



МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)
«МАИ»



Кафедра 308 «Информационные технологии»

Дипломная работа

**«Разработка системы мониторинга состояния ЛА
(Integrated System Health Management) на основе
методов интеллектуального анализа данных
(Data Mining)»**

Выполнил: студент гр. 03-617 Панченко Владимир Владимирович

Руководитель: к.т.н., доцент каф. 308 Пискунов Вячеслав Алексеевич

Москва, 2013 год

Введение в предметную область

Методы контроля состояния систем

Традиционные

- интенсивная работа множества экспертов;
- глубокое знание предметной области и специфики задачи;
- сложность и большой размер построенных моделей, что затрудняет работу в режиме реального времени.

Основанные на интеллектуальном анализе данных

- минимальное участие экспертов;
- не требуют знания предметной области и устройства системы;
- компактность построенных моделей и лёгкость их интерпретации;
- позволяют извлекать знания из накопленных архивных данных.

Цели и задачи

- Разработать метод мониторинга состояния ЛА по данным телеметрии на основе методов интеллектуального анализа данных.
- Реализовать программную систему для ПЭВМ, использующую данный метод.

Требования к системе

Система должна:

- строить модель объекта контроля только на основе телеметрии при различных режимах его работы, без априорных данных о предметной области, его назначении, составе, конструкции;
- обладать способностью классифицировать аномалии в работе объекта контроля;
- в случае, если текущее поведение объекта контроля не было представлено в обучающей выборке, давать оператору численную характеристику отклонения от номинальных режимов;
- обрабатывать большие массивы входных данных (несколько десятков тысяч точек) за конечное время;
- учитывать как непрерывные, так и дискретные параметры;
- не иметь ограничений на закон распределения входных данных;
- быть устойчивой к аномалиям в обучающей выборке;
- быть устойчивой к отсутствию значений каких-либо параметров во входных данных;
- определять состояние объекта контроля в режиме реального времени.

Обзор существующих методов

- Orca
- GritBot (C4.5)
- Модель гауссовых смесей (GMM)
- Динамические байесовские сети (DBN)
- Одноклассовый метод опорных векторов (SVM)
- Inductive Monitoring System (IMS)

Сравнение существующих методов

Критерий / Метод	Orca	GritBot	GMM	DBN	SVM	IMS
Построение модели системы без априорных данных о предметной области, назначении системы, её составе, конструкции	+	+	±	+	–	+
Классификация аномалий	–	–	–	+	–	–
Численная характеристика аномалии	+	–	–	–	+	+
Обработка больших выборок за конечное время	±	–	±	+	+	+
Работа с дискретными параметрами	+	+	–	±	–	–
Отсутствие ограничений на закон распределения входных данных	+	+	–	±	+	+
Устойчивость к аномалиям в обучающей выборке	+	+	–	–	–	–
Устойчивость к отсутствию значений параметров во входных данных	+	+	–	–	–	+
Работа в режиме реального времени	–	–	+	+	+	+

Предложенный метод

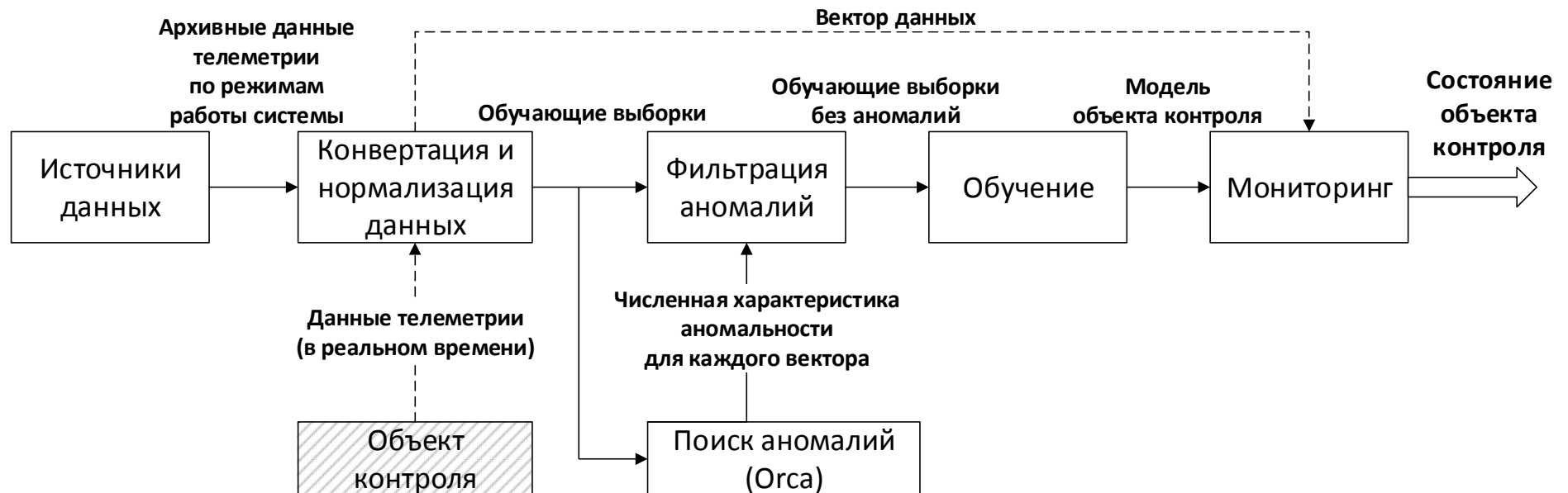
Шаг 1. Конвертация и нормализация данных.

Шаг 2. Поиск аномалий в обучающих выборках.

Шаг 3. Фильтрация аномалий в обучающих выборках.

Шаг 4. Обучение.

Шаг 5. Мониторинг.



Предложенный метод

Входные данные:

- Данные телеметрии для каждого режима работы объекта контроля;
- Пороговое значение ε ;
- Весовые коэффициенты параметров (опционально).

Выходные данные:

- Текущий режим работы объекта контроля;
- Если поведение объекта контроля аномально, ближайший режим и численная характеристика отклонения от него.

Метрика пространства:

$$d(x, y, \Omega) = \sqrt{\sum_{i=1}^p \left[\omega_i \left(a_i^{(x)} - a_i^{(y)} \right)^2 \right]} + \sum_{j=1}^q \left[\begin{cases} 0, & b_j^{(x)} = b_j^{(y)}, \\ \omega_{j+p}, & b_j^{(x)} \neq b_j^{(y)} \end{cases} \right]$$

Конвертация и нормализация данных

$$\hat{x} = \frac{x - \min(X)}{\max(X) - \min(X)}, x \in X \quad \text{— минимаксная нормализация}$$

$$\hat{x} = \frac{x - \bar{x}}{\sigma_x} \quad \text{— нормализация с помощью СКО}$$

До:

Время	Давление	Состояние клапана	Температура	Вибрация	Скорость потока
16:30:00	2857.2	Открыт	49.8	0.21	1.5



После:

