

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.89+519.237.8

Сычѳв
Алексей Анатольевич

Исследование методов интеллектуального анализа данных функционирования
технических систем

АВТОРЕФЕРАТ
на соискание степени магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Скобцов В. Ю.
к.т.н., доцент

Минск 2019

КРАТКОЕ ВВЕДЕНИЕ

Большое количество информации, поступающей и накапливающейся в специализированных банках данных на всех этапах жизненного цикла бортовой аппаратуры малых космических аппаратов (БА МКА), может быть эффективно использовано для совершенствования процесса анализа надежности, живучести и работоспособности БА и ее подсистем. Данные телеметрии БА МКА представляют собой, как правило, разнородные нерегулярные многомерные данные. Поэтому актуальным является разработка методик и алгоритмов анализа таких данных с целью построения кластерных, классификационных и прогностических моделей для оценки состояния работоспособности БА МКА.

В данной работе рассматривается задача идентификации наличия/отсутствия отказа в процессе функционирования БА МКА на основе анализа данных телеметрии БА МКА. Данная задача решается с применением метода логического анализа данных, который позволяет построить логическую классификационную модель на основе базовых концепций дифференциации и интеграции содержащихся в наборе данных. Логический анализ данных проводится на основе минимизированного набора бинарных или бинаризованных атрибутов из вектора телеметрических данных. Формируются два отдельных набора положительных и негативных шаблонов, совокупность которых составляет классификационную модель. Предложенный метод будет апробироваться на данных телеметрической информации с реальной БА МКА.

В работе использован инструментарий языка Java, фреймворк JavaFX, среда разработки IntelliJ IDEA 14.0.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью данной диссертационной работы является исследование методов интеллектуального анализа данных и создания программного обеспечения для решения задачи идентификации наличия/отсутствия отказа в процессе функционирования БА МКА (бортовой аппаратуры малого космического аппарата) на основе анализа данных телеметрии БА МКА.

Для выполнения поставленной цели требуется решить нижеследующие задачи:

1. Разработать методы и алгоритмы обработки телеметрических данных для системы идентификации наличия/отсутствия отказов.
2. Разработать архитектуру программного средства идентификации наличия/отсутствия отказов с функциями чтения и записи в файл промежуточных служебных данных.
3. Реализовать часть программного средства в виде многопоточной системы для вычисления, а также чтения и записи служебных данных.
4. Провести экспериментальные исследования разработанной системы.

Объектом исследования являются проблемно ориентированные алгоритмы логического анализа данных и идентификации наличия/отсутствия отказов.

Предметом исследования является программное обеспечение компьютерных систем для решения задач идентификации наличия/отсутствия отказов, методы и алгоритмы логического анализа данных для поддержки принятия решений.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность использования алгоритма логического анализа телеметрических данных для построения идентификационных шаблонов, позволяющих установить наличие либо отсутствие отказов на основе переданных входящих данных.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и в рамках задания 1.4 НИР «Разработать комплекс методик и программных средств для оценки надежности бортовой аппаратуры маломассогабаритных космических аппаратов при ее проектировании, наземных испытаниях и эксплуатации» (шифр «Надежность») программы Союзного государства «Мониторинг-СГ».

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя В. Ю. Скобцова, заключается в формулировке целей и задач исследования, а также в предоставлении экспериментальных данных для использования их в обучающих и тестовых целях.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на Республиканской научно-технической конференции «Информационные технологии и системы: проблемы, методы, решения» (ИТС – 2018) (Минск, Беларусь, 2018); 55-й научной конференции аспирантов, магистрантов и студентов БГУИР (Минск, Беларусь, 2019); 32-й Петербургской научной мультikonференции «Математические методы в технике и технологиях» (Санкт-Петербург, Россия, 2019).

Опубликованность результатов диссертации

По теме диссертации опубликовано 2 печатные работы, из них 1 статья в сборниках публикаций, 1 работа в сборниках трудов и материалах международной конференции.

Структура и объем диссертации

Данная диссертационная работа включает в себя введение, общую характеристику работы, три главы, заключение, список использованных источников, список публикаций автора и приложения.

