

УДК: 519.246.8

## Двуслойные интервальные взвешенные графы в оценке рыночных рисков

А. В. Дмитриев<sup>1,а</sup>, Н. В. Марков<sup>2</sup>

<sup>1</sup>Национальный исследовательский университет «Высшая школа экономики»,  
факультет бизнес-информатики,  
Россия, 105187, г. Москва, ул. Кирпичная, д. 33

<sup>2</sup>Открытое акционерное общество «Новолипецкий металлургический комбинат»,  
Россия, 398040, г. Липецк, пл. Металлургов, д. 2

E-mail: <sup>а</sup>a\_v\_dmitriev@mail.ru

Получено 10 сентября 2013 г.,  
после доработки 17 ноября 2013 г.

Данная работа посвящена применению двуслойных интервальных взвешенных графов в прогнозировании нестационарных временных рядов и оценке по полученным прогнозам рыночных рисков. Первый слой графа с интервальными вершинами, формируемый во время первичного обучения системы, отображает все возможные флуктуации системы в отрезке времени, в котором обучали систему. Интервальные вершины второго слоя графа (надстройка над графом первого слоя), отображающие степень ошибки моделируемых значений временного ряда, соединены ребрами с вершинами графа первого слоя. Предложенная модель апробирована на получении 90-дневного прогноза цен на стальные билеты. Средняя ошибка прогноза составила 2,6 %, что меньше средней ошибки авторегрессионных прогнозов.

Ключевые слова: рыночные риски, прогнозирование, нестационарные временные ряды, двуслойные интервальные взвешенные графы

### Double layer interval weighted graphs in assessing the market risks

A. V. Dmitriev<sup>1</sup>, N. V. Markov<sup>2</sup>

<sup>1</sup>National Research University High School of Economics, Faculty of Business Informatics, 33 Kirpichnaya st., Moscow, 105187, Russia

<sup>2</sup>NOVOLIPETSK (parent company), 2 Metallurgov pl., Lipetsk, 398040, Russia

**Abstract.** — This scientific work is dedicated to applying of two-layer interval weighted graphs in non-stationary time series forecasting and evaluation of market risks. The first layer of the graph, formed with the primary system training, displays potential system fluctuations at the time of system training. Interval vertexes of the second layer of the graph (the superstructure of the first layer) which display the degree of time series modeling error are connected with the first layer by edges. The proposed model has been approved by the 90-day forecast of steel billets. The average forecast error amounts 2,6 % (it's less than the average forecast error of the autoregression models).

Keywords: market risks, forecasting, non-stationary time series, double layer interval weighted graphs

Citation: *Computer Research and Modeling*, 2014, vol. 6, no. 1, pp. 159–166 (Russian).

## Введение

В современных рыночных условиях бизнес ежедневно подвергается различного рода рискам, которые негативно влияют на финансовое состояние предприятия. Причина тому — наличие многочисленных факторов вероятностного характера, учет которых часто не представляется возможным из-за незнания закономерностей распределения случайных величин или громоздкости вычислительных процедур. С другой стороны, фактор риска является сильным стимулом для экономии средств и ресурсов предприятия, оптимизации всех бизнес-процессов.

Особое положение среди рисков для большинства предприятий занимают рыночные риски, связанные с ценами на различного рода товары, услуги, финансовые инструменты и многое другое. В связи с этим в условиях кризисного рынка, характеризующегося резкими колебаниями цен, одной из важнейших проблем становится выбор математической модели для оценки, анализа и прогнозирования нестационарных временных рядов.

Нестационарность рынка является наиболее характерным его свойством. Наиболее заметно нестационарность проявляется для волатильности рынка. При прогнозировании цены в условиях кризисного рынка возникает ряд проблем, главная из которых — невозможность адекватно спрогнозировать значения с помощью общепринятых математических моделей из-за резко нестационарных участков ряда.

Существует немало количество статистических и структурных моделей анализа и прогнозирования временных рядов [Jingfei, 2006], однако вследствие малой эффективности большинства из них в условиях нестационарных рынков возникают весьма интересные и нестандартные подходы к прогнозированию. Такими, например, являются модели, основанные на нечеткой логике и цепях Маркова [Song, Chissom, 1993; Chen, 1996; Jilani, Burney, Ardil, 2007]. Основная цель создания как статистических, так и структурных моделей прогнозирования — минимизация ошибки предсказания [Brockwell, Davis, 2002].

