

# Комплексный подход к анализу данных мониторинга высокопроизводительных вычислительных установок

С.С. Колюхов<sup>1</sup>, А.А. Московский<sup>1</sup>, Е.А. Рябинкин<sup>2</sup>, В.Е. Велихов<sup>2</sup>

Группа компаний РСК<sup>1</sup>, НИЦ «Курчатовский институт»<sup>2</sup>

Рост масштабов современных высокопроизводительных вычислительных систем предъявляет все более высокие требования к автоматизации управления этими вычислительными комплексами. В данной работе рассматривается подход, основанный на использовании современных методов статистического анализа наборов однотипных данных, временных рядов, к анализу показаний численных сенсоров состояния вычислительной системы. При помощи как стандартных, так и вновь разрабатываемых методов удастся автоматически определять режим работы системы (классифицировать поведение), детектировать аномальные состояния, решать задачи визуализации.

## 1. Введение

Современные высокопроизводительные вычислительные комплексы представляют собой сложные системы. Многолетней тенденцией является устойчивый рост среднего числа узлов в высокопроизводительном вычислительном кластере, наблюдаемый, например, как 38% среднегодовой прирост числа процессорных сокетов для машин в рейтинге суперЭВМ «ТОР 500» [1]. В настоящее время эксплуатируются системы с характерным числом узлов 1000-10000. Нарастание числа узлов ведет к усложнению задачи управления (администрирования) вычислительного комплекса.

В настоящее время схема работы службы эксплуатации построена на использовании программных комплексов мониторинга состояния эксплуатируемых установок [2-5]. Программные комплексы собирают и визуализируют информацию о состоянии вычислительного комплекса (ВК).

В большинстве случаев в системах мониторинга кластерных систем отслеживается состояние индивидуальных узлов. Критическое состояние детектируется для индивидуальных компонент либо по превышению пороговых значений (например, перегрев по превышению температуры), либо по достаточно долгому отсутствию обновления данных. В то же время для современного оборудования количество показателей, определяемых на каждом отсчете, может достигать порядка 100, что делает задачу определения пороговых значений и детектированию аномалий достаточно нетривиальной.

Современные системы мониторинга масштабируемы и легко настраиваемы под нужды каждой конкретной установки. В то же время, в случае нештатной ситуации, решение о начале каких-либо действий по исправлению принимается на основе анализа данных мониторинга оператором. При быстром развитии событий скорость реакции человека может оказаться недостаточной, равно как и трудно обеспечить выбор верного варианта поведения в нестандартной ситуации. За счет большой выделяемой тепловой мощности, современные установки могут получить серьезные повреждения при отказе инженерных подсистем. Например, в случае отказа или падения мощности подсистемы охлаждения, на принятие решения об остановке вычислительных узлов остается обычно несколько минут. С наблюдаемым быстрым увеличением масштаба установок необходимость в автоматизации процесса принятия решений по управлению вычислительными комплексами будет только расти.

Систему жизнеобеспечения вычислительного кластера можно рассматривать как сложную динамическую систему, изменение состояний которой можно попытаться описать, если в качестве динамических параметров определить показания сенсорных датчиков, осуществляющих мониторинг вычислительного комплекса.

К настоящему времени сложилось два основных подхода к проблеме описания эволюции и асимптотического поведения динамических систем. Первый подход основан на априорном знании законов, лежащих в основе динамики изучаемого явления и описывающих переходы си-

стемы из одного состояния в другое. Примером такого подхода может служить рассмотрение динамической системы как синонима системе дифференциальных уравнений (автономной или нет, в зависимости от того, является ли изучаемая динамическая система стационарной или нет). Успех этого подхода в большой степени зависит от степени адекватности применяемой аналитической модели рассматриваемому явлению

Второй, вероятностный или статистический подход используется в тех случаях, когда отсутствует точная информация о внутренних или внешних причинах, определяющих динамическое поведение изучаемой системы, или же когда взаимосвязи между динамическими параметрами столь сложны, что не представляется возможным построить более-менее адекватную и пригодную в практических вычислениях аналитическую модель. Этот подход позволяет, не вдаваясь в детали описания динамики системы, тем не менее с той или иной степенью успеха описывать динамику системы в целом.

