

К ПРОБЛЕМЕ СИНТЕЗА РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

¹Зеленков П.В., ¹Каюков Е.В., ²Царев Р.Ю., ²Штарик Е.Н., ²Штарик А.В.

¹ФГБОУ ВПО «Сибирский государственный аэрокосмический университет
имени академика М.Ф. Решетнева», Красноярск, e-mail: zelenkow@rambler.ru;

²ФГАУ ВПО «Сибирский федеральный университет», Красноярск, e-mail: tsarev.sfu@mail.ru

Рассмотрена проблема синтеза распределенных информационно-аналитических систем поддержки принятия решений для обеспечения эффективного поиска и обработки информации. Приведена модифицированная модель формирования структуры распределенных информационно-аналитических систем, обеспечивающая применение узкоспециализированных тезаурусов при контекстном поиске информации в открытых источниках сети интернет и локальных информационных ресурсах. Предложены способы повышения надежности распределенной информационно-аналитической системы поддержки принятия решений путем введения избыточности данных в локальных подсистемах. Для решения задачи повышения эффективности информационного поиска разработан метод учета мультилингвистичности документов, представленных в мировой сети открытых источников информации. При этом учтена возможность использования мультилингвистических запросов при поиске мультилингвистической информации как в одной, так и нескольких поисковых процедурах, выполненных с применением разных языковых подмножеств.

Ключевые слова: система поддержки принятия решений, тезаурус, синтез распределенных систем

TO THE PROBLEM OF SYNTHESIS OF DISTRIBUTED INFORMATION-ANALYTICAL DECISION MAKING SUPPORT SYSTEMS

¹Zelenkov P.V., ¹Kayukov E.V., ²Tsarev R.Y., ²Shtarik E.N., ²Shtarik A.V.

¹Siberian state aerospace university named

after academician M.F. Reshetnev, Krasnoyarsk, e-mail: zelenkow@rambler.ru;

²Siberian federal university, Krasnoyarsk, e-mail: tsarev.sfu@mail.ru

The article considers the problem of synthesis of distributed information-analytical decision making support systems to provide effective search and information processing. The modified model of the distributed information-analytical system structure formation is proposed. This model allows the use of highly specialized thesaurus while context searching in open Internet and local network resources. The ways of improving the reliability of distributed information-analytical decision making support system by introducing data redundancy in local subsystems are proposed. To solve the problem of efficiency increasing of information processing it is developed the accounting method of multiple language documents submitted in the global network of open information sources. It is taken into account the use of multiple language queries when searching for multiple language information, both in one and several search procedures implemented using different language subsets.

Keywords: decision making support system, thesaurus, synthesis of distributed systems

Анализ открытых источников показывает, что наиболее эффективным средством поиска информации в современных информационно-телекоммуникационных технологиях является сеть интернет, в рамках которой существуют поисковые системы. Разработанные методы поддержки принятия решений позволяют снизить трудозатраты при обработке больших массивов данных и значительно повысить скорость принятия решения.

Можно утверждать, что для создания систем, обеспечивающих эффективный поиск и обработку информации, необходимо решить задачу синтеза структуры распределенной информационно-аналитической системы поддержки принятия решений. Для решения поставленной задачи в современных условиях необходимо производить учет мультилингвистичности документов, представленных в мировой сети открытых источников информации [5]. Не менее

важно помнить, что в ходе поиска мультилингвистической информации появляется возможность формирования мультилингвистических запросов как в рамках одной поисковой процедуры, так и рамках нескольких поисковых процедур, выполненных с применением разных языковых подмножеств [6].

Задача синтеза структуры распределенной информационно-аналитической системы поддержки принятия решений

Пусть $S_j = (T_j, D_j, M_{ji}, \delta_j)$, $j = \overline{1, v}$ – локальные информационные системы, предназначенные для обеспечения функционирования сложной информационно-аналитической системы поддержки принятия решений, состоящей из совокупности $K = \{1, \dots, \phi\}$ объектов [3]; i – коэффициент, отвечающий за языковую принадлеж-

ность документа или запроса в рассматриваемой модели; T_j – множество тезаурусов с дескрипторными множествами $T_{ji}^{D_j}$, описывающими состояния объектов (K); D_{ji} – коллекция возможных мультилингвистических документов, которые требуются органам управления для принятия решения; M_{ji} – множество допустимых мультилингвистических запросов со стороны органов управления; $\delta_j : M_{ji} \rightarrow 2^{D_{ji}}$ – отображение, сопоставляющее каждому мультилингвистическому вопросу множество мультилингвистических документов. При этом мультилингвистическая информация с каждой локальной информационной системы поступает в центр сбора информации для дальнейшей передачи ее управляющему органу. Одна локальная информационная система может обслуживать несколько объектов.

