ОЦЕНКА ТОЧНОСТИ НЕПРЕРЫВНОЙ И ДИСКРЕТНОЙ МОДЕЛЕЙ МАРКОВА

Михайлова Т.В.
Кафедра ПМиИ ДонНТУ
tanya@r5.dgtu.donetsk.ua

Abstract

Michailova T. Estimation of exactness of the continuous and discrete models Markov`s. In the work there is estimations of exactness of continuous and discrete models.

Введение

Одной из проблем при использовании дискретных и непрерывных моделей является уточнение границы их использования. Дискретные Марковские модели, по сравнению с непрерывными, отражают работу вычислительной системы более точно, поскольку ВС имеют дискретный характер работы, и в вычислительной среде присутствует элемент вероятности. Однако непрерывные модели менее трудоемки. В различных работах приводятся количественные оценки только для непрерывных моделей [1], ссылки на оптимальное соотношение между эффективностью и точностью непрерывных моделей [2], оценки погрешности укрупненных состояний для непрерывных моделей [3]. Определим точность непрерывной и дискретной моделей и границы их использования. Сравним непрерывную и дискретную модели на примере вычислительной системы из K одинаковых устройств, которая обслуживает не более M заявок.

Непрерывная модель

Пусть длительность обслуживания заявки имеет экспоненциальное распределение с математическим ожиданием $\mu = V/\Theta$ c^{-1}, где Θ-трудоемкость обслуживаемой заявки [единиц обслуживания], V-быстродействие обслуживания [единиц обслуживания в сек.] на каждом из устройств.

Построим непрерывную модель. Граф переходов представлен на рис.1. За состояние примем количество обрабатываемых на устройствах заявок m $(m=\overline{0,M})$. Т.о. множество состояний $S=\{(m) \mid m=\overline{0,M}\}$. Количество состояний $L=M+1$.

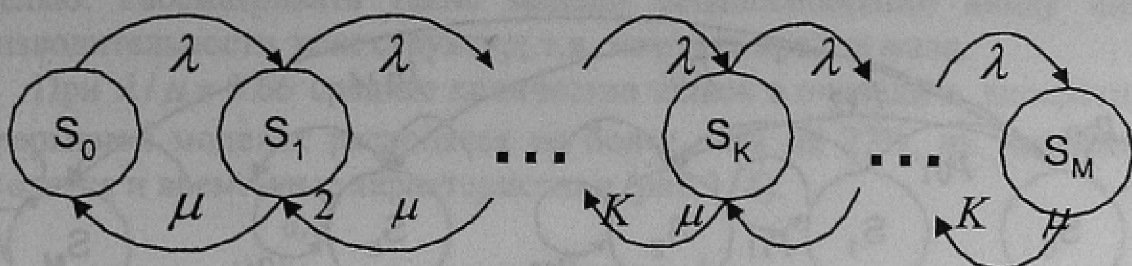

Рисунок 1– Граф переходов непрерывной модели

Для определения вектора стационарных вероятностей $\overline{\pi} = (\pi_0, \pi_1, ..., \pi_M)$ составим уравнения баланса:

$$\begin{cases} \lambda * \pi_0 + \mu * \pi_1 = 0, \\ \lambda * \pi_{i-1} - (\lambda + (i-1) * \mu) * \pi_i + i * \mu * \pi_{i+1} = 0, \quad если\ i = \overline{2, K}, \\ \lambda * \pi_{i-1} - (\lambda + K * \mu) * \pi_i + K * \mu * \pi_{i+1} = 0, \quad если\ i = \overline{K+1, K}. \end{cases}$$

Одно из уравнений заменим условием нормировки [4].

Характеристики, получаемые с помощью непрерывной модели [4]: среднее число занятых устройств в узле, загрузка устройств, среднее число задач в узле, среднее число задач, находящихся в очереди, среднее время пребывания в узле.

Дискретная Марковская модель

Для построения дискретной марковской модели, соответствующей той же системе, введем:

$q_0 = \lambda * \tau$ $(r_0 = 1 - q_0)$-вероятность того, что в следующий такт времени (где τ- такт времени), заявка поступит (не поступит) на обслуживание в ВС;

$q_1 = \tau * \mu$ $(r_1 = 1 - q_1)$- вероятность того, что в следующий такт времени, устройство завершит (не завершит) обслуживание.

