

АНАЛИТИЧЕСКАЯ СИСТЕМА ИДЕНТИФИКАЦИИ РЫНОЧНОЙ СИТУАЦИИ

Батай И. А. (МИЭМ), Теребулин С. С. (к.э.н., ОАО РАО «ЕЭС России»)

Наметившаяся тенденция расширения рынка ценных бумаг, связанная с активным ростом количества систем интернет-трейдинга и увеличением доли сделок частных инвесторов обуславливает повышение интереса к системам идентификации и прогнозирования ситуации на рынке. Авторами разработана и протестирована предлагаемая аналитическая система, предназначенная в первую очередь для использования неопытным в тонкостях технического анализа мелким инвесторам.

В силу особенностей рыночных отношений можно выделить ряд основополагающих свойств изменения цены актива во времени $C(t)$.

- *Ограниченность.* Цена актива в любой момент времени является положительной величиной, а с другой стороны всегда ограничена сверху (в предельном случае максимальным объемом средств, участвующих в торгах).
- *Цикличность.* Основополагающим свойством всех без исключения товарных и финансовых рынков, не вызывающим полемики среди исследователей, является свойство цикличности [5]. Концепция цикличности утверждает, что цена произвольного рыночного актива $C(t)$ формируется суммой циклов $S_v(t)$, $v = 1, 2, \dots$, отличающихся друг от друга периодами, амплитудами и фазами. Изучению цикличности рыночных цен посвящено много исследований (например, Э. Дьюи [6], Дж. Херст [7]), в результате которых, в частности, было установлено, что для прогностических целей реальной ценностью обладают только так называемые доминирующие циклы, описывающие действующие на рынке тенденции (т.е. рыночную ситуацию). При этом в каждый момент времени на большинстве рынков наблюдается несколько доминирующих циклов, проявление которых искажается непрогнозируемыми возмущениями. Таким образом, цена $C(t)$ может быть записана в виде:

$$C(t) = \sum_v S_v(t) + N(t), \quad (1)$$

где $N(t)$ – случайные возмущения.

Значения как амплитуд, так и периодов входящих в формулу (1) циклов, вообще говоря, неизвестны.

- *Дискретный характер.* Рыночная цена фиксируется в момент заключения сделки и представляет собой множество точек на плоскости Время – Цена. График ценового движения – ломанная, соединяющая все точки в хронологическом порядке.

Перечисленные характеристики позволяют воспользоваться эффективным вспомогательным приемом, часто применяемым при построении аналитических систем – последовательными снятиями направленности [5]. Заложенный в его основу принцип состоит в следующем:

- на первом этапе наблюдаемый процесс $C(t)$ аппроксимируется функцией известного вида $L_1(t)$ и вычисляется их разность $C^1(t) = C(t) - L_1(t)$, представляющая собой некомпенсированные остатки аппроксимации;
- на втором этапе процесс $C^1(t)$, в свою очередь, аппроксимируется функцией $L_2(t)$ и вычисляются некомпенсированные остатки $C^2(t) = C^1(t) - L_2(t)$;
- и т.д.

В результате, после применения последовательности итераций исходный процесс может быть представлен в виде

$$C(t) = \sum_{v=1}^V L_v(t) + C^V(t). \quad (2)$$

В качестве аппроксимирующей функции для максимального устранения негативного влияния случайных возмущений принято использовать механизмы сглаживания ценового ряда. Наиболее известные методы сглаживания связаны с использованием скользящих средних [4]:

– Простая (simple) скользящая средняя

$$\tilde{L}_i(t_n) = \frac{1}{n} \sum_{k=1}^n C_i(t_k)$$

– Взвешенная (Weighted) скользящая средняя

$$\tilde{L}_i(t_n) = \frac{1}{n} \sum_{k=1}^n w_k C_i(t_k), \text{ где } w_k \text{ – весовые коэффициенты.}$$

– Экспоненциальная (exponential) скользящая средняя

$$\tilde{L}_i(t_n) = \tilde{L}_i(t_{n-1}) + (s (C_i(t_n) - \tilde{L}_i(t_{n-1}))), \quad \text{где}$$

s – сглаживающий фактор (smoothing factor). Наиболее распространенным сглаживающим фактором является $s = \frac{2}{n} + 1$.

