

УДК 004.75

Д.Ю. Воронин, аспирант*Севастопольский национальный технический университет**ул. Университетская 33, СевНТУ, Севастополь, Украина, 99053**E-mail: dima@voronins.com***СЕРВИС-ОРИЕНТИРОВАННАЯ КЛАСТЕРИЗАЦИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ ДИАГНОСТИКИ**

Рассматриваются вопросы, связанные с сервис-ориентированной кластеризацией оптимизационных задач (ОЗ) в распределенных системах диагностики (РСД). Описанный механизм позволяет свести управление нагрузкой при решении ОЗ в РСД к частным случаям, для которых имеются эффективные эвристики, полученные в ходе экспериментов, проводимых на базе многоверсионного комплекса имитационных моделей информационных процессов, протекающих в РСД.

Ключевые слова: *распределенная система, управление нагрузкой, сервис-ориентированная кластеризация, решение оптимизационных задач, имитационное моделирование.*

Широкое использование распределенных вычислений в качестве опорного средства большинства информационных технологий (в здравоохранении, энергетике, банковских системах, бизнес процессах и т.д.), а также стремительно растущие требования к обеспечению гарантоспособности вычислительных сервисов, резко обострили проблему эффективного управления нагрузкой в РСД при решении оптимизационных задач [1]. Необходимо отметить, что эта задача достаточно сложна и приводит к графовым отображениям, сводимым к многостадийной задаче составления расписаний и многоиндексной оптимизационной задаче булевского программирования [2,3].

Целью статьи является разработка и внедрение подхода, позволяющего на основании складывающейся информационной ситуации провести кластеризацию ОЗ и тем самым свести управление нагрузкой в РСД к частным случаям, для которых имеются эффективные эвристики, полученные в ходе экспериментов, проводимых на базе многоверсионного [4] комплекса имитационных моделей информационных процессов, протекающих в РСД.

Под РСД будем понимать систему удаленной кардиологической диагностики пациентов, имеющих нарушения в работе сердечно-сосудистой системы. Специфика диагностирования такого рода нарушений такова, что большую часть исследований необходимо выполнять, моделируя повседневную жизнь испытуемого. Существует множество различных методов диагностики кардиологических пациентов. Среди них особо выделяют эхокардиографию (ультразвуковое исследование сердца), нагрузочную пробу и холтеровское мониторирование. УЗИ и метод нагрузочных проб, безусловно, имеют свои преимущества, однако при проведении таких исследований пациент находится в пределах лечебного учреждения и ведет не совсем типичный для него ритм жизни. Таким образом, несмотря на достаточно обширный объем информации, получаемой амбулаторными методами, часто существует необходимость в дополнительных исследованиях, посвященных поиску причинно-следственных связей между состоянием сердечно-сосудистой системы пациента и событиями его повседневной жизни [5]. Сформулируем технические требования к РСД.

1. РСД имеет иерархическую структуру, с множеством подсистем, среди которых следует особо выделить подсистему обработки информации, управляющие воздействия для которой формируют сервисы по управлению нагрузкой в РСД.

2. Все данные, описывающие качество функционирования сердечно-сосудистой системы пациента, можно разделить на две большие группы: статические и динамические. К динамическим отнесем информацию, постоянно приходящую от ЭКГ-регистратора холтеровского мониторирования, а также аудио-поток, описывающий субъективное состояние пациента в режиме реального времени. ЭКГ-регистратор фиксирует объективное состояние сердечно-сосудистой системы пациента, а аудио-поток определяет его субъективное самочувствие. К статическим данным относится история болезни (содержащая также информацию полученную от РСД на предыдущих тактах работы), данные по нагрузочной пробе и видео-поток УЗИ, записанный в амбулаторных условиях еще до начала исследования. Все эти данные используются для прогнозирования состояния пациента и постановки диагноза. При работе системы из частных прогнозов формируется интегральный, который и будет влиять как на диагноз так и на решения по оперативному вмешательству специалистов (срочная госпитализация, удаленное лечение, дальнейшее наблюдение).

3. Все множество задач обработки информации, генерируемое k -ым пациентом в одном цикле работы РСД образует пакет задач, содержательная структура которого представлена на рисунке 1. Направленные дуги представляют собой отношение предшествования между задачами пакета. Прямоугольники соответствуют задачам пакета.

