УДК 004.89

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ БАЗ ЭКСПЕРТНЫХ ЗНАНИЙ, ПОЛУЧЕННЫХ ПРИ РАЗРЕШЕНИИ ИНЦИДЕНТОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

Карасев А.А., ведущий инженер, Дирекция ИТ НИУ ВШЭ, E-mail: akarasev@hse.ru
Старых В.А., к.т.н., профессор, МИЭМ НИУ ВШЭ, E-mail: vstarykh@hse.ru
Москва, Россия

Аннотация. Статья посвящена рассмотрению вопросов сбора и классификации знаний, полученных в процессе эксплуатации информационных систем, с целью их дальнейшей организации в базы экспертных знаний. В качестве перспективного подхода предлагается использование онтологической теории, обеспечивающей базис для построения математической модели представления знаний; организация и построение иерархии объектов осуществляется с использованием классификаторов, основанных на структурном подходе и модели OSI/ISO. В статье описывается программная реализация предложенных методов; приводятся результаты апробации полученных инструментальных средств.
Ключевые слова: приобретение знаний; инцидент; базы экспертных знаний; онтология; классификация; структурный подход; модель OSI/ISO.

SOFTWARE AND MATHEMATICAL SUPPORT OF ORGANIZING EXPERT KNOWLEDGE ACQUIRED DURING INCIDENT RESOLUTION IN INFORMATION SYSTEMS

Karasev A.A., senior engineer, HSE, E-mail: akarasev@hse.ru
Starykh V.A., Ph.D., Associate Prof., HSE, E-mail: vstarykh@hse.ru
Moscow, Russia

Abstract. This article is devoted to a solution of the problem of collecting and classification of the expert knowledge acquired during the operation of information systems aimed to organize expert knowledge bases. As perspective approach ontological theory use is offered; it provides a basis for building mathematical model of knowledge representation. Organization of objects hierarchy is carried out with use of classification based on structural approach and the OSI/ISO reference model. The article also outlines the basic principles of software implementation of offered methods and results of its testing.
Keywords: knowledge acquisition; incident; expert knowledge base; ontology; classification; structural approach; OSI/ISO reference model.

В процессе эксплуатации современных информационных систем (ИС) используются как нормативные, так и экспертные знания об особенностях функционирования конкретной системы, накопленные в процессе эксплуатации и являющиеся для нее уникальными. Нормативные знания описывают штатный режим работы ее компонентов на основании технической и эксплуатационной документации; при этом их использование в большинстве случаев не позволяет учесть различные аспекты взаимодействия конкретных аппаратных и программных компонентов в

составе действующей ИС. Экспертные знания с этой точки зрения обладают большей ценностью, однако из-за отсутствия эффективных программных средств сбора, обработки и накопления возможность их дальнейшего использования ограничена.

Таким образом, эффективность функционирования ИС (определяемая, в том числе, количеством возникающих инцидентов[1] и временем, затрачиваемым на их обнаружение и разрешение) во многом определяется возможностью наполнения баз экспертных знаний, позволяющих повторно использовать знания, полученные в ходе обработки ранее возникавших инцидентов, для сокращения времени разрешения аналогичных и однотипных инцидентов в дальнейшем. В то же время отсутствие таких баз приводит к возникновению целого ряда факторов, отрицательно сказывающихся на функционировании ИС, в том числе:

- принятие решений по обеспечению функционирования ИС на основе предположений, а не ранее зафиксированных фактов, вызванное недостатком управляющей информации;
- необходимость повторного поиска способа разрешения одинаковых инцидентов вместо использования стандартной, ранее задокументированной и внесенной в базу знаний последовательности действий;
- потеря значительного объема незафиксированных экспертных знаний в случае увольнения ключевых сотрудников.

