

ИЗ РАБОЧЕЙ ТЕТРАДИ  
ИССЛЕДОВАТЕЛЯ

СУПЕРКОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В ОБЩЕСТВЕННЫХ НАУКАХ:  
АГЕНТ-ОРИЕНТИРОВАННЫЕ ДЕМОГРАФИЧЕСКИЕ МОДЕЛИ

© 2016 г. В.Л. Макаров<sup>а</sup>, А.Р. Бахтизин<sup>а</sup>, Е.Д. Сушко<sup>а</sup>, В.А. Васенин<sup>б</sup>,  
В.А. Борисов<sup>б</sup>, В.А. Роганов<sup>б</sup>

<sup>а</sup> Центральный экономико-математический институт РАН, Москва, Россия

<sup>б</sup> Московский государственный университет им. М.В. Ломоносова, Москва, Россия

e-mail: makarov@cemi.rssi.ru; albert.bakhtizin@gmail.com; sushko\_e@mail.ru; vassenin@msu.ru;  
wazzy.msu@gmail.com; radug-a@ya.ru

Поступила в редакцию 04.08.2015 г.

Настоящая статья является продолжением работы, опубликованной в “Вестнике РАН” (2016, № 3). В предыдущей статье был проанализирован международный опыт подготовки и использования агент-ориентированных моделей и технические наработки по их реализации на суперкомпьютерах. Здесь описываются две мультиагентные демографические модели, построенные в Центральном экономико-математическом институте РАН. Модели различаются уровнем детализации при имитации репродуктивного поведения людей. Анализируются этапы и методы эффективного отображения счётного ядра мультиагентной системы на архитектуру современного суперкомпьютера с использованием разработанной авторами технологии поддержки агент-ориентированного моделирования для суперкомпьютеров – STARS (Supercomputer Technology for Agent-oriented Simulation).

**Ключевые слова:** агент-ориентированные модели, суперкомпьютерные технологии, параллельные вычисления, демографические модели.

DOI: 10.7868/S086958731605008X

В марте 2011 г. на суперкомпьютере “Ломоносов” была запущена агент-ориентированная модель (АОМ), имитирующая развитие социально-экономической системы России на протяжении ближайших 50 лет. Реализованная АОМ основана на взаимодействии 100 млн. агентов, условно представляющих социально-экономическую среду России. Поведение каждого агента задано набором алгоритмов, которые описывают его действия и взаимодействие с другими агентами в реальном мире. В проекте участвовали авторы настоящей статьи. Данные для моделирования были предоставлены Федеральной службой государственной статистики и Российским мониторингом экономического положения и здоровья населения. Модель для обычного компьютера была построена в 2009 г., в 2011 г. она конвертирована в суперкомпьютерную версию [1]. В этой версии отсутствовало межагентное взаимодействие, что, с одной стороны, не позволяло в пол-

ной мере использовать преимущества агентного подхода, а с другой – значительно упрощало распараллеливание программного кода. В последней версии агенты общаются между собой, что повлекло изменение технологии распараллеливания модели, а также других программных библиотек.

К стоявшим перед нами задачам относилось не только рассмотрение различных методов распараллеливания мультиагентных систем для запуска их на суперкомпьютерах, но и сопоставление показателей эффективности этих методов в зависимости от архитектуры мультиагентной системы, численности популяции агентов и числа используемых ядер вычислительных кластеров. Следовательно, на первом этапе исследования необходимо было разработать такие мультиагентные модели, которые:

- моделировали бы реально происходящие в социально-экономических системах процессы,

МАКАРОВ Валерий Леонидович – академик РАН, директор ЦЭМИ РАН. БАХТИЗИН Альберт Рауфович – доктор экономических наук, заведующий лабораторией ЦЭМИ РАН. СУШКО Елена Давидовна – кандидат экономических наук, ведущий научный сотрудник ЦЭМИ РАН. ВАСЕНИН Валерий Александрович – доктор физико-математических наук, заведующий отделом НИИ механики МГУ им. М.В. Ломоносова. БОРИСОВ Василий Александрович – студент механико-математического факультета МГУ им. М.В. Ломоносова. РОГАНОВ Владимир Александрович – старший научный сотрудник НИИ механики МГУ им. М.В. Ломоносова.

где действуют самостоятельные различающиеся между собой акторы;

- нуждались бы в использовании суперкомпьютеров для проведения численных экспериментов, поскольку моделируемые процессы являются массовыми и требуют для адекватного представления в модели создания популяции агентов очень большой численности;

- отличались бы друг от друга какими-либо особенностями архитектуры, которые имеют существенное значение для выбора методов их распараллеливания и реализации на суперкомпьютере.

