

Учреждение образования  
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

УДК 004.75:004.777

НГУЕН  
Хань Куок

**МОДЕЛИ И АЛГОРИТМЫ КОМПОЗИЦИИ ПАКЕТОВ ВЕБ-СЕРВИСОВ**

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата технических наук

по специальности 05.13.15 – Вычислительные машины,  
комплексы и компьютерные сети

Минск 2014

**Работа** выполнена в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники».

**Научный руководитель** **Иванов Николай Николаевич**, кандидат физико-математических наук, доцент, доцент кафедры электронных вычислительных машин учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

**Официальные оппоненты:** **Котов Владимир Михайлович**, доктор физико-математических наук, профессор, заведующий кафедрой дискретной математики и алгоритмики Белорусского государственного университета

**Куликов Святослав Святославович**, кандидат технических наук, доцент, менеджер отдела функционального тестирования №2 ИООО «ЭПАМ Системз»

**Оппонирующая организация** Белорусский национальный технический университет

Защита состоится «19» июня 2014 г. в 14.00 на заседании совета по защите диссертаций Д 02.15.04 при учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, г. Минск, ул. П. Бровки, 6, корп. 1, ауд. 232, тел. 293-89-89, e-mail: dissovet@bsuir.by

## КРАТКОЕ ВВЕДЕНИЕ

С развитием технологий надежного высокоскоростного доступа в интернете приложения в сфере науки, образования, бизнеса, социальные общения людей перемещаются во Всемирную паутину. Веб-сервисы реализуют интерфейсы обмена данными между различными приложениями, которые могут быть созданы на разных языках и распределены на узлах сети. Пользователю для решения сложных задач требуется использовать функциональность нескольких веб-сервисов, в ходе развития технологии возникло понятие композиции веб-сервисов. Задача композиции состоит в определении набора веб-сервисов, согласованное взаимодействие которых обеспечивает функциональность, эквивалентную запросу пользователя. Веб-сервисы делятся на элементарные, которые не содержат других веб-сервисов и сложные, которые представлены композицией взаимосвязанных элементарных веб-сервисов. Элементарный веб-сервис называется атомом, а сложный веб-сервис будем называть просто сервисом.

Композицию атомов можно выполнить вручную по описанию их входных и выходных параметров. Автоматизацию композиции атомов можно организовать различными путями, например, исследуя их семантические модели. В результате применения семантических моделей большинство исследователей предлагают построение ориентированного ациклического взвешенного графа с единственными начальным и завершающим узлами, такой граф далее будем называть сетью. Дуги сети соответствуют атомам, а узлы сети соответствуют событиям. Каждый полный путь построенной сети соответствует композиции атомов, решающих задачу пользователя. Если сеть содержит несколько полных путей, то ставится задача построения оптимальной композиции, то есть выбора оптимального полного пути в сети. Целевая функция задачи учитывает функциональные критерии атомов и нефункциональные критерии пользователя, соответствующие его предпочтениям, таким как стоимость использования сервиса, репутация провайдера сервисов и т.п.

В диссертации расширяется задача композиции сервиса. Зачастую пользователю требуется решение нескольких задач, для каждой из которых строится композиция сервиса. Такая ситуация может возникать при коллективном обсуждении, когда участники ставят несколько задач. Для такого случая предлагается формировать и посылать на выполнение пакет сервисов, для каждого из которых необходимо построить композицию из атомов.

В диссертации на основе методов решений оптимизационных задач на графах и результатов теории расписания были разработаны точные и эвристические алгоритмы решения задачи композиции пакета сервисов.

Некоторые атомы одновременно входят в сети разных сервисов из пакета. Это происходит в ситуациях, когда атом, используемый различными сервисами, обращается к единому ресурсу, захватывая его на время выполнения операции. Сети, построенные для задач пакета, объединяются по общим атомам–дугам, так создается объединенная сеть композиции пакета сервисов. При решении задачи оптимизации на объединенной сети необходимо составить оптимальное расписание использования единого ресурса. Решением задачи композиции пакета сервисов является оптимальный набор взаимосвязанных полных путей объединенной сети, каждый из путей создает композицию одного из сервисов пакета. При этом некоторые атомы–дуга входят в оптимальные пути нескольких сервисов.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Связь работы с крупными научными программами и темами**

Исследования проводились на кафедре электронных вычислительных машин учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» в рамках научно-исследовательской госбюджетной темы ГБ № 11-2019 «Методы и алгоритмы параллельной обработки изображений и видеопоследовательностей, распознавания объектов в режиме реального времени».

### **Цель и задачи исследования**

Цель диссертационной работы состоит в разработке математических моделей и алгоритмов решения задачи композиции пакета сервисов.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать математическую модель задачи композиции пакета сервисов, включающую модель использования сервисами общего атома.
2. Разработать точный и эвристический алгоритмы построения плана использования сервисами общего атома.
3. Разработать точный и эвристический алгоритмы решения задачи композиции пакета сервисов.
4. Сгенерировать тестовые задачи для вычислительного эксперимента, провести сравнительный анализ построенных алгоритмов.
5. Провести эксперимент по применению задач композиции пакета сервисов на одной из систем сервис-ориентированной архитектуры.

Объектами исследования являются веб-сервисы, обеспечивающие распределенную обработку информации в системах и сетях.

Предметом исследования являются архитектура веб-сервисов, методы построения композиции веб-сервисов.

### **Положения, выносимые на защиту**

1. Математическая модель задачи композиции пакета сервисов. В отличие от модели композиции единственного сервиса разработанная модель комбинирует модели задачи календарного планирования, транспортной задачи и задачи построения расписания в одностадийной системе и позволяет решать задачу композиции нескольких сервисов.

2. Точный алгоритм решения задачи композиции пакета сервисов на объединенной сети как модификация алгоритма Беллмана–Форда для решения задачи построения кратчайшего пути на сети с ограничениями. Отличие предложенного алгоритма состоит в том, что в алгоритм Беллмана–Форда встраивается алгоритм решения задачи построения расписания в одностадийной системе, необходимый для разрешения коллизии использования общего атома несколькими сервисами.

3. Генетический алгоритм для решения задачи композиции пакета сервисов на объединенной сети. Отличие алгоритма заключается в том, что он включает эвристическую процедуру для составления расписания использования сервисами общего атома. Генетический алгоритм позволяет решать задачи большой размерности, для которых точные алгоритмы требуют значительных затрат времени. Однако генетический алгоритм не гарантирует оптимальности решения.

4. Точный и эвристический алгоритмы построения расписания использования сервисами общего атома. Точный алгоритм имеет сложность  $O(n^5)$ . Сложность эвристического алгоритма равна  $O(n \log n)$ .

