерывно осуществляет передачу данных, что соответствует обслуживанию пользователей в соответствии с некоторой стратегией, в зависимости от которой передаваемый БС вектор комплексных информационных символов $s(t)$ будет предназначаться одному или нескольким пользователям одновременно. В статье рассмотрены три стратегии [6], согласно которым на n -м такте, $n \geq 1$, на обслуживание выбирается пользователь, который провел в системе больше всего времени (стратегия 1 обслуживания в порядке поступления пользователей в систему); пользователь, радиоканал которого обеспечивает максимальную пропускную способность на приемной стороне (стратегия 2 обслуживания пользователя с наилучшей пропускной способностью канала); подмножество пользователей с максимальной суммарной пропускной способностью (стратегия 3 обслуживания подмножества пользователей с наилучшей суммарной пропускной способностью).

Обслуживание пользователей системы MIMO можно описать случайным процессом в дискретном времени $X_n = (\xi_n, \mathbf{F}_n, \mathbf{A}_n)$, где случайная величина ξ_n — число заявок в системе; $\mathbf{F}_n = (F_{1n}, F_{2n}, \dots, F_{\xi_n n})$ — вектор остаточных длин заявок, находящихся в системе; $\mathbf{A}_n = (A_{1n}, A_{2n}, \dots, A_{\xi_n n})$ — вектор времен пребывания заявок, находящихся в системе, на n -м такте, $n \geq 1$. Заявка соответствует пользователю системы MIMO, а под остаточной длиной заявки подразумевается случайный объем данных, предназначенный конкретному пользователю, выраженный в каких-либо условных единицах, т. е. размер запрошенного пользователем файла (в момент поступления заявки) или оставшаяся для передачи часть файла (на последующих тактах).

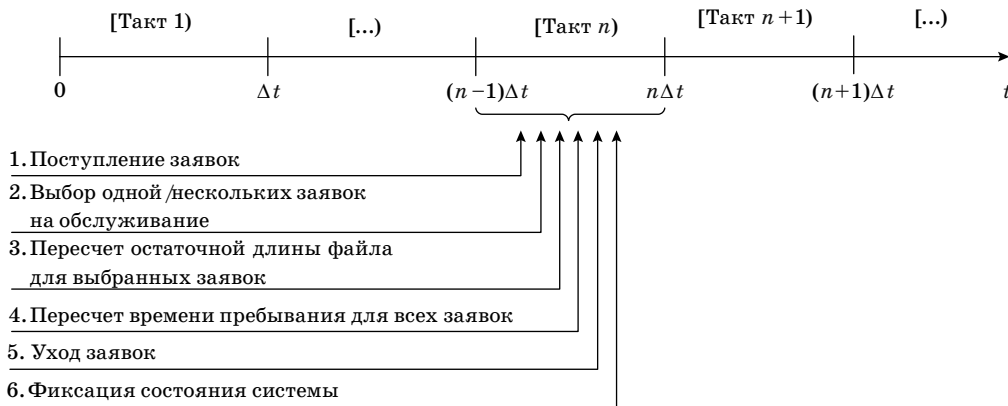
Будем считать, что согласно протоколу передачи данных в многопользовательской системе MIMO на интервале $[(n-1)\Delta t, n\Delta t]$, соответствующем n -му такту, $n \geq 1$, в системе последовательно совершаются действия, показанные на рис. 5.

В введенных обозначениях три рассмотренные стратегии обслуживания пользователей [6] принимают следующий вид. В описанной выше последовательности событий на n -м такте, $n \geq 1$, выбирается:

— согласно стратегии 1 — пользователь $i(n)$, который провел в системе больше всего времени:

$$i(n) = \arg \max_{1 \leq i \leq \xi_n} A_{in},$$

при этом пропускная способность системы $C(\mathbf{H})$ определяется по формуле (1):



■ **Рис. 5.** Временная диаграмма последовательности событий в СМО в дискретном времени
 ■ **Fig. 5.** Event sequence diagram of discrete time queuing system

$$C(\mathbf{H}_{i(n)}(n)) = \sum_{k=1}^m \log_2 \left(1 + \frac{\rho_k}{\sigma^2} \lambda_k \right);$$

— согласно стратегии 2 — пользователь $i(n)$, канал которого обладает максимальной пропускной способностью:

$$i(n) = \arg \max_{1 \leq i \leq \xi_n} C(\mathbf{H}_i(n)),$$

при этом $C(\mathbf{H})$ по-прежнему задается формулой (1);
 — согласно стратегии 3 — подмножество пользователей $N^\infty(n)$ с максимальной суммарной пропускной способностью:

$$N^\infty(n) = \arg \max_{P\{1,2,\dots,\xi_n\}} C_{SumRate}(\mathbf{H}_k(n), k=1, \dots, \xi_n),$$

при этом пропускная способность системы $C(\mathbf{H})$ определяется формулой (2):

$$C_{SumRate}(\mathbf{H}_k(n), k=1, \dots, \xi_n) = \max_{\sum_j(n) > 0, \sum_j(n) Tr(\sum_j(n)) \leq P_T} \log_2 \left| \mathbf{I} + \sum_j \mathbf{H}_j(n) \Sigma_j(n) \mathbf{H}_j'(n) \right|.$$

В обозначениях построенного случайного процесса последовательность событий на n -м такте принимает следующий вид.

1. Поступление в систему новой заявки/группы заявок в соответствии с геометрическим распределением с параметром a , где a — вероятность поступления заявки, $0 < a < 1$. При поступлении на n -м такте новой заявки число заявок/групп заявок в системе увеличивается на единицу: $\xi_n = \xi_{n-1} + 1$. Генерация случайной остаточной длины $F_{\xi_n n}$ для поступившей заявки/группы заявок. Время пребывания поступившей заявки/группы заявок в системе $A_{\xi_n n} = 1$. Генерация случайных элементов канальной матрицы $\mathbf{H}_{\xi_n}(n)$ для поступившей заявки/группы заявок.

2. Генерация случайных элементов канальной матрицы $\mathbf{H}_i(n)$ для каждой из уже обслуживавшихся заявок/групп заявок в системе, $1 \leq i \leq \xi_{n-1}$.

3. Расчет функции $C(\mathbf{H})$ для каждой заявки (стратегии 1 и 2) или для всех подмножеств заявок в системе (стратегия 3).

4. Выбор заявки $i(n)$ или подмножества заявок $N^\infty(n)$ на обслуживание в соответствии с одной из трех стратегий.