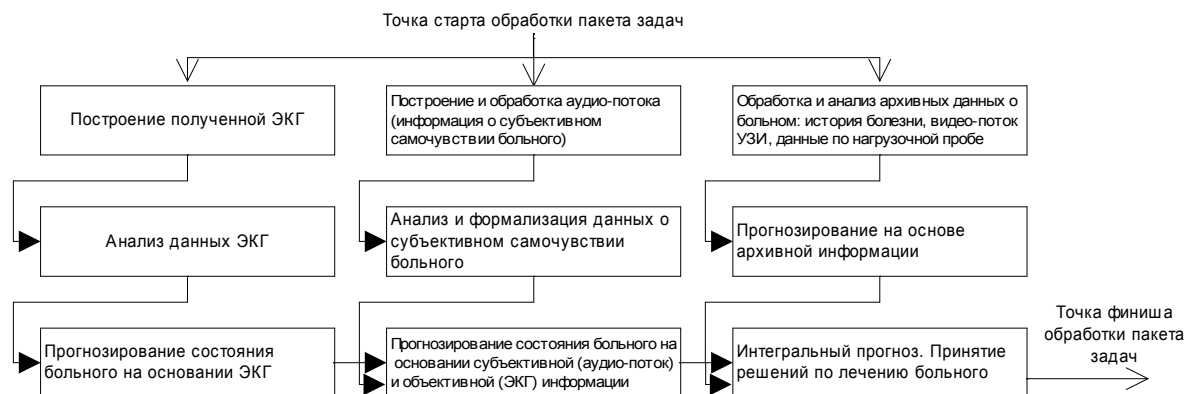


Рисунок 1 – Содержательная структура пакета задач k-ого пациента

Информационная структура пакета задач показана на рисунке 2. В каждом пакете находится входная и результирующая информация, полученная при обработке задачи. На рисунке 2 используются следующие обозначения: а – информация от ЭКГ-регистратора; b – данные о полноте требуемого анализа ЭКГ; с – аудио-поток; d – архивные данные; е – отчет об анализе ЭКГ; f – отчет-прогноз на основании анализа ЭКГ; g – формализованный аудио-поток; h – отчет-прогноз на основании информации о субъективном и объективном состоянии пациента; у – отчет-прогноз на основании архивной информации; z – интегральный отчет-прогноз, сформированное обновление истории болезни пациента.

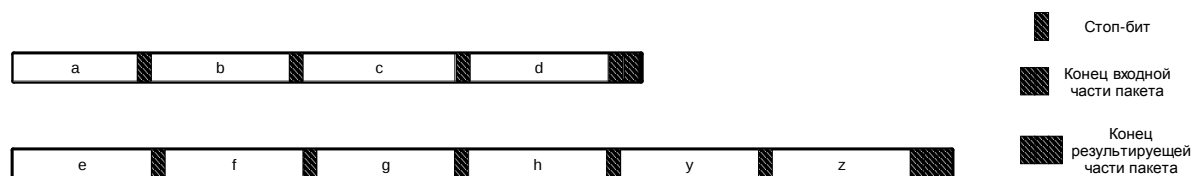


Рисунок 2 – Информационная структура пакета задач

Функционирование диспетчерских сервисов происходит в условиях различного объема априорной информации о РСД и обрабатываемых в ней заданиях. Наряду с точными методами (теория расписаний, булевское программирование и т.п.), применимыми для детерминированного случая, в условиях стохастичности априорной информации приходится прибегать к поиску субоптимальных решений (эвристика, методы искусственного интеллекта, приближенное решение на основе имитационного моделирования [6], нейронных сетей и т.п.). Таким образом, идентификацию информационных ситуаций следует проводить исходя из уровня априорной информации о пакете обрабатываемых заданий. Для всех заданий пакета могут быть априорно заданы следующие характеристики: время обработки всех заданий пакета для всех узлов РСД с учетом доставки и пересылки данных; информация об ограничениях на территориальную распределенность; ограничения на максимальное время доставки и пересылки данных; требования к ресурсам обрабатывающих узлов РСД; потребность в данных (связи между заданиями пакета); модель обмена данными; необходимая платформа для развертывания РСД; список зарезервированных ресурсов, время ожидания в очереди; повышенные требования к гарантоспособности используемого узла РСД.

Выделим восемь информационных ситуаций $Inf @ \xrightarrow{H_0} \{InfS_d\}$, $d = 0 \dots 7$, характеризующих уровень априорной информации о пакете обрабатываемых заданий.