Математическая модель, основанная на нечеткой логике, предполагает использование нестандартных способов приближенных вычислений. В такой модели универсальное множество, сформированное из значений временного ряда, принято делить на интервалы. Полученные интервалы представляют нечеткими множествами ( $A_i$ ) с определенными для них логическими отношениями. Недостатком такого подхода является отсутствие вполне определенного, обоснованного для конкретной задачи алгоритма оценки оптимального количества интервалов. Пользователь модели сам определяет, на какое количество интервалов разделить амплитуду колебаний значений временного ряда. Кроме того, длины интервалов разбиения универсального множества влияют на точность прогнозирования [Chen, Hsu, 2004; Huarng, 2004].

Для устранения указанного недостатка нами предложено деление на интервалы универсального множества относительных отклонений между соседними значениями временного ряда, что позволило нам ограничиться от фиксированной амплитуды значений. Кроме того, для разбиения универсального множества мы воспользовались шкалой «1-9», предложенной Т. Саати [Саати, 1993]. Обоснованное использование шкалы «1-9» в условиях нашей задачи позволило разбить универсальное множество на 18 интервалов, что в дальнейшем (после реализации программного продукта) дало приемлемое соотношение длительности расчета, используемого процессорного времени, а также точности моделирования и прогнозирования. Однако четких рекомендаций относительно условия деления на интервалы нет. Отметим, что девятиуровневое ранжирование Т. Саати применял для анализа иерархий в задачах принятия решений, а не для анализа временных рядов.

Следующим этапом моделирования нечетких временных рядов является формирование нечетких логических отношений ( $A_i \rightarrow A_j$ ) и их группировка по схеме

$$\left. \begin{array}{l} A_i \rightarrow A_{j_1} \\ A_i \rightarrow A_{j_2} \\ \dots \end{array} \right\} \Rightarrow A_i \rightarrow A_{j_1}, A_{j_2}, \dots$$

Вид данной группировки логических отношений позволил нам представить отношения между значениями лингвистической переменной, представленные набором интервалов отклонений, в форме ориентированного графа. Ребрами такого графа являются отношения между значениями лингвистической переменной, вершинами — интервалы, что не противоречит определению графа [Татт, 1988].

Для каждой из таких вершин мы однозначно определяем переход путем взвешенной оценки сгруппированных отношений. Таким образом мы определяем взвешенный граф. Обучение системы проводили в два этапа, поэтому ее структура изображается в виде двуслойного интервального взвешенного графа.

Все эти соображения положены в основу нашей математической модели, алгоритм которой представлен ниже.

## Прогнозирование временных рядов двуслойными интервальными взвешенными графами

Основу предложенной модели прогнозирования составляет обучение системы исходя из структурных особенностей временного ряда.

### Первичное обучение системы

#### 1. Определение универсального множества.

Имеем нестационарный временной ряд  $\{F(t_0), F(t_1), \dots, F(t_n)\}$ . Нестационарность ряда обусловлена высокой амплитудой колебаний финансовых показателей нестабильного кризисного рынка, характеризуемого резкими перепадами и скачками цен. Универсальное множество задачи  $\{\Delta F_r(t_1), \Delta F_r(t_2), \dots, \Delta F_r(t_{n+1})\}$  формируем из значений относительных отклонений между соседними значениями временного ряда, которые рассчитываем по формуле

$$\Delta F_r(t_{i+1}) = \frac{F(t_{i+1}) - F(t_i)}{F(t_i)}, \quad i = \overline{0, n}.$$

#### 2. Разбиение универсального множества на интервалы и определение нечетких множеств.

Для разбиения амплитуды колебаний временного ряда на интервалы находим отрицательное по модулю и положительное максимальные относительные отклонения:

$$\Delta F_{\max}^- = \min\{\Delta F_r(t_1), \Delta F_r(t_2), \dots, \Delta F_r(t_{n+1})\}, \quad \Delta F_{\max}^+ = \max\{\Delta F_r(t_1), \Delta F_r(t_2), \dots, \Delta F_r(t_{n+1})\}. \quad (1)$$