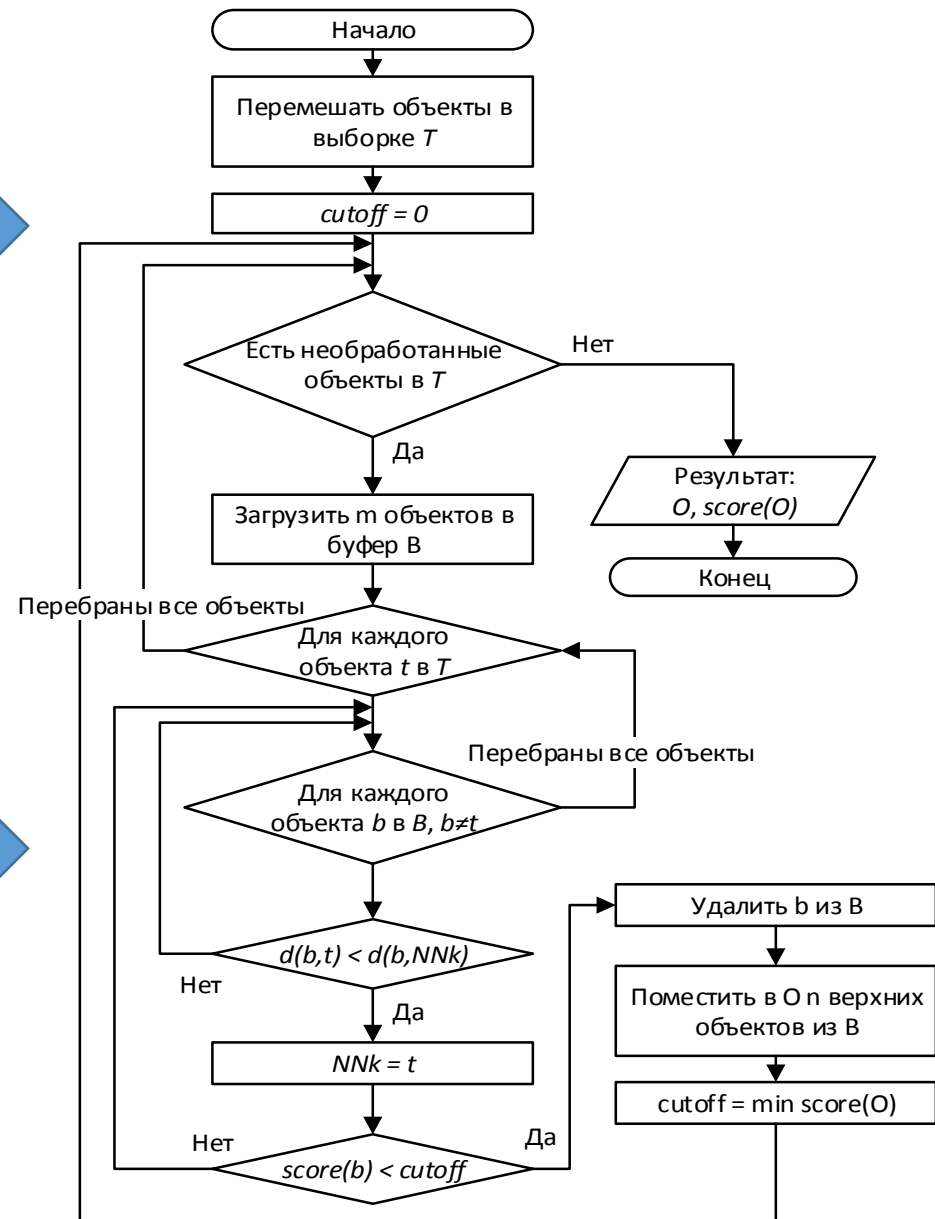
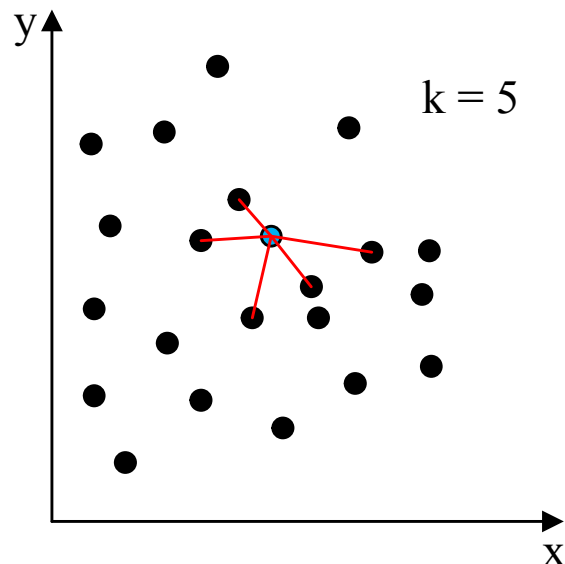
a_1	a_2	a_3	a_4	b_1
0.91	0.38	0.99	0.54	1

Поиск аномалий

Алгоритм Orca:

- рандомизация входных данных
- вложенные циклы
- правило отсечения

Принцип ближайшего соседа:



Фильтрация аномалий

Фильтр по правилу трёх сигм:

$$\forall o \in O: score(o) > M[o] + 3\sigma$$

Пороговый фильтр:

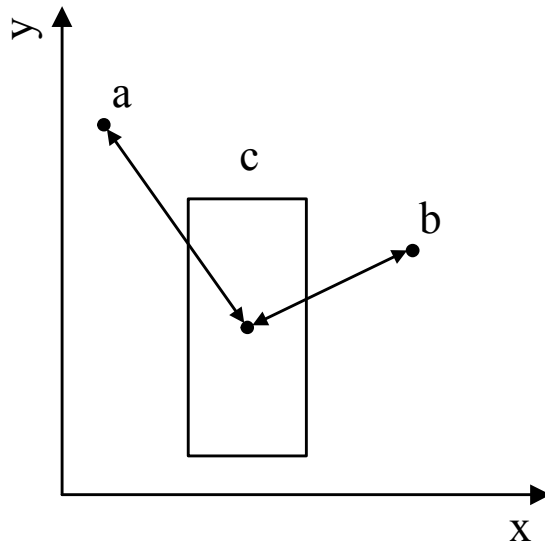
$$\forall o \in O: score(o) - \min O > \rho$$

Разностный фильтр:

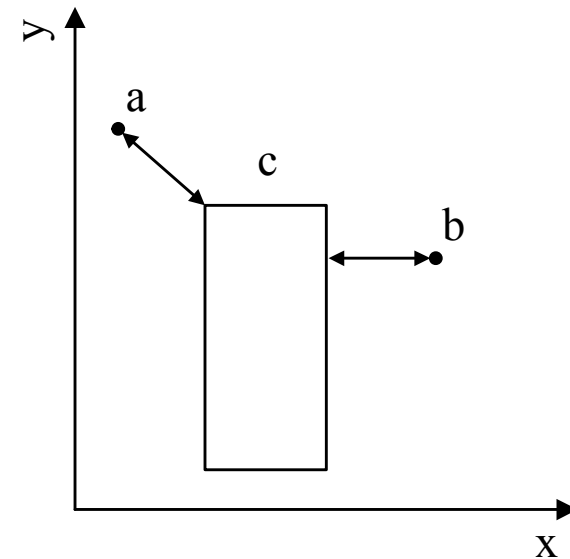
$$\forall o_i, o_j \in O, i > j: score(o_i) - score(o_j) > \Delta \rightarrow o_1 \dots o_i$$