Процесс идентификации информационной ситуации может быть формализован при помощи графа переходов, изображенного на рисунке 3. Состояния в графе упорядочены таким образом, что при движении слева направо уровень априорной информации возрастает от полной неопределенности до детерминированного случая, к которому уже может быть применен аппарат теории расписаний.

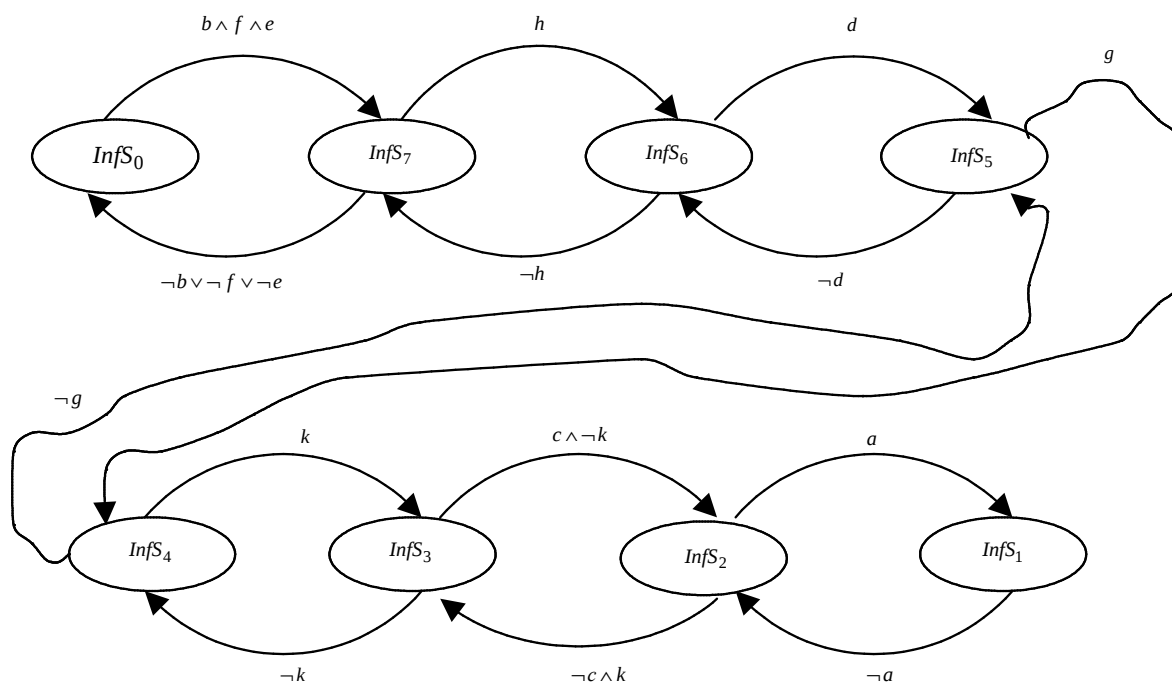


Рисунок 3 – Граф переходов между информационными ситуациями:

- а) время обработки всех заданий пакета для всех узлов РСД с учетом доставки и пересылки данных; б) информация об ограничениях на территориальную распределенность; в) ограничения на максимальное время доставки и пересылки данных; г) требования к ресурсам обрабатывающих узлов РСД; д) потребность в данных (связи между заданиями пакета); е) модель обмена данными; г) необходимая платформа для развертывания РСД; ж) список зарезервированных ресурсов; з) повышенные требования к гарантированности используемого узла РСД

Анализ литературных источников показал, что задача нахождения отображения $G_j \xrightarrow{v_i} G_i$, $v_i : \{v_i \in V_{IIP}; ((\Phi_{отбр} | v_i) \rightarrow \max)\}$ в общем виде является крайне сложной и практически неразрешимой задачей, однако предлагаемый механизм кластеризации позволит свести ее к частным случаям, для которых имеются эффективные процедуры решения, основанные на использовании многоверсионного комплекса имитационных моделей информационных процессов в РСД. Для того чтобы выделить характерные классы ОЗ, решаемых в РСД необходимо разработать систему решающих правил Ψ , позволяющую на основании ИнфС провести кластеризацию большинства ОЗ: каждому классу ОЗ соответствует группа бинарных идентификаторов, объединенных в информационный профиль, представляющий собой геометрический образ в четырехмерном пространстве, которому можно сопоставить множество элементов M_i , причем $M_i \subset RAD$. На рисунках 4,а – 4,в изображены информационные профили для классов $\langle 1,0,1,0 \rangle$, $\langle 0,0,0,0 \rangle$, $\langle 1,1,1,1 \rangle$.

