



КОМПЬЮТЕРНЫЙ ЭКСПЕРИМЕНТ ПО ЛОКАЛИЗАЦИИ ДЕФЕКТОВ В ПОДСИСТЕМАХ КИС

*Моженкова Елена Викторовна,
Белорусский государственный университет информатики
и радиоэлектроники, г. Минск, Республика Беларусь,*

E-mail: elena.mohenkova@gmail.com

*Парамонов Антон Иванович,
Белорусский государственный университет информатики
и радиоэлектроники, г. Минск, Республика Беларусь,*

E-mail: anton_paramonov@tut.by

Аннотация. Данная статья посвящена анализу результатов применения подхода обработки банков данных для возможности локализации дефекта бизнес-процесса корпоративной информационной системы на стороне разработчика.

Ключевые слова: локализация дефекта, обработка данных, корпоративная информационная система.

Одной из главных проблем в задачах сопровождения корпоративных информационных систем (КИС) является локализация дефектов в ее информационных потоках и подсистемах. С целью минимизировать затраты отдела сопровождения на обработку банков данных (БНД) предприятия и обеспечения информационной прочности КИС предлагается комплексный подход к решению проблемы локализации дефектов. В основе подхода к обработке банков данных КИС используется формализованное описание дефекта [1] и метод агрегации потоков данных [2]. Предполагается, что применение данного решения позволит повысить эффективность пересылки информации [3] между заказчиком и разработчиком путем сокращения объемов как обрабатываемых, так и передаваемых данных. Для апробации предложенного подхода была проведена серия компьютерных экспериментов в рамках КИС одного предприятия, результаты которых представлены в данной статье.

База данных предприятия состоит из набора таблиц, которые описывают сущности КИС в динамике изменений их свойств за весь период работы системы. Для локализации различного рода дефектов предлагается использовать шаблоны вида [дефект-сущности]. Предварительно для каждого класса дефектов устанавливается свой набор сущностей. Кроме того, среди набора таблиц заданных сущностей задается «менеджер таблиц» – таблица, которая является

первичной в обработке метаданных. «Менеджер таблиц» считается приоритетным источником дефекта. На основе связей этой таблицы формируется выборка слепка данных. Объем выборки слепка данных может быть уменьшен за счет установки фильтра в разрезе записей менеджера таблиц. АИС получает список таблиц и связей между ними согласно указанному шаблону. В случае если пользователем установлены дополнительные временные характеристики, то АИС при выгрузке данных также установит временные рамки существования данных. В зависимости от типа обрабатываемых сущностей к слепку данных могут быть применимы фильтры значений полей менеджера таблиц и временных характеристик. Таким образом, можно определить модель дефекта как кортеж записей данных из набора взаимосвязанных таблиц сущностей за актуальный временной период.

С целью апробации подхода обработки банка данных предлагается рассмотреть и проанализировать результаты обработки двух подсистем: аутентификации пользователей и расчета заработной платы сотрудников КИС.

Локализация дефекта в подсистеме аутентификации пользователей КИС. Общий объем базы данных КИС – 41472 МБ. Число пользователей – 33. КИС внедрена на предприятии в 2016 году. Эксперимент проводился 25.01.2021.

Для локализации дефекта не требуется обрабатывать весь объем данных. Согласно схеме классификации проблем для локализации дефекта, класс аутентификации реализуется двумя сущностями – «Пользователи» и «Роли» (рисунок 1). Сущность «Пользователи» представлена таблицей User (содержит сведения о зарегистрированных в системе пользователях). Сущность «Роли» представлена таблицами Roles (содержит информацию о ролях, которые можно назначить пользователю) и Roles_User (содержит информацию о соответствии назначенных ролей пользователям).