Распределенная информационная система $S = (T, D, M, \delta)$ определяется через локальные составляющие:

1. $T = \bigcup_j T_j$; $K_j = K \cap (T_{0j} \times T_{0j})$.
2. $D_i = \bigcup_j D_{ji}$.

$$\delta_j(m_{ji}) = \bigcap_{q=1}^k \bigcup_{l=1}^p \delta_{jq}^l(\bar{m}_{ji}) = \bigcap_{q=1}^k \bigcup_{l=1}^p \{d_{jq} : d_{jq} \in D_{jq}^l \wedge m_{jq}\},$$

где $\bar{m}_{ji} = \{m_{1ji}, \dots, m_{kji}\}$, $\bar{m}_{jq} = \{m_{jq}\}$, p – число объектов, обслуживаемых локальной информационной системой.

Объем передаваемой информации на запрос \bar{m}_{ji} из j -й локальной информационной системы равен

$$\alpha_{\bar{m}_{ji}} = F(\delta_j(\bar{m}_{ji})),$$

где F – оператор преобразования информации к виду, предназначенному для передачи в каналы связи.

Сформулируем задачу распределения объектов по локальным информационным подсистемам при множестве допустимых запросов в распределенной системе $M_i = \{m_{1i}, \dots, m_{ki}\}$, на которые ответы формируются последовательно без повторения запросов. Для удобства положим, что каждый запрос описывается одним дескриптором.

Для формализации этой задачи введем обозначения: c_{jl} – стоимость сбора информации об l -м объекте j -й локальной информационной системой; b_{jl} – стоимость передачи единицы информации об l -м объекте в центр из j -й информационной локальной системы; x_{jl} – булева переменная, равная 1,

$$3. K = \bigcup_j K_j.$$

$$4. \forall m_{ji} \in M_{ji} \quad \delta_j(m_{ji}) = \{d_{ji} : d_{ji} \in D_{ji} \wedge m_{ji}\}.$$

Распределенная информационная система выступает в роли центра сбора и обработки информации, т.е. обработки ответов на запросы пользователя [4].

Описание состояния каждого объекта представляется в виде своей информационной модели:

$$S_j^l = (T_j^l, D_{ji}^l, M_{ji}^l, \delta_j^l),$$

$$l \in K, j \in \{1, \dots, v\}.$$

Индекс j указывает, что информационная модель l -го объекта сформирована в j -й локальной информационной системе.

Информация об объектах представляется независимыми мультилингвистическими информационными моделями и вместе с тем она сосредоточена в одном месте, в одной «точке» [7]. Таким образом, информационные модели объектов являются подсистемами локальной системы, обслуживающей эти объекты.

В случае использования запросов сложного типа, включающих в себя несколько дескрипторов, выражение имеет вид

если l -й объект обслуживается j -й локальной информационной системой, и 0 – в противном случае.

Требуется найти

$$\min_{\{x_{jl}\}} \sum_j \sum_l c_{jl} x_{jl} + \sum_q \sum_j \sum_l b_{jl} F(\delta_j^l(m_{qi})) x_{jl},$$

при ограничениях

$$\sum_j x_{jl} = 1, \quad \forall l \in K; \quad (1)$$

$$\sum_l x_{jl} \geq N_j, \quad \forall j \in \{1, \dots, v\}. \quad (2)$$

Ограничения (1)–(2) требуют обслуживания каждого объекта только одной информационной системой. Условие ограничивает количество объектов, подлежащих обслуживанию локальными подсистемами, либо, напротив, требует, чтобы их было не меньше заданного числа.

Усложним требования к распределенной системе поддержки принятия решений. Потребуем, чтобы органы управления получали ответы на запросы даже в случае, если любая локальная информационная система перестанет функционировать. Иначе говоря, возникает потребность в дублировании информации об объектах [8].