5. Пересчет остаточной длины $F_{i(n) n}$ для выбранной заявки

$$F_{i(n) n} = F_{i(n) n-1} - C(\mathbf{H}_{i(n)}(n))\Delta t$$

или остаточных длин $F_{k n}$ для каждой заявки из подмножества $N^\infty(n)$

$$F_{k n} = F_{k n-1} - C(\mathbf{H}_k(n))\Delta t, k \in N^\infty(n).$$

6. Пересчет времен пребывания всех заявок в системе — каждая компонента вектора \mathbf{A}_n увеличивается на единицу: $\mathbf{A}_n = \mathbf{A}_{n-1} + \mathbf{1}$.

7. Сбор статистики — скорость передачи данных на приемной стороне $C(\mathbf{H})$; время загрузки файла пользователем $T = A_{i n}$, если $\max(0, F_{i n}) = 0$.

8. Окончание обслуживания и уход из системы заявок, для которых $\max(0, F_{i n}) = 0$.

9. Фиксация состояния системы.

Построенный случайный процесс в дискретном времени $X_n = (\xi_n, \mathbf{F}_n, \mathbf{A}_n)$ может быть исследован с помощью аппарата цепей Маркова, при этом могут быть получены стационарные характеристики процесса. Далее для трех исследованных стратегий обслуживания пользователей проведено моделирование системы ММО с помощью известных моделей СМО в дискретном времени с групповым поступлением и ординарным обслуживанием $Geo^{[X]}|Geo|1$, а также с помощью моделей СМО в непрерывном времени $M|G|1$, которое

позволило получить в замкнутой форме формулы для оценки среднего времени загрузки файла пользователем и провести сравнение исследуемых стратегий обслуживания пользователей системы ММО.

Стратегия 1 обслуживания в порядке поступления пользователей в систему

Модель 1: $Geo_{\lambda\Delta t}^{[X]}|Geo_{EC(N)\Delta t}|1|_{\infty.e}$ Предположим, что на каждом такте с вероятностью $a = \lambda\Delta t$, $0 < a < 1$, на БС системы ММО поступает запрос пользователя, которому необходимо передать файл случайного объема. При этом в каждом файле содержится i ($i \geq 1$) бит (или целое положительное число других условных единиц количества информации) с вероятностью c_i . Считаем, что файл передается БС побитово в соответствии с геометрическим распределением времени обслуживания с параметром $b = \mu\Delta t$, $0 < b \leq 1$. Величины $\lambda > 0$ и $\mu > 0$ введены для того, чтобы в дальнейшем перейти к рассмотрению аналогичной системы в непрерывном случае, в которой эти параметры будут иметь смысл интенсивностей поступления и обслуживания соответственно.

Таким образом, систему ММО можем смоделировать СМО $Geo^{[X]}|Geo|1$, в которую поступает группа заявок (файл), а заявки обслуживаются прибором по одной (побитово). Функционирование СМО можно описать цепью Маркова, образованной числом заявок в системе непосредственно после окончания тактов $\{N(t), t \geq 0\}$, причем, как известно [9], стационарные вероятности $p_i = \lim_{t \rightarrow \infty} P(N(t) = i)$, $i > 0$, выражаются рекуррентно, а вероятность p_0 , как и некоторые вероятностно-временные характеристики системы, можно выразить с помощью производящих функций. Введем следующие обозначения для производящих функций распределения числа заявок в группе и стационарного распределения вероятностей p_i :

$$C(z) = \sum_{i=1}^{\infty} c_i z^i,$$

$$P(z) = \sum_{i=0}^{\infty} p_i z^i, \beta(z) = \sum_{i=1}^{\infty} b_i z^i = \sum_{i=1}^{\infty} b(\bar{b})^{i-1} z^i = \frac{bz}{1-\bar{b}z},$$

где $\bar{b} = 1 - b$, а также среднего числа заявок в группе:

$$\hat{c} = \sum_{i=0}^{\infty} i c_i.$$

Тогда решение системы уравнений равновесия с помощью производящих функций примет вид [8]

$$P(z) = b(\bar{a} + aC(z)) \left(b\bar{a} - \frac{a(z - (\bar{b}z + b)C(z))}{1-z} \right)^{-1} p_0,$$

откуда, пользуясь тем, что $P(1) = 1$, получаем $p_0 = 1 - a\hat{c}/b$, $a\hat{c}/b = \rho$ — загрузка системы ($\rho < 1$ — необходимое и достаточное условие стационарности системы). Стационарное распределение времени ожидания начала обслуживания первой заявки из группы или, что эквивалентно, времени ожидания начала обслуживания группы заявок имеет производящую функцию следующего вида [9]:

$$w(z) = \frac{1}{\beta(z)} (b(\beta(z) - 1)p_0 + (\bar{b}\beta(z) + b)P(\beta(z))).$$

Тогда среднее время ожидания начала обслуживания группы заявок

$$w = w'(1) = 1 - \rho + \frac{1}{b}(N + \bar{b} - 1),$$

где среднее число заявок в системе

$$N = P'(1) = \frac{a \ 2\hat{c}(1 - a\hat{c}) + C''(1)}{b \ 2(1 - \rho)},$$

а среднее время пребывания группы заявок в системе, соответствующее времени загрузки файла, определяется выражением $E[T] = w + \hat{c}\hat{b}$, т. е.

$$E[T] = 1 - \rho + \frac{1}{b}(N + \bar{b} - 1) + \hat{c}\hat{b}. \tag{3}$$

Заметим, что в данном случае

$$\hat{b} = \frac{1}{b} = \frac{1}{E[C(N)]\Delta t}, \tag{4}$$

причем величина $C(N)$ должна быть выражена в соответствующих целых условных единицах, например, битах в секунду [бит/с].

Если допустить, что число заявок в группе имеет геометрическое распределение с параметром c , $\bar{c} = 1 - c$, то

$$\hat{c} = \frac{1}{c}, C''(1) = \frac{2\bar{c}}{c^2}.$$

Отметим, что выбор СМО $Geo^{[X]}|Geo|1|_{\infty}$ в качестве модели ММО дает некоторое преимущество в части возможности анализа различных функций распределений длины группы заявок (размера файла) и влияния этих вариаций на такие важные характеристики функционирования системы, как среднее время пребывания заявки в системе и средняя длина очереди [11]. Анализ

чувствительности этих параметров к виду функции распределения длины сообщения может стать предметом дальнейших исследований.

