

УДК 004.032.26

Д.Ю. Воронин, ассистент*Севастопольский национальный технический университет**ул. Университетская 33, СевНТУ, Севастополь, Украина, 99053**E-mail: dima@voronins.com***ОБЕСПЕЧЕНИЕ ВЫСОКОЙ ТЕРМИНАЛЬНОЙ ГОТОВНОСТИ НА ОСНОВЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ**

Рассматриваются основные вопросы, связанные с обеспечением высокой готовности медицинских информационных систем критического применения. Предлагаются информационные технологии вариантного анализа распределения вычислительных ресурсов, которые в отличие от существующих основаны на процедурах обратного заполнения в совокупности с комбинаторной и нейросетевой оптимизацией.

Ключевые слова: *высокая готовность, управление вычислительными ресурсами, версионно-модельная избыточность, адаптация, нейронные сети, имитационное моделирование*

По результатам анализа современных достижений в области медицинской информатики можно сделать вывод о том, что в настоящее время существенным образом усилился интерес к использованию медицинских информационных систем (МИС) в лечебных учреждениях. МИС является организационно-административной и программно-технической системой, автоматизирующей процессы сбора, обработки, доставки и хранения потоков информации, необходимой лечебному учреждению при реализации своих базовых функций. Одним из наиболее важных и бурно развивающихся направлений использования МИС являются критические лечебные учреждения. Особенностью применения МИС в данной предметной области является низкая прогнозируемость вычислительной ситуации, обусловленная нестационарным характером ухудшения состояния критического пациента, что в свою очередь чрезвычайным образом затрудняет наступление регулярных событий по информационному обслуживанию медицинских потоков данных. В соответствии с требованиями современных приложений, наиболее перспективным направлением обеспечения высокой готовности МИС является ситуативная политика распределения вычислительных ресурсов, основанная на применении различных эвристик, нейронных сетей, генетических алгоритмов, адаптивного подхода и введении версионно-модельной избыточности. Таким образом, рассматриваемая задача является актуальной.

Целью статьи является дальнейшее развитие информационных технологий вариантного анализа распределения вычислительных ресурсов, которые в отличие от существующих должны быть основаны на процедурах обратного заполнения в совокупности с комбинаторной и нейросетевой оптимизацией, что позволит обеспечить высокую готовность вычислительных сервисов МИС при нестационарной вычислительной нагрузке, характерной для обслуживания большинства лечебных учреждений.

Под *информационной моделью* будем понимать кортеж $\langle \text{ИМ}^X, \text{ИМ}^Y, \text{ИМ}^R \rangle$, где ИМ^X — априорный кортеж, ИМ^Y — результирующий кортеж, ИМ^R — кортеж параметрической настройки.

$\text{ИМ}^X = \langle M, G, D, f, n_{i,j} \rangle$, где M — общий объем вычислительных ресурсов МИС;
 $G = \left(Z, \Gamma_Z = |Z|^2 \right)$ — граф, задающий отношение связности между заданиями в потоке, Z — множество вершин, которым поставлены в соответствие функционально-логические связи между заданиями; множество дуг Γ_Z характеризует трудоемкость обработки заданий (модель дуга-работа);
 $D = \left(Q, \Gamma_N = |Q|^2 \right)$ — граф структуры распределенной системы обработки информации в МИС, здесь Q — множество вершин графа (соответствуют узлам МИС), а Γ_N — множество дуг, которые поставлены в соответствие каналам распределения ресурсов МИС; $f = \{f_{i,j} : (i, j)\}$ — множество функций $f_{i,j}$ (описаны ниже); $n_{i,j} = \{n_{i,j,r} \mid r = 1, \dots, |n_{i,j}|\}$ — множество возможных значений ресурса, выделяемого на дугу (i, j) , причем, $n_{i,j}^{\min} = n_{i,j,1}$, $n_{i,j}^{\max} = n_{i,j,|n_{i,j}|}$.