На втором этапе следовало реализовать и апробировать разработанные конструкции моделей на обыкновенном персональном компьютере для популяции агентов с максимально возможной численностью. Наконец, на третьем этапе нужно было, используя различные методы распараллеливания, каким-то образом преобразовать созданные модели для переноса их на суперкомпьютер, открывая тем самым перспективу значительного увеличения популяции агентов. Первоначальные модели и результаты проведенных на них экспериментов, во-первых, играют роль контрольных примеров, позволяющих существенно сократить время на отладку и верификацию суперкомпьютерных моделей, во-вторых, служат базой для сравнения показателей эффективности выбранных методов распараллеливания.

Классическим примером массовых социально-экономических процессов, в которых каждый из участников под влиянием своих интересов и в соответствии со своими возможностями может действовать самостоятельно, следует признать демографические процессы. Неслучайно поэтому в литературе можно найти описания многочисленных мультиагентных систем, ориентированных на отражение самых разнообразных аспектов основных демографических процессов, таких как смертность, рождаемость и миграция [2–5].

## АГЕНТ-ОРИЕНТИРОВАННЫЕ МОДЕЛИ ДЕМОГРАФИЧЕСКОГО ПЕРЕХОДА

**Концепция разработанных демографических моделей.** Предложенные нами демографические мультиагентные модели предназначены для рассмотрения такого известного явления, как демографический переход, и связанных с ним структурных изменений в обществе. Демографическим переходом принято называть переход от традиционного типа воспроизводства населения с высокими показателями смертности и рождаемости к современному — с низкими уровнями этих показателей. Особенности и этапы данного феномена подробно рассматривались в работах многих ведущих демографов, как зарубежных, так и отечественных, здесь достаточно сослаться на известную работу А.Г. Вишневого [6].

Наблюдаемыми признаками демографического перехода являются существенные изменения в показателях смертности и рождаемости. С развитием общества, повышением его благосостояния и улучшением системы здравоохранения показатели смертности (особенно младенческой) снижаются. Происходящее одновременно изменение уровня рождаемости вызвано изменением стиля жизни, в первую очередь женщин — их эмансипацией, повышением уровня образования, появлением личностных целей, не связанных с созданием семьи. Одним из ярких симптомов этого может служить набирающее силу в развитых и благополучных в экономическом отношении странах движение “child free” — полный отказ от рождения детей. В итоге в современном обществе суммарный коэффициент рождаемости (среднее число детей, рождаемых женщиной в течение репродуктивного периода) зачастую снижается до уровня, не обеспечивающего простое воспроизводство населения. Следует вместе с тем подчеркнуть, что изменение уровней смертности и рождаемости происходит хотя и параллельно, но не синхронно — снижение рождаемости отстаёт от снижения смертности на одно-два поколения.

Описанные процессы влекут за собой не только сокращение общей численности населения (депопуляция), но и кардинальную перестройку его возрастной структуры — так называемое старение населения (увеличение доли людей старших возрастных когорт). Обычно такого рода процессы рассматриваются в связи с междустрановыми сопоставлениями, однако они могут наблюдаться и внутри одной страны, если отдельные группы её населения придерживаются разных репродуктивных стратегий. Различия в поведении людей, придерживающихся стратегий, присущих традиционному или, наоборот, современному типу воспроизводства проявляются в выборе времени вступления в брак, а также в среднем числе рождаемых каждой женщиной детей.

Особенности этих процессов в нашей стране хорошо представлены, например, в сборнике, посвящённом результатам уникального социально-демографического обследования “Родители и дети, мужчины и женщины в семье и обществе” [7]. Оно проводилось в рамках международной программы “Поколения и гендер”, объединившей исследователей-демографов из более чем 30 стран мира. Репродуктивные стратегии женщин современной России изучались также О.М. Шубат и А.П. Багировой [8].

Использование результатов названных работ позволяет перенести фокус моделирования демографических процессов на микроуровень и применить агент-ориентированный подход. Кроме того, эти результаты позволяют создать достаточное число интеллектуальных, интенциональных агентов [9], то есть агентов, наделённых собственными механизмами мотивации. Для подоб-

ных агентов моделируются внутренние убеждения, желания, намерения и мотивы, порождающие цели, которые и определяют их действия. В наших моделях роль “убеждений” будут выполнять желаемое максимальное число детей, а также распределение рождения этих детей в течение репродуктивного периода.