Используя (1) шкалу «1–9» Т. Саати разбиваем универсальное множество на 18 интервалов:

$$A_1 = [\Delta F_{\max}^-, \Delta F_{\max}^- - d_1^-], \quad A_2 = [\Delta F_{\max}^- - d_1^-, \Delta F_{\max}^- - 2d_1^-], \quad \dots, \quad A_9 = [\Delta F_{\max}^- - 8d_1^-, 0],$$

$$A_{10} = [0, \Delta F_{\max}^+ - 8d_1^+], \quad \dots, \quad A_{18} = [\Delta F_{\max}^+ - d_1^+, \Delta F_{\max}^+]. \quad (2)$$

В разбиении (2) длины интервалов:

$$d_1^- = \frac{\Delta F_{\max}^-}{9}, \quad d_1^+ = \frac{\Delta F_{\max}^+}{9}.$$

Цифра «9» в знаменателе непосредственно связана со шкалой «1–9» Т. Саати, определяющей степень существенности отклонения значения от нуля. В таком случае значения лингвистической переменной:  $A_1$  (максимальное отклонение),  $A_2$  (очень большое отклонение),

$A_3$  (большое отклонение),  $A_4$  (отклонение выше среднего),  $A_5$  (среднее отклонение),  $A_6$  (отклонение ниже среднего),  $A_7$  (малое отклонение),  $A_8$  (очень малое отклонение),  $A_9$  (минимальное отклонение). Данное определение значений лингвистической переменной непосредственно связано со спецификой решаемой нами задачи: определение степени риска и принятие решения по модели.

Каждому элементу  $\Delta F_r(t_{i+1})$ ,  $i = \overline{0, n}$ , универсального множества присваиваем соответствующий значению интервал  $A_i$ .

### 3. Определение интервального графа.

Определяем структуру орграфа  $G_A = \langle IV_A, E_A \rangle$ , вершинами которого являются полученные интервалы:  $IV_A = \{A_1, A_2, \dots, A_{18}\}$ . Ребра графа задаем логическими отношениями

$$E_A = \{(A_i, A_j) \mid A_i \rightarrow A_j\}. \quad (3)$$

Отношения (3) описывают последовательность относительных отклонений в структуре временного ряда: отклонение на шаге  $(k + 1)$ , которому соответствует интервал  $A_j$ , следует за отклонением на шаге  $k$ , которому соответствует интервал  $A_i$ . Соответственно, получаем ребро  $E_A = \{(A_i, A_j) \mid A_i \rightarrow A_j\}$ , свидетельствующее о том, что на  $(k + 1)$  шаге за интервалом  $A_i$  следует интервал  $A_j$ . Таких упорядоченных последовательностей из пары интервалов во временном ряде может быть несколько, поэтому каждую из них будем именовать переходом. При возможном множестве переходов из одной интервальной вершины в другую соответствующее им ребро в графе будет только одно.

Для определения отношений между соседними значениями  $\Delta F_r(t_j)$  и упрощения расчетов по модели будем использовать сгруппированные отношения  $R_A = \{(A_i, (A_{j_1} \dots A_{j_n})) \mid A_i \rightarrow A_{j_1}, \dots, A_{j_n}\}$ .

### 4. Определение первичных прогнозных значений.

Моделируемые значения временного ряда определяем по формуле

$$\tilde{F}(t_{i+1}) = F(t_i) + F(t_i) \cdot \phi(A_i). \quad (4)$$

В формуле (4)  $A_i$  — интервал на шаге  $t_i$ ,  $\phi(A_i)$  — значение моделируемого отклонения для вершины  $A_i$ .

Величину  $\phi(A_i)$  определяем по формуле

$$\phi(A_i) = A_{mid} \cdot \mu(A_i). \quad (5)$$

В (5)  $A_{mid} = (\bar{A}_1 \dots \bar{A}_{18})$  — вектор-строка средних значений каждого интервала,  $\mu(A_i)$  — значение весовой функции для вершины  $A_i$ .

