

## ОБРАБОТКА ИНФОРМАЦИИ

## DATA PROCESSING

Научная статья  
УДК 681.324  
doi: 10.17223/19988605/74/8

## Влияние ошибок вытеснения на вероятность попадания в кэш

Сергей Эрикович Воробейчиков<sup>1</sup>, Александр Викторович Корсун<sup>2</sup>, Анна Сергеевна Морозова<sup>3</sup>,  
Сергей Петрович Сущенко<sup>4</sup>, Алексей Сергеевич Шкуркин<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup> Национальный исследовательский Томский государственный университет, Томск, Россия

<sup>4</sup> Томский филиал Федерального исследовательского центра  
информационных и вычислительных технологий, Томск, Россия

<sup>1</sup> sev@mail.tsu.ru

<sup>2</sup> aleksnfs15@gmail.com

<sup>3</sup> annamo12@yandex.ru

<sup>4</sup> ssp.inf.tsu@gmail.com

<sup>5</sup> shkurkin@mail.tsu.ru

**Аннотация.** Рассмотрен процесс заполнения ассоциативного кэша самыми востребованными адресуемыми объектами. Исследование основано на стратегии идеального вытеснения с ошибками, моделируемой многомерной марковской цепью. В аналитическом виде получены оценки вероятности попадания в кэш произвольной ассоциативности. Показана область применимости полученных оценок.

**Ключевые слова:** кэш; стратегии вытеснения; ошибки вытеснения; вероятность попадания в кэш; распределение вероятностей востребованности адресуемых объектов; цепь Маркова.

**Для цитирования:** Воробейчиков С.Э., Корсун А.В., Морозова А.С., Сущенко С.П., Шкуркин А.С. Влияние ошибок вытеснения на вероятность попадания в кэш // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2026. № 74. С. 82–94. doi: 10.17223/19988605/74/8

Original article  
doi: 10.17223/19988605/74/8

## The effect of preemption errors on the probability of cache hits

Sergey E. Vorobeychikov<sup>1</sup>, Alexander V. Korsun<sup>2</sup>, Anna S. Morozova<sup>3</sup>,  
Sergey P. Sushchenko<sup>4</sup>, Alexey S. Shkurkin<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup> National Research Tomsk State University, Tomsk, Russian Federation

<sup>4</sup> Tomsk Branch of the Federal Research Center for Information and Computing Technologies, Tomsk, Russian Federation

<sup>1</sup> sev@mail.tsu.ru

<sup>2</sup> aleksnfs15@gmail.com

<sup>3</sup> annamo12@yandex.ru

<sup>4</sup> ssp.inf.tsu@gmail.com

<sup>5</sup> shkurkin@mail.tsu.ru

**Abstract.** The process of filling an associative cache with the most frequently accessed objects is examined. The study is based on a perfect eviction strategy with errors, modeled by a multivariate Markov chain. Estimates of the cache hit probability for arbitrary associativity are obtained analytically. The applicability of these estimates is demonstrated.

**Keywords:** cache; eviction strategies; eviction errors; cache hit probability; addressable object demand probability distribution; Markov chain.

**For citation:** Vorobeychikov, S.E., Korsun, A.V., Morozova, A.S., Sushchenko, S.P., Shkurkin, A.S. (2026) The effect of preemption errors on the probability of cache hits. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 74. pp. 82–94. doi: 10.17223/19988605/74/8

## Введение

Важнейшей характеристикой кэш-памяти вычислительных систем является вероятность попадания, в значительной мере определяющая скорость доступа к адресуемым объектам. Основа применения кэша для балансировки скоростей обработки потока команд процессором и доступа к адресуемым объектам в оперативной памяти – принцип пространственно-временной локальности [1]. Основным фактором, определяющим вероятность попадания в ассоциативный кэш, является механизм вытеснения уже загруженных в кэш блоков памяти при возникновении конфликта адресов [2, 3]. Обычно вытеснению подлежат самые неиспользуемые процессором блоки памяти. Такое вытеснение реализует идеальная стратегия [2]. Для реализации идеальной стратегии вытеснения требуется знание вероятностей востребованности блоков оперативной памяти с загруженными в нем приложениями. Данные вероятности оцениваются на основе статистики обращения процессора к загруженным в кэш блокам памяти. При этом возможны ошибки оценивания, которые приводят к вытеснению «нужных» процессору блоков памяти. Как правило, ошибки обусловлены несовершенством реализуемых стратегий вытеснения, степенью рассеяния распределения востребованности адресуемых объектов процессором и недостаточной ассоциативностью кэша, приводящей к преждевременному вытеснению «нужных» объектов [4, 5]. Сравнительное исследование эффективности различных стратегий вытеснения и трудоемкости их реализации проводилось в работах [4–10]. Повышение эффективности работы кэша программными методами локализации размещения массивов и структур обрабатываемых данных рассмотрено в [4]. Авторы [5, 6] выполнили логический анализ различных стратегий вытеснения и ресурсоемкости алгоритмизации стратегий (затраты на память и время выполнения), а также экспериментальные измерения их эффективности. В работе [7] исследованы методы нейтрализации конфликта адресов ассоциативным кэшированием и стратегиями замещения в кэше загруженных объектов. В [8] проведено имитационное моделирование интенсивности попадания в кэш для различных приложений с широким спектром показателей локальности. Авторы [9] предложили способ натуральных измерений быстродействия элементов иерархической памяти и политик вытеснения адресуемых объектов. В [10] проанализирована зависимость быстродействия иерархической памяти от ассоциативности и коэффициента неблокируемости транзакций доступа к адресуемым объектам при промахах в кэше. Однако анализ зависимости эффективности стратегий и ресурсоемкости их алгоритмизации от ассоциативности кэша, глубины его неблокируемости и объема отображаемой на кэш оперативной памяти выполнялся только численно. Исследование идеальной стратегии вытеснения наименее востребованного блока с ошибками проводилось в [2], но только для кэша с коэффициентом ассоциативности, равным двум, и существенно ограниченным числом блоков памяти, отображаемых на группу кэша. В данной работе выполнен анализ влияния таких ошибок на вероятность попадания в кэш произвольной ассоциативности и емкости отображаемой памяти.

### 1. Модель кэша для вытесняющей стратегии с ошибками

Известные исследования эффективности ассоциативного кэширования [2] опираются на математическую модель кэша с идеальной стратегией вытеснения самого неиспользуемого процессором блока памяти (с наименьшей вероятностью востребованности). При этом предполагается, что номера блоков памяти  $m$ , отображаемых на заданную группу кэша ассоциативности  $A > 1$ , упорядочены по убыванию вероятности востребованности процессором  $f_m$ ,  $m = \overline{0, M-1}$ . Модель основана на многомерной (размерности  $A$ ) марковской цепи, описывающей процесс замещения блоков памяти, загруженных в кэш, при конфликте адресов. Состояниям цепи соответствует набор номеров блоков памяти, содержащихся в группе кэша [2]. Переходы между состояниями задаются вероятностями востребованных процессором

отсутствующих в заполненной группе кэша блоков памяти. В стационарных условиях марковская цепь распадается на  $A!$  неприводимых подцепей [2]. При идеальном вытеснении множество стационарных состояний цепи Маркова имеет вид, приведенный на рис. 1. Если допустить, что ошибки вытеснения возможны в любом состоянии марковской цепи, то получаем неразрешимую в аналитическом виде систему уравнений равновесия.