$IM^Y = \langle \Pi_i, W, N, N_m^k \rangle$, где $\Pi_i = \langle t_k, z_k, n_r \rangle$ — поток заданий, связанных с информационным обслуживанием i -ого критического больного, t_k — момент времени, в течение которого при выделении ресурса объемом n_r будет обрабатываться множество заданий $z_k \subseteq Z$;

$W(v_k, G, D)$ — отображение графа G на D , где $v_k \in V$ представляет собой k -ый вариант отображения графовых структур (перераспределения ресурсов между узлами МИС в процессе обработки заданий), причем $V = \bigtimes_{m=1}^{\Delta |G_Z|} n_m = V_O \cup V_{IP}$, $V_O \cap V_{IP} = \emptyset$; V — множество всех допустимых вариантов отображений, V_f — слабо эффективные отображения, $V_{f\delta}$ — эффективные отображения; $N = \{n_{i,j} : (i,j) \in G_Z\}$ — объем вычислительных ресурсов, перераспределяемый на основе генератора вариантов Gk ;

$IM^R = \langle n_{i,j,r}, t_{i,j,r}, f_{i,j}, N_m^k \rangle$, где $n_{i,j,r} = \{n_{i,j,r} = 1, \dots, |n_{i,j}|\}$ определяет объем ресурса вида r , выделяемого для обработки задания (i, j) ; $t_{i,j,r}$ — длительность обработки задания (i, j) в узле МИС при выделении ресурса в количестве $n_{i,j,r}$; $f_{i,j}$ — функция, устанавливающая соответствие между $n_{i,j,r}$ и $t_{i,j,r}$, $r = 1, \dots, |n_{i,j}|$; $N_m^k = \{n_{i,j} : (i,j) \in G_Z\}$ — суммарное количество ресурса, распределяемого m -му заданию при k -ом распределении.

Под *операционной моделью* $\Phi_{отбр}$, определяющей эффективность отображения W , будем понимать:

$$\Phi_{отбр}(W) = E \left(\Pi_{ДК} \left[\Omega \left(K_{\Phi M}, \xi_{t_k} \left[\Psi(K_{БМ}), MD \right], P_0 \left[\Xi(Inf@, InfS), TSM \right], \Delta, \kappa, \Phi_{t_k}^{\Delta}, \Phi_{t_k}^* \right), t_k \right] \right),$$

где E — критерий эффективности функционирования МИС (выбирается ЛПР); $\Pi_{ДК}$ — вектор-функция состояний МИС (определяется на основе программно-диалогового комплекса поддержки принятия решений по обеспечению высокой готовности МИС); Ω — оператор адаптивного выбора функциональной модели распределения вычислительных ресурсов в узлах МИС для обеспечения ее высокой готовности; $K_{\Phi M}$ — комплекс функциональных моделей обеспечения высокой готовности МИС; ξ_{t_k} — оператор на основе MD получает скалярную величину, характеризующую прогнозные значение потерь, формируемых в результате выбора определенной модели из $K_{\Phi M}$; Ψ — оператор выбора ЛПР оптимального модуля из блока $K_{БМ}$ на основе оценочной матрицы поддержки принятия решения; $K_{БМ}$ — многоверсионный блок модулей вычислительных сервисов МИС; MD — множество моделей принятия решений о свертке векторного критерия в скалярный; P_0 — оператор, рассчитывающий вектор начальных предпочтений для очередной фазы адаптивного выбора функциональной модели обеспечения высокой готовности МИС; Ξ — оператор нейросетевой идентификации информационных ситуаций, связанных с распределением вычислительных ресурсов МИС; $Inf@$ — базовые системные характеристики информационного обслуживания пациентов критического лечебного учреждения; $InfS$ — множество типов информационных ситуаций; TSM — матрица соответствия между $InfS$ и вектором начальных предпочтений для очередной фазы адаптивного выбора функциональной модели из $K_{\Phi M}$; Δ — порог чувствительности при адаптивном выборе функциональной модели распределения вычислительных ресурсов в узлах МИС; κ — управляющий параметр частоты контроля качества адаптации; $\Phi_{t_k}^{\Delta}$ — оператор расчета суммарного уменьшения потерь за κ шагов адаптации; $\Phi_{t_k}^*$ — оператор расчета разброса между экспериментальным и прогнозным значением потерь на каждом 2κ -ом шаге адаптации; $t_k, k \in \{0, 1, 2, \dots\}$ — моменты принятия решений. Следовательно, вариант отображения $v_k \in V$ эффективен, если не существует $v_f \in V$ такого, что $(\Phi_{отбр}|v_f) \geq (\Phi_{отбр}|v_k)$. Вариант $v_k \in V$ слабо эффективен, если не существует $v_f \in V$ такого, что $(\Phi_{отбр}|v_f) > (\Phi_{отбр}|v_k)$. Среди $v_k, k = \overline{1, |V|}$, $\exists v_i : \{v_i \in V_{IP}, ((\Phi_{отбр}|v_i) \rightarrow \max)\}$.