Для моделирования таких демографических процессов, как смертность и рождаемость (миграция в модели не учитывается), и расчёта численности населения на перспективу мы воспользовались методом передвижки возрастов [10]. Этот метод опирается на представление, что выжившие в соответствии с половозрастными коэффициентами выживаемости индивиды ежегодно становятся на год старше и участвуют в репродуктивном процессе (в соответствии с данными о рождаемости для женщин фертильного возраста). На основе этого подхода и с использованием реальных статистических данных в Центральном экономико-математическом институте РАН были разработаны две мультиагентные демографические модели, различающиеся уровнем детализации при имитации репродуктивного поведения людей. Причём направление усложнения поведения агентов было выбрано в том числе с учётом значимости вносимых изменений для выбора алгоритмов распараллеливания моделей при их переносе на суперкомпьютер. С этой точки зрения принципиальным отличием является то, что в первой модели агенты не образуют супружеские пары (нет взаимодействия с другими агентами) — агенты-женщины репродуктивного возраста самостоятельно определяют число желаемых детей и время их рождения (создания новых агентов). Во второй модели агенты-люди между собой взаимодействуют, образуя пары и согласовывая рождение общих детей.

Модели представляют собой программные продукты, реализованные в среде AnyLogic\* и обладающие интерфейсом для демонстрации пользователю происходящих во времени изменений основных демографических характеристик популяции агентов условного региона. Общая схема работы моделей показана на рисунке 1. Охарактеризуем конструкцию первой модели, а затем покажем изменения отдельных элементов конструкции во второй модели.

**Конструкция первой модели** предполагает разделение агентов на две группы, различающиеся репродуктивными стратегиями — с высокой и низкой рождаемостью соответственно.

Компонентами агентов-людей являются:

- тип репродуктивного поведения;
- пол;

- возраст;
- максимальное (желаемое) число детей;
- фактическое число детей;
- родственная связь с агентом-матерью;
- процедура включения в коллекцию женщин, желающих родить ребёнка.

К компонентам среды относятся:

- текущий год;
- общая численность популяции агентов;
- доля агентов традиционного типа;
- половозрастная структура популяции агентов;
- доля агентов традиционного типа в каждой возрастной когорте;
- коэффициенты смертности, дифференцированные по полу и возрасту;
- суммарные коэффициенты рождаемости для двух типов агентов;
- параметры вероятностных распределений, с помощью которых случайным образом определяется число детей для агентов-представителей каждого типа;
- распределение рождений по возрасту матери для двух типов агентов;
- коллекции агентов-женщин, желающих родить ребёнка (отдельно для каждого типа агентов и для каждой возрастной когорты, соответствующей репродуктивному периоду женщин — от 15 до 49 лет);
- процедура имитации смертности;
- процедура имитации рождаемости;
- процедуры связи с внешними файлами (чтение исходных данных и запись результирующих показателей).

При имитационном моделировании большое значение имеет не только правильная организация процессов, происходящих в течение модельного времени, но и точное воссоздание в начале работы модели стартового состояния системы (блоки A2–A5 на рисунке 1). Для этого из базы данных (таблиц формата Excel) считывается массив исходной информации. Для воссоздания структуры популяции агентов используется исходная информация, представленная в виде двух стандартных возрастно-половых пирамид, соответствующих двум типам воспроизводства (для популяции агентов традиционного типа характерна относительно большая численность новорождённых и детей, поэтому пирамида имеет более широкое основание).

Параметрами модели служат общая численность популяции агентов, а также доля в этой численности агентов традиционного типа. В соответствии со значениями параметров и соотношениями, заданными возрастно-половыми пирамидами, создаётся множество агентов и распределяются такие их характеристики, как тип репродуктивного поведения, пол и возраст. То есть каждому агенту присваиваются такие значения возраста, пола и типа, чтобы структура создаваемой популяции точно воспроизводила рассчитан-

\* AnyLogic — инструмент имитационного моделирования, который поддерживает все подходы к созданию имитационных моделей: процессно-ориентированный (дискретно-событийный), системно-динамический и агентный, а также любую их комбинацию. Более подробно см.: <http://www.anylogic.ru>.

