

**Е.А. Гайворонская**

Новосибирский национальный исследовательский государственный университет, Новосибирск

**А.А. Цыплаков**

Новосибирский национальный исследовательский государственный университет; Институт экономики и организации промышленного производства СО РАН, Новосибирск

## **Использование модифицированного алгоритма Эрева—Рота в агент-ориентированной модели рынка электроэнергии<sup>1</sup>**

**Аннотация.** Один из важных инструментов анализа и предсказания функционирования рынков электроэнергии — это агент-ориентированные модели, в которых имитируется поведение децентрализованных агентов (например, производителей и покупателей), у каждого из которых собственные цели и возможности. В этих моделях большую роль играет самообучение агентов, подающих ценовые заявки на оптовом рынке. В процессе повторяющихся взаимодействий агент адаптируется к окружающей обстановке и поведению других агентов, учится предсказывать результаты своих действий. В статье представлена модификация классического алгоритма обучения с подкреплением Эрева—Рота, принимающая во внимание расстояния между альтернативами. Предложенный модифицированный алгоритм был применен для обучения агентов в агент-ориентированной модели оптового рынка электроэнергии России (сибирская ценовая зона) в рамках рынка на сутки вперед. Показано, что он обладает рядом существенных преимуществ по сравнению с исходным алгоритмом. В частности, алгоритм хорошо поддается интерпретации, устойчив к выбору величины шага дискретизации, инвариантен к сдвигу шкалы выигрышей. В целом алгоритм более гибок, чем исходный. При использовании модифицированного алгоритма наблюдается хорошее соответствие между динамикой модельной цены и наблюдаемой динамикой цены на рынке.

**Ключевые слова:** *агент-ориентированные модели, оптовый рынок электроэнергии, рынок на сутки вперед, алгоритмы обучения, алгоритм Эрева—Рота.*

Классификация JEL: C63, D43, D44, D83, L1, L94.

DOI: 10.31737/2221-2264-2018-39-3-3

### **1. Введение**

Традиционные экономические модели базируются на предположениях о полной информации, рациональности поведения экономических агентов и равновесности состояния экономики. Также часто используются предположения об однородности агентов и симметричности равновесия. Классический подход к моделированию экономических систем очень удобен, но всегда был предметом справедливой критики. В частности, можно сослаться на результаты исследований, которые демонстрируют, что в реальной жизни люди могут вести себя совсем по-другому, чем предполагается в традиционных моделях (см., например, (Kahneman, 2003)).

<sup>1</sup> Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 18-010-00728).

Одно из перспективных направлений, не базирующееся на предположениях о полноте информации, рациональности агентов, равновесности экономической системы и однородности агентов – использование так называемых агент-ориентированных моделей (АОМ) (Макаров, Бахтизин, 2009; Макаров, 2012). В АОМ для анализа и предсказания реальных явлений используются компьютерные эксперименты с искусственными обществами. Автономные единицы, составляющие эти общества, называются агентами. Они принимают решения по определенным алгоритмам и могут характеризоваться ограниченной рациональностью, способностью обучаться, адаптироваться к окружающей обстановке и воспроизводить наиболее успешные действия.

Большие преимущества агент-ориентированного моделирования очевидны, в частности при моделировании зарубежных рынков электроэнергии (Weidlich, 2008; Weidlich, Veit, 2008; Guerci, Rastegar, Cincotti, 2010; Ringler, Keles, Fichtner, 2016). При моделировании оптового рынка электроэнергии важно учитывать, что агенты (производители и покупатели электроэнергии) автономны и взаимодействуют друг с другом, причем взаимодействия являются достаточно сложными. На таких рынках агенты обладают определенной рыночной властью и учитывают это в своих действиях. Агенты восприимчивы к окружающей обстановке, собирают информации о собственном состоянии и ситуации на рынке. Кроме того, они активны и могут экспериментировать с различными стратегиями поведения, стремясь увеличить свою прибыль. Также агенты способны в результате повторяющихся взаимодействий обучаться предсказанию последствий своих действий.

Агент-ориентированный подход является очень гибким и позволяет отразить указанные особенности рынка в рамках единой модели, не делая при этом сильных допущений. Важное преимущество заключается в том, что АОМ позволяет заменить трудноосуществимые и дорогостоящие натурные эксперименты с энергетическими системами доступными компьютерными экспериментами. Например, АОМ можно использовать для анализа последствий различных событий, таких как изменения в правилах рыночной торговли, кризисные ситуации и т.п.

Одним из ключевых аспектов АОМ является моделирование процесса адаптации и обучения агентов в ходе повторяющихся взаимодействий. Под адаптивным понимается агент, который имеет некоторый целевой показатель и который действует таким образом, чтобы данный показатель увеличивался (Holland, Miller, 1991). В АОМ свободных рынков электроэнергии агенты (поставщики или покупатели), как правило, обучаются прогнозированию цен на рынке или максимизации своей прибыли.

В литературе встречается много различных подходов к моделированию процессов адаптации и обучения. В данной статье рассматри-

вается алгоритм обучения с подкреплением, предложенный в (Erev, Roth, 1998; Roth, Erev, 1995). Это один из самых популярных алгоритмов обучения, используемых в АОМ рынков электроэнергии (см., например, (Nicolaisen et al., 2001; Koesrindartoto, 2002; Sun, Tesfatsion, 2007; Rupérez Micola et al., 2008; Veit et al., 2009)). В данной статье усовершенствованный алгоритм использован в АОМ российского рынка электроэнергии (сибирская ценовая зона) в рамках рынка на сутки вперед.

Существующие модели российского рынка электроэнергии обычно основываются на анализе моделей несовершенной конкуренции (Айзенберг и др., 2013) и рассматривают вопросы рыночной власти фирм на рынке (Васин, Дайлова, 2013). Мы используем другой подход – имитационное моделирование бесконечно повторяющегося двустороннего аукциона (бесконечной динамической игры) с обучением агентов. Кроме того, имеющиеся модели рассматривают рынок электроэнергии России агрегированно, т.е. не учитывают поведения отдельных игроков и их неоднородность. В то же время в построенной нами модели спрос и предложение представлены дезагрегированно, посредством неоднородных самообучающихся агентов, которые адаптируются к работе рынка. Следовательно, она может помочь разобраться в микрооснованиях работы рынка электроэнергии. Подобного рода модель в перспективе может быть использована для решения практических задач как регуляторами (например, для оценки последствий изменения правил торговли), так и участниками рынка (для выбора торговых стратегий с помощью симулятора рынка).

Заметим, что возможности верификации построенной модели ограничиваются характером имеющихся данных, которые доступны нам лишь в обезличенном виде. Недоступность детальных данных не позволяет сопоставить поведение модельных агентов с поведением реальных компаний, участвующих в работе рынка. Кроме того, наша модель трактует оптовый рынок электроэнергии упрощенно, не рассматривая возможного влияния сектора свободных двусторонних договоров и балансирующего рынка на рынок на сутки вперед. Также на получаемые результаты может повлиять используемое нами предположение об отсутствии ограничений по перетоку электроэнергии во второй ценовой зоне.

Конечно, из-за упрощений модель может давать неточные прогнозы поведения рынка. Недостаточная степень верификации заставляет осторожно относиться к получаемым выводам. Однако мы уверены, агент-ориентированное моделирование в целом и агент-ориентированное моделирование рынков электроэнергии имеют хорошие перспективы, поэтому важно накапливать опыт работ в этой области для российского рынка при всех неизбежных на данном этапе недостатках.

В разд. 2 дается обзор работ, посвященных агент-ориентированному моделированию свободного рынка электроэнергии с акцентом

на моделях обучения. В разд. 3 анализируются недостатки алгоритма Эрева–Рота и предлагается его модификация с устранением данных недостатков. В разд. 4 описаны результаты расчетов.

## **2. Существующие подходы к агент-ориентированному моделированию рынков электроэнергии**

### **2.1. Адаптация и обучение агентов**

В (Haasdijk, Eiben, Winfield, 2013), говоря об адаптации в популяции агентов, выделяют три фундаментальных механизма: эволюцию, индивидуальное обучение и социальное обучение. Как правило, при описании рынков электроэнергии речь идет об индивидуальной адаптации и обучении. То есть в АОМ таких рынков обычно не рассматривается адаптация за счет эволюционного отбора наиболее удачных агентов или процессы социального обучения, когда агенты перенимают стратегии поведения друг у друга.

В имеющейся литературе можно выделить несколько различных подходов. В частности, в ряде АОМ рынков электроэнергии поведение агентов определяется жесткими моделями окружающей обстановки, заложенными в них исследователями исходя из здравого смысла или имеющихся у них экспертных знаний о моделируемой системе. При этом агент обучается в очень малой степени и только настраивает небольшое число параметров внутренней модели. Такие алгоритмы неуниверсальны и зачастую не применимы за пределами той АОМ, для которой разработаны. В качестве примеров можно назвать работы (Bower, Bunn, 2000; Liu et al., 2012), где используются специфические правила изменения цены заявки в зависимости от различных факторов.

При этом многие подходы к адаптации агентов являются довольно общими и не применяют каких-либо жестких моделей окружающей среды (см., например, Q-обучение, генетические алгоритмы и другие алгоритмы, о которых пойдет речь ниже).

Большинство алгоритмов обучения можно отнести к обучению с подкреплением (Kaelbling, Littman, Moore, 1996; Sutton, Barto, 2016). При использовании таких алгоритмов агенты учатся выбирать действия в зависимости от ситуации методом проб и ошибок. Удачные действия подкрепляются некоторым вознаграждением от окружающей среды.

Один из самых распространенных подходов к обучению с подкреплением в АОМ рынков электроэнергии основывается на алгоритме Эрева–Рота (Erev, Roth, 1998; Roth, Erev, 1995), который применим к выбору из конечного числа действий. Использование алгоритма Эрева–Рота было популяризировано работой (Nicolaisen et al., 2001) (см. также (Koesrindartoto, 2002; Weidlich, Veit, 2006; Rupérez et al., 2008; Veit et al., 2009). В (Sun, Tesfatsion, 2007; Veit et al., 2009) авторы применили функцию мягкого максимума (softmax) для задания вероят-

ностей выбора действий. За счет выбора коэффициента эта функция позволяет варьировать степень случайности выбора действий (степень экспериментирования).

В алгоритмах обучения с подкреплением может учитываться влияние действий не только на немедленное вознаграждение, но и на будущую ситуацию и поток будущих вознаграждений. При этом агент может просчитывать свои действия на несколько шагов вперед. Самый известный алгоритм такого рода – это Q-обучение, предложенное в (Watkins, 1989), которое применяется к марковскому процессу принятия решений. Всем парам «состояние» – «действие» сопоставляется некоторая оценка дисконтированных будущих вознаграждений, называемая Q-ценность.