Обучение

Функции расстояния между точкой и кластером



До центра кластера



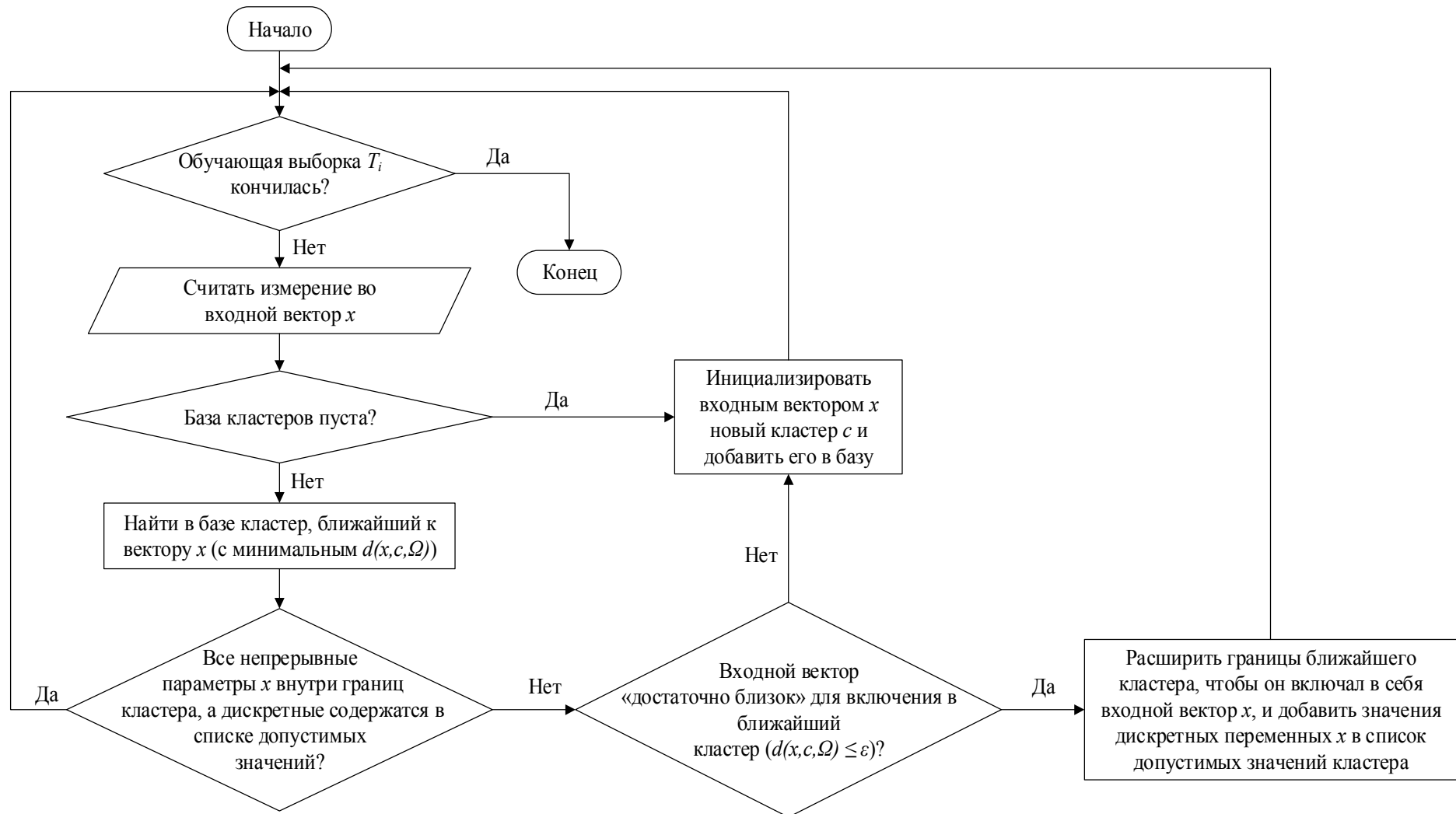
До ближайшей точки кластера

Пример кластера

	a_1	a_2	a_3	a_4	b_1
Верхняя граница	0.91	0.38	0.85	0.64	—
Нижняя граница	0.45	0.45	0.55	0.14	—
Допустимые значения	—	—	—	—	1

