

УДК 004.023

ОБОСНОВАНИЕ ПРОГНОЗИРОВАНИЯ ЦЕН ПОЛИМЕРОВ ПОСРЕДСТВОМ НОВОСТНОГО ПОТОКА

И. А. Черенков

Аспирант*

Контактный тел.: 063-416-54-07

E-mail: igor.cherenkov@gmail.com

С. В. Орехов

Кандидат технических наук, доцент*

Контактный тел.: (057) 707-64-74

E-mail: osv@kpi.kharkov.ua

*Кафедра автоматизированных систем управления
Национальный технический университет «Харьковский
политехнический институт»
ул. Фрунзе, 21, г. Харьков, Украина, 61002

Розглянута специфіка цінового прогнозування. Виділені ключові показники зміни цін. Дана постановка завдачі прогнозування. Наведені можливі методи розв'язку

Ключові слова: прогнозування цін, полімери, інтернет новинний потік

Рассмотрена специфика ценового прогнозирования. Выделены ключевые показатели изменения цен. Дана постановка задачи прогнозирования. Приведены возможные методы решения

Ключевые слова: прогнозирование цен, полимеры, интернет новостной поток

A Price forecasting specific was described. Key indicators of price change were marked out. Problem definition and possible approaches for it solution were given

Keywords: price forecasting, polymers, internet news flow

Введение

Неопределённость в завтрашней цене продукта присуща любому рынку, и её прогнозирование является всё ещё неразрешённой проблемой. Наиболее ярким примером является валютный рынок, для которого постоянно появляются новые системы автоматического прогнозирования цен. Однако, ни одна из них не даёт реальных результатов по причине того, что используемые для прогнозирования цены алгоритмы являются преимущественно модельными: они распознают изменение в цене валюты исключительно по форме тренда и поэтому являются мало объективными [1]. Это связано с тем, что на цену влияет множество факторов и её непосредственное изменение происходит вследствие какого-либо события, т.к. форма тренда не отражает сложность процесса формирования цены, а лишь констатирует её изменение. Это доказывает важность дополнения таких прогнозов экспертным мнением о рынке.

Вместе с тем, возможности экспертного оценивания ограничены человеческим фактором. Эксперт субъективен и не может одновременно охватить все множество причин изменения цен. Кроме того, практически отсутствует обмен информацией между экспертами в силу конкурентной рыночной среды. Все эти причины говорят об актуальности проблемы прогнозирования цены.

Состояние проблемы

Ценовое прогнозирование для рынка полимеров Украины связано с ключевыми субрынками, которыми соответственно являются рынки полиэтилена (ПЭ), полипропилена (ПП), поливинилхлорида (ПВХ) и полистирола (ПС). Основными игроками этого рынка являются производители полимеров, трейдеры (дистрибьюторы полимеров) и переработчики (производители готовых изделий из полимеров).

Для наглядности рассмотрим пример ценового тренда (рис. 1). По оси X отображено время (неделя, месяц, квартал), а по оси ординат цена, например, на ПВХ.

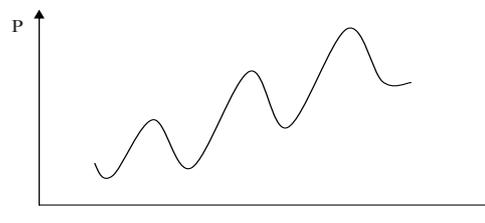


Рис. 1. Примерный график ценового тренда

Основной круг лиц, заинтересованных в прогнозах изменения цен: аналитик (представитель компании по маркетинговым исследованиям), продакт менеджер

(производитель полимера), менеджер по продажам (трейдер), менеджер по закупкам полимера (переработчик). Каждый из них преследует собственные цели, но всех их интересуют одни и те же данные: закономерности поведения тренда, точки перегиба, экстремумы на определённом промежутке времени, список факторов, повлиявших на данное состояние тренда.

Поиск точек перегиба предназначен для идентификации локальных моментов оптимальных для продаж или покупок соответственно. Поиск закономерностей тренда необходим для определения шаблонов поведения цены. Поиск экстремумов, ценовых пиков, необходим для понимания перспектив, границ закономерностей, использования в качестве сигналов значительных перемен на рынке.

Дерево целей, соответствующее интересам этих групп лиц представлено на рис. 2.

Другими словами, необходимо решить часть задачи ситуационного управления, а именно, задачу определения ситуации текущей или полной.