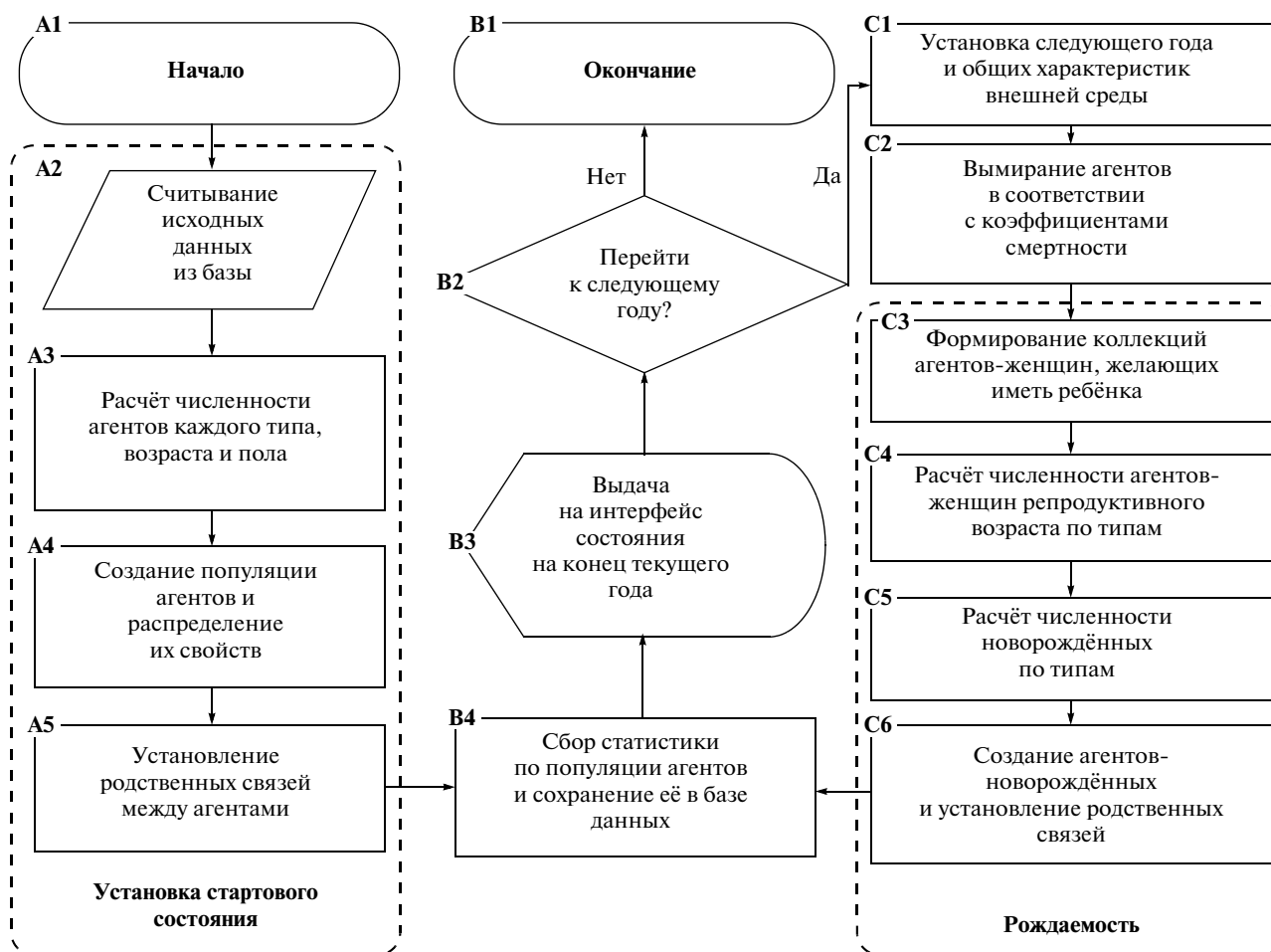


Рис. 1. Общая схема работы демографической модели

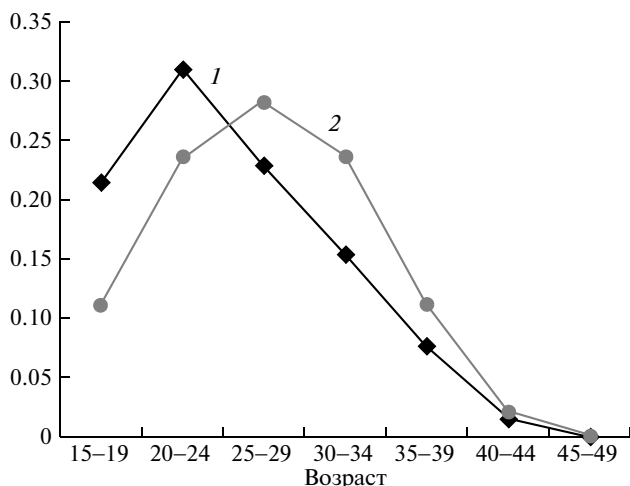
ную на основе исходных данных. Каждому агенту-женщине присваивается выбранное случайным образом с помощью  $\beta$ -распределения желаемое число детей в интервале от минимального до максимального для её типа. Затем устанавливаются связи мать–ребёнок между созданными агентами, для чего используются данные о распределении рождений по возрасту матери. Фактически для каждого агента моложе 20 лет случайным образом с помощью  $\beta$ -распределения определяется возраст его матери, а затем из когорты полученного возраста выбирается агент-женщина того же типа, желаемое число детей которой больше числа рождённых. Другими словами, предполагается, что тип ребёнка совпадает с типом матери. Агент-ребёнок и агент-мать запоминают друг друга, а кроме того, к числу детей агента-женщины добавляется единица. Параметры используемых здесь  $\beta$ -распределений подобраны таким образом, чтобы получаемые частоты выпадения возрастов соответствовали эмпирическим данным о распределении рождений по возрасту матери для двух типов агентов, представленных на рисунке 2 (где доли общей численности новорождённых