Обучение

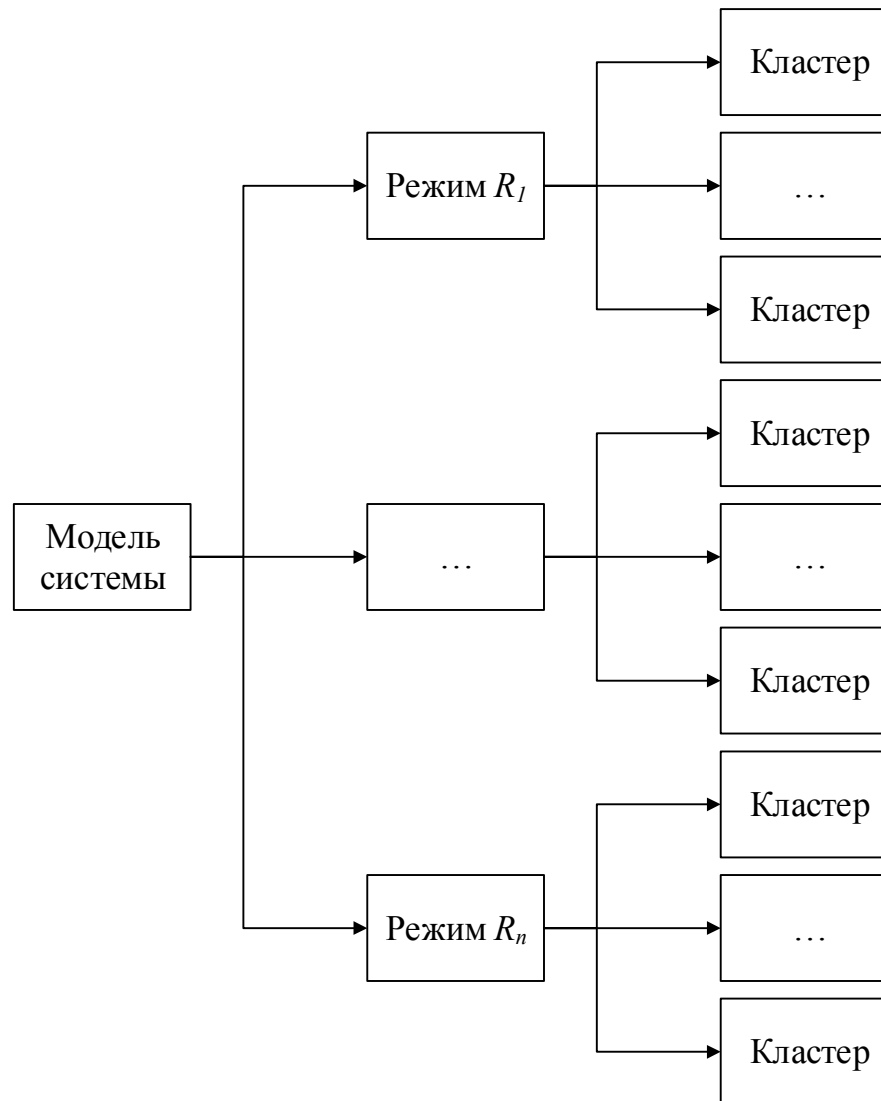
Блок-схема алгоритма обучения



Обучение

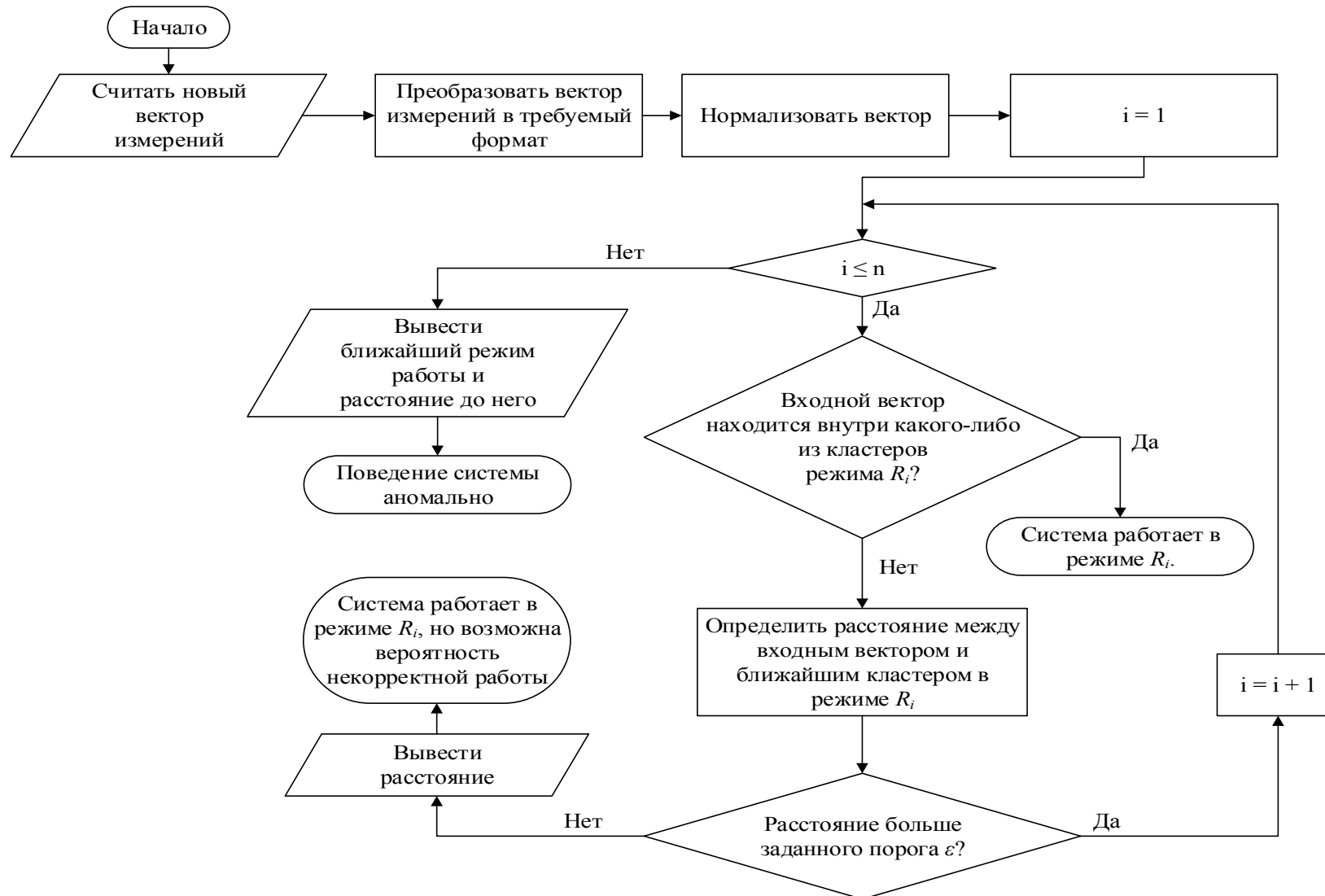
Шаг 4

Модель объекта контроля



Мониторинг

Блок-схема алгоритма мониторинга



Программная реализация

Программная реализация разработана на языке программирования C# 5.0 с использованием платформы .NET 4.5.

Ключевые особенности:

- Модульность: каждый модуль можно использовать отдельно, в том числе в составе других систем.
- Объектно-ориентированный дизайн: простота расширения системы (возможность добавления новых источников данных, типов фильтров, метрик пространства и т.д.)
- UNIX-тип приложения: лёгкость интеграции (поддержка текстовых потоков, настройка через параметры запуска)

Тестирование

Тестирование системы проводилось на основе реальных данных телеметрии гиросилового комплекса управления МКС. В качестве объекта контроля был взят один из гиродинов.

Четыре режима работы (два штатных и два нештатных):

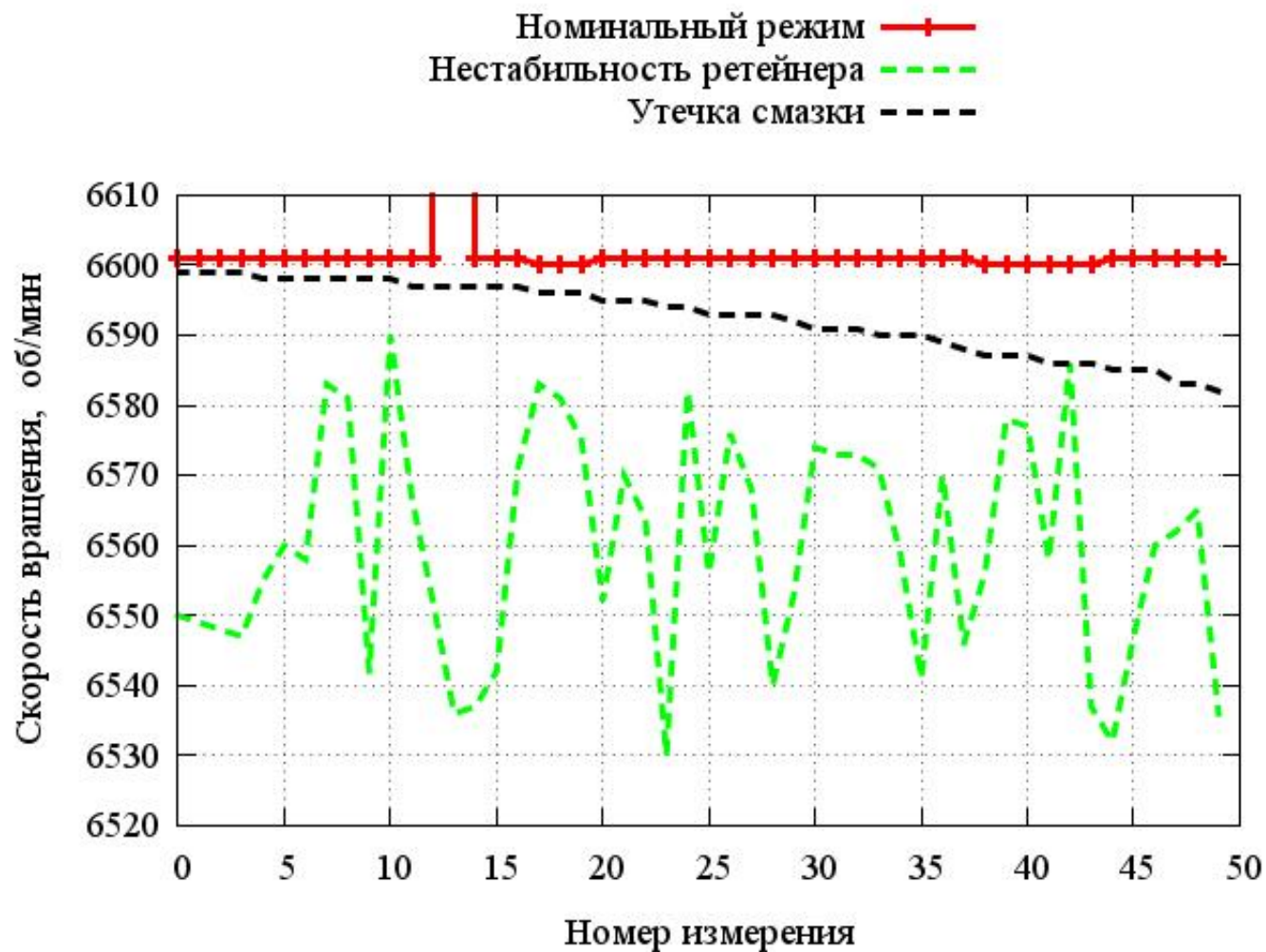
- номинальный;
- гиродин выключен из комплекса;
- утечка смазки;
- нестабильность крепления.