В стандартном алгоритме Q-обучения агент каждый раз выбирает действие с наибольшей Q-ценностью (так называемый жадный выбор). В (Harp et al., 2000) применен случайный выбор по функции мягкого максимума в сочетании с имитацией отжига (больцмановский процесс выбора). Снижение параметра температуры приводит к тому, что выбор все более приближается к жадному, изменяя соотношение между исследованием (экспериментированием) и использованием в пользу последнего. Аналогичная модификация (SA-Q-обучение) применена в (Bakirtzis, Tellidou, 2006), где агент с определенной вероятностью может выбирать не наилучшее, а случайное действие, причем вероятность зависит от Q-ценности двух действий и показателя температуры. В работе (Krause, Andersson, 2006) авторы использовали упрощенное Q-обучение с  $\epsilon$ -жадным выбором без учета состояния.

Еще один популярный подход основывается на генетических алгоритмах (GA), где правила поведения представляются как хромосомы (например, кодируются последовательностью битов). Адаптация агентов к окружающей обстановке происходит в результате эволюционного отбора наиболее приспособленных правил поведения. При этом привлекаются такие механизмы эволюции, как наследование, скрещивание, мутация, отбор. Важно отметить, что рассматривание в популяции правил поведения для одного агента означает социальное обучение, а для популяции агентов – индивидуальное. Для моделирования рынков электроэнергии генетические алгоритмы обучения применялись, например, в работах (Richter, Sheblé, 1998; Nicolaisen, et al., 2000; Cau, Andersen, 2002).

Представляет интерес использование в АОМ обучающихся систем классификаторов (learning classifier system, LCS), которые совмещают обычное обучение с подкреплением и генетические алгоритмы (Holland et al., 2000; Urbanowicz, Moore, 2009). Основное отличие LCS от обычного обучения с подкреплением – это способность к обобщениям. По сравнению с Q-обучением алгоритмы LCS более гибкие, поскольку могут применяться не только для марковских процессов принятия решений и не требуют наличия у агентов конкретной

цели. В (Bagnall, Smith, 2000; Bagnall, Smith, 2005) для моделирования оптового рынка электроэнергии Великобритании был взят вариант LCS, называемый XCS, в котором приспособленность классификаторов определяется не их силой, а точностью классификации (Wilson, 1995). Недостатком использования алгоритмов LCS в АОМ является то, что LCS обладают довольно сложной архитектурой. Так, например, описание XCS в (Butz, Wilson, 2002) показывает, что алгоритм требует настройки большого числа параметров.

## 2.2. Моделирование поведения поставщиков

При агент-ориентированном моделировании свободных рынков электроэнергии преобладают статьи, где в качестве агентов участвуют поставщики, которые нацелены на максимизацию прибыли, а спрос является экзогенным. В данном пункте мы рассмотрим именно такие работы. Однако небольшая группа исследователей обратила внимание и на покупателей. Эти модели рассмотрены в п. 2.3. Другие обзоры АОМ свободного рынка электроэнергии можно найти в работах (Weidlich, 2008; Weidlich, Veit, 2008; Zhou et al., 2007; Sensfuß et al., 2007; Guerci, Rastegar, Cincotti, 2010; Лисин и др., 2013).

Дж. Бауэр и Д. Банн (Bower, Bunn, 2000, 2001) исследовали рынок электроэнергии Англии и Уэльса. Целью было сравнение поведения цен на рынке на сутки вперед и на краткосрочном рынке двусторонних договоров, который на тот момент планировали создать. В (Bower, Bunn, 2001) та же модель применялась к оптовому рынку электроэнергии в Германии. Авторы анализировали влияние слияний компаний на цены. Был сделан вывод, что цены на электроэнергию в ответ на слияния и поглощения будут расти значительно, чем до реформы. Новый рынок позволил взаимодействовать поставщикам и покупателям напрямую, устанавливая индивидуальные объемы поставок и цены на сутки вперед. Это должно было позволить части сделок совершаться по более низким ценам. Результаты проведенных имитационных экспериментов указали на ошибочность введения нового рынка, так как в полученной модели цены не снижались, как это ожидалось.

В работе (Visudhiphan, 2003) проведен поиск оптимальной стратегии подачи ценовых заявок на рынке на сутки вперед при использовании трех альтернативных алгоритмов обучения и продемонстрировано, что повторяющиеся подачи ценовых заявок играют важную роль в динамике рынка, в то время как традиционные статические модели рынка не в состоянии отразить эту динамику.

В (Weidlich, Veit, 2006) построена одновременная модель рынка на сутки вперед и балансирующего рынка. Для обучения поставщиков был применен модифицированный алгоритм Эрева–Рота, как в работе (Nicolaisen et al., 2001). В статье (Veit et al., 2009) авторы продолжают свое исследование и применяют модель из статьи (Sun, Tesfatsion, 2007) к оптовому рынку электроэнергии Германии.

В работе (Krause, Andersson, 2006) изучались рыночная концентрация и политика системного оператора по управлению рынком на сутки вперед с использованием алгоритма Q-обучения для поставщиков.

В (Bagnal, Smith, 2005) представлена АОМ свободного рынка Великобритании до нового торгового соглашения (NETA). В статье анализировалось, способны ли агенты-поставщики, использующие иерархический алгоритм XCS, обучиться оптимальной стратегии, когда они конкурируют с необучаемыми агентами. Также рассматривается, возможна ли кооперация между агентами без прямого сговора. Авторы сделали акцент на сложность алгоритма обучения, но не продемонстрировали, как эта сложность улучшает результаты по сравнению с другими способами обучения.

В (Liu et al., 2012) сравниваются поведение цен и объемов сделок для двух механизмов ценообразования на оптовом рынке электроэнергии: каждый платит, сколько заявил, или все платят одинаковую равновесную цену.

### 2.3. Моделирование поведения поставщиков и покупателей

Д. Банн и Ф. Оливьера (Bunn, Oliveira, 2001) применили агент-ориентированное моделирование к новому рынку NETA в Великобритании. В отличие от статей Дж. Бауэра и Д. Банна здесь моделируется не только поведение поставщиков, но и поведение покупателей в рамках рынка на сутки вперед и балансирующего рынка. Был рассмотрен механизм установления равновесия на новом рынке и получен ряд прогнозов относительно его функционирования.

И. Праса с соавторами в своих статьях также рассматривают покупателей как отдельных агентов, но не наделяют их переговорной силой, а совокупный спрос моделируют как неэластичный по цене. В статье (Praça et al., 2003) исследуется реакция агентов на изменения условий рынка, а в (Praça et al., 2004) сравнивается поведение агентов при разных механизмах ценообразования.

В статье (Nicolaisen et al., 2001) авторы для обучения агентов применяли алгоритм Эрева–Рота. В работе изучается влияние на рынок электроэнергии изменения концентрации и мощностей поставщиков. Полученный результат сопоставим с более ранними исследованиями, где покупатели и продавцы используют генетические алгоритмы обучения (Nicolaisen et al., 2000). Авторы этих статей разработали программный комплекс AMES для экспериментального изучения оптовых рынков электроэнергии, описанный в статьях (Tesfatsion, 2008; Li, Tesfatsion, 2009). Обучаться в модели могут лишь поставщики, но зато она учитывает физические ограничения. Имеется ряд работ по построению имитационной модели рынка с применением обучения Эрева–Рота, например (Sun, Tesfatsion, 2007; Koesrindartoto, 2002).

В (Cau, Anderson, 2002) в рамках условной модели свободного рынка электроэнергии сделали акцент на обучении. Они использовали для моделирования подачи ценовых заявок агентами генетический алгоритм. Был сделан вывод, что учет состояния окружающей среды в виде дискретной переменной способствует лучшему обучению.

В (Vubnitz et al., 2014) построена модель оптового рынка электроэнергии Германии с неопределенностью (возможность отключения электростанций и колебания поставок от источников возобновляемой энергии). Авторы сделали вывод, что результаты имитаций на основе этой модели хорошо соответствуют историческим рыночным данным.

В целом анализ литературы по агент-ориентированному подходу показывает, что исследователи редко рассматривают двусторонние договоры (регулируемые или свободные). Также редко принимаются во внимание затраты на транспортировку и ограничения сети, хотя потенциально эти факторы могут повлиять на исход торговли. В типичной АОМ рынка электроэнергии имитируется торговая система, через которую осуществляется торговля на реальном рынке. Администратор торговой системы контролирует процесс торгов и определяет равновесные цены. У агентов есть дискретное или непрерывное множество альтернатив (наборов цена—объем), и они выбирают альтернативы, ориентируясь на прибыль, зависящую от оптовой цены, установившейся на рынке, и от собственных характеристик, таких как издержки производства электроэнергии, розничные цены и т.п.

### **3. Алгоритм Эрева—Рота и его модификация**

#### **3.1. Выбор алгоритма обучения**

Для использования в АОМ следует выбирать алгоритмы обучения, соответствующие особенностям агент-ориентированного подхода. Например, существуют мощные системы искусственного интеллекта, которые позволяют решать небанальные задачи, такие как управление автомобилем, игра в шахматы и т.д., но применения подобных систем в АОМ следует избегать. Это связано с тем, что, во-первых, модель со сложно устроенными агентами трудно отлаживать, во-вторых, сложные алгоритмы замедляют расчеты по модели.

Мы считаем, что простые и интуитивные одношаговые алгоритмы обучения с подкреплением хорошо подходят для модели оптового рынка электроэнергии. На таких рынках агентам не нужно просчитывать последствия своих действий на несколько периодов вперед, а значит, нет смысла использовать Q-обучение или XCS на основе Q-обучения. Если делать выбор между генетическими алгоритмами и алгоритмом Эрева—Рота, то последний кажется более предпочтительным, поскольку он проще при сопоставимых прочих характеристиках. Именно по указанным причинам во многих АОМ рынка электроэнергии применяется алгоритм Эрева—Рота.

### 3.2. Исходный алгоритм

Алгоритм обучения с подкреплением Эрева–Рота был предложен в статьях (Roth, Erev, 1995; Erev, Roth, 1998) в качестве модели поведения людей в игровых ситуациях. Авторы продемонстрировали, что алгоритм хорошо согласуется с реальным поведением в экспериментальных играх, он достаточно прост и в то же время имитирует психологические особенности приема решений.

Идея, лежащая в основе любого обучения с подкреплением, заключается в том, что склонность реализовывать действие должна повышаться, если это действие дало положительные результаты, и снижаться, если результаты отрицательные. В алгоритме Эрева–Рота дополнительно учитываются такие психологические явления, как экспериментирование, забывание, влияние собственного опыта на склонность выбирать те или иные действия и на стремление экспериментировать.