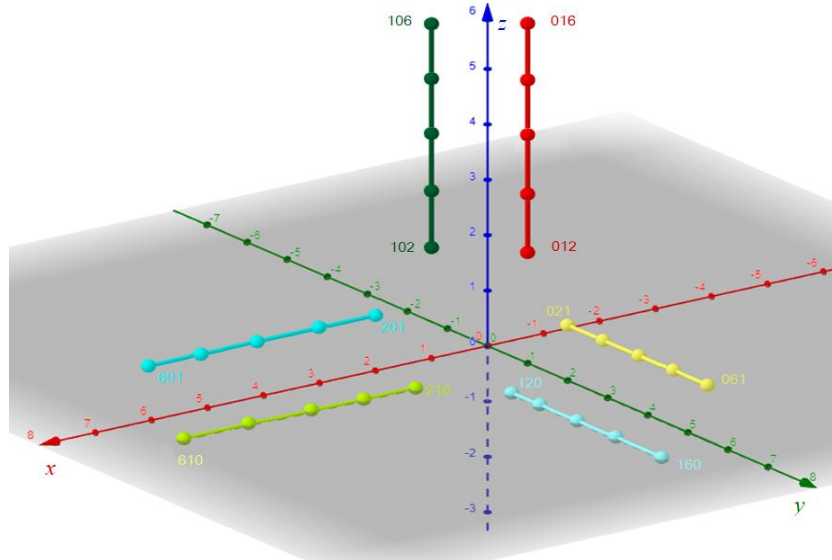


Рис. 1. Цепь Маркова, описывающая динамику идеального вытеснения блоков памяти при  $A = 3, M = 6$   
 Fig. 1. Markov chain describing the dynamics of ideal memory block eviction under  $A = 3, M = 6$

Для получения аналитического решения допустим, что ошибки, обусловленные недостаточной статистикой обращения к адресуемым объектам, имеют место только в состоянии, в котором группа кэша содержит  $A$  самых востребованных блоков памяти (с номерами  $m = \overline{0, A-1}$ ). Полагаем, что с вероятностью ошибки  $e_a, a = \overline{0, A-2}, \sum_{a=0}^{A-2} e_a \leq 1$ , в данном состоянии из кэша вытесняются блоки памяти с номером  $a$ . Очевидно, что неидеальное вытеснение блоков в этом состоянии приведет к самому значимому снижению вероятности попадания в кэш в силу вытеснения объективно наиболее востребованных блоков памяти. Полагаем, что в остальных состояниях цепи Маркова механизм вытеснения работает идеально.

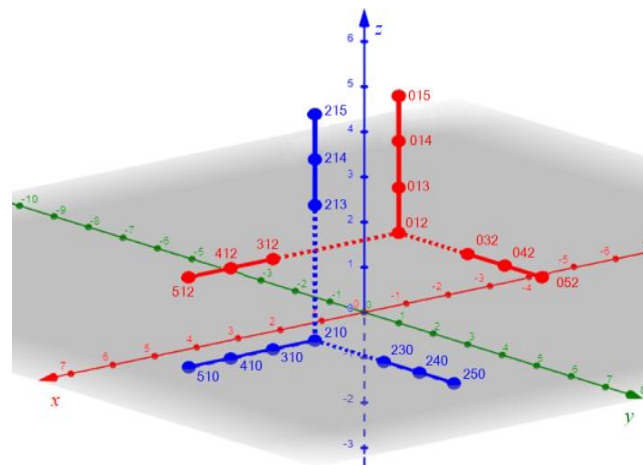


Рис. 2. Две из шести подцепей цепи Маркова, описывающей динамику неидеального вытеснения блоков памяти при  $A = 3, M = 5$   
 Fig. 2. Two of the six subchains of the Markov chain describing the dynamics of non-ideal eviction of memory blocks under  $A = 3, M = 5$

Считаем, что известны объективные вероятности востребованности процессором блоков памяти вычислителя и в зависимости от механизма, реализующего вытеснение блоков при конфликте адресов, заданы вероятности ошибки, возможные при оценке вероятности востребованности адресуемых объектов  $f_m$ . Полагаем, что данные ошибки приводят к некорректному вытеснению адресуемых объектов. Цепь Маркова при этом распадается на факториальное число ( $A!$ ) структурно одинаковых неприводимых подцепей (рис. 2), в пределах которых переходные вероятности примут вид:

$$\pi_{IJ} = \begin{cases} f_m \left( 1 - \sum_{l=0}^{A-2} e_l \right), I = \{i_0, i_1, \dots, i_{A-1}\}, i_a = \overline{0, M-1}, a = \overline{0, A-1}, J = \{j_0, j_1, \dots, j_{A-1}\}, j_k = m, k = \arg \left\{ \frac{\max m_a}{a = \overline{0, A-1}} \right\}, \\ j_a = i_a, a = \overline{0, A-1}, a \neq k; \\ f_m e_l, I = \{i_0, i_1, \dots, i_{A-1}\}, i_a = \overline{0, A-1}, a = \overline{0, A-1}, J = \{j_0, j_1, \dots, j_{A-1}\}, j_k = m, k = \arg \left\{ \frac{m_a = l}{a = \overline{0, A-1}} \right\}, \\ j_a = i_a, a = \overline{0, A-1}, a \neq k; \\ f_m, I = \{i_0, i_1, \dots, i_{A-1}\}, i_a = \overline{0, M-1}, a = \overline{0, A-1}, J = \{j_0, j_1, \dots, j_{A-1}\}, j_k = m, k = \arg \left\{ \frac{\max m_a}{a = \overline{0, A-1}} \right\}, \\ j_a = i_a, a = \overline{0, A-1}, a \neq k. \end{cases}$$

Предположим, что  $P(m_0, m_1, \dots, m_{A-1})$  – вероятность состояния, в котором кэш с коэффициентом ассоциативности  $A$  содержит блоки памяти с номерами  $m_0, m_1, \dots, m_{A-1}$ ,  $m_a = \overline{0, M-1}$ ,  $a = \overline{0, A-1}$ ,  $m_i \neq m_j$ ,  $i, j = \overline{0, A-1}$ . Тогда оценка вероятности обнаружения  $m$ -го блока памяти в кэше определится соотношением [2]

$$\hat{\Pi}_m(A, \vec{F}) = \sum_{m_0=0}^{M-1} \sum_{m_1=m_0+1}^{M-1} \dots \sum_{m_{A-2}=m_{A-3}+1}^{M-1} AP(m, m_0, \dots, m_{A-2}), m = \overline{0, M-1}. \quad (1)$$

Здесь  $\vec{F} = \{f_0; f_2; \dots; f_{M-1}\}$  – вектор упорядоченных по убыванию вероятностей востребованности процессором блоков памяти, отображаемых на ассоциативный кэш. Оценка вероятности попадания адресуемого объекта в кэш ассоциативности  $A$  задается выражением [2]

$$\hat{\Pi}(A, \vec{F}) = \sum_{m=0}^{M-1} \hat{\Pi}_m(A, \vec{F}) f_m. \quad (2)$$