Модель 2: $Geo_{\lambda\Delta t} | G_{E[C(H)\Delta t/E[F]]} | 1 | \infty$. Предположим, что поток поступающих в систему ММО запросов пользователей является геометрическим с вероятностью поступления запроса на такте n , равной $0 < a < 1$, и с дискретной функцией распределения $B(t)$ времени обслуживания запроса с средним значением \hat{b} .

Тогда систему ММО, функционирующую в дискретном времени с условием обслуживания пользователей в порядке поступления, можем смоделировать с помощью дискретной СМО $Geo|G|1|\infty$. В этом случае среднее время пребывания заявки в системе описывается выражением [9, 10]

$$E[T] = \hat{b} + \frac{ab^{(2)} - \rho}{2(1-\rho)}, \quad (5)$$

где $\rho = a\hat{b} < 1$ — загрузка системы; $b^{(2)}$ — второй момент времени обслуживания заявки. Поскольку время обслуживания каждого пользователя зависит от размера передаваемой информации и скорости ее передачи, указанные характеристики можем оценить следующим образом:

$$\hat{b} = \frac{E[F]}{E[C(H)\Delta t}; \quad b^{(2)} = \frac{E[F]^2}{E[C(H)]^2 \Delta t^2},$$

где F — случайная величина объема информации, передаваемой пользователю.

Кроме того, можем оценить среднее число пользователей, находящихся в системе, по моментам окончания обслуживания:

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

Заметим, что фактически величины $E[F]$ и \hat{b} являются тождественными, так как характеризуют средний объем передаваемой информации. Однако в силу того, что в СМО $Geo^{[X]}|Geo|1|\infty$ ($M^{[X]}|M|1|\infty$ — в непрерывном случае) величина \hat{b} определяет среднее число заявок в поступающей группе, в соответствующих моделях на случайную величину размера файла накладывается очевидное требование дискретности и целочисленности ее значений. Данное ограничение объясняет выбор «бита» в качестве единицы измерения объема передаваемой информации в описанных моделях, но при этом не исключает возможности выбора других единиц информации с соблюдением указанных требований. Что касается величины F , то на нее подобные ограничения не распространяются, поэтому она может быть и непрерывной.

Стратегия 2 обслуживания пользователя с наилучшей пропускной способностью канала

В случае выбора на обслуживание на каждом такте пользователя с максимальной пропускной способностью логично предположить, что системе ММО возможно смоделировать с помощью СМО $Geo^{[X]}|Geo|1|\infty$ с случайным выбором заявки из очереди (Random Order of Service — ROS [12], или Service in Random Order — SIRO [13]), т. е. заявка, находящаяся в очереди, может быть выбрана для обслуживания на приборе с вероятностью $1/i$, зависящей только от числа заявок в очереди.

Рассмотрим модели, в которых обслуживание не может прерываться на каждом такте, т. е. выбранная заявка находится на приборе до тех пор, пока ее обслуживание не будет завершено. Заметим, что допущение об отсутствии прерывания обслуживания должно давать хорошее приближение в случае, если среднее время обслуживания заявки (одного бита или одной условной единицы количества информации) будет равно длительности такта.

Модель 1: $Geo_{\lambda\Delta t}^{[X]} | Geo_{E[\max C(H)\Delta t]} | 1 | \infty$ с случайным выбором заявки из очереди. Стационарное распределение числа заявок в системе в случае дисциплины случайного выбора из очереди не отличается от стационарного распределения в случае стандартной дисциплины обслуживания заявок — в порядке поступления, хотя стационарные распределения времен пребывания заявок в системе будут отличаться [9, 12, 13]. Тем не менее средние времена ожидания начала обслуживания и пребывания заявки так же, как и группы заявок, в системе будут принимать одинаковые значения в обоих случаях в силу формулы Литтла. Поэтому выражение для среднего времени пребывания заявки в системе сохранит вид (3) за исключением того, что среднее время обслуживания одной заявки в силу условия максимальной пропускной способности будет равно

$$\hat{b} = \frac{1}{b} = \frac{1}{E[\max C(H)\Delta t]}. \quad (6)$$

Таким образом, в среднем скорость обслуживания группы заявок в случае выбора на обслуживание пользователя с максимальной пропускной способностью будет выше, а соответственно, и время пребывания заявки в системе будет лучше (меньше), чем в случае обслуживания пользователей в порядке их поступления.

Модель 2: $Geo_{\lambda\Delta t} | G_{E[\max C(H)\Delta t/E[F]]} | 1 | \infty$ с случайным выбором заявки из очереди. В случае СМО $Geo|G|1|\infty$ с случайным выбором заявки из очереди ситуация будет аналогична описанной для модели 1, т. е. стационарные вероятности числа заявок в системе будут совпадать с стационарными вероятностями в стандартной СМО $Geo|G|1|\infty$, так

же как и среднее время пребывания заявки в системе. Таким образом, для времени пребывания заявки в системе будет справедлива формула (4), где \hat{b} и $b^{(2)}$ можно оценить следующим образом:

$$\hat{b} = \frac{1}{b} = \frac{E[F]}{E[\max C(\mathbf{H})]\Delta t}; \quad b^{(2)} = \frac{E[F]^2}{E[\max C(\mathbf{H})]^2 \Delta t^2}.$$

Стратегия 3 обслуживания подмножества пользователей с наилучшей суммарной пропускной способностью

Модель 1: $Geo_{\lambda \Delta t}^{[X]} | Geo_{E[C_{SumRate}(\mathbf{H})]\Delta t} | 1 | \infty$ с случайным выбором заявки из очереди. В качестве модели ММО с заявленной стратегией естественно выбрать СМО типа $Geo^{[X]} | Geo | m | \infty$ с случайным выбором на обслуживание, поскольку в данном случае подразумевается, что обслуживаться могут уже не один, а несколько пользователей одновременно. Тем не менее здесь возникает серьезный вопрос о подборе основных параметров упомянутой СМО, в частности — числа приборов и интенсивностей обслуживания на них. Так, для оценки числа приборов СМО можно выбрать среднее число пользователей, которые могут обслуживаться одновременно в системе ММО, причем их количество не должно превышать M_T , а в качестве интенсивности обслуживания на каждом из приборов — величину, равную отношению $E[C_{SumRate}(\mathbf{H})]\Delta t$ к выбранному числу приборов, либо средние индивидуальные интенсивности обслуживания из работы [14]. Данная аппроксимация представляется допустимой в силу естественного предположения о том, что частота события $\{E[C_{SumRate}(\mathbf{H})] \geq E[\max C(\mathbf{H})]\}$ будет гораздо выше частоты противоположного ему события. Однако к решению проблемы можно подойти существенно эффективнее и ограничиться выбором СМО с $m = 1$ прибором, среднее время обслуживания на котором

$$\hat{b} = \frac{1}{b} = \frac{1}{E[C_{SumRate}(\mathbf{H})]\Delta t}. \quad (7)$$

Таким образом, за счет увеличения интенсивности обслуживания на одном приборе будет моделироваться ситуация параллельного обслуживания пользователей с гораздо меньшей интенсивностью на каждом из них. В результате оценка среднего времени пребывания заявки в системе будет иметь знакомый вид (3).