В первой главе анализируется предметная область. Выявляются основные существующие проблемы, связанные с тематикой магистерской диссертации, разбираются направления их решения.

Вторая глава посвящена разработке методов и алгоритмов интеллектуального анализа данных функционирования технических систем.

В третьей главе приводится архитектура программного средства и его практическая реализация, представлены результаты экспериментальных исследований и практического применения разработанного программного обеспечения.

Общий объем работы составляет 61 страница, из которых основного текста – 51 страница, 13 рисунков на 13 страницах, 1 таблицы на 1 странице, список использованных источников из 34 наименований на 3 страницах, 1 приложение на 45 страницах.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** определена предметная область исследования, приведена актуальность темы диссертационной работы, кратко рассмотрены исследуемые вопросы, обоснована практическая важность работы.

В **первой главе** проведен анализ предметной области. Кратко рассмотрены исследуемые технические системы. Проведен анализ данных функционирования телеметрии. В связи с тем, что проводимые исследования были направлены на расширение возможностей и функционала комплекса методик и программных средств для оценки надежности бортовой аппаратуры малых космических аппаратов, в частности, для программного модуля интеллектуального анализа данных (ПМ ИАД), решения задачи идентификации и прогнозирования отказов по данным телеметрии МКА, был приведен краткий обзор методов и подходов, реализованных в рамках ПМ ИАД данного комплекса.

Был проведен анализ теоретических основ машинного обучения с учителем, дано понятие классификации, исследованы некоторые классификационные алгоритмы, проанализированы их достоинства и недостатки.

Также была разобрана литературная теоретическая основа данной работы. Было выяснено, что метод ЛАД и его теоретическое ядро является масштабируемым на большие размеры данных и позволяет найти свое применение для решения задач, поставленных перед данной работой.

Результаты исследований, проведенных в этих направлениях, отражены в работах П. Л. Хаммера (P. L. Hammer), Т. О. Бонатеса (T. O. Bonates), Т. Ибараки (T. Ibaraki), А. Когана (A. Kogan), А. Ядава (A. Yadav), А. Сингха (A. Singh), П. Х. Шета (P. H. Sheth), С. Гангопадхья (S. Gangopadhyay), Е. Бороса (E. Boros), М. Энтони (M. Anthony), Дж. Рэтсэби (J. Ratsaby) и др.

Вторая глава посвящена разработке методов интеллектуального анализа данных функционирования технических систем. Исходя из проведенного в **первой главе** анализа, учитывая положительные и отрицательные стороны проанализированных классификаторов было принято решение исследовать возможности одного из перспективных современных методов классификации – логического анализа данных (ЛАД), для решения задачи анализа телеметрии МКА в контексте проблемы идентификации потенциальных отказов.

Для решения поставленной задачи были использованы реальные телеметрические данные, которые были собраны в процессе функционирования бортовой аппаратуры малого космического аппарата АИСТ-1 Самарского университета им. С.П. Королева.

Данные телеметрии МКА АИСТ-1 были получены в рамках меморандума о научно-техническом сотрудничестве ОИПИ НАН Беларуси и Самарского университета им. С.П. Королева.

Данные телеметрии представляют из себя таблицу, строками (элементами) которой являются наборы показателей датчиков за единицу времени. Столбцами же данной таблицы являются атрибуты (значения вышеупомянутых датчиков). Последний столбец данных содержит в себе бинарный параметр, значение которого, будучи “0”, означает, что отказа не было, а “1” – означает то, что отказ наблюдался.

Были рассмотрены и исследованы основные этапы работы метода логического анализа данных.

В **третьей главе** приводится программная реализация разработанных методов интеллектуального анализа данных функционирования технических систем. Рассматривается общая архитектура программной реализации. Приводится обоснование выбора средств программной реализации – основным ориентиром при выборе средств реализации являлось то, что комплекс программных средств и методик для оценивания и анализа показателей надежности и живучести БА МКА и интеллектуального анализа данных телеметрии БА МКА, для которого с целью расширения функциональных возможностей разрабатывалось данный программный модуль, имеет сервис-ориентированную архитектуру и веб-реализацию. Бекенд веб-реализации разработан на языке Java с использованием сервера Tomcat.