## 2. Вероятности состояний цепи Маркова

Для ассоциативности  $A = 2$  цепь Маркова распадается на две неприводимые подцепи. Одна подцепь содержит набор состояний  $(0, m)$ ,  $m = \overline{1, M-1}$ , и  $(m, 1)$ ,  $m = \overline{2, M-1}$ , а другая – набор  $(m, 0)$ ,  $m = \overline{1, M-1}$ , и  $(1, m)$ ,  $m = \overline{2, M-1}$ . Уравнения равновесия для первой подцепи принимают вид:

$$P(0, 1) \sum_{m=2}^{M-1} f_m = f_0 \sum_{m=2}^{M-1} P(m, 1) + f_1 \sum_{m=2}^{M-1} P(0, m); \quad P(0, n) \sum_{m=1, m \neq n}^{M-1} f_m = f_n \left( P(0, 1)(1 - e_0) + \sum_{m=2, m \neq n}^{M-1} P(0, m) \right); \\ P(n, 1) \sum_{m=0, m \neq n}^{M-1} f_m = f_n \left( P(0, 1)e_0 + \sum_{m=2, m \neq n}^{M-1} P(m, 1) \right), \quad n = \overline{2, M-1}.$$

Вторая подцепь в стационарных условиях описывается симметричными уравнениями

$$P(1, 0) \sum_{m=2}^{M-1} f_m = f_0 \sum_{m=2}^{M-1} P(1, m) + f_1 \sum_{m=2}^{M-1} P(m, 0); \quad P(n, 0) \sum_{m=1, m \neq n}^{M-1} f_m = f_n \left( P(1, 0)(1 - e_0) + \sum_{m=2, m \neq n}^{M-1} P(m, 0) \right); \\ P(1, n) \sum_{m=0, m \neq n}^{M-1} f_m = f_n \left( P(1, 0)e_0 + \sum_{m=2, m \neq n}^{M-1} P(1, m) \right), \quad n = \overline{2, M-1}.$$

С учетом условия нормировки отсюда получаем вероятности состояний цепи Маркова

$$P(0,1) = P(1,0) = \frac{f_0 f_1}{2U_2}; \quad P(0,m) = P(m,0) = \frac{f_0 f_m (1 - e_0)}{2U_2}; \quad P(1,m) = P(m,1) = \frac{f_0 f_m e_0}{2U_2}, \quad m = \overline{2, M-1}; \quad (3)$$

$$U_2 = f_0(1 - f_0) - e_0(f_0 - f_1)(1 - f_0 - f_1).$$

При  $A = 3$  цепь Маркова в стационарных условиях распадается на шесть неприводимых подцепей, две из которых проиллюстрированы на рис. 2. Уравнения равновесия для одной из подцепей имеют вид:

$$P(0,1,2) \sum_{m=3}^{M-1} f_m = f_0 \sum_{m=3}^{M-1} P(m,1,2) + f_1 \sum_{m=3}^{M-1} P(0,m,2) + f_2 \sum_{m=3}^{M-1} P(0,1,m);$$

$$P(0,1,n) \sum_{m=2, m \neq n}^{M-1} f_m = f_0 \sum_{m=n+1}^{M-1} P(m,1,n) + f_1 \sum_{m=n+1}^{M-1} P(0,m,n) + f_n \left[ (1 - e_0 - e_1) P(0,1,2) + \sum_{m=3, m \neq n}^{M-1} P(0,1,m) \right];$$

$$P(0,n,2) \left( f_1 + \sum_{m=3, m \neq n}^{M-1} f_m \right) = f_0 \sum_{m=n+1}^{M-1} P(m,n,2) + f_2 \sum_{m=n+1}^{M-1} P(0,n,m) + f_n \left[ e_1 P(0,1,2) + \sum_{m=3, m \neq n}^{M-1} P(0,m,2) \right];$$

$$P(n,1,2) \left( f_0 + \sum_{m=3, m \neq n}^{M-1} f_m \right) = f_1 \sum_{m=n+1}^{M-1} P(n,m,2) + f_2 \sum_{m=n+1}^{M-1} P(n,1,m) + f_n \left[ e_0 P(0,1,2) + \sum_{m=3, m \neq n}^{M-1} P(m,1,2) \right], \quad n = \overline{3, M-1}.$$

Отсюда с учетом того, что возвратными [11] являются только состояния  $(0, 1, m)$ ,  $m = \overline{2, M-1}$ ;  $(0, m, 2)$ ,  $(m, 1, 2)$ ,  $m = \overline{3, M-1}$ , получаем вероятности состояний цепи Маркова

$$\begin{cases} P(0,1,2) = P(0,2,1) = P(1,0,2) = P(1,2,0) = P(2,0,1) = P(2,1,0) = \frac{f_0 f_1 f_2}{6U_3}; \\ P(0,1,m) = P(0,m,1) = P(1,0,m) = P(1,m,0) = P(m,0,1) = P(m,1,0) = \frac{f_0 f_1 f_m (1 - e_0 - e_1)}{6U_3}, \quad m = \overline{3, M-1}; \\ P(m,1,2) = P(1,2,m) = P(1,m,2) = P(2,1,m) = P(2,m,1) = P(m,2,1) = \frac{f_1 f_2 f_m e_0}{6U_3}, \quad m = \overline{3, M-1}; \end{cases} \quad (4)$$

$$U_3 = f_0 f_1 (1 - f_0 - f_1) - (1 - f_0 - f_1 - f_2) (e_0 f_1 (f_0 - f_2) + e_1 f_0 (f_1 - f_2)).$$

При  $e_a = 0$ ,  $a = \overline{0,1}$ , вероятности состояний цепи Маркова принимают известный вид [2]:

$$P(0,1,m) = P(0,m,1) = P(1,0,m) = P(1,m,0) = P(m,0,1) = P(m,1,0) = \frac{f_m}{6(1 - f_0 - f_1)}, \quad m = \overline{2, M-1}.$$

Стационарный режим для ассоциативности  $A = 4$  описывается следующей системой уравнений равновесия одной из 24 неприводимых подцепей Маркова:

$$P(0,1,2,3) \sum_{m=4}^{M-1} f_m = f_0 \sum_{m=4}^{M-1} P(m,1,2,3) + f_1 \sum_{m=4}^{M-1} P(0,m,2,3) + f_2 \sum_{m=4}^{M-1} P(0,1,m,3) + f_3 \sum_{m=4}^{M-1} P(0,1,2,m);$$

$$P(0,1,2,n) \sum_{m=3, m \neq n}^{M-1} f_m = f_0 \sum_{m=n+1}^{M-1} P(m,1,2,n) + f_1 \sum_{m=n+1}^{M-1} P(0,m,2,n) + f_2 \sum_{m=n+1}^{M-1} P(0,1,m,n) + f_n \left[ (1 - e_0 - e_1 - e_2) P(0,1,2,3) + \sum_{m=4, m \neq n}^{M-1} P(0,1,2,m) \right], \quad n = \overline{4, M-1};$$

$$P(0,1,n,3) \left( f_2 + \sum_{m=4, m \neq n}^{M-1} f_m \right) = f_0 \sum_{m=n+1}^{M-1} P(m,1,n,3) + f_1 \sum_{m=n+1}^{M-1} P(0,m,n,3) + f_3 \sum_{m=n+1}^{M-1} P(0,1,n,m) + f_n \left[ e_2 P(0,1,2,3) + \sum_{m=4, m \neq n}^{M-1} P(0,1,m,3) \right], \quad n = \overline{4, M-1};$$

