

Михайлов М.В. Анализ отклонений как метод верификации транзакций в страховании. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей IX Междунар. научно-техн. конф. – Пенза: ПДЗ, 2009. – С. 338-340.

## АНАЛИЗ ОТКЛОНЕНИЙ КАК МЕТОД ВЕРИФИКАЦИИ ТРАНЗАКЦИЙ В СТРАХОВАНИИ

М.В. Михайлов

Ярославский государственный университет им. П.Г. Демидова,  
г. Ярославль, Россия

Показано, что анализ отклонений может решить задачу верификации транзакций в страховании и выявлять случайные и умышленные ошибки в большом информационном массиве. При этом некоторые алгоритмы анализа отклонений идентичны алгоритмам поиска ассоциативных правил.

### **Mikhailov M.V. The analysis of deviations as a method of transactions» verification in insurance.**

It is shown that the analysis of deviations can solve the task of transactions» verification in insurance and can reveal accidental and intentional errors in a big information array.

Развитие информационных технологий приносит изменения в цикл принятия управленческих решений. Этапы сбора, подготовки и анализа данных при этом наиболее заметно претерпевают перемены.

В страховании при обработке больших информационных массивов страховых случаев (запись о страховом случае назовём транзакцией) зачастую требуется выявить нетипичные транзакции, которые в дальнейшем экспертами могут быть расценены как опечатки, мошенничество или начало новой тенденции в системе. В теории интеллектуального анализа данных поиск таких транзакций называется анализом отклонений.

Некоторые алгоритмы интеллектуального анализа данных и анализа отклонений приведены в специальной литературе [1]. Зачастую анализ отклонений осуществим теми же алгоритмами, которые совершают поиск ассоциативных правил, только при других начальных условиях.

Рассмотрим следующие агрегированные величины:

*поддержка* (Support) – отношение количества транзакций, поддерживающих данное правило, к общему количеству транзакций:

$$Supp_{x \rightarrow y} = Supp_F = \frac{|D_F = xuy|}{|D|};$$

*достоверность* (Confidence) – отношение числа транзакций, содержащих наборы X и Y, к числу транзакций, содержащих набор X:

$$Conf_{x \rightarrow y} = \frac{|D_F = xuy|}{|D_x|} = \frac{Supp X \cup Y}{Supp X};$$

*улучшение* (Improvement) – отношение числа транзакций, содержащих наборы X и Y, к произведению количества транзакций, содержащих набор X, и количества транзакций, содержащих набор Y:

$$\text{Im } pr_{x \rightarrow y} = \frac{|D_F = xiy|}{|D_x| \cdot |D_y|} = \frac{\text{Supp}_{xiy}}{\text{Supp}_x \cdot \text{Supp}_y}.$$

Если в случае поиска ассоциативных правил входные значения *Supp*, *Conf* и *Impr* задаются большими для поиска наиболее достоверных ассоциативных правил, то при анализе отклонений эти же величины в этом же алгоритме принимают маленькие значения, чтобы алгоритм выполнял поиск наиболее редких транзакций.

Рассмотрим систему лекарственного страхования, которая в разных формах имеет распространение в некоторых государствах, и выясним возможности применения анализа отклонений. Страховой случай в данной системе происходит при выписке застрахованному в системе лицу допустимого лекарственного препарата врачом, участвующим в системе. При этом процент от стоимости лекарственного препарата (или полная стоимость) возмещается системой. Подобная система для льготных категорий граждан функционирует в России.

Построим пространственную модель данных [2] системы лекарственного страхования для анализа:

1. Факт: выписка и оплата медицинского препарата (при соблюдении определённых условий).

2. Иерархические измерения:

«ВРЕМЯ»: Год -> Квартал -> Месяц -> Декада -> День -> Час,

«ВРАЧ»: Район -> Учреждение здравоохранения -> Врач,

«ПРЕПАРАТ»: Международное непатентованное наименование -> номенклатурный код лекарственного средства,

«АПТЕКА»: Район -> Аптека (аптечный киоск)-> Аптекарь

«ДИАГНОЗ»: Класс МКБ10 -> Блок МКБ10 -> Код МКБ10

3. Простые измерения «СПЕЦИАЛЬНОСТЬ ВРАЧА», «ДОЛЖНОСТЬ ВРАЧА», «ПАЦИЕНТ(район проживания, пол, возраст и др.)» и др.

4. Показатели: «СТОИМОСТЬ ПРЕПАРАТОВ», «КОЛИЧЕСТВО УПАКОВОК», «ВРЕМЯ ЗАДЕРЖКИ ВЫДАЧИ ПРЕПАРАТА».

Каждая транзакция содержит сочетания показателей и значений измерений. Задача работы алгоритма поиска ассоциативных правил и анализа исключений – найти правила вида

ЕСЛИ A1 [ и A2 и A3 ...] ТО B1 [ и B2 и B3 ...]

при разных входных значениях *Supp*, *Conf* и *Impr*

Применение анализа отклонений может решить задачу верификации транзакций в страховании и выявлять случайные и умышленные ошибки в большом информационном массиве. При этом некоторые алгоритмы анализа отклонений идентичны алгоритмам поиска ассоциативных правил.

#### Библиографический список

1. Дюк В., Самойленко А. Data Mining : учебный курс. – СПб. : Питер, 2001. – 368 с.

2. Шаши Ш, Санжей Ч. Основы пространственных баз данных. – М. : Кудиц-образ, 2004. – 330 с.