для матерей из разных возрастных групп по каждому типу в сумме составляют единицу).

Анализ текущего состояния (блоки В2–В4) предполагает сбор статистики по популяции агентов – общая численность агентов, численность агентов-мужчин и агентов-женщин по возрастным группам и типам. Собранный статистика сохраняется в таблицах выходного файла формата Excel для последующей обработки и анализа.

Интерфейс модели демонстрирует пользователю “фактическое” состояние популяции в текущем году, а также графики, показывающие изменение отслеживаемых результатов работы модели в течение модельного времени. После отображения на экране текущего состояния программа ждёт решения пользователя о прекращении или продолжении симуляции. В первом случае происходит остановка работы модели, во втором – переход к следующему шагу её работы. На каждом последующем шаге (блоки С1–С6) выполняются следующие действия:

- к значению переменной года прибавляется единица;



**Рис. 2.** Доли новорожденных для матерей из разных возрастных групп в общей численности новорожденных

1 — традиционный тип, 2 — современный тип

- имитация смертности (блок С2) — на основе дифференцированных по полу и возрасту коэффициентов смертности рассчитывается вероятность умереть для каждого агента популяции, после чего вероятностным образом определяется его судьба, часть агентов уничтожается, а оставшиеся становятся на год старше (коэффициенты смертности для агентов двух типов приняты одинаковыми и постоянными на протяжении всего периода симуляции);

- имитация рождаемости (блоки С3–С6).

На этом этапе агенты-женщины репродуктивного возраста в соответствии со своим возрастом и типом репродуктивной стратегии пополняют коллекцию желающих родить, если их желаемое число детей больше фактического. Именно этот выбор и представляет собой действие, доступное агентам в модели. Затем для каждого типа отдельно рассчитывается общее число агентов-женщин репродуктивного возраста и общее число младенцев, которых они должны родить в текущем году, соответствующее заданным суммарным коэффициентам рождаемости. Для каждого типа создается рассчитанное количество агентов нулевого возраста, а пол каждого новорожденного агента определяется случайным образом (женский — с вероятностью 0.488 [11]). Наконец, каждому новорожденному агенту подбирается мать из коллекций агентов-женщин репродуктивного возраста, желающих в этом году родить ребенка. Если при этом число рожденных агентом-матерью детей сравнивается с числом желаемых, она удаляется из коллекции желающих родить и больше не будет участвовать в репродуктивном процессе.

**Результаты апробации первой модели.** С использованием модели были проведены эксперименты по прогнозированию изменения численности популяции агентов, населяющих условный

регион, возрастной структуры этой популяции, а также соотношения численности агентов разных типов для основных возрастных групп и для популяции в целом. Были проведены 20 экспериментов при следующих значениях параметров:

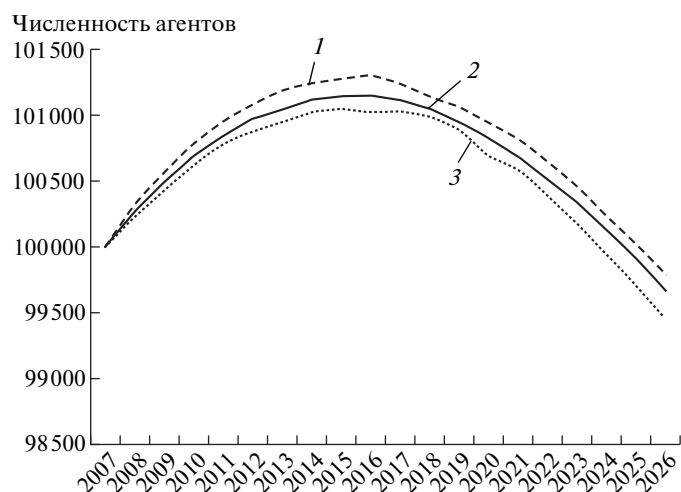
- общая численность популяции агентов — 100000;
- доля агентов традиционного типа — 10%;
- в качестве показателей смертности для агентов двух типов выбраны общероссийские значения;
- суммарный коэффициент рождаемости для агентов-женщин традиционного типа — 3, для современного — 1.4;
- минимальное число желаемых детей для агентов традиционного типа — 2, максимальное — 10, для агентов современного типа — 0 и 3 соответственно;
- базовый год — 2007;
- период симуляции — 20 шагов (лет).