Для выявления основных методов сбора и последующего использования экспертных знаний в процессе функционирования ИС был проведен сравнительный анализ ряда программных средств[2]. Анализ проводился по таким критериям как источники наполнения баз данных, наличие предустановленных баз знаний, возможности создания и поддержания в актуальном состоянии связей между записями об инцидентах, а также наличие встроенных классификаторов. В результате был выявлен ряд ограничений и недостатков существующих программных средств и применяемых в них методов, в том числе:

- Недостаточная формализация. Собранные знания хранятся в виде самостоятельных, не связанных друг с другом информационных статей, содержащих слабоструктурированную информацию, что не позволяет эффективно использовать информационные технологии для ее поиска и обработки.

[1] Событий, оказывающих отрицательное воздействие на функционирование ИС, и могущих привести к ухудшению рабочих параметров отдельных ее подсистем или качества предоставляемых пользователям услуг [1]

[2] Рассматривались модули организации и наполнения баз экспертных знаний в составе ПО IBM Tivoli Service Request Manager, Microsoft System Center Service Manager, HP Service Manager, BMC Remedy IT Service Management Suite, OTRS ITSM

- Отсутствие встроенных средств классификации. В большинстве программных средств отсутствуют классификаторы, в соответствии с которыми могли бы распределяться новые записи об инцидентах. В результате этого усложняется их первоначальная привязка и определение причин возникновения; повышается вероятность появления записей с однотипным содержанием.

- Отсутствие связей. В рассмотренных программных средствах не реализованы инструменты создания, и редактирования связей между инцидентами и объектами в составе ИС. Это приводит к отсутствию у пользователя возможности видеть причинно-следственные связи между событиями, приводящими к нарушению штатного режима функционирования ИС.

Выявленные недостатки и ограничения не позволяют рассматривать существующие программные средства в качестве инструментов организации баз экспертных знаний об инцидентах в ИС.

Для устранения перечисленных ограничений в данной работе используется онтологический подход [2]. Термин «онтология» (от др. греч. онтос – сущее, логос – учение, наука) в информационных технологиях и интеллектуальных информационных системах обозначает систему понятий заданной предметной области, представленной в виде набора существенных понятий, отношений, возможных между ними, и их значимых свойств. Онтологии используются для формального определения сущностей, их отношений и свойств, формирующих концептуальную модель предметной области.

В состав онтологий входит словарь (тезаурус) имен сущностей для представления и обмена знаниями в рассматриваемой предметной области и набор отношений, которые можно выделить между сущностями [3]; отношения могут обозначаться как универсальными связями, например, «часть – целое», «причина – следствие», так и специфическими для данной области. Сущности, как и связи, могут обладать различными свойствами, необходимыми для отражения и актуализации свойств предметной области в ее модели.

В соответствии с определением онтологии экспертные знания, полученные при обработке инцидентов в ИС, в данной работе организуются в виде таксономии – иерархической системы понятий, связанных друг с другом отношениями с определенной семантикой, что позволяет структурно организовывать сущности в составе онтологии в виде графа.

Основной задачей при проектировании онтологии является разработка принципов классификации и структурирования знаний. В данной работе применяются два независимых классификатора объектов предметной области обработки инцидентов в

ИС [4]. Первый основан на структурном подходе, позволяющем учитывать универсальные иерархические связи между объектами. Второй использует эталонную модель взаимодействия открытых систем ISO/OSI для определения специфичных для предметной области связей между объектами и их свойств.

Структурный подход базируется на утверждении о том, что любая ИС может рассматриваться как совокупность своих подсистем. Подсистемы могут относиться к общим – формирующим основу ИС, прикладным – предназначенным для решения задач в определенной области и обеспечивающим – позволяющим организовать надежную работу прикладных.

Далее в каждой подсистеме выделяется множество программно-аппаратных компонентов, из которых она состоит. Такими компонентами являются физическое или виртуальное оборудование, программное обеспечение. Следующим этапом декомпозиции является выделение в составе каждого аппаратно-программного компонента объектов, являющихся атомарными с точки зрения установления взаимосвязей с инцидентами. Атомарный объект либо является причиной возникновения инцидентов, либо оказывается затронутым инцидентами, вызванными другими объектами. При необходимости объекты в составе аппаратно-программного комплекса могут объединяться в группы по тем или иным признакам.