Матрица переходных вероятностей $P = \left\{p_{ij}, i = \overline{0, M}, j = \overline{0, M}\right\}$ и СЛАУ для определения стационарных вероятностей строятся в соответствии с методикой [4].

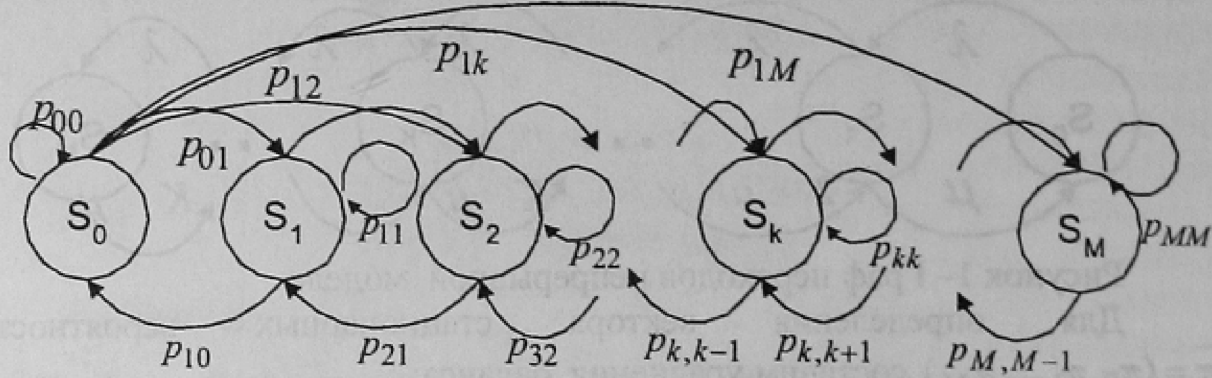

Рисунок 2– Граф переходов дискретной модели

Элементы матрицы переходных вероятностей P определяются следующим образом:

$$p_{ji} = \begin{cases} r_0 C_K^{j-i} q^{j-i} r^{K-j+i} + q_0 C_K^{j-i+1} q^{j-i+1} r^{K-j}, & \textit{если} \quad j-i<K, \\ r_0 C_j^{j-i} q^{j-i} r^{i} + q_0 C_j^{j-i+1} q^{j-i+1} r^{i}, & \textit{если} \quad i<j<K, \\ r_0 r^{j} + q_0 C_j^{1} q^{1} r^{j-1}, & \textit{если} \quad i=j<K, \\ r_0 r^{K} + q_0 C_K^{1} q^{1} r^{K-1}, & \textit{если} \quad i=j\geq K, \\ q_0 r^{j}, & \textit{если} \quad i+1=j<K, \\ q_0 r^{K}, & \textit{если} \quad i+1=j\geq K, \\ 0, & \textit{иначе.} \end{cases}$$

Для определения вектора стационарных вероятностей необходимо решить СЛАУ $\overline{\pi} = \overline{\pi} P$, заменив одно из уравнений условием нормировки.

Характеристики ВС, вычисляемые с помощью дискретной модели, рассчитываются аналогично [4].

Сравнительный анализ моделей

Сравним эти две модели. Зафиксируем количество устройств *K*, отношение λ/μ и исследуем некоторые характеристики вычислительной среды (время обслуживания задачи, среднее количество задач в очереди, загрузку) в зависимости от количества задач *M*. Выберем те значения *M*, при которых ВС не будет терять заявки.

Пусть в узле два устройства. При $\lambda/\mu/K \leq 0.4$ характеристики, вычисляемые дискретной и непрерывной моделями совпадают. Поэтому для анализа эффективности ВС можно пользоваться непрерывной моделью. Рассматривать такие модели нецелесообразно ввиду низкой производительности этих структур, т.к. загрузка крайне мала.

При $\lambda/\mu = 0.66$ среднее количество заявок в очереди в дискретной и непрерывной моделях расходятся не более, чем на 27%, на столько же расходятся и временные характеристики (рис.3, 4).