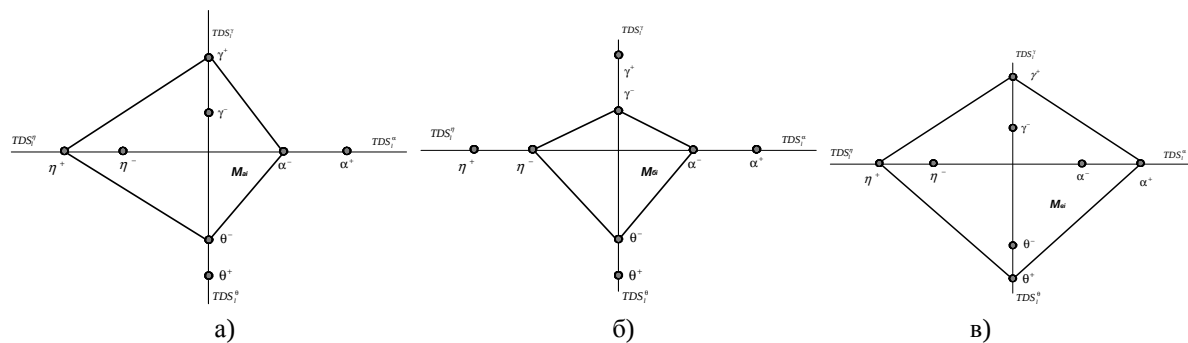


Рисунок 4 – Информационные профили:

- а) класс $\langle 0,0,1,0 \rangle$ б) класс $\langle 0,0,0,0 \rangle$ в) класс $\langle 1,1,1,1 \rangle$

На основании данных о информационной ситуации построим систему правил Ψ , выполняющую кластеризацию ОЗ и имеющую следующий вид:

$$InfS_0^\eta, InfS_0^\alpha, InfS_0^\gamma, InfS_0^\theta \xrightarrow{\Psi} KZD = \eta^\varnothing \alpha^\varnothing \gamma^\varnothing \theta^\varnothing$$

$$InfS_1^\eta, InfS_2^\eta, InfS_3^\eta, InfS_4^\eta \xrightarrow{\Psi} KZD^\eta = \eta^+$$

$$InfS_5^\eta, InfS_6^\eta, InfS_7^\eta \xrightarrow{\Psi^\eta} KZD^\eta = \eta^-$$

Здесь KZD^η – бинарный идентификатор класса ОЗ.

$$InfS_i^\alpha, i = 1 \dots 7 \xrightarrow{\Psi} KZD^\alpha = \begin{cases} \alpha^+, \text{ если } Inf @ \rightarrow pRVS \approx C_{\alpha+}, \\ \alpha^-, \text{ если } Inf @ \rightarrow pRVS \approx C_{\alpha-}, \end{cases}$$

где KZD^α – бинарный идентификатор класса ОЗ, $pRVS$ – используемый тип планирования, $C_{\alpha+}$ – внешнее планирование (присутствует в глобальной среде Grid), $C_{\alpha-}$ – локальное планирование (характерно для кластеров, мультимпьютеров, SMP и DSM архитектур).

$$InfS_i^\gamma, i = 1 \dots 7 \xrightarrow{\Psi} KZD^\gamma = \begin{cases} \gamma^+, \text{ если } Inf @ \rightarrow dRVS \geq C_\gamma, \\ \gamma^-, \text{ если } Inf @ \rightarrow dRVS < C_\gamma, \end{cases}$$

где KZD^γ – бинарный идентификатор класса ОЗ, $dRVS$ – противоречивость критериев эффективности качества управления нагрузкой в РСД, C_γ – порог возникновения конфликта между критериями оценивания эффективности диспетчерских решений.

$$InfS_i^\theta, i = 1 \dots 7 \xrightarrow{\Psi^\theta} KZD^\theta = \begin{cases} \theta^+, \text{ если } Inf @ \rightarrow \exists C_\theta, \\ \theta^-, \text{ иначе} \end{cases},$$

где KZD^θ – бинарный идентификатор класса ОЗ, C_θ – обработка ОЗ производится с дообслуживанием заданий в узле РСД.