Постановка задачи

Задачу определения текущей ситуации сформулируем следующим образом: на основе системы факторов и значений сигнальных показателей определить, к какому классу принадлежит текущая рыночная ситуация. Для простоты будет рассматривать три класса ситуаций: а) нормальная, когда данная конфигурация факторов и значение сигнальных показателей не приводит к изменению цены; б) «аларм-минус», когда данная конфигурация факторов и значений сигнальных показателей приведет к

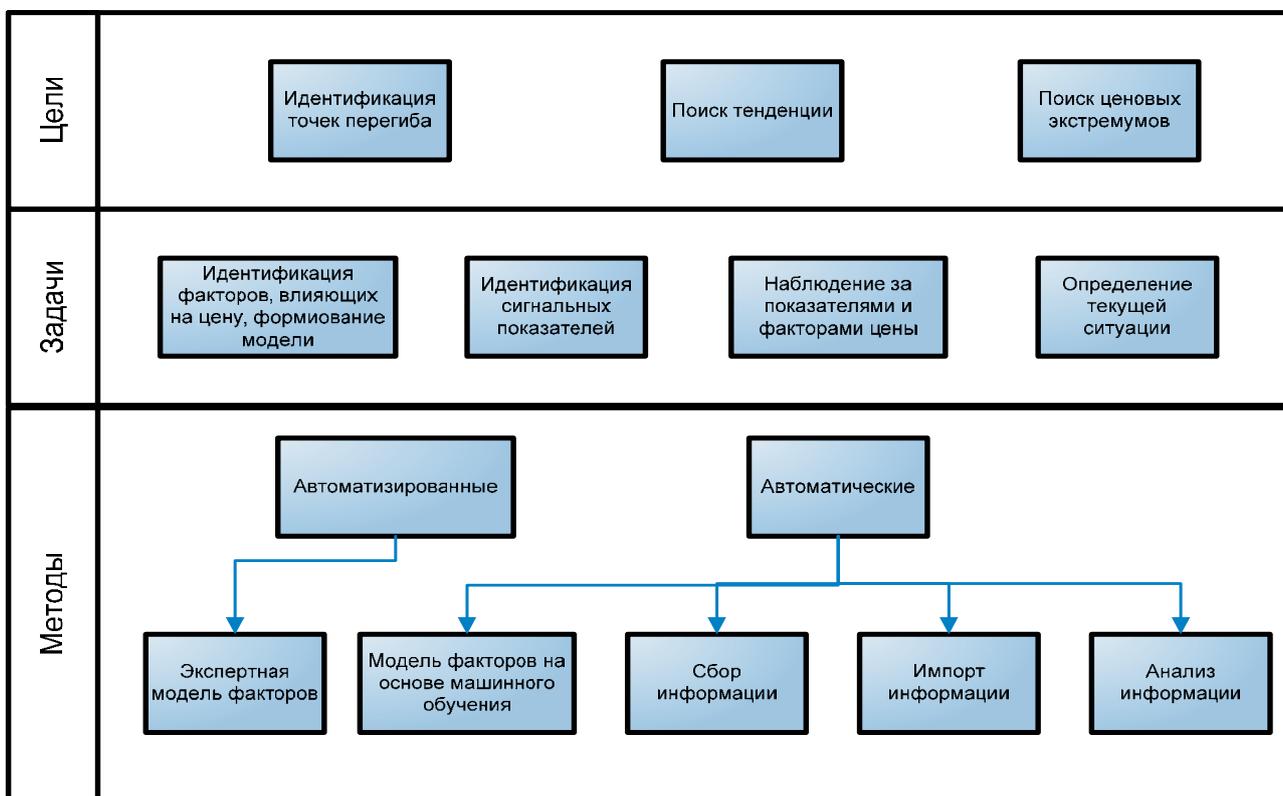


Рис. 2. Дерево целей

Для реализации всех вышеописанных целей необходима реализация схожего набора задач: определить факторы, влияющие на цену, определить сигнальные показатели для своевременного реагирования на изменения цены, осуществлять непрерывное наблюдение за факторами и сигнальными показателями. Факторы в совокупности с сигнальными статистическими показателями помогут построить так называемую систему алармов. «Аларм» (alert) – это событие или ситуация, которая по своим характеристикам выходит за допустимые рамки. Поэтому необходимо построить модель определения таких «алармов».

снижению цены; в) «аларм-плюс», когда данная совокупность критериев приведет к повышению цены на полимер.

Анализ постановки задачи показывает, что ее можно разбить на следующие подзадачи – рис. 3.

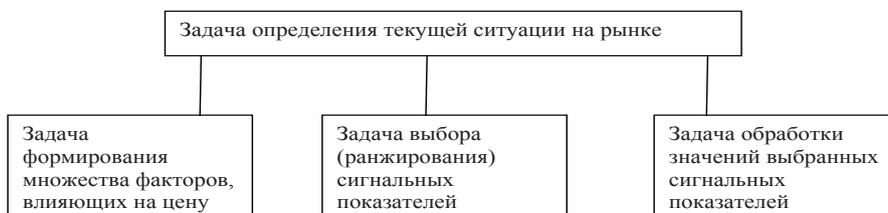


Рис. 3. Дерево подзадач

Во-первых, совокупность факторов, которые возможно могут повлиять на цену полимера, в каждый момент времени разная. Следовательно, существует задача формирования множества факторов.

