

ЭЛЕКТРОННАЯ КОМПОНЕНТНАЯ БАЗА, ФОТОНИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ

УДК 629.78

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ДИАГНОСТИКИ СПУТНИКОВЫХ СБОЕВ

© 2021 г. Ю.Л. Лобков, А.А. Парамонов, Д.А. Иванов*, Д.С. Савельев

ФГАУ «Военный инновационный технополис «ЭРА», Анапа, Россия

*E-mail: Ivanov.dmitry0425@gmail.com

В данном исследовании поднимается тема использования методов машинного обучения для диагностики неисправностей спутниковых подсистем по параметрам их телеметрии. Данные телеметрии кластеризуются с использованием алгоритма кластеризации k -средних в сочетании с функцией t -распределенного стохастического вложения соседей (t -СВС) для уменьшения размерности. Обсуждается классификация данных с помощью логического анализа данных (ЛАД), для генерации шаблонов для каждого класса отказов, которые используются для определения вероятной причины отказа для каждого параметра телеметрии. Эти вероятности позволяют с помощью анализа дерева отказов (АДО), определить наиболее вероятную причину отказа спутника.

ВВЕДЕНИЕ

Космические системы дороги из-за высокой стоимости разработки и запуска в космос. Потери таких систем являются постоянными и обычно происходят из-за постепенной деградации компонентов системы и/или устройств. Во избежание непредвиденных обстоятельств и естественных неисправностей желательно, чтобы бортовая система диагностики неисправностей космических аппаратов (КА) была способна обнаруживать, изолировать, идентифицировать и классифицировать неисправности в системе [1].

Анализ и понимание ухудшения и отказов выбравших подсистем на орбите, то есть ее реальной надежности, обеспечивает полезную обратную связь для производителей космических аппаратов и является важным первым шагом на пути к повышению общей надежности космических аппаратов [2].

В последние годы для обнаружения и диагностики аномалий в работе КА предлагаются методы машинного обучения и интеллектуального анализа данных. Распространенный способ устранения аномалий – анализ истории данных телеметрии, чтобы определить корень проблемы. Специалистами активно изучается проблема обнаружения, изоляции и идентификации неисправностей, рассматриваются подходы к диагностике неисправностей на основе моделей для аэрокосмических систем.

Эти подходы включают экспертные системы, нейронные сети, машину опорных векторов (МОВ), метод главных компонент (МГК), оценку параметров, фильтры Калмана (ФК).

В этой работе исследуются методы машинного обучения, которые можно использовать для диагностики неисправностей КА. Использовался метод интеллектуального анализа данных, чтобы изучить возможную производительность, выраженную любым параметром телеметрии, указывающим на работоспособность определенного бортового устройства. Это позволяет оператору спутника контролировать его общее состояние, чтобы снизить риск отказа с помощью точных и автоматизированных методов. Предлагаемая методика существенно поможет улучшить принятие решений в аварийных ситуациях и временной потере спутников.

АНАЛИЗ ПОВЕДЕНИЯ КОСМИЧЕСКИХ АППАРАТОВ

В контексте этой работы анализ поведения спутника означает: «оценку функциональности спутника при определенном рабочем режиме», где есть набор спутниковых режимов, таких как режим съемки, нормальный рабочий режим и полностью рабочий режим. На основе телеметрических измерений возможно определить режим работы спутника и соответствующие функциональные возможности. Для этого распаковываются данные из

необработанных файлов телеметрии и преобразуются в формат временных рядов, который соответствует предполагаемой методике.

Теория МОВ сводит к минимуму ожидаемую ошибку обучающейся машины, что снижает проблему переобучения. Это явление известно, как принцип минимизации структурного риска (МСР), который используется для задач классификации и регрессии. МОВ имеет два типа: классификация опорных векторов (КОВ) и регрессия опорных векторов (РОВ) [3]. КОВ классифицирует данные на две или более группы на основе введенных функций, максимизируя поля между введенными классами, такими как категоризация текста. РОВ используется с данными временных рядов для прогнозирования следующих значений путем минимизации суммы расстояний от точек данных до гиперплоскости, например, прогнозируемых цен на акции. Приближенная функция используется для нелинейной МОВ, называемой функцией ядра, которая используется для нахождения приближения функций с действительными значениями на основе условия Мерсера, таких как линейное ядро, полиномиальное ядро, сигмовидное ядро и радиальная базисная функция (РБФ). В этом исследовании использовалась функция ядра РБФ:

$$K(x, x') = \exp\left(\frac{-\|x - x'\|^2}{2^2}\right), \quad (1)$$

где РБФ имеет параметр r , и необходимо определить значение r до эксперимента.

