

DOI:

МОНИТОРИНГ ИНФОРМАЦИОННОЙ СИСТЕМЫ НА ОСНОВЕ АНАЛИЗА ТЕКСТОВ ОБРАЩЕНИЙ ПОЛЬЗОВАТЕЛЕЙ¹

Баканов А.С.,

Институт психологии РАН

Россия, 129366 Москва, ул. Ярославская, д.13

arsb2000@pochta.ru

Ташев Т.Д.,

Институт информационных и коммуникационных технологий БАН

Болгария, София 1113, ул. „Акад. Г.Бончев” бл.2, ИИКТ-БАН, тел.: (02) 979-32-40

ttashev@iit.bas.bg

Баканова Н.Б.

Институт прикладной математики им. Келдыша РАН

Россия, 125047, Москва, Миусская пл., д.4,

nina@keldysh.ru

Аннотация: Статья посвящена разработке метода мониторинга распределенной информационной системы. Метод основан на выявлении проблематики документов, обращений и предложений, поступающих в информационную систему (ИС) и посвященных вопросам и проблемам, связанным с функционированием информационной системы. В соответствии с нормативами документооборота на каждый документ входного потока готовятся указания по исполнению документа. Работа требует высокого профессионализма, хорошего знания проблемной области, сферы ответственности исполнителей. Автоматизированное выявление проблематики документов, обращений и предложений, посвященных вопросам и проблемам, связанным с качеством функционирования информационной системы, позволит оперативно реагировать на неполадки, возникающие в работе ИС, значительно ускорит процесс реагирования на обращения пользователей, повысит качество мониторинга ИС.

Ключевые слова: мониторинг, критерии, проблематика документов и обращений.

Введение

Актуальность и важность задач, связанных с мониторингом крупномасштабных распределенных информационных систем, обуславливает востребованность использования новых методов и способов для осуществления мониторинга и анализа функционирования информационных систем. В настоящей статье предлагается разработка метода анализа текстов документов и обращений, поступающих в крупномасштабную распределенную информационную систему от пользователей. Предлагаемый метод направлен на выявление проблематики документа в соответствии со специально подготовленным тезаурусом.

1 Использование анализа текстовых обращений для мониторинга ИС

Эволюция систем передачи данных в крупномасштабные, распределенные информационные сети и системы, увеличение количества и качества сервисов, предоставляемых такими сетями, их востребованность - все это привело к росту количества пользователей таких сетей. Обработка документов и обращений пользователей является трудоемким процессом и занимает много времени. Среди большого потока документов и обращений, направляемых пользователями администраторам сетей, встречаются документы связанные с качеством функционирования сетей. Предварительный анализ таких документов позволит ускорить сортировку и обработку поступающих обращений, а также повысит эффективность мониторинга функционирования крупномасштабных информационных сетей.

В рамках рассматриваемого метода предполагается создание интеллектуальных средств поддержки принятия решений в области мониторинга функционирования распределенных систем. Современные достижения в области интеллектуальных технологий, включая базы знаний, экспертные системы, системы поддержки принятия решений позволяют разработать интеллектуальные средства для сортировки и структурирования поступающих обращений.

Вопросы, связанные с интеграцией информационных систем и систем поддержки принятия решений в различных прикладных направлениях активно обсуждается на международных конгрессах, конференциях и семинарах, им посвящено большое количество научных работ [1, 9]. В рамках

¹ Статья подготовлена при поддержке гранта РФФИ №18-07-00833А

предлагаемого метода предполагается проведение анализа текста из поступающих в службу поддержки документов и обращений пользователей. На основе извлеченных данных предполагается сортировка и структуризация по определенным тематическим направлениям. Будет организован информационный поиск методом контент-анализа на соответствие с терминами проблемно-ориентированного тезауруса, анализ полученных результатов и представление их в виде удобном для оперативного принятия решения по поступившему обращению [3].

Метод контент-анализа предполагает формализованное исследование содержания текстовых массивов в целях количественной оценки с целью последующей интерпретации выявленных закономерностей. Посредством контент-анализа, в предлагаемом методе будет обработан массив текстов, в результате чего будет произведена количественная оценка обращения или документа [4, 7].

Тезаурус – тип словаря, описывающий лексическую семантику, в котором слова сгруппированы в соответствии с понятийной классификацией. В предлагаемом методе, использование тезауруса позволит соотнести слова с другими понятиями и их группами, благодаря чему может быть получена некоторая структурированная модель поступающего обращения или документа.

Таким образом, в предлагаемом методе предлагается объединить два метода работы с текстовой информацией для создания нового метода автоматизированной обработки текстовых обращений и документов, поступающих к персоналу, отвечающему за обслуживание распределенной информационной системы. Также в рамках предлагаемого метода предусматривается разработка и интеграция с модулем автоматизированной подготовки проектов указаний и правил по решению проблем, связанных с обслуживанием информационной системы [5, 6].