$$P(0,n,2,3) \left( f_1 + \sum_{m=4, m \neq n}^{M-1} f_m \right) = f_0 \sum_{m=n+1}^{M-1} P(m,n,2,3) + f_2 \sum_{m=n+1}^{M-1} P(0,n,m,3) + f_3 \sum_{m=n+1}^{M-1} P(0,n,2,m) +$$

$$\begin{aligned}
 & + f_n \left[ e_1 P(0,1,2,3) + \sum_{m=4, m \neq n}^{M-1} P(0,m,2,3) \right], n = \overline{4, M-1}; \\
 P(n,1,2,3) \left( f_0 + \sum_{m=4, m \neq n}^{M-1} f_m \right) & = f_1 \sum_{m=n+1}^{M-1} P(n,m,2,3) + f_2 \sum_{m=n+1}^{M-1} P(n,1,m,3) + f_3 \sum_{m=n+1}^{M-1} P(n,1,2,m) + \\
 & + f_n \left[ e_0 P(0,1,2,3) + \sum_{m=4, m \neq n}^{M-1} P(m,1,2,3) \right], n = \overline{4, M-1}.
 \end{aligned}$$

С учетом условия нормировки решение данной системы уравнений примет вид:

$$\left\{ \begin{aligned}
 P(m_0, m_1, m_2, m_3) & = \frac{1}{24U_4} \prod_{a=0}^3 f_a, m_a = \overline{0,3}, a = \overline{0,3}, m_i \neq m_j, i, j = \overline{0,3}, i \neq j; \\
 P(m_0, m_1, n_a, m_3) & = \frac{1}{24U_4} f_{n_a} \prod_{i=0}^2 f_i \left( 1 - \sum_{j=0}^{A-2} e_j \right), n_a = \overline{4, M-1}, a = \overline{0,3}, m_i = \overline{0,2}, m_i \neq m_j, i, j = \overline{0,3}, i \neq j; \\
 P(m, l, n_a, k) & = \frac{1}{24U_4} e_b f_{n_a} \prod_{i=0}^2 f_i, n_a = \overline{4, M-1}, a = \overline{0,3}, m, l, k = \overline{0,3}, m \neq l \neq k;
 \end{aligned} \right. \quad (5)$$

$$U_4 = \left( 1 - \sum_{i=0}^2 f_i \right) \prod_{i=0}^2 f_i - \left( 1 - \sum_{i=0}^3 f_i \right) \left[ \sum_{a=0}^{A-2} e_a (f_a - f_3) \prod_{i=0, i \neq a}^2 f_i \right].$$

Здесь  $b$  – разность двух множеств,  $b = B \setminus B, B = \{0,1,2,3\}, B = \{m,l,k\}$ . При нулевых ошибках ( $e_a = 0, a = \overline{0,2}$ ) вероятности состояний преобразуются к виду[2]:

$$P(m_0, \dots, m_1, \dots, m_2) = \frac{f_m}{24(1 - f_0 - f_1 - f_2)}, m = \overline{3, M-1}, m_a = \overline{0,2}, a = \overline{0,2}.$$

В случае произвольной ассоциативности для одной из неприводимых подцепей Маркова условия равновесия описываются уравнениями

$$\begin{aligned}
 P(0,1, \dots, A-1) \sum_{m=A}^{M-1} f_m & = \sum_{a=0}^{A-1} f_a \sum_{k_a=A}^{M-1} P(0,1, \dots, k_a, \dots, A-1); \\
 P(0,1, \dots, A-2, n) \sum_{m=A-1, m \neq n}^{M-1} f_m & = \sum_{a=0}^{A-2} f_a \sum_{m_a=n+1}^{M-1} P(0,1, \dots, m_a, \dots, A-2, n) + \\
 & + f_n \left[ \left( 1 - \sum_{a=0}^{A-2} e_a \right) P(0,1, \dots, A-1) + \sum_{m=A, m \neq n}^{M-1} P(0,1, \dots, A-2, m) \right], n = \overline{A, M-1}; \\
 P(0,1, \dots, n_a, \dots, A-1) \left[ f_a + \sum_{m=A, m \neq n_a}^{M-1} f_m \right] & = \sum_{i=0, i \neq a}^{A-1} f_i \sum_{m_i=n_a+1}^{M-1} P(0,1, \dots, m_i, \dots, n_a, \dots, A-1) + \\
 & + f_n \left[ e_a P(0,1, \dots, A-1) + \sum_{m_i=A, m_i \neq n_a}^{A-1} P(0,1, \dots, k_i, \dots, A-2, n) \right], a = \overline{0, A-2}, n = \overline{A, M-1}.
 \end{aligned}$$

Решение с учетом условия нормировки принимает вид:

$$\left\{ \begin{aligned}
 P(0,1, \dots, A-1) & = \prod_{a=0}^{A-1} f_a / A! U_A; \\
 P(0,1, \dots, A-2, m) & = \left( 1 - \sum_{a=0}^{A-2} e_a \right) f_m \prod_{i=0}^{A-2} f_i / A! U_A, m = \overline{A, M-1}; \\
 P(0,1, \dots, a-1, m, a+1, \dots, A-2) & = e_a f_m \prod_{i=0}^{A-2} f_i / A! U_A, m = \overline{A, M-1};
 \end{aligned} \right. \quad (6)$$

$$U_A = \left( 1 - \sum_{a=0}^{A-2} f_a \right) \prod_{i=0}^{A-2} f_i - \left( 1 - \sum_{a=0}^{A-1} f_a \right) \left[ \sum_{a=0}^{A-2} e_a (f_a - f_{A-1}) \prod_{i=0, i \neq a}^{A-2} f_i \right].$$

### 3. Вероятность попадания адресуемых объектов в кэш

Найдем оценки вероятности попадания в кэш востребованных процессором блоков памяти. Согласно (1) при  $A = 2$  для оценок вероятностей попадания в кэш отдельных блоков памяти из (3) получаем:

$$\hat{\Pi}_0(2, \vec{F}) = \frac{f_0 [1 - f_0 - e_0 (1 - f_0 - f_1)]}{U_2}, \quad \hat{\Pi}_1(2, \vec{F}) = \frac{f_1 [f_0 + e_0 (1 - f_0 - f_1)]}{U_2},$$

$$\hat{\Pi}_m(2, \vec{F}) = \frac{f_m [f_0 - e_0 (f_0 - f_1)]}{U_2}, \quad m = \overline{2, M-1}.$$

Оценка вероятности попадания в кэш (2) составит

$$\hat{\Pi}(2, \vec{F}) = \frac{1}{U_2} \left\{ f_0 (f_0 (1 - f_0) + f_1^2) - e_0 (1 - f_0 - f_1) (f_0 - f_1) (f_0 + f_1) + [(f_0 - e_0 (f_0 - f_1))] \sum_{m=2}^{M-1} f_m^2 \right\}.$$

Нетрудно убедиться, что при отсутствии ошибок вытеснения ( $e_0 = 0$ ) вероятность попадания в кэш отдельных блоков памяти принимает известный вид [2]:  $\Pi_0(2, \vec{F}) = 1$ ,  $\Pi_m(2, \vec{F}) = \frac{f_m}{1 - f_0}$ ,  $m = \overline{1, M-1}$ .