В последнее время применение вид МОВ для прогнозирования временных рядов, называемый регрессией опорных векторов, также показал много прорывов и правдоподобных результатов, таких как прогнозирование финансового рынка, прогнозирование цен на электроэнергию, оценка мощности потребления и реконструкция хаотических систем.

В Китае разработали обнаружение аномалий для данных телеметрии КА с использованием регрессионной модели машины опорных векторов наименьших квадратов. Регрессионная модель используется для выявления аномального поведения космических аппаратов на орбите.

Также существует подход к обнаружению неисправностей и диагностике космических аппаратов, основанный на анализе главных компонент (АГК) и машинах опорных векторов. АГК используется для выбора наиболее важных функций и со-

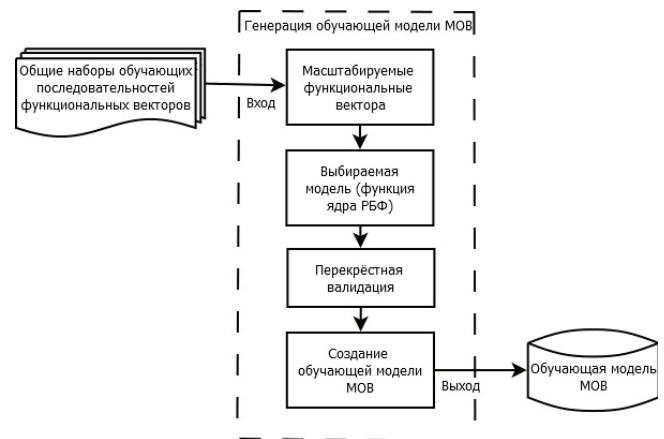


Рис. 1. Блок-схема машины опорных векторов

кращения входных данных до векторов признаков с низкой размерностью, которые применяются к мультиклассовой модели классификации МОВ с целью определения типа неисправности.

Перед выполнением прогнозирования тестовых последовательностей создается обучающая модель МОВ, как показано на рис. 1. Подготавливается «Общий обучающий набор», чтобы сгенерировать обучающую модель для прогнозирования телеметрии. Все последовательности в общем обучающем наборе сформированы обучающими последовательностями.

ДИАГНОСТИКА ОТКАЗОВ КОСМИЧЕСКИХ АППАРАТОВ

Раннее обнаружение, как правило, важно для уменьшения программного воздействия потенциального сбоя. В этом разделе представляется методология, предложенная для построения спутниковой системы диагностики неисправностей, основанную на нескольких последовательных этапах обработки данных, уменьшения размеров, классификации, генерации шаблонов и анализа данных.

ЛАД – это подход алгоритмов классификации, который используется для классификации двух или нескольких классов. ЛАД классифицирует наблюдения на положительные или отрицательные, где положительные наблюдения означают ошибочные данные или что-то неправильное, в то время как отрицательный означает, что данные корректны. Двухклассный ЛАД состоит из трех этапов: бинаризация данных, генерация паттернов и формирование теории.

На этапе бинаризации каждый входной объект преобразуется в набор двоичных атрибутов (переменных). После этапа бинаризации ЛАД генерирует набор логических правил, называемых паттернами (шаблонами); Они отличаются от наблюдений, принадлежащих к одному классу, а не к другому. Такие шаблоны называются позитивными или негативными шаблонами, где позитивные шаблоны описывают ошибочные данные. Паттерны играют важную роль в ЛАД, поскольку они допускают четкую интерпретацию экспертами людьми. Каждая положительная (отрицательная) точка наблюдения покрывается, по крайней мере, одним положительным (отрицательным) шаблоном и не покрывается никакими отрицательными (положительными) шаблонами, которые были сгенерированы. Кроме того, отдельные шаблоны также могут указывать на положительный (отрицательный) характер новых точек наблюдения, и поэтому можно ожидать, что правильно выбранный набор шаблонов может быть использован для построения общего правила классификации. Это правило является расширением частично определенной булевой функции и называется формацией теории.

В 2016 году была предложена методология прогнозирования, основанная на надежности и использующая данные мониторинга состояния, которые могут иметь дело с любым количеством индикаторов состояния, без выбора наиболее значимых, как предлагают многие методы. Используя метод Каплана-Мейера в качестве метода оценки, управляемого временем, в сочетании с логическим анализом данных в качестве метода диагностики, управляемого событиями, для отражения влияния условий эксплуатации на срок эксплуатации контролируемого оборудования.