### **Личный вклад соискателя**

Основные результаты и положения, выносимые на защиту, получены лично автором. Математические модели и алгоритмы, рассмотренные в работе, включают алгоритмы решения задачи композиции пакета сервисов; алгоритмы построения расписания использования сервисами общего атома. Они были разработаны и экспериментально исследованы автором самостоятельно. Научный руководитель кандидат физико-математических наук, доцент Н. Н. Иванов принимал участие в постановке целей исследования, их предварительном анализе, а также в обсуждении полученных результатов.

## **Агробация результатов диссертации**

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих международных конференциях: VII Международная научно-практическая конференция молодых исследователей «Содружество наук. Барановичи-2011» (г. Барановичи, 2011); The 11th International Conference (PRIP-2011) «Pattern Recognition and Information Processing» (г. Минск, 2011); Международная научная конференция (ИТС 2011) «Информационные технологии и системы» (г. Минск, 2011); Международная заочная научно-практическая конференция «Актуальные вопросы современной информатики» (г. Коломна, Россия, 2012); 48-я научная конференция аспирантов, магистрантов и студентов БГУИР «Компьютерные системы и сети» (г. Минск, 2012); VII Международная научная конференция «Средства медицинской электроники и новые медицинские технологии» (г. Минск, 2012); VII Международная научно-практическая конференция «Актуальные проблемы экономики, управления, образования и права» (г. Коломна, Россия, 2012); I Международная научно-техническая конференция «Алгоритмические и программные средства в информационных технологиях, радиоэлектронике и телекоммуникациях» (г. Тольятти, Россия, 2013); 49-я научная конференция аспирантов, магистрантов и студентов БГУИР «Компьютерные системы и сети» (г. Минск, 2013); II научно-практическая конференция «Прикладные проблемы оптики, информатики, радиофизики и физики конденсированного состояния» (г. Минск, 2013); Научно-практическая конференция «Современные информационные компьютерные технологии mcIT-2013» (г. Гродно, 2013).

## **Опубликованность результатов диссертации**

По теме диссертационной работы опубликовано 18 печатных работ. Из них 7 статей в научно-технических журналах, включенных в перечень научных изданий Республики Беларусь для опубликования результатов диссертационных исследований, и иностранных рецензируемых журналах (3,1 авторских листов), 11 докладов в сборниках материалов международных конференций (3,0 авторских листов).

## **Структура и объем диссертации**

Диссертация состоит из введения, общей характеристики работы, четырех глав, заключения, библиографического списка и приложения. Общий объем диссертации составляет 125 страниц, из них 99 страниц основного текста, 28

иллюстраций на 9 страницах, 5 таблиц на 4 страницах, библиографический список из 88 наименований на 7 страницах, список собственных публикаций из 18 наименований на 2 страницах и 1 приложение на 4 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ

Во *введении* обоснована актуальность темы диссертационной работы, сформулированы цель и задачи исследования.

В *первой главе* рассмотрены технология веб-сервиса, его общая модель и стандарты. Рассмотрена задача композиции сервиса, дан аналитический обзор моделей задачи оптимальной композиции сервиса и методов ее решения. Поставлена задача композиции пакета сервисов, в которой пользователь посылает запрос на композицию нескольких сервисов. В этой главе как пример применения результатов исследована система архивации и передачи медицинских изображений. Приводится обзор системы и выяснены ее возможности.

*Вторая глава* посвящена разработке математической модели задачи композиции пакета сервисов и ее подзадач. Описаны модели задачи композиции единственного сервиса и модели задачи построения расписания использования сервисами общего атома.

Математическая модель задачи композиции пакета сервисов представляет собой комбинацию задачи календарного планирования, транспортной задачи и задачи построения расписания в одностадийной системе.

Пользователь посылает пакет заданий на выполнение сервисов  $\{S_1, S_2, \dots, S_N\}$ . Для каждого  $k = 1, 2, \dots, N$  задана сеть  $G_k = (V_k, A_k)$  из атомов которой создается композиция сервиса  $S_k$ . Из  $N$  сетей  $G_k$  склеиванием совпадающих атомов создается объединенная сеть  $G = (V, A)$ , где  $V$  – узлы,  $A$  – дуги сети. Узлы обозначим буквами  $i$  и  $j$ , дугу, связывающую узлы  $i$  и  $j$ , обозначим  $(i, j)$ , дуга направлена от  $i$  к  $j$ . Сеть  $G$  имеет единственный начальный узел  $s$  и единственный конечный узел  $f$ .

В сетевой постановке получаем задачу оптимизации целевой функции, нелинейной в общем случае, на потоке заданной величины с ограничениями на дугах. Допустимыми решениями задачи являются  $N$  полных путей  $C_k$  по числу сетей  $G_k$ . Любой полный путь  $C_k$  на сети  $G_k$  выполняет сервис  $S_k$ , точнее этот сервис выполняется последовательностью атомов, образующих путь  $C_k$ .

Определим атомарную функцию  $x_{ij}^k(t)$  для дуг  $(i, j)$  сети  $G_k$ , которая определена на интервале времени  $[0, T_x]$ , который выбирается достаточно большим так, чтобы к моменту  $T_x$  закончилось выполнение всех сервисов  $S_k$ . Атомарная функция  $x_{ij}^k(t)$  фиксирует моменты начала и завершения использования сервисом  $S_k$  атома  $(i, j)$

$$x_{ij}^k(t) = \begin{cases} 1, & t \in [T_0^k(i, j), T_1^k(i, j)), \\ 0 & \text{иначе,} \end{cases} \quad (1)$$

где  $T_0^k(i, j)$ ,  $T_1^k(i, j)$  моменты начала и завершения использования атома  $(i, j)$  соответственно. Если в сети  $G_k$  существует полный путь  $C$ , состоящий из  $(m - 1)$  дуг  $((i_1, i_2), \dots, (i_{m-1}, i_m))$ , где  $i_1 = s$ ,  $i_m = f$ , то, зная моменты начала и окончания использования сервисом  $S_k$  этих дуг, путь  $C$  можно задать  $(m - 1)$  атомарными функциями (1). По смыслу параметров

$$T_1^k(i, j) - T_0^k(i, j) = T_{\text{exec}}^k(i, j), \quad (2)$$

где  $(i, j)$  – любая дуга сети  $G_k$ ,  $T_{\text{exec}}^k(i, j)$  – продолжительность выполнения атома  $(i, j)$ , она зависит от атома, но не зависит от сервиса. Начатое выполнение сервиса не прерывается. Дуги пути  $C$  задают порядок, это выражается условиями, наложенными на границы интервалов использования

$$T_1^k(i_n, i_{n+1}) \leq T_0^k(i_{n+1}, i_{n+2}), \quad (3)$$

это неравенство справедливо для всех последовательных дуг пути  $C$ , то есть для  $n = 1, 2, \dots, m - 2$ .