Основным достоинством метода скользящей средней является его простота. Вместе с тем, данный метод представляется недостаточно точным для целей достоверного прогнозирования. Одним из наиболее современных методов приближенного описания регулярной (неслучайной) составляющей любого наблюдаемого в дискретные моменты времени процесса является метод скользящей линейной регрессии [1, 3, 5]. Результаты его применения при оптимально подобранных параметрах по данным компьютерного моделирования почти в два раза превышают аналогичные показатели использования нерегрессионных скользящих. Результаты анализа, проведенного авторами по результатам торгов на ММВБ за 2001-2002 гг оказались близки по значениям, приводимым в ряде исследований [1] и посвященным анализу торгов на РТС. Методики оценивались на одинаковых наборах исходных данных по критерию максимизации вероятности идентификации направления движения рынка. Исследования, проводимые авторам затрагивали ценовые показатели 65 ценных бумаг (корпоративные акции и государственные облигации). Результаты авторских исследований и представленных в ряде публикаций представлены в таблице 1.

Таблица 1. Результаты анализа различных методик

Метод	РТС (2001)	ММВБ (2002)
Простая скользящая средняя	20%	22%
Взвешенная скользящая средняя	14%	13%
Экспоненциальная скользящая средняя	24%	18%
Скользящая линейная регрессия	42%	47%

Для произвольной функции $C(t)$, представленной своими выборками $C(t_k)$, взятыми в дискретные равноотстоящие друг от друга моменты времени t_k ($k=0, 1, 2, \dots$), линейной регрессией называется линейная функция $\tilde{L}(t, n) = a_n + b_n t$, удовлетворяющая по отношению к функции $C(t)$ критерию наименьших квадратов [3]. Параметр n называют длиной регрессии или размером «окна»; временной интервал, состоящий из используемых при вычислениях моментов времени, называют «окном»; полученную в результате вычислений линейную функцию $\tilde{L}(t, n)$, называют линейной регрессией длины n , в которой коэффициенты a_n и b_n определяются по формулам:

$$\begin{cases} b_n = \frac{A_n B_n - n C_n}{A_n^2 - n D_n} \\ a_n = \frac{B_n - b_n A_n}{n} \end{cases} \quad (3)$$

где

$$\begin{aligned} A_n &= \sum_{k=0}^{n-1} t_k, \\ B_n &= \sum_{k=0}^{n-1} C(t_k), \\ C_n &= \sum_{k=0}^{n-1} t_k C(t_k), \\ D_n &= \sum_{k=0}^{n-1} t_k^2. \end{aligned}$$

Суть метода скользящей линейной регрессии состоит в том, что для выбранного размера «окна» циклически повторяются построения линейной регрессии $\tilde{L}(t, n)$ процесса $C(t)$ и вычисление ее значения в точке t_n с тем отличием, что для каждого последующего построения «окно» сдвигается на один временной интервал. В результате для дискретных моментов времени t_n, t_{n+1}, \dots мы имеем новую, вообще говоря, нелинейную функцию $L(t, n)$, определяемую соотношением $L(t_p, n) = \tilde{L}(t_p, n)$, где $p = n, n + 1, \dots$. В дальнейшем построенную таким образом функцию будем называть *функцией линейной регрессии* длины n процесса $C(t)$.

Именно функции линейной регрессии будут использоваться в дальнейшем в качестве аппроксимирующих функций $L_v(t)$ в методе последовательных снятий направленности. При надлежащем подборе длин функций линейной регрессии каждая из функций $L_v(t)$ будет описывать доминирующий цикл (уменьшение длины выборки будет соответствовать циклам с меньшими периодами). Нескомпенсированный остаток $C^V(t)$ в таком случае будет представлять сумму неучтенных циклов и случайных возмущений.