Исходя из поставленных задач и особенностей веб-реализации основного программного комплекса было принято решение о целесообразности использования языка программирования Java.

Были приведены основные использованные в ходе разработки программного средства возможности языка Java.

Также были разобраны основные методологии разработки и приведено обоснование выбора одной из них.

Далее были описаны особенности программной реализации основных этапов работы метода логического анализа данных. Приведена структурная схема разработанного программного средства. Также были приведены результаты проверки работоспособности программного средства. По итогам проверки было определено, что программное средство полностью реализует алгоритм логического анализа данных и может применяться для идентификации наличия или отсутствия отказов, на основе телеметрических данных.

ЗАКЛЮЧЕНИЕ

В ходе работы над данной диссертацией был проведен анализ и исследована литература по основным методам машинного обучения - алгоритмами классификации и кластеризации. По итогам анализа первого этапа было проведено основное исследование по публикациям, связанных с методом логического анализа данных.

Также были разобраны основные понятия, которые использовались в ходе разработки программного средства либо во время описания программы в данной работе.

Кроме этого был проведен анализ существующих языков программирования и предназначенных для них средств разработки. Особое внимание уделялось возможности легкого использования многопоточности, кроссплатформенности.

Разработанная реализация метода логического анализа данных применим для автоматизированного анализа данных телеметрии БА МКА с целью решения задач анализа их надежности. Метод ЛАД может быть применен как самостоятельно так и в составе ансамбля методов классификации/прогнозирования отказов на основе данных телеметрии БА МКА.

Предложена архитектура программного модуля для решения задачи идентификации наличия/отсутствия отказов на основе данных телеметрии и функционирующая на компьютерах общего назначения, в том числе с использованием веб-сервисов.

В ходе рассмотрения данного алгоритма была проведена разработка программного модуля, позволяющего определять возможные отказы работы бортового оборудования малого космического аппарата.

Проведено тестирование работоспособности программного средства. В качестве проверки данное программное средство апробировалось в сфере автоматизированного анализа данных телеметрии с целью улучшения надежности аппаратуры. Как упоминалось выше, телеметрические данные реальные и были собраны в процессе функционирования бортовой аппаратуры малого космического аппарата АИСТ-1 Самарского университета им. С. П. Королева.

В процессе вычислительных экспериментов было установлено, что точность обнаружения наиболее распространенных отказов (18 записей) составляет около 94%, менее распространенных отказов (13 записей) – точность обнаружения составляла около 76%.

В качестве проверки данное программное средство апробировалось в сфере автоматизированного анализа данных телеметрии с целью улучшения надежности аппаратуры. Как упоминалось выше, телеметрические данные

реальные и были собраны в процессе функционирования бортовой аппаратуры малого космического аппарата АИСТ-1 Самарского университета им. С. П. Королева.

Результаты работы данного программного средства были представлены на 55-й Юбилейной Научно-Практической Конференции студентов и магистрантов, а также на конференции «Информационные технологии и системы: проблемы, методы, решения» (ИТС – 2018) и 32-й Петербургской научной мультikonференции.

Рекомендации по практическому использованию результатов

1. Полученные результаты планируется интегрировать в комплекс методик и программных средств для оценки надежности бортовой аппаратуры маломассогабаритных космических аппаратов, в частности, в программный модуль интеллектуального анализа данных телеметрии МКА, для решения задачи идентификации и прогнозирования отказов по данным телеметрии.

2. Результаты работы могут использоваться при подготовке кадров, входящих в состав команды наземных комплексов управления МКА

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Сычѳв, А. А. Имплементация метода логического анализа данных / А. А. Сычѳв // Информационные технологии и управление: материалы 55-ой научной конференции аспирантов, магистрантов и студентов. – Минск: БГУИР, 2019. – с. 65.

2-А. Скобцов, Ю. А. Логический анализ данных телеметрии бортовой аппаратуры малых космических аппаратов / Ю. А. Скобцов, В. Ю. Скобцов, А. А. Сычѳв // Математические методы в технике и технологиях: материалы 32-й Петербургской научной мультikonференции. - Санкт-Петербург: СПбПУ, 2019. – пленарный доклад.