

# Разработка статистического метода обработки траекторной информации на основе алгоритма классификации

А. С. Гоголевский, Р. Е. Трепков,  
С. А. Трепкова

Военно-космическая академия  
имени А.Ф. Можайского  
vka@mil.ru

В. А. Соколова

Санкт-Петербургский государственный университет  
промышленных технологий и дизайна  
sokolova\_vika@inbox.ru

**Аннотация.** Рассмотрен статистический метод обработки траекторной информации на основе метода опорных векторов, примененный в рамках одноклассовой классификации и в фазовом пространстве. Предлагаемый метод наиболее точно строит разделяющую функцию при малой обучающей выборке, что позволяет исключить из траекторной информации аномальные измерения и в следствии этого эффективнее производить другие процессы траекторной обработки.

**Ключевые слова:** классификация, радиолокация, траекторная информация

## I. ВВЕДЕНИЕ

Обнаружение и исключение аномальных измерений при обработке траекторной информации является одной из задач статистического машинного обучения, которая заключается в идентификации новых или неизвестных данных [1].

Особенностью задачи обнаружения аномальных измерений является отсутствие информации о классах, которым принадлежат объекты соответствующих наблюдений. С этой точки зрения данную задачу можно рассматривать как задачу классификации, а точнее как одноклассовую классификацию (ОКК), в которой имеется только один класс, а все измерения, не принадлежащие этому классу, являются аномальными. Стоит отметить, что в радиолокационной системе измерения имеют строгую последовательность, зависящую от времени, с определенным шагом измерения. Сами измерения снимаются с устройств (антенные решетки), которые имеют точностную характеристику, т.е. имеют погрешность. Также обучающая выборка имеет малую размерность. Из этого следуют, что каждое измерение имеет неточную величину. В свою очередь, это приводит к тому, что вектор признаков (параметров) является неточным.

Требуется разработать метод обнаружения аномальных измерений устойчивый к неточной (неопределенности) обучающей выборке.

Для учета аномальных измерений было предложено большое количество робастных моделей классификации [2]. Робастность в задачах классификации предполагает, что положение (или координата, значение признака) элемента обучающей выборки на самом деле является неточным и находится в некоторой области. Робастные

модели классификации различаются в зависимости от вида области и ее размера.

Для решения поставленной задачи предлагается модификация метода опорных векторов для ОКК с Гауссовым ядром за счет использования робастной модели  $\epsilon$ -засорения.

## II. ФОРМАЛЬНАЯ ПОСТАНОВКА ЗАДАЧИ ОБНАРУЖЕНИЯ АНОМАЛЬНЫХ ИЗМЕРЕНИЙ

Для наглядности разрабатываемого метода проанализируем радиолокационную систему с одним измеряемым параметром.

Для проведения статистического анализа перенесем результаты измерений на числовую ось. Также выделим множество значений, по которым будем разрабатывать и тестировать метод. Далее данное множество будем называть обучающей выборкой, а все остальные данные будут применены для проверки разработанного метода.

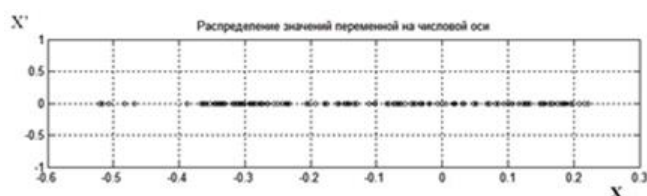


Рис. 1. Обучающая выборка на числовой оси

На рис. 1 представлена обучающая выборка  $n$  длины на числовой оси. Каждое измерение имеет строгую последовательность относительно времени, т.е. измерения происходят с определенным шагом. Значение шага измерения является регулируемой величиной. Здесь применение методов машинного обучения не эффективно, так как при поступлении новых данных велика вероятность ошибки и слишком мала точность измерения. Поэтому требуется разработать правило для получения дополнительной информации об измерениях и проверить его экспериментально.

Основная идея правила заключается в том, чтобы получить дополнительную информацию нужно произвести анализ в фазовом пространстве [3].

Согласно физическому энциклопедическому словарю, фазовым пространством или полным фазовым пространством является геометрический образ, представленный множеством всевозможных состояний

физической системы, наделенных естественным понятием близости.