В силу того, что рассматриваемая стратегия подразумевает выполнение условия $E[C_{SumRate}(\mathbf{H})] \geq E[\max C(\mathbf{H})]$, очевидно, что стратегия выбора на обслуживание подмножества пользователей с наибольшей суммарной пропускной способностью будет наилучшей среди трех вышеупомянутых стратегий с точки зрения оценки среднего времени пребывания заявки в системе.

Модель 2: $Geo_{\lambda \Delta t} | G_{E[C_{SumRate}(\mathbf{H})]\Delta t / E[F]} | 1 | \infty$ с случайным выбором заявки из очереди. В случае СМО $Geo | G | 1 | \infty$ с случайным выбором заявки из очереди ситуация будет аналогична описанной выше, т. е. стационарные вероятности числа заявок в системе будут совпадать с стационарными вероятностями в стандартной СМО $Geo | G | 1 | \infty$, так же как и среднее время пребывания заявки в системе. Таким образом, для времени пребывания заявки в системе справедлива формула (4), где \hat{b} и $b^{(2)}$ вычисляются как

$$\hat{b} = \frac{1}{b} = \frac{E[F]}{E[C_{SumRate}(\mathbf{H})]\Delta t}; \quad b^{(2)} = \frac{E[F]^2}{E[C_{SumRate}(\mathbf{H})]^2 \Delta t^2}.$$

Моделирование системы ММО в непрерывном времени

Предположим, что система ММО функционирует в непрерывном времени, тогда все определенные выше величины $y_i(t)$, $\mathbf{H}_i(t)$, $\mathbf{s}(t)$, $\mathbf{v}_i(t)$, $C(\mathbf{H}_i(t))$, $C_{SumRate}(\mathbf{H}_k(t), k = 1, 2, \dots)$ изменяют значения в произвольный момент t ($t \geq 0$).

Стратегия 1 обслуживания в порядке поступления пользователей в систему

Модель 3: $M_{\lambda}^{[X]} | M_{E[C(\mathbf{H})]} | 1 | \infty$. Рассмотрим непрерывный аналог СМО $Geo^{[X]} | Geo | 1 | \infty$ — $M^{[X]} | M | 1 | \infty$. Для этой системы среднее время пребывания заявки в системе, т. е. сумма времени ожидания начала обслуживания первой заявки из группы и времени обслуживания всей группы заявок, вычисляется по формуле

$$E[T] = \frac{\hat{c} + c^{(2)} \lambda}{2(1-\rho) \mu^2} + \frac{\hat{c}}{\mu}, \quad (8)$$

где λ — интенсивность входящего потока групп заявок; μ — интенсивность обслуживания одной заявки, которая характеризуется формулой

$$\frac{1}{\mu} = \frac{1}{E[C(\mathbf{H})]};$$

$\rho = \hat{c}\lambda/\mu < 1$ — загрузка системы; \hat{c} — среднее число заявок в группе; $c^{(2)}$ — второй момент числа заявок в группе. В случае геометрического распределения числа заявок в группе величины \hat{c} и $c^{(2)}$ определяются следующим образом:

$$\hat{c} = \frac{1}{c}, \quad c^{(2)} = \frac{2-c}{c^2}.$$

Модель 4: $M_{\lambda} | G_{E[F]/E[C(\mathbf{H})]} | 1 | \infty$. Поведение системы ММО в случае обслуживания в порядке поступления пользователей можем смоделировать СМО $M | G | 1 | \infty$, в которой среднее время пребы-

вания заявки в системе описывается известной формулой

$$E[T] = \frac{1}{\mu} + \frac{\lambda m^{(2)}}{2(1-\rho)}, \quad (9)$$

где $\rho = \lambda\mu^{-1} < 1$ — загрузка системы, среднее время обслуживания μ^{-1} можно оценить следующим образом:

$$\frac{1}{\mu} = \frac{E[F]}{E[C(H)]} = \frac{\int_0^{\infty} sf(s)ds}{E[C(H)]},$$

а второй момент случайной величины времени обслуживания

$$m^{(2)} = \frac{E[F]^2}{E[C(H)]^2} = \frac{\int_0^{\infty} s^2 f(s)ds}{E[C(H)]^2},$$

F — непрерывная случайная величина объема данных, передаваемых пользователю, с плотностью распределения $f(\cdot)$.

Стратегия 2 обслуживания пользователя с наилучшей пропускной способностью канала

Модель 3: $M_{\lambda}^{[X]}|M_{E[\max C(H)]}|1|_{\infty}$ с случайным выбором заявки из очереди. Для СМО $M^{[X]}|M|1|_{\infty}$ с случайным выбором заявки из очереди среднее время пребывания заявки в системе определяется формулой (8), где среднее время обслуживания одной заявки определяется выражением

$$\frac{1}{\mu} = \frac{1}{E[\max C(H)]}.$$

Модель 4: $M_{\lambda}|G_{E[F]/E[\max C(H)]}|1|_{\infty}$ с случайным выбором заявки из очереди. Среднее время пребывания заявки в системе $M|G|1|_{\infty}$ с случайным выбором заявки из очереди будет принимать такое же значение, как и в условиях классической СМО $M|G|1|_{\infty}$ в силу формулы Литтла. Поэтому выражение для среднего времени пребывания заявки в системе сохранит вид (9) за исключением того, что в силу условия максимальной пропускной способности для μ^{-1} будет справедливо

$$\frac{1}{\mu} = \frac{E[F]}{E[\max C(H)]}, \quad (10)$$

а для второго момента времени обслуживания

$$m^{(2)} = \frac{E[F]^2}{E[\max C(H)]^2} = \frac{\int_0^{\infty} s^2 f(s)ds}{E[\max C(H)]^2},$$

т. е. в среднем скорость обслуживания будет выше, а соответственно, и время пребывания заявки в системе будет меньше.