На рис. 3 приведена зависимость оценок вероятностей обнаружения отдельных блоков памяти в кэше от вероятности корректного вытеснения блоков памяти при  $A = 2$  и  $M = 4$  для распределения востребованности адресуемых объектов вида:  $\vec{F} = \{0, 4; 0, 3; 0, 2; 0, 1\}$ .

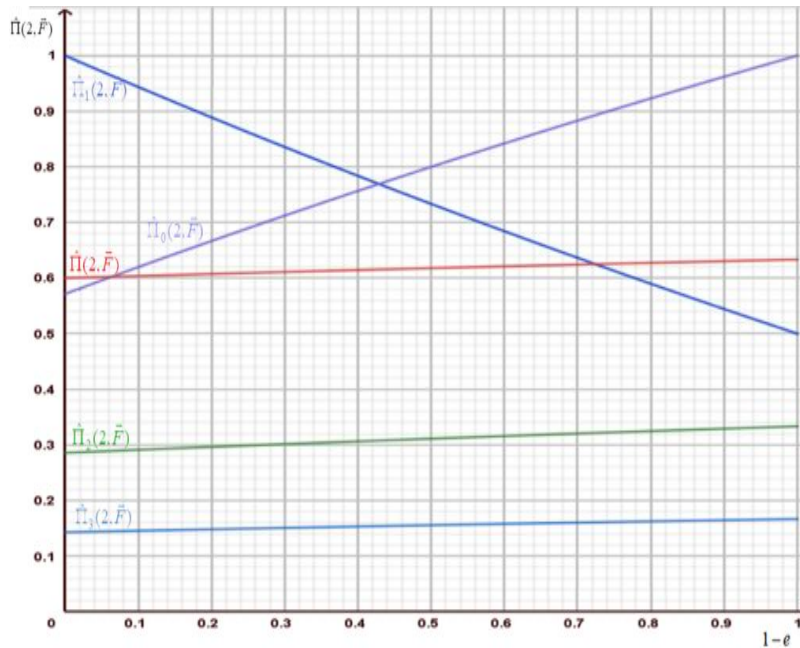


Рис. 3. Зависимость оценок вероятностей обнаружения отдельных блоков памяти в кэше  $\hat{\Pi}_m(2), m = \overline{0, 3}$  и оценки попадания в кэш от вероятности правильного решения при  $A = 2, M = 4$

Fig. 3. Dependence of estimates of the probabilities of detecting individual memory blocks in the cache  $\hat{\Pi}_m(2), m = \overline{0, 3}$  and the estimate of hitting the cache on the probability of a correct decision when  $A = 2, M = 4$

Для кэша с коэффициентом ассоциативности  $A = 3$  вероятности обнаружения в нем блоков памяти и вероятности попадания в кэш из (1), (2) и (4) определяются соотношениями

$$\hat{\Pi}_0(3, \vec{F}) = f_0 [f_1 (1 - f_0 - f_1) - (1 - f_0 - f_1 - f_2) (e_0 f_1 + e_1 (f_1 - f_2))] / U_3,$$

$$\hat{\Pi}_1(3, \vec{F}) = f_1 [f_0 (1 - f_0 - f_1) - (1 - f_0 - f_1 - f_2) (e_0 (f_0 - f_2) + e_1 f_0)] / U_3,$$

$$\hat{\Pi}_2(3, \vec{F}) = f_2 [f_0 f_1 + (1 - f_0 - f_1 - f_2) (e_0 f_1 + e_1 f_0)] / U_3,$$

$$\hat{\Pi}_m(3, \bar{F}) = f_m [f_0 f_1 - e_0 f_1 (f_0 - f_2) - e_1 f_0 (f_1 - f_2)] / U_3, \quad m = \overline{3, M-1}.$$

$$\hat{\Pi}(3, \bar{F}) = \frac{1}{U_3} \left\{ f_0 f_1 [(1 - f_0 - f_1)(f_0 + f_1) + f_2^2] - (1 - f_0 - f_1 - f_2) \times \right.$$

$$\times \left\{ e_0 f_1 [(f_0 - f_2)(f_0 + f_1 + f_2)] + e_1 f_0 [(f_1 - f_2)(f_0 + f_1 + f_2)] \right\} +$$

$$\left. + [f_0 f_1 - e_0 f_1 (f_0 - f_2) - e_1 f_0 (f_1 - f_2)] \sum_{m=3}^{M-1} f_m^2 \right\}.$$

Вид зависимости оценок обнаружения отдельных блоков памяти в кэше ассоциативности, равной трем, для распределения востребованности  $\bar{F} = \{0, 4; 0, 3; 0, 15; 0, 1; 0, 05\}$  иллюстрируется на рис. 4.

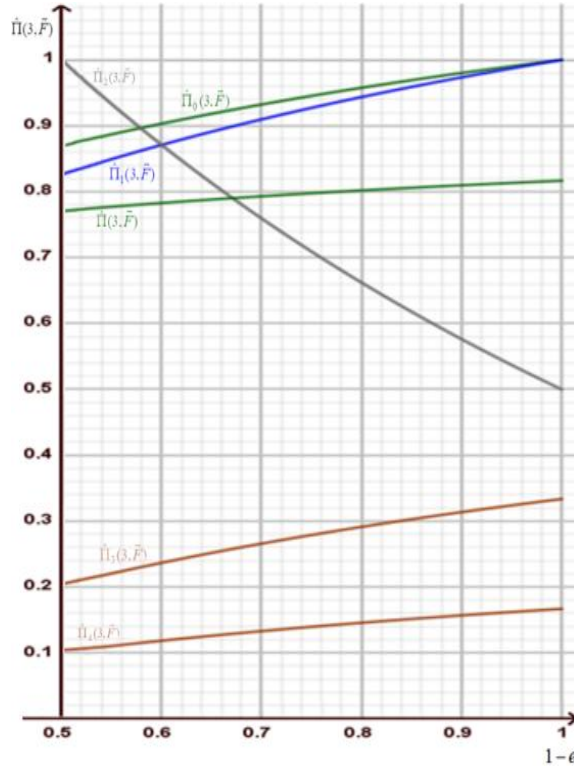


Рис. 4. Зависимость оценок вероятностей обнаружения отдельных блоков памяти в кэше  $\hat{\Pi}_m(3), m = \overline{0, 3}$  и оценки попадания в кэш от вероятности правильного решения при  $A = 3, M = 5$  и равновероятных ошибках  
 Fig. 4. Dependence of estimates of the probabilities of detecting individual memory blocks in the cache  $\hat{\Pi}_m(3), m = \overline{0, 3}$  and the cache hit rate from the probability of a correct solution when  $A = 3, M = 5$  and equally probable errors