В настоящей работе, рассматривающей динамическое изменение состояний системы жизнеобеспечения ВК, из-за наличия в ней большого числа динамических параметров с крайне разнообразными характеристиками, которые включают в себя как экзогенные факторы вроде температуры или влажности воздуха на выходе из системы охлаждения, напряжения и силы тока, подаваемого на вычислительные узлы, пользовательской нагрузки на ВК и многое др., так и эндогенные факторы: потребляемая вычислительным кластером мощность, характеристики воздушного потока, создаваемого системой кондиционирования, температура и влажность воздуха в межстоечном пространстве и многое др., представляется разумным применение именно второго, статистического подхода. Конечной целью работы является создание набора методик, позволяющего решать практические задачи по автоматическому выделению аномалий в поведении вычислительных комплексов, визуализации состояния ВК для оператора (администратора) установки.

## 2. Теоретическая часть - Метод анализа данных

В нашей работе алгоритм анализа данных мониторинга состоит в построении адекватных статистических моделей: модели, описывающей набор собранных ранее данных, т. н. «множество нормальных состояний» и модели, позволяющей вычислять возможные будущие состояния ВК, модели регрессора, и в применении модели «множества нормальных состояний» для классификации состояний, полученных в модели регрессора.

### 2.1 Представление данных

Состояние вычислительного комплекса в каждый момент времени представляется в виде случайного вектора:

$$S(t) = (r_1(t), r_2(t), \dots, r_N(t)),$$

где  $r_i(t)$  – показания сенсоров установки в момент времени  $t$  ( $i = 1, \dots, N$ ),  $N$  - число сенсоров, осуществляющих мониторинг вычислительного комплекса.

В такой постановке задача анализа данных мониторинга может быть сформулирована как анализ статистических свойств многомерного временного ряда, где под значениями временного ряда в конкретный момент времени понимается вектор состояния, компоненты которого есть набор показаний сенсорных датчиков в данный момент времени.

### 2.2 Регуляризация набора данных

Перед началом работы всего алгоритма данные требуют процедуры предварительной предобработки, т.н. регуляризации, Это связано с рядом причин: нерегулярность, асинхронность поступления данных измерений от отдельных датчиков и появление отдельных выпадающих значений, что мешает построению адекватных статических моделей

Во многих работах проводится регуляризация на основе интерполяции значений таким образом, чтобы измерения были доступны через одинаковые интервалы – такой подход, например, по умолчанию применяется в популярном средстве хранения данных RRDB [8]. В таком случае

возможно исчезновение «пиковых» значений измерений после регуляризации, что может быть в некоторых ситуациях нежелательным.

Регуляризация в нашем случае проводится по следующей схеме:

1) оценивается интервал времени, содержащий показания со всех сенсорных датчиков;  
2) полученный на шаге 1 интервал времени разбивается на заданное пользователем число подинтервалов одинаковой длины;

3) каждому из полученных на шаге 2 значений времени приписывается набор показаний сенсорных датчиков в ближайший слева момент времени, несколько начальных значений времени из исходной нерегуляризованной выборки при этом отбрасываются, если для них не определены все компоненты многомерного вектора состояний ВК.

Дополнительно при построении обеих моделей может проводиться отброс выпадающих значений, которыми считаются все значения, не попавшие в интервал допустимых значений, определяемый индивидуально для каждого типа сенсорных датчиков. Если такой интервал не определен, то отбрасываются те значения,  $z$ -координата которых больше 3.29 (5% уровень значимости для нормального распределения).