Правила построения функции линейной регрессии и ограниченность процесса $C(t)$ обуславливают затухание (снижение уровня) нескомпенсированных остатков с ростом числа итераций. Критерием останова итеративного процесса аппроксимации может служить условие неуменьшения остатка, описываемое выражением:

$$\left\| \left(C(t) - \sum_{v=1}^V L_v(t) \right) - \left(C(t) - \sum_{v=1}^{V+1} L_v(t) \right) \right\| < \delta \quad (4)$$

где δ – допустимый уровень погрешности аппроксимации.

Пусть имеется V аппроксимирующих функций $L_v(t)$, представляющих собой функции линейной регрессии длины n_v . При этом достоверным является соотношение

$$C(t) \approx \sum_{v=1}^V L_v(t) \quad (5)$$

Для условия

$$C'(t) \approx \sum_{v=1}^V L'_v > 0, \quad (6)$$

описывающего положительное направление развития рынка, более жестким является требование одновременного выполнения условий неотрицательности сумм пар производных соседних функций (в смысле длины, и следовательно индекса):

$$\begin{cases} L'_{v_1}(t) + L'_{v_2}(t) > 0 \\ L'_{v_3}(t) + L'_{v_4}(t) > 0, \\ \dots \end{cases} \quad (7)$$

где $v_1 < v_2 < v_3 < \dots$.

В свою очередь, понижительный тренд характеризуется выполнением обратного формулам (7) условия, заключающегося в отрицательности хотя бы одной из сумм соседних пар производных:

$$\begin{cases} L'_{v_1}(t) + L'_{v_2}(t) < 0 \\ L'_{v_3}(t) + L'_{v_4}(t) < 0, \\ \dots \end{cases} \quad (8)$$

где $v_1 < v_2 < v_3 < \dots$.

При этом точки на временной оси, в которых происходит изменение условий (7) на (8) и обратно, будем считать моментами соответствующих изменений направления развития рынка – точками сигналов к продаже или покупке.

По парное сложение производных функций линейной регрессии разной длины и сравнение полученных сумм с нулем обеспечивает уменьшение величины задержки формируемых сигналов по отношению к реально существующему на рынке экстремуму. При сложении «длинной» и «короткой» производных вблизи локальных минимумов и максимумов их встречные движения уравниваются (очевидно, что более «длинная» производная дает более поздний сигнал изменения направления развития рынка). Таким способом обеспечивается более высокая точность и своевременность прогноза. С другой стороны, одновременное применение нескольких пар просуммированных производных различной длины, по существу, является фильтром искажающих возмущений.

Значительное повышение характеристик идентификации рыночной ситуации при использовании описанного идентификатора локальных экстремумов объясняется следующими соображениями. Проведенный авторами графический анализ показал, что на рынках ценных бумаг и валют понижительная тенденция (после достижения ценой локального максимума) в большинстве случаев развивается опережающими темпами по сравнению с повышательной (после достижения ценой локального минимума). Как следствие, идентификация понижительной тенденции требует применения более оперативных механизмов. С другой стороны, более длительный и неустойчивый характер зарождения повышательной ценовой тенденции обуславливает необходимость более жесткой фильтрации случайных отклонений и несущественных для модели непродолжительных коррекций рынка.

В дополнение к предложенному механизму фильтрации авторы ввели в модель требование ограничения отклонения прогностической функции от реального значения цены. Сигнал о необходимости совершения операций генерируется системой только в случае выполнения условия смены направления рынков (критерии (7), (8)) и достоверности отношения:

$$\left| C(t) - \sum_{v=1}^V L_v(t) \right| \leq \beta\sigma \quad (9)$$

где $\sigma^2 = \frac{1}{n} \sum_{k=0}^{n-1} \left(C(t_k) - \sum_{v=1}^V L_v(t_k) \right)^2$ – среднеквадратическое отклонение цены актива от его

аппроксимирующей функции;