АДО – это дедуктивный нисходящий подход, который определяет потенциальный отказ, а затем определяет все события, которые могут привести к отказу. АДО предоставляет два типа результатов: качественные и количественные. Целью качественного анализа является сокращение структурированного дерева отказов. Целью количественного анализа является расчет вероятности отказа главного события на основе качественного анализа.

В качественном анализе законы булевой алгебры используются для удаления избыточных событий отказа компонентов из списка основных событий для получения минимальных наборов сокращений. Наборы минимальных сокращений представляют собой комбинацию минимальных (необходимых и достаточных) состояний отка-

за компонентов, которые вызывают режим отказа системы. Подход с использованием минимальных наборов сокращений – это классический подход, который предоставляет важную качественную, а также количественную информацию.

Для количественной оценки вероятности ключевого события в дереве отказов должна быть представлена вероятность каждого базового события. После этого вероятности базовых событий распространяются вверх до ключевого события с использованием логических отношений для дерева отказов. Количественные данные необходимо вводить для базовых событий; входные данные состоят из вероятностей этих событий дерева отказов.

Вероятность главного события – это вероятность общего отказа системы. Она рассчитывается из дерева отказов в зависимости от вероятностей, которые применяются к базовым событиям. Вероятность промежуточных событий, также называемая вероятностью перехода, вычисляется как вероятность ключевого события из ее поддерева. Вероятность главного события рассчитывается как:

$$Q(t) = 1 - e^{-\lambda t}, \quad (2)$$

где $Q(t)$ – это ключевое событие недоступности или ненадежности, а λ – постоянная интенсивность отказов компонента.

Наиболее важным результатом АДО является набор показателей важности для всех событий в дереве отказов; эти показатели рассчитаны для главного события. Как промежуточным, так и базовым событиям можно назначить приоритеты в соответствии с их мерой важности. Наивысшая важность может использоваться для распределения ресурсов, использующихся для минимизации вероятности отказа при ключевом событии.

Чтобы выполнить процесс обучения для ЛАД и сгенерировать паттерн(ы) классификации, который(е) соответствует одному классу событий, использовались классифицированные данные К-средних. Паттерн представляет собой комбинацию условно меняющихся параметров. Для каждого паттерна рассчитывается его вероятность как умножение вероятности каждого параметра паттерна на основе АДО. Более того, кластеризация является необходимым шагом перед применением ЛАД, чтобы выполнить этап обучения с необходимой точностью.

Получение вероятности паттерна позволяет нам судить о возможности появления последовательных комбинированных параметров и возможной

причине. Применение этой процедуры к каждой спутниковой подсистеме приводит к вероятностному подходу к причине отказа. Чтобы вычислить вероятность отказа подсистемы, используется анализ дерева отказов на сгенерированных паттернах. АДО используется для определения вероятности того, какой компонент подсистемы является причиной отказа спутника.

ОЦЕНКА

Спутник имеет большое количество параметров телеметрии, которое может превышать несколько тысяч. Были выбраны параметры, которые отражают производительность критических подсистем, таких как источник питания, блок кондиционирования, связь с земной станцией (S -диапазон), температура внутри спутника и экранирование солнечных панелей. Напряжение на шине питания является одним из наиболее важных параметров для различных подсистем, таких как сама шина питания, коммуникационные и бортовые компьютерные системы. Необходимо использовать первый набор значений для выполнения процесса обучения МОВ, в то время как последующий набор данных будет использоваться для сравнения предсказанных значений с реальными значениями для оценки алгоритма. В процессе используются файлы телеметрии за разные периоды жизни спутника. Можно использовать две трети значений телеметрии для обучения алгоритма, а оставшуюся треть для его оценки. Прогнозируемые значения сопоставляются с реальными значениями параметров, извлеченных из файла телеметрии. Затем оценивается среднеквадратичная ошибка для предсказанных значений, чтобы оценить производительность алгоритма. Предлагаемая методика представлена на рис. 2.