С одной стороны, такой фильтр значений позволяет снизить влияние отдельных выпадающих значений на качество прогноза, и наш опыт эксплуатации ВК показывает, что наиболее важным фактором является наблюдение за трендами в показаниях сенсорных датчиков. Вариант предподготовки данных с отбрасыванием выпадающих значений это допускает. Но с другой стороны, именно выявление аномального поведения показаний сенсорных датчиков является основной целью данной работы, поэтому этот вопрос требует дальнейшего исследования.

## 2.3 Анализ временных рядов – методы прогноза

Существует множество методов для представления и анализа временных рядов [6, 7]. Наиболее простыми являются параметрические модели, как линейные, так и нелинейные. К числу первых относятся такие широко используемые методы как векторный метод авторегрессионной модели (VAR) или метод пространства состояний (SS), а также многие другие. К числу вторых можно отнести методы на основе многомерных сплайнов, которые с практической точки зрения представляют определенную сложность как в использовании, так и в трактовке результатов.

В последнее время все больший интерес исследователей стали привлекать т. н. непараметрические методы статистического анализа, наиболее перспективным из которых является построение модели регрессии на основе метода опорных векторов [9].

Мы в своей дальнейшей работе предполагаем исследовать вопрос применимости как параметрического, так и непараметрического подходов. В качестве первого нами была выбрана линейная модель «векторной авторегрессии с экзогенными переменными» (vector-autoregressive model with exogenous variables, VARX), а в качестве второго – метод  $k$ -кратной регрессии на основе метода опорных векторов ( $k$ -fold SVR).

В общем виде модель VARX можно представить в следующем виде:

$$A(L)y_t = \epsilon_t + B(L)x_t,$$

где  $A(L)$ ,  $B(L)$  матричные функции от оператора лага  $L$ ,  $x_t$  и  $y_t$  – экзогенные и эндогенные переменные,  $\epsilon_t$  – вектор ошибок, обладающий определенными заранее заданными свойствами.

Построение модели VARX может осуществляться одним из способов, используемых для построения параметрических моделей, например, либо с помощью стандартного метода наименьших квадратов, либо с помощью метода максимального правдоподобия. На основе модели VARX можно продолжить временной ряд и вычислить прогноз для нескольких следующих временных шагов.

После построения модели VARX происходит оценка ее качества и качества предсказания. Для этих целей используется набор стандартных статистических тестов и информационных критериев.

Основным достоинством непараметрических методов является их универсальность и независимость от априорных знаний. Метод  $k$ -fold SVR заключается в последовательном применении метода SVR к каждой из компонент многомерного вектора. Метод SVR состоит в построе-

нии такой линейной функции, которая наилучшим образом приближала бы значения искомой функции. Для поиска такого приближения используется принцип минимизации структурного риска (SRM principle), т.е. минимизация следующего функционала структурного риска:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n L(y(i), f(x(i), w)),$$

где  $L - \epsilon$  – лосс-функция того или иного вида.

## 2.4 Обнаружение аномального поведения ВК на основе решения задачи об обнаружении выпадающих значений методом SVM

Выявление аномального поведения ВК происходит путем соотнесения результатов прогнозирования или набора данных за последние моменты времени с классом всевозможных состояний ВК, наблюдавшихся в прошлом и признанных экспертами<sup>1</sup> как типичные и соответствующие нормальной работе ВК. Данное множество строится с помощью подхода на основе ядерных методов.

Фундаментальной идеей ядерных методов, к которым относится метод опорных векторов (SVM), является отображение входных данных, заданных в произвольном множестве, т.н. входном пространстве (input space), в такое гильбертово пространство, часто называемое характеристическим (feature space), в котором можно использовать хорошо изученные методы линейной алгебры и евклидовой геометрии [10].

Такое неявное преобразование  $F$  принято называть ядерной функцией (kernel function). Теоретической основой для подобного подхода является теорема Мерсера, утверждающая, что любая непрерывная симметричная полуположительно определенная ядерная функция  $k(x, y)$  может быть выражена как скалярное произведение в некотором евклидовом пространстве подходящей размерности

$$k(x, y) = \langle F(x), F(y) \rangle.$$

Одной из самых простых операций, которые можно выполнить в евклидовом пространстве, является построение плоскости, которая бы отделяла друг от друга точки двух разных множеств, классов. Другими операциями могут быть кластеризация данных или организация их по какому-либо другому принципу.