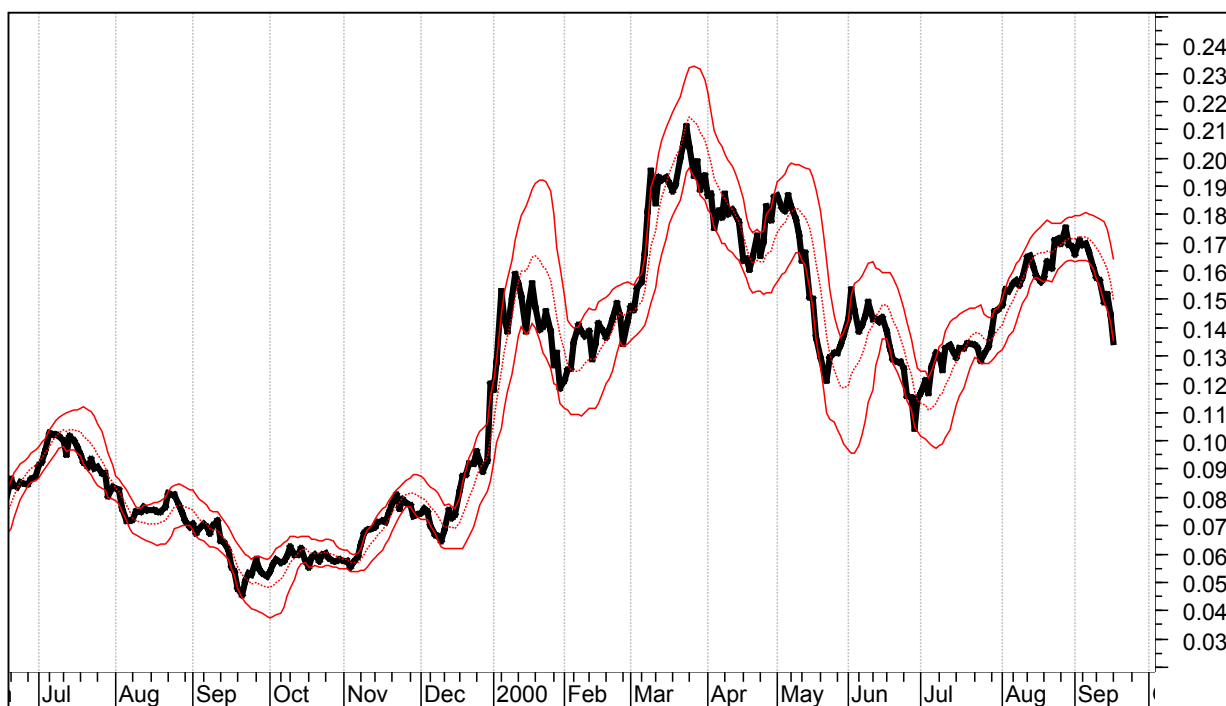
β – коэффициент доверия.

Необходимо заметить, что надлежащим выбором размера «окон» n_v обеспечивается для растущего рынка тяготение цены актива $C(t)$ к линии $\sum_{v=1}^V L_v(t) + \beta\sigma$, называемой верхней

границей коридора, а для падающего рынка – к нижней границе коридора $\sum_{v=1}^V L_v(t) - \beta\sigma$.

Наличие ценового коридора активов отмечается многими авторами (например, Доброчеевым О. В. [1]), что подтверждает необходимости введения дополнительного механизма фильтрации.

Таким образом, мы получили модель идентификации рыночной ситуации, в качестве оптимизируемых параметров которой выступают длины «окон» функций линейной регрессии n_1, n_2, \dots и величина коэффициента доверия β . Сигналом необходимости совершения сделки служит изменение достоверности критериев (7) и (8) при выполнении условия (9). Пример работы системы для котировок акций РАО ЕЭС на ММВБ за 2002 г представлен на рисунке 1.