Для того, чтобы задать в модели полный пути  $C$  на сети  $G_k$  используется условие доставки по сети продукта в единичном количестве с условием баланса на узлах сети. Для этого достаточно на каждой дуге сети ввести булеву переменную  $u_{ij}^k$ , значение 1 переменной означает, что дуга входит в путь  $C$ , 0 – не входит. Выразим  $u_{ij}^k$  через атомарную функцию

$$u_{ij}^k = \begin{cases} 0, & x_{ij}^k(t) \equiv 0, t \in [0, T_x] \\ 1, & \text{иначе.} \end{cases} \quad (4)$$

Используя переменные  $u_{ij}^k$ , условие баланса записывается как

$$\sum_{\forall (i,n) \in A_k} u_{in}^k - \sum_{\forall (n,j) \in A_k} u_{nj}^k = \begin{cases} -1, & n = s, \\ 1, & n = f, \\ 0 & \text{иначе,} \end{cases} \quad (5)$$

где  $n$  пробегает все узлы сети  $G_k$ .

Функциональные критерии QoS и нефункциональные критерии пользователя могут выражаться в виде ограничений от атомарных функций, или входить в целевой функционал. В общем виде ограничение формулируется так:

$$\mathcal{F}^k(X^k) \in F^k, \quad (6)$$

где  $\mathcal{F}^k$  – функционал на  $X^k$ ;  $X^k$  – вектор атомарных функций  $x_{ij}^k(t)$  на сети  $G_k$ ;



$F^k$  – интервал вещественных чисел, задающий ограничение. Векторные критерии агрегированы в единый целевой функционал  $Q^k$ , определенный на векторе  $X^k$ . Требуется найти оптимальное решение

$$\min Q^k(X^k), \quad (7)$$

при ограничениях (2) – (6).

Модель использования сервисами общего атома – это модель задачи построения расписания в одностадийной системе. При решении этой задачи на общем атоме  $(i, j)$  определяются значения функций  $x_{ij}^k(t)$ , то есть границы интервала  $[T_0^k(i, j), T_1^k(i, j)]$ . Модель задачи композиции пакета сервисов включает в себя ограничения (2) – (6) для каждой из сетей  $G_k$ , а также модель задачи построения расписания в одностадийной системе  $1/r_j, p_j = p, d_j/\Sigma U_j$  и оптимизирует

$$\min \sum_{k=1}^N w_k Q^k(X^k), \quad (8)$$

где  $w_k$  – весовые коэффициенты, отражающие приоритеты пользователя.

В нашем случае целевые функционалы (7) – это функции, линейные по значениям на дугах; соотношения (6) – ограничения по директивному срокам.

Во второй главе строится архитектура брокера композиции пакета сервисов, схема которой представлена на рисунке 1.

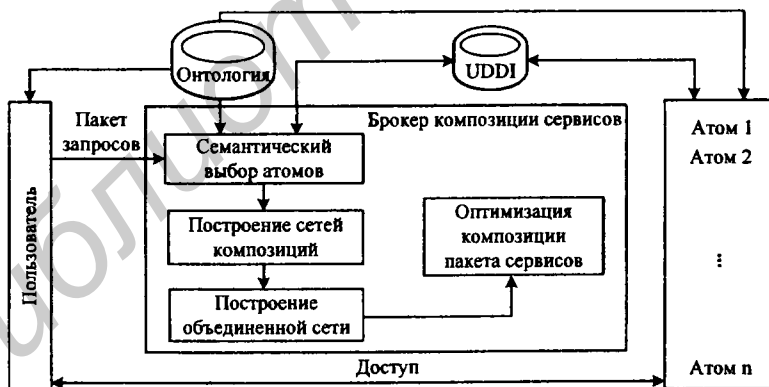


Рисунок 1 – Архитектура брокера композиции пакета сервисов

Блок «Семантический выбор атомов» получает запрос от пользователя на композицию сервисов  $\{S_1, S_2, \dots, S_N\}$ . С помощью базы онтологии и расширенного сервисного реестра UDDI он обрабатывает этот семантический запрос и выбирает атомы для выполнения задач пользователя.

Блок «Построение сетей композиций» для каждого сервиса  $S_k$ ,  $k = 1, \dots, N$  создает сеть  $G_k$ . Если для какого-то сервиса такая сеть уже существует в хранилище, то используется она. После получения набора сетей  $G_k$  для всех сервисов  $S_k$  блок посылает запрос блоку «Построение объединенной сети», который склеиванием общих атомов создает объединенную сеть  $G$ .

Основная задача композиции пакета сервисов возложена на блок «Оптимизация композиции пакета сервисов», он использует разработанные алгоритмы для решения задачи оптимизации композиции пакета сервисов.

В *третьей главе* предложены точный и генетический алгоритмы решения задачи композиции пакета сервисов на объединенной сети. Точный алгоритм представляет собой модификацию алгоритма Беллмана–Форда для построения кратчайшего пути на сети с ограничениями. На рисунках 2 и 3 даны блок-схемы расстановки пометок и обратного хода точного алгоритма.