Задача обеспечения высокой готовности МИС в общем виде формулируется следующим образом: необходимо найти $v_i : \{(\Phi_{\text{отбр}} | v_i) \rightarrow \max\}$, при ограничениях ресурса $\sum_{p^k} N_m^k \leq M$ и выполнении условия $\forall v_i \in V_{\text{лр}}$.

Под готовностью будем понимать качественную характеристику, отражающую способность системы выполнить заданный комплекс задач в установленное время. Рассмотренные в [1] недостатки общепринятых методов распределения вычислительных ресурсов МИС сформировали требования к комплексу функциональных моделей $K_{\Phi M} = \{\Phi M_1, \Phi M_2, \Phi M_3\}$, где ΦM_1 — функциональная модель, базирующаяся на алгоритме обратного заполнения (Backfill), модифицированного в направлении оценивания трудоемкости выполнения заданий в вычислительных узлах МИС; ΦM_2 — функциональная модель решения модифицированной нейросетевой задачи распределения вычислительных ресурсов МИС [2]; ΦM_3 — функциональная модель, определяющая приоритет заданий в очереди на упреждающую пересылку [3] с последующим распределением вычислительных ресурсов на основе модифицированной векторной задачи о назначении (реализована с учетом повышения отказоустойчивости МИС).

Необходимо заметить, что применение адаптивных технологий к выбору функциональной модели ΦM_i , $i = 1 \dots 3$ позволяет парировать неопределенность, связанную с динамически изменяющейся интенсивностью поступления требований на обслуживание заданий в вычислительных узлах МИС. Математической базой для предлагаемого подхода послужили работы Л.А. Растригина, А.В. Назина, А.С. Позняка, однако оперативный характер решений по распределению вычислительных ресурсов в МИС потребовал повышенной скорости сходимости процедур стохастической аппроксимации, реализованной при использовании следующих информационных технологий:

- применение во внутреннем контуре адаптации многоверсионного блока модулей $K_{\text{БМ}}$ для получения прогнозной оценки функции потерь;

- обнуление вектора вероятностей выбора функциональных моделей ΦM_i , $i = 1 \dots 3$ при детектировании изменения типа информационной ситуации в МИС.

Текущая информационная ситуация, идентифицируема средствами, описанными в [4], и влияет на вектор начальных предпочтений, ΦM_i , эквивалентны. Алгоритм адаптивного выбора функциональной модели распределения вычислительных ресурсов в узлах МИС состоит в следующем.

1. Задать V — множество допустимых вариантов управления распределением вычислительных ресурсов МИС. Загрузить множество эквивалентных функциональных моделей ΦM_i , $i = 1 \dots 3$, реализующих отображение $V \mapsto V^{\text{лр}}$, $\varsigma : \{ |V^{\text{лр}}| < |V| \}$.