K -средние используются для классификации данных непересекающихся групп (кластеры) на основе их сходства. Метод K -средних использует центроиды для представления кластеров путем оптимизации функции квадратичной ошибки. Алгоритм кластеризации K -средних зависит от евклидова расстояния для вычисления центроида каждой группы, поэтому точки данных в одном кластере считаются принадлежащими одному классу, а точки в разных кластерах считаются принадлежащими к разным классам. Рис. 3–5 представляют собой применение метода K -средних к параметрам из разных спутниковых подсистем для получения двух кластеров, введенных в ЛАД.

Для большого количества функций в наборе данных для K -средних потребуется алгоритм

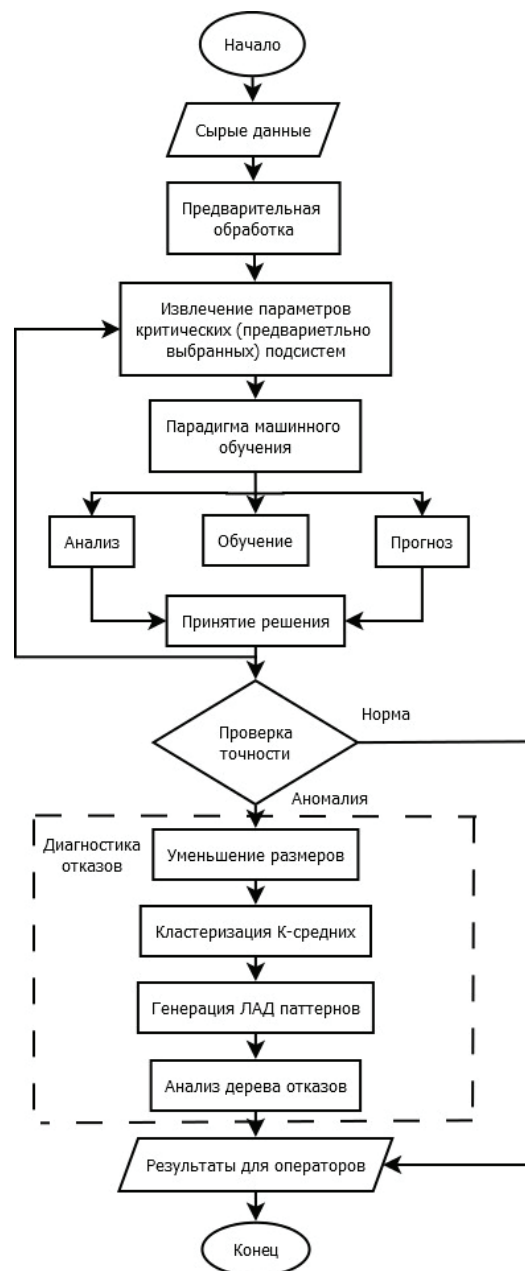


Рис. 2. Блок-схема предлагаемой методики

уменьшения размерности, такой как t -распределенное стохастическое вложение соседей. t -СВС – это метод нелинейного уменьшения размерности, в котором стохастическое вложение соседей начинается с преобразования многомерных евклидовых расстояний между точками данных в условные вероятности, которые представляют собой сходства [4]. t -СВС часто используется не только для уменьшения размерности, но и для визуализации данных, потому что его основная цель – сокращение набора данных функций до 2-D или 3-D, чтобы обеспечить более визуальное представление