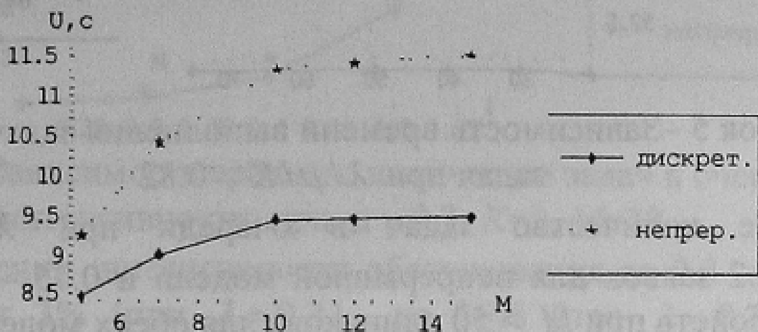

Рисунок 3 –Зависимость времени выполнения задачи от количества задач при $\lambda/\mu/K = 0.66$

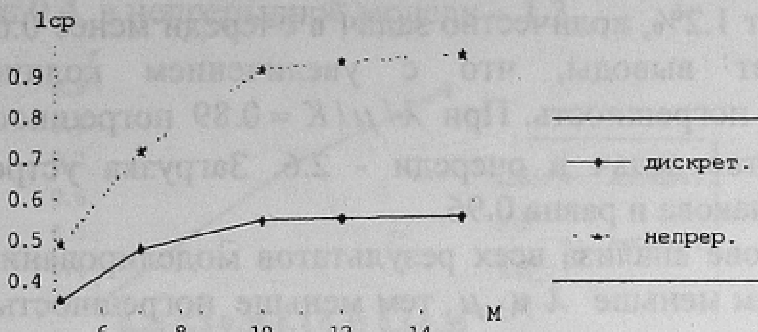

Рисунок 4 –Зависимость среднего количества заявок в очереди от количества задач при $\lambda/\mu/K = 0.66$

Загрузка устройств в дискретной и непрерывной моделях при количестве задач более 10 остается одинаковой и равна 0.67, при M=5 – равна 0.62. В этом случае анализировать эффективность ВС целесообразнее с использованием дискретных Марковских моделей, если требуется точность временных характеристик.

Если далее увеличивать соотношение $\lambda/\mu/K$, загрузка увеличивается, возрастает количество задач в очереди, следовательно, увеличивается погрешность. В случае предельных загрузок вычислительные структуры нецелесообразно исследовать.

При увеличении количества устройств до 10 получаются следующие зависимости. При $\lambda/\mu/K = 0.66$ погрешность временных характеристик не превышает 4%, загрузка устройств при этих данных равна 0.66, а количество задач в очереди не превышает 0.3. Т.о., с увеличением количества устройств при одинаковх соотношениях $\lambda/\mu/K$, уменьшается погрешность. При $\lambda/\mu/K = 0.82$ погрешность характеристик не превышает 18% (рис. 5).

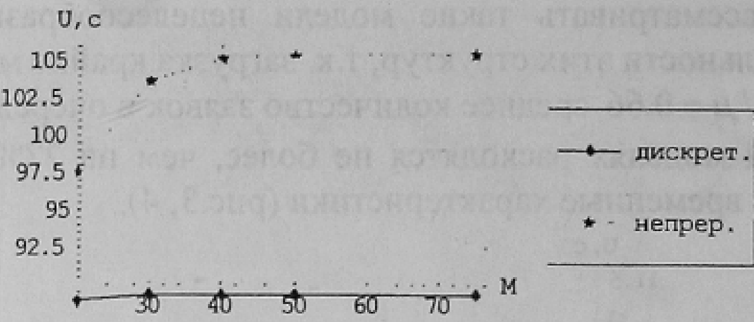

Рисунок 5 –Зависимость времени выполнения задачи от количества задач при $\lambda/\mu/K = 0.82$

Среднее количество задач в очереди при $\lambda/\mu/K = 0.82$ не превышает 2.2 заявок для непрерывной модели и 0.75 – для дискретной. Загрузка устройств при $M > 50$ одинакова для обеих моделей и равна 0.82.

Увеличим количество устройств до 50. Получаются следующие зависимости. При $\lambda/\mu/K \leq 0.82$ погрешность временных характеристик не превышает 1.2%, количество задач в очереди менее 0.6. Эти результаты подтверждают выводы, что с увеличением количества устройств уменьшается погрешность. При $\lambda/\mu/K = 0.89$ погрешность не превышает 5%, количество задач в очереди - 2.6. Загрузка устройств для обеих моделей одинакова и равна 0.96.

На основе анализа всех результатов моделирования можно сделать вывод, что чем меньше λ и μ, тем меньше погрешность.