Весовую функцию в (5) для каждой вершины задаем в следующем виде:

$$\mu(A_i) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{18} \end{pmatrix},$$

где  $\lambda_k = \frac{c_k}{N_i}$ ,  $c_k$  — количество переходов  $A_i \rightarrow A_k$ ,  $N_i$  — количество переходов из интервальной вершины  $A_i$ .

### Вторичное обучение системы

Шаги 1 и 2 алгоритма вторичного обучения аналогичны соответствующим этапам первичного обучения системы. Единственное отличие заключается в использовании универсального множества  $\{\Delta\tilde{F}_r(t_1), \Delta\tilde{F}_r(t_2), \dots, \Delta\tilde{F}_r(t_n)\}$  из относительных отклонений между фактическими и смоделированными значениями временного ряда:

$$\Delta\tilde{F}_r(t_i) = \frac{\tilde{F}(t_i) - F(t_i)}{F(t_i)}, \quad i = \overline{0, n}.$$

Используется разбиение, аналогичное (2), универсального множества на 18 интервалов:  $\{B_1, \dots, B_{18}\}$ .

#### 3. Определение двуслойного интервального взвешенного графа.

Определяем надстройку (второй слой) для графа  $G_A$ , которая состоит из множества интервальных вершин  $IV_B = \{B_1, \dots, B_{18}\}$ . В таком случае связи между вершинами графа первичного обучения  $G_A$  и вершинами надстройки  $G_B$  следующие:

$$E_{AB} = \{(E_{A_i}, B_j) | E_{A_i} \rightarrow B_j\}.$$

Далее для удобства вычислений используем сгруппированные отношения между вершинами слоев:

$$R_{ABi} = \{(E_{A_i}, (B_{j_1} \dots B_{j_n})) | E_{A_i} \rightarrow B_{j_1}, \dots, B_{j_n}\}.$$

#### 4. Определение прогнозных значений.

С учетом двуслойности интервального графа моделируемые значения временного ряда определяем по формуле

$$\tilde{F}(t_{i+1}) = F(t_i) + (\phi(A_i) - \phi(B_i)) \cdot F(t_i), \quad (6)$$

где  $A_i$  — интервал относительного отклонения на шаге  $t_i$ ,  $B_i$  — интервал ошибки на шаге  $t_i$ ,  $\phi(B_i) = B_{mid} \cdot \mu(B_i)$  — значение моделируемого отклонения для вершины  $B_i$ ,  $B_{mid} = (\bar{B}_1 \dots \bar{B}_{18})$  — вектор-столбец средних значений каждого интервала разбиения.

Весовую функцию в формуле (6) для каждой вершины графа задаем в следующем виде:

$$\mu(B_i) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{18} \end{pmatrix},$$

где  $\lambda_k = \frac{c_k}{N_i}$ ,  $c_k$  — количество переходов  $E_{A_i} \rightarrow B_k$ ,  $N_i$  — количество переходов из  $E_{A_i}$ .

Таким образом, предложенную модель прогнозирования можно представить в виде двух слоев: слоя первичного и слоя вторичного обучения системы. Математическая модель прогнозирования представима в виде двуслойного интервального взвешенного графа (ДИВГ) (см. рис. 1).

Первый слой формируется во время первичного обучения и представляет собой ориентированный граф, вершинами которого являются интервалы (интервальные вершины). Данный граф отображает все возможные флуктуации системы в отрезке времени, в котором обучали систему.

Второй слой модели представлен надстройкой над графом первого слоя. Интервальные вершины надстройки показывают степень ошибки моделируемых значений временного ряда.

Данные вершины не соединены между собой, однако соединены ребрами с вершинами графа первого слоя. Эти соединения показывают, каким интервалам принадлежат погрешности моделирования после первичного обучения системы. Таким образом минимизируется погрешность моделирования значений временного ряда, по которым проводилось обучение системы.