Модель 5: $M_{\lambda}|G_{E[F]/E[\max C(H)]}|1|_{\infty}$ с разделением процессора. Непрерывным аналогом дискретной СМО с выбором на каждом такте пользователя с максимальной пропускной способностью может стать СМО $M|G|1$ — EPS, т. е. СМО с справедливым разделением процессора. Такой выбор может быть обусловлен тем, что предельным вариантом циклической дисциплины кругового доступа (round-robin) является дисциплина обслуживания, известная как справедливое разделение процессора [15–17]. Кроме того, заметим, что в определенных условиях можно говорить о равенстве времен пребывания пользователей в СМО с разделением процессора и с случайным выбором заявки из очереди [18, 19].

Таким образом, пользуясь известной формулой для среднего времени пребывания в системе $M|G|1$ — EPS заявки с средней длиной $E[F]$, получим простое выражение для интересующей нас оценки среднего времени пребывания заявки в системе, соответствующего времени получения услуги пользователем системы ММО:

$$E[T] = \frac{E[F]}{1-\rho}, \quad (11)$$

где $\rho = \lambda/\mu < 1$, а среднее время обслуживания μ^{-1} , в свою очередь, можем оценить с помощью выражения (10).

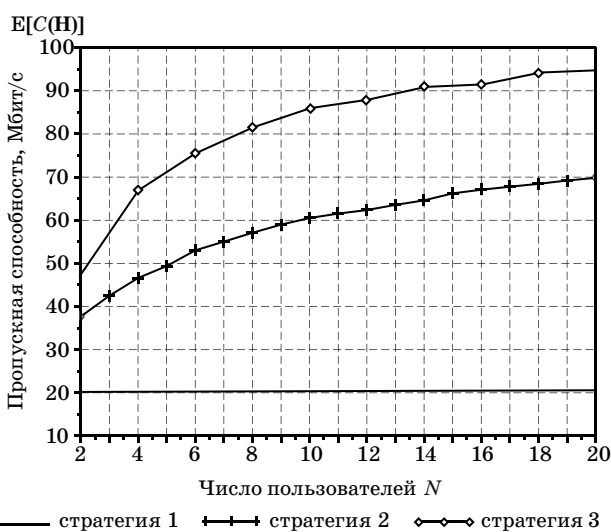
Стратегия 3 обслуживания подмножества пользователей с наилучшей пропускной способностью

Ситуацию выбора для обслуживания подмножества пользователей с наилучшей пропускной способностью можно смоделировать с помощью трех СМО, построенных в предыдущем подразделе для стратегии 2 обслуживания пользователя с наилучшей пропускной способностью канала. Для времени пребывания заявки в каждой из систем будут справедливы соответствующие формулы (8), (9) и (11), в которых при определении величин μ^{-1} и $m^{(2)}$ будет использоваться величина $E[C_{SumRate}(H)]$ вместо $E[\max C(H)]$ или $E[C(H)]$. Поскольку ожидаемо, что $E[C_{SumRate}(H)] \geq E[\max C(H)]$, то среднее время пребывания заявки в системе будет наилучшим (наименьшим) среди всех трех стратегий.

Прежде чем переходить к численному моделированию и сравнению с аналитическими результатами, необходимо отметить, что при большой нагрузке, а точнее перегрузке ($E[F] > E[T]$), моделирование системы ММО с помощью СМО с разделением процессора или случайным выбором на обслуживание уже не будет являться объективным, т. е. фактически не будет применимо [15, 16], поэтому при большом числе пользователей в системе более адекватной моделью станет СМО $M|G|_{\infty}$ ($Geo|G|_{\infty}$ — в дискретном случае).

Численный эксперимент

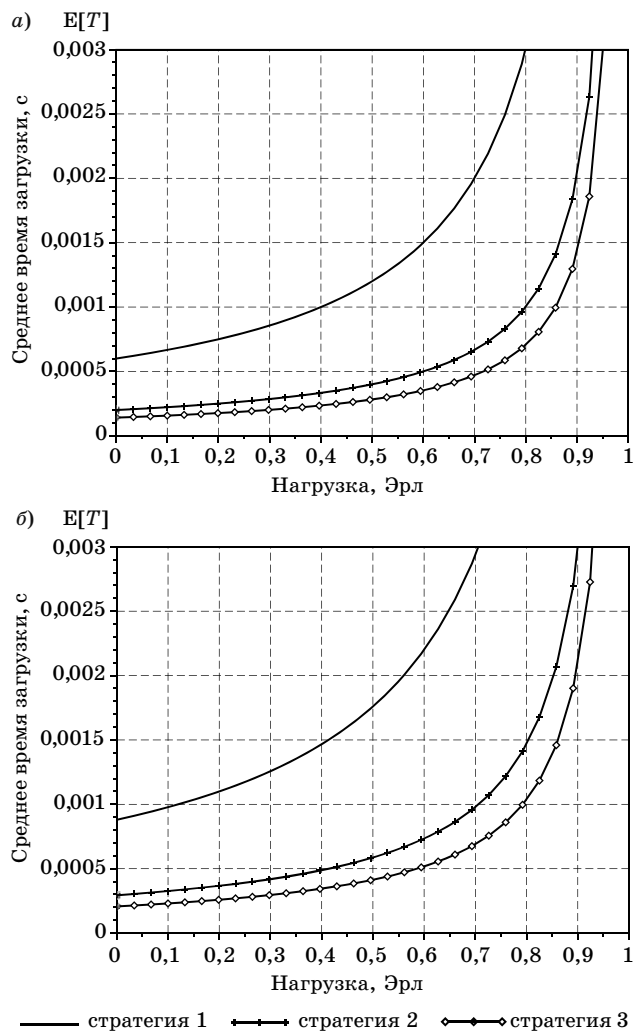
Организация численного эксперимента базируется на характеристиках многопользовательской передачи MIMO стандарта 802.11 WLANs [20] со спецификацией построения сетей высокой и сверхвысокой пропускной способности (High Throughput, Very High Throughput) в частотных диапазонах 2,4 и 5 ГГц, достигающих максимальной скорости передачи данных до 7 Гбит/с. Одно из важнейших отличий стандарта IEEE 802.11 — это возможность объединения нескольких MAC-пакетов в один пакет физического уровня A-MPDU (Aggregated MAC Protocol Data Unit), причём один пакет A-MPDU может агрегировать несколько пакетов MSDU (MAC Service Data Unit), принадлежащих различным приложениям и обладающих различными требованиями к качеству предоставления услуг (QoS). При разработке исходных данных для численного эксперимента по выбору стратегии загрузки файлов для наилучшей пропускной способности системы в случае обслуживания одного или нескольких пользователей с помощью формул (1), (2) при заданной средней пропускной способности $E[C(\mathbf{H})] = 20$ Мбит/с для ширины полосы 20 МГц были получены средние значения пропускной способности $E[\max C(\mathbf{H})]$ для стратегии 2 и $E[C_{SumRate}(\mathbf{H})]$ для стратегии 3 в зависимости от числа пользователей системы MIMO (рис. 6). При этом для моделирования перемещения пользователей системы MIMO на каждом такте разыгрывалось случайное приращение элементов канальной матрицы \mathbf{H} , имеющее комплексное нормальное распределение CN(0, 1) [21].