Основными свойствами фазового пространства являются:

1. Время не является одной из координат фазового пространства.
2. Через каждую точку проходит одна и только одна фазовая траектория.

Сформулируем правило получения дополнительной информации. В общем виде анализируются все параметры. Предполагается, что структура этого процесса следующая:

$$X(t) = x(t) + \varepsilon(t)$$

В этом соотношении  $x(t)$  – истинное значение измеряемого параметра, которое является обычной (неслучайной) функцией времени, а величина  $\varepsilon(t)$  представляет случайную составляющую, которая моделирует искажения (в идеале их можно считать белым шумом).

Основная идея правила заключается в том, чтобы измерения с  $n$ -параметрами в фазовом пространстве дополнить еще и другими параметрами, а именно, приращениями данных параметров.

В результате, фазовое пространство будет представлять собой множество параметров в определенный момент времени и приращения данных параметров.

Исходя из свойств фазового пространства, через каждую точку проходит одна, и только одна фазовая траектория, поэтому фазовое пространство разбивается на непересекающиеся фазовые траектории. Этот геометрический образ – фазовое пространство, заполненное непересекающимися фазовыми траекториями, называется фазовым портретом.

Таким образом, в фазовом пространстве получаем дополнительную информацию об измерениях.

Теперь требуется разработать метод, позволяющий идентифицировать аномальные измерения и учитывающий, что вектор признаков (параметров) является неточным (неопределенным).

### III. Постановка задачи одноклассовой классификации

Анализ литературы [4] показывает, что под задачей обнаружения аномальных измерений имеют в виду задачу одноклассовой классификации.

Сам смысл обнаружения аномальных измерений заключается в идентификации новых данных относительно исходных. Другими словами, предполагается, что данные сконцентрированы в некоторой компактной области, которая называется областью нормальных измерений, а все остальные данные, расположенные на некотором удалении от этой области, считаются аномальными.

По результатам исследований в качестве базового метода был выбран метод (модель) обнаружения аномалий Шелкопфа, в основе которого лежит метод опорных векторов.

В соответствии с работой [5], стандартная модель обнаружения аномальных измерений Шелкопфа подразумевает определение параметрической функции  $f$ , которая принимает значение +1 в небольшой области, покрывающей большинство точек обучающей выборки, и -1 вне этой области. В связи с тем, что в данной работе рассматривается система, в которой функция  $f$  является нелинейной, этот факт существенно затрудняет ее определение и вычисление ее параметров. Поэтому основная идея, заложенная в этой модели обнаружения аномальных измерений или одноклассовой классификации, заключается в отображении точек обучающей выборки в некоторое пространство  $G$  (рис. 2), в котором разделяющая функция  $f$  становится линейной. Такое отображение также называется спрямляющим.

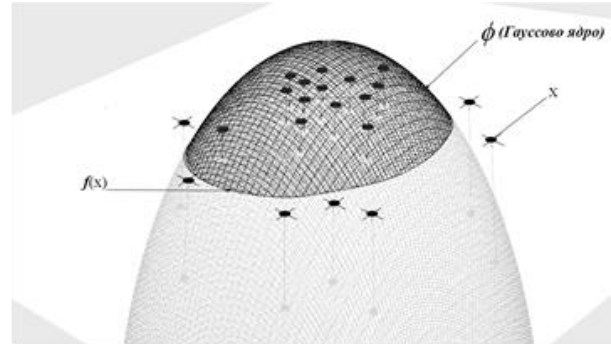


Рис. 2. Отображение точек обучающей выборки в спрямляющем пространстве  $G$

Пусть имеется обучающая выборка  $x_1, \dots, x_n \subset X$ , где  $x_i$  – вектор состояния системы с двумя параметрами: измеряемый параметр и его приращение;  $n$  – количество измерений в обучающей выборке в фазовом пространстве;  $X$  – множество состояний мобильного измерительного комплекса, также его можно представить в виде компактного подмножества  $R^n$ .

Пусть  $\phi$  – такое отображение состояний системы  $X \rightarrow G$ , при котором точки обучающей выборки отображаются в пространство признаков большей размерности  $G$ . В задачах одноклассовой классификации отображение  $\phi$  обычно переводит пространство-прообраз в часть сферы с центром в точке  $0 \in G$ .