Формат данных телеметрии:

Параметр	Описание	Единица измерения
T	Дата и время измерения	дд/чч:мм:сс
S	Состояние (используется/выключен из комплекса)	—
ω	Скорость вращения ротора	об/мин
I	Ток, потребляемый двигателем ротора	А
a	Вибрация	g м/с ²
t_b	Температура на креплении ротора	°С
t_h	Температура на датчике скорости вращения	°С

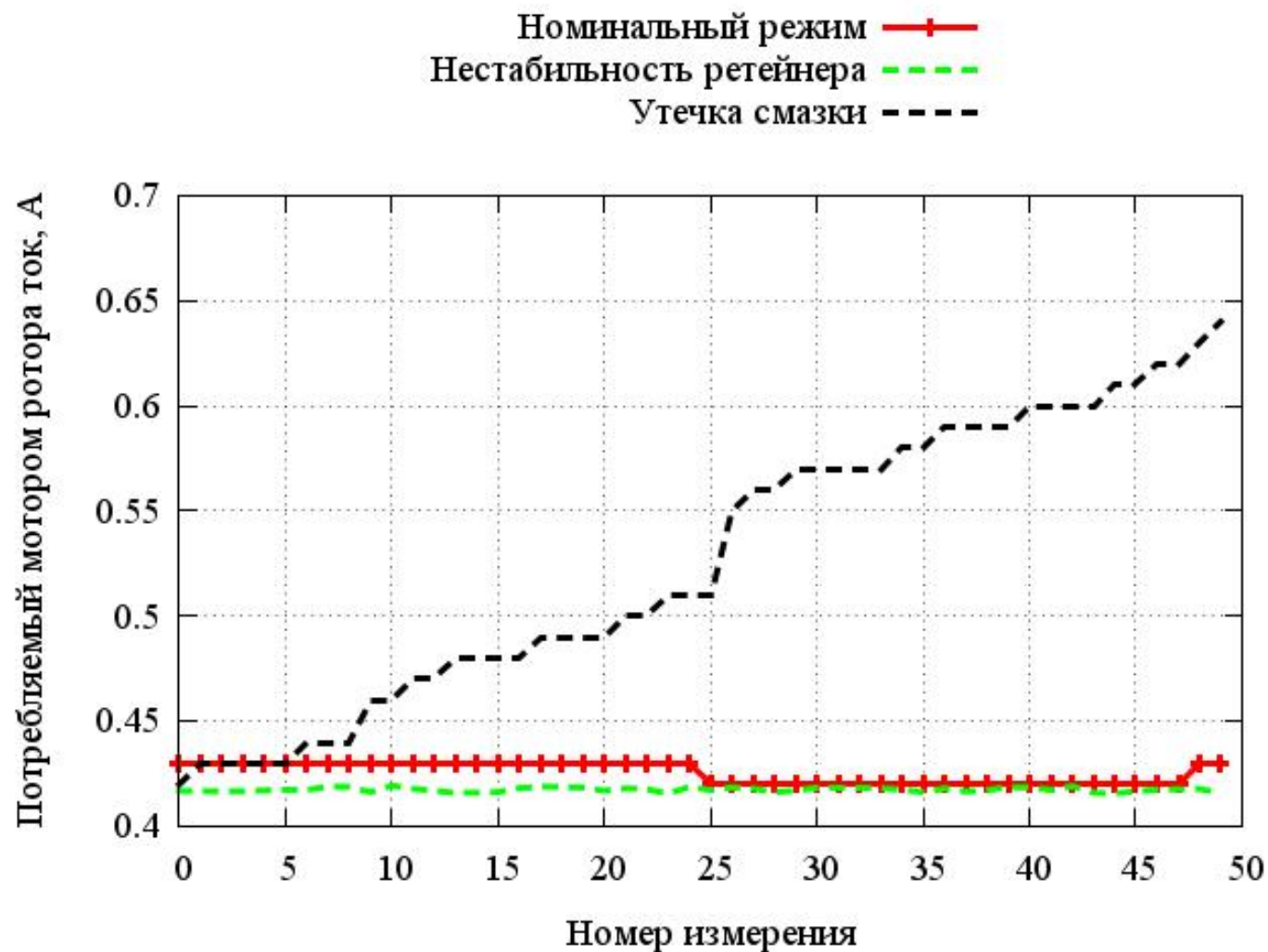
Тестирование

Обучающие выборки: скорость вращения ротора



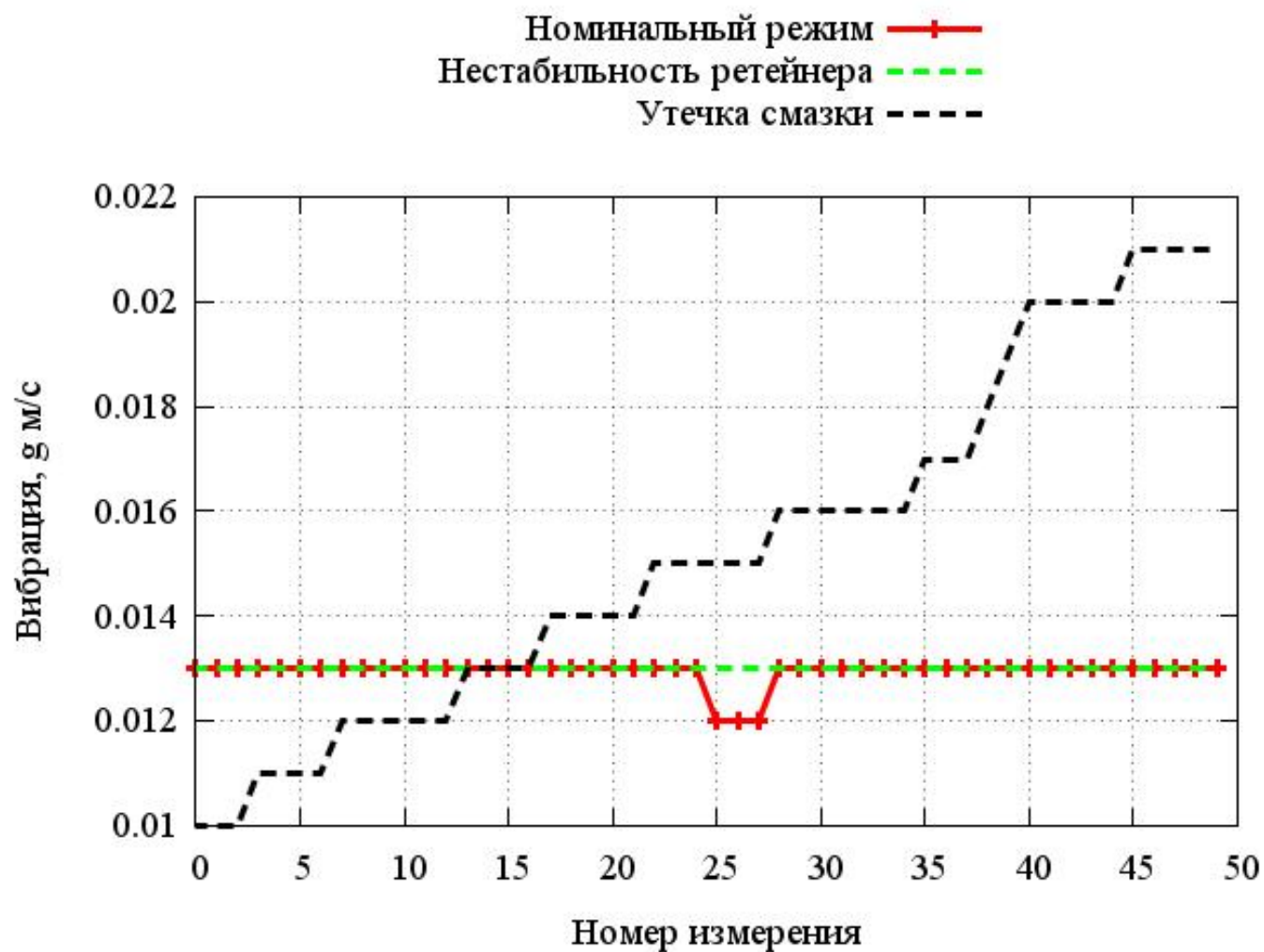
Тестирование