Для четырехходового кэша ( $A = 4$ ) из (1) и (2) согласно (5) получаем оценки обнаружения блоков памяти в кэше и вероятности попадания в кэш

$$\hat{\Pi}_0(4, \bar{F}) = f_0 [f_1 f_2 (1 - f_0 - f_1 - f_2) - (1 - f_0 - f_1 - f_2 - f_3)(e_0 f_1 f_2 + e_1 f_2 (f_1 - f_3) + e_2 f_1 (f_2 - f_3))] / U_4,$$

$$\hat{\Pi}_1(4, \bar{F}) = f_1 [f_0 f_2 (1 - f_0 - f_1 - f_2) - (1 - f_0 - f_1 - f_2 - f_3)(e_0 f_2 (f_0 - f_3) + e_1 f_0 f_2 + e_2 f_0 (f_2 - f_3))] / U_4,$$

$$\hat{\Pi}_2(4, \bar{F}) = f_2 [f_0 f_1 (1 - f_0 - f_1 - f_2) - (1 - f_0 - f_1 - f_2 - f_3)(e_0 f_1 (f_0 - f_3) + e_1 f_0 (f_1 - f_3) + e_2 f_0 f_1)] / U_4,$$

$$\hat{\Pi}_3(4, \bar{F}) = f_3 [f_0 f_1 f_2 + (1 - f_0 - f_1 - f_2 - f_3)(e_0 f_1 f_2 + e_1 f_0 f_2 + e_2 f_0 f_1)] / U_4,$$

$$\hat{\Pi}_m(4, \bar{F}) = f_m [f_0 f_1 f_2 - e_0 f_1 f_2 (f_0 - f_3) - e_1 f_0 f_2 (f_1 - f_3) - e_2 f_0 f_1 (f_2 - f_3)] / U_4, \quad m = \overline{4, M-1}.$$

$$\hat{\Pi}(4, \bar{F}) = \frac{1}{U_4} \left\{ f_0 f_1 f_2 [(1 - f_0 - f_1 - f_2)(f_0 + f_1 + f_2) + f_3^2] - (1 - f_0 - f_1 - f_2 - f_3) \times \right.$$

$$\times \{e_0 f_1 f_2 (f_0 - f_3)(f_0 + f_1 + f_2 + f_3) + e_1 f_0 f_2 (f_1 - f_3)(f_0 + f_1 + f_2 + f_3) + e_2 f_0 f_1 (f_2 - f_3)(f_0 + f_1 + f_2 + f_3)\} + [f_0 f_1 f_2 - e_0 f_1 f_2 (f_0 - f_3) - e_1 f_0 f_2 (f_1 - f_3) - e_2 f_0 f_1 (f_2 - f_3)] \sum_{m=4}^{M-1} f_m^2 \}.$$

Зависимость полученных оценок при  $A = 4$  от уровня равновероятных ошибок приведена на рис. 5 для вектора вероятностей востребованности  $\vec{F} = \{0,35; 0,25; 0,15; 0,1; 0,09; 0,06\}$ .

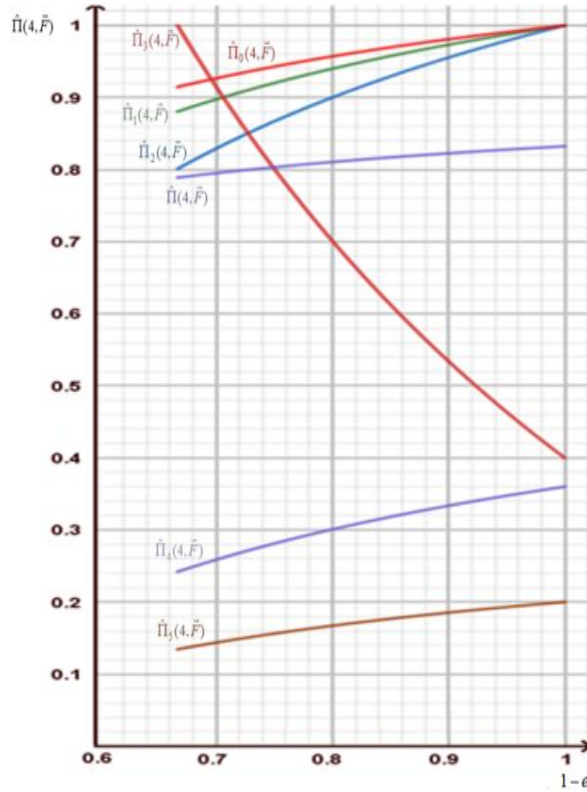


Рис. 5. Зависимость оценок вероятностей обнаружения отдельных блоков памяти в кэше  $\hat{P}_m(4), m = \overline{0,5}$  и оценки попадания в кэш от вероятности правильного решения при  $A = 4, M = 6$  и равновероятных ошибках  
Fig. 5. Dependence of estimates of the probabilities of detecting individual memory blocks in the cache  $\hat{P}_m(4), m = \overline{0,5}$  and the cache hit rate from the probability of a correct solution when  $A = 4, M = 6$  and equally probable errors

Сравнительные зависимости оценок вероятности попадания в кэш для различных ассоциативностей  $A = \overline{2,4}$ , вектора востребованности  $\vec{F} = \{0,3; 0,25; 0,15; 0,1; 0,05; 0,05; 0,05; 0,05\}$  и равновероятных ошибок вытеснения блоков памяти приведены на рис. 6. Нетрудно видеть, что для  $\sum_{a=0}^{A-2} e_a \leq 0,3$  оценка вероятности попадания в кэш  $\hat{P}(A, \vec{F})$  незначительно отклоняется от идеального варианта вытеснения.

Для произвольной ассоциативности из (1), (2) и (6) получаем оценки вероятности обнаружения заданного блока памяти в кэше и вероятности попадания в кэш

$$\hat{P}_a(A, \vec{F}) = \frac{f_a}{U_A} \left[ \left( 1 - \sum_{i=0}^{A-2} f_i \right) \prod_{i=0, i \neq a}^{A-2} f_i - \left( 1 - \sum_{i=0}^{A-1} f_i \right) \left( e_a \prod_{i=0, i \neq a}^{A-2} f_i + \sum_{n=0, n \neq a}^{A-2} e_n (f_n - f_{A-1}) \prod_{i=0, i \neq a, i \neq n}^{A-2} f_i \right) \right], a = \overline{0, A-2},$$

$$\hat{P}_{A-1}(A, \vec{F}) = \frac{f_{A-1}}{U_A} \left[ \prod_{i=0}^{A-2} f_i + \left( 1 - \sum_{a=0}^{A-1} f_a \right) \sum_{a=0}^{A-2} e_a \prod_{i=0, i \neq a}^{A-2} f_i \right], \hat{P}_m(A, \vec{F}) = \frac{f_m}{U_A} \left[ \prod_{i=0}^{A-2} f_i - \sum_{a=0}^{A-2} e_a (f_a - f_{A-1}) \prod_{i=0, i \neq a}^{A-2} f_i \right], m = \overline{A, M-1},$$

$$\hat{P}(A, \vec{F}) = \frac{1}{U_A} \left\{ \prod_{i=0}^{A-2} f_i \left[ f_{A-1}^2 + \left( 1 - \sum_{i=0}^{A-2} f_i \right) \sum_{i=0}^{A-2} f_i \right] - \left( 1 - \sum_{i=0}^{A-1} f_i \right) \left\{ \sum_{a=0}^{A-2} e_a \prod_{i=0, i \neq a}^{A-2} f_i \left[ (f_a - f_{A-1}) \sum_{i=0}^{A-1} f_i \right] \right\} \right\} +$$

$$+ \left[ \prod_{i=0}^{A-2} f_i - \sum_{a=0}^{A-2} e_a (f_a - f_{A-1}) \prod_{i=0, i \neq a}^{A-2} f_i \right] \sum_{m=4}^{M-1} f_m^2 \Bigg\}. \quad (7)$$