■ Рис. 6. Средняя пропускная способность в зависимости от числа заявок в системе

■ Fig. 6. Ergodic capacity depending on the number of requests in the queuing system

На рис. 7, а и б приведены графики зависимости среднего времени пребывания заявки в системе, полученные с помощью формулы (3), на примере трафика UDP с негарантированным типом услуги Best Effort [22]. При $N = 10$ для заданной средней пропускной способности $E[C(\mathbf{H})] = 20$ Мбит/с и найденных средних $E[\max C(\mathbf{H})] = 60$ Мбит/с и $E[C_{SumRate}(\mathbf{H})] = 85$ Мбит/с при длине одного такта $\Delta t = 10^{-8}$ с параметр b_1 геометрического распределения для стратегии 1 обслуживания в порядке поступления по формуле (4) равен $b_1 = 0,2$. Параметры $b_2 = 0,6$ и $b_3 = 0,85$ определяются для стратегии 2 по формуле (6) и для стратегии 3 по формуле (7) аналогично. Из условия стационарности следует, что при $\rho \in (0; 1)$ вероятность $a \in \left[0; \frac{b_{1,2,3}}{\hat{c}}\right)$. Для численного эксперимента



■ Рис. 7. Сравнение среднего времени загрузки файла для трех стратегий при $\hat{c} = 1,5$ КБ (а) и $\hat{c} = 2,2$ КБ (б)

■ Fig. 7. Comparison of the average file transfer delay for three strategies for $\hat{c} = 1,5$ КБ (а) и $\hat{c} = 2,2$ КБ (б)

та используем значения средней длины заявки \hat{c} , равные 1,5 и 2,2 КБ [22–24].

Напомним, что в соответствии с анализом построенных аналитических моделей в виде СМО наилучшей стратегией с точки зрения наименьшего времени пребывания заявки в системе является стратегия с выбором на обслуживание подмножества пользователей с максимальной пропускной способностью, что подтверждается графиками на рис. 7.

Однако, как ни парадоксально, обслуживание пользователей в порядке поступления в некоторых ситуациях может оказаться выгоднее обслуживания пользователей с наилучшей пропускной способностью канала с точки зрения снижения стоимости обслуживания и уменьшения времени пребывания в системе. Прийти к такому выводу можно, если продолжать проводить аналогию между системой ММО (в условиях стратегии 2 обслуживания) и СМО с случайным выбором на обслуживание [12, 25] при не слишком большой разнице между значениями $E[C(\mathbf{H})]$ и $E[\max C(\mathbf{H})]$, а, возможно, даже наблюдая близость значений $E[C(\mathbf{H})]$ и $E[C_{SumRate}(\mathbf{H})]$ в условиях малого числа антенн у БС. В этих случаях естественно предположить, что чем больше времени пользователь пребывает в системе, тем меньше для него становится вероятность попасть на обслуживание при том, что вероятность поступления очередного пользователя становится выше. Таким образом, разброс реального времени пребывания для конкретного пользователя в системе довольно велик. То есть речь идет о том, что, помимо значения среднего времени пребывания заявки в системе, важен еще и второй момент этой случайной величины, который, как известно, для стратегии обслуживания пользователей в порядке их поступления будет меньше, чем, в частности, для СМО с обслуживанием в случайном порядке [12, 25] $E[T_{FCFS}]^2 \leq E[T_{SIRO}]^2$. Поскольку возможность такой ситуации не исключена и сделанное предположение имеет под собой довольно веские

основания, дальнейшие исследования в этом направлении представляют интерес для получения более точных рекомендаций для выбора параметров функционирования реальной физической системы ММО.

Заключение

В работе проанализированы три стратегии обслуживания пользователей с помощью математических моделей в виде СМО в дискретном и непрерывном времени. Применение теории дискретных систем обслуживания к моделированию процесса планирования [26] способствовало построению симулятора для приближенного анализа методом Монте-Карло, а также упростило задачу получения аналитических результатов с помощью предложенных формул в замкнутой форме. Построенные модели дают возможность оценить среднее время пребывания заявки в системе, характеризующее один из важнейших показателей качества обслуживания многопользовательской системы ММО — время получения услуги пользователем. Полученные характеристики позволяют сделать вывод об эффективности каждой из исследованных стратегий. В дальнейшем предполагается построить математические модели, учитывающие возможность прерывания обслуживания на каждом такте, для получения более точных оценок, проанализировать второй момент времени пребывания заявки в системе, а также исследовать влияние различных распределений длины группы заявок (размера передаваемых файлов) на время пребывания заявки в системе.

Финансовая поддержка

Исследование выполнено при финансовой поддержке Министерства образования и науки Российской Федерации, проект № 2.3397.2017/4.6.