Исходный алгоритм Эрева–Рота (АЭР) выглядит следующим образом. В период  $t=1$  для каждого возможного действия  $j=1, \dots, J$  устанавливается некоторый первоначальный уровень показателя  $S_j^t$  – склонность выбрать это действие (propensity). Если в результате выбора в период  $t$  действия  $k^t$  получено подкрепление (вознаграждение)  $R^t \geq 0$ , то склонности меняются по правилу

$$S_j^{t+1} = (1-f)S_j^t + \begin{cases} R^t(1-e), & j = k^t; \\ R^t e / (J-1), & j \neq k^t. \end{cases} \quad (1)$$

Здесь  $f \in [0; 1]$  – параметр новизны (recency),  $e$  – параметр склонности экспериментировать (насколько сильно агент любит применять пробные действия для накопления собственного опыта). Э. Рот и И. Эрев предложили использовать значения  $f = 0,1$  и  $e = 0,2$ .

Параметр новизны  $f$  определяет, насколько сильно агент склонен изменить свой выбор по сравнению с предыдущим периодом. Из-за множителя  $(1-f)$  происходит забывание – влияние прошлого опыта со временем падает до нуля.

При данных  $S_j^t$  действие выбирается случайно:  $k^t = j$  с вероятностью

$$p_j^t = S_j^t / \sum_{i=1}^J S_i^t. \quad (2)$$

Чтобы показатели склонности  $S_j^t$  не убывали со временем, а вероятности  $p_j$  оставались положительными, подкрепление  $R^t$  должно быть неотрицательным. В рамках теоретико-игровой постановки вознаграждение определяется выигрышем в игре. Авторы предложили взять некоторую возрастающую функцию  $R^t = R(\pi^t)$ , где  $\pi^t$  – это выигрыш в предыдущем периоде, и использовали формулу  $R^t = \pi^t - \pi_{min}$ , где  $\pi_{min}$  – это наименьший возможный выигрыш.

### 3.3. Недостатки исходного алгоритма

Несмотря на привлекательные свойства, алгоритм Эрева–Рота обладает рядом недостатков.

1. Он не инвариантен к изменению вознаграждения  $R^t$  на постоянную величину. В идеале выбор действий не должен зависеть от  $\pi_{min}$ , поскольку во многих ситуациях такой выигрыш может соответствовать только исключительно неудачно выбранным действиям. В то же время при использовании (1), (2) и  $R^t = \pi^t - \pi_{min}$  такая зависимость имеет место.

2. Как замечено в (Nicolaisen et al., 2001), если  $\pi^t = \pi_{min}$ , то  $S_j^t$  для всех действий сокращаются пропорционально, так что вероятности  $p_j^t$  не меняются, в то время как хотелось бы, чтобы вероятность неудачно выбранных действий, приведших к очень низкому выигрышу  $\pi_{min}$ , уменьшалась. Авторы данной статьи предложили модификацию формулы (1):

$$S_j^{t+1} = (1-f)S_j^t + \begin{cases} R^t(1-e), & j = k^t; \\ S_j^t e / (J-1), & j \neq k^t. \end{cases} \quad (3)$$

Однако подобная модификация представляется слишком уж произвольной. К тому же она не решает других проблем АЭР.

3. Одномерная числовая величина  $S_j^t$  одновременно выполняет две функции. С одной стороны, она отражает, насколько часто применялось действие  $j$  (назовем это «опытность»). С другой стороны, ее величина тем больше, чем более высокий выигрыш получается в среднем при выборе действия  $j$ . Как следствие, не очень удачные, но часто выбиравшиеся в прошлом действия могут получать неоправданное преимущество.

4. Величину  $S_j^t$  трудно интерпретировать, так как она имеет те же единицы измерения, что и вознаграждение  $R^t$ , но что именно она показывает, сложно понять.

5. По замыслу Э. Рота и И. Эрва, чтобы поддерживать склонность экспериментировать, следует подкреплять не только выбранное действие, но и альтернативные действия. Это заложено в формуле (1). Авторы алгоритма предполагали, что в некоторых ситуациях его можно модифицировать, дополнительно подкрепляя не все действия, а только те, которые в некотором смысле близки к выбранному. В частности, если альтернативные действия  $j=1, \dots, J$  упорядочены по степени близости, то вместо (1) можно использовать формулу с подкреплением двух смежных действий (Erev, Roth, 1998):

$$S_j^{t+1} = (1-f)S_j^t + \begin{cases} R^t(1-e), & j = k^t; \\ R^t e / 2, & j = k^t \pm 1; \\ 0 - & \text{иначе.} \end{cases} \quad (4)$$

Как бы то ни было, авторы алгоритма не предлагают общего подхода к решению указанной проблемы.

6. Если потенциально возможные действия – величина непрерывная, ее приходится дискретизировать. Если шаг дискретизации большой, получающаяся решетка будет очень грубой. При маленьком шаге значение  $J$  будет велико, и агент будет обучаться слишком мед-

ленно. Особенно острой проблема становится, когда агент выбирает сразу несколько непрерывных величин. Например, если взять прямоугольную решетку из 10 уровней цены и 10 уровней выпуска, то  $J = 100$  и требуется проводить обучение на протяжении многих сотен периодов, чтобы достигнуть приемлемой степени обучения.

### 3.4. Модифицированный алгоритм

В предлагаемом алгоритме каждому правилу (действию)  $j$  сопоставляется два разных показателя. Это *ожидаемый выигрыш*  $\Pi_j^t$  и *опытность*  $W_j^t$ . Разделение на два показателя напоминает идею, лежащую в основе алгоритма XCS, где (если сравнивать с традиционными алгоритмами LCS) введен дополнительный параметр опытности классификаторов (experience).

Пусть в период  $t$  было использовано правило  $k^t$  и был получен выигрыш  $\pi^t$ . Тогда новое значение ожидаемого выигрыша для каждого правила  $j$  равно среднему взвешенному фактического выигрыша  $\pi^t$  с весом  $\lambda_j^t$  и предыдущего ожидаемого значения  $\Pi_j^t$  с весом  $1 - \lambda_j^t$ :

$$\Pi_j^{t+1} = \lambda_j^t \pi^t + (1 - \lambda_j^t) \Pi_j^t. \quad (5)$$

Коэффициент  $\lambda_j^t \in [0; 1]$  регулирует скорость обучения. Ожидаемый выигрыш меняется тем быстрее, чем ближе  $\lambda_j^t$  к единице. В соответствии с идеей Эрева–Рота обучение происходит не только для выбранного действия, но и для альтернативных действий с разной скоростью обучения для разных действий.

Коэффициент  $\lambda_j^t$  вычисляется по формуле

$$\lambda_j^t = \delta(j, k^t) / W_j^{t+1}. \quad (6)$$

Здесь  $\delta(j, k^t) \in [0; 1]$  – прирост опытности для правила  $j$ . Он зависит от степени близости между правилами  $j$  и  $k^t$ . Предполагается, что прирост опытности наибольший для правила  $k^t$ :

$$\delta(k^t, k^t) = 1. \quad (7)$$

Показатели опытности всех правил меняются по формуле

$$W_j^{t+1} = (1 - \varphi) W_j^t + \delta(j, k^t), \quad (8)$$

где  $\varphi$  – показатель забывчивости, равный нулю или небольшому положительному числу. Если  $\varphi = 0$ , опытность может накапливаться до бесконечности. В противном случае она ограничена сверху величиной

$$1 + (1 - \varphi) + (1 - \varphi)^2 + \dots = 1 / \varphi. \quad (9)$$

Соответственно, если одно и то же правило  $j$  выбирается достаточно много периодов, то для него  $\lambda_j^t$  приближается к  $\varphi$ .

Чем больше показатель  $W_j^t$ , тем меньше полученный выигрыш будет влиять на ожидаемый выигрыш:

$$\Pi_j^{t+1} = \Pi_j^t + \frac{\delta(j, k^t)}{(1 - \varphi) W_j^t + \delta(j, k^t)} (\pi^t - \Pi_j^t). \quad (10)$$

Таким образом, можно понять, за счет чего обучается правило: за счет того что оно само или соседнее с ним было только что выбрано ( $\delta(j, k^t)$  около 1) или за счет того что оно еще мало обучено и его нужно обучить (т.е. значение  $W_j^t$  достаточно мало).

Если сравнивать с исходным алгоритмом, здесь забывчивость применяется только к опыту, но не к ожидаемому выигрышу, что вполне естественно и исправляет отмеченные недостатки неинтерпретируемого комплексного показателя  $S_j^t$ .

Структура функции  $\delta(j, k^t)$  зависит от характера модели. Можно предложить достаточно универсальную формулу

$$\delta(j, k^t) = \exp(-\gamma L(j, k^t)), \quad (11)$$

где  $\gamma$  – положительный коэффициент;  $L(j, k^t)$  – некоторая подходящая функция, задающая расстояние между правилами  $j$  и  $k^t$  ( $L \geq 0$ ,  $L(k^t, k^t) = 0$ ). Например, следуя логике формулы (4), предложенной Эрвем и Ротом, мы могли бы взять  $L(j, k^t) = 0$  для пары смежных правил  $j = k^t \pm 1$  и  $L(j, k^t) = +\infty$  для всех остальных.

Меняя коэффициент  $\gamma$ , можно регулировать степень воздействия выигрыша, полученного для правила  $k^t$ , на остальные правила  $j$ . Чем больше  $\gamma$ , тем такое воздействие меньше.

Далее, важным аспектом алгоритма обучения является принцип выбора правила. Естественно сделать так, что чем выше ожидаемый выигрыш, тем выше вероятность выбора правила. В некоторых ситуациях выигрыш может быть отрицательным, поэтому в общем случае нельзя сделать вероятности пропорциональными ожидаемым выигрышам. Достаточно общий подход – сделать вероятность выбора правила  $j$  пропорциональной величине  $\exp(\alpha^t \Pi_j^t)$ . При этом, что немаловажно, вероятности будут инвариантны к сдвигу выигрышей на постоянную величину.

Таким образом, можно взять следующие вероятности:

$$p_j = \exp(\alpha^t \Pi_j^t) / \sum_{i=1}^J \exp(\alpha^t \Pi_i^t). \quad (12)$$

В АОМ рынка электроэнергии подобная функция использована, в частности, в (Sun, Tesfatsion, 2007) в модификации алгоритма Эрва–Рота. По форме это формула для вероятности выбора альтернативы в модели мультиномиального логита. В литературе по алгоритмам обучения данная функция называется функцией мягкого максимума (softmax). Коэффициент  $\alpha^t \geq 0$  в (12) определяет степень нерациональности и экспериментирования при выборе. При  $\alpha^t = 0$  это будет равновероятный выбор, а при  $\alpha^t \rightarrow \infty$  выбирается правило с максимальным значением ожидаемого выигрыша.