Во-вторых, необходимо решить задачу выбора (ранжирования) и обработки значений сигнальных показателей из всего статистического потока, который наблюдается на рынке полимеров. Конфигурация показателей также неоднозначно влияет на изменение цены. Поскольку значение одного показателя в совокупности со значениями других может быть полностью «перекрыто». Следовательно, влияние последнего сводится к нулю.

Анализ возможных методов решения

Решать выделенные задачи можно либо автоматизировано, либо автоматически. Очевидно, что автоматизированный подход не является перспективным по ряду причин: ограниченность человеческого восприятия в количестве поступающей информации, невозможность одновременного анализа всех факторов, субъективизм. Альтернативный данному подходу будет использование автоматических методов. На текущем уровне развития информационных технологий не представляет трудности реализовать процесс сбора и импорта информации автоматически. Кроме того, в совокупности с технологиями машинного обучения это позволит уменьшить влияние субъективизма человеческого фактора [2, 3, 4].

Отдельно необходимо выделить проблематику идентификации факторов. Поскольку множество факторов влияющих на цену продукта сильно варьируются (это могут быть и государственная политика, и геополитическая ситуация, и действия конкурентов), то их полный список никогда не будет включён в некоторую экспертную модель.

Необходимо помнить, что на изменение цены влияют непосредственно не сами факторы, а сигналы об изменениях в факторах, т.е. новости. Попытки прогнозирования на основе новостного потока для тех же валютных рынков осуществляют постоянно, однако большинство прототипов в лучшем случае лишь немногим превосходят порог положительного эффекта при случайном выборе, таким образом, сохраняется актуальность данной проблемы [1]. На данный момент полностью автоматический сбор информации о факторах изменения цен на данный момент возможен лишь посредством интернета, т.к. ни один другой источник не обладает такой оперативностью и полнотой [4].

Следует отметить, что и со сбором интернет-новостного потока есть ряд проблем. Все они технического характера и разрешимы посредством индивидуальной подгонки средств сбора каждого из источников, т.е. нет абсолютно универсального средства сбора новостного потока, и для каждого сайта источники новости необходимо будет создать отдельный механизм извлечения новости, функционирование которого уже будет автоматическим [4].

Кроме того, сложной задачей является обработка значений и сигнальных показателей. Например, в качестве такого показателя можно рассматривать цены на нефть. Нефть – это основное сырье для про-

изводства полимеров. Во-первых, сама цена на нефть зависит от такого аморфного множества факторов. Во-вторых, данные о цене на нее постоянно изменяются – возможный шаг изменения одна минута, так как система торгов ведется в режиме реального времени. В-третьих, существуют так называемые граничные значения данного показателя, переходя через которые рынок полимеров может кардинально измениться. Например, если цена на нефть достигнет определенного значения, то производство полимеров в заданной стране станет абсолютно невыгодным. Таким образом, это означает, что какое-то множество игроков рынка просто исчезает. Возникает система показателей с граничными значениями. Но если, например, один показатель достиг своего предела, а десять других еще в норме, то это никакого существенного влияния не окажет. То есть исследовать необходимо именно комбинации различных показателей и их значения. Чтобы реализовать такое исследование необходимо анализировать предысторию. Следовательно, нужен подход к автоматической обработке этих исторических данных (алгоритм, персонал и ИТ). Перспективными в этой ситуации являются технологии добычи и хранилищ данных [2].

Выводы

1. Для прогнозирования цены на рынке полимеров необходимо использовать решение задачи по определению текущей рыночной ситуации на основе новостного потока и системы сигнальных показателей.

2. Учёт многочисленных факторов, влияющих на прогнозирование цены полимеров, возможно с анализом подзадач: а) с выбором и обоснованием множества факторов, влияющих на цену, через поток новостей; б) с обработкой данных о показателях рынка с последующей оценкой комбинации их значений и граничных условий; в) с собственно распознаванием текущей ситуации как «норма», «аларм-минус» или «аларм-плюс».

Литература

1. Чубукова, И. А. Data Mining: учебное пособие / И. А. Чубукова – М.: БИНОМ, 2006. – 382 с.
2. Text Mining Systems for Market Response to News: A Survey [Электронный ресурс] / М.-А. Mittermayer, G. Knolmayer // Institute of Information Systems, University of Bern. – Working paper №184. – 2006. – 17 p. – Режим доступа : <http://www.ie.iwi.unibe.ch/publikationen/beric-hite/resource/WP-184.pdf>.
3. Web Mining: Clustering Web Documents a Preliminary Review [Электронный ресурс] / К. Hammouda // Department of Systems Design Engineering, University of Waterloo. – 2001. – 13p. – Режим доступа : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.4076&rep=rep1&type=pdf>.
4. Ландэ, Д. В. Поиск знаний в Internet [Текст] / Д. В. Ландэ – К. :Диалектика, 2005. – 272 стр.