В качестве прогнозных значений характеристик популяции были взяты их соответствующие средние по всем экспериментам, при этом модель продемонстрировала высокую устойчивость. Так, для общей численности популяции отклонения от среднего значения на протяжении 20 лет располагались в интервале от  $-0.23\%$  до  $0.16\%$ , что можно считать очень хорошим результатом.

Анализ результатов экспериментов позволяет утверждать, что модель адекватно имитирует такие наблюдаемые в реальной жизни процессы, как снижение общей численности населения (депопуляция), а также его старение. На рисунке 3 показаны снижение численности агентов начиная с 2016 г. и максимальные и минимальные значения численности агентов по годам. Старение населения отражено на рисунке 4: численность агентов моложе 50 лет неуклонно снижается на протяжении всего периода, а численность агентов старше 80 лет возрастает.

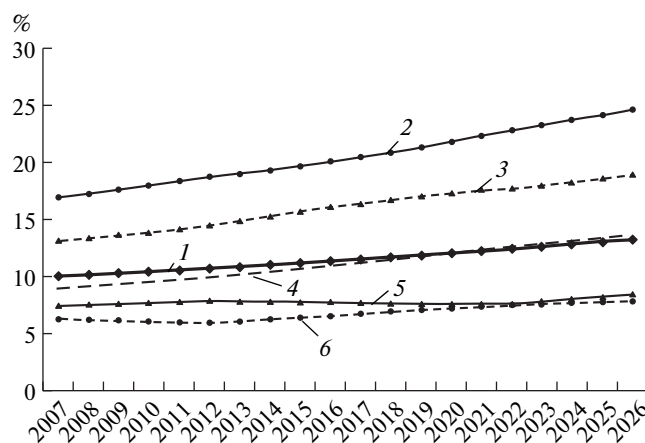
Помимо этого, модель демонстрирует ещё один эмпирически наблюдаемый эффект, а именно, что соотношение численности агентов двух типов будет неодинаковым в разных возрастных группах (рис. 5). При общем незначительном росте доли агентов традиционного типа на каждом шаге моделирования, она тем выше, чем моложе возрастная группа. Более того, чем моложе агенты, тем больше скорость увеличения этой доли. Так, если в среднем по популяции доля агентов традиционного типа за 20 лет, для которых строился прогноз, выросла на  $3.2\%$ , то в группе самых младших агентов (от 0 до 14 лет) эта доля выросла на  $7.7\%$ , а в следующей возрастной группе агентов (от 15 до 24 лет) — уже только на  $5.8\%$ .

**Отличия второй модели.** Исходные данные для второй модели те же, что и для первой. Общая последовательность имитируемых демографических процессов также совпадает со схемой первой модели. С содержательной точки зрения отличие заключается в более детализированном и реали-



**Рис. 3.** Динамика численности популяции агентов  
1 – максимальные значения, 2 – средние значения, 3 – минимальные значения

стичном описании процессов воспроизводства населения. Так, рождению детей предшествует образование семейных пар, которые “обсуждают” число желаемых детей, при этом процедура поиска партнёра полностью контролируется самим агентом и зависит от его принадлежности к тому или иному типу репродуктивного поведения. “Обсуждение” в модели реализуется в виде обмена сообщениями между агентами, ищущими себе партнёра. Взаимодействие агентов при поиске партнёра является именно тем отличием новой модели, которое существенным образом влияет на организацию распараллеливания её работы для запуска на суперкомпьютере.