2 Интеграция контент-анализа и направленного поиска

Идея метода автоматизированной подготовки проектов указаний и правил по решению проблем, связанных с обслуживанием информационной системы, базируется на использовании справочника распределения должностных полномочий, подготовленного в виде проблемно-ориентированного тезауруса, и адаптации метода контент – анализа для алгоритмической обработки текстов поступающих документов и обращений. Метод включает интеграцию контент – анализа и направленного поиска на основе специализированного тезауруса. Как исследовательский метод контент - анализ используется практически во всех областях науки, так или иначе практикующих анализ текстовых источников – в теории массовой коммуникации, в социологии, политологии, истории и других. Метод контент – анализа, применяется для качественно-количественного анализа содержания документов, с целью выявления или измерения различных фактов и зависимостей, отраженных в них. Возможность использования метода связана с выявлением единиц анализа, которые варьируются в зависимости от целей, типа информационного массива, а также ряда дополнительных причин. Для решения задачи «подготовки проектов указаний» в качестве единиц анализа выбраны понятийно-тематические единицы [6, 7, 9] справочника распределения должностных полномочий обслуживающей организации и должностных обязанностей персонала, обслуживающего распределенную информационную систему.

На первом этапе работ по разработке метода мониторинга распределенной информационной системы необходимо выполнить следующие работы:

- 1) принципиальная разработка метода анализа текстов, поступающих в информационную систему;
- 2) разработка технологии представления справочника должностных полномочий в виде проблемно – ориентированного тезауруса

На втором этапе работ, на основе справочника распределения должностных полномочий, который будет представлен в виде тезауруса, будет осуществлена разработка алгоритмов поддержки принятия решений при подготовке проектов указаний на основе разработанного метода анализа текстов документов.

На третьем этапе планируется разработка программной модели, обеспечивающей реализацию метода и интеграцию его в распределенную информационную систему, а также разработка интерфейсов, обеспечивающих удобные режимы работы при подготовке проектов указаний;

В результате в результате реализации данного метода мониторинга будет реализована следующая технологическая цепочка по обработке поступающих в систему документов/обращений:

- 1) выявление проблематики документа, поступающего в систему;
- 2) анализ поступивших проблемно-ориентированных текстовых массивов на основе проблемно – ориентированного тезауруса;

3) сортировка обращений, визуализация обращений, а также автоматизированная подготовка проектов указаний и правил по решению проблем, связанных с обслуживанием информационной системы;

Описанный в настоящей работе метод основан на использовании справочника распределения должностных полномочий обслуживающей организации и должностных обязанностей персонала, обслуживающего распределенную информационную систему, и метода контент – анализа, позволяющего выявлять определенные тематически – понятийные единицы в тексте документа. Реализация описанного метода и интеграция его с информационными системами, позволит повысить эффективность мониторинга распределенных систем, снизит трудоемкость работы обслуживающего персонала и значительно уменьшит время реагирования на проблемные ситуации.

Литература

1. *Баканов А.С.* Аспекты интеграции систем электронного документооборота с системами поддержки принятия решений // *Электросвязь* № 1. 2015 с. 23-25.
2. *Баканов А.С., Зеленова М.Е.* Когнитивно-стилевые детерминанты успешности профессиональной деятельности // *Социальная психология и общество*. 2015. Т.6. №2. – С. 61-75.
3. *Баканова Н.Б.* Использование программно-технических комплексов для повышения эффективности контроля в системах документооборота // *«Электросвязь»*. 2007. № 6. С. 51–53.
4. *Петровский А.Б.* Многокритериальное принятие решений по противоречивым данным: подход теории мультимножеств. // *Информационные технологии и вычислительные системы*. 2004. №2. С. 56-66.
5. *Степановская И.А.* Концепция цифрового интернет-мониторинга / *Материалы 11-й Международной конференции «Управление развитием крупномасштабных систем» (MLSD'2018, Москва)*. М.: ИПУ РАН, 2018. т. 2. С. 426-429.
6. *Фархадов М.П., Петухова Н.В., Васьковский С.В., Фархадова М.Э.* Повышение эффективности речевого интерфейса с применением когнитивных и лингвистических знаний // *Управление большими системами*. 2019. Вып. 81. С. 90-112.
7. *Фархадов М.П.* Распознавание речи в системах массового обслуживания населения // *Труды СПИИРАН*. 2011. вып. 4 (19). С. 65-86.
8. *Цвиркун А.Д.* Основы синтеза структуры сложных систем. – М.: Наука, 1997. – 256с.
9. *Bakanov A., Atanasova T. and Bakanova N.*, "Cognitive Approach to Modeling Human-Computer Interaction with a Distributed Intellectual Information Environment," 2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE), Sofia, Bulgaria, 2019, pp. 1-4. doi: 10.1109/BdKCSE48644.2019.9010597
10. *Tashev T., Atanasova T.* Computer Simulation of MIMA Algorithm for Input Buffered Crossbar Switch // *International Journal "Information Technologies & Knowledge"*, 2011, Vol.5, No. 2, pp.183-189.