Для построения отображения  $\phi$  может использоваться произвольное ядро Мерсера

$$K(\mathbf{x}, \mathbf{y}) = (\phi(\mathbf{x}), \phi(\mathbf{y})).$$

Существуют несколько типов ядер, однако в ряде работ предпочтение отдается, по ряду причин, гауссову ядру, в связи с его эффективностью. Гауссово ядро имеет вид:

$$K(\mathbf{x}, \mathbf{y}) = \exp(-\sigma \cdot \|\mathbf{x} - \mathbf{y}\|^2).$$

Здесь  $\sigma$  – параметр ядра, определяющий геометрическую структуру выборки в пространстве  $G$  и принимающий значение  $\sigma > 0$ .

В работе [5] отмечается, что задача выбора подходящего параметра  $\sigma$  является чрезвычайно важной. Поэтому, выбор параметра  $\sigma$  в данной работе определяется эмпирически.

Данная модификация метода опорных векторов для одноклассовой классификации на основе фазового пространства эффективней, она наиболее точно строит разделяющую функцию  $f$  при малой обучающей выборке. Данный факт был доказан эмпирическим методом.

#### IV. ОБОБЩЕННАЯ СТРАТЕГИЯ ДЛЯ РОБАСТНОГО МЕТОДА

Для учета неопределенности в обучающей выборке используются методы статистической робастности.

Основы робастных методов были разработаны Дж. Тьюки, П. Хьюбера, Ф. Хампеля [6]. Проведя анализ данной работы, можно сделать вывод, что наиболее популярными из них, вследствие их наибольшей эффективности, являются различные модели засорения. В работе предлагается использовать модель  $\varepsilon$ -засорения [7].

Идея данной модели заключается в «размытии» вероятности точек (состояний системы) в соответствии с робастной моделью  $\varepsilon$ -засорения.

Главным отличием от ряда известных робастных моделей классификации, где каждая точка может находиться внутри шара, заключается в том, что в модели рассматривается возможность нахождения вероятности каждой точки в единичном шаре. В этом заключается основная идея построения робастной модели одноклассовой классификации, предлагаемого в работе.

Для построения робастной модели обнаружения аномальных измерений необходимо определить правило или стратегию, по которой выбирается оптимальное распределение из множества  $\mathbf{M}(\varepsilon)$ .

#### V. ОБОБЩЕННАЯ СТРАТЕГИЯ ДЛЯ РОБАСТНОГО МЕТОДА

Одной из распространенных стратегий при принятии решения является стратегия, выбора наихудшего распределения. Данная стратегия получила название минимаксной или пессимистической стратегией.

Согласно этой стратегии, выбирается такое распределение вероятностей из  $\mathbf{M}(\varepsilon)$ , что функционал риска достигает максимума для каждого фиксированного набора  $\mathbf{w}$ ,  $\rho$ . Следует отметить, что оптимальное распределение вероятностей может быть различным для различных параметров  $\mathbf{w}$  и  $\rho$ .

Другой противоположной стратегией является миниминная или оптимистическая стратегия. Согласно которой выбирается распределение вероятностей из  $\mathbf{M}(\varepsilon)$ , минимизирующее функционал риска  $R(\mathbf{w}, \rho)$  для различных параметров  $\mathbf{w}$  и  $\rho$ .

В работе рассмотрена более общая стратегия, являющаяся линейной комбинацией двух представленных ранее стратегий, с некоторым коэффициентом  $\gamma$ , который называется параметром «осторожности». Эта стратегия в литературе называется осторожной [8].

Используя метод опорных векторов для вычисления параметров  $\mathbf{w}$  и  $\rho$ , добавим член регуляризации  $\frac{1}{2}\langle \mathbf{w}, \mathbf{w} \rangle$  к целевой функции и введем новые переменные

$$\xi_i = \max \{0, \rho - \langle \mathbf{w}, \phi(\mathbf{x}) \rangle\}, \quad G = \max_{i=1, \dots, n} \xi_i.$$

Это приводит к задаче квадратичного программирования

$$R_\gamma(\mathbf{w}_{\text{opt}}, \rho_{\text{opt}}) = \frac{1}{2} \langle \mathbf{w}, \mathbf{w} \rangle + \frac{(1-\varepsilon)}{n} \sum_{i=1}^n \xi_i + \varepsilon \gamma G - \rho \nu + \varepsilon(1-\gamma) \min_{i=1, \dots, n} \xi_i \rightarrow \min_{\mathbf{w}, \rho, G, \xi_i},$$