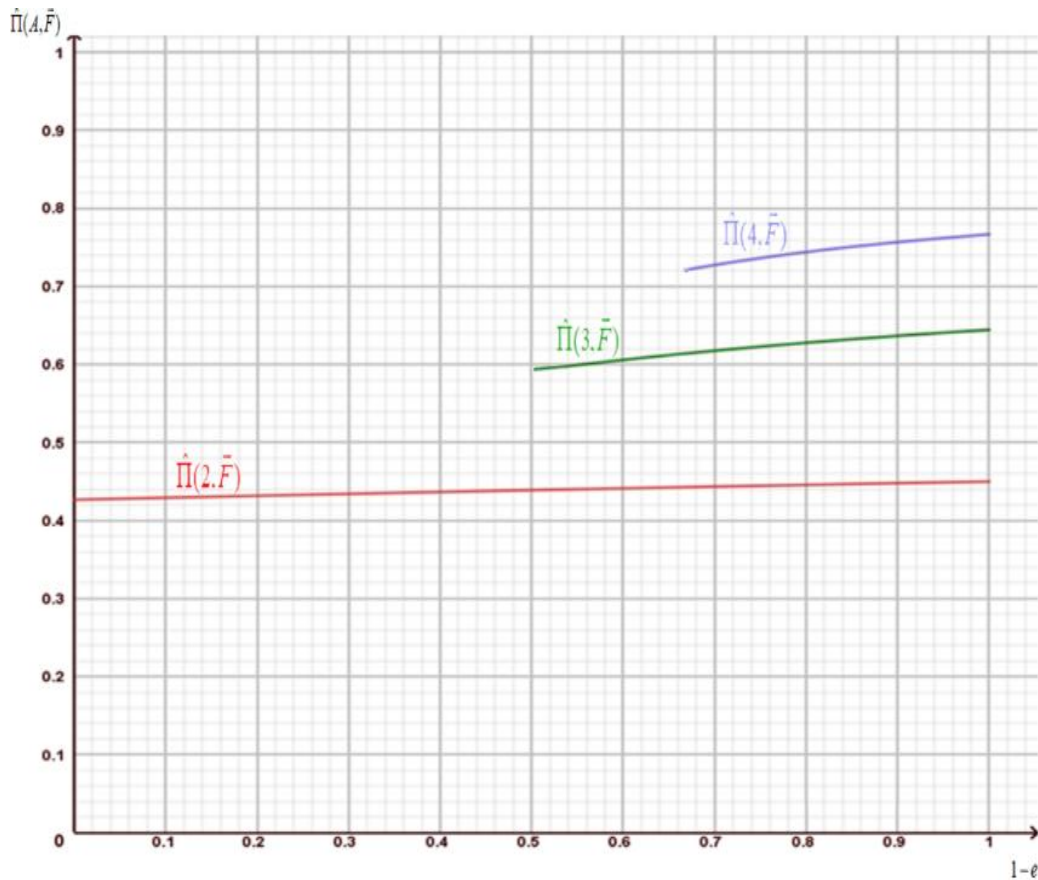


Рис. 6. Сравнительные кривые зависимостей оценок вероятностей попадания в кэш  $\hat{P}(A)$  при  $M = 8$  и равновероятных ошибках

Fig. 6. Comparative curves of dependencies of cache hit probability estimates  $\hat{P}(A)$  at  $M = 8$  and equally probable errors

#### 4. Чувствительность оценки вероятности попадания в кэш к виду распределения вероятностей востребованности адресуемых объектов

В реальных вычислительных системах ошибки при различных стратегиях вытеснения адресуемых объектов могут происходить в произвольных состояниях ассоциативного кэша. Проанализируем влияние реальных процессов вытеснения на адекватность полученных оценок вероятности попадания в кэш. Рассмотрим кэш ассоциативности  $A = 2$ , на который отображается пять блоков памяти ( $M = 5$ ). Полагаем, что ошибки вытеснения могут быть в любом состоянии моделирующей цепи Маркова с одинаковой вероятностью. Результирующая система уравнений равновесия допускает только численное решение. В таблице приведены сравнительные значения вероятности попадания в кэш  $P(2, \bar{F}_k), k = \overline{1,3}$  при ошибках в произвольных состояниях цепи Маркова и их оценки  $\hat{P}(2, \bar{F}_k), k = \overline{1,3}$  для различных уровней ошибок и трех распределений востребованности процессором блоков памяти с адресуемыми объектами соответственно:

$$\begin{aligned} \bar{F}_1 &= \{f_0 = 0,47; f_1 = 0,26; f_2 = 0,15; f_3 = 0,09; f_4 = 0,03\}, \\ \bar{F}_2 &= \{f_0 = 0,35; f_1 = 0,3; f_2 = 0,2; f_3 = 0,1; f_4 = 0,05\}, \\ \bar{F}_3 &= \{f_0 = 0,24; f_1 = 0,22; f_2 = 0,2; f_3 = 0,18; f_4 = 0,16\}. \end{aligned}$$

Сравнительные значения вероятности попадания в кэш и ее оценки

$P(2, \overline{F_k})/e$	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$P(2, \overline{F_1})$	0,657	0,640	0,624	0,607	0,590	0,572	0,551	0,528	0,498	0,456	0,359
$\hat{P}(2, \overline{F_1})$	0,657	0,653	0,648	0,643	0,638	0,633	0,627	0,621	0,615	0,608	0,601
$P(2, \overline{F_2})$	0,569	0,558	0,546	0,534	0,521	0,506	0,489	0,469	0,442	0,403	0,326
$\hat{P}(2, \overline{F_2})$	0,569	0,567	0,566	0,564	0,562	0,56	0,558	0,556	0,554	0,552	0,55
$P(2, \overline{F_3})$	0,433	0,426	0,42	0,415	0,411	0,406	0,401	0,396	0,39	0,383	0,372
$\hat{P}(2, \overline{F_3})$	0,433	0,431	0,43	0,429	0,428	0,426	0,425	0,424	0,422	0,421	0,419

Из таблицы нетрудно видеть, что при движении вектора вероятностей востребованности адресуемых объектов к равномерному распределению (от  $\overline{F_1}$  к  $\overline{F_3}$ ) вероятность попадания в кэш и ее оценка для уровня ошибок  $e \leq 0,3$  сближаются. Очевидно, что с ростом ассоциативности увеличивается разность значений вероятности между «наиболее» и «наименее» востребованными процессором адресуемыми объектами, загруженными в кэш. Тогда в ассоциативном кэше ошибки вытеснения различных объектов, загруженных в кэш, будут связаны между собой неравенством  $e_a \leq e_{a+1}, a = 0, A-3$ . Как следствие, при этом снижается уровень ошибочных действий, связанных с вытеснением востребованных блоков памяти (с номерами  $m < A-1$ ). Тогда для интегрального уровня ошибок  $\sum_{a=0}^{A-2} e_a \leq 0,3$  оценка (7) может быть использована при расчетах вероятности попадания в кэш. Таким образом, с ростом ассоциативности влияние степени рассеяния вероятностей востребованности процессором адресуемых объектов на оценку вероятности попадания в кэш также снижается. Аналогичная картина имеет место с увеличением числа блоков памяти, отображаемых на кэш заданной ассоциативности  $A$ .