Зафиксируем количество задач *M*, количество устройств *K* интенсивность μ и исследуем характеристики вычислительной среды (время обслуживания задачи, среднее количество задач в очереди, загрузку) при изменении интенсивности входного потока.

Пусть $\mu = 0.3$, *K*=2, *M*=20. Загрузка изменяется от 0.3 до 0.92. Погрешность временных характеристик (рис.6) и среднего количества задач в очереди (рис. 7) не превышает 35%, загрузки для обеих моделей одинаковы (рис.8). При $\lambda \leq 0.2$ (загрузка не превышает 0.4) погрешность менее 5%.

Уменьшим интенсивность обслуживания μ до 0.1. Погрешность в этом случае не превышает 9% (рис.9, 10). Загрузка изменяется от 0.25 до 0.93.

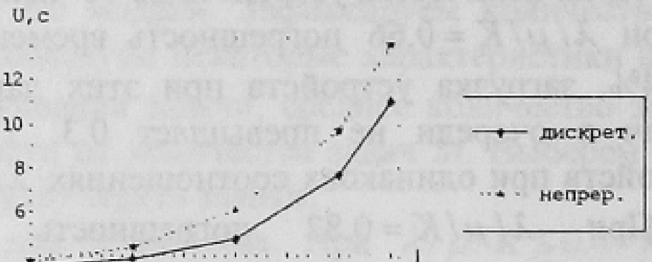

Рисунок 6 –Зависимость времени выполнения задачи от интенсивности при $\mu = 0.3$, *K*=2, *M*=20

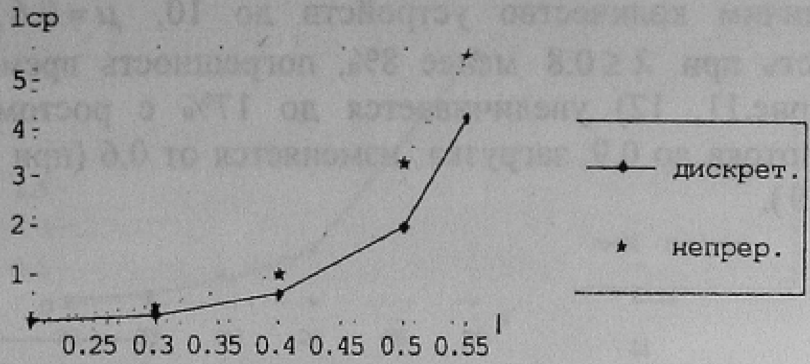

Рисунок 7 –Зависимость среднего количества задач в очереди от интенсивности при $\mu = 0.3$, K=2, M=20

При увеличении интенсивности обслуживания до 0.5 погрешности увеличиваются от 3% (при $\lambda = 0.2$) до 36% (при $\lambda = 0.7$) с ростом интенсивности входного потока, загрузка приборов изменяется от 0.2 (при $\lambda = 0.2$) до 0.7 (при $\lambda = 0.7$), количество задач в очереди при $\lambda = 0.7$ в дискретной модели 0.4, в непрерывной модели – 1.3.

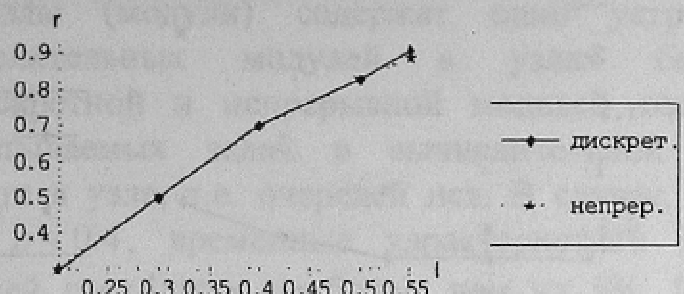

Рисунок 8 –Зависимость загрузки от интенсивности при $\mu = 0.3$, K=2, M=20

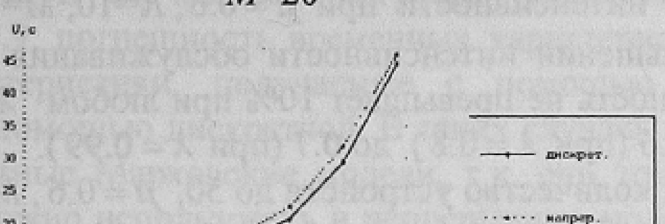

Рисунок 9 –Зависимость времени выполнения задачи от интенсивности при $\mu = 0.1$, K=2, M=20