Было разработано множество различных вариантов метода SVM с учетом специфики решаемой ими задач. Например, изменив постановку классической задачи о бинарной классификации данных методом SVM, авторам [11] удалось создать простой и надежный алгоритм классификации с одним классом, который можно использовать для решения т.н. задачи об определении новизны (novelty detection) или об определении выпадающих значений (outliers detection). В этом методе тестовое изучаемое множество в характеристическом пространстве отделяется гиперплоскостью от начала координат.

Схожим образом в методе описания областей с помощью метода опорных векторов (support vector density description, SVDD) [12] изучаемое множество в характеристическом пространстве заключается в гиперсфере минимального радиуса. Для наиболее распространенного класса ядерных функций, т.н. RBF ядерных функций (radial-based function RBF). Эти два подхода дают схожие результаты.

Также стоит упомянуть перспективный подход на основе ядерных функций, несвязанный с нахождением опорных векторов, т.н. нелинейный анализ главных компонент (kernel PCA). В работе [13] было показано, что при определенных условиях этот метод может давать лучшие результаты в определении выпадающих значений по сравнению с методами на основе метода опорных векторов. Однако к недостаткам этого метода следует отнести его большую вычислительную стоимость, обусловленную вычислением обратных матриц и собственных значений, а также неустойчивость к помехам при обучении алгоритма.

В данной работе для описания «множества нормальных состояний»  $X$  мы использовали классическую задачу о бинарной классификации, где в качестве дополнительного второго

---

<sup>1</sup> в их роли, как правило, выступают операторы систем управления ВК или инженерно-технический персонал, обслуживающий работу ВК

множества было взято центрально-симметричное к множеству  $X$  множество  $-X$ . В такой постановке эта задача эквивалентна задаче о классификации с одним классом [11]. Такой подход позволил нам использовать формулы для расчета апостериорных вероятностей, полученных в работах [14] и [15].

### 3. Программная реализация

Экспериментальное ПО, разработанное и реализованное в рамках данной работы, включает в себя следующие модули:

- модуль сбора данных мониторинга;
- пользовательский интерфейс;
- система обнаружения внештатных ситуаций;
- модуль внешней интеграции.

Модуль сбора данных поддерживает как работу с RRDB из пакета RRDTool, так и средства сбора данных из SNMP источников. Реализована поддержка сбора данных из ПО APC InfrastruXure Central [16] и Netbotz [17], которые осуществляют мониторинг различных инфраструктурных сервисов, таких как прецизионные и межстоечные кондиционеры различных производителей, температурные и влажностные датчики, размещенные в различных точках инфраструктуры залов, распределители питания, индивидуальные компоненты серверной инфраструктуры.

Данные пакеты осуществляют сбор и визуализацию данных как с SNMP-совместимых источников данных, так и с датчиков, доступных по различным промышленным протоколам: Modbus, one-wire и другим, что позволяет получать информацию от широкого количества датчиков универсальным образом.

Пользовательский интерфейс нашего экспериментального ПО реализован в виде веб-приложения, которое предоставляет пользователю возможность удаленно осуществлять загрузку данных мониторинга ВК и запуск алгоритмов их анализа. Модуль реализован с помощью фреймворка Symfony [18], который представляет собой набор библиотек функций и скриптов на языках программирования php и JavaScript, а также html-шаблонов для создания содержимого веб-сайтов.

Модуль обнаружения внештатных ситуаций реализован при помощи существующих библиотек методов статистического анализа из пакета программ R-software [19, 20].