при ограничениях

$$\xi_i \geq \rho - \langle \mathbf{w}, \phi(\mathbf{x}) \rangle, \quad \xi_i \geq 0, \quad i = 1, \dots, n, \\ G \geq \xi_i, \quad i = 1, \dots, n.$$

Очевидно, что целевая функция может быть записана как множество локальных целевых функций. Функционал риска определяется как

$$R_\gamma(\mathbf{w}_{\text{opt}}, \rho_{\text{opt}}) = \min_{k=1, \dots, n} R_{\gamma, k}.$$

Таким образом, необходимо решить  $n$  задач оптимизации.

Вместо минимизации целевой функции  $R_\gamma$  можно записать двойственную задачу оптимизации, называемую лагранжианом.

В результате вычислений данная целевая функция не отличается от целевой функции, полученной в стандартном методе опорных векторов с эмпирическим распределением вероятностей. Однако, основное отличие содержится в ограничениях.

В результате опыта в одном случае решается стандартный метод опорных векторов ( $\varepsilon = 0$ ). В другом случае рассматривается полная неопределенность, когда  $\varepsilon = 1$ . В этом случае только одна из точек обучающей выборки с наибольшей функцией потерь определяет решающую функцию.

Оба случая нельзя использовать в разрабатываемом методе обнаружения аномальных измерений. Так как, в первом случае исходные данные полностью определяются, как аномальные измерения, а во втором случае система становится не чувствительной к ошибкам в обучающей выборке. Поэтому, в рамках работы эмпирическим методом решается задача нахождения оптимального параметра  $\varepsilon$ .

В итоге, принято решение использовать крайний вариант «осторожной» стратегии при  $\gamma=1$  для выбора оптимального распределения из множества  $\mathbf{M}(\varepsilon)$ , другими словами, использовать минимаксную стратегию. Это обусловлено тем, что использовать «осторожную» стратегию с неизвестным параметром  $\gamma$  в реальных условиях, по условиям поставленной задачи, не представляется возможным, так как этот процесс выбора оптимального параметра осторожности  $\gamma$  является трудоемким и неоднозначным. Также нельзя использоваться миниминную стратегию, так как система будет полностью не устойчива к ошибкам в исходных данных.

#### VI. ВЫБОР ПАРАМЕТРОВ ГАУССОВА ЯДРА $\sigma$ И $C$

В результате анализа литературы был выбран параметр  $C$  равный 0,8. Была сгенерирована модельная обучающая выборка  $n=100$ , где  $n$  – число измерений.

Были выбраны семь значений  $\sigma$  в диапазоне  $[0,9;0,005]$ : 0,9; 0,3; 0,1; 0,07; 0,05; 0,01; 0,005. Данные значения параметра ядра  $\sigma$  были выбраны с помощью алгоритма кросс-валидации LOO [9].

В табл.1 представлены результаты анализа разработанного метода машинного обучения с помощью алгоритма кросс-валидации LOO. Выбор параметра  $\sigma$  осуществляется по максимальному значению LOO, и принимает значение 0,07 (рис. 3).

ТАБЛИЦА I.

$\sigma$	0,9	0,3	0,1	<b>0,07</b>	0,05	0,01	0,005
ACC <sub>LOO</sub>	0,501	0,545	0,686	<b>0,717</b>	0,69	0,687	0,273

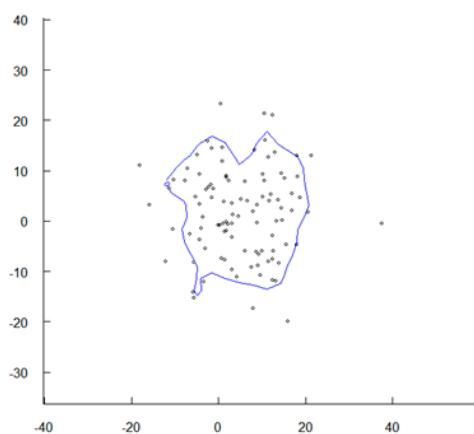


Рис. 3. Визуальное представление решения задачи обнаружения аномальных измерений со значением параметра  $\sigma = 0.07$

В результате исследований были также установлены следующие закономерности:

1. При возрастании  $\sigma$  область нормальных измерений уменьшается и разбивается на множество областей. При максимальных значениях данные области образуют точку. Получается эффект кластеризации;

2. При уменьшении  $\sigma$  область нормальных измерений увеличивается и в итоге охватывает всю область X.