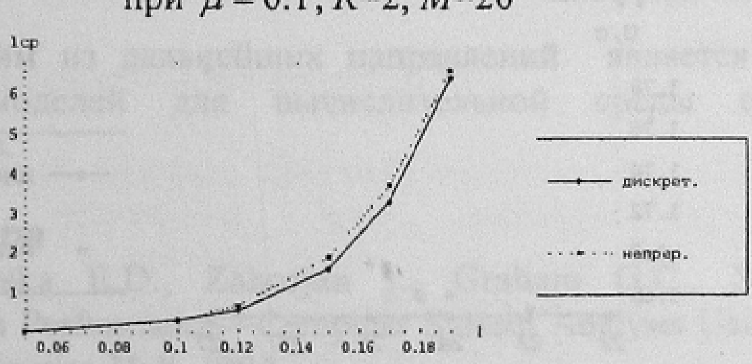

Рисунок 10 –Зависимость среднего количества задач в очереди от интенсивности при $\mu = 0.1$, K=2, M=20

Увеличим количество устройств до 10, $\mu = 0.6$, K=10, M=100. Погрешность при $\lambda \le 0.8$ менее 8%, погрешность времени выполнения задачи (рис.11, 12) увеличивается до 17% с ростом интенсивности входного потока до 0.9, загрузка изменяется от 0.6 (при $\lambda = 0.6$) до 0.83 (при $\lambda = 0.9$).

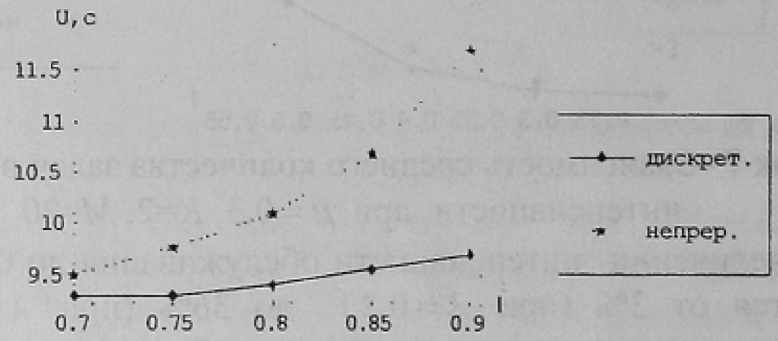

Рисунок 11 –Зависимость времени выполнения задач от интенсивности при $\mu = 0.6$, K=10, M=100

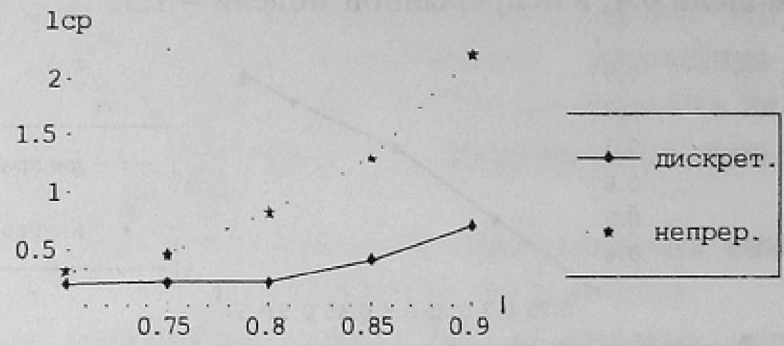

Рисунок 12 –Зависимость среднего количества задач в очереди от интенсивности при $\mu = 0.6$, K=10, M=100

При уменьшении интенсивности обслуживания до 0.1 при тех же данных погрешность не превышает 10% при любом λ, при этом загрузка изменяется от 0.6 (при $\lambda = 0.8$) до 0.7 (при $\lambda = 0.99$).

Увеличим количество устройств до 50, $\mu = 0.6$, M=500. Погрешность времени выполнения задачи не превышает 8% (рис.13, 14), загрузка изменяется от 0.7 (при $\lambda = 22$) до 0.9 (при $\lambda = 27$). Дальнейшее увеличение интенсивности входного потока влечет увеличение погрешности и нагрузки.