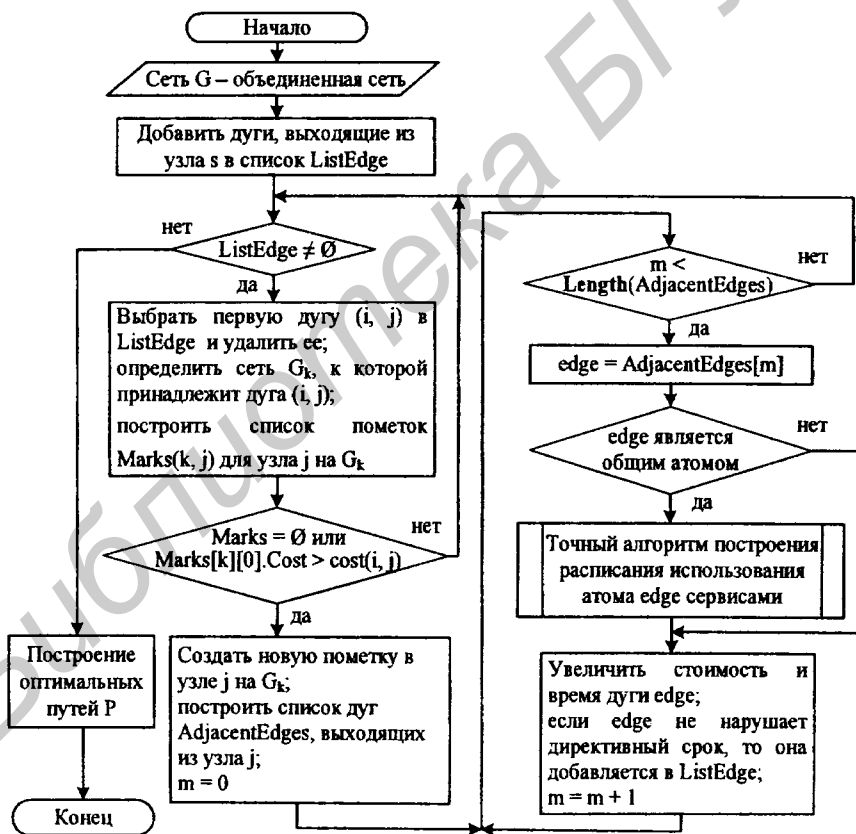


Рисунок 2 – Построение пометок в задаче композиции пакета сервисов

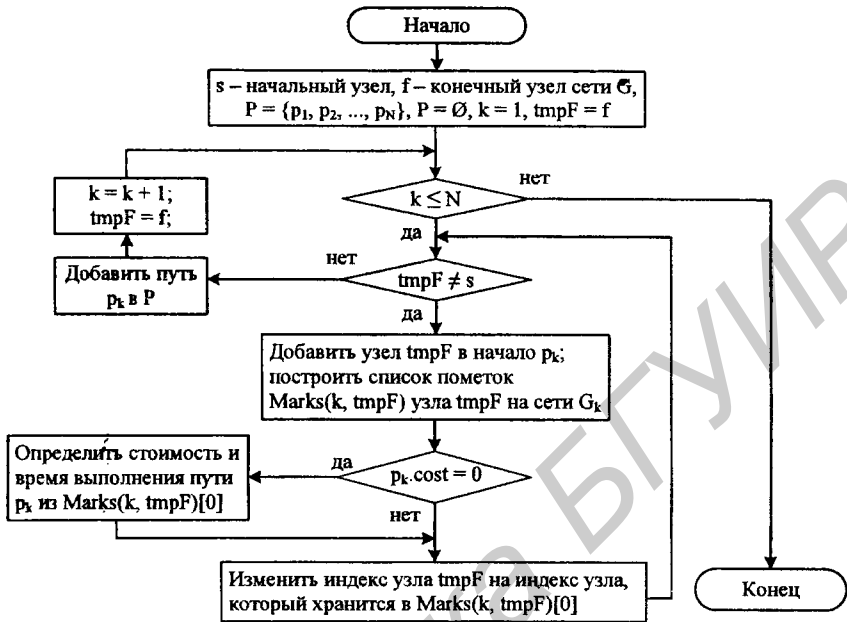


Рисунок 3 – Построение оптимальных путей P

На первом этапе алгоритм, начиная из начального узла  $s$ , последовательно расставляет пометки на узлах и дугах сети  $G$ . При переходе по дуге к следующему узлу, выполняется проверка, представляет ли дуга общий атом. Если это так, то точным алгоритмом строится расписание использования сервисами этого атома. Входными параметрами алгоритма построения расписания являются время поступления  $g_{ij}^k$  сервиса  $S_k$  на общий атом  $(i, j)$  и директивный срок  $d_{ij}^k$  выполнения сервиса  $S_k$  на  $(i, j)$ . Значения  $g_{ij}^k$  получаются из пометок узла  $i$ . Значение  $d_{ij}^k$  представляет наиболее поздний срок наступления события  $j$  для сервиса  $S_k$ . Оно вычисляется алгоритмом Дейкстры, который определяет кратчайший путь от конечного узла сети  $f$  до текущего узла  $j$  сети  $G_k$ . Такая оценка наиболее позднего срока корректна ввиду предположения о том, что на каждом полном пути сети  $G_k$  существует не более одного общего атома.

На втором этапе обратным ходом из конечного узла  $f$  в начальный узел  $s$  по полученным пометкам строятся оптимальные пути, то есть решение задачи композиции сервисов.

Для задач большой размерности, конкретные параметры которых обоснованы в главе 4, предлагается генетический алгоритм.

Начальная популяция – это случайный набор полных путей на сетях  $G_k$  без гарантий выполнения директивных сроков. Пути начальной популяции строятся парами, на сети  $G_k$  случайно выбирается внутренний узел  $i$  и через него строится полный путь. Если возможно, то через этот же узел  $i$  строится второй полный путь. Если это невозможно, то на полученном пути ищется узел, через который строится второй полный путь. Выбрав другой узел на сети  $G_k$ , можно построить еще пару пересекающихся полных путей. Так строится желаемое количество пересекающихся полных путей сети  $G$  для начальной популяции. Опыт показывает, что достаточно построить число путей, равное половине числа узлов сети, хотя это зависит от структуры сети.

В качестве хромосом в задачах о путях на графах обычно выбирают последовательность дуг, входящих в путь. Такое представление выбрано и в предлагаемом алгоритме.

Для операции кроссовера выбираются два пути сети  $G_k$ , имеющие хотя бы один общий узел. Хромосома разрывается по общему узлу и обменом хвостов строятся два новых пути. Если хотя бы один из новых путей будет по своей приспособленности лучше, чем ее родители, то она выбирается как особь нового поколения.

В каждом поколении на каждой сети  $G_k$  с вероятностью  $P$  проводится мутация хромосом. Мутация осуществляется заменой фрагмента хромосомы. Выбирается непрерывный фрагмент пути, фиксируются начальный  $t_1$  и конечный узел  $t_2$  этого фрагмента и строится новый путь, соединяющий узлы  $t_1$  и  $t_2$ , этим создается мутация хромосомы. Мутация позволяет во многих случаях выбраться из области локального оптимума решения. Для проверки критерия останова в каждом поколении из каждой сети  $G_k$  выбирается наилучший путь. Если эти пути имеют общие атомы, то применяется эвристический алгоритм для составления расписаний их использования. Для этих путей вычисляется функция приспособленности. Она сравнивается с аналогичным значением функции приспособленности предыдущего поколения. Из этих двух значений сохраняется наименьшее и соответствующие ему пути.

В качестве функции приспособленности хромосомы для каждой сети  $G_k$  была выбрана взвешенная сумма целевой функции (8) и квадратичного отклонения времени выполнения путей от директивного срока (6). Эксперименты, описанные в главе 4 показали, генетический алгоритм в большинстве случаев дает решение близкое к оптимальному, но для больших размерностей оно и неоптимально и нарушает директивные сроки.

В этой главе также предложены точный и эвристический алгоритм построения расписания использования сервисами общего атома.

Описание точного алгоритма:

Вход. Множество сервисов  $S = \{S_1, S_2, \dots, S_M\}$ , для каждого  $S_i$  заданы

параметры  $r_i$ ,  $d_i$ , параметр  $p$  продолжительности использования общего атома.