### 4. Первые результаты

Для тестирования разработанного ПО нами были собраны данные мониторинга установки - вычислительного кластера. Всего набор включает данные от более чем 1000 сенсоров. Наблюдаемые характеристики включали:

- температуры процессоров и модулей памяти вычислительных узлов;
- силу тока и рабочие напряжения на шкафных блоках распределителей питания;
- влажность воздуха на входе в шкаф;
- температура воздуха на входе в шкаф;
- характеристики кондиционеров: охлаждаемая мощность, параметры теплоносителя: давление, входная и выходная температура.

#### 4.1 Визуализация набора данных

Показания сенсорных датчиков с точки зрения их статистической обработки представляют собой довольно сложный объект. Временные ряды, соответствующие им, как правило, являются нестационарными с резко и хаотически меняющимися статистическими свойствами. На Рис. 1 представлены два типичных примера таких данных, которые условно можно отнести к двум разным классам статистических данных. На этом рисунке приведены как измеренные данные для зависимости сила тока (в А) на шкафных блоках распределителей питания от времени (в мс), так и полученные с помощью модели регрессора k-fold SVR.

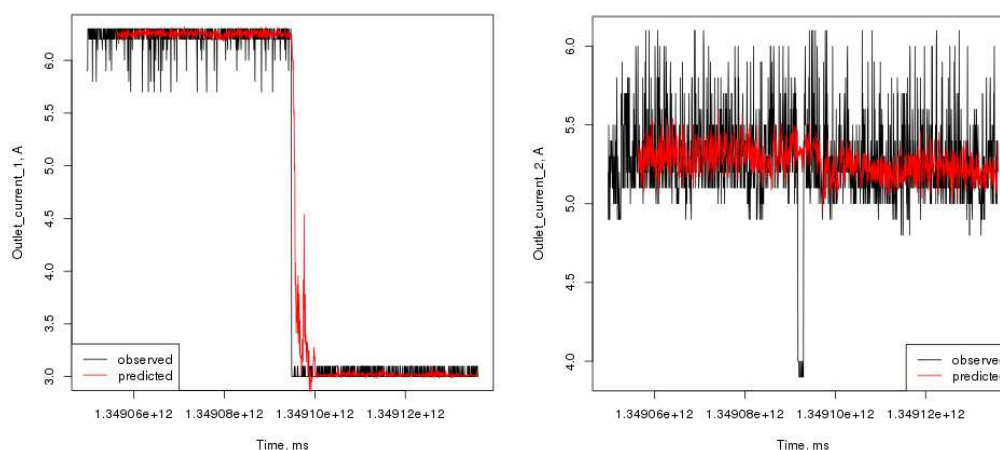


Рис. 1. Типичные зависимости данных мониторинга от времени.

В общем случае для визуализации многомерных данных принято использовать метод анализа главных компонент, имеющий для этого целый ряд оптимальных свойств [21].

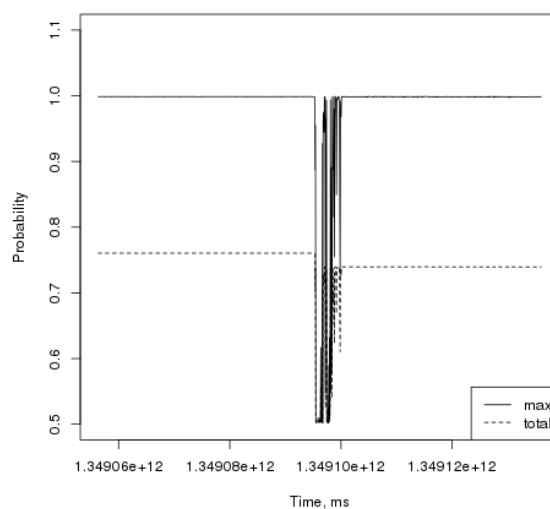
## 4.2 Выделение аномалий

В настоящий момент нами был реализован следующий базовый алгоритм. После регуляризации данных мониторинга происходит построение модели регрессора и осуществление на ее основе прогнозирования будущих возможных состояний ВК. Полученный результат прогноза поступает на вход классификатору, построенному на основе метода SVM, который производит классификацию результата прогноза, т.е. отнесение его к множеству типичных значений, а также вероятности попадания в такое множество, состоящее как правило из нескольких кластеров, каждое из которых обрабатывается отдельно. Обучающими данными этого классификатора служат показания сенсорных датчиков в прошлом.