2. Задать цель управления, например, найти последовательность $v_i \in V_{\text{лр}}$, минимизирующую значения функции средних текущих потерь $\Phi_n = \frac{1}{n} \sum_{t=1}^n \xi_t$, $n = 1, 2, \dots, K$, где K — число тактов адаптации,

а Φ_n характеризует качество адаптации на предыдущих n тактах.

3. При помощи ЛПР необходимо:

- задать Δ , κ — управляющие параметры адаптивного поиска вариантов распределения вычислительных ресурсов в МИС;

- указать значение коэффициента χ , задающего уровень доверия прогнозным потерям;

- для получения прогнозных оценок значения функции потерь выбрать модуль из многоверсионного блока $K_{\text{БМ}}$;

4. Определить вектор начальных предпочтений вероятностей выбора ΦM_i в соответствии с идентифицированным типом информационной ситуации (на основе нейросетевого идентификатора);

5. В каждый из дискретных моментов времени t_n ($n = 1, 2, \dots, K$; $t_{n+1} > t_n$):

- 5.1. При помощи модуля из $K_{\text{БМ}}$ получить прогнозное значение функции потерь ξ_n , которое является последствием реализации варианта распределения ресурсов v_i , полученного в результате использования функциональной модели ΦM_i , выбранной на предыдущем такте адаптации (при использовании рандомизированной стратегии, которая описана ниже).

В условиях неполной информации об интенсивности поступления требований по информационному обслуживанию заданий в МИС, использование детерминированных правил выбора вариантов управления существенно затруднено. Это привело к необходимости использования

стохастического автомата с переменной структурой, использующего рандомизированную стратегию *RndStr* вида:

$$p_{n+1}(\Phi M_i, i=1\dots 3) = F \left(p_1, \dots, p_n; n, \Delta, \kappa, \gamma_n, \left(\frac{1}{n} \sum_{t=1}^n \xi_n[\Psi(K_{BM}), MD] \right), P_0[\Xi(Inf@, InfS), TSM] \right),$$

$$n=1, 2, \dots, K; \quad p_{n+1} = \pi_{\varepsilon_{n+1}}^N [p_n - \gamma_n | \xi_n[\Psi(K_{BM}), MD] - \Delta | \text{sign}(\Delta p_n)],$$

причем $\gamma_n \geq 0, \varepsilon_n \in (0; N^{-1}), \Delta > 0$,

где F — вектор-функция; p_n — вектор условных вероятностей выбора функциональных моделей ΦM_i в момент времени t_n ; n — номер шага адаптивного поиска; Δ — порог чувствительности, γ_n — длина шага поиска; κ — управляющий параметр частоты контроля качества адаптации; $\xi_n[\Psi(K_{BM}), MD]$ — потери на интервале времени (t_n, t_{n+1}) ; P_0 — оператор, рассчитывающий вектор начальных предпочтений для очередной фазы адаптивного выбора функциональной модели обеспечения высокой готовности МИС; $\pi_{\varepsilon_{n+1}}^N$ — оператор проектирования.

5.2. Если $\Phi_n^\Delta < \Delta$, то переход к пункту 4, иначе — к 5.3. Φ_n^Δ — суммарное уменьшение прогнозных потерь за последние κ шагов;

5.3. Если $\Phi_n^* > \chi \cdot \Delta$, то необходимо выдать ЛПР рекомендацию о выборе другого модуля из многоверсионного блока K_{BM} ; Φ_n^* — разброс между экспериментальным и прогнозным значением функции потерь, рассчитываемый каждые 2κ шагов адаптации.

6. Пока $n \neq K$ переход к пункту 5.1, иначе — конец.

Упрощенная алгоритмическая схема адаптивной подсистемы поиска оптимальной функциональной модели обеспечения высокой готовности МИС представлена на рисунке 1.