На рисунке 1 представлена схема, иллюстрирующая алгоритм привязки инцидента к атомарным объектам в соответствии с рассмотренным классификатором.

Эталонная модель взаимодействия открытых систем ISO/OSI (Open Systems Interconnection) позволяет регламентировать процесс совместной работы ИС, отличающихся по составу используемого аппаратного и программного обеспечения. Модель последовательно описывает правила и процедуры организации взаимодействия на семи уровнях, начиная с оборудования и физических средств связи и заканчивая уровнем приложений. Назначение каждого уровня стандартизируется, определяется набор протоколов, функционирующих в его рамках, описываются интерфейсы, использующиеся для передачи данных от одного уровня к другому.

Распределение инцидентов в соответствии с предложенным классификатором может выполняться в зависимости как от того, к какому уровню модели относится объект в составе ИС, связанный с ними[3], так и от того, на каком уровне было нарушено взаимодействие в результате возникновения инцидента.

[3]Штатное функционирование которого было нарушено инцидентом, либо само привело к возникновению инцидента

Рисунок 1. **Схема алгоритма привязки инцидента к атомарным объектам в соответствии с классификатором, основанном на структурном подходе**

В отличие от классификатора, основанного на структурном подходе, распределение и привязка инцидентов в соответствии с которым осуществляется только на основании затронутых атомарных объектов, данный классификатор не использует связи объектов с инцидентами как основные.

Предложенные классификаторы позволили разработать математическую модель организации базы экспертных знаний об инцидентах в ИС на основе используемого онтологического подхода [5]. В разработанной модели выделяется три уровня (рисунок 2); к первому – объектному – уровню, обозначаемому $L_{об}$, относится совокупность множеств понятий-сущностей $X_{об}$, соответствующих таким объектам предметной области, как подсистемы ИС (почтовая система, виртуальная инфраструктура, серверное оборудование), их аппаратно-программные компоненты, атомарные объекты

и их группы, а также множеств понятий-отношений $R_{об}$, описывающих иерархические связи между ними.

Распределение понятий-сущностей и отношений на объектном уровне $L_{об}$ осуществляется с использованием классификатора, основанного на использовании структурного подхода.

К третьему уровню – уровню инцидентов, обозначаемому $L_и$, – относятся множества зафиксированных в процессе эксплуатации ИС инцидентов $X_и$ и причинно-следственных связей между ними $R_{пс}$, специфических для предметной области обработки инцидентов в ИС:

$$L_и = X_и \cup R_{пс}, \text{ где } R_{пс} = R(X_и, X_и)$$

Объекты на уровне инцидентов распределяются в соответствии с классификатором, основанном на модели взаимодействия открытых систем ISO/OSI.

Особенностью уровня инцидентов $L_и$ в построенной модели является наличие свойств у объектов множеств инцидентов $X_и$ и причинно-следственных связей между ними $R_{пс}$. Данные множества обозначаются, соответственно, как $P(X_и)$ и $P(R(X_и, X_и))$ и расширяют определение уровня инцидентов $L_и$, описание которого принимает следующий вид: $L_и = X_и \cup P(X_и) \cup R(X_и, X_и) \cup P(R(X_и, X_и))$

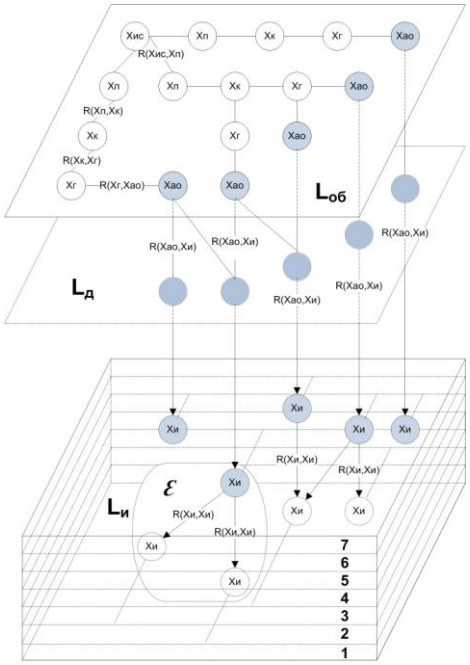

Рисунок 2. **Трехуровневая модель организации базы экспертных знаний об инцидентах в ИС**