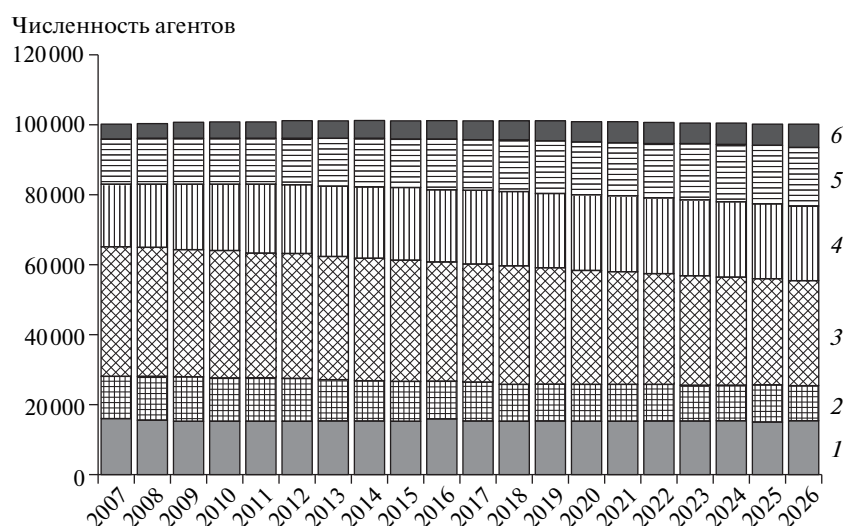
**Рис. 5.** Доли агентов традиционного типа по популяции агентов в целом (1) и по основным возрастным группам

Возрастные группы: 2 – 0–14 лет, 3 – 15–24 года, 4 – 25–49 лет, 5 – 50–64 года, 6 – 65–79 лет

В соответствии с обозначенной спецификой появляются новые компоненты агентов (место жительства, все родственные связи, процедура поиска партнёра, процедура рождения ребёнка) и новые компоненты среды (коллекции одиноких агентов – мужчин и женщин, желающих создать семью, и модифицированная процедура имитации рождаемости).

Установка стартового состояния системы включает следующие шаги:

- вся территория, на которой расселяются агенты в рабочем окне модели, делится на кварталы (квадраты), и каждому агенту присваивается его “адрес” – номер квадрата;



**Рис. 4.** Динамика возрастной структуры популяции агентов

Возрастные группы: 1 – 0–14 лет, 2 – 15–24 года, 3 – 25–49 лет, 4 – 50–64 года, 5 – 65–79 лет, 6 – 80 лет и старше

- желаемое число детей присваивается всем агентам, независимо от пола;
- агенты-дети наделяются обоими родителями;
- устанавливаются все родственные связи агентов — партнёр, родители, дети, братья и сёстры;
- для каждой возрастной когорты в интервале от 15 до 48 лет для женщин и от 18 до 65 лет для мужчин формируются коллекции (списки) одиноких мужчин и одиноких женщин (отдельно), желающих создать семью;
- агенты из коллекций одиноких женщин пытаются найти одинокого мужчину для образования пары;
- агенты-женщины репродуктивного возраста, имеющие партнёра и ещё не родившие столько детей, сколько они планировали, пополняют коллекции женщин, желающих родить ребенка.

Модификации подверглась и процедура имитации рождаемости:

- рассчитывается общее количество агентов-женщин репродуктивного возраста каждого типа, а также их количество для пятилетних возрастных когорт;
- рассчитывается общее количество агентов-новорождённых каждого типа с учётом соответствующих общих коэффициентов рождаемости;
- рассчитывается количество агентов-новорождённых, приходящихся на группы агентов-женщин из пятилетних возрастных когорт (с учётом заданного распределения новорождённых по возрасту матерей каждого типа);
- для каждой возрастной когорты и каждого типа рассчитывается вероятность рождения ребёнка агентом из коллекций женщин, желающих родить;
- для каждого агента из указанных коллекций вероятностным образом решается его судьба, и если ответ положительный, то выполняется внутренняя процедура рождения ребёнка этим агентом.

Алгоритм действий агента в рамках процедуры поиска партнёра включает:

- для каждой одинокой женщины-агента случайным образом определяется возраст будущего претендента (максимальная разница в возрасте — 20 лет);
- в коллекции одиноких агентов-мужчин данного возраста и того же типа, что и агент-женщина, находится такой, который проживает в том же квартале и не является её родственником;
- с этим агентом-претендентом агент-женщина обменивается сообщениями о желаемом числе детей;
- если разница между этими числами не превышает двух, то образуется пара — возникают партнёрские связи, а также родственные связи с уже имевшимися у супругов детьми (если они были). Кроме того, оба партнёра удаляются из коллекций одиноких агентов, а агент-женщина добавляется в соответствующую коллекцию жен-

щин, желающих родить (если она ещё не имеет столько детей, сколько хотела).