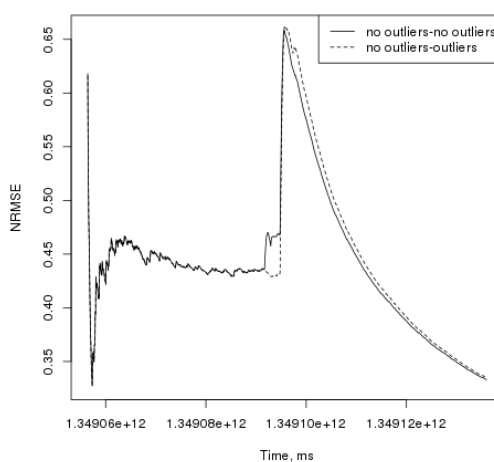
На Рис. 1-3 приведены результаты моделирования для тестовой системы, состоящей из двух сенсорных датчиков силы тока с использованием метода k-fold SVR в качестве модели регрессора (размерность входного вектора = 20, число обучающих точек = 100).

На Рис. 2 представлена полученная зависимость полной и максимальной вероятности попадания в «множество нормальных состояний». Резкое падение значений вероятностей в середине приведенных графиков обусловлена резким изменением режима работы ВК, о чем также свидетельствуют данные, приведенные на Рис. 1.

На Рис. 3 приведены зависимости нормированной среднеквадратичной ошибки NRMSE [22] для различных способов предподготовки данных. Как видно из этого рисунка довольно высокое значение NRMSE возникает в те моменты времени, когда сильно меняется режим работы ВК. Здесь стоит отметить, что использование процедуры отброса выпадающих значений позволяет уменьшить ошибку прогноза. По-видимому наиболее компромиссным вариантом является отбрасывание отдельных выпадающих значений только при построении модели регрессора, именно для этого варианта предобработки данных приводятся результаты на Рис. 1-2.



**Рис. 2.** Зависимости от времени максимальной и полной вероятностей попадания результатов прогноза в «множество нормальных состояний».



**Рис. 3.** Зависимости от времени нормированной среднеквадратичной ошибки прогнозирования. Надписи: no outliers - no outliers и no outliers – outliers соответствуют двум вариантам предобработки данных: без отсева выпадающих значений и с отсевом выпадающих значений при построении модели регрессора.

На тестовом примере видно, что применение специальных статистических методов в принципе позволяет правильно классифицировать состояния вычислительной установки. Момент смены режима работы ВК выявляется довольно однозначно.

В общем случае сложность подобной задачи в первую очередь связана с ее высокой размерностью. Типичное число сенсорных датчиков в современных ВК порядка нескольких тысяч, поэтому можно надеяться, что сведение такой задачи как наблюдение за всеми сенсорными датчиками к наблюдению за одним-двумя параметрами (вероятности попадания множество нормальных состояний», ошибки прогноза) позволит решить проблему мониторинга ВК.

## 5. Заключение

В данном исследовании показано, что ряд стандартных статистических методов анализа и классификации данных в принципе может быть успешно использован для углубленного анализа данных мониторинга высокопроизводительных вычислительных комплексов. В отличие от традиционного подхода, рассматривающего независимо отдельные компоненты системы (вычислительные узлы, агрегаты) в данной работе обоснован комплексный подход, позволяющий

оценить состояние системы в целом с учетом взаимодействия различных компонент друг с другом.

В то же самое время, построенная базовая модель требует дальнейшего тестирования на наборах данных близких к реально используемым в мониторинге ВК с целью определения ее недостатков и их устранения, а также отработки методик работы с ней и трактовки получаемых с ее помощью результатов.