Уровень действий является промежуточным по отношению к уровням инцидентов и объектов и содержит множество действий, выполняемых над объектами в составе ИС для разрешения связанных с ними инцидентов. При этом действие или

последовательность действий, определенная и выполненная при первоначальном разрешении инцидента и зафиксированная в базе знаний для последующего использования, определяется благодаря причинно-следственным связям между инцидентами и атомарными объектами $R(X_{ao}, X_{и})$. Данные связи позволяют не только определить возможные способы разрешения инцидента, но и идентифицировать атомарные объекты в составе ИС, ставшие причиной возникновения инцидента или инциденты, повлиявшие на их функционирование.

Предложенная математическая модель позволяет описывать все доступное множество объектов предметной области обработки инцидентов в ИС; содержит все необходимые знания в рамках базы экспертных знаний.

Общее количество объектов разных классов в онтологии обработки инцидентов в ИС, основанной на предложенной математической модели, приведено в таблице 1.

Таблица 1. **Количественные показатели для объектов в онтологии обработки инцидентов в ИС**

Класс объекта	Количество экземпляров класса
Понятия-сущности	
Подсистема ИС	10
Аппаратно-программный компонент	31
Группа объектов	60
Атомарный объект	109
Понятия-отношения	
Аппаратно-программные компоненты в составе подсистемы ИС	42
Группы объектов в составе аппаратно-программного компонента	75
Объекты в составе группы	115

В качестве инструмента программной реализации в данной работе используется комплект средств разработки Thinkmap SDK [6]. Он обеспечивает поддержку ряда современных технологий (Java EE, HTML5, XML) и предназначен для создания приложений, решающих задачу организации и визуализации больших объемов данных.

Апробация разработанной системы приобретения и представления знаний для организации и наполнения базы экспертных знаний осуществлялась в процессе эксплуатации аппаратно-программного комплекса Федерального центра информационно-образовательных ресурсов. ИС ФЦИОР представляет собой сложный гетерогенный аппаратно-программный комплекс, включающий в себя уровни аппаратного обеспечения, виртуализации, операционных и прикладных систем. На каждом уровне используется оборудование и программное обеспечение различных

производителей, сконфигурированное и настроенное в соответствии со спецификой решаемых задач.

Разработанная система использовалась для организации и наполнения базы экспертных знаний, полученных при обработке инцидентов, возникающих в процессе эксплуатации ИС ФЦИОР. За время опытной эксплуатации средствами системы зафиксированы и внесены в базу знаний описания 98 инцидентов, 128 связей типа «инцидент – атомарный объект», 38 связей типа «инцидент – инцидент».

Для оценки параметров функционирования ИС до и после внедрения разработанной программной системы использовались следующие метрики:

- процент инцидентов, разрешенных на первой линии поддержки;
- процент инцидентов, разрешенных с первого раза;
- процент инцидентов, правильно связанных с объектами с первого раза;
- число инцидентов, разрешенных с использованием известных решений;
- процент инцидентов, не связанных с известными проблемами.

На рисунке 3 представлена диаграмма, позволяющая сравнить зафиксированные значения метрик для соответствующих периодов наблюдения.