Постановка задачи меняется в части, касающейся ограничения (1). Ограничение примет вид:

$$\sum_j x_{jl} = 2, \forall l \in K.$$

Расширим постановку задачи. Будем считать, что после закрепления объектов за локальными информационными системами в задаче минимизации перераспределения объектов не происходит. Однако органам управления необходимо иметь информацию о предыдущих состояниях объектов даже после прекращения функционирования любой из локальных подсистем. Это означает, что происходит дублирование информационных моделей об объектах постоянно в процессе их функционирования и информация локальных информационных систем перераспределяется между ними.

Для ясности изложения упростим ситуацию, полагая, что перераспределение ин-

$$\min_{\{x_{jl}, y_{jst}\}} \sum_j \sum_l c_{jl} x_{jl} + \sum_q \sum_j \sum_s b_{jl} F(\delta_j^l(m_{qj})) x_{jl} + \sum_j \sum_s \sum_l a_{js} [F(T_j^l, D_{jl}^l, M_{jl}^l, \delta_j^l) x_{jl}] y_{jst}$$

при ограничениях

$$\begin{aligned} \sum_j x_{jl} &= 1, \quad \forall l \in K; \\ \sum_l x_{jl} &\geq N_j, \quad \forall j \in \{1, \dots, v\}; \\ \sum_s y_{jst} &= 1, \\ \forall l \in K_j &= \{l: x_{jl} = 1\}, \\ \forall j \in \{1, \dots, v\}. \end{aligned} \quad (3)$$

Ограничение (3) требует перераспределения информации об l -м объекте из j -й локальной информационной системы.

На основе задачи минимизации можно сформулировать ряд задач, учитывающих те или иные требования относительно структуры распределенной системы. Для этого необходимо сформулировать требования в виде ограничений и ввести их в описание задачи [2].

Рассмотрим, что нового в постановку задач синтеза распределенных систем поддержки принятия решений привносит использование модели системы на тезаурусе с весами. Главное достоинство этой модели состоит в возможности учета распределения информации по дескрипторам, описывающим документ, и тем самым, по сути дела, возможности оценки неопределенности в описании ответов на запросы со стороны органов управления.

В практических приложениях, учитывающих мультилингвистичность представ-

формации осуществляется один раз. Хотя в реальной ситуации информация о состоянии объектов передается по мере ее поступления на всем интервале времени функционирования локальных информационных систем, указанное ограничение не снижает общности рассуждений, так как полная постановка задачи потребует просто дополнительного суммирования по дискретным моментам времени.

Введем следующие обозначения: a_{js} – стоимость передачи единицы информации из j -й локальной информационной системы в s -ю; y_{jst} – булева переменная, равная 1, если информация об l -м объекте перераспределяется из j -й локальной информационной системы в s -ю [1].

Окончательно задача синтеза структуры распределенной информационно-аналитической системы поддержки принятия решений формулируется так:

ления информации в системе, возможны и другие требования к точности получаемых ответов, зависящие от выбранного языка (или множества языков). Как и ранее, необходимо только всегда четко представлять себе целесообразность использования тех или иных моделей в разных языковых множествах и конечный эффект от их мультилингвистической реализации.

Заключение

В статье представлены разработанные модели синтеза структуры распределенной информационно-аналитической системы поддержки принятия решений. Выполнена модификация моделей синтеза структуры распределенных систем поддержки принятия решений, обеспечивающая применение узкоспециализированных тезаурусов при контекстном поиске информации в открытых источниках сети интернет и локальных информационных ресурсах. Применение тезаурусов в поисковых процедурах позволяет существенно повысить качество текстового поиска за счет того, что при поиске исключено точное сравнение терминов поисковой строки, приводящее к расширенной выборке терминов, относящихся к заданной предметной области.

На основе анализа существующих моделей синтеза информационно-аналитических систем поддержки принятия решений предложена модель, использующая элементы теории множеств и обеспечивающая синтез структуры распределенной системы на основе узкоспециализированных темати-

ческих тезаурусов. Одним из преимуществ данной модели является возможность ее применения в двух режимах: с учетом частоты и без использования частотных характеристик термов.

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации, соглашение 16.740.11.0750 «Мультилингвистическая информационно-аналитическая система выявления перспективных направлений научных исследований».