## VII. ВЫБОР ОПТИМАЛЬНОГО ПАРАМЕТРА РОБАСТНОСТИ

В рамках работы осуществлен выбор оптимального параметра робастности  $\varepsilon$ . Также произведено сравнение метода опорных векторов без применения робастной модели и с применением

Окончательное принятие решения о выборе эффективного метода обнаружения аномальных измерений осуществлялось с помощью кросс-валидации LOO.

Была сгенерирована модельная обучающая выборка  $n=100$ , где  $n$  – число измерений. Для метода обнаружения аномальных измерений (метода опорных векторов) были выбраны следующие параметры ядра:  $\sigma=0,07$ ;  $C=0,8$ . Были выбраны шесть значений  $\varepsilon$  в диапазоне  $[0;1]$ : 0; 0,2; 0,4; 0,6; 0,8; 1.

В табл.2 представлены результаты проведенных исследований. В первом столбце представлены значения  $\varepsilon$ , во втором столбце приведены значения относительные коэффициенты LOO с применением метода опорных векторов с модель робастности, а в третьем столбце значение относительного коэффициента LOO метода опорных векторов без модели робастности. Последняя строчка приводит наилучший коэффициент.

ТАБЛИЦА II.

$\varepsilon$	SVM- $\varepsilon$	SVM
0	0,727	0,717
0,2	0,801	-
<b>0,4</b>	<b>0,856</b>	-
0,6	0,701	-
0,8	0,486	-
0,9	0,201	-
<b>LOO<sub>max</sub></b>	<b>0,856</b>	0,717

По результатам исследований наиболее эффективным методом обнаружения аномалий по LOO является робастный метод с параметром равным 0,4.

## VIII. ЗАКЛЮЧЕНИЕ

В данной работе предложен статистический робастный метод обработки траекторной информации на основе метода опорных векторов, примененный в рамках одноклассовой классификации и в фазовом пространстве в условиях неточной обучающей выборке (неопределенности).

Данный метод нуждается в проверке его эффективности на реальных данных. Но предварительно при проверке на модельных данных метод наиболее точно строит разделяющую функцию при малой обучающей выборке и неопределенности в обучающей выборке, что позволяет исключить из траекторной информации аномальные измерения и в следствии этого эффективнее производить другие процессы траекторной обработки.

## СПИСОК ЛИТЕРАТУРЫ

- [1] Жук Ю.А. Робастные модели классификации и их анализ / Информационные системы и технологии: теория и практика. СПб: СПбГЛТУ, 2014. Вып.6. 85 с.
- [2] Андрейчиков А.В., Андрейчикова О.Н. Интеллектуальные информационные системы: учебник. Москва: Финансы и статистика, 2004. 424 с.
- [3] Гоголевский А.С. Метод преобразования данных для обнаружения аномального поведения системы // Материалы международной конференции по мягким вычислениям и измерениям. СПб: ЛЭТИ, 2013. С. 164–166.
- [4] Мерков А.Б. Введение в методы статистического обучения. Москва: Эдиториал УРСС, 2011. 256 с.
- [5] Scott Ferson, Vladik Kreinovich, Lev Ginzburg, David S. Myers, and Kari Sentz, Constructing Probability Boxes and Dempster-Shafer Structures, Sandia National Laboratories, Report SAND2002-4015, January 2003.
- [6] Хампель Ф. Робастность в статистике. Москва: Мир, 1989. 512 с.
- [7] Huber P.J. Robust estimation of a location parameter / Ann. Math. Statist. 1964. № 1. Vol. 35. P. 73–101.
- [8] Уткин Л.В. Анализ риска и принятие решений при неполной информации. СПб.: Наука, 2007. 404 с.
- [9] Mullin M., Sukthankar R. Complete Cross-Validation for Nearest Neighbor Classifiers // Proceedings of International Conference on Machine Learning. 2000. P. 1137–1145.