Если со временем изменять коэффициент  $\alpha^t$  так, чтобы он неограниченно возрастал (например,  $\alpha^t = \alpha^0 t$ ), агент постепенно перестает экспериментировать и останавливается на правиле, которое считает наилучшим. Такое поведение соответствует известному алго-

ритму оптимизации – имитации отжига. Также этот принцип выбора часто называют больцмановским (Kaelbling, Littman, Moore, 1996; Sutton, Barto, 2016) из-за сходства с распределением Больцмана. При этом  $1/\alpha^t$  интерпретируется как температура.

#### 4. Модель оптового рынка электроэнергии России

##### 4.1. Структура модели

Имеется два типа агентов:  $N_G$  – поставщики и  $N_C$  – покупатели. Они торгуют через рынок (двусторонний аукцион), организуемый администратором торговой системы (АТС), указывая в своих заявках цены и объемы. Поскольку АТС не обладает свободой действий (действия задаются формулой расчета равновесной цены и индивидуальных равновесных объемов), то в дальнейшем мы не будем считать его отдельным агентом.

Агенты характеризуются следующими неизменными параметрами: минимальные и максимальные цены и объемы, которые агент может указывать в ценовой заявке, средние издержки производства (для поставщиков) и ожидаемая розничная цена (для покупателей). Выбор этих параметров соответствует сибирской ценовой зоне рынка на сутки вперед в рамках ОРЭМ России и описан в (Рашидова, 2017). Всего в модели  $N_G = 53$  поставщиков и  $N_C = 21$  покупателей.

Имитации в модели идут в дискретном времени:  $t = 1, 2, \dots$  На реальном рынке торговля ведется относительно поставок электроэнергии в течение одного часа данного дня. Поскольку в течение суток условия меняются, считаем, что в модели зафиксирован определенный типичный час, а  $t$  – номер дня.

В начале периода  $t$  все агенты должны подать АТС свои заявки:  $(p_i^c, q_i^c)$ ,  $i = 1, \dots, N_C$  от покупателей и  $(p_i^g, q_i^g)$ ,  $i = 1, \dots, N_G$  от поставщиков (чтобы упростить обозначения, индекс времени опустим). Заявки задают ступенчатые функции совокупного спроса и предложения, пересечение которых определяет равновесный объем  $Q^*$  (общий объем покупки/продажи электроэнергии) и равновесную цену  $P^*$  для всех сделок по купле–продаже электроэнергии в заданный час (рис. 1). Поставщики, предлагающие цену выше равновесной, и покупатели, предлагающие цену ниже равновесной, не продают (не покупают) электроэнергию в этот час.

Таким образом, АТС по заявкам вычисляет равновесную цену  $P^*$  и сообщает ее агентам. Также поставщикам сообщаются объемы  $q_i^{g*}$ , которые они

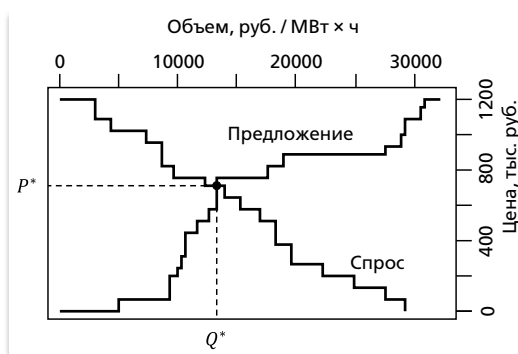


Рис. 1

Поиск равновесия для условного примера

должны поставить, а покупателям – объемы  $q_i^{c*}$ , которые они должны выкупить. На основе этой информации агенты считают свою прибыль: для поставщика –

$$q_i^{s*} (P^* - C_i), \quad (13)$$

для покупателя –

$$q_i^{c*} (r_i - P^*). \quad (14)$$

Подача агентами заявок определяется описанным выше модифицированным алгоритмом Эрева–Рота (для упрощения индекс агента в дальнейшем будем опускать). Вообще говоря, каждый из агентов может выбирать произвольную заявку цена–объем  $(p, q)$  с неотрицательными  $p$  и  $q$ . Однако в алгоритме обучения требуется конечное число правил, поэтому мы дискретизируем интервалы  $[P_{min}, P_{max}]$  и  $[Q_{min}, Q_{max}]$ , деля их на 20 и 8 равноотстоящих значений соответственно. Таким образом, множество правил поведения  $(p_j, q_j)$ ,  $j = 1, \dots, J$ , из которого выбирает агент, образует прямоугольную решетку  $20 \times 8$  и  $J = 160$ .

Выигрыш агента  $\pi^t$  в модели отождествляется с его прибылью. Соответственно, каждому правилу при инициализации сопоставляется некоторый исходный уровень ожидаемой прибыли  $\Pi_j^1$ . Опытность правил исходно устанавливается на низком уровне  $W_j^1 = 10^{-10}$ . В формуле (8) использовалось  $\varphi = 1/50$ . Коэффициент  $\gamma$  в формуле (11) был выбран равным 15. Расстояние между правилами  $j$  и  $k^t$  рассчитывалось по формуле

$$L(j, k^t) = \sqrt{\left(\frac{p_j - p_{k^t}}{P_{max} - P_{min}}\right)^2 + \left(\frac{q_j - q_{k^t}}{Q_{max} - Q_{min}}\right)^2}. \quad (15)$$

Вероятности (12) рассчитываются по ожидаемым прибылям с коэффициентом  $\alpha^t = \alpha^0 t$ , где  $\alpha^0 = 0,000001$ .

Далее мы рассмотрим имитационные эксперименты на основе описанной модели. Технически модель была реализована на языке R. Каждый эксперимент охватывает 100 периодов (шагов). Нами будет рассмотрено два вида обучения: исходный алгоритм Эрева–Рота (АЭР) и модифицированный (МАЭР) – и два разных режима для заявок покупателей: ценопринимающие заявки и произвольные заявки. Здесь мы подробно описали только по одному прогону для каждого варианта, но в целом они отражают общую ситуацию, что показывает анализ 100 прогонов для каждого рассмотренного варианта.

#### 4.2. Компьютерные имитации: ценопринимающие заявки

При ценопринимающих заявках покупатель выбирает только объем, соглашаясь на любую цену. Соответственно, спрос неэластичен по цене (рис. 2).

При использовании МАЭР цена, объем и прибыль колеблются меньше и в среднем выше, чем при АЭР (рис. 3–4). Для АЭР прибыль не имеет явного тренда, и тенденция к стабилизации колебаний

отсутствует, что свидетельствует о том, что агенты не обучились и не сократили набора правил, которые они считают наиболее выгодными.

Для МАЭР доля поставщиков в совокупной прибыли в среднем составляет около 63% (рис. 4). Что касается АЭР, то доля поставщиков колеблется очень широко, и тенденция к стабилизации колебаний практически отсутствует, в среднем она равна 44%.

Также интересно посмотреть на процесс обучения отдельных агентов. Для примера рассмотрим поставщика с номером 25. При использовании МАЭР его прибыль положительна и превышает при-

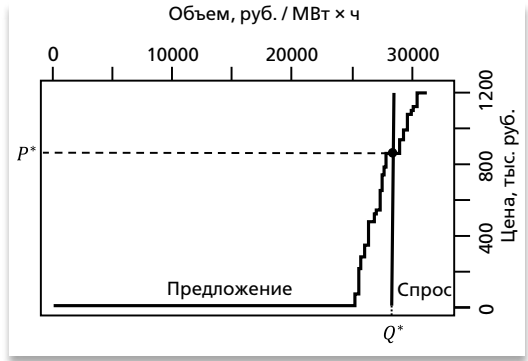


Рис. 2

Поиск равновесия для условного примера с ценопринимающими заявками

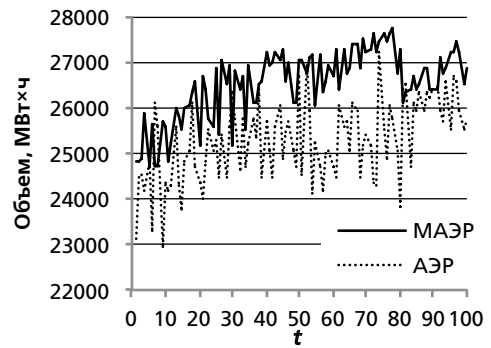
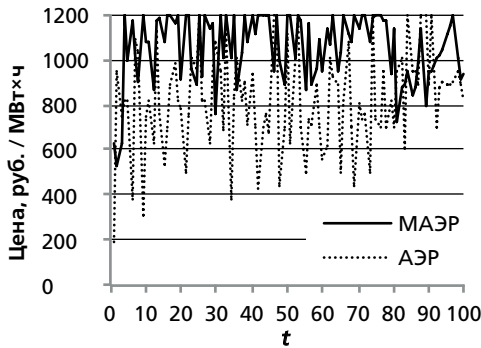


Рис. 3

Динамика цены и объема

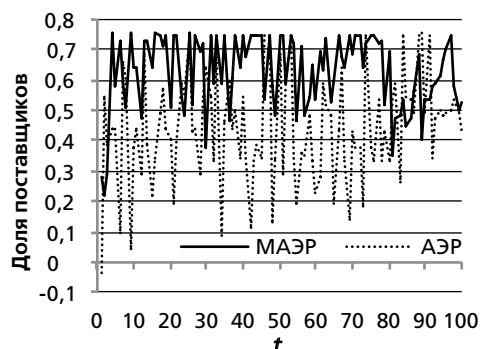
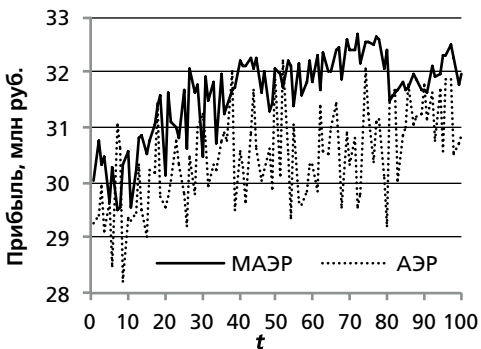


Рис. 4

Динамика совокупной прибыли и доли поставщиков в совокупной прибыли

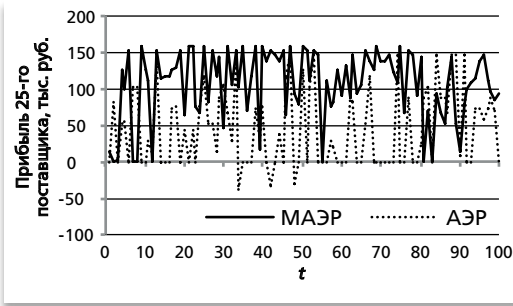


Рис. 5

Динамика прибыли 25-го поставщика

В табл. 1 указаны уровни цены и объема, по которым строится набор правил для данного поставщика.