Обучающие выборки: потребляемый мотором ротора ток



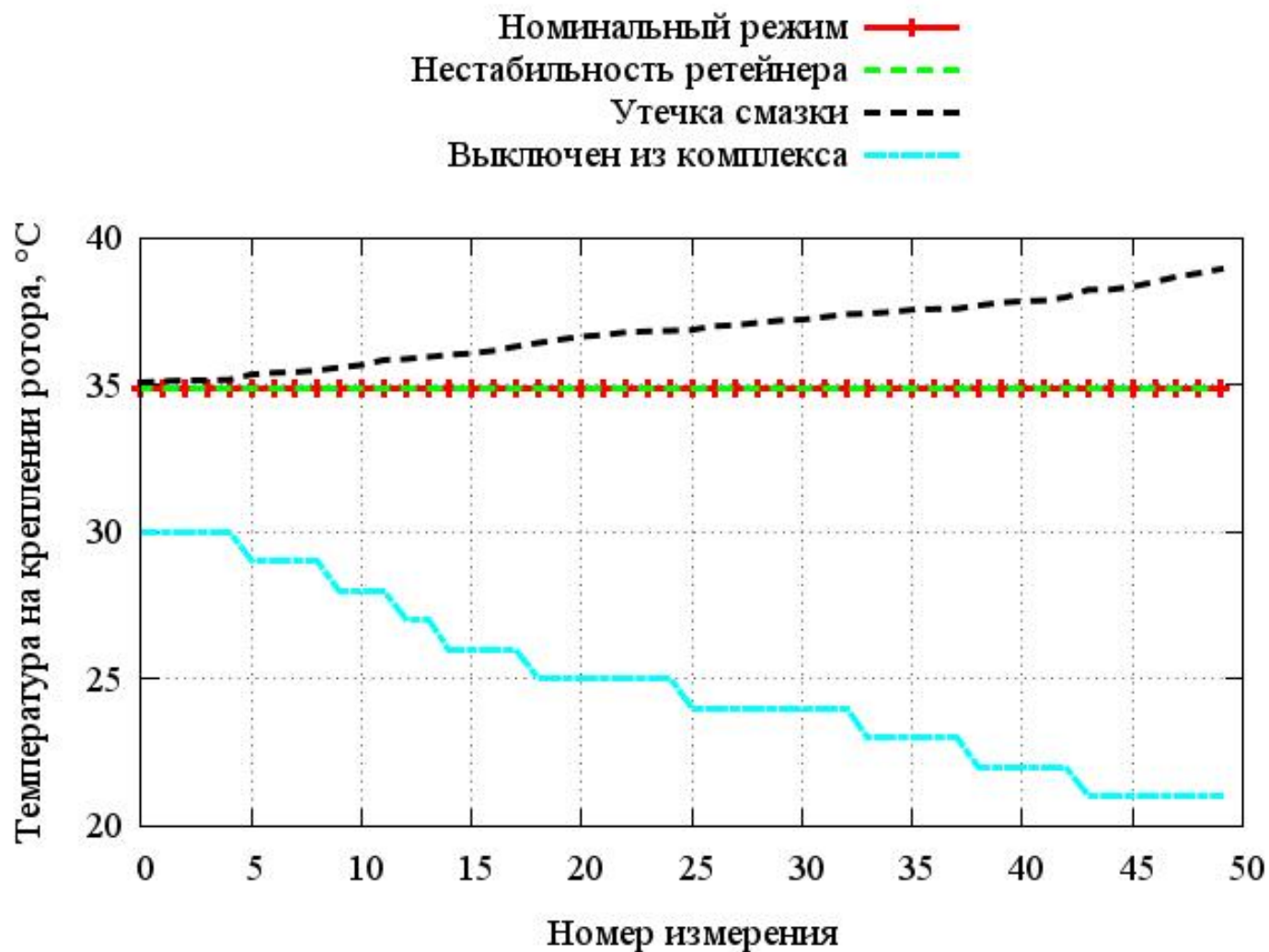
Тестирование

Обучающие выборки: вибрация



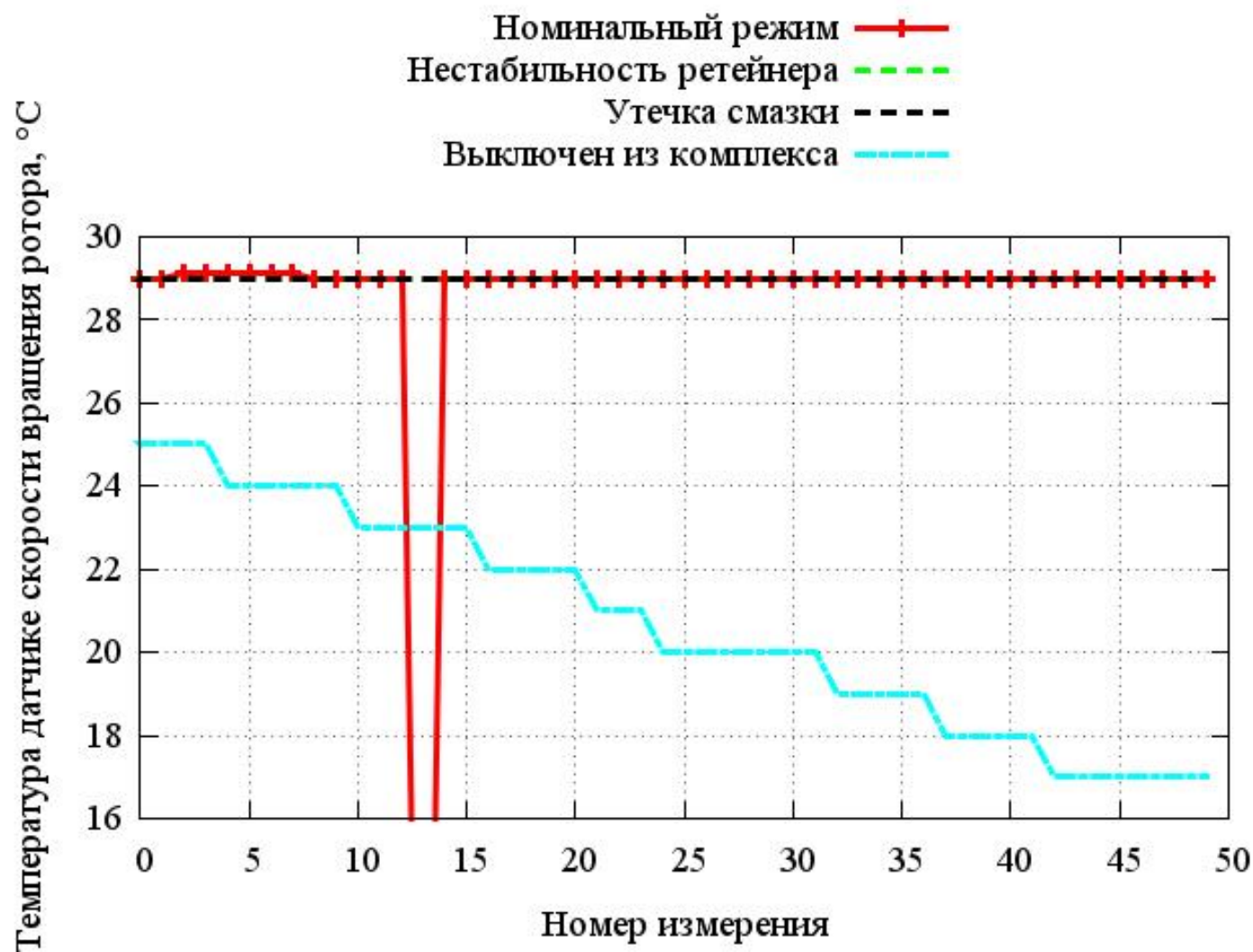
Тестирование

Обучающие выборки: температура на креплении ротора



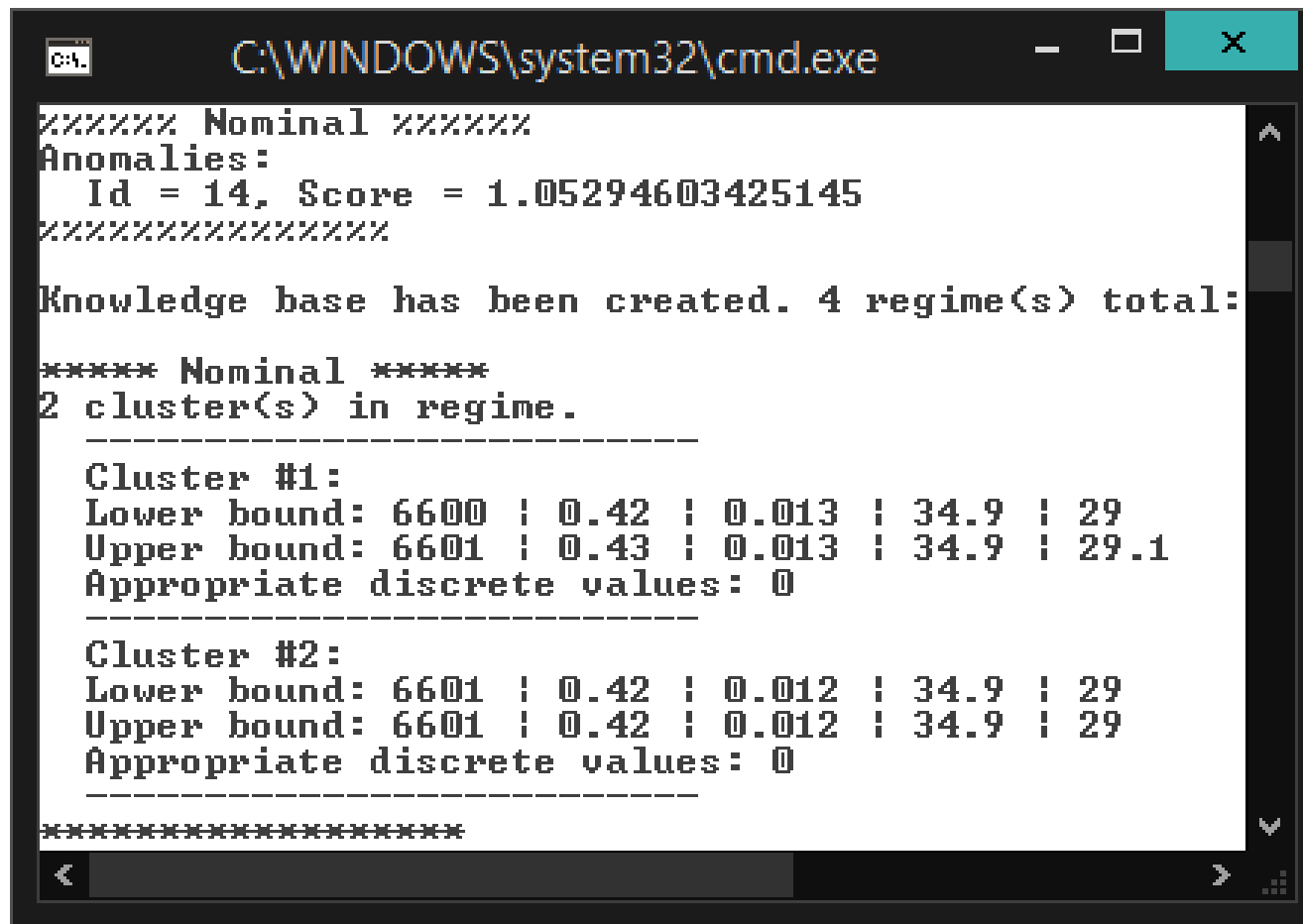
Тестирование

Обучающие выборки: температура на датчике крепления ротора



Тестирование

Результаты работы системы: найденные аномалии и база кластеров для номинального режима работы



```
C:\WINDOWS\system32\cmd.exe

##### Nominal #####
Anomalies:
  Id = 14, Score = 1.05294603425145
#####

Knowledge base has been created. 4 regime(s) total:

***** Nominal *****
2 cluster(s) in regime.
-----
Cluster #1:
Lower bound: 6600 | 0.42 | 0.013 | 34.9 | 29
Upper bound: 6601 | 0.43 | 0.013 | 34.9 | 29.1
Appropriate discrete values: 0
-----
Cluster #2:
Lower bound: 6601 | 0.42 | 0.012 | 34.9 | 29
Upper bound: 6601 | 0.42 | 0.012 | 34.9 | 29
Appropriate discrete values: 0
-----
*****
```