**Шаг 1:** Сортировать сервисы по неубыванию директивных сроков, то есть, если  $i < j$ , то  $d_i \leq d_j$ .

**Шаг 2:** Задать множество  $T$  с элементами

$$T = \{r_i + c.p, i \in \{1, \dots, M\}; c \in \{-1, 0, \dots, M\}\}.$$

Задать множество  $V$  с элементами

$$V = \{V^k(a, m), k \in \{0, \dots, M\}; a \in T; m \in \{0, \dots, M\}\}.$$

Для каждого  $a \in T$ ,  $k \in \{0, \dots, M\}$  и  $m \in \{k+1, \dots, M\}$  присвоить

$$V^k(a, m) = \infty.$$

Для каждого  $a \in T$ ,  $k \in \{0, \dots, M\}$  присвоить  $V^k(a, 0) = a + p$ .

**Шаг 3:** Для каждого  $k \in \{1, \dots, M\}$  выполнить:

Для каждого  $a \in T$  выполнить:

Для каждого  $m \in \{1, \dots, k\}$  выполнить:

$$b_{\min} = V_{a, m}^{k-1}.$$

Для каждого  $x \in \{0, 1, \dots, m-1\}$  выполнить:

$$y = m - x - 1; t = \max\{r_k, V_{a, x}^{k-1}\}.$$

Если  $t + p \leq d_k$ , выполнить:

$$b = V_{t, y}^{k-1}; b_{\min} = \min\{b_{\min}, b\}.$$

$$V_{a, m}^k = b_{\min}.$$

**Шаг 4:** Найти максимальное число сервисов из  $1, \dots, M$ , которые выполняются в директивный срок, то есть найти  $m_{\max}$ :

$$a_{\min} = \min_{a \in T} \{a\}; m_{\max} = \max\{m\}, \text{ где } V_{a_{\min}, m}^M \neq \infty$$

и найти минимальное время выполнения этих сервисов:

$$h = V_{a_{\min}, m_{\max}}^M.$$

Пока  $h > 0$ , выполнять:

$$h = h - p.$$

$$\text{Найти } S_j = \arg \max_{S_i \in S} \{r_k, \text{ где } r_k \leq h, h + p \leq d_k\}.$$

$$\text{Положить: } E = E \cup \{S_j\}; S = S \setminus \{S_j\}.$$

**Выход.** Расписание:  $\pi = E \cup S$ .

Предложенный эвристический алгоритм на базе идеи Л. Шраге, он состоит в использовании хвоста  $q$  сервиса, для каждого сервиса  $S_i$  хвост определяется по формуле  $q_i = K - d_i$ , где  $K$  – постоянная, превышающая

директивные сроки всех сервисов. Затем среди сервисов, которые имеют наименьшее время поступления, выбирается сервис с максимальным хвостом. Процедура повторяется до тех пор, пока в расписание не будут включены все сервисы.

Описание эвристического алгоритма:

**Вход.** Множество сервисов  $S = \{S_1, S_2, \dots, S_M\}$ , для каждого  $S_i$  заданы параметры  $r_i, d_i$ , параметр  $p$ .

**Шаг 1:** Определить хвост  $q_i$  для каждого сервиса  $q_i = K - d_i$  и минимум времени поступления сервисов:

$$t = \min_{S_j \in S} (r_j).$$

**Шаг 2:** Определить набор сервисов  $S'$ , которые имеют время поступления, не большее чем  $t$ :

$$S' = \{S_j | S_j \in S \text{ и } r_j \leq t\}.$$

**Шаг 3:** Определить в наборе  $S'$  сервис  $S_j$  с наибольшим хвостом:

$$S_j = \arg \max_{S_k \in S'} \{q_k\}.$$

**Шаг 4:** Добавить сервис  $S_j$  в конец расписания и удалить его из  $S$ :

$$E = E \cup \{S_j\}; S = S \setminus \{S_j\}.$$

**Шаг 5:** Определить время поступления следующего сервиса как наименьшее для сервисов из  $S$ :

$$t = \max\{t + p, \min_{S_m \in S} \{r_m\}\}.$$

**Шаг 6:** Повторить шаги от 2 до 5, пока в расписание не будут включены все сервисы.

**Выход.** Расписание:  $\pi = E$ .

В *четвертой главе* экспериментально исследовано время работы точного и генетического алгоритмов решения задачи композиции пакета сервисов. Результат экспериментов представлен на рисунке 4 и в таблице 1.

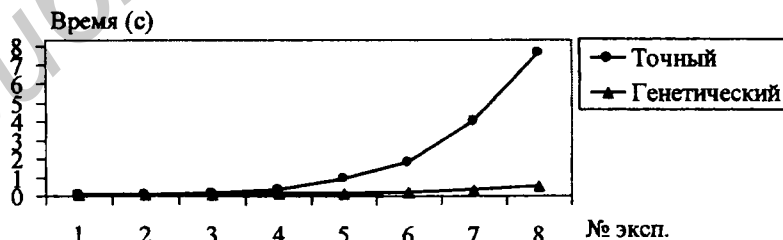


Рисунок 4 – Время работы алгоритмов решения задачи композиции пакета сервисов

Таблица 1 – Параметры экспериментов алгоритмов решения задачи композиции пакета сервисов

№ эксперимента	Число сервисов	Число дуг	Число общих дуг	Число опытов	Точный алгоритм		Генетический алгоритм		
					Время выполнения (мс)	Среднеквадр. отклонение	Время выполнения (мс)	Среднеквадр. отклонение	к-оптимальность решения
1	10–15	25–50	5–10	20	54.3	1.33	46.2	1.73	1
2	15–20	50–80	7–15	20	95.2	2.13	63.7	2.46	1
3	20–25	80–100	9–20	30	159.6	14.56	80.2	9.66	1.3
4	25–30	100–120	12–25	30	319.5	24.32	102.8	10.07	3.3
5	30–35	120–140	15–30	50	913.9	51.17	116.1	9.34	4.8
6	35–40	140–170	17–35	50	1804.5	99.39	162.3	14.88	4.9
7	40–45	170–200	18–40	50	3993.6	193.74	293.3	25.58	6.1
8	45–50	210–250	22–45	50	7584.7	848.81	503.4	53.69	5.7

Результат экспериментов показывает, что точный алгоритм решения задачи композиции пакета сервисов можно применять для решения задачи композиции пакета до 30 сервисов на заданных сетях с числом дуг до 120, имеющих до 25 общих атомов. Для больших размерностей рекомендуется использовать эвристический алгоритм.

Экспериментально проведено сравнение точного и эвристического алгоритмов построения расписания на атоме (рисунок 5, таблица 2).