Список литературы

1. Антамошкин А.Н. Определение оптимальной структуры мультиверсионного программного обеспечения при ограничениях по времени и стоимости / А.Н. Антамошкин, И.В. Ковалев // Вестник САА. – 2000. – С. 111–124.
2. Божем Б.У. Инженерное проектирование программно-обеспечения. – М.: Радио и связь, 2005. – 512 с.
3. Ежеманская С.Н. Надежность модульных структур N-вариантных программных систем. – Красноярск: ВСФ РГУИТТ, НИИ СУВИПТ, 2005. – С. 169–174.
4. Завьялова О.И. Минимизация межмодульного интерфейса при формировании мультиверсионного программного обеспечения / О.И. Завьялова, Д.В. Капулин, Р.Ю. Царев // Системы управления и информационные технологии. – 2011. – № 3.1 (45). – С. 140–143.
5. Ковалев И.В. Кроссплатформенная поисковая мультиагентная система / И.В. Ковалев, М.А. Русаков, Р.Ю. Царев // Научно-техническая информация. Сер. 2. Информационные процессы и системы / Всероссийский институт научной и технической информации РАН. – 2010. – № 2. – С. 15–17.
6. Орлов С.А. Технологии разработки программного обеспечения: разработка сложных программных средств. – СПб.: Питер, 2002. – 464 с.
7. Оценка транзакционной надежности современных систем управления и обработки информации / Р.Ю. Царев, А.В. Штарик, Е.Н. Штарик, О.И. Завьялова // Приборы и системы. Управление, контроль, диагностика. – 2012. – № 6. – С. 29–32.
8. Практическая реализация надежностного анализа архитектуры программной системы / Е.В. Гражданцев, О.И. Завьялова, М.А. Русаков, Р.Ю. Царев // Вестник СибГАУ – 2008. – № 1. – С. 37–40.

References

1. Antamoshkin A.N., Kovalev I.V. *Opredelenie optimal'noj struktury mul'tiversionnogo programmnogo obespe-*

chenija pri ogranichenijah po vremeni i stoimosti [Detection of the optimal structure multiversion software under constraints of time and cost]. Bulletin SAA, 2000, pp. 111–124.

2. Bojem B.U. *Inzhenernoe proektirovanie programmnogo obespechenija* [Engineering design of software]. Moscow, Radio and communications, 2005. 512 p.

3. Ezhemanskaja S.N. *Nadezhnost' modul'nyh struktur N-variantnyh programmnyh sistem* [Reliability of modular structures of N- versions software systems]. Krasnojarsk, VSF RGUIPT, NII SUVPT, 2005, pp. 169–174.

4. Zavyalova O. I., Kapulin D.V., Tsarev R.Yu. *Minimizacija mez modul' nogo interfejsa pri formirovanii mul'tiversionnogo programmno go obespechenija* [Minimization of intermodule interface when forming multiversion software]. Control systems and information technologies, 2011, no. 3.1 (45), pp. 140–143.

5. Kovalev I.V., Rusakov M.A., Tsarev R.Yu. *Krossplatformennaja poiskovaja mul'tiagentnaja sistema* [Cross-platform multi-agent search system]. Scientific and technical information. Ser. 2. Information processes and systems. Russian institute of scientific and technical information of Russian Academy of Sciences, 2010, no. 2, pp. 15–17.

6. Orlov S.A. *Tehnologii razrabotki programmno go obespechenija: razrabotka slozhnyh programmnyh sredstv* [Software engineering: the development of complex software]. St. Petersburg, Piter, 2002, 464 p.

7. Tsarev R.Yu., Shtarik A.V., Shtarik E.N., Zavyalova O.I. *Ocenka tranzakcionnoj nadezhnosti sovremennyh sistem upravlenija i obrabotki informacii* [Evaluation of transaction reliability of modern control and information processing systems]. Devices and systems. Control, monitoring, diagnostics, 2012, no. 6, pp. 29–32.

8. Grazhdancev E.V., Zavyalova O.I., Rusakov M.A., Tsarev R.Yu. *Prakticheskaja realizacija nadezhnostnogo analiza arhitektury programmnoj sistemy* [Practical implementation of reliability analysis the architecture of program system]. Krasnojarsk, Bulletin SibGAU, 2008, no. 1, pp. 37–40.

Рецензенты:

Ченцов С.В., д.т.н., профессор, зав. кафедрой «Системы автоматизи, автоматизированного управления и проектирования», зам. директора по учебной работе Института космических и информационных технологий Сибирского федерального университета, г. Красноярск;

Ступина А.А., д.т.н., профессор кафедры «Системный анализ и исследование операций», Сибирский государственный аэрокосмический университет, г. Красноярск.

Работа поступила в редакцию 04.02.2013.