Таблица 1

Уровни цены и объема для 25-го поставщика

Цена, руб./МВт×ч	1	2	3	4	5	6	7	8	9	10
	0,0	63,2	126,3	189,5	252,6	315,8	378,9	442,1	505,3	568,4
	11	12	13	14	15	16	17	18	19	20
Объем, МВт×ч	631,6	694,7	757,9	821,1	884,2	947,4	1010,5	1073,7	1136,8	1200,0
	1	2	3	4	5	6	7	8		
	178,4	189,7	201,0	212,3	223,7	235,0	246,3	257,6		

Для МАЭР степень влияния прибыли, полученной в результате выбора правила  $k'$ , на ожидаемую прибыль остальных правил зависит от показателя  $\delta(j, k')$ . Выбранное правило  $k'$  имеет  $\delta = 1$ . Для ближайших правил этот показатель достаточно велик, а для дальних слабо отличается от нуля, что иллюстрирует рис. 6. Деления на осях «Цена» и «Объем» обозначают номера возможных цен и объемов в матрице правил.

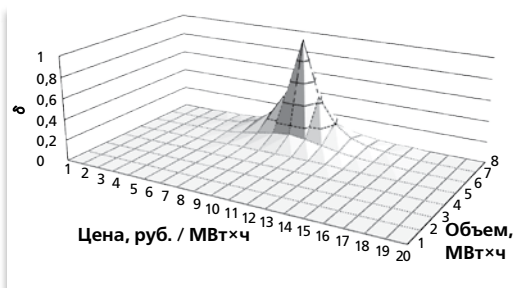


Рис. 6

Коэффициент  $\delta$  для правила на пересечении строки 7 и столбца 10 для 25-го поставщика

быль при АЭР (рис. 5). Также в случае АЭР прибыль колеблется в более широких пределах. Это можно объяснить тем, что в АЭР правила не разделяются на ближайшие и дальние, так что обучение не очень сфокусированное.

Напоминаем, что в нашем случае у каждого агента 160 индивидуальных правил (возможных сочетаний 20 уровней цены и 8 уровней объема), и их можно представить матрицей  $8 \times 20$ .

Особый интерес представляет показатель опытности  $W$ , который отражает, насколько сильно обучились правила. Матрица показателя опытности правил 25-го поставщика на сотом шаге показана на рис. 7 и в табл. 2. Самое опытное правило находится на пересечении строки 7 и столбца 10.

Таблица 2

Фрагмент матрицы показателя опытности правил 25-го поставщика

Объем	Цена						
	8	9	10	11	12	13	14
6	3,06	3,21	3,63	3,29	2,55	2,58	2,19
7	2,48	3,81	5,59	5,53	4,39	2,93	2,43
8	2,37	3,90	4,94	3,15	3,15	2,63	3,53

Показатель  $W$  для самого опытного правила в пределе должен стремиться к  $1/\varphi=50$ . На сотом шаге он достиг лишь уровня 5,59 (рис. 8). Таким образом, понятно, что ста шагов не хватит, чтобы агент хорошо обучился. Продлив эксперимент до 500 шагов, можно увидеть, что примерно на трехсотом шаге лучшее правило почти обучилось, но затем агент переключился на другое правило. На левом графике рис. 8 изображено лучшее правило на сотом шаге, на правом графике – это же правило на пятисотом шаге.

Можно ускорить обучение агентов, задав  $\alpha^0=0,0001$ . Увеличение коэффициента  $\alpha^0$  положительно влияет на скорость обучения агентов. Когда  $\alpha^0$  небольшая, график  $W$  – это ломаная линия, которая убывает на некоторых участках. При увеличении  $\alpha^0$  в 100 раз  $W$  возрастает на всем временном промежутке по логарифму. Однако если оставить  $\alpha^0$  достаточно большим, то нельзя будет сделать однозначный вывод о том, что при модифицированном алгоритме упрощается прогнозирование итоговой цены, так как ее колебания силь-

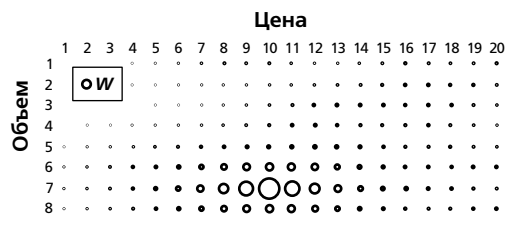


Рис. 7

Матрица показателя опытности для 25-го поставщика; площадь окружностей пропорциональна опытности

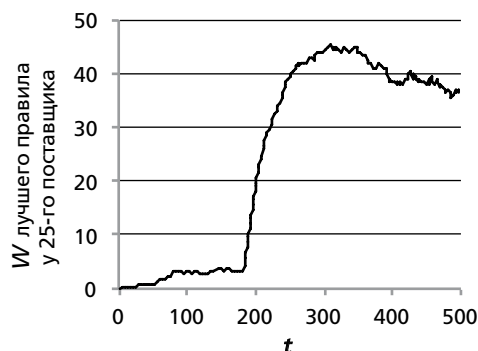
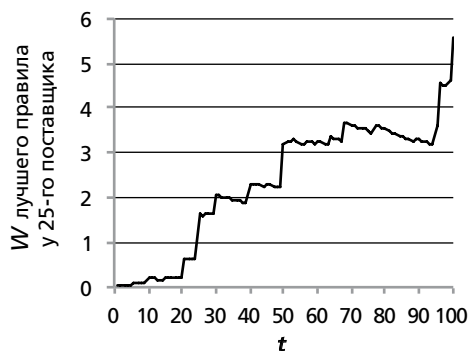


Рис. 8

Обучение 25-го поставщика при модифицированном алгоритме, 100 и 500 шагов

нее, чем при алгоритме Эрева–Рота. Поэтому при расчетах мы брали  $\alpha^0 = 0,000001$ , что замедляет обучение, но при этом делает весь модифицированный алгоритм не хуже для предсказаний, чем алгоритм Эрева–Рота.

#### 4.3. Компьютерные имитации: произвольные заявки

Теперь посмотрим, что изменится, если потребители будут подавать произвольные заявки. Можно увидеть (рис. 9), что при МАЭР равновесные объемы после шага 20 почти не колеблются, в отличие от случая с ценопринимающими заявками, когда они имели тенденцию к росту. Возможно, это объясняется тем, что теперь при таких заявках покупатели по переговорной силе не уступают поставщикам. При АЭР цены колеблются значительно меньше, чем при ценопринимающих заявках.

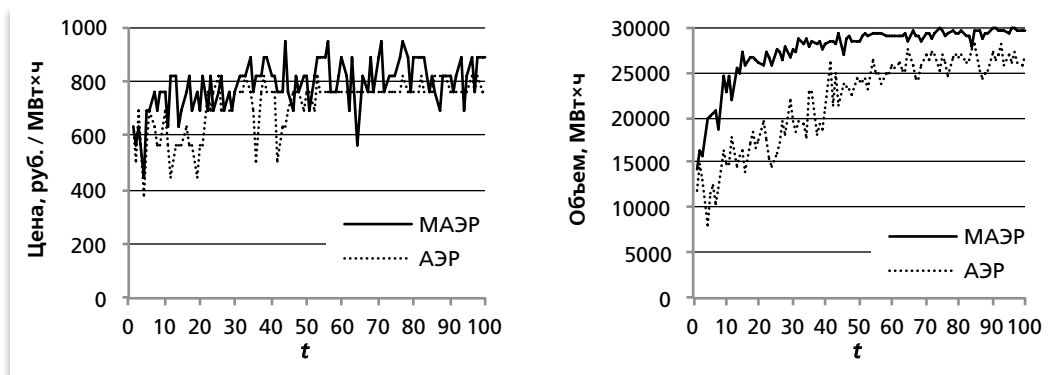


Рис. 9

*Динамика цены и объема*

При МАЭР объем примерно на шаге 20 стабилизируется, а при АЭР объем ниже и постепенно приближается к тому же уровню. Это говорит о том, что агенты обучились быстрее при использовании МАЭР.

При МАЭР совокупная прибыль также быстро выходит на определенный уровень и колеблется незначительно (рис. 10), а значит, агенты быстрее выбирают какое-то конкретное правило и придерживаются его. Если сравнивать этот выбор с выбором при ценопринимающих заявках, то уровень совокупной прибыли на последних итерациях немного выше. При эластичном спросе и предложении и обучении с помощью АЭР ситуация изменилась. Теперь у прибыли появилась тенденция к росту, которая обусловлена ростом объема продаж.

При таких произвольных заявках покупатели обладают большей переговорной силой, что снижает долю поставщиков в общей прибыли (в среднем до 44% при МАЭР; см. рис. 10).

Показатель  $W$  для 25-го поставщика на сотом шаге приблизился к 5, т.е. его скорость обучения осталась примерно той же (рис. 11). Прибыль снизилась, но стала более предсказуемой.

Таким образом, если поставщикам и покупателям позволено подавать произвольные заявки, распределение выигрышей меняется и становится более справедливым по отношению к покупателям, а колебания цены и объема сокращаются. Равновесие и уровень совокупной прибыли становятся более предсказуемыми.

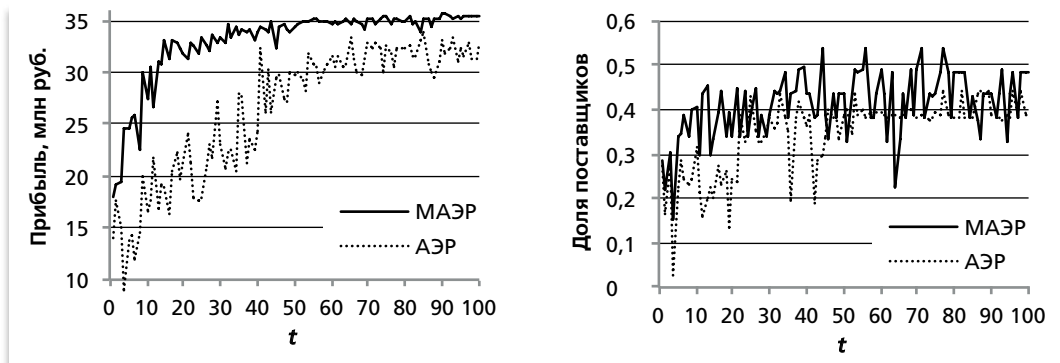


Рис. 10

*Динамика совокупной прибыли и доли поставщиков в совокупной прибыли*

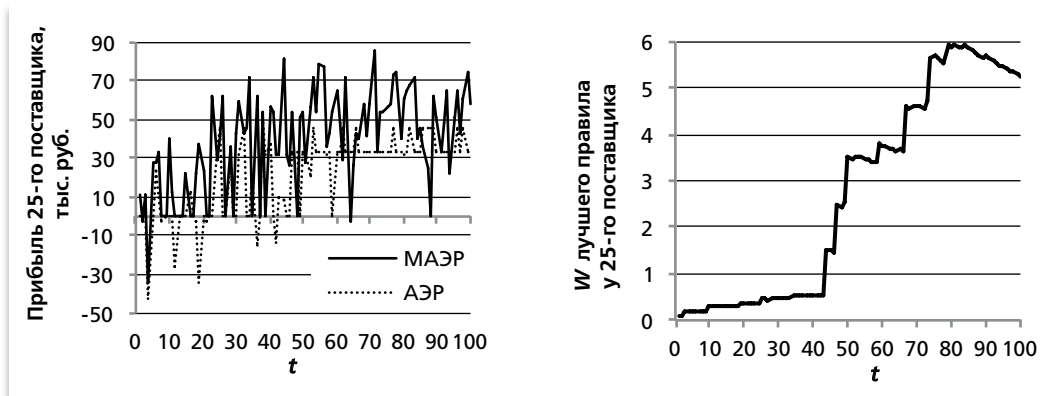


Рис. 11

*Динамика прибыли для 25-го поставщика и его обучение при МАЭР*

#### 4.4. Обсуждение результатов

Экономический смысл предложенной модификации алгоритма обучения Эрева–Рота заключается в том, что для используемых агентами заявок вида «цена–объем» вводится зависимость обучения от расстояния. Одно дело, когда параметры заявки немного, например на 1%, меняются по сравнению со вчерашними, – агент просто локально экспериментирует. Другое дело, когда параметры коренным образом