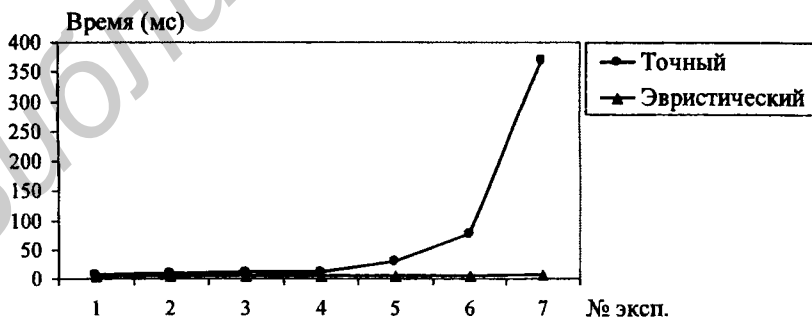


Рисунок 5 – Время работы точного и эвристического алгоритмов построения расписания использования общего атома

Таблица 2 – Параметры экспериментов алгоритмов построения расписания использования общего атома

№ эксперимента	Число сервисов	Число опытов	Точный алгоритм			Эвристический алгоритм			
			Время выполнения (мс)	Среднеквадр. отклонение	Число запаздываний	Время выполнения (мс)	Среднеквадр. отклонение	Число запаздываний	Количество неоптимальных решений
1	2	20	6.6	0.81	0.35	3.5	0.51	0.35	0
2	4	20	7.5	0.47	0.60	3.7	0.47	0.65	1
3	6	30	9.3	0.48	0.63	3.9	0.76	0.73	3
4	8	30	11.3	0.60	1.10	4.1	0.55	1.27	3
5	12	50	28.5	3.94	1.02	4.4	0.57	1.12	5
6	16	50	75.4	6.48	1.70	4.8	0.68	1.88	7
7	22	50	369.5	47.08	2.48	6.7	0.78	2.70	8

Эксперимент показывает, что при числе сервисов до 22 оба алгоритма работают достаточно быстро. При числе сервисов больше чем 22, у эвристического алгоритма скорость работы почти не изменяется, но точный алгоритм работает гораздо медленнее. Вывод по эксперименту: если пользователь сервисов не требует жестких ограничений по времени выполнения, то для задачи расписания на атоме может применяться точный алгоритм.

В *заключении* сформулированы основные научные результаты диссертации и рекомендации по их практическому применению.

*Приложение* содержит акты внедрения результатов диссертации.

## ЗАКЛЮЧЕНИЕ

### Основные научные результаты диссертации

1. Построена математическая модель задачи композиции пакета сервисов. Модель представляет собой комбинацию задачи календарного планирования, транспортной задачи и задачи построения расписания в одностадийной системе. Отличие предлагаемой модели от модели композиции единственного сервиса состоит в том, что сеть, на которой решается задача композиции пакета, содержит дуги-атомы, общие для нескольких сервисов. Для



оптимизации использования общего атома в модель композиции встроена модель задачи построения расписания в одностадийной системе [1, 5, 7].

2. Разработан точный алгоритм решения задачи композиции пакета сервисов. Алгоритм предназначен для решения задачи композиции с сетевой математической моделью. Алгоритм основан на методе Беллмана–Форда для построения кратчайшего пути на сети с ограничениями. Отличие разработанного алгоритма от алгоритма Беллмана–Форда заключается в том, что как подзадача в предлагаемом алгоритме используется построения расписания в одностадийной системе. Предполагается, что на каждом полном пути объединенной сети существует не более одного общего атома. Разработанный точный алгоритм за приемлемое время позволяет решить задачу композиции пакета до 30 сервисов на объединенной сети с числом дуг до 120, и до 25 общих атомов [5, 6, 7, 12].

3. Разработан генетический алгоритм решения задачи композиции пакета сервисов. В качестве хромосомы выбрана последовательность дуг полных путей, представляющих решение задачи. Для составления расписания использования сервисами общего атома в алгоритме применяется предложенный эвристический метод. Разработанный алгоритм тестировался на задачах с количеством сервисов в пакете до 50, количеством дуг объединенной сети до 250 и количеством общих атомов до 45 [2, 4, 7, 10].

4. Предложены точный и эвристический алгоритмы построения расписания использования сервисами общего атома. Это задача построения расписания в одностадийной системе. Точный алгоритм решения задачи основан на модификации метода динамического программирования и использует идею Ж. Карлиера. Метод имеет сложность  $O(n^5)$ . Эвристический алгоритм построения расписания использования сервисами общего атома применяет идею Л. Шраге, его сложность  $O(n \log(n))$ . В экспериментах установлено, что эвристический алгоритм следует применять для задач с числом требований более 50. Если допустимы нарушения директивных сроков, то эвристический алгоритм рекомендуется применять и для задач меньшей размерности [6, 7, 10, 13].

5. Предложены точный и генетический алгоритмы решения задачи композиции сервиса из атомов на единственной сети. Точный алгоритм построен на основе алгоритма Беллмана–Форда. Для генетического алгоритма выбраны способы построения его компонент. Эксперимент показал, что точный алгоритм эффективно решает задачу композиции сервисов из атомов на сети с числом узлов до 60 и числом дуг до 400. Для задач больших размерностей следует применять генетический алгоритм [7, 13, 17].

6. Спланированы и выполнены эксперименты по сравнению точного и генетического алгоритмов решения задач композиции пакета. Спланированы и

выполнены эксперименты по сравнению точного и эвристического алгоритмов решения задач построения расписаний использования общего атома [5, 6].

7. Представлена демонстрационная версия системы на реальной задаче композиции пакета сервисов для медицинских целей [3, 4, 8, 9, 14]. Выбор набора необходимых атомов и создание сетей выполнялось вручную, в эксперименте учитывалось только время решения задачи композиции пакета сервисов. Затраты машинного времени на применение предложенных алгоритмов композиции оказались значительно ниже, чем затраты времени на непосредственное выполнение сервисов, особенно это было заметно при обращении к базам данных, соотношение затрат времени было порядка 1:1000.

### **Рекомендации по практическому использованию результатов**

Результаты исследований использованы для практической реализации прикладной системы в проекте «Разработка удаленной системы анализа изображений» СООО «Элилинк Консалтинг» (Минск). Применение разработанных алгоритмов в системе позволяет построить оптимальный сервис и компоновать их для выполнения сложной задачи в удаленной обработке изображений. По итогам этого проекта получен акт от СООО «Элилинк Консалтинг» о практическом использовании результатов исследования [10, 12, 18].

Была разработана демонстрационная версия системы APACS. Она была инсталлирована в офисе СООО «Элилинк Консалтинг» и в госпитале Министерства транспорта (Ханой, Вьетнам). Использовалась база данных медицинских изображений сервера DICOM этого госпиталя. Были протестированы новые сервисы и новые функции системы. Демо-версия системы показала работоспособность системы APACS с дополнительными функциями [10, 11, 13, 16]. В результате этого применения получены два акта о практическом использовании результатов исследования.