Литература

1. Solomitckii D., Petrov V., Nikopour H., Akdeniz M., Orhan O., Himayat N., Talwar S., Andreev S., Koucheryavy Y. Ray-based evaluation of dual-polarized MIMO in (Ultra)-dense millimeter-wave urban deployments. *IEEE Vehicular Technology Conference*, 2018, pp. 1–7. doi: 10.1109/VTCSpring.2018.8417788
2. Бакулин М. Г., Варукина Л. А., Крейнделин В. Б. *Технология ММО: принципы и алгоритмы*. М., Горячая линия-Телеком, 2014. 244 с.
3. Borst S. User-level performance of channel-aware scheduling algorithms in wireless data networks. *IEEE INFOCOM 2003. Twenty-second Annual Joint*

Conference of the IEEE Computer and Communications Societies, 2003, vol. 1, pp. 321–331. doi:10.1109/INFCOM.2003.1208684

4. Bellalta B., Oliver M. A space-time batch-service queueing model for multi-user MIMO communication systems. *Proceedings of the 12th International Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems*, Tenerife, 2009, pp. 357–364. doi:10.1145/1641804.1641866
5. Bellalta B., Daza V., Miquel O. An approximate queueing model for multi-rate multi-user MIMO systems. *IEEE Communications Letters*, 2011, vol. 15, no. 4, pp. 392–394. doi:10.1109/LCOMM.2011.020311.102434

6. Airy M., Shakkattai S., Heath R. W. Spatially greedy scheduling in multi-user MIMO wireless systems. *The Thirty-Seventh Asilomar Conference on Signals, Systems Computers*, 2003, vol. 1, pp. 982–986. doi:10.1109/ACSSC.2003.1292064
7. Jankiraman M. *Space-Time Codes and MIMO Systems*. Artech House, 2004. 327 p.
8. Telatar E. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 1999, vol. 10, pp. 585–595. <https://doi.org/10.1002/ett.4460100604>
9. Печинкин А. В., Разумчик Р. В. *Системы массового обслуживания в дискретном времени*. М., Физматлит, 2018. 432 с.
10. Alfa Attahiru S. *Applied Discrete-Time Queues*. Springer-Verlag, New York, 2016. 383 p.
11. Гайдамака Ю. В., Зарипова Э. Р., Орлов Ю. Н. К анализу эффектов группового поступления сигнальных сообщений на время ожидания начала обслуживания. *Информационные технологии и вычислительные системы*, 2015, № 3, с. 74–80.
12. Norikazu K., Takagi H., Yutaka T., Hong S.-J., Toshiharu H. Waiting time analysis for $M^{[X]}|G|1$ queues with/without vacations under random order of service discipline. *Journal of the Operations Research Society of Japan*, 2000, vol. 43, no. 4, pp. 455–468. doi:10.15807/jorsj.43.455
13. Takagi H. A note on the response time in $M|G|1$ queues with service in random order and bernoulli feedback. *Journal of the Operations Research Society of Japan*, 1996, vol. 39, no. 4, pp. 486–500. doi:10.15807/jorsj.39.486
14. Gesbert D., Kountouris M., Heath R., Chae C.-B., Sälzer T. From single user to multiuser communications: shifting the MIMO paradigm. *IEEE Signal Process*, 2007, vol. 24, pp. 36–46.
15. Yashkov S. F. Processor-sharing queues: Some progress in analysis. *Queueing Systems*, 1987, vol. 2, no. 1, pp. 1–17.
16. Яшков С. Ф. Математические вопросы теории систем обслуживания с разделением процессора. *Итоги науки и техники. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика»*, 1990, т. 29, с. 3–82.
17. Яшков С. Ф., Яшкова А. С. Эгалитарное разделение процессора. *Информационные процессы*, 2006, т. 6, № 4, с. 396–444.
18. Borst S. C., Voxma O. J., Morrison J. A., Núñez Queija R. The equivalence between processor sharing and service in random order. *Operations Research Letters*, 2003, vol. 31, no. 4, pp. 254–262. doi:10.1016/S0167-6377(03)00006-3
19. Yashkov S. F. On random order of service and processor sharing. *Information Processes*, 2006, vol. 6, no. 2, pp. 160–163.
20. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, IEEE std 802.11, 2016.
21. Foschini G., Gans M. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 1998, no. 6, p. 311. <https://doi.org/10.1023/A:1008889222784>
22. Daldoul Y., Meddour D., Ksentini A. IEEE 802.11n/ac data rates under power constraints. *IEEE International Conference on Communications*, Kansas City, USA, May 20–24, 2018. doi:10.1109/ICC.2018.8422697
23. Gast M. S. *802.11ac: A Survival Guide: Wi-Fi at Gigabit and Beyond*. O'Reilly Media, 2013. 154 p.
24. Sharon O., Alpert Y. A new aggregation based scheduling method for rapidly changing IEEE 802.11ac wireless channels. *Wireless Sensor Network*, 2016, vol. 8, no. 8, pp. 145–165. doi:10.4236/wsn.2016.88013
25. Saaty T. L. *Elements of Queuing Theory with Applications*. Dover Publications, 1983. 423 p.
26. Sopin E., Ageev K., Samouylov K. Approximate analysis of the limited resources queuing system with signals. *Proceedings — European Council for Modelling and Simulation ECMS*, 2019, vol. 33, no. 1, pp. 462–465.

UDC 621.39

doi:10.31799/1684-8853-2019-4-69-81

Effective user service strategies in a multi-user MIMO systemGorbunova A. V.^a, PhD, Phys.-Math., Associate Professor, orcid.org/0000-0001-7073-7022Medvedeva E. G.^{a,b}, Assistant Professor, orcid.org/0000-0001-9104-708X, medvedeva-eg@rudn.ruGaidamaka Yu. V.^{a,b}, Dr. Sc., Phys.-Math., Professor, orcid.org/0000-0003-2655-4805Shorgin V. S.^b, PhD, Tech., Senior Scientist, orcid.org/0000-0002-1090-544XSamouylov K. E.^{a,b}, Dr. Sc., Tech., Professor, orcid.org/0000-0002-6368-9680, samouylov-ke@rudn.ru^aPeoples' Friendship University of Russia (RUDN University), 13, Miklukho-Maklaya St., 117198, Moscow, Russian Federation^bThe Institute of Informatics Problems of the Federal Research Center «Computer Science and Control» of the RAS, 44–2, Vavilov St., 119333, Moscow, Russian Federation**Introduction:** One of the main advantages of multiple antenna systems is the support of multi-user connections, simultaneously transmitting parallel data streams. This can significantly increase the overall system throughput. However, the spectral efficiency

can be additionally increased through the usage of various dynamic strategies which take into account the current state of the system. Here we propose a method to evaluate the effectiveness of such strategies with the help of appropriate mathematical models. **Purpose:** Development of mathematical models for a multi-user MIMO system which would allow you to evaluate the average delay of a file transfer. **Results:** Mathematical models of the MIMO system are presented in the form of discrete/continuous time queuing systems for three user service strategies. We provide formulas for estimating the average file transfer delay in a closed form, which simplify numerical analysis as compared to the well-known methods using mathematical models in continuous time. This method allows you to find the most effective strategy with less computational complexity. Another advantage of the constructed models is the opportunity of further research on the effect of file length distributions on the file transfer delay. The conclusions we made about the effectiveness of each strategy are consistent with our earlier work, which is confirmed by a numerical example. **Practical relevance:** The developed mathematical models for a multi-user MIMO system can be used to set the operating parameters of a real wireless MIMO system.