Мы предполагаем, что дальнейшее развитие предложенного подхода позволит разработать эффективные и востребованные средства анализа данных мониторинга ВК. В последующей работе внимание будет уделено не только усовершенствованию моделей анализа данных, но также сбору большего объема данных и развитию программных средств анализа, включая анализ в реальном времени и интерактивные средства взаимодействия с пользователем.

## Литература

1. Dongarra J. Supercomputers and Clusters and Grids, Oh My!  
URL: <http://www.netlib.org/utk/people/JackDongarra/SLIDES/osu-0207.pdf>  
(дата обращения: 12.11.2012).
2. Ganglia Monitoring System.  
URL: <http://ganglia.sourceforge.net> (дата обращения: 12.11.2012).
3. Nagios - The Industry Standard in IT Infrastructure Monitoring.  
URL: <http://www.nagios.org> (дата обращения: 12.11.2012).
4. Clumon: The Cluster Monitor System.  
URL: <http://clumon.ncsa.illinois.edu> (дата обращения: 12.11.2012).
5. Supermon.  
URL: <http://supermon.sourceforge.net> (дата обращения: 12.11.2012).
6. Хенан Э. Многомерные временные ряды. М., Мир, 1974. 576 с.
7. Hannan E. J., Deistler M. The Statistical Theory of Linear System. J. Wiley & Sons, 1988. 380 p.
8. RRDtool.  
URL: <http://oss.oetiker.ch/rrdtool> (дата обращения: 12.11.2012).
9. Sapankevych N.I. Time Series Prediction using Support Vector Machines: A Survey // IEEE Computational Intelligence Magazine. May 2009. P. 25-40.
10. Cristianini N., Schölkopf B. Support Vector Machines and Kernel Methods // AI Magazine. 2002. Vol. 23, No. 3. P. 31-41.
11. Scholkopf B., Platt J.C., Shawe-Taylor J., Smola A.J., Williamson R.C., Estimating the Support of a High-Dimensional Distribution // Neural Computation. 2001. Vol. 13, P. 1443-1471.
12. Tax D.M.J., Duin R.P.W. Support vector domain description // Pattern Recognition Letters. 1999. Vol. 20. P. 1191-1199.
13. Hoffmann H. Kernel PCA for Novelty Detection // Pattern Recognition. 2007. Vol. 40. P. 863-874.
14. Platt J. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods // Smola A.J., Bartlett P.L., Scholkopf B., and Schurmans D., editors. Advances in Large Margin Classifiers. Cambridge, MA. MIT Press, 2000. P. 21-30.
15. Wu T.-F. Probability Estimates for Multi-class Classification by Pairwise Coupling // Journal of Machine Learning Research. 2004. Vol. 5. P. 975-1005.



16. APC's InfraStruXure Central, Change Manager and Capacity Manager Enable Comprehensive Datacenter Management with Predictive Simulation and Modeling.  
URL: <http://www.apc.com> (дата обращения: 12.11.2012).
17. NetBotz - Remote Environmental Monitoring, Early Detection, Instant Alerting, Temperature Sensors, Humidity Sensors, Fluid Detectors, Chemical Sensors, IT Physical Security, Unified Security, Critical Infrastructure Protection.  
URL: <http://www.netbotz.com> (дата обращения: 12.11.2012).
18. High Performance PHP Framework for Web Development – Symfony.  
URL: <http://symfony.com> (дата обращения: 12.11.2012).
19. The R Project for Statistical Computing.  
URL: <http://www.r-project.org> (дата обращения: 12.11.2012).
20. Chang Ch.-Ch., Lin Ch.-J. LIBSVM – A Library for Support Vector Machines // ACM Transactions on Intelligent Systems and Technology. 2011. Vol. 2, No. 3. P. 27:1–27:27.
21. Айвазян С. А. Прикладная статистика. Основы эконометрики. Том 2. М., Юнити-Дана, 2001. 432 с.
22. Vlachos I., Kugiumtzis D. State Space Reconstruction for Multivariate Time Series Prediction.  
URL: <http://arxiv.org/abs/0809.2220> (дата обращения: 12.11.2012).