меняются, например в два раза, — это уже пересмотр стратегии ценообразования. Если заявки меняются мало, агенту следует ожидать прибыль, похожую на ту, что он получал за последнее время. При сильном изменении прибыль скорее всего существенно поменяется. Эта логика формализуется модифицированным алгоритмом, в котором у каждой потенциальной заявки (т.е. каждого правила) есть показатель опытности, изменение которого зависит от расстояния между правилами. В отличие от обычного алгоритма Эрева–Рота процесс обучения построен так, что воздействие на правила, расположенные ближе к выбранному, более сильное.

Модифицированный алгоритм более динамичен, так как склонность к эксперименту не зафиксирована, как в алгоритме Эрева–Рота, а изменяется во времени по мере обучения: чем больше времени прошло, тем меньше агент хочет слепо экспериментировать, так как он уже чему-то научился. То есть мы добавили алгоритму Эрева–Рота черты, присущие обучению человека.

Теоретические достоинства алгоритма и соответствие экономическому содержанию моделируемого явления — это не все требования, которые к нему следует предъявлять. Еще одно важное требование — это соответствие получаемых по модели результатов тому, что мы наблюдаем на реальных рынках. Проверка данного соответствия называется верификацией.

В рамках верификации модели и лежащего в ее основе алгоритма обучения мы можем сравнить цены, получаемые в прогонах модели, и реальные оптовые цены на электроэнергию (рис. 12). При сравнении мы использовали модифицированный алгоритм, неэластичный спрос, 1000 итераций, из которых взяты последние 100.

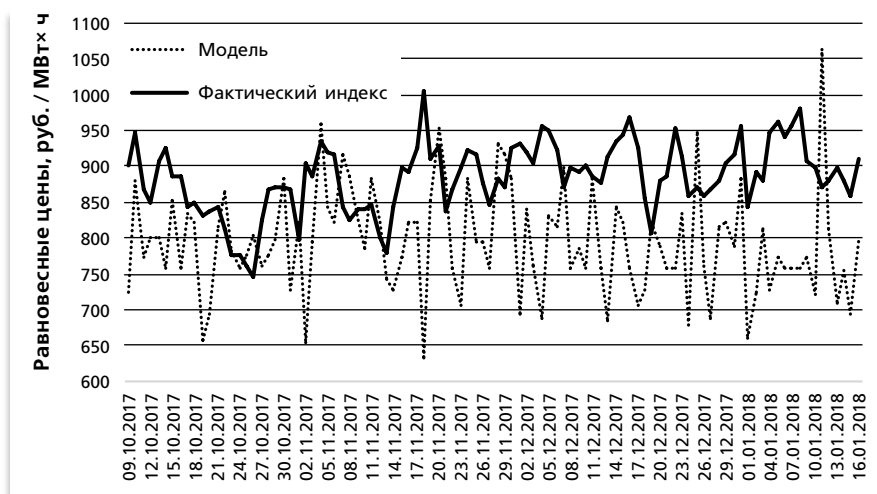


Рис. 12

*Динамика равновесной цены по модели и фактического индекса равновесных цен на покупку электроэнергии за период 09.10.2017–17.01.2018*

Колебания в ценах имитационной модели более интенсивные, однако общий характер их поведения близок к реальному. Причем дневной индекс равновесных цен на покупку электроэнергии – это некоторая усредненная величина, а колебания по дням цены для конкретного часа должны быть более выраженными и похожими на те, что мы получили в модели.

Таким образом, нам удалось построить имитационную модель рынка, которая дает результаты, достаточно схожие с реальными. Эта модель позволяет экспериментировать со свойствами агентов и правилами рынка и предсказывать реакцию агентов на изменение правил и то, как эта реакция влияет на рынок.

Сейчас покупатели второй ценовой зоны подают ценопринимающие заявки, т.е. соглашаются на любую равновесную цену, которая установится на рынке. Проведя эксперимент с изменением формы подачи ценовых заявок покупателей, мы обнаружили, что чем более эластичен спрос, тем стабильнее становится цена на рынке. Поэтому если стимулировать покупателей указывать в заявках те цены, по которым они хотели бы купить электроэнергию, то интенсивность колебаний на рынке должна снизиться.

С практической точки зрения это означает, что за более активной позицией покупателей последует увеличение объемов на другом сегменте – на рынке свободных двухсторонних договоров, где покупатели и продавцы договариваются о ценах и объемах напрямую, а не через обезличенные заявки. Также при активной позиции покупателей на рынке есть потенциал для увеличения суммарной прибыли, а значит, с точки зрения общественного благосостояния необходимо увеличивать ценовую эластичность рыночного спроса.

Построенная модель позволяет получать качественно новые результаты, так как с ее помощью можно изучать микрооснования работы рынка. В частности, можно отвечать на вопросы о последствиях изменений правил торговли, введения новых законов, регулирующих оптовый рынок. Существенным преимуществом предложенной модели является то, что она позволяет проводить анализ того, как рынок приходит в равновесие, как агенты реагируют на изменения и как проходит адаптация. Вообще говоря, любое изменение может повлечь за собой пересмотр торговых стратегий. С этой точки зрения учет в модели обучения агентов важен не только при изучении становления рынка, но и при выяснении последствий введения новых правил, появления на рынке новых игроков или внедрения новых технологий.

## 5. Заключение

В данной статье мы попытались продвинуться в понимании того, как следует моделировать поведение самообучающихся агентов в АОМ. Изучение литературы позволило сделать вывод, что принятие

за основу алгоритма Эрева–Рота является неплохим вариантом для исследования, но алгоритм нуждается в модификации.

Основная особенность предложенного здесь модифицированного алгоритма заключается в том, что разная скорость обучения для разных правил, используемых агентом, определяется опытностью этих правил и расстоянием от выбранного правила. В отличие от исходного алгоритма Эрева–Рота здесь мы можем отслеживать, насколько обучено каждое правило на каждом шаге и что способствует обучению.

Важной особенностью модифицированного алгоритма является то, что он более гибок, так как склонность агента к экспериментированию меняется по мере обучения — чем больше опыта накоплено агентом, тем менее он склонен экспериментировать. Кроме того, модифицированный алгоритм позволяет уйти от таких абстрактных понятий, как «подкрепление» и «склонность», и перейти к более понятной «ожидаемой прибыли», не теряя при этом основную идею.

Предложенный алгоритм помог нам разработать агент-ориентированную модель оптового рынка электроэнергии России. На построенной модели мы сравнили новый и исходный алгоритмы Эрева–Рота и получили ряд выводов. Главный вывод состоит в том, что агенты обучаются быстрее при использовании модифицированного алгоритма. Прибыль и объем меньше подвержены случайным колебаниям, двигаются вверх по тренду и быстрее достигают высокого уровня. Значительные случайные колебания, наблюдаемые при исходном алгоритме даже после большого числа периодов, означают, что агенты очень медленно ищут идеальные для себя правила.

Общую эффективность работы рынка можно анализировать на основе совокупной прибыли. Имитационные эксперименты показали, что модифицированный алгоритм обучения позволяет достигнуть более высокой экономической эффективности.

АОМ разрешает ставить эксперименты для сравнения различных режимов работы рынка. Мы проанализировали последствия выбора покупателями заявленной цены вместо выставления ценопринимающих заявок. Эксперимент показал, что это привело к снижению доли поставщиков в совокупной прибыли и рынок стал более стабильным, а поведение агентов — более предсказуемым.

#### ЛИТЕРАТУРА

- Айзенберг Н.И., Зоркальцев В.И., Киселева М.А.** (2013). Модели несовершенной конкуренции применительно к анализу электроэнергетического рынка Сибири // *Журнал Новой экономической ассоциации*. № 2 (18). С. 62–87.
- Васин А.А., Дайлова Е.А.** (2013). Анализ краткосрочной эффективности механизмов оптового рынка электроэнергии // *Журнал Новой экономической ассоциации*. № 2 (18). С. 35–60.

- Лисин Е.М., Стриелковски В., Григорьева А.Н., Анисимова Ю.А.** (2013). Современные подходы к разработке моделей рынков электроэнергии и исследованию влияния рыночной силы на конъюнктуру энергорынка // *Вектор науки ТГУ*. № 1 (23). С. 188–197.
- Макаров В.Л.** (2012). Искусственные общества // *Экономика и математические методы*. Т. 48. № 3. С. 3–20.
- Макаров В.Л., Бахтизин А.Р.** (2009). Новый инструментарий в общественных науках – агент-ориентированные модели: общее описание и конкретные примеры // *Экономика и управление*. № 12. С. 13–25.
- Рашидова Е.А.** (2017). Агент-ориентированное моделирование оптового рынка электроэнергии России // *Мир экономики и управления*. Т. 17. № 1. С. 70–85.
- Bagnall A.J., Smith G.D.** (2000). An Adaptive Agent Model for Generator Company Bidding in the UK Power Pool // *Artificial Evolution 4<sup>th</sup> European Conference* (AE'99). P. 191–203.
- Bagnall A.J., Smith G.D.** (2005). A Multiagent Model of the UK Market in Electricity Generation // *IEEE Transactions on Evolutionary Computation*. Vol. 9. No. 5. P. 522–536.
- Bakirtzis A.G., Tellidou A.C.** (2006). Agent-Based Simulation of Power Markets under Uniform and Pay-as-Bid Pricing Rules using Reinforcement Learning // *Proceeding of the IEEE Power System Conference and Exposition* (PSCE '06). P. 1168–1173.
- Bower J., Bunn D.W.** (2000). Model-Based Comparison of Pool and Bilateral Markets for Electricity // *Energy Journal*. Vol. 21. No. 3. P. 1–21.
- Bower J., Bunn D.W.** (2001). Experimental Analysis of the Efficiency of Uniform-Price Versus Discriminatory Auctions in the England and Wales Electricity Market // *Journal of Economic Dynamics and Control*. Vol. 25. No. 3–4. P. 561–592.
- Bublitz A., Genoese M., Fichtner W.** (2014). An Agent-Based Model of the German Electricity Market with Short-Time Uncertainty Factors. In: “*11<sup>th</sup> International Conference on the European Energy Market (EEM'14)*”. P. 1–5.
- Bunn D.W., Oliveira F.S.** (2001). Agent-Based Simulation – An Application to the New Electricity Trading Arrangements of England and Wales // *IEEE Transactions on Evolutionary Computation*. Vol. 5. No. 5. P. 493–503.
- Butz M.V., Wilson S.W.** (2002). An Algorithmic Description of XCS // *Soft Computing*. Vol. 6. P. 144–153.
- Cau T.D.H., Anderson E.J.** (2002). A Co-Evolutionary Approach to Modelling the Behavior of Participants in Competitive Electricity Markets // *IEEE Power Engineering Society Summer Meeting*. Vol. 3. P. 1534–1540.
- Erev I., Roth A.E.** (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria // *American Economic Review*. Vol. 88. No. 4. P. 848–881.
- Guerci E., Rastegar M.A., Cincotti S.** (2010). Agent-Based Modeling and Simulation of Competitive Wholesale Electricity Markets. In: “*Handbook of Power Systems II, Energy Systems*”. S. Rebennack et al. (eds). Berlin, Heidelberg: Springer. P. 241–286.

- Haasdijk E., Eiben A.E., Winfield A.F.** (2013). Individual, Social and Evolutionary Adaptation in Collective Systems. In: “*Handbook of Collective Robotics — Fundamentals and Challenges*” Kernbach S. (ed.). Pan Stanford. P. 411–469.
- Harp S.A., Brignone S., Wollenberg B.F., Samad T.** (2000). SEPIA. A Simulator for Electric Power Industry Agents // *IEEE Control Systems*. Vol. 20. Issue 4. P. 53–69.
- Holland J.H., Booker L.B., Colombetti M., Dorigi M., Goldberg D.E., Forrest S., Riolo R.L., Smith R.E., Lanzi P.L., Stolzmann W., Wilson S.W.** (2000). What Is a Learning Classifier System? In: “*Learning Classifier Systems: From Foundations to Applications*”. Chapter 1. Lanzi P.L., Stolzmann W., Wilson S.W. (eds). Berlin, Heidelberg: Springer. P. 3–32.
- Holland J.H., Miller J.H.** (1991). Artificial Adaptive Agents in Economic Theory // *American Economic Review*. Vol. 81. No. 2. P. 365–370.
- Kaelbling L.P., Littman M.L., Moore A.W.** (1996). Reinforcement Learning: A Survey // *Journal of Artificial Intelligence Research*. Vol. 4. P. 237–285.
- Kahneman D.** (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics // *American Economic Review*. Vol. 93. Issue 5. P. 1449–1475.
- Koesrindartoto D.P.** (2002). Discrete Double Auctions with Artificial Adaptive Agents: A Case Study of an Electricity Market using a Double Auction Simulator. Iowa State University Department of Economics Working Papers Series No. 02005.
- Krause T., Andersson G.** (2006). Evaluating Congestion Management Schemes in Liberalized Electricity Markets using an Agent-Based Simulator. In: “*IEEE Power Engineering Society General Meeting*”. Montreal.
- Li H., Tesfatsion L.** (2009). The AMES Wholesale Power Market Test Bed: A Computational Laboratory for Research, Teaching, and Training. In: “*Power & Energy Society General Meeting (PES '09)*”. P. 1–8.
- Liu Z., Yan J., Shi Y., Zhu K., Pu G.** (2012). Multi-Agent Based Experimental Analysis on Bidding Mechanism in Electricity Auction Markets // *International Journal of Electrical Power & Energy Systems*. Vol. 43. No. 1. P. 696–702.
- Nicolaisen J., Petrov V., Tesfatsion L.** (2001). Market Power and Efficiency in a Computational Electricity Market with Discriminatory Double-Auction Pricing // *IEEE Transactions on Evolutionary Computation*. Vol. 5. No. 5. P. 504–523.
- Nicolaisen J., Smith M., Petrov V., Tesfatsion L.** (2000). Concentration and Capacity Effects on Electricity Market Power. In: “*Proceeding of the 2000 Congress on Evolutionary Computation*”. La Jolla. P. 1041–1047.
- Praça I., Ramos C., Vale Z., Cordeiro M.** (2003). An Agent-Based Simulator for Electricity Markets: Seller, Buyer, and Trader Players // *Lecture Notes in Computer Science*. Vol. 2744. P. 290–301.
- Praça I., Ramos C., Vale Z., Cordeiro M.** (2004). Intelligent Agents for the Simulation of Competitive Electricity Markets // *International Journal of Modelling and Simulation*. Vol. 24. No. 2. P. 73–79.
- Richter C.W., Sheblé G.B.** (1998). Genetic Algorithm Evolution of Utility Bidding Strategies for the Competitive Marketplace // *IEEE Transactions on Power Systems*. Vol. 13. Issue 1. P. 256–261.

- Ringler P., Keles D., Fichtner W.** (2016). Agent-Based Modelling and Simulation of Smart Electricity Grids and Markets – A Literature Review // *Renewable and Sustainable Energy Reviews*. Vol. 57. P. 205–215.
- Roth A. E., Erev I.** (1995). Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term // *Games and Economic Behavior*. No. 8. P. 164–212.
- Rupérez Micola A., Banal-Estañol A., Bunn D.W.** (2008). Incentives and Coordination in Vertically Related Energy Markets // *Journal of Economic Behavior & Organization*. Vol. 67. P. 381–393.
- Sensfuß F., Ragwitz M., Genoese M., Möst D.** (2007). Agent-Based Simulation of Electricity Markets: A Literature Review // *Energy Studies Review*. Vol. 15. No. 2. P. 1–29.
- Sun J., Tesfatsion L.** (2007). Dynamic Testing of Wholesale Power Market Designs: An Open-Source Agent-Based Framework // *Computational Economics*. Vol. 30. No. 3. P. 291–327.
- Sutton R.S., Barto A.G.** (2016). Reinforcement Learning: An Introduction. [Электронный ресурс] Режим доступа: <https://webdocs.cs.ualberta.ca/~sutton/book/bookdraft2016sep.pdf>, свободный. Загл. с экрана. Яз. англ. (дата обращения: январь 2017 г.).
- Tesfatsion L.** (2008). The AMES Wholesale Power Market Test Bed as a Stochastic Dynamic State-Space Game. [Электронный ресурс] Режим доступа: <http://www2.econ.iastate.edu/tesfatsi/AMESStateSpaceModel.pdf>, свободный. Загл. с экрана. Яз. англ. (дата обращения: январь 2017 г.).
- Urbanowicz R.J., Moore J.H.** (2009). Learning Classifier Systems: A Complete Introduction, Review, and Roadmap // *Journal of Artificial Evolution and Applications*. Vol. 2009. Article ID 736398.
- Veit D., Weidlich A., Kraft J.** (2009). An Agent-Based Analysis of the German Electricity Market with Transmission Capacity Constraints // *Energy Policy*. Vol. 37. No. 10. P. 4132–4144.
- Visudhiphan P.** (2003). An Agent-Based Approach to Modeling Electricity Spot Markets. Ph.D. dissertation. Massachusetts Institute of Technology. [Электронный ресурс] Режим доступа: <http://hdl.handle.net/1721.1/28270>, свободный. Загл. с экрана. Яз. англ. (дата обращения: январь 2017 г.).
- Watkins C.J.C.H.** (1989). Learning from Delayed Rewards. Ph.D. dissertation. [Электронный ресурс] Cambridge University. Режим доступа: [http://www.cs.rhul.ac.uk/~chrisw/new\\_thesis.pdf](http://www.cs.rhul.ac.uk/~chrisw/new_thesis.pdf), свободный. Загл. с экрана. Яз. англ. (дата обращения: январь 2017 г.).
- Weidlich A.** (2008). Engineering Interrelated Electricity Markets: An Agent-Based Computational Approach. Heidelberg: Physica-Verlag.
- Weidlich A., Veit D.** (2006). Bidding in Interrelated Day-Ahead Electricity Markets: Insights from an Agent-Based Simulation Model. In: “*Proceedings of the 29<sup>th</sup> IAEE International Conference*”. Potsdam. P. 1–10.
- Weidlich A., Veit D.** (2008). A Critical Survey of Agent-Based Wholesale Electricity Market Models // *Energy Economics*. Vol. 30. No. 4. P. 1728–1759.

**Wilson S.W.** (1995). Classifier Fitness Based on Accuracy // *Evolutionary Computation*. Vol. 3. No. 2. P. 149–175.

**Zhou Z., Chan W.K., Chow J.H.** (2007). Agent-Based Simulation of Electricity Markets: A Survey of Tools // *Artificial Intelligence Review*. Vol. 28. No. 4. P. 305–342.

Поступила в редакцию 15 февраля 2018 г.

#### REFERENCES (with English translation or transliteration)

**Ayzenberg N.I., Kiseleva M.A., Zorkaltsev V.I.** (2013). Models of Imperfect Competition in Analysis of Siberian Electricity Market. *Journal of the New Economic Association*, 2 (18), 62–88 (in Russian).

**Bagnall A.J., Smith G.D.** (2000). An Adaptive Agent Model for Generator Company Bidding in the UK Power Pool. In: “*Artificial Evolution 4<sup>th</sup> European Conference (AE’99)*”, 191–203.

**Bagnall A.J., Smith G.D.** (2005). A Multiagent Model of the UK Market in Electricity Generation. *IEEE Transactions on Evolutionary Computation*, 9, 5, 522–536.

**Bakirtzis A.G., Tellidou A.C.** (2006). Agent-Based Simulation of Power Markets under Uniform and Pay-as-Bid Pricing Rules using Reinforcement Learning. In: “*Proceeding of the IEEE Power System Conference and Exposition (PSCE ‘06)*”, 1168–1173.

**Bower J., Bunn D.W.** (2000). Model-Based Comparison of Pool and Bilateral Markets for Electricity. *Energy Journal*, 21, 3, 1–21.

**Bower J., Bunn D.W.** (2001). Experimental Analysis of the Efficiency of Uniform-Price versus Discriminatory Auctions in the England and Wales Electricity Market. *Journal of Economic Dynamics and Control*, 25, 3–4, 561–592.

**Bublitz A., Genoese M., Fichtner W.** (2014). An Agent-Based Model of the German Electricity Market with Short-Time Uncertainty Factors. In: “*11<sup>th</sup> International Conference on the European Energy Market (EEM’14)*”, 1–5.

**Bunn D.W., Oliveira F.S.** (2001). Agent-Based Simulation – An Application to the New Electricity Trading Arrangements of England and Wales. *IEEE transactions on evolutionary computation*, 5, 5, 493–503.

**Butz M.V., Wilson S.W.** (2002). An Algorithmic Description of XCS. *Soft Computing*, 6, 144–153.

**Cau T.D.H., Anderson E.J.** (2002). A Co-Evolutionary Approach to Modelling the Behavior of Participants in Competitive Electricity Markets. *IEEE Power Engineering Society Summer Meeting*, 3, 1534–1540.

**Erev I., Roth A.E.** (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review*, 88, 4, 848–881.

**Guerci E., Rastegar M.A., Cincotti S.** (2010). Agent-Based Modeling and Simulation of Competitive Wholesale Electricity Markets. In: “*Handbook of Power Systems II, Energy Systems*”. Rebenack S. et al. (eds). Berlin, Heidelberg: Springer, 241–286.

**Haasdijk E., Eiben A.E., Winfield A.F.** (2013). Individual, Social and Evolutionary Adaptation in Collective Systems. In: “*Handbook of Collective Robotics — Fundamentals and Challenges*” Kernbach S. (ed.). Pan Stanford, 411–469.

- Harp S.A., Brignone S., Wollenberg B.F., Samad T.** (2000). SEPIA. A Simulator for Electric Power Industry Agents. *IEEE Control Systems*, 20, 4, 53–69.
- Holland J.H., Booker L.B., Colombetti M., Dorigi M., Goldberg D.E., Forrest S., Riolo R.L., Smith R.E., Lanzi P.L., Stolzmann W., Wilson S.W.** (2000). Chapter 1. What Is a Learning Classifier System? In: *“Learning Classifier Systems: From Foundations to Applications”*. Lanzi P.L., Stolzmann W., Wilson S.W. (eds). Berlin, Heidelberg: Springer, 3–32.
- Holland J.H., Miller J.H.** (1991). Artificial Adaptive Agents in Economic Theory. *American Economic Review*, 81, 2, 365–370.
- Kaelbling L.P., Littman M.L., Moore A.W.** (1996). Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research*, 4, 237–285.
- Kahneman D.** (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93, 5, 1449–1475.
- Koesrindartoto D.P.** (2002). Discrete Double Auctions with Artificial Adaptive Agents: A Case Study of an Electricity Market using a Double Auction Simulator. Iowa State University Department of Economics Working Papers Series, 02005.
- Krause T., Andersson G.** (2006). Evaluating Congestion Management Schemes in Liberalized Electricity Markets using an Agent-Based Simulator. In: *“IEEE Power Engineering Society General Meeting”*. Montreal.
- Li H., Tesfatsion L.** (2009). The AMES Wholesale Power Market Test Bed: A Computational Laboratory for Research, Teaching, and Training. In: *“Power & Energy Society General Meeting (PES ‘09)”*, 1–8.
- Lisin E.M., Strielkovski V., Grigor’eva A.N., Anisimova Ju.A.** (2013). Modern Approaches to Modelling Electricity Markets and Research of Influence of Market Forces the Situation on Energy Market. *Vector of Sciences. Togliatti State University*, 1 (23), 188–197.
- Liu Z., Yan J., Shi Y., Zhu K., Pu G.** (2012). Multi-Agent Based Experimental Analysis on Bidding Mechanism in Electricity Auction Markets. *International Journal of Electrical Power & Energy Systems*, 43, 1, 696–702.
- Makarov V.L.** (2012). Artificial Societies. *Economics and Mathematical Methods*, 48, 3, 3–20 (in Russian).
- Makarov V.L., Bakhtizin A.R.** (2009). New Instruments in Social Sciences – Agent-Oriented Models: General Description and Specific Examples. *Economics and Management*, 12, 13–25 (in Russian).
- Nicolaisen J., Petrov V., Tesfatsion L.** (2001). Market Power and Efficiency in a Computational Electricity Market with Discriminatory Double-Auction Pricing. *IEEE Transactions on Evolutionary Computation*, 5, 5, 504–523.
- Nicolaisen J., Smith M., Petrov V., Tesfatsion L.** (2000). Concentration and Capacity Effects on Electricity Market Power. In: *“Proceeding of the 2000 Congress on Evolutionary Computation”*. La Jolla. USA, 1041–1047.
- Praça I., Ramos C., Vale Z., Cordeiro M.** (2003). An Agent-Based Simulator for Electricity Markets: Seller, Buyer, and Trader Players. *Lecture Notes in Computer Science*, 2744, 290–301.
- Praça I., Ramos C., Vale Z., Cordeiro M.** (2004). Intelligent Agents for the Simulation of Competitive Electricity Markets. *International Journal of Modelling and Simulation*, 24, 2, 73–79.

- Rashidova E.A.** (2017). Agent-Based Modeling of Wholesale Electricity Market. *World of Economics and Management*, 17, 1, 70–85 (in Russian).
- Richter C.W., Sheblé G.B.** (1998). Genetic Algorithm Evolution of Utility Bidding Strategies for the Competitive Marketplace. *IEEE Transactions on Power Systems*, 13. Issue 1, 256–261.
- Ringler P., Keles D., Fichtner W.** (2016). Agent-Based Modelling and Simulation of Smart Electricity Grids and Markets – A Literature Review. *Renewable and Sustainable Energy Reviews*, 57, 205–215.
- Roth A. E., Erev I.** (1995). Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, 8, 164–212.
- Rupérez Micola A., Banal-Estañol A., Bunn D.W.** (2008). Incentives and Coordination in Vertically Related Energy Markets. *Journal of Economic Behavior & Organization*, 67, 381–393.
- Sensfuß F., Ragwitz M., Genoese M., Möst D.** (2007). Agent-Based Simulation of Electricity Markets: A Literature Review. *Energy Studies Review*, 15, No. 2, 1–29.
- Sun J., Tesfatsion L.** (2007). Dynamic Testing of Wholesale Power Market Designs: An Open-Source Agent-Based Framework. *Computational Economics*, 30, 3, 291–327.
- Sutton R.S., Barto A.G.** (2016). Reinforcement Learning: An Introduction. Available at: <https://webdocs.cs.ualberta.ca/~sutton/book/bookdraft2016sep.pdf> (accessed: January 2017).
- Tesfatsion L.** (2008). The AMES Wholesale Power Market Test Bed as a Stochastic Dynamic State-Space Game. Available at: <http://www2.econ.iastate.edu/tesfatsi/AMESStateSpaceModel.pdf> (accessed: January 2017).
- Urbanowicz R.J., Moore J.H.** (2009). Learning Classifier Systems: A Complete Introduction, Review, and Roadmap. *Journal of Artificial Evolution and Applications*, Article ID 736398.
- Vasin A.A., Daylova E.A.** (2013). Analysis of the Short-Term Efficiency of Mechanisms of the Wholesale Electricity Market. *Journal of the New Economic Association*, 2 (18), 35–60 (in Russian).
- Veit D., Weidlich A., Kraft J.** (2009). An Agent-Based Analysis of the German Electricity Market with Transmission Capacity Constraints. *Energy Policy*, 37, 10, 4132–4144.
- Visudhiphan P.** (2003). An Agent-Based Approach to Modeling Electricity Spot Markets. Ph.D. dissertation. Massachusetts Institute of Technology. Available at: <http://hdl.handle.net/1721.1/28270> (accessed: January 2017).
- Watkins C.J.C.H.** (1989). Learning from Delayed Rewards. Ph.D. dissertation. Cambridge University. Available at: [http://www.cs.rhul.ac.uk/~chrisw/new\\_thesis.pdf](http://www.cs.rhul.ac.uk/~chrisw/new_thesis.pdf) (accessed: January 2017).
- Weidlich A.** (2008). Engineering Interrelated Electricity Markets: An Agent-Based Computational Approach. Heidelberg: Physica-Verlag.
- Weidlich A., Veit D.** (2006). Bidding in Interrelated Day-Ahead Electricity Markets: Insights from an Agent-Based Simulation Model. *Proceedings of the 29<sup>th</sup> IAEE International Conference*. Potsdam, 1–10.

- Weidlich A., Veit D.** (2008). A Critical Survey of Agent-Based Wholesale Electricity Market Models. *Energy Economics*, 30, 4, 1728–1759.
- Wilson S.W.** (1995). Classifier Fitness Based on Accuracy. *Evolutionary Computation*, 3, 2, 149–175.
- Zhou Z., Chan W.K., Chow J.H.** (2007). Agent-Based Simulation of Electricity Markets: A Survey of Tools. *Artificial Intelligence Review*, 28, 4, 305–342.

Received 15.02.2018

**E.A. Gaivoronskaia**

Novosibirsk State University, Novosibirsk, Russia

**A.A. Tsyplakov**

Novosibirsk State University; Institute of Economics and Industrial Engineering of SB RAS, Novosibirsk, Russia

## **Using a Modified Erev—Roth Algorithm in an Agent-Based Electricity Market Model<sup>2</sup>**

**Abstract.** One of the important tools for the analysis and prediction of operation of electricity markets are agent-based models, which simulate the behavior of decentralized agents (for example, producers and buyers), each with its own objectives and means. In these models learning of agents submitting price bids to a wholesale market plays an important role. In the process of repeated interaction an agent adapts to the environment and to the behavior of other agents, learns to predict the results of its own actions. The paper presents a modification of the classical Erev—Roth reinforcement learning algorithm which takes into account the distance between alternatives. The proposed modified algorithm is used to represent agents' learning in an agent-based model of the Russian wholesale electricity market (Siberian pricing zone) within the bounds of the day-ahead market. It is shown that it has some significant advantages as compared to the original algorithm. In particular, the algorithm is naturally interpretable, is robust to the choice of discretization step, is invariant to a shift in payoffs scale. On the whole, the algorithm is more flexible than the original one. When the modified algorithm is used, one observes good coherence between the dynamics of model price and the observable dynamics of the price in the market.

**Keywords:** *agent-based models, wholesale electricity market, day-ahead market, learning algorithms, Erev—Roth algorithm.*

JEL Classification: C63, D43, D44, D83, L1, L94.

DOI: 10.31737/2221-2264-2018-39-3-3

---

<sup>2</sup>This study was supported by the Russian Foundation for Basic Research (project 18-010-00728).