Keywords — multiple antenna system, multi-user MIMO, mathematical model of a MIMO system, MIMO capacity, file transfer delay.

For citation: Gorbunova A. V., Medvedeva E. G., Gaidamaka Yu. V., Shorgin V. S., Samouylov K. E. Effective user service strategies in a multi-user MIMO system. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2019, no. 4, pp. 69–81 (In Russian). doi:10.31799/1684-8853-2019-4-69-81

References

- Solomitckii D., Petrov V., Nikopour H., Akdeniz M., Orhan O., Himayat N., Talwar S., Andreev S., Koucheryavy Y. Ray-based evaluation of dual-polarized MIMO in (Ultra)-dense millimeter-wave urban deployments. *IEEE Vehicular Technology Conference*, 2018, pp. 1–7. doi:10.1109/VTCSpring.2018.8417788
- Bakulin M. G., Varukina L. A., Kreyndelin V. B. *Tekhnologiya MIMO: principy i algoritmy* [MIMO Technology: Principles and Algorithms]. Moscow, Goryachaya liniya-Telekom Publ., 2014. 244 p. (In Russian).
- Borst S. User-level performance of channel-aware scheduling algorithms in wireless data networks. *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies*, 2003, vol. 1, pp. 321–331. doi:10.1109/INFCOM.2003.1208684
- Bellalta B., Oliver M. A space-time batch-service queueing model for multi-user MIMO communication systems. *Proceedings of the 12th International Symposium on Modeling Analysis and Simulation of Wireless and Mobile Systems*, Tenerife, 2009, pp. 357–364. doi:10.1145/1641804.1641866
- Bellalta B., Daza V., Miquel O. An approximate queueing model for multi-rate multi-user MIMO systems. *IEEE Communications Letters*, 2011, vol. 15, no. 4, pp. 392–394. doi:10.1109/LCOMM.2011.020311.102434
- Airy M., Shakkattai S., Heath R. W. Spatially greedy scheduling in multi-user MIMO wireless systems. *The Thirty-Seventh Asilomar Conference on Signals, Systems Computers*, 2003, vol. 1, pp. 982–986. doi:10.1109/ACSSC.2003.1292064
- Jankiraman M. *Space-Time Codes and MIMO Systems*. Artech House, 2004. 327 p.
- Telatar E. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 1999, vol. 10, pp. 585–595. https://doi.org/10.1002/ett.4460100604
- Pechinkin A. V., Razumchik R. V. *Sistemy massovogo obsluzhivaniya v diskretnom vremeni* [Discrete Time Queuing Systems]. Moscow, Fizmatlit Publ., 2018. 432 p. (In Russian).
- Alfa Attahiru S. *Applied Discrete-Time Queues*. Springer-Verlag, New York, 2016. 383 p.
- Gaidamaka Yu. V., Zaripova E. R., Orlov Y. N. Analyzing effects of SIP-messages batch arrival on queuing time for server processing. *Journal of Information Technologies and Computing Systems*, 2015, no. 3, pp. 74–80 (In Russian).
- Norikazu K., Takagi H., Yutaka T., Hong S.-J., Toshiharu H. Waiting time analysis for M[X]G1 queues with/without vacations under random order of service discipline. *Journal of the Operations Research Society of Japan*, 2000, vol. 43, no. 4, pp. 455–468. doi:10.15807/jorsj.43.455
- Takagi H. A note on the response time in M/G/1 queues with service in random order and bernoulli feedback. *Journal of the Operations Research Society of Japan*, 1996, vol. 39, no. 4, pp. 486–500. doi:10.15807/jorsj.39.486
- Gesbert D., Kountouris M., Heath R., Chae C.-B., Sälzer T. From single user to multiuser communications: shifting the MIMO paradigm. *IEEE Signal Process*, 2007, vol. 24, pp. 36–46.
- Yashkov S. F. Processor-sharing queues: Some progress in analysis. *Queueing Systems*, 1987, vol. 2, no. 1, pp. 1–17.
- Yashkov S. F. Mathematical problems in the theory of processor-sharing queues. *Itogi nauki i tekhniki. Seriya "Teoriya veroyatnosti. Matematicheskaya statistika. Teoreticheskaya kibernetika"*, 1990, vol. 29, pp. 3–82 (In Russian).
- Yashkov S. F., Yashkova A. S. Egalitarian processor sharing. *Information Processes*, 2006, vol. 6, no. 4, pp. 396–444 (In Russian).
- Borst S. C., Boxma O. J., Morrison J. A., Núñez Queija R. The equivalence between processor sharing and service in random order. *Operations Research Letters*, 2003, vol. 31, no. 4, pp. 254–262. doi:10.1016/S0167-6377(03)00006-3
- Yashkov S. F. On random order of service and processor sharing. *Information Processes*, 2006, vol. 6, no. 2, pp. 160–163.
- Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, IEEE std 802.11, 2016.
- Foschini G., Gans M. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 1998, no. 6, p. 311. https://doi.org/10.1023/A:1008889222784
- Daldoul Y., Meddour D., Ksentini A. IEEE 802.11n/ac data rates under power constraints. *IEEE International Conference on Communications*, Kansas City, USA, May 20–24, 2018. doi:10.1109/ICC.2018.8422697
- Gast M. S. *802.11ac: A Survival Guide: Wi-Fi at Gigabit and Beyond*. O'Reilly Media, 2013. 154 p.
- Sharon O., Alpert Y. A new aggregation based scheduling method for rapidly changing IEEE 802.11ac wireless channels. *Wireless Sensor Network*, 2016, vol. 8, no. 8, pp. 145–165. doi:10.4236/wsn.2016.88013
- Saaty T. L. *Elements of Queuing Theory with Applications*. Dover Publications, 1983. 423 p.
- Sopin E., Ageev K., Samouylov K. Approximate analysis of the limited resources queuing system with signals. *Proceedings — European Council for Modelling and Simulation ECMS*, 2019, vol. 33, no. 1, pp. 462–465.