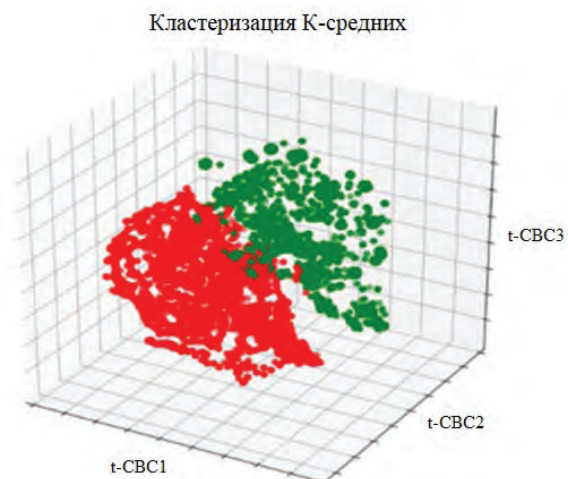


Рис. 3. Блоковая кластеризация

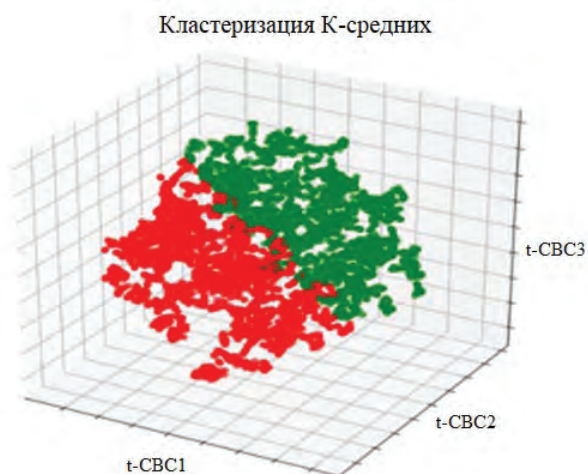


Рис. 4. Батарейная кластеризация

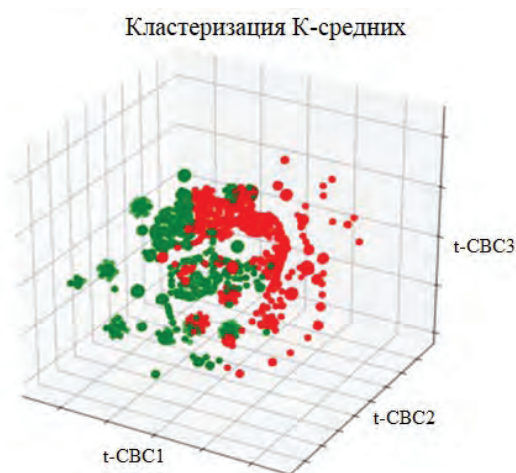


Рис. 5. S-Band кластеризация

данных. Если необходимо более трех измерений, t -СВС может не пригодиться.

Данные извлекаются и форматируются, что позволяет легко манипулировать значениями. Можно использовать первые 20% значений телеметрии для процесса обучения, 30% – для тестирования и, наконец, остальные 50% – для классификации, параметры телеметрии, извлеченные из необработанных данных, полученных со спутника, форматируются так, чтобы их можно было использовать в качестве входных данных для ЛАД.

ЗАКЛЮЧЕНИЕ

Рекомендуется использовать более простые методы регрессии такие как регрессия опорных векторов для реализации интеллектуального анализа данных спутниковой телеметрии с низкой околоземной орбиты. Данные классифицируются (отказ и исправность) в соответствии с различными режимами работы с использованием метода классификации K -средних в сочетании с функцией уменьшения размерности t -СВС. Использовались K -средние для классификации данных, а затем введение этих данных для обучения в логический модуль метода логического анализа данных для получения некоторых паттернов, которые указывают значения условных параметров. Затем каждый паттерн переводится в количество условий/событий определенных параметров. Затем используется анализ дерева отказов для оценки вероятности и зависимости каждого диапазона параметров от ключевого события (отказа спутника). ЛАД обнаруживает общие закономерности отказа, в то время как АДО определяет, какое базовое событие (соответствует определенной схеме) вызывает отказ спутника.

Использование предлагаемой цепочки подходов « K -средние – ЛАД – АДО» имеет преимущество, заключающееся в том, что система автономно способна обнаруживать и определять все возможные схемы отказа независимо, что обеспечивает лучшую точность диагностики неисправностей.

СПИСОК ЛИТЕРАТУРЫ

1. Управление космическими полетами: учеб. пособие: в 2 ч. В.А. Соловьев, Л.Н. Лысенко, В.Е. Любинский; под общ. ред. Л.Н. Лысенко. М.: Изд-во МГТУ им. Н.Э. Баумана, 2015. Ч. 2. – 426 с.

2. Системы контроля и испытания вооружения и военной техники, военная метрология: элек-

тронное учебное пособие / В.Н. Арсеньев [и др.]. – СПб.: ВКА имени А.Ф. Можайского, 2016. – 600 с.

3. Сайт machinelearning.ru [Электронный ресурс]. – Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных. Режим доступа <http://www.machine-learning.ru/wiki/index.php?title=SVM#:~:text=Машина%20опорных%20век->

[торов%20-%20является,названием%20SVM%20\(Support-%20Vec-tor%20Machine\)](http://www.machine-learning.ru/wiki/index.php?title=SVM#:~:text=Машина%20опорных%20век-), свободный – Загл. с экрана (дата обращения 06.04.2021).

4. Сайт Data Science – наука о данных по-русски [Электронный ресурс] – Сайт посвященный интеллектуальному анализу данных. Режим доступа: <http://datascientist.one/k-means-algorithm/>, свободный – Загл. с экрана (дата обращения 09.04.2021).