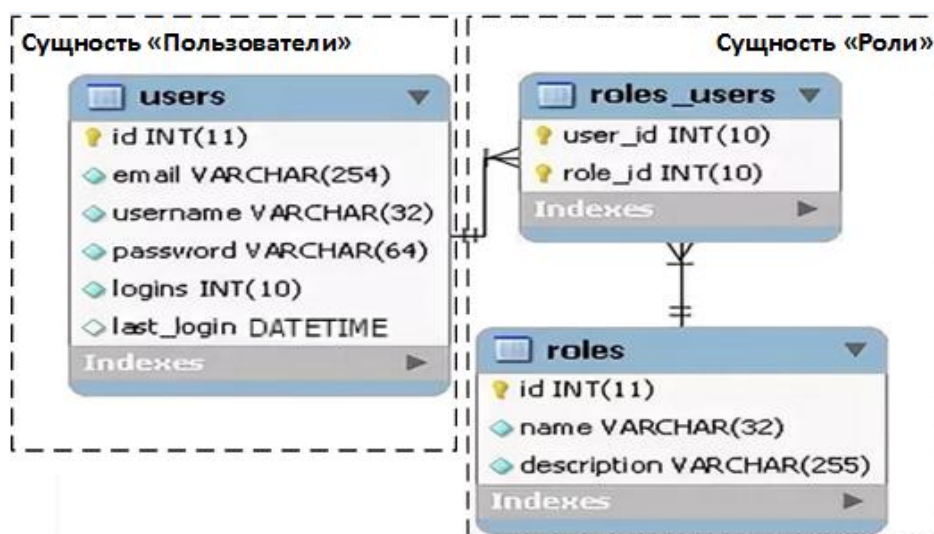


Рис. 1 Схема сущностей «Пользователи» и «Роли» БД КИС

SCIENCE TIME

Общий объем обрабатываемой БД составляет $V^* = 41472$ МБ.

Рассмотренные сущности будут представлены как:

$$S_1 = \langle \{users\}, t_1, k_1^d, r_1^d \rangle,$$

$$S_2 = \langle \{roles, roles_users\}, t_2, k_2^d, r_2^d \rangle.$$

Рассчитаем объем обрабатываемой информации на базе характеристик таблиц подсистемы аутентификации КИС V_A (таблица 1).

Таблица 1

Характеристики таблиц подсистемы аутентификации КИС

Наименование таблицы	Размер записи (Б)	Количество записей в таблице (k_i)	Количество связей (r^i)	Объем (Б)	Объем (МБ)
Users	243	33	1	8019	0,00771523
Roles	44	64	1	2816	0,00268555
roles_users	104	112	2	11648	0,011108
Итого				22554	0,021509

$$V_A = V_{S1} + V_{S2} = 0,00771523 + (0,00268555 + 0,011108) = 0,021509 \text{ (МБ)}$$

Проблема произошла в определенный временной интервал (за последние сутки, т.е. 25.01.2021). К слепку данных, которые могли повлиять на возникновение дефекта, можно применить временной интервал к полю «Дата последнего подключения» – *users.last_login*.

$$S_1 = \langle \{users\}, "25.02.2020", k_1^d, r_1^d \rangle,$$

$$S_2 = \langle \{roles, roles_users\}, t_2, k_2^d, r_2^d \rangle.$$

С учетом временного интервала объем обрабатываемой информации V_{At} сократится на количество записей не попавших во временной интервал (известно, что за последние сутки с КИС работало 10 пользователей):

$$V_{At} = 0,00231743 + (0,00268555 + 0,011108) = 0,01611098 \text{ (МБ)}$$

Объем обрабатываемой информации уменьшился на 0,005398 (МБ) или 1,33 раза. Если клиенту известно имя пользователя, по которому возникла проблемная ситуация, можно применить фильтр менеджера таблиц. В качестве входных данных была сгенерирована ошибка по Иванову И.И. Зная имя пользователя для данной схемы БД можно ограничить слепок данных, установив фильтр для поля *users.username*:

$$Filter = \langle "users", ("username", "Иванов И.И.") \rangle$$

Зная, что фильтр применим к таблице *users*, дополним данными из таблицы 1.

$$S_1 = \langle \{users\}, "07.01.2020", 1, 1 \rangle,$$

$$S_2 = \langle \{roles, roles_users\}, t_2, 176, 3 \rangle.$$

Рассчитаем объем информации.

$$V_{AtFiltr} = 0,00023174 + (0,00268555 + 0,011108) = 0,01402529 \text{ (МБ)}$$

В таблице 2 представлен анализ объемов данных КИС.

Таблица 2

Анализ объемов данных КИС

Показатель	Объем (МБ)
V^*	41472
V_A	0,021509
V_{At}	0,01611098
$V_{AtFiltr}$	0,01402529

Применение фильтра позволило сократить объем обрабатываемой информации в 1,5 раза или на 0,007484 (МБ).

Локализация дефекта в подсистеме расчета заработной платы сотрудников КИС. Общий объем базы данных КИС – 50380,8 МБ. Число сотрудников – 7310. КИС внедрена на предприятии в 2009 году. Эксперимент проводился 30.01.2021.

Так как бизнес-процесс расчета заработной платы сложный, в схему классификации для данного класса проблем предлагается выделить сущность «Лицевой счет» и класс «Дополнительные сущности» (рисунок 2).