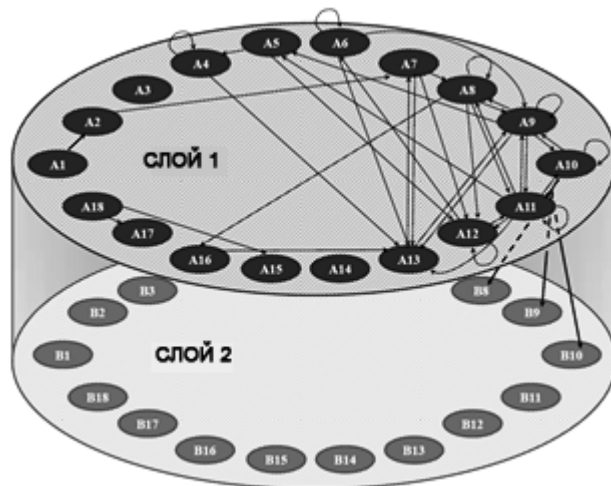


Рис. 1. Двуслойный интервальный взвешенный граф

## Оценка рыночных рисков двуслойными интервальными взвешенными графами

### *Прогнозирование цены на стальные билеты*

Для апробации предложенной математической модели нами получен 90-дневный прогноз цен на стальные билеты по ежедневным отчетным данным с лондонской металлургической биржи LME в период с 24.07.2008 г. по 09.02.2011 г. Данные предоставлены ОАО «НЛМК». Практическая значимость такого прогноза обусловлена необходимостью разработки стратегии хеджирования.

График динамики цен на стальные билеты представлен на рисунке 2, который изображает резко нестационарный временной ряд с большой амплитудой колебаний.



Рис. 2. Динамика цен на стальные билеты

Наибольшая цена за данный промежуток времени зафиксирована в середине 2008 года, после чего следует резкий спад. С четвертого квартала 2008 года наблюдается относительно устойчивое повышение цен на стальные билеты, которое сопровождается значительными локальными скачками цены.

При тестировании данная модель показала следующие результаты: средняя ошибка прогнозирования без слоя минимизации невязки — 3,82 % (однослойный интервальный взвешенный граф — без вторичного обучения); при использовании слоя минимизации невязки — 2,09 % (двуслойный интервальный взвешенный граф).

Результат 90-дневного ДИВГ-прогноза цен на стальные билеты представлен на рисунке 3. Очевидна положительная динамика цен. Через 90 дней цена на стальные билеты составит 546,8 \$ при фактической цене 558,5 \$. В таком случае погрешность прогноза — 2,09 %. С точки зрения экономики 2,09% достаточно большая погрешность (разница между прогнозируемой и фактической ценой 12 \$), однако спрогнозированные значения полностью удовлетворяют тренду фактических значений, результирующее реальное значение не упало ниже прогнозируемого. В случае заключения форвардной сделки на 90-дневный срок ОАО «НЛМК», вероятнее всего, получило бы прибыль.