Тестирование

Результаты работы системы: базы кластеров для остальных режимов

```
C:\WINDOWS\system32\cmd.exe
***** Disabled *****
10 cluster(s) in regime.

Cluster #1:
Lower bound: 0 | 0 | 0 | 30 | 24
Upper bound: 0 | 0 | 0 | 30 | 25
Appropriate discrete values: 1

Cluster #2:
Lower bound: 0 | 0 | 0 | 29 | 24
Upper bound: 0 | 0 | 0 | 29 | 24
Appropriate discrete values: 1

Cluster #3:
Lower bound: 0 | 0 | 0 | 28 | 23
Upper bound: 0 | 0 | 0 | 28 | 24
Appropriate discrete values: 1

Cluster #4:
Lower bound: 0 | 0 | 0 | 27 | 23
Upper bound: 0 | 0 | 0 | 27 | 23
Appropriate discrete values: 1

Cluster #5:
Lower bound: 0 | 0 | 0 | 26 | 22
Upper bound: 0 | 0 | 0 | 26 | 23
Appropriate discrete values: 1

Cluster #6:
Lower bound: 0 | 0 | 0 | 25 | 20
Upper bound: 0 | 0 | 0 | 25 | 22
Appropriate discrete values: 1

Cluster #7:
Lower bound: 0 | 0 | 0 | 24 | 19
Upper bound: 0 | 0 | 0 | 24 | 20
Appropriate discrete values: 1

Cluster #8:
Lower bound: 0 | 0 | 0 | 23 | 18
Upper bound: 0 | 0 | 0 | 23 | 19
Appropriate discrete values: 1

Cluster #9:
Lower bound: 0 | 0 | 0 | 22 | 17
Upper bound: 0 | 0 | 0 | 22 | 18
Appropriate discrete values: 1

Cluster #10:
Lower bound: 0 | 0 | 0 | 21 | 17
Upper bound: 0 | 0 | 0 | 21 | 17
Appropriate discrete values: 1

*****
```

```
C:\WINDOWS\system32\cmd.exe
***** Retainer instability *****
1 cluster(s) in regime.

Cluster #1:
Lower bound: 6530 | 0.4155455 | 0.013 | 34.9 | 29
Upper bound: 6590 | 0.4193636 | 0.013 | 34.9 | 29
Appropriate discrete values: 0

*****

***** Lubrication starvation *****
5 cluster(s) in regime.

Cluster #1:
Lower bound: 6597 | 0.42 | 0.01 | 35.09 | 29
Upper bound: 6599 | 0.47 | 0.012 | 35.887 | 29
Appropriate discrete values: 0

Cluster #2:
Lower bound: 6597 | 0.48 | 0.013 | 35.954 | 29
Upper bound: 6597 | 0.48 | 0.013 | 36.168 | 29
Appropriate discrete values: 0

Cluster #3:
Lower bound: 6593 | 0.49 | 0.014 | 36.311 | 29
Upper bound: 6596 | 0.51 | 0.015 | 36.874 | 29
Appropriate discrete values: 0

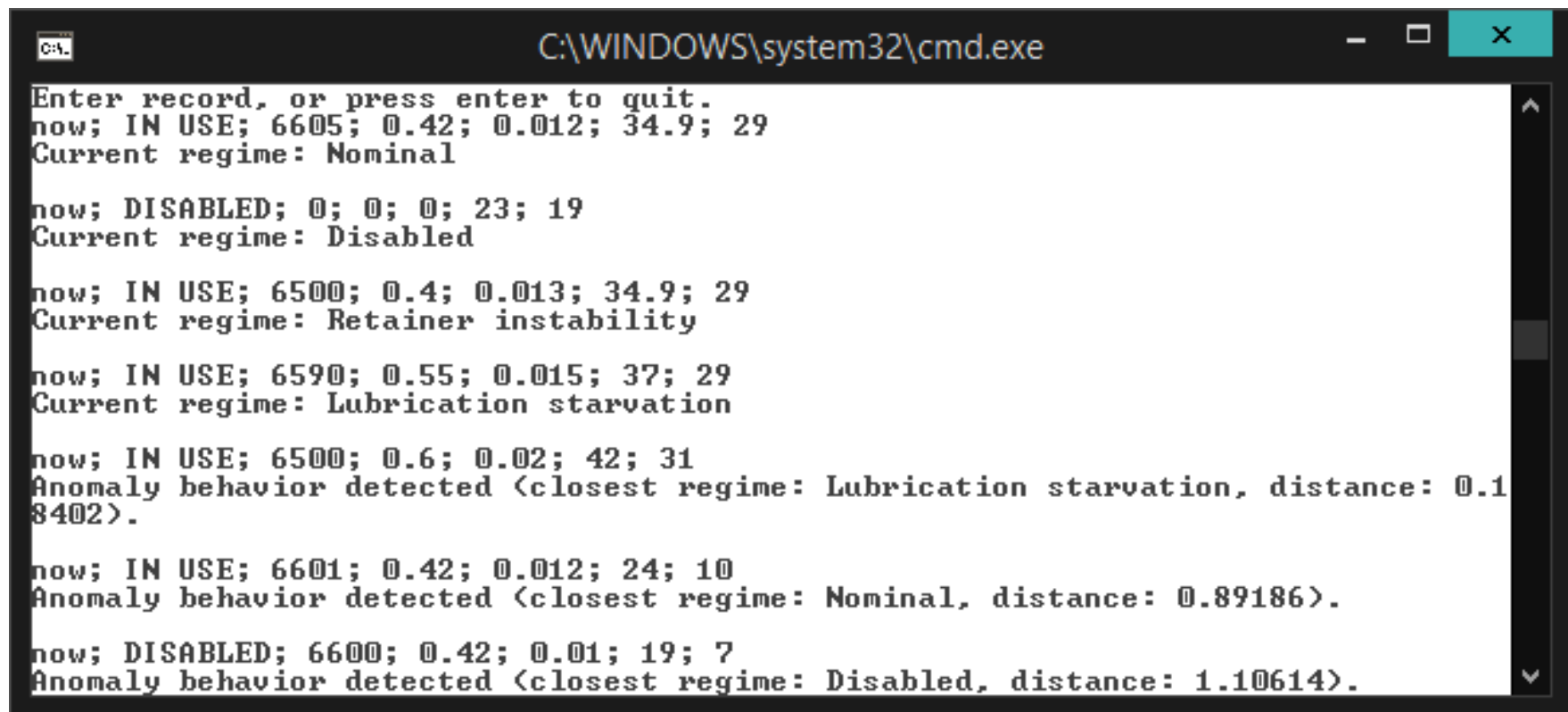
Cluster #4:
Lower bound: 6587 | 0.55 | 0.015 | 37.005 | 29
Upper bound: 6593 | 0.59 | 0.019 | 37.8 | 29
Appropriate discrete values: 0

Cluster #5:
Lower bound: 6582 | 0.6 | 0.02 | 37.851 | 29
Upper bound: 6587 | 0.64 | 0.021 | 38.941 | 29
Appropriate discrete values: 0

*****
```


Тестирование

Результаты работы системы: режим мониторинга

A screenshot of a Windows command prompt window titled "C:\WINDOWS\system32\cmd.exe". The window contains several lines of text representing system monitoring data. The text is as follows:

```
Enter record, or press enter to quit.  
now; IN USE; 6605; 0.42; 0.012; 34.9; 29  
Current regime: Nominal  
  
now; DISABLED; 0; 0; 0; 23; 19  
Current regime: Disabled  
  
now; IN USE; 6500; 0.4; 0.013; 34.9; 29  
Current regime: Retainer instability  
  
now; IN USE; 6590; 0.55; 0.015; 37; 29  
Current regime: Lubrication starvation  
  
now; IN USE; 6500; 0.6; 0.02; 42; 31  
Anomaly behavior detected (closest regime: Lubrication starvation, distance: 0.18402).  
  
now; IN USE; 6601; 0.42; 0.012; 24; 10  
Anomaly behavior detected (closest regime: Nominal, distance: 0.89186).  
  
now; DISABLED; 6600; 0.42; 0.01; 19; 7  
Anomaly behavior detected (closest regime: Disabled, distance: 1.10614).
```

Результаты

- Произведён анализ существующих алгоритмов выявления аномалий без учителя, выделены их достоинства и недостатки.
- Разработан метод диагностики состояния ЛА и его программная реализация, позволяющая строить модель произвольной системы для последующего контроля её состояния только на основе данных телеметрии при различных режимах её работы. Метод не требует априорных знаний о предметной области, назначении системы, её составе, конструкции.
- Программная реализация метода обладает способностью обучаться на больших массивах исходных данных, учитывает как дискретные, так и непрерывные параметры, устойчива к аномалиям в обучающих выборках и может осуществлять контроль даже при отсутствии некоторых параметров в измерениях.
- Разработанное ПО обладает достаточным быстродействием для работы в режиме реального времени.

В результате разработки все требования задания были полностью удовлетворены.



МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)
«МАИ»



Кафедра 308 «Информационные технологии»

Спасибо за внимание!

Дипломная работа

**«Разработка системы мониторинга состояния ЛА (Integrated System Health Management)
на основе методов интеллектуального анализа данных (Data Mining)»**

Выполнил: студент гр. 03-617 Панченко Владимир Владимирович

Руководитель: к.т.н., доцент каф. 308 Пискунов Вячеслав Алексеевич

Москва, 2013 год