

## ОБ ИСПОЛЬЗОВАНИИ ПОДХОДОВ БАЗИРУЮЩИХСЯ НА ТЕХНОЛОГИИ DATA MINING

© 2020 И. Я. Львович, Н. Е. Кравцова, Ю. Л. Чупринская

*Воронежский институт высоких технологий (Воронеж, Россия)*

*Работа посвящена рассмотрению возможностей применения методов технологии обработки информации data mining для безопасности в информационных системах.*

*Ключевые слова: data mining, защита информации, моделирование.*

Во многих случаях исследователям приходится сталкиваться с необходимостью проведения анализа по экспериментальным данным.

Для компьютерной обработки их представляют в табличном виде. Таблиц может быть несколько, они взаимосвязаны [1, 2]. Задачи при этом различные:

1. Выявление возможных взаимосвязей среди составляющих.
2. Ведение прогнозирования по новым фактам.

Также могут быть поставлены задачи по проведению оценок заданных признаков, исходя из описания соответствующего набора переменных.

Несколько десятилетий назад в экспериментальных исследованиях использовали методику Р.Фишера. Она применялась в рамках теории разового эксперимента.

Потом на практике все чаще стали сталкиваться с необходимостью анализа разнородных данных [3].

Данные стали содержаться в специальных хранилищах и обрабатываться комплексным образом. Применяется понятие многомерного куба, значительным образом облегчающего извлечение необходимой информации [4, 5].

В технологии Data Mining среди большого количества данных идёт поиск небольшого количества требуемых.

В зависимости от конкретных предметных областей формировались разные

направления внутри данной технологии. Нельзя говорить о существовании универсального подхода или анализе данных.

Тогда исследователи утверждают, что осуществляется исполнение принципа, в котором рассматривается множество моделей [6, 7].

В ходе анализа данных исследователи опираются на статистические модели.

Существуют проблемы, связанные с тем, каким образом сделать выбор алгоритма, который будет в определённом смысле оптимальным, когда извлекаются данные [8].

В моделях, использующих обучение с учителем, исходят из варьирования функций, описывающих модели и остатки.

Модели должны быть такие, в которых прогнозы являются надёжными, точными и несмещёнными.

Модели определяются законами распределения рассматриваемых данных и их типом.

Поиск закономерностей в подходах с учителем связан с задачами кластеризация и организации.

В современных условиях при проведении статистического анализа данных необходимо использовать соответствующие компьютерные программы [9, 10].

Среди некоммерческих проектов в данном направлении можно отметить среду R.

Её в настоящее время применяют многие разработчики.

На рисунке показаны возможности применения технологии data mining для анализа воздействий на сложные технические системы.

---

Львович Игорь Яковлевич – Воронежский институт высоких технологий, доктор, техн., наук, профессор, office@vivt.ru.

Кравцова Нина Евгеньевна – Воронежский институт высоких технологий, студент, kravtrovni@yandex.ru.

Чупринская Юлия Леонидовна – Воронежский институт высоких технологий, студент, chuprimy9ul@yandex.ru.