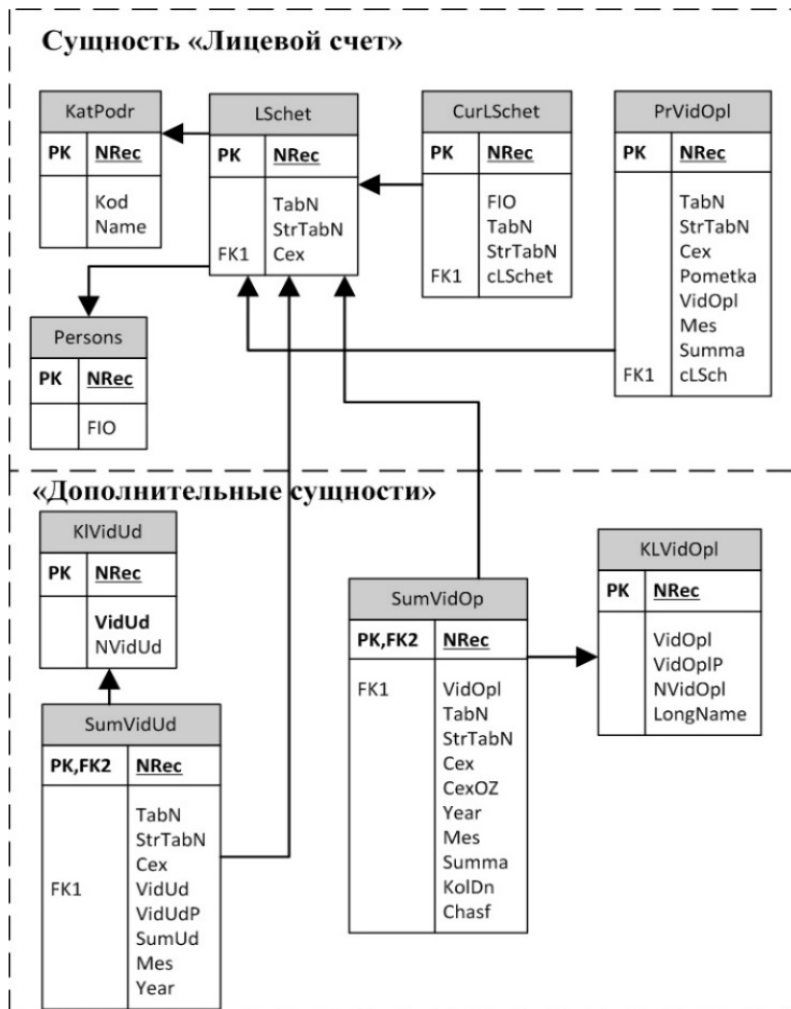


Рис. 2 Схема сущностей подсистемы расчета заработной платы

Сущность «Лицевой счет» представлена таблицами: LSchet (сведения о лицевых счетах сотрудников), KatPodr (подразделения предприятия), Persons (физические лица), CurLSchet (работники из текущего подразделения), PrVidOpl (межрасчетные выплаты). Класс «Дополнительные сущности» представлен таблицами: KlvidUd (классификатор видов удержаний), SumVidUd (начисления по видам удержаний), KlVidOpl (классификатор видов оплат), SumVidOpl (начисления).

Общий объем обрабатываемой БД составляет $V^* = 50380,8$ МБ. Рассмотренные сущности будут представлены как:

$$S_1 = \langle \{Lschet, KatPodr, CurLschet, PrVidOpl\}, t_1, k_1^d, r_1^d \rangle,$$

$$S_2 = \langle \{KlvidUd, SumVidUd, KlVidOpl, SumVidOpl\}, t_2, k_2^d, r_2^d \rangle.$$

Рассчитаем объем обрабатываемой информации на базе характеристик таблиц подсистемы расчета заработной платы V_A (таблица 3).