На рисунке 2 приведены результаты сравнительного анализа эффективности обеспечения высокой готовности МИС для различных типов информационных ситуаций, как при применении отдельных функциональных моделей $\Phi M_1 - \Phi M_3$, так и при их альтернативном использовании в рамках единого комплекса $K_{\Phi M}$. В результате анализа полученных данных можно сделать вывод о том, что средняя эффективность использования $K_{\Phi M}$ выше, чем любой из моделей $\Phi M_1 - \Phi M_3$ и, следовательно, парирование ситуативной неопределенности при использовании $K_{\Phi M}$ помогает повысить высокую готовность МИС на (12...28) %.

Рассмотрены функциональные модели $\Phi M_1 - \Phi M_3$, которые являются эквивалентными, направлены на обеспечение высокой готовности МИС при распределении вычислительных ресурсов для обработки информационных потоков, порождаемых процессами лечения и обследования критических пациентов нефрологического стационара. Использование $\Phi M_1 - \Phi M_3$ альтернативно: эффективность их применения варьируется в зависимости от информационной ситуации, идентифицируемой на основании доступной априорной информации об обрабатываемых заданиях.

Рассмотренные функциональные модели обладают следующими особенностями:

1. ΦM_1 нуждается в предварительной оценке трудоемкости обрабатываемых заданий, что накладывает повышенные требования к объему априорно доступной информации и существенно сужает ее применимость для медицинских систем критического назначения. Таким образом, для большинства ИнфС в векторе начальных предпочтений P_0 вероятность выбора ΦM_1 существенно занижена. Однако в фоновом режиме функционирования МИС, при достаточном объеме априорной информации, эффективность метода *Backfill* стремительно возрастает.

2. ΦM_2 наиболее универсальна и показывает хорошие результаты для большинства информационных ситуаций, при функционировании МИС в сверхоперативном режиме и малом объеме доступной априорной информации об обрабатываемых заданиях.

3. ΦM_3 имеет повышенную актуальность для слабосвязанных графов, в особо сложных информационных ситуациях, при нестационарной интенсивности запросов на обработку заданий, носящих точки перегрузки трафика.

