



$$M(K^{U.D.C.}) = \begin{pmatrix} n_{11} & n_{12} & \dots & n_{1l} \\ n_{21} & n_{22} & \dots & n_{2l} \\ \dots & \dots & \dots & \dots \\ n_{k1} & n_{k2} & \dots & n_{kl} \end{pmatrix}.$$

Элементами n_{ij} этой матрицы являются частота слова в пределах отдельного документа, т.е. $TF(w, d)$. Количество столбцов и строк в этой матрице соответствует количеству слов и документов в корпусе.

Использование словаря синонимов позволяет существенно уменьшить количество элементов матрицы путем наложения частот синонимов, встречающихся в текстах. После суммирования частот синонимов снижение размерности матрицы пропорционально уменьшению количества столбцов.

Обозначив слова-синонимы символом p , получаем:

$$n'_{ij} = \begin{cases} p = p_1 + p_2 + p_3 + \dots + p_n, & \text{если } w_{ij} \text{ имеет синонимы} \\ n_{ij}, & \text{если } p = 0, \text{ синонимы отсутствуют} \end{cases}.$$

Таким образом, объединение синонимических рядов дает возможность повысить оперативность анализа электронных документов, классифицируемых по УДК.

Список литературы:

1. Мисуно И.С. Поиск текстовой информации с помощью векторных представлений [Текст] / И.С. Мисуно, Д.А. Рачковский, С.В. Слипченко, А.М. Соколов. – Проблемы программирования, 2005, №4 – С. 50–59
2. http://www.ruslang.ru/agens.php?id=text_noss2_title

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ МЕТОДОМ SSA С УЧЕТОМ РИСКА

Чистякова А. А., Шамша Б. В.

Харьковский национальный университет радиоэлектроники

В настоящее время эффективность функционирования предприятий существенным образом зависит от процессов получения, накопления, передачи данных и выявления в них закономерностей, которые существенно помогут построить модель прогнозирования основных технико-экономических показателей. К таким предприятиям следует отнести: банки, медицинские учреждения, предприятия телекоммуникации и связи, страховые компании, метеорологические станции и др.

В докладе основное внимание уделяется вопросам построения модели прогнозирования при помощи разрабатываемой информационной технологии, которая позволит не только оценить закономерности данных, но и формализовано подойти к выбору метода прогнозирования.

В работе отмечается, что на сегодняшний день множество методов построения моделей временных рядов имеют свои предпосылки и предположения их использования. Так, в частности, метод экспоненциального сглаживания требует стационарности ряда, метод ARIMA – стационарности временного ряда при некоторой трансформации, которая выражается взятием разности определенного порядка, методы регрессионного анализа требуют выполнения ряда своих специальных требований. Поэтому возникает проблема



Секция 1. Информационные системы и технологии: опыт создания, модели, инструменты, проблемы

формализованного подхода к выбору того или другого метода для анализа нестационарных временных рядов.

В настоящей работе представлены основные методы трансформации временных рядов, к которым следует отнести взятие разности с целью приведения к классу TS- рядов, логарифмирование для уменьшения дисперсии и др. Однако, существуют временные ряды, которые несмотря на их трансформацию не приводятся к стационарным, поэтому такие методы как экспоненциальное сглаживание, ARIMA использовать нельзя.

В этом случае предлагается использовать метод сингулярного спектрального анализа (англ. Singular Spectrum Analysis - SSA), который не требует задания определенных характеристик ряда и может быть применен к временным рядам различных классов. Метод SSA включает элементы классического анализа временных рядов, многомерного статистического анализа, многомерной геометрии, статистического анализа линейных динамических систем и обработки сигналов. Предложенный метод позволяет идентифицировать структуру временного ряда, разложив его на аддитивные составляющие, такие как тренд, компонента гармоник и случайная составляющая. Выделяют четыре этапа метода: вложение, сингулярное разложение, группировка и восстановление. Достоинством метода SSA является возможность выделения гармонических составляющих с изменяющимися амплитудами и частотами, что выгодно отличает его от методов, в основе которых лежит преобразование Фурье.

В докладе представляется информационная технология построения модели прогнозирования нестационарных временных рядов, которая включает решения ряда задач:

- выбор параметров метода SSA на основании статистических характеристик ряда;
- построение одного или нескольких фазовых портретов ряда с использованием выделенных главных компонент;
- расчет коэффициентов линейно-рекуррентной формулы и коэффициентов корреляции оценок модели;
- определение ошибки модели и ее статистических характеристик;
- построение модели прогнозирования нестационарного временного ряда с оценкой риска.

Доклад посвящен информационной технологии построения модели прогнозирования нестационарных временных рядов с оценкой риска с использованием SSA, ЛРФ, ARCH методов.

СТРУКТУРА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ КАЧЕСТВОМ ПРОЦЕССА ВЫРАЩИВАНИЯ МОНОКРИСТАЛЛОВ ПОЛУПРОВОДНИКОВ

Шевченко И.В.

Кременчугский национальный университет имени Михаила Остроградского

Эффективность реализации бизнес-процессов производства непосредственно связана с оптимизацией технических, технологических и административных решений, что обеспечивает условия получения продукции требуемого качества при минимуме затрат. Фактически это означает, что целям управления качеством подчиняется весь комплекс научно-исследовательских и опытно-конструкторских разработок по созданию и модернизации технологических процессов, технологических установок, их оснащению всех видов, информационному и программному обеспечению и автоматизированному регулированию. Решение этих задач может обеспечить объединение разрозненных мероприятий по управлению производственным процессом в единую систему [1].

Понятие «управление качеством» можно сформулировать как «установление, обеспечение и поддержание необходимого уровня качества продукции при ее разработке, производстве и эксплуатации или потреблении, осуществляемое путем систематического контроля качества и целенаправленного воздействия на влияющие на него условия и