### Заключение

Предложена модель идеального ассоциативного кэша с ошибками вытеснения востребованных процессором блоков памяти в виде цепи Маркова. Для получения аналитической оценки вероятности попадания в кэш введено допущение о неидеальном вытеснении только в состоянии, когда кэш заполнен самыми востребованными блоками памяти. Получены оценки вероятностей обнаружения отдельных блоков памяти в кэше и оценка вероятности попадания в кэш произвольной ассоциативности (7). Показано, что применение полученных оценок возможно при отклонении от стратегии идеального вытеснения с ошибками, не превышающими уровня 0,3. Отмечается, что с ростом ассоциативности кэша и объема отображаемой на кэш памяти влияние ошибок вытеснения снижается. Направлением дальнейших исследований возможен анализ с помощью имитационного моделирования или натуральных измерений вероятностей ошибочных замещений для различных стратегий вытеснения адресуемых объектов и широкого спектра приложений с различной степенью локализации.

### Список источников

1. Танненбаум Э., Остин Т. Архитектура компьютера. 6-е изд. СПб.: Питер, 2021. 816 с.
2. Сущенко М.С., Сущенко С.П. Математические модели иерархической памяти вычислительных систем. Новосибирск: Наука; Томск: Изд-во НТЛ, 2012. 148 с.
3. Сущенко С.П. Архитектура вычислительных систем: учеб. пособие. Томск: СКК-Пресс, 2006. 198 с.
4. Герасимов С.В., Машечкин И.В., Петровский М.И., Попов И.С., Терехин А.Н., Чернов А.В. Организация кэширования: учеб.-метод. пособие. М.: Изд. отд. факультета ВМК МГУ, 2011. 26 с.
5. Song J., Xiaodong Z. LIRS: An efficient low inter-reference recency set replacement policy to improve buffer cache performance // ACM Sigmetrics Performance Evaluation Review. 2002. V. 30, is. 1. P. 31–42.
6. Zhou Y., Philbin J.F. The Multi-Queue Replacement Algorithm for Second Level Buffer Caches // Proc. of the General Track: 2001 USENIX Annual Technical Conference. 2001. P. 91–104.
7. Лускинд Ю.И. Буферные запоминающие устройства типа кэш // Зарубежная радиоэлектроника. 1990. № 4. С. 29–42.

8. Толмачев В.В., Булах С.С., Клычков И.А. Результаты экспериментального исследования кэш-памяти микропроцессорных систем на основе имитационного моделирования // T-Comm: телекоммуникации и транспорт. 2017. Т. 11, № 11. С. 41–47.
9. Ключев А.О., Антонов А.А. Измерение производительности компонентов подсистемы памяти для гетерогенных систем на кристалле // Программные продукты и системы. 2016. Т. 29, № 4. С. 78–84.
10. Сущенко М.С., Сущенко С.П. Моделирование иерархической памяти вычислительных систем неблокирующего типа // Вестник Томского государственного университета. 2000. № 269. С. 87–90.
11. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979. 432 с.

#### References

1. Tannenbaum, E. & Ostin, T. (2021) *Arkhitektura komp'yutera* [Computer Architecture]. 6th ed. St. Petersburg: Piter.
2. Sushchenko, M.S. & Sushchenko, S.P. (2012) *Matematicheskie modeli ierarkhicheskoy pamyati vychislitel'nykh sistem* [Mathematical models of hierarchical memory of computing systems]. Novosibirsk: Nauka; Tomsk: NTL.
3. Sushchenko, S.P. (2006) *Arkhitektura vychislitel'nykh sistem* [Architecture of Computing Systems]. Tomsk: SKK-Press.
4. Gerasimov, S.V., Mashechkin, I.V., Petrovsky, M.I., Popov, I.S., Terekhin, A.N. & Chernov, A.V. (2011) *Organizatsiya keshirovaniya* [Organization of Caching]. Moscow: Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University.
5. Song, J. & Xiaodong, Z. (2002) LIRS: An efficient low inter-reference recency set replacement policy to improve buffer cache performance. *ACM Sigmetrics Performance Evaluation Review*. 30(1). pp. 31–42.
6. Zhou, Y. & Philbin, J.F. (2001) The Multi-Queue Replacement Algorithm for Second Level Buffer Caches. *Proceedings of the General Track: 2001 USENIX Annual Technical Conference*. pp. 91–104.
7. Luskind, Yu.I. (1990) Bufernye zapominayushchie ustroystva tipa kesh [Cache Buffer Storage Devices]. *Zarubezhnaya radioelektronika*. 4. pp. 29–42.
8. Tolmachev, V.V., Bulach, S.S. & Klychkov, I.A. (2017) Rezul'taty eksperimental'nogo issledovaniya kesh-pamyati mikroprotse-sornykh sistem na osnove imitatsionnogo modelirovaniya [Results of Experimental Study of Cache Memory of Microprocessor Systems Based on Simulation Modeling]. *T-Comm: Telekommunikatsii i transport*. 11(11). pp. 41–47.
9. Klyuchev, A.O. & Antonov, A.A. (2016) Izmerenie proizvoditel'nosti komponentov podsystemy pamyati dlya geterogennykh sistem na kristalle [Measuring the Performance of Memory Subsystem Components for Heterogeneous Systems on a Chip]. *Programmnye produkty i sistemy*. 29(4). pp. 78–84.
10. Sushchenko, M.S. & Sushchenko, S.P. (2000) Modelirovanie ierarkhicheskoy pamyati vychislitel'nykh sistem neblokiryuyushchego tipa [Modeling of Hierarchical Memory of Non-Blocking Computing Systems]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 269. pp. 87–90.
11. Kleinrock, L. (1979) *Teoriya massovogo obsluzhivaniya* [Queueing Theory]. Moscow: Mashinostroenie.

#### Информация об авторах:

**Воробейчиков Сергей Эрикович** – доктор физико-математических наук, профессор кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: sev@mail.tsu.ru

**Корсун Александр Викторович** – аспирант кафедры прикладной информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: aleksnfs15@gmail.com

**Морозова Анна Сергеевна** – доцент, кандидат физико-математических наук, доцент кафедры прикладной информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: annamo12@yandex.ru

**Сущенко Сергей Петрович** – профессор, доктор технических наук, заведующий кафедрой прикладной информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета; Томский филиал Федерального исследовательского центра информационных и вычислительных технологий (Томск, Россия). E-mail: ssp.inf.tsu@gmail.com

**Шкуркин Алексей Сергеевич** – доцент, кандидат технических наук, доцент кафедры прикладной информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: shkurkin@mail.tsu.ru

**Вклад авторов:** все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

#### Information about the authors:

**Vorobeychikov Sergey E.** (Doctor of Physical and Mathematical Sciences, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: sev@mail.tsu.ru

**Korsun Alexander V.** (Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: aleksnfs15@gmail.com

**Morozova Anna S.** (Candidate of Physical and Mathematical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: annamo12@yandex.ru

**Sushchenko Sergey P.** (Doctor of Technical Sciences, Professor, National Research Tomsk State University; Tomsk Branch of the Federal Research Center for Information and Computing Technologies Tomsk, Russian Federation). E-mail: ssp.inf.tsu@gmail.com

**Shkurkin Alexey S.** (Candidate of Technical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: shkurkin@mail.tsu.ru

*Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.*

*Поступила в редакцию 10.11.2025; принята к публикации 05.03.2026*

*Received 10.11.2025; accepted for publication 05.03.2026*