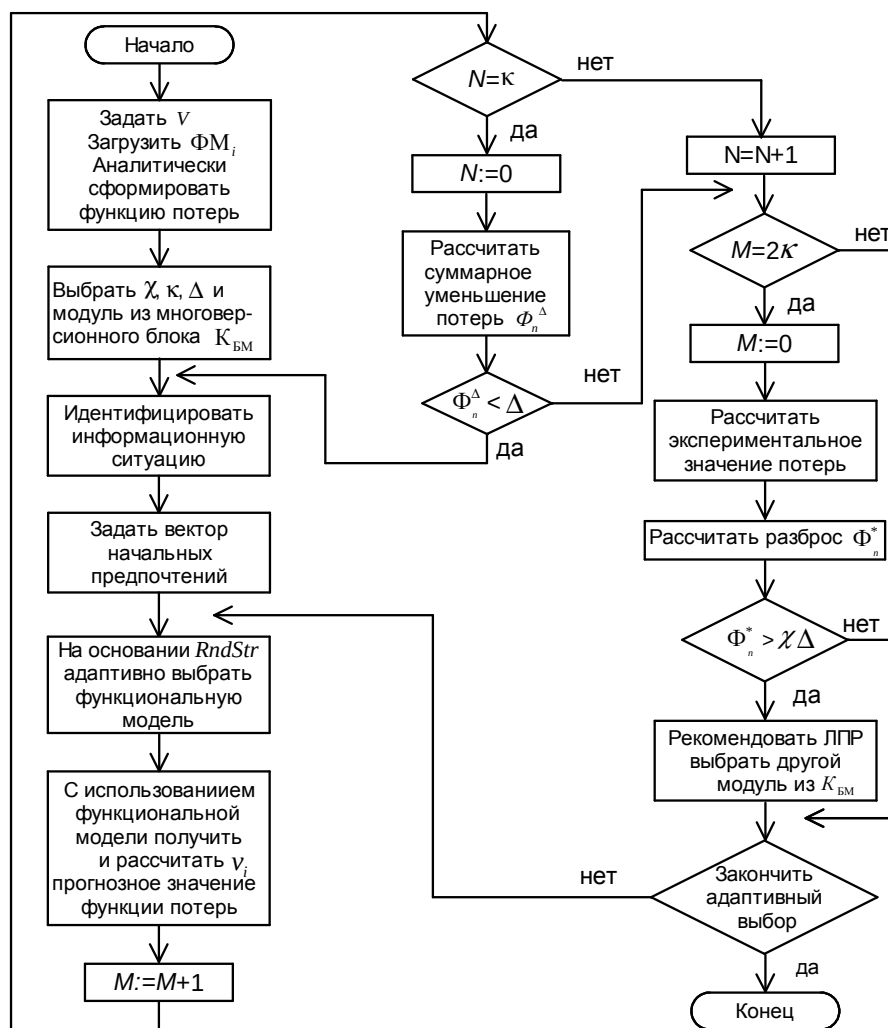


Рисунок 1 — Алгоритмическая схема процедуры адаптивного принятия решений по обеспечению высокой готовности медицинских информационных систем



Рисунок 2 — Результаты сравнительного анализа эффективности применения функциональных моделей и комплекса по обеспечению высокой готовности МИС для различных типов информационных ситуаций

Таким образом, решение по выбору функциональной модели для складывающейся информационной ситуации нужно принимать адаптивно, основываясь на векторе начальных предпочтений P_0 и значениях функции потерь, формируемой при помощи многоверсионного блока модулей вычислительных сервисов МИС.

В качестве направления дальнейших исследований планируется повысить вычислительную устойчивость МИС при обслуживании требований в режиме реального времени и внедрить прилагаемый подход в программно-диалоговый комплекс поддержки принятия решения по обеспечению высокой готовности МИС нефрологического стационара.

Бibliографический список использованной литературы

1. Скатков А.В. Обеспечение гарантоспособности распределенной вычислительной системы на основе моделей диспетчеризации / А.В. Скатков, Д.Ю. Воронин // Радиоэлектронные и компьютерные системы. — 2009. — № 5 (39). — С. 31–38.
2. Воронин Д.Ю. Оперативная диспетчеризация в распределенных вычислительных системах на основе нейросетевых моделей / Д.Ю. Воронин // Радиоэлектронные и компьютерные системы. — 2010. — № 6 (47). — С. 73–77.
3. Скатков А.В. Управление ресурсами узла распределенной системы с использованием нейросетевой оптимизации / А.В. Скатков, Д.Ю. Воронин // Новые технологии. — 2007. — № 4 (18). — С. 63–65.
4. Воронин Д.Ю. Сервис-ориентированная кластеризация оптимизационных задач в распределенных системах диагностики / Д.Ю. Воронин // Оптимизация производственных процессов: сб. науч. тр. — Севастополь, 2010. — Вып. 12. — С. 136–141.

Поступила в редакцию 05.11.2010 г.

Воронін Д.Ю. Забезпечення високої готовності на основі інформаційних технологій розподілу обчислювальних ресурсів

Розглядаються основні питання, що пов'язані із забезпеченням високої готовності медичних інформаційних систем критичного застосування. Пропонуються інформаційні технології варіантного аналізу розподілу обчислювальних ресурсів, які на відміну від існуючих засновані на процедурах зворотнього заповнення в сукупності з комбінаторною та нейромережевою оптимізацією.

Ключові слова: висока готовність, управління обчислювальними ресурсами, нейронні мережі, версійно-модельна надмірність, адаптація, імітаційне моделювання.

Voronin D.Y. Ensuring of systems high availability on the basis of information technologies of resource management

The main issues related to ensuring high availability of medical information systems of critical applications are considered. The information technology of the computing resources distribution is proposed. It is based on the inverse filling in the totality with neural network and combinatorial optimization and, thus, is opposed to existing procedures.

Keywords: high availability, computational resource management, neural networks, decision-making, adaptation, simulation.