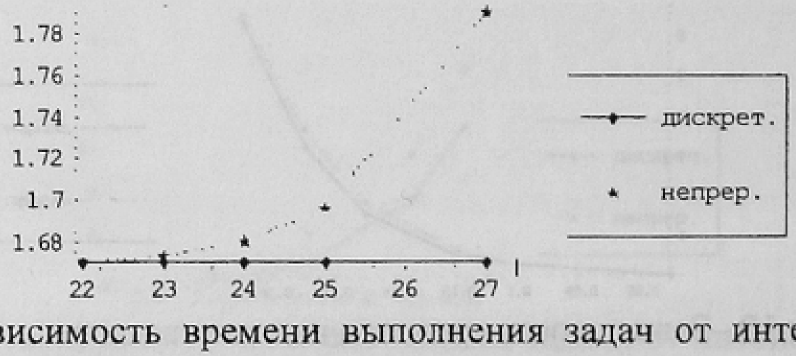

Рисунок 13 –Зависимость времени выполнения задач от интенсивности при $\mu = 0.6$, K=50, M=500

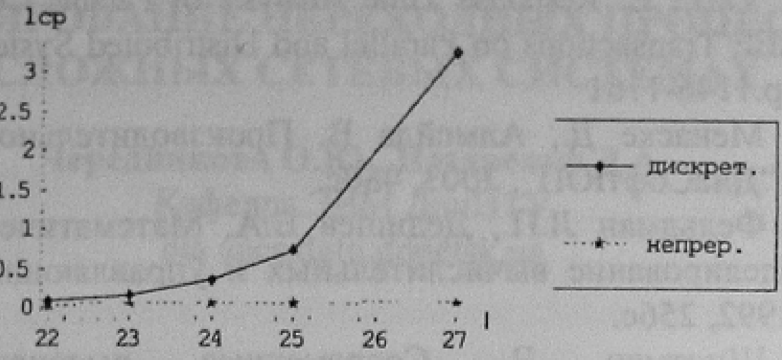

Рисунок 14 –Зависимость среднего количества задач в очереди от интенсивности при $\mu = 0.6$, K=50, M=500

Выводы

На основе анализа результатов моделирования с помощью дискретной и непрерывной моделей можно сделать выводы: характеристики, получаемые обеими моделями идентичны в случаях, если все вычислительные узлы (модули) содержат одно устройство. Если количество вычислительных модулей в узлах более одного, характеристики дискретной и непрерывной моделей совпадают, если количество обрабатываемых задач в вычислительной среде менее количества устройств в узле, т.е. очередей нет. В случае, если загрузки узлов небольшие $\rho < 0.4$, временные характеристики дискретной и непрерывной моделей расходятся не более, чем на 5%. В этом случае эффективнее для исследования ВС использовать непрерывные модели, т.к. они менее трудоемки. В случае, если загрузка устройств находится в пределах от 0.4 до 0.9, погрешность временных характеристик достигает 35%, причем характеристики, получаемые с помощью непрерывной модели, хуже, чем с помощью дискретной. В таких случаях целесообразно использовать дискретные Марковские модели, т.к. они точнее отражают работу ВС, однако можно использовать и непрерывные модели, если надо получить только характеристики, обусловленные загрузкой.

Кроме этого, при прочих равных условиях, чем больше количество устройств в узле, тем меньше расхождения в дискретной и непрерывной моделях.

Поэтому одним из дальнейших направлений является разработка приближенных моделей для вычислительной среды с большим количеством узлов.

Литература

1. Lazowska E.D., Zahorjan J., Graham G.C., Sevcik K.C. Quantitative System Performance—Computer System Analysis Using Queueing Network Models. Prentice-Hall, 1984

2. Varki E. Response Time Analysis of Parallel Computer and Storage Systems //IEEE Transactions on Parallel and Distributed Systems, vol. 12, №11, nov.2001, pp.1146-1161

3. Менаске Д., Алмейда В. Производительность WEB-служб.- СПб: ООО “ДиаСофтЮП”, 2003, 480с.

4. Фельдман Л.П., Дедищев В.А. Математическое обеспечение САПР. Моделирование вычислительных и управляющих систем. –Киев: УМК ВО, 1992, 256с.

5. Шнитман В. Современные высокопроизводительные компьютеры. Информационно-аналитические материалы центра информационных технологий, 1996: http//hardware/app_kis.

6. Корнеев В.В. Параллельные вычислительные системы. –М., 1999, 312с.

7. Столингс У. Структурная организация и архитектура компьютерных систем.- М.: Вильямс, 2002, 893с.

Дата надходження до редакції 06.05.2005 р.