Рассмотрим вопросы построения комплекса программных моделей РАД (КПМ), сужающего область поиска решений по управлению нагрузкой в РСД таким образом, что качество получаемых решений не ухудшается. Рассматривая функционирование КПМ на одном из самых верхних уровней абстракции, можно утверждать, что он преобразует множество элементов входного алфавита $\alpha_i \in A$ во множество элементов выходного алфавита $\beta_i \in B$ (отображено на рисунке 5).

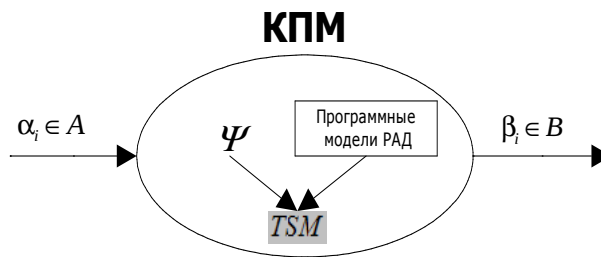


Рисунок 5 – Схема функционирования КПМ в рамках серого ящика

Такое преобразование становится возможным благодаря организованному взаимодействию между таблицей соответствия TSM , программными моделями РАД и системой правил кластеризации ОЗ. Для выполнения функций КПМ для основных классов ОЗ были найдены наиболее подходящие программные модели РАД. Представленные в таблице №1 данные получены в результате множества исследований, проведенных автором при помощи имитационного и натурного моделирования, а также путем глубокого анализа соответствующих литературных источников. В таблице 1 используются следующие условные обозначения: «?» – нет информации о применимости программной модели РАД для данного класса ОЗ, «*» – программная модель РАД актуальна для данного класса ОЗ и может быть применена, « $\vee$ » – логическое ИЛИ: следует выбрать из множества альтернатив помеченных « $\vee$ » только одну; « $\wedge$ » – логическое И: реализация программной модели РАД, помеченная «* $\wedge$ » является для текущего класса ОЗ опорной.

Таблица 1 – Таблица TSM , задающая соответствие между классом ОЗ и ПМ РАД

ПМ РАД класс ОЗ	$ПМ_1$	$ПМ_2$	$ПМ_3$	$ПМ_4$	$ПМ_5$	$ПМ_6$	$ПМ_7$	$ПМ_8$	$ПМ_9$
$\eta^+ \alpha^- \gamma^- \theta^-$	*√	*√							
$\eta^+ \alpha^- \gamma^- \theta^+$	*√	*√		*∧					
$\eta^+ \alpha^- \gamma^+ \theta^-$			*						
$\eta^+ \alpha^- \gamma^+ \theta^+$			*	*∧					
$\eta^- \alpha^+ \gamma^- \theta^-$					*√	*√			*√
$\eta^- \alpha^+ \gamma^- \theta^+$				*∧	*√	*√			*√
$\eta^- \alpha^- \gamma^- \theta^-$							*√	*√	
$\eta^- \alpha^- \gamma^- \theta^+$				*∧			*√	*√	
$\eta^0 \alpha^0 \gamma^0 \theta^0$	?	?	?	?	?	?	?	?	?

Таким образом, таблица соответствия TSM является информационным обеспечением для КПМ: в качестве исходных данных выступает информация о классе ОЗ, а результатом является множество программных моделей РАД, связанных между собой логическими функциями.

Программным моделям РАД поставлено в соответствие множество ПМ, эти модели имеют следующее функционально-целевое назначение: $ПМ_1$ – оптимизирует использование ресурсов на основе теории расписаний и математического аппарата сетевого планирования [2]; $ПМ_2$ – решает задачу о назначении в условиях ограниченности ресурсов; $ПМ_3$ – решает векторную задачу о назначении; $ПМ_4$ – реализует нейросетевую оптимизацию процесса дообслуживания задания в узле РСД; $ПМ_5$ – реализует метод метадиспетчера (интерфейс между службами Grid и локальными планировщиками); $ПМ_6$ – решает задачу диспетчеризации методом брокеров (агентов уровня приложений, занимающихся поиском свободных ресурсов. Пример: проект AppLes); $ПМ_7$ – организует функционирование метода совместного, связанного планирования (gang-scheduling); $ПМ_8$ – решает задачу диспетчеризации методом обратного заполнения (backfilling) и справедливого распределения ресурсов (fairshare) Пример: Maui; $ПМ_9$ – реализует метод Топоркова, состоящий в согласованном выделении ресурсов в РСД.