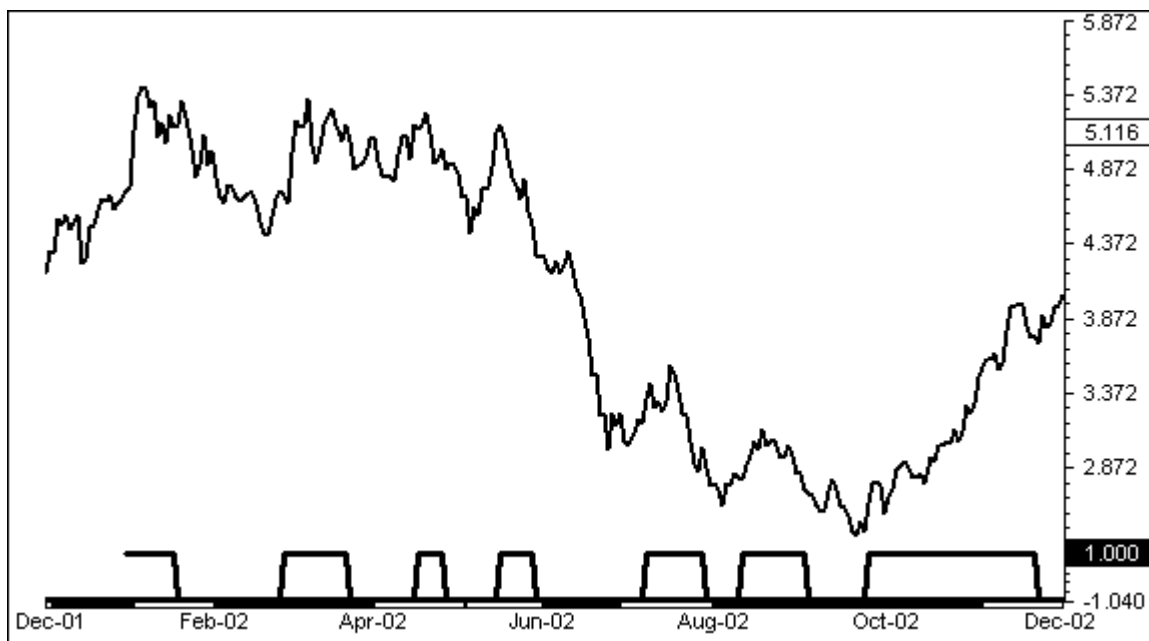


Рисунок 1 Пример работы модели на данных о торгах акций «РАО ЕЭС» на ММВБ

Настройка системы должна регулярно обновляться и соответствовать пожеланиям и правилам работы каждого инвестора.

К устанавливаемым клиентом параметрам следует отнести:

- Частоту перенастройки параметров (не чаще раза в день);
- Количество аппроксимируемых циклов (может быть рассчитано автоматически или установлено вручную);
- δ – допустимый уровень погрешности аппроксимации (в случае автоматического определения количества циклов);
- Длительность анализируемого периода;
- Максимальный и минимальный размеры окон регрессионных функций;
- Критерии оптимизации: вероятностная либо экономическая эффективность описанной торговой системы.

В результате тестирования системы по критерию вероятности правильной идентификации рыночной ситуации установлено преимущество предлагаемой модели перед общепринятыми схемами. При этом вероятность достоверного прогноза достигала на отдельных наборах котировок значений 0,73.

Совместное использование реализации предлагаемой авторами модели с системами интернет-трейдинга, широко представленным в настоящее время на российском рынке, позволит частному инвестору сократить убытки и оптимизировать свою стратегию, воспринимая сигналы модели как рекомендательные, а в некоторых случаях и как указания к действию.

ЛИТЕРАТУРА

1. Бурькин Д. В. Компьютерное моделирование на фондовом рынке // В сборнике «Компьютерное моделирование. Экономика». – М.: «Вузовская книга», 2000.
2. Доброчеев О. В. Глобальный кризис. Российский сценарий // «НГ-СЦЕНАРИИ», №1 (35), 1999 г.
3. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. – М., 1968. – 720 с.

4. Ле Бо Ч., Лукас Д.В. Компьютерный анализ фьючерсных рынков: Пер. с англ. – М.: Издательский дом «Альпина», 1998. – 304 с.
5. Мерфи Д.Д. Технический анализ фьючерсных рынков: теория и практика. – М.: Сокол, 1996. – 592 с.
6. Dewey E., Mandino O. Cycles – The Mysterious Forces that Trigger Events. – N.Y.: Hawthorne, 1971.
7. Hurst J. M. The Profit Magic of StockTransaction Timing. – Englewood Cliffs. N.J.: Prentice Hall, 1970.