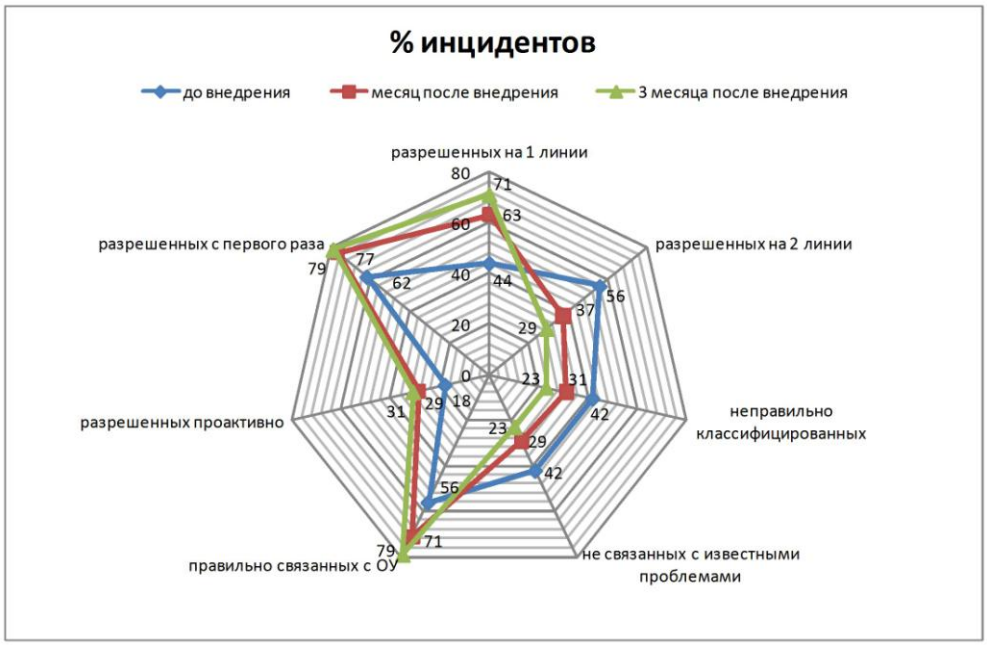

Рисунок 3. **Значения выбранных метрик до и после внедрения разработанной системы**

Значения метрик, определяющих такие факторы при обработке инцидентов в ИС, как правильность и скорость их разрешения с использованием базы экспертных знаний, значительно увеличиваются в первый месяц после внедрения системы; диаграмма при этом смещается влево. Дальнейший рост оцениваемых показателей (при относительно

постоянном количестве возникающих в ИС инцидентов) обусловлен постепенным наполнением базы знаний.

Заключение. Применение онтологической теории, разработанных моделей и методов формализации и классификации экспертных знаний об инцидентах в ИС позволило учесть недостатки существующих программных средств и реализовать соответствующие функции, в том числе классификацию и разрешение инцидентов с учетом выявленных причинно-следственных связей, возможность получать информацию о причинах, структуре и уровне их возникновения в соответствии с эталонной моделью OSI/ISO.

Разработанная система приобретения и представления знаний предназначена для организации баз экспертных знаний, полученных в процессе эксплуатации ИС. Подобные базы знаний могут быть как централизованными, так и распределенными; использоваться в рамках одной организации или предоставлять интерфейсы для совместной работы с внешними пользователями, объединяясь в открытые базы экспертных знаний предметной области обработки инцидентов в ИС.

Список литературы

1. Глоссарий терминов и определений ITIL V3/ ITIL V3 Glossary / Пер. с англ. ITIL V3 Translation Project, 2009. – 146 с.

2. Карасев А.А. Онтологическое обеспечение процессов администрирования информационных систем / Л.С. Болотова, А.А. Карасёв, С.С. Смирнов, В.А. Старых. – Качество. Инновации. Образование. – 2013. – №12. – М.: Европейский центр по качеству, 2013. – с. 88–94

3. Соловьев В.Д. Онтологии и тезаурусы: учебно-методическое пособие / В.Д. Соловьев, Б.В. Добров, В.В. Иванов, Н.В. Лукашевич. – Казань, Москва, 2006. – 157 с.

4. Карасев А.А. Формализация экспертных знаний для управления инцидентами информационных систем на основе онтологического подхода / Л.С. Болотова, А.А. Карасев, В.А. Старых. – Информационные технологии. – 2014. – №6. – М.: Новые технологии, 2014. – с. 3–10

5. Гаврилова Т.А. Формирование прикладных онтологий / Т.А. Гаврилова // КИИ-2006: труды X национальной конференции по искусственному интеллекту, Обнинск, 25-28 сентября 2006 г. М.: Физматлит, 2006. – Том 2

6. URL: http://www.thinkmap.com