Подходы к построению моделей $ПМ_5$ – $ПМ_9$ являются общеизвестными, однако задача разработки $ПМ_1$ – $ПМ_4$ обладает элементами научной новизны и состоит в модификации известных методов принятия детерминированных решений по диспетчеризации заданий в сосредоточенных системах. Например, задача гарантоспособной диспетчеризации в РСД дополнительно расширяется следующей оптимизационной подзадачей [6].

На рисунке 6 показана зависимость времени решения задачи управления нагрузкой в РСД от размера пакета с использованием кластеризации и без нее.

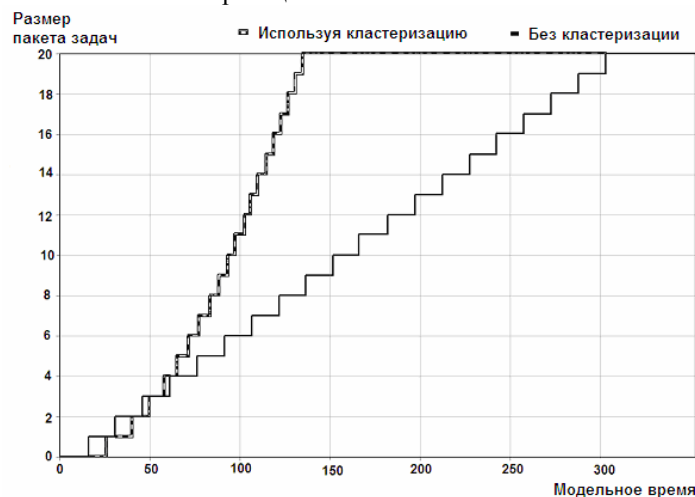


Рисунок 6 – Зависимость времени решения задачи управления нагрузкой в РСД от размера пакета с использованием кластеризации и без нее

Проанализировав приведенные на графике (рисунок 6) результаты можно сделать вывод о том, что эффективность использования предлагаемого подхода значительным образом возрастает при увеличении размерности задачи.

Планируется внедрить предлагаемый подход по кластеризации ОЗ в аппаратно-программный комплекс, осуществляющий ситуативное многоверсионное моделирование информационных процессов в РСД, поиск эффективных стратегий обработки ОЗ в РСД и оценку получаемых решений, что и составит перспективы дальнейшего исследования в данном направлении.

Бibliографический список

1. Скатков А.В. Распределенные системы: структурный анализ, классификация, экстремальные задачи на графах / А.В. Скатков, Д.Ю. Воронин, Д.Н. Данильчук // Вестник национального Хмельницкого университета. — 2007. — №3 Т.1(93) — С. 77–82.
2. Воронин Д.Ю. Системотехнический анализ процессов диспетчеризации в распределенных вычислительных системах графах / Д.Ю. Воронин // Вестник СевГТУ. Сер. Информатика, электроника, связь: сб. науч. тр. — Севастополь, 2008. — С. 38–42
3. Цвиркун А.Д. Основы синтеза структуры сложных систем. / А.Д. Цвиркун. — М.: Наука, 1982. — 200 с.
4. Волковой А.В. Многоверсионные системы и технологии для критических приложений / А.В. Волковой, И.В. Лысенко, В.С. Харченко, О.В. Шурыгин; под ред. В.С. Харченко. — Харьков: ХАИ, 2008. — 224 с.
5. Моисеева Н.М. Оценка ригидности магистральных артерий по данным бифункционального суточного мониторирования АД и ЭКГ прибором BPLab / Н.М. Моисеева, Ю.А. Пономарев, А.Н. Рогоза // Артериальная гипертензия. — К.: Заславский, 2007. — № 1. — С. 61–68.
6. Скатков А.В. Обеспечение гарантоспособности распределенной вычислительной системы на основе моделей диспетчеризации / А.В. Скатков, Д.Ю. Воронин // Радиоэлектронные и компьютерные системы. — Харьков: ХАИ, 2009. — № 5 (39). — С. 31–38.

Поступила в редакцию 02.10.2010 г.