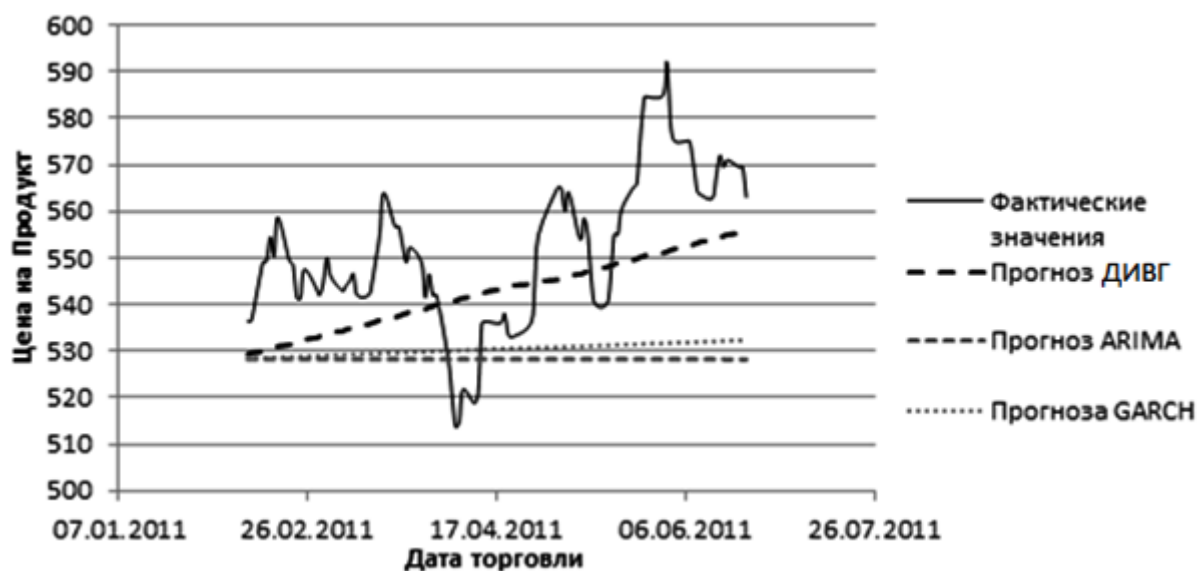


Рис. 3. Фактические и прогнозные значения цен на стальные билеты

### Сравнение точности ДИВГ-прогноза с точностью ARIMA- и GARCH-прогнозов

Следующим этапом апробации предложенной модели является сравнение точности ДИВГ-прогноза с точностью авторегрессионного прогноза цен на стальные билеты методами ARIMA(1,1,1) и GARCH(1,1), определенных на всем объеме выборки, представленном ценами на стальные билеты с 24.07.2008 г. по 09.02.2011 г. Выбор этих методов прогнозирования не является случайным: из всего многообразия математических методов прогнозирования временных рядов ARIMA и GARCH являются наиболее используемыми [Бокс, Дженкинс, 1974; Marenzia, 2008; Alfares, Nazeeruddin, 2002].

В результате построения моделей и прогнозов на 90 дней были получены результаты, представленные на рисунке 3. Сравнение точности прогноза проводили сравнением средней ошибки прогноза для каждой модели. Наилучший результат оценки средней ошибки прогноза показала модель ДИВГ — 2,56 %. Средняя ошибка GARCH-прогноза составила 4,07 %, ARIMA-прогноза — 4,40 %. Следовательно, авторегрессионные модели с заметно меньшей точностью прогнозируют значения нестационарных временных рядов.

## Список литературы

- Бокс Дж., Дженкинс Г. М.* Анализ временных рядов, прогноз и управление. — М.: Мир, 1974. — 406 с.
- Саати Т.* Принятие решений. Метод анализа иерархий. — М.: Радио и связь, 1993.
- Татт У.* Теория графов. — М.: Мир, 1988. — 424 с.
- Alfares H. K., Nazeeruddin M.* Electric load forecasting: literature survey and classification of methods // International Journal of Systems Science. — 2002. — Vol. 33. — P. 23–24.
- Brockwell P. J., Davis R. A.* Introduction to Time Series and Forecasting. 2nd edn. — Springer-Verlag, 2002. — 328 p.
- Chen S. M.* Forecasting enrollments based on fuzzy time series // Fuzzy Sets and Systems. — 1996. — Vol. 81. — P. 311–319.
- Chen S. M., Hsu C. C.* A New Method to Forecast Enrollments using Fuzzy Time Series // Int. Journal Applied Science and Engineering. — 2004. — Vol. 2. — P. 234–244.
- Huang K.* Effective Lengths of Intervals to Improve Forecasting in Fuzzy Time Series // Fuzzy Sets and Syst. — 2004. — Vol. 123. — P. 387–394.
- Jilani T. A., Burney S. M. A., Ardil C.* Fuzzy Metric Approach for Fuzzy Time Series Forecasting based on Frequency Density Based Partitioning. Proceedings of World Academy of Science, Engineering and Technology. — 2007. — Vol. 23. — P. 333–338.
- Jingfei Y. M.* Power System Short-term Load Forecasting // Thesis for Ph. Degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universitat, 2006. — 139 p.
- Marengia D. H.* Forecasting Spot Electricity Market Prices Using Time Series Models // Thesis for the Degree of Master of Science in Electric Power Engineering. Gothenburg, Chalmers University of Technology, 2008. — 89 p.
- Song Q., Chissom B. S.* Fuzzy time series and its models // Fuzzy Sets and Systems. — 1993. — Vol. 54. — P. 269–277.