Таблица 3

Характеристики таблиц подсистемы расчета заработной платы

Наименование таблицы	Размер записи (Б)	Количество записей в таблице (k_i)	Количество связей (r^i)	Объем (Б)	Объем (МБ)
Lschet	1871	7310	2	13677010	13,0434
KatPodr	992	3567	1	3538464	3,3745
CurLschet	205	7310	1	1498550	1,4291
PrVidOpl	897	405063	1	363341511	346,51
KlvidUd	789	116	1	91524	0,087284
SumVidUd	784	3319814	2	2602734176	2482,16
KlVidOpl	1026	373	1	382698	0,364969
SumVidOpl	769	3452368	2	2654870992	2531,88
Итого				5640134925	5378,849253

$$V_A = V_{S1} + V_{S2} = 364,357 + 5014,492253 = 5378,849253 \text{ (МБ)}$$

Проблема произошла в определенный временной интервал (за последние сутки, т.е. 30.01.2021). К слепку данных, которые могли повлиять на возникновение дефекта, можно применить временной интервал к полям «Месяц», «Год» расчета – $SumVidUd.Mes$, $SumVidUd.Year$, $SumVidOpl.Mes$, $SumVidOpl.Year$.

$$S_1 = \langle \{Lschet, KatPodr, CurLschet, PrVidOpl\}, t_1, k_1^d, r_1^d \rangle,$$

$$S_2 = \langle \{KlvidUd, SumVidUd, KlVidOpl, SumVidOpl\},$$

$$\{Year = "2021", Mes = "1" \}, k_2^d, r_2^d \rangle.$$

С учетом временного интервала объем обрабатываемой информации V_{At} сократится на количество записей не попавших во временной интервал (сокращается количество записей таблиц SumVidOpl, SumVidUd – $V_{SumVidOpl} = 75,0537$, $V_{SumVidUd} = 76,5176$):

$$V_{At} = 364,357 + 152,023553 = 516,380553 \text{ (МБ)}$$

Объем обрабатываемой информации уменьшился на 4862,4687 (МБ) или 9,6 раз. Если клиенту известно табельный номер сотрудника, по которому возникла проблемная ситуация, можно применить фильтр менеджера таблиц. В качестве входных данных была сгенерирована ошибка по табельному номеру 561039. Зная табельный номер сотрудника для данной схемы БД можно ограничить слепок данных, установив фильтр для поля *Lschet.TabN*:

$$Filter = \langle "Lschet", ("TabN", "561039") \rangle.$$

Зная, что фильтр применим к таблице *Lschet* и связанным таблицам, дополним данными из таблицы 2.

$$S_1 = \langle \{Lschet, KatPodr, CurLschet, PrVidOpl\}, t_1, 7, 4 \rangle,$$

$$S_2 = \langle \{KlvidUd, SumVidUd, KlvidOpl, SumVidOpl\},$$

$$\{Year="2021", Mes="1"\}, 205169, 4 \rangle.$$

Рассчитаем объем информации.

$$V_{AtFiltr} = 0,00634764 + 152,023553 = 152,0299006 \text{ (МБ)}$$

В таблице 4 представлен анализ объемов данных КИС.

Таблица 4

Анализ объемов данных КИС

Показатель	Объем (МБ)
V^*	50380,8
V_A	5378,849253
V_{At}	516,380553
$V_{AtFiltr}$	152,0299006

Применение фильтра позволило сократить объем обрабатываемой информации в 35 раз или на 5226,819352 (МБ).

Результаты проведенных экспериментов показали значительный эффект от применения предложенного подхода. Оценивая результаты уменьшения объема обрабатываемой и передаваемой информации можно сделать предположение, что чем больший объем обрабатываемых данных, тем больший эффект достигается. Особенно если обращения к базе будут частыми и однотипными. В ходе дальнейших экспериментов оцениваются возможности подхода при обработке банков данных с разнородной информацией.

Литература:

1. Моженкова Е.В. Модель представления информации в корпоративных информационных системах / Е.В. Моженкова, А.И. Парамонов // Информационные технологии и системы 2018 (ИТС 2018) = Information Technologies and Systems 2018 (ITS 2018): материалы международной научной конференции, Минск, 25 октября 2018 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол.: Л.Ю. Шилин [и др.]. – Минск, 2018. – С. 228-229.
2. Моженкова Е.В. Метод агрегации потоков данных в задачах локализации проблем / Моженкова Е. В., Парамонов А. И. // Информационные технологии и системы 2019 (ИТС 2019) = Information Technologies and Systems 2019 (ITS 2019): материалы международной научной конференции, Минск, 30 октября 2019 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол.: Л.Ю. Шилин [и др.]. – Минск, 2019. – С. 242-243.
3. Моженкова Е.В. Схема генерации и восстановления данных при локализации дефекта / Е.В. Моженкова, А.И. Парамонов // Компьютерные системы и сети: 55-я юбилейная научная конференция аспирантов, магистрантов и студентов, Минск, 22-26 апреля 2019 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2019. – С. 132-133.