С использованием второй модели были реализованы те же эксперименты, что и с использованием первой, и получены аналогичные результаты. Таким образом, можно заключить, что разработанные агент-ориентированные модели, несмотря на очевидное упрощение действительности, верно воспроизводят как начальное состояние населения условного региона, включая его половозрастную и социальную структуру, так и динамику основных демографических характеристик.

## АДАПТАЦИЯ МОДЕЛЕЙ ДЛЯ ЗАПУСКА НА СУПЕРКОМПЬЮТЕРЕ

Для запуска описанных моделей на суперкомпьютерах использовалась разработанная авторами технология поддержки AOM для суперкомпьютеров STARS (Supercomputer Technology for Agent-oriented Simulation). Важно отметить, что кроме общей схемы распараллеливания, которая определяет способ преобразования исходного кода модели, обязательным является соответствие базовым требованиям современных суперкомпьютеров, для которых де-факто стандартом выступает ориентация на MPI (программный интерфейс для обмена сообщениями между процессами, выполняющими задачу в параллельном режиме). При этом, к сожалению, нельзя рассчитывать ни на общую память узлов суперкомпьютера, ни на поддержку стандарта OpenMP, в чём авторы убедились ещё несколько лет назад. Желаемого результата также не принесёт ни установка на суперкомпьютер среды моделирования AnyLogic, ни прочие возможности типа библиотек для работы с файлами в формате MS Excel. Нежелательно использование стандартной среды исполнения Java, поскольку она редко присутствует на суперкомпьютере. Таким образом, вся специфика, свойственная исходной среде исполнения AnyLogic и стандартному Java-окружению, должна быть адаптирована к суперкомпьютеру, а взаимодействия между вычислительными узлами основываться на стандарте MPI.

**Имеющийся опыт и анализ новых программных технологий.** Уже опробованная авторами при распараллеливании демографической модели России в 2011 г. [1] библиотека для многоагентного моделирования ADEVS показала себя достаточно хорошо. Последние версии ADEVS, кроме того, в определённом объёме поддерживают Java, что тоже является плюсом. Однако разработчики ADEVS до сих пор не реализовали параллельную работу на суперкомпьютере, за исключением технологии OpenMP для мультипроцессоров, в связи с чем в предыдущей нашей работе потребовалось значительное количество доработок в части поддержки MPI.

При распараллеливании предыдущей, достаточно простой модели она была переписана на

C++ целиком, что можно считать избыточным: пред- и постобработка данных, а также создание начального состояния многоагентной среды не относятся к критическим по времени операциям. Для суперкомпьютера обычно достаточно распараллелить лишь вычислительное ядро алгоритма, то есть в данном случае фазу пересчёта состояния популяции. В текущей реализации модели фаза пересчёта не использует планировщик AnyLogic, так как требуется обеспечить строго заданную очередность обработки состояния агентов.

Анализ новейших программных технологий показал, что в последнее время активно развиваются встраиваемые средства для исполнения Java-программ, которые используют так называемую AOT (Ahead-of-Time) компиляцию. При этом результатом работы AOT-компилятора является обычный автономный исполняемый модуль, содержащий машинный код для целевой платформы. Интересно отметить, что подобный подход используется в новых версиях операционной системы Android, что, на наш взгляд, не случайно: и для встраиваемых систем, и для суперкомпьютера эффективность исполнения кода становится основным фактором. Эксперименты с одним из подобных продуктов — AOT-компилятором Avian — позволили заключить, что, во-первых, он позволяет получить автономный исполняемый модуль в виде MPI-приложения для суперкомпьютера, при этом на C++ легко реализуется произвольный дополнительный код, включая инициализацию и привязку к коммуникационной библиотеке MPI; во-вторых, скорость работы полученного программного модуля (к примеру, на классической игре Д. Конвея “Жизнь” [12]) приблизительно соответствует скорости работы ADEVS. Это позволило переложить значительную часть работы на AOT-компилятор и реализовать на C++ только самое необходимое, закрепив за ADEVS функцию поддержки ускоренных стадий со сложным межагентным взаимодействием.

**Технические решения для распараллеливания модели.** В качестве программного средства для организации межузлового взаимодействия выбрана парадигма активных сообщений (мобильных сериализованных объектов). Она широко и с успехом применяется в этом качестве, поскольку обладает неоспоримым преимуществом — избавляет программиста от необходимости анализировать поступающие сообщения. Достигается это благодаря тому, что сообщения пересылаются в виде сериализованных объектов, при приёме и восстановлении объектного вида которых просто вызывается их метод *handle*. То есть о том, что должно происходить при приёме сообщения, заботится исключительно формирующая и посылающая его сторона.

На базе активных сообщений для встраиваемