

А.В. СУВОРОВА  
**ГИБРИДНЫЕ МОДЕЛИ ОЦЕНКИ ПАРАМЕТРОВ  
СОЦИАЛЬНО-ЗНАЧИМОГО ПОВЕДЕНИЯ ПО СВЕРХМАЛОЙ  
НЕПОЛНОЙ СОВОКУПНОСТИ НАБЛЮДЕНИЙ**

---

*Суворова А.В. Гибридные модели оценки параметров социально-значимого поведения по сверхмалой неполной совокупности наблюдений*

**Аннотация.** Предложен обзор возможных подходов к построению моделей социально-значимого поведения по сверхмалой неполной совокупности наблюдений и оценке параметров таких моделей. Рассмотрены такие методы как агентное моделирование, статистический подход, включая методы анализа малой выборки, методология временных рядов и их приложения к решению указанной задачи. Выявлены преимущества использования моделей в рамках теории вероятностных графических моделей.

**Ключевые слова:** оценка интенсивности, модели поведения, последние эпизоды, неопределенность.

*Suvorova A.V. Hybrid models for parameter estimates of socially significant behavior on the base of supershort incomplete set of observations.*

**Abstract.** We provide a description of the methods for socially significant behavior modeling and estimation on the base of supershort incomplete set of observations. We consider agent-based modeling, statistical approach including small samples analysis, time series analysis and their application to the described problem. Finally, we describe advantages of probabilistic graphical models for representation of socially significant behavior.

**Keywords:** rate estimate, behavior models, last episodes, uncertainty.

---

**1. Введение.** Во многих отраслях социологических, психологических, маркетинговых исследований возникают задачи оценивания интенсивности социально-значимого поведения респондентов [48, 59, 60, 61]. Например, в настоящее время наиболее острой методологической проблемой в эпидемиологии является оценка риска передачи и приобретения такой опасной и неизлечимой инфекции как инфекция вирусом иммунодефицита человека (ВИЧ) в зависимости от особенностей инъекционного и сексуального поведения индивида. Наиболее точно такой риск характеризуется инцидент-показателем [100] — числом заразившихся за определенный период среди лиц, подвергавшихся риску заражения, отнесенным к человеко×месяцам наблюдения.

Прямые измерения во многих случаях невозможны или требуют значительных финансовых, временных и трудовых ресурсов. В частности, «дневниковый метод» подразумевает запись всех действий респондента в течение дня, после чего полученные данные, как правило, накапливавшиеся несколько месяцев, поступают на обработку эксперта, который подсчитывает число эпизодов поведения определенного

вида за данный период. Однако такой вид исследований достаточно дорог, его сложно организовать и долго выполнять. Наиболее доступными исходными данными для анализа поведения выступают самоотчеты респондентов о его поведении, то есть ответы в анкете на блок вопросов или результаты проведения интервью. Требуется предложить математические модели, позволяющие выполнить более дешевые косвенные измерения интенсивности социально-значимого поведения на основе ответов респондентов [46–52, 59, 60].

Сейчас для оценки интенсивности поведения респондента по его «одномоментному» самоотчету используются два метода, каждый из которых имеет недостатки [60, 61]. Первый метод — прямые вопросы: «Сколько раз Вы делали так в течение последнего месяца (трех, шести, года)?». На такие вопросы респонденты обычно дают практически не соотносящиеся с реальностью ответы.

Второй метод — Лайкерт-шкалы — опросники, в которых используются качественные, а не количественные варианты: «Никогда», «Редко», «Иногда», «Часто», «Всегда» и подобные им возможности для ответа [78, 80, 108]. Вопрос ставится легко, ответ тоже получить несложно, однако эти ответы не несут никаких полезных сведений относительно числа эпизодов: то, что «Часто» для одного человека, может быть «Редко» для другого, а то, что «Часто» в одном виде поведения, может быть «Редко» для другого вида поведения. Кроме того, «расстояние» между «Всегда» и «Очень часто» совершенно не обязательно совпадает с расстоянием между «Редко» и «Никогда». На практике шкалы арифметизируют, но за этой арифметизацией не стоит никакой достоверной гипотезы; получающиеся расчеты ситуацию с риском не характеризуют вообще никак.

Одной из возможных альтернатив Лайкерт-шкал представляется опрос респондента об эпизодах социально-значимого поведения [32, 36, 44, 46–52, 59, 60, 61]. Такой опрос позволяет судить об интервалах между последними эпизодами социально-значимого поведения, о минимальном и максимальном интервалах между последовательными эпизодами, а так же об интервале между временем опроса и последним эпизодом. Однако из-за особенностей припоминания таким образом можно получить сведения лишь о небольшом числе (3–5) последних эпизодов и о рекордных интервалах между ними, что приводит к необходимости анализа неполных данных.

Цель данной работы — описать возможные подходы к построению моделей социально-значимого поведения по сверхмалой неполной совокупности наблюдений и оценке параметров таких моделей.

**2. Агентное моделирование.** Задачи моделирования поведения человека ставятся и решаются в различных областях, включая исследования в психологии, экономике, социологии, медицине, эпидемиологии, компьютерной безопасности и т.д. [2, 4, 9, 12, 20, 22, 23, 29, 32, 33, 39, 41–43, 48, 57, 58, 62, 63, 68, 82, 102]. Часть исследований представляют собой попытки всестороннего моделирования поведения человека как индивида, фокусирующиеся на причинах, мотивах, установках человека (экономическая модель человека [12, 37, 38]). С другой стороны, вопросы моделирования поведения человека можно отнести к более широкой, комплексной области исследований, посвященной анализу и моделированию поведения в контексте интеллектуальных систем [11, 31, 35, 45]. В частности, если рассматривать поведение как взаимодействие с внешней средой, реакцию в зависимости от внешних факторов, то к возможным способам моделирования поведения человека можно отнести, например, такое направление в искусственном интеллекте как агентное (многоагентное) моделирование.

Агентное моделирование (agent-based model (ABM))— метод имитационного моделирования, исследующий поведение децентрализованных агентов и то, как такое поведение определяет поведение всей системы в целом [17–19, 40, 54, 75]. Агентами могут быть люди, компании, транспортные средства, животные и т.д., другими словами, некоторая сущность, которая воспринимает среду с помощью датчиков, воздействует на эту среду, способна к обучению и обладает автономностью [40]. Поведение агентов описывается на индивидуальном уровне, а поведение системы возникает как результат деятельности множества агентов. Агентная модель представляет реальный мир в виде многих отдельно специфицируемых агентов, каждый из которых взаимодействует с другими агентами. В процессе функционирования агент может изменить как внешнюю среду, так и свое поведение.

Один из примеров [75] использования агентного моделирования — потребительский рынок. В сложной среде рынка выбор покупателя зависит от индивидуальных особенностей, сети контактов, а также внешних влияний, которые лучше всего описываются с помощью агентного моделирования. Другой стандартный пример — это эпидемиология. Здесь агенты это люди, которые могут быть носителями инфекции, переболевшими или восприимчивыми к болезни. Агентное моделирование поможет описать социальные сети, контакты между людьми и в итоге получить объективные прогнозы распространения инфекции.

В настоящее время агентные и мультиагентные системы получили широкое применение в таких областях как системы телекоммуникации, поисковые системы, логистика, компьютерные игры, САПР, системы управления и контроля [1, 17–19, 25, 30, 40, 53–55, 76, 88, 93, 94, 98, 103, 111].

**3. Статистические модели.** Теоретические модели, построенные только на предположениях исследователя, более грубы, учитывают меньшее число факторов, требуют определенных допущений и упрощений. Однако результаты расчета по ним легче обозримы, они более явно отражают присущие явлению основные закономерности, такие модели более приспособлены для поиска оптимальных решений [26]. Статистические модели более соответствуют данным, не требуют столь грубых допущений, позволяют учесть большое число факторов. Но и у них есть свои недостатки: громоздкость, плохая обозримость, невозможность включения фактора в модель при отсутствии данных, описывающих этот фактор. Классические подходы к построению статистических моделей рассмотрены, например, в работах [65, 66, 97] и многих других.

Примеров моделирования поведения на основе статистических данных много (например, [24, 50, 91, 105]). Одним из наиболее распространенных способов построения моделей для объяснения изменчивости некоторой переменной  $Y$  в зависимости от изменений предикторов  $X$  являются регрессионные модели. В общем случае регрессионная модель представляется в виде  $Y \approx f(X)$ , где  $f(X)$  может быть как линейной (тогда говорят о линейной регрессии), так и нелинейной функцией. Целью регрессионного анализа является построение подходящей модели и определение коэффициентов, соответствующих данным. Существуют различные методы как выбора модели, так и оценки ее параметров. Однако использование регрессионных моделей имеет ряд ограничений. В частности, применение методов возможно только при определенных предположениях в отношении используемых данных. Например, в случае линейной регрессии предполагается, что данные распределены нормально, ошибки являются независимыми нормально распределенными случайными величинами с нулевым математическим ожиданием [8, 65, 91, 95].

Как отмечено выше, используемый способ получения данных об эпизодах поведения обладает рядом достоинств, но таким образом невозможно собрать сведения о большом количестве эпизодов поведения одного респондента, в отличие от технических систем, позволяющих производить измерения с требуемой периодичностью. Таким образом,

востребованной оказывается математическая методология для небольшой выборки. В работах [73, 101, 110], описывающих методы анализа на основе небольших выборок, «небольшой» считается выборка, включающая 10–30 наблюдений, в то время как в описанной во введении проблеме речь идет о 3–5 эпизодах. Существующие решения проблемы малой выборки в большинстве случаев представляют собой модификации конкретного метода, решающего определенный узкий круг проблем [73, 101, 110]. Кроме того, ограниченное число ответов приводит к неполноте данных. Исходные данные, то есть сведения об интервалах между несколькими последними эпизодами поведения, о минимальном и максимальном интервалах между последовательными эпизодами поведения, являются членами вариационного ряда. Однако в этом случае мы получаем вариационный ряд «с пропусками» — известен его первый член (минимальный интервал  $T_{\min}$ ), последний (максимальный интервал  $T_{\max}$ ) и несколько элементов между ними, причем с неизвестными номерами. В случае трех интервалов мы рассматриваем элементы  $T_{i:n}$ ,  $T_{j:n}$ ,  $T_{k:n}$ . причем общее число  $n$  эпизодов поведения за исследуемый промежуток времени  $T$  нам также неизвестно (если данные согласованы, можно считать, что  $n \in \left[ \frac{T}{T_{\max}}, \frac{T}{T_{\min}} \right]$ ).

Таким образом, анализ необходимо проводить в условиях неполноты исходных данных.

**4. Временные ряды.** Уточним, что целью данного обзора не является рассмотрение всех подходов к моделированию поведения человека, такая область очень широка и трудноопределима: модели как способ описания предметной области, описания имеющихся данных возникают при решении любой задачи анализа [41, 76, 82, 102]. Рассматривается более узкая задача: изучается поведение особого вида, имеющее четко определенные проявления — эпизоды поведения. В качестве примера можно привести рискованное поведение, например, поведение, связанное с риском передачи и получения такого неизлечимого инфекционного заболевания как ВИЧ. Для такого поведения есть четко определенные эпизоды, во время которых может произойти передача заболевания: незащищенный половой контакт, потребление инъекционных наркотиков и т.д. [72, 83] Таким образом, поведение рассматривается как последовательность эпизодов и анализируются данные об эпизодах.

В такой постановке исследуемая задача близка к задачам, возникающим при анализе временных рядов.

Временной ряд (ВР) — это последовательность дискретных упорядоченных в неслучайные равноотстоящие моменты времени измерений (показателей, наблюдений)  $y(t_1), y(t_2), \dots, y(t_N)$ , характеризующих уровни состояний изучаемого процесса, протекающего в условиях неопределенности [70].

Анализ временного ряда позволяет понять причины, обусловившие поведение изучаемого процесса, и построить модели временных рядов, используемые как для объяснения поведения процесса, так и для прогноза его развития [3, 7, 27, 34, 69–71, 82]. Анализ временных рядов применяется для описания поведения показателей качества во времени, и определения применимости данных для точного прогнозирования изменений. Кроме того, он используется для выявления неожиданного изменения формы временного ряда. Другими словами, анализ временных рядов представляет набор методов для изучения последовательных во времени наблюдений.

Анализ временных рядов полезен в планировании, конструировании, определении изменений процесса, в измерении эффекта определенного внешнего вмешательства или действия, для понимания поведения процесса при определенных условиях и выборе регулировки, способной повлиять на процесс для достижения его цели или уменьшения изменчивости процесса [5, 6, 28, 56, 69–71, 77, 79, 81, 82, 85–87, 90, 92, 96, 104, 107, 109].

Задача моделирования временных рядов в общем виде может быть сформулирована следующим образом [70]. Пусть заданы значения временного ряда  $Y = \{y(1), y(2), \dots, y(N)\}$ , где  $y(t)$  — наблюдение в момент времени  $t$ ,  $t = 1, 2, \dots, N$ . Необходимо построить оценки  $\hat{Y} = \{\hat{y}(N+1), \hat{y}(N+2), \dots, \hat{y}(N+\tau)\}$ , где  $\tau$  ( $1 \leq \tau \leq N$ ) — горизонт прогнозирования.

Существует несколько подходов к моделированию временных рядов [70]: статистический, нейросетевой, нечеткий. Статистический подход отражает статистическую зависимость в наблюдаемом числовом временном ряду, то есть строится модель вида  $y_t = f(x_t, a) + \varepsilon_t$ , где  $f(x_t, a)$  — систематическая компонента,  $a$  — параметр,  $\varepsilon_t$  — ошибки, удовлетворяющие ограничениям модели. Предполагается, что закономерность изменений, выявленная для временного ряда в прошлом, сохранится на некотором отрезке времени в будущем.

Разработаны и применяются множество методов анализа временных рядов и прогнозирования. Часть из этих методов относятся к отдельным приемам или процедурам прогнозирования, другие представляют набор отдельных эвристических алгоритмов [70]. Одной из наиболее распространенных методологий является АРПСС или АRI-МА («авторегрессии проинтегрированного скользящего среднего») [7, 27], которая позволяет выделить регулярные компоненты временного ряда и построить прогноз. Модель класса АRIМА позволяет упростить процесс разработки модели временного ряда, получить модели широкого класса временных рядов (стационарных и нестационарных) с приемлемыми показателями точности, снизить требования к уровню квалификации пользователя и реализована в распространенных статистических пакетах [70].

В нейросетевом подходе [10, 21, 67, 92, 96, 109] задача прогнозирования временных рядов формулируется как задача распознавания образов, причем под распознаванием образов подразумевается способность к обобщению и выделению скрытых зависимостей. Моделирование временных рядов в рамках данного подхода сводится к задаче наилучшего приближения функции  $\hat{y}_{k+1} = \varphi(y_k, \dots, y_{k-n+1}) + \varepsilon_{k+1}$ , где  $\hat{y}_{k+1}$  — прогнозируемое значение,  $y_k, \dots, y_{k-n+1}$  — наблюдаемые значения,  $\varphi(y_k, \dots, y_{k-n+1})$  — некоторая нелинейная функция, параметрической моделью которой служит нейронная сеть,  $\varepsilon_{k+1}$  — ошибка прогноза [70]. Каждый формальный нейрон такой нейронной сети осуществляет преобразование входных сигналов следующим образом [70]:  $y = F(WX)$ , где  $X = (x_1, x_2, \dots, x_n)^T$  — вектор входных сигналов,  $W = (w_1, w_2, \dots, w_n)$  — вектор весовых коэффициентов, который оценивается во время обучения нейронной сети,  $F$  — функция нелинейного преобразования. Архитектура нейронной сети, количество скрытых слоев и число нейронов на скрытых слоях обычно определяются на практике. В частности, если размерность скрытого слоя слишком большая, можно столкнуться с переобучением сети, если слишком мала, то в качестве решения можно получить локальный минимум [70]. Предложены способы настройки числа нейронов, один из них описан в [16].

Согласно [70], использование нейронных сетей при моделировании временных рядов позволяет снизить требования к квалификации разработчика, снижают трудоемкость создания модели. Однако полу-

ченные модели невозможно объяснить в терминах предметной области, для изучения которой получен временной ряд.

Как было отмечено выше, исходные данные выражены на естественном языке, т.е. являются неточными и неполными. Рассмотренные подходы позволяют учитывать только вероятностную неопределенность исходных данных. Рассмотрение же класса нечетких временных рядов позволяет учесть различные виды неопределенности данных. Подробное описание видов неопределенности, возникающей в задаче моделирования социально-значимого поведения респондентов и оценке интенсивности такого поведения, и возможных способах ее обработки можно посмотреть в [46, 64]

Нечеткое моделирование временных рядов позволяет [5, 6, 56, 71, 85, 86] оперировать не только количественными, но и качественными данными, использовать информацию, поступающую в текстовом виде, представлять результаты анализа также в качественной форме.

Согласно [70], коллекция  $Y_t$  называется нечетким временным рядом, если задано универсальное множество  $X_t \subset R^1$ , ( $t = 1, 2, \dots$ ), на котором определены нечеткие множества  $y_t^i$  ( $i = 1, 2, \dots$ ) — элементы коллекции  $Y_t$ . В большинстве случаев последовательные элементы временного ряда зависимы, так что

$$R = \{(y_t, y_{t-1}), (y_{t-1}, y_{t-2}), \dots\} \subseteq Y_t \times Y_{t-1},$$

где  $Y_t$  — переменные,  $y_t$  — наблюдаемые значения этих переменных. Моделью нечеткого временного ряда называют систему отношений  $R$  в выражении  $y_t^j = y_{t-1}^i \circ R_{ij}(t, t-1)$ , где  $\circ$  — операция композиции из теории нечетких множеств,  $y_t^i \in Y_t$ ,  $y_{t-1}^i \in Y_{t-1}$ .

В [69, 70] представлен подробный обзор различных методов моделирования нечетких временных рядов, проблем, возникающих в таком моделировании, подходов к их решению и дальнейших направлений исследования данной тематики.

Однако, несмотря на удобство использования методологии моделирования и анализа временных рядов (особенно при работе с качественными данными), применение этих подходов к решению задачи, связанной с моделированием социально-значимого поведения, сталкивается с рядом ограничений. В частности, такие модели характеризуются невысоким качеством при моделировании коротких временных рядов (причем короткими считаются ряды с количеством наблюдений меньше 40) [7].



Предложены методы работы и с короткими временными рядами, например, метод анализа и прогноза временных рядов «гусеница» [13–15]. Базовый вариант метода состоит в преобразовании одномерного ряда в многомерный с помощью однопараметрической процедуры сдвига, исследовании полученной многомерной траектории с помощью анализа главных компонент и восстановлении ряда по выбранным главным компонентам [13]. Результатом применения метода является разложение временного ряда на простые компоненты: медленные тренды, сезонные и другие периодические или колебательные составляющие, а также шумовые компоненты. Однако даже этот метод не позволяет построить качественную модель при 3–5 наблюдениях. Другое ограничение описанного подхода связано с рассмотрением рядов с равноотстоящими элементами, в то время как интервалы между эпизодами социально-значимого поведения не являются одинаковыми, в общем случае они представляют собой случайные величины.

**5. Гибридные модели.** Во многих случаях при изучении поведения кроме данных об эпизодах поведения становятся известными также дополнительные сведения — психологические, демографические, социальные характеристики, позволяющие лучше описать поведение. Кроме того, могут быть обоснованные предположения о характере поведения, о связях между параметрами. В частности, рискованное поведение во многих исследованиях рассматривается как пуассоновский процесс [74, 84, 89, 99, 104]. Включение таких теоретических предположений позволяет построить более полную модель.

К настоящему моменту при решении задачи оценки интенсивности социально-значимого поведения такое поведение рассматривается нами как случайный процесс определенного класса. При таком предположении были рассмотрены различные типы исходных данных [32, 36, 44, 46–52, 59, 60, 61]. Так выявлены психологические характеристики респондентов, ассоциированные с рискованным поведением [24, 50, 60]. Предложены методы оценки параметров поведения по данным о последних эпизодах [59, 60] и — отдельно — по сведениям о рекордных интервалах между эпизодами [36, 47, 49]. Разработаны модификации методов, учитывающие неточность, неполноту и гранулярность данных [32, 48, 49]. Однако попытки построить модель с одновременным учетом как данных о последних эпизодах, так и сведений о рекордных интервалах между эпизодами приводят к сложной модели [44], которая в свою очередь для ее применения к решению практических задач требует дальнейшего исследования и разработки (или адаптации) методов для численной оценки полученных выражений.

Кроме того, в такую модель очень сложно добавлять новые факторы, новые предположения о предметной области (даже просто использование другого случайного процесса в качестве модели поведения) требуют полного пересчета компонентов модели.

Одним из возможных решений является анализ рассмотренной задачи в рамках подхода, связанного с вероятностно-графическими моделями [51, 52]. Сведение исходной задачи к построению байесовской сети доверия позволит воспользоваться уже существующим мощным алгоритмическим аппаратом теории байесовских сетей доверия. В этом случае появление большего числа наблюдений, скрытых переменных, необходимость отобразить более сложные связи сведется к частичной модификации построенной ранее модели: к появлению большего числа узлов, установлению новых связей, пересчету некоторых условных вероятностей.

**6. Заключение.** Необходимость оценки интенсивности поведения и нахождения ее связей с другими параметрами исследуемых объектов возникает в современных науках о человеке и обществе при решении многих задач. В данной работе предложен обзор возможных подходов к построению моделей социально-значимого поведения по сверхмалой неполной совокупности наблюдений и оценке параметров таких моделей. Рассмотрены такие методы как агентное моделирование, статистический подход, включая методы анализа малой выборки, методология временных рядов и их приложения к решению указанной задачи. Выявлены преимущества использования моделей в рамках теории вероятностных графических моделей.

## Литература

1. *Азаров А.А., Тулупьева Т.В., Тулупьев А.Л.* Агентоориентированный подход к моделированию комплекса «Информационная Система – Персонал – Злоумышленник» в задачах оценки защищенности от социо-инженерных атак // Список-2012: Материалы всероссийской научной конференции по проблемам информатики (25–27 апреля 2012 г., Санкт-Петербург). СПб.: ВВМ, 2012. С. 374–377.
2. *Азаров А.А., Тулупьева Т.В., Фильченко А.А., Тулупьев А.Л.* Вероятностно-реляционный подход к представлению модели комплекса «информационная система – персонал – критичные документы» // Труды СПИИРАН. 2012. Вып. 1(20). С. 57–71.
3. Анализ временных рядов URL: <http://www.statsoft.ru/home/textbook/modules/st-timser.html> (дата обращения 15.01.2013)
4. *Афанасьев И.В.* Возможности математического моделирования поведения аудитории с помощью динамических математических моделей // Актуальные проблемы современной науки. 2006. № 4. С. 212–218.
5. *Афанасьева Т.В., Ярушклина Н.Г.* Анализ эффективности модели нечеткой тенденции в прогнозировании временных рядов // Автоматизация процессов управления. 2011. № 4. С. 43–49.

6. *Берштейн Л.С., Ковалев С.М., Муравский А.В.* Модели представления нечетких темпоральных знаний в базах временных рядов // Известия Российской академии наук. Теория и системы управления. 2009. № 4. С. 130–141.
7. *Бокс Дж., Дженкинс Т.* Анализ временных рядов. Прогноз и управление. М.: Мир, 1974. 406 с.
8. *Боровков А.А.* Математическая статистика. Новосибирск: Наука; Издательство института информатики, 1997. 772 с.
9. *Былова Н.К., Посвалок Н.Э., Савин С.З., Юсупова М.В.* Информационное моделирование медико-социальных аспектов поведения потребителей психоактивных веществ // Успехи современного естествознания № 11 2004. С. 88–89
10. *Бэйтс Д.Э., Ван ден Берг В.М., Вуд Д.* Нейронные сети и финансовые рынки: принятие решений в торговых операциях. Москва: ТВП, 1997. 236 с.
11. *Виноградов А.Н., Жилкова Л.Ю., Осипов Г.С.* Динамические интеллектуальные системы. Моделирование целенаправленного поведения // Известия Российской академии наук. Теория и системы управления. 2003. № 1. С. 87–94.
12. *Галочкин И.В.* Социальные предпочтения в экономическом поведении: методы измерения и моделирования // Экономика и математические методы. 2010. Т. 46. № 3. С. 82–92.
13. Главные компоненты временных рядов: метод «Гусеница» / Под редакцией Данилова Д.Л. и Жиглявского А.А.. СПб.: СПбГУ, 1997. 308 с.
14. *Голяндина Н.Э.* Метод "Гусеница"-SSA: анализ временных рядов: Учебное пособие. СПб.: , 2004. 76 с.
15. *Голяндина Н.Э.* Метод "Гусеница"-SSA: прогноз временных рядов: Учебное пособие. СПб.: , 2004. 52 с.
16. *Горбань А. Н.* Обучение нейронных сетей. М.: СП «ПараГраф», 1990. 159 с.
17. *Городецкий В.И.* Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор (часть 1) // Искусственный интеллект и принятие решений. 2011. № 2. С. 19–30.
18. *Городецкий В.И.* Теория, модели, инфраструктуры и языки спецификации командного поведения автономных агентов. Обзор (часть 2) // Искусственный интеллект и принятие решений. 2011. № 3. С. 34–47.
19. *Городецкий В.И., Карсаев О.В., Самойлов В.В., Серебряков С.В.* Открытые сети агентов // Труды СПИИРАН. 2007. № 4. С. 11–35.
20. *Гребенников Р.В.* Разработка индивидуальных характеров поведения при моделировании толпы // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2008. № 2. С. 107–110.
21. *Ежов А.А., Шумский С.А.* Нейрокомпьютинг и его применения в экономике и бизнесе. М.: МИФИ, 2002. 306 с.
22. *Жилкин С.Д.* Результаты применения алгоритмов моделирования программного обеспечения с целью выявления аномалий поведения // Информационное противодействие угрозам терроризма. 2010. № 14. С. 106–110.
23. *Задков А.В.* Потребительское поведение: составляющие, методы изучения и моделирования // Маркетинг и маркетинговые исследования. 2009. № 3. С. 240–247.
24. *Зельтерман Д., Суворова А.В., Пащенко А.Е., Мусина В.Ф., Тулупьев А.Л., Тулупьева Т.В., Гро Л.Е., Хаймер Р.* Диагностика регрессионных уравнений в анализе интенсивности рискованного поведения по его последним эпизодам // Труды СПИИРАН. 2011. Вып. 17. С. 33–46.
25. *Иващенко А.В., Карсаев О.В., Скобелев П.О., Царев А.В., Юсупов Р.М.* Мультиагентные технологии для разработки сетевых систем управления // Известия Южного федерального университета. Технические науки. 2011. Т. 116. № 3. С. 11–23.

26. *Каледин О.Е.* Математическое моделирование управляемых процессов на основе статистических данных: дисс. ... канд. ф.-м. наук: 05.13.18. Саранск, 2011. 127 с.
27. *Канторович Г.Г.* Анализ временных рядов // Экономический журнал ВШЭ. 2002. № 1–2.
28. *Ковалев С.М.* Гибридные коннекционистские модели извлечения темпоральных знаний // Интегрированные модели и мягкие вычисления в искусственном интеллекте. V-я Международная научно-практическая конференция. Сборник научных трудов. В 2-х т. Т. 1. С. 30–40.
29. *Корниенко А.А., Корниенко А.В.* Логико-лингвистическое моделирование поведения социально-экономических субъектов // Известия Томского политехнического университета. 2005. Т. 308. № 3. С. 173–177.
30. *Котенко И.В.* Многоагентное моделирование для исследования механизмов защиты информации в сети Интернет // Имитационное моделирование. Теория и практика (ИММОД). Сборник докладов. СПб., 2009. С. 38–47. Также доступно URL: <http://www.spiras.nw.ru/files/conferences/IMMOD-2009/IMMOD-2009-1.pdf> (дата обращения 15.01.2013)
31. *Кузнецов О.П.* Интеллектуализация поддержки управляющих решений и создание интеллектуальных систем // Проблемы управления. 2009. № 3. С. 64–72.
32. *Лавренко А.В., Суворова А.В., Пащенко А.Е., Тулупьева Т.В., Тулупьев А.Л.* Вероятностно-нечеткая модель социально-значимого поведения в случае неточных данных об его эпизодах, предшествующих интервью // Гибридные и синергетические интеллектуальные системы: теория и практика. Материалы 1-го международного симпозиума. Т. 2. Калининград: Изд-во БФУ им. И.Канта, 2012. С. 19–29.
33. *Лукашова Е.В.* Постановка проблемы моделирования потребительского поведения с позиции системного анализа // Известия Санкт-Петербургского университета экономики и финансов. 2008. № 1. С. 116–119.
34. Методы и модели анализа временных рядов : метод. указания к лаб. работам / сост. С.И. Татаренко. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2008. 32 с.
35. *Осипов Г.С.* Планирование и моделирование целенаправленного поведения в интеллектуальных системах // Вестник Российского университета дружбы народов. Серия: Прикладная и компьютерная математика. 2002. № 1. С. 51–61.
36. *Пащенко А.Е., Суворова А.В., Тулупьева Т.В., Тулупьев А.Л.* Вероятностные распределения порядковых статистик в анализе сверхкоротких нечетких и неполных временных рядов // Труды СПИИРАН. 2009. Вып. 10. СПб.: Наука, 2009. С. 184–207.
37. *Петросян Д.С.* Интегративная модель поведения человека // Общественные науки и современность. 2008. № 3. С. 39–51.
38. *Петросян Д.С.* Концептуальные и математические модели поведения человека как экономического агента // Аудит и финансовый анализ, 2009. № 1.
39. *Плавинский С.Л., Баранова А.Н., Разнатовский К.И.* Сексуальное поведение, венерические болезни и гетеросексуальная эпидемия ВИЧ-инфекции — некоторые результаты математического моделирования // Российский семейный врач. 2007. Т. 11. № 3. С. 30–38.
40. *Рассел С., Норвиг П.* Искусственный интеллект: современный подход, 2-е изд. М.: Издательский дом «Вильямс», 2006. 1408 с.
41. *Самарский А.А., Михайлов А.П.* Математическое моделирование: Идеи, методы, примеры. М.: ФИЗМАТЛИТ, 2005. 320 с.
42. *Скажун С.В.* Математическое моделирование поведения пользователей компьютерных систем // Математичні машини і системи, 2005. № 2. С. 122–129.

43. *Соколов С.Н.* Моделирование поведения пользователей интернет-ресурсов на основе смеси цепей Маркова // *Естественные и технические науки.* 2009. № 5. С. 302-305.
44. *Степанов Д.В., Мусина В.Ф., Суворова А.В., Тулупьев А.Л., Сироткин А.В., Тулупьева Т.В.* Функция правдоподобия с гетерогенными аргументами в идентификации пуассоновской модели рискованного поведения в случае информационного дефицита // *Труды СПИИРАН.* 2012. Вып. 23. С. 157–184.
45. *Стефанюк В.Л., Жожикашвили А.В.* Наследование свойств категории при переходе от статических систем, использующих знания, к динамическим // *Искусственный интеллект и принятие решений.* 2008. № 1. С. 4–14.
46. *Суворова А.В.* Подходы к представлению и обработке неопределенности данных и знаний о поведении индивидов // *Труды СПИИРАН.* 2012. Вып. 23. С. 206–222.
47. *Суворова А.В., Пащенко А.Е., Тулупьева Т.В.* Оценка характеристик сверхкороткого временного ряда по гранулярным данным о рекордных интервалах между собитиями // *Труды СПИИРАН.* 2010. Вып. 12. С. 170–181.
48. *Суворова А.В., Тулупьев А.Л., Пащенко А.Е., Тулупьева Т.В., Красносельских Т.В.* Анализ гранулярных данных и знаний в задачах исследования социально значимых видов поведения // *Компьютерные инструменты в образовании.* №4. 2010. С. 30–38.
49. *Суворова А.В., Тулупьев А.Л., Тулупьева Т.В., Пащенко А.Е.* Эвристическая оценка интенсивности поведения по рекордным интервалам между эпизодами: обработка неточности ответов респондентов // *XV Международная конференция по мягким вычислениям и измерениям. SCM-2012.* (25–27 июня 2012 г. Санкт-Петербург). Сборник докладов. 2012. Т. 2. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2012. С. 101–104.
50. *Суворова А.В., Тулупьева Т.В., Тулупьев А.Л.* Обобщенная линейная регрессионная модель для прогноза временного интервала между последним эпизодом рискованного поведения и моментом интервью на основе социально-демографических и психологических особенностей // *Труды СПИИРАН.* 2012. Вып. 2 (21). С. 80–94.
51. *Суворова А.В., Тулупьева Т.В., Тулупьев А.Л., Сироткин А.В.* Применение байесовских сетей доверия для моделирования угрожающего поведения индивида по неполным и неточным данным // *Тринадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2012* (16–20 октября 2012 г., г. Белгород). Труды конференции. Т. 3. Белгород: Изд-во БГТУ, 2012. С. 292–299.
52. *Суворова А.В., Тулупьева Т.В., Тулупьев А.Л., Сироткин А.В., Пащенко А.Е.* Вероятностные графические модели социально-значимого поведения индивида, учитывающие неполноту информации // *Труды СПИИРАН.* 2012. Вып. 22. С. 101–112.
53. *Тарасов В.Б.* От логических к диалогическим решеткам и бирешеткам: использование в теории агентов // *Вестник Ростовского государственного университета путей сообщения.* 2011. № 3. С. 129–141.
54. *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям. М.: Эдиториал УРСС, 2002. 352 с.
55. *Тарасов В.Б., Афонин П.В., Картежников Д.* Модель формирования структур объединений виртуальных предприятий на основе популяции эволюционирующих агентов // *Известия Южного федерального университета. Технические науки.* 2007. Т. 73. № 1. С. 55–58.
56. *Теория и практика нечетких гибридных систем/Под ред. Н.Г. Ярушкиной.* М.: ФИЗМАТЛИТ, 2007. 208 с.
57. *Трифопова Е.Н.* Учет объективных и субъективных факторов в процессе моделирования поведения потребителей продовольственных товаров // *Региональные агрорисистемы: экономика и социология.* 2008. № 1. С. 10–10.

58. *Тулупьев А.Л., Азаров А.А., Пащенко А.Е.* Информационные модели компонент комплекса «Информационная система – персонал», находящегося под угрозой социоинженерных атак // Труды СПИИРАН. 2010. Вып. 3 (14). С. 50–57.
59. *Тулупьева Т.В., Пащенко А.Е., Мусина В.Ф., Тулупьев А.Л., Азаров А.А., Жук С.Н., Казакова О.С., Красносельских Т.В., Сироткин А.В., Суворова А.В., Фильченков А.А.* Отчет о научно-исследовательской работе «Классификация ответов респондентов о последних эпизодах рискованного поведения и косвенная оценка его интенсивности» (заключительный), инвентарный № 02201259425 от 2012.06.26, по теме «Моделирование и измерение количественных характеристик ВИЧ-рискованного поведения на основе обработки ответов респондентов», регистрационный № 01201262070. СПб.: СПИИРАН, 2012. 34 с. (Депонировано в ЦИТиС.)
60. *Тулупьева Т.В., Пащенко А.Е., Мусина В.Ф., Тулупьев А.Л., Жук С.Н., Азаров А.А., Суворова А.В., Сироткин А.В., Фильченков А.А.* Отчет о научно-исследовательской работе «Опросный инструментарий для выявления особенностей рискованного поведения в контексте адаптивных стилей и анализ результатов пилотного исследования» (промежуточный), инвентарный № 02201259423 от 2012.06.26, по теме «Взаимосвязь адаптивных стилей ВИЧ-инфицированных и степени рискованности их поведения», регистрационный № 01201262071. СПб.: СПИИРАН, 2012. 78 с. (Депонировано в ЦИТиС.)
61. *Тулупьева Т.В., Пащенко А.Е., Тулупьев А.Л., Красносельских Т.В., Казакова О.С.* Модели ВИЧ-социально-значимого поведения в контексте психологической защиты и других адаптивных стилей. СПб.: Наука, 2008. 140 с.
62. *Фильченков А.А.* Алгебраическая байесовская сеть как основа для медицинской диагностической модели // «Математическое и компьютерное моделирование в биологии и химии. Перспективы развития». Сборник трудов I Международной интернет-конференции. Казань: Из-во «Казанский университет». 2012. С. 162–166.
63. *Фильченков А.А.* Математическое моделирование диагностической модели защищенности информационной системы на основе комбинирования неполной и неточной аналитической информации // VII Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России (ИБРР-2011)». Материалы конференции. СПб.: СПОИСУ. 2011. С. 175–176.
64. *Фильченков А.А.* Меры истинности и вероятностные графические модели для представления знаний с неопределенностью // Труды СПИИРАН. 2012. Вып. 23. С. 254–295.
65. *Ширяев А.Н.* Вероятность: Учеб. пособ. для вузов. 2-е изд. М.: Наука. Гл. ред. физ.-мат. лит., 1989. 640 с.
66. *Ширяев Н.* Основы стохастической финансовой математики В 2-х т.-М.: ФАЗИС, 1998.
67. *Шугай Ю.С.* Разработка нейросетевых алгоритмов анализа многомерных временных рядов и их применение при исследовании солнечно-земных связей: дисс... к.ф.-м.н.: 05.13.18. М., 2006. 137 с.
68. *Юсупов Р.М., Соколов Б.В.* Имитационное моделирование и его применение в науке и промышленности // Вестник Российской академии наук. 2008. Т. 78. № 5. С. 471–472.
69. *Ярушкина Н.Г.* Современный интеллектуальный анализ нечетких временных рядов // Интегрированные модели и мягкие вычисления в искусственном интеллекте. V-я Международная научно-практическая конференция. Сборник научных трудов. В 2-х т. Т. 1. С. 19–29.
70. *Ярушкина Н. Г., Афанасьева Т. В., Перфильева И. Г.* Интеллектуальный анализ временных рядов: Учебное пособие. Ульяновск: УлГТУ, 2010. 320 с.

71. Ярушкіна Н.Г., Перфильева И.Г., Афанасьева Т.В. Интегральный метод нечеткого моделирования и анализа нечетких тенденций // Автоматизация процессов управления. 2010. № 2. С. 59–63.
72. Adams J., Moody J., Morris M. Sex, Drugs, and Race: How Behaviors Differentially Contribute to the Sexually Transmitted Infection Risk Network Structure // American Journal of Public Health. 2013, Vol 103, No. 2. P. 322–329.
73. Andonie R., Fabry-Asztalos L., Badi' Abdul-Wahid C., Abdul-Wahid S., Barker G.I., Magill L.C. Fuzzy ARTMAP Prediction of Biological Activities for Potential HIV-1 Protease Inhibitors Using a Small Molecular Data Set // IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2011, Vol. 8, No. 1. P. 80–93.
74. Anupindi R., Dada M., Gupta S. Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products // Marketing Science. 1998. Vol. 17, No. 4. P. 406–423.
75. Anylogic Simulation Software. URL: <http://www.anylogic.com/> (дата обращения 15.01.2013)
76. Axelrod R., Tesfatsion L. On-Line Guide for Newcomers to Agent-Based Modeling in the Social Sciences URL: <http://www2.econ.iastate.edu/tesfatsi/abmread.htm>
77. Azadeh A., Saberi M., Gitiforouz A. An integrated simulation-based fuzzy regression-time series algorithm for electricity consumption estimation with non-stationary data // Journal of the Chinese Institute of Engineers, 2011, 34:8. P. 1047–1066.
78. Bailly N., Maitre I., Amanda M., Herve C., Alaphilippe D. The Dutch Eating Behaviour Questionnaire (DEBQ). Assessment of eating behaviour in an aging French population // Appetite, 2012. Vol. 59. P. 853–858.
79. Batyrshin I. Up and Down Trend Associations in Analysis of Time Series Shape Association Patterns. Lecture Notes on Computer Science. 2012. Vol. 7329. P. 246–254.
80. Bennett J., Greene G., Schwartz-Barcott D. Perceptions of emotional eating behavior. A qualitative study of college students // Appetite. 2013. Vol. 60. P. 187–192.
81. Bernat D., Maldonado-Molina M., Hyland A., Wagenaar A. Effects of Smoke-Free Laws on Alcohol-Related Car Crashes in California and New York: Time Series Analyses From 1982 to 2008 // American Journal of Public Health; 2013, Vol. 103 Issue 2. P. 214–219.
82. Bezruchko B., Smirnov D. Extracting Knowledge From Time Series: An Introduction to Nonlinear Empirical Modeling. Berlin: Springer, 2010. 407 p. Также доступно URL: <http://link.springer.com/book/10.1007/978-3-642-12601-7/page/1> (дата обращения 15.01.2013)
83. Case K.K., Ghys P.D., Gouws E., Eaton J.W., Borquez A., Stover J., Cuchi P., Abu-Raddad L.J., Garnett G.P., Hallett T.B. Understanding the modes of transmission model of new HIV infection and its use in prevention planning // Bulletin of the World Health Organization. 2012. Vol. 90. No. 11. P. 831–838A.
84. Cruyff M., Bockenholt U., Hout A. van den , Heijden P. van der Accounting for Self-Protective Responses in Randomized Response Data from a Social Security Survey Using the Zero-Inflated Poisson Model // The Annals of Applied Statistics, 2008, Vol. 2, No. 1, P. 316–331.
85. Domanska D., Wojtylak M. Application of fuzzy time series models for forecasting pollution concentrations // Expert Systems with Applications, 2012, 39. P. 7673–7679.
86. Dubois D. Ranking Fuzzy Numbers in the Setting of Possibility Theory // Information Sciences, 1983, 30. P. 183–224.
87. Duchene F., Garbay C., Rialle V. Learning recurrent behaviors from heterogeneous multivariate time-series // Artificial Intelligence in Medicine, 2007, 39. P. 25–47.

88. *Eagena B., Caron R., Abdul-Kader W.* An Agent-Based Modelling Tool (ABMT) for scheduling diagnostic imaging machines // *Technology & Health Care*. 2010. Vol. 18. Issue 6. P. 409–415.
89. *Holmes W.C., Foa E.B., Sammel M.D.* Men's Pathways to Risky Sexual Behavior: Role of Co-Occurring Childhood Sexual Abuse, Posttraumatic Stress Disorder, and Depression Histories // *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 2005, Vol. 82, No. 1, Supplement 1. doi:10.1093/jurban/jti028.
90. *Kadirov D.* Macro-Systems Role of Marketing: Do We Trade Environment for Welfare? // *Journal of Macromarketing*, 2011, 31(4). P. 359–375.
91. *Kutner M., Neter J., Nachtsheim C., Li W.* *Applied Linear Statistical Models*. Fifth edition. McGraw-Hill Inc., 2004. 1370 pp.
92. *Lohani A.K., Kumar R., Singh R.D.* Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques // *Journal of Hydrology*, 2012.
93. *Moore D., Dray A., Green R., Hudson S.L., Jenkinson R., Siokou C., Perez P., Bammer G., Maher L., Dietze P.* Extending drug ethno-epidemiology using agent-based modelling // *Addiction*. 2009. Vol. 104. Issue 12. P.1991–1997.
94. *Motsa S., Sibanda P.* Agent-Based Modeling and Genetic Algorithm Simulation for the Climate Game Problem // *Mathematical Problems in Engineering*. 2012. Vol. 2012. Special section. P.1–14.
95. *Muller A., Mitchell J., Crosby R., Cao L., Johnson J., Claes L., Zwaan M.* Mood states preceding and following compulsive buying episodes: an ecological momentary assessment study // *Psychiatry Research*, 2012. No.200. P.575–580.
96. *Nayaka P.C., Sudheer K.P., Ranganc D.M., Ramasastri K.S.* A neuro-fuzzy computing technique for modeling hydrological time series // *Journal of Hydrology*, 2004, 291. P. 52–66.
97. *Oksendal B.* *Stochastic Differential Equations: An Introduction with Applications*. Berlin: Springer, 2003.
98. *Perrin D., Ruskin H.J., Crane M.* Model refinement through high-performance computing: an agent-based HIV example // *Immunome Research*. 2010 Supplement 1, Vol. 6. P.1–9.
99. *Rhodes W., Kling R., Johnston P.* Using Booking Data to Model Drug User Arrest Rates: A Preliminary to Estimating the Prevalence of Chronic Drug Use // *J Quant Criminol*, 2007, 23. P. 1–22.
100. *Rothman K.J.* *Epidemiology: An Introduction*. Oxford etc.: Oxford University Press, 2002. 223 p.
101. *Rubin D.B.* A Small Sample Correction for Estimating Attributable Risk in Case-Control Studies // *The International Journal of Biostatistics*, 2010, Vol. 6, Iss. 1, Article 32. DOI: 10.2202/1557-4679.1252.
102. *Saaty T.L., Alexander J.M.* *Thinking with models: Mathematical models in the Physical, Biological and Social Sciences*. N.Y.: Pergamon Press, 1981. 181 p.
103. *Sengupta R.R., Bennett D.A.* Agent-based modelling environment for spatial decision support // *International Journal of Geographical Information Science*. 2003. Vol. 17. Issue 2. P.157–180.
104. *Tan J.Y., Bong D.B., Rigit A.R.* Time Series Prediction using Backpropagation Network Optimized by Hybrid K-means-Greedy Algorithm // *Engineering Letters*. 2012. Vol. 20, Issue 3. P. 203–210.
105. *Trujols J., Guardia J., Pero M., Freixa M., Sinol N., Tejero A., Cobos J.* Multi-episode survival analysis: An application modelling readmission rates of heroin dependents at an inpatientdetoxification unit // *Addictive behavior*, 2007. Vol. 32. P. 2391–2397.



106. *Van Vliet C., Van der Ploeg C., Kidula N., Malonza I., Tyndall M., Nagelkerke N.* Estimating Sexual Behavior Parameters from Routine Sexual Behavior Data // *The Journal of Sex Research*, 1998, Vol. 35, No. 3. P. 298–305.
107. *Vanli O.A., Zhang C., Wang B.* An adaptive Bayesian approach for robust parameter design with observable time series noise factors // *IE Transactions*, 2013, 45:4. P. 374–390.
108. *Wardle J.* Eating style. A validation study of the dutch eating behaviour questionnaire in normal subjects and women with eating disorders // *Journal of Psychosomatic Research*, 1987. Vol. 31, No. 2. P. 161–169.
109. *Wei S., Zuo D., Song J.* Improving prediction accuracy of river discharge time series using a Wavelet-NAR artificial neural network // *Journal of Hydroinformatics*. 2012. Vol. 14 Issue 4. P. 974–992.
110. *Winter J. C. F., Dodou D., Wieringa P.A.* Exploratory Factor Analysis With Small Sample Sizes // *Multivariate Behavioral Research*, 2009, 44:2. P. 147–181.
111. *Xenitidou M., Elsenbroich C.* Construct Validity and Theoretical Embeddedness of Agent-based Models of Normative Behaviour // *International Journal of Interdisciplinary Social Sciences*. 2010. Vol. 5. Issue 4. P.67–79.

**Суворова Алена Владимировна** — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН, аспирант математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов: математическая статистика, теория вероятности, применение методов математического моделирования в эпидемиологии. Число научных публикаций — 40. suvalv@gmail.com; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научный руководитель — А.Л. Тулупьев.

**Suvorova Alena Vladimirovna** — junior researcher, Laboratory of Theoretical and Interdisciplinary Computer Science, SPIIRAS, PhD student, Faculty of Mathematics and Mechanics of St. Petersburg State University (SPbSU). Research interests: mathematical statistics, probability theory, application of mathematical modeling in epidemiology. The number of publications — 40. suvalv@gmail.com; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientific advisor — A.L. Tulupjev.

**Поддержка исследований.** Статья содержит материалы исследований, частично поддержанных грантами РФФИ 09-01-00861-а, 10-01-00640-а, 12-01-00945-а. Также исследования поддержаны грантом для молодых ученых и кандидатов наук Комитета по науке и высшей школе Правительства Санкт-Петербурга «Модели и алгоритмы анализа сверхкоротких неточных временных рядов на основе гранулярных данных и знаний», руководитель — А.В. Суворова (2012).

Рекомендовано ТИМПИ СПИИРАН, зав. лаб. А.Л. Тулупьев, д.ф.-м.н., доцент.  
Статья поступила в редакцию 25.01.2013.

## РЕФЕРАТ

### *Суворова А.В.* **Гибридные модели оценки параметров социально-значимого поведения по сверхмалой неполной совокупности наблюдений.**

Во многих отраслях социологических, психологических, маркетинговых исследований возникают задачи оценивания интенсивности социально-значимого поведения респондентов. При этом встает задача оценки интенсивности социально-значимого поведения респондента по ответам на блок вопросов. Однако из-за особенностей припоминания таким способом можно получить сведения лишь о небольшом числе последних эпизодов и о рекордных интервалах между ними, что приводит к необходимости анализа неполных данных. Исходные данные, то есть сведения об интервалах между несколькими последними эпизодами поведения, о минимальном и максимальном интервалах между последовательными эпизодами поведения, являются членами вариационного ряда. Однако в этом случае мы получаем вариационный ряд «с пропусками» — известен его первый член (минимальный интервал), последний (максимальный интервал) и несколько элементов между ними, причем с неизвестными номерами.

Цель данной работы — описать возможные подходы к построению моделей социально-значимого поведения по сверхмалой неполной совокупности наблюдений и оценке параметров таких моделей.

Рассмотрены такие методы как агентное моделирование, статистический подход, включая методы анализа малой выборки, методология временных рядов и их приложения к решению указанной задачи.

При изучении возможности применения статистического подхода востребованной оказывается математическая методология для небольшой выборки. Однако, в работах, описывающих методы анализа на основе небольших выборок, «небольшой» считается выборка, включающая 10–30 наблюдений, в то время как в описанной во введении проблеме речь идет о 3–5 эпизодах. Такое же ограничение появляется и при рассмотрении методологии анализа временных рядов.

Кроме того, в рассматриваемые модели очень сложно добавлять новые факторы, новые предположения о предметной области (психологические, демографические, социальные характеристики, позволяющие лучше описать поведение). Одним из возможных решений такой проблемы является анализ рассмотренной задачи в рамках подхода, связанного с вероятностно-графическими моделями. Сведения исходной задачи к построению байесовской сети доверия позволит воспользоваться уже существующим мощным алгоритмическим аппаратом теории байесовских сетей доверия.

## SUMMARY

### *Suvorova A. V.* **Hybrid models for parameter estimates of socially significant behavior on the base of supershort incomplete set of observations.**

In many fields of sociological, psychological and marketing research, we face the problem of socially significant behavior rate or frequency estimate. We need to estimate behavior rate on the base of the responses to the questionnaire or the results of the interview. Questionnaire produces data about several behavior episodes only and about extreme intervals between them. As a result we have to analyze incomplete data. Initial data (several last episodes and extreme intervals between episodes) are items in ordered sample, but it is ordered sample with gaps: we know the first item (minimum interval), the last one (maximum intervals) and several items with unknown order.

We provide a description of the methods for socially significant behavior modeling and estimation on the base of supershort incomplete set of observations. We consider agent-based modeling, statistical approach including small samples analysis, time series analysis and their application to the described problem.

Application of the statistical approaches is limited with the fact of very small sample size. Classical methods for small sample analysis are applicable to the samples with 10–30 items but procedures for rate estimate on the basis of the data about behavior episodes deal with 3–5 items. Time series analysis has the same restriction.

Note that these models do not provide a simple way to include additional factors such as psychosocial, demographical characteristics. One of possible ways to solve this problem is to consider a problem of socially significant behavior rate estimates in terms of probabilistic graphical models. Such formal description of the problem allows applying powerful methods and developed algorithms of the theory of Bayesian belief networks. We can use existing software to make computational simulations and apply the model to solve practical tasks.