В будущем планируется внедрение системы APACS во вьетнамских учреждениях здравоохранения, в частности, в госпитале Министерства транспорта, военном госпитале № 103, а также применение результатов исследований для создания виртуальной лаборатории, основанной на COA, и разработанного метода композиции веб-сервисов для обучения студентов [14, 15].

# СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ ПО ТЕМЕ ДИССЕРТАЦИИ

## *Статьи в научных журналах*

1. Нгуен, Х.К. Вопросы построения интегрированной системы на основе принципов сервис-ориентированной архитектуры / Х.К. Нгуен // Доклады БГУИР. – 2011. – № 3(57). – С. 63–66.

2. Нгуен, Х.К. Расширение системы PACS архивации и передачи изображений методами архитектуры SOA / Х.К. Нгуен // Нейрокомпьютеры: разработка, применение. – 2013. – № 4. – С. 48–54.

3. Нгуен, Х.К. Архитектура виртуальной обучающей лаборатории, интегрированной с системой PACS / Х.К. Нгуен // Современные проблемы науки и образования. – 2013. – № 4. – С. 1–8.

4. Nguyen, K.Q. Extension of picture archiving and communication system on the base of SOA / K.Q. Nguyen, N. Ivanov // Journal of Information and Communication Technology (JICT). – 2013. – Vol. 3, № 1. – P. 32–42.

5. Нгуен, Х.К. Задача композиции веб сервисов при пакетном запросе / Х.К. Нгуен, Н.Н. Иванов // Современные проблемы науки и образования. – 2013. – № 6. – С. 1–8.

6. Nguyen, K.Q. Web services sharing in batch service composition / K.Q. Nguyen, N. Ivanov // International Journal of Computer Application (IJCA). – 2014. – Vol. 1, № 4. – P. 208–219.

7. Нгуен, Х.К. Алгоритмы построения расписания обработки на элементарном сервисе в задаче пакетной композиции сервисов / Х.К. Нгуен, Н.Н. Иванов // Современные проблемы науки и образования (приложение «Технические науки»). – 2014. – № 6. – С. 10–12.

## *Материалы конференций*

8. Нгуен, Х.К. Основные фазы проектирования системы сервисно-ориентированной архитектуры / Х.К. Нгуен // Содружество наук. Барановичи-2011: материалы VII Международной научно-практической конференции, Барановичи, Беларусь, 19–20 мая 2011 г. – Барановичи, 2011. – С. 144–145.

9. Nguyen, K.Q. Security requirements for SOA on the base of security standards of Web-service / K.Q. Nguyen // Proceedings of the International Conference on Pattern recognition and information processing (PRIP'2011), May 18–20, 2011, Minsk, Belarus. – Minsk, 2011. – P. 280–282.

10. Нгуен, Х.К. Сервисы удаленной обработки биомедицинских изображений / Х.К. Нгуен // Информационные технологии и системы (ИТС)

2011: материалы Международной научной конференции, Минск, Беларусь, 26 ноября 2011 г. – Минск, 2011. – С. 218–219.

11. Нгуен, Х.К. Модель обеспечения безопасности для расширенной системы архивации и передачи изображений на основе СОА / Х.К. Нгуен // Актуальные вопросы современной информатики: материалы Международной заочной научно-практической конференции, Коломна, Россия, 1–15 апреля 2012 г. – Коломна, 2012. – С. 113–116.

12. Нгуен, Х.К. Модуль обмена информации экспертами в расширенной системе архивации и передачи изображения / Х.К. Нгуен // Компьютерные системы и сети: материалы 48-й научной конференции аспирантов, магистрантов и студентов БГУИР, Минск, Беларусь, 7–11 мая 2012 г. – Минск, 2012. – С. 15–16.

13. Нгуен, Х.К. Расширение PACS дополнительными сервисами / Х.К. Нгуен // «Медэлектроника – 2012» – средства медицинской электроники и новые медицинские технологии: материалы VII Международной научной конференции, Минск, Беларусь, 13–14 декабря 2012 г. – Минск, 2012. – С. 88–90.

14. Нгуен, Х.К. Архитектура учебной виртуальной лаборатории как расширение системы PACS / Х.К. Нгуен // Актуальные проблемы экономики, управления, образования и права: материалы VII Международной научно-практической конференции, Коломна, Россия, 30 ноября 2012 г. – Коломна, 2012. – С. 117–123.

15. Нгуен, Х.К. Учебная виртуальная лаборатория на основе новых компонент системы PACS и СОА-подхода / Х.К. Нгуен // Алгоритмические и программные средства в информационных технологиях, радиоэлектронике и телекоммуникациях: материалы I Международной научно-технической конференции, Тольятти, Россия, январь 2013 г. – Тольятти, 2013. – С. 194–200.

16. Нгуен, Х.К. Организация безопасности системы PACS, интегрированной с виртуальной лабораторией / Х.К. Нгуен // Компьютерные системы и сети: материалы 49-й научной конференции аспирантов, магистрантов и студентов БГУИР, Минск, Беларусь, 06–10 мая 2013 г. – Минск, 2013. – С. 38–39.

17. Нгуен, Х.К. Расширение системы PACS функциями коллективного обсуждения / Х.К. Нгуен // Прикладные проблемы оптики, информатики, радиофизики и физики конденсированного состояния: материалы конференции, Минск, Беларусь, 27–28 февраля 2013 г. – Минск, 2013. – С. 10–11.

18. Нгуен, Х.К. Новые компоненты консилуим и консультант в системе архивации и передачи изображений PACS / Х.К. Нгуен // Современные информационные компьютерные технологии mcIT-2013: материалы научно-практической конференции, Гродно, Беларусь, 22–25 апреля 2013 г. – Гродно, 2013. – С. 47–50.



## РЭЗІЮМЭ

Нгуен Хань Куок

Мадэлі і алгарытмы кампазіцыі пакетаў вэб-сэрвісаў

*Ключавыя словы:* вэб-сэрвіс, кампазіцыя пакетаў вэб-сэрвісаў, алгарытмы кампазіцыі пакетаў вэб-сэрвісаў, PACS.

*Мэта працы:* распрацоўка матэматычных мадэляў і алгарытмаў рашэння задачы кампазіцыі пакета вэб-сэрвісаў.

*Метады даследавання:* метады размеркаваных вылічэнняў на кампутарных сетках, метады камбінаторнай аптымізацыі на сетках, генетычны метады аптымізацыі, метады дынамічнага праграмавання, тэорыя графаў, тэорыя раскладаў.