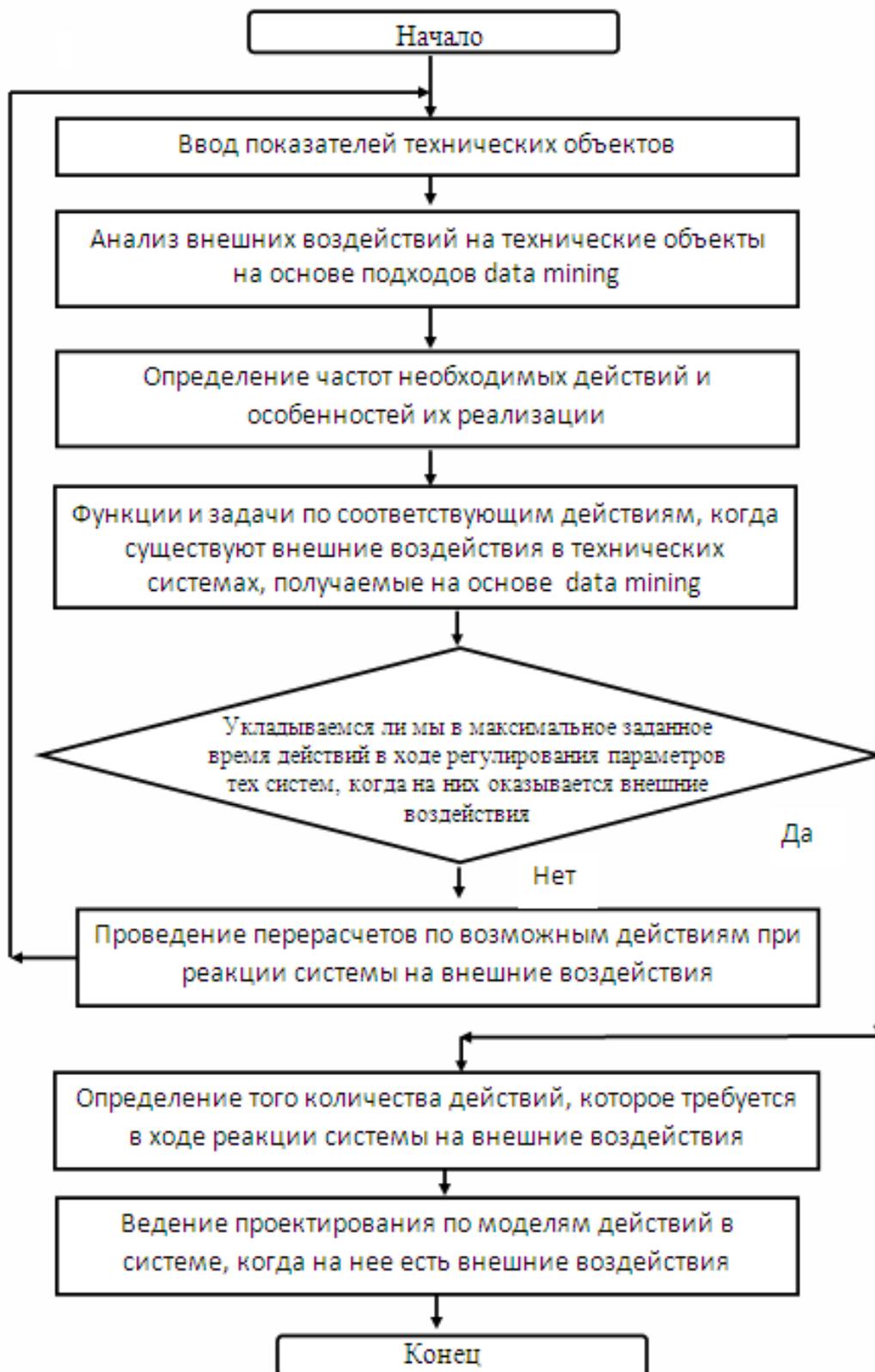


Рисунок. Иллюстрация возможностей анализа воздействий на сложные технические системы при помощи подходов data mining

#### ЛИТЕРАТУРА

1. Львович, И. Я. Основы информатики / И. Я. Львович, Ю. П. Преображенский, В. В. Ермолова. – Воронеж, Издательство: Воронежский институт высоких технологий (Воронеж). – 2014. – 339 с.

2. Львович, Я. Е. Проблемы построения корпоративных информационных систем на основе web-сервисов / Я. Е. Львович, И. Я. Львович, Н. В. Волкова // Вестник Воронежского государственного технического университета. – 2011. – Т. 7. – № 6. – С. 8-10.

3. Преображенский, Ю. П. Проблемы кодирования информации в каналах связи / Ю. П. Преображенский // Современные инновации в науке и технике Сборник научных трудов 8-й Всероссийской научно-технической конференции с международным участием. Ответственный редактор А. А. Горохов. – 2018. – С. 180-182.

4. Кайдакова, К. В. Проблемы защиты информации в современных электронных документах / К. В. Кайдакова // Успехи современного естествознания. – 2012. – № 6. – С. 107-108.

5. Преображенский, Ю. П. Проблемы цифровизации в современном обществе / Ю. П. Преображенский // Инновационные доминанты социально-трудовой сферы: экономика и управление. Материалы ежегодной международной научно-практической конференции по проблемам социально-трудовых отношений. Редакционная коллегия: А. А. Федченко, О. А. Колесникова. – 2019. – С. 243-245.

6. Преображенский, Ю. П. О характеристиках защиты в сетевых информационных

системах / Ю. П. Преображенский // Современные материалы, техника и технология. Сборник научных статей 9-й Международной научно-практической конференции. В 2-х томах. Ответственный редактор А. А. Горохов. – 2019. – С. 160-163.

7. Преображенский, Ю. П. Информационные технологии в системе интернет-вещей / Ю. П. Преображенский // Интеллектуальные информационные системы Труды Международной научно-практической конференции. В 2-х частях. – 2019. – С. 25-26.

8. Гончарова, Н. П. Применение методов аппроксимации данных / Н. П. Гончарова // Успехи современного естествознания. – 2011. – № 7. – С. 97а.

9. Преображенский, Ю. П. О видах информационных систем в организации / Ю. П. Преображенский // Молодежь и системная модернизация страны Сборник научных статей 3-й Международной научной конференции студентов и молодых ученых. В 4-х томах. Ответственный редактор А. А. Горохов. – 2018. – С. 131-134.

## **ABOUT USE OF APPROACHES BASED ON DATA MINING TECHNOLOGIES**

© 2020 I. Ya. Lvovich, N. E. Kravtsova, Yu. L. Chuprinskaya

*Voronezh Institute of High Technologies (Voronezh, Russia)*

*The work is devoted to the consideration of the possibilities of using methods of data mining information processing technology for security in information systems.*

*Key words: data mining, information protection, modeling.*