*Атрыманая вынікі і іх навізна:* распрацавана матэматычная мадэль задачы кампазіцыі пакета сэрвісаў. Мадэль уяўляе сабою камбінацыю задачы каляндарнага планавання, транспартнай задачы і задачы пабудовы раскладу ў аднастадыйнай сістэме тэорыі раскладаў. У адпаведнасці з гэтым мадэль аб'ядноўвае тры згаданых кірункі даследавання. Распрацаваны дакладны алгарытм рашэння задачы кампазіцыі пакета сэрвісаў на аб'яднанай сетцы. Алгарытм уяўляе сабою мадыфікацыю алгарытму Беллмана–Форда для рашэння задачы найкарацейшага шляху на сетцы з абмежаваннямі. Адрозненне прапанаванага алгарытму ў тым, што пры кампазіцыі пакета сэрвісаў у алгарытм Беллмана–Форда ўбудаваны алгарытм рашэння задачы пабудовы раскладу ў аднастадыйнай сістэме тэорыі раскладаў, неабходны для разважання калізій выкарыстання сэрвісамі агульнага атама. Распрацаваны генетычны алгарытм, для рашэння задачы кампазіцыі пакета сэрвісаў на аб'яднанай сетцы. Гэты алгарытм дазваляе вырашаць задачы вялікай мернасці, для якіх дакладныя алгарытмы патрабуюць значных выдаткаў часу. Прапанаваны дакладны і з'ўрыстычны алгарытмы пабудовы раскладу выкарыстання сэрвісамі агульнага атама. Дакладны алгарытм вырашае задачу з фіксаванымі абмежаваннямі, ён мае паліномную складанасць. Для з'ўрыстычнага алгарытму абмежаванні саслабляюцца, ён ужываецца або на задачах вялікай мернасці, або ў тым выпадку, калі дапушчальныя парушэнні абмежаванняў задачы.

*Вобласць ужывання:* распрацоўка размеркаваных сістэм, заснаваных на САА і тэхналогіі вэб-сэрвісу; аўтаматызаваныя сістэмы кіравання ў карпарацыях з тэрытарыяльна выдаленымі падраздзяленнямі, турыстычных кампаніях, банках, медыцынскіх і адукацыйных установах.

## РЕЗЮМЕ

Нгуен Хань Куок

Модели и алгоритмы композиции пакетов веб-сервисов

*Ключевые слова:* веб-сервис, композиция пакетов веб-сервисов, алгоритмы композиции пакетов веб-сервисов, PACS.

*Цель работы:* разработка математических моделей и алгоритмов решения задачи композиции пакета веб-сервисов.

*Методы исследования:* методы распределенных вычислений на компьютерных сетях, методы комбинаторной оптимизации на сетях, генетический метод оптимизации, метод динамического программирования, теория графов, теория расписаний.

*Полученные результаты и их новизна:* разработана математическая модель задачи композиции пакета сервисов. Модель представляет собой комбинацию задачи календарного планирования, транспортной задачи и задачи построения расписания в одностадийной системе теории расписаний. В соответствии с этим модель объединяет три упомянутых направления исследования. Разработан точный алгоритм решения задачи композиции пакета сервисов на объединенной сети. Алгоритм представляет собой модификацию алгоритма Беллмана–Форда для решения задачи кратчайшего пути на сети с ограничениями. Отличие предложенного алгоритма в том, что при композиции пакета сервисов в алгоритм Беллмана–Форда встроено алгоритм решения задачи построения расписания в одностадийной системе теории расписаний, необходимый для разрешения коллизий использования сервисами общего атома. Разработан генетический алгоритм, решающий задачу композиции пакета сервисов на объединенной сети. Этот алгоритм позволяет решать задачи большой размерности, для которых точные алгоритмы требуют значительных затрат времени. Предложены точный и эвристический алгоритмы построения расписания использования сервисами общего атома. Точный алгоритм решает задачу с жесткими ограничениями, он имеет полиномиальную сложность. Для эвристического алгоритма ограничения ослабляются, он применяется либо на задачах большой размерности, либо в том случае, когда допустимы нарушения ограничений задачи.

*Область применения:* разработка распределенных систем, основанных на SOA и технологии веб-сервиса; автоматизированные системы управления в корпорациях с территориально удаленными подразделениями, туристических компаниях, банках, медицинских и образовательных учреждениях.

## RESUME

Nguyen Khanh Quoc

Models and algorithms of batch composition for web services

*Keywords:* web service, batch composition of web services, algorithms of batch composition of web services, PACS.

*The purpose of this thesis:* development of mathematical models and algorithms for solving the problem of batch composition of web services.

*Research methods:* methods of distributed calculations on computing networks, methods of combinatorial optimizations on networks, genetic method of optimization, graph theory, scheduling theory.

*The obtained results and their novelty:* a mathematical model of batch composition for web services. The model combines PERT approach, the transportation problem, and the single machine scheduling. In distinction of composition problem for single task, the problem of composition for several tasks is stated here. Accordingly, the model combines three mentioned research domains. An exact algorithm for the problem of batch composition for web services on integrated networks is developed. The algorithm modifies Bellman-Ford method for the solving shortest path problem with restrictions on a network. The distinction the proposed algorithm with Bellman-Ford method is in embedding a single machine scheduling module into the method as the need has arisen in solving collision of tasks on shared element of batch development. Genetic algorithm producing heuristic solution for batch composition of web services is designed. It applies to large size problems when exact methods need enormous running time. Exact and heuristic algorithms are proposed for finding scheduling for tasks on the shared web service. This exact algorithm solves the problem with rigorous restrictions, its computation complexity is polynomial. For heuristic algorithm the restrictions are reduced, it is applied in the case when problem has large size or when some violations of restrictions are allowed.

*Field of application:* distributed computing systems based on SOA and web service technology design, automated management system for travel agencies, banks, corporations with distributed departments, healthcare organizations, educational institutes.

*Научное издание*

**НГУЕН Хань Куок**

**МОДЕЛИ И АЛГОРИТМЫ КОМПОЗИЦИИ ПАКЕТОВ ВЕБ-СЕРВИСОВ**

**Автореферат диссертации на соискание ученой степени  
кандидата технических наук**

**Специальности 05.13.15 – Вычислительные машины, комплексы и  
компьютерные сети**

---

Подписано в печать 13.05.2014.  
Гарнитура «Таймс».  
Уч.-изд. л. 1,4.

Формат 60×84 1/16  
Отпечатано на ризографе.  
Тираж 60 экз.

Бумага офсетная.  
Усл. печ. л. 1,63.  
Заказ 166.

---

Издатель и полиграфическое исполнение: учреждение образования  
«Белорусский государственный университет информатики и радиоэлектроники».  
Свидетельство о государственной регистрации издателя, изготовителя,  
распространителя печатных изданий №1/238 от 24.03.2014, №2/113 от 07.04.2014.  
ЛП № 02330/264 от 14.04.2014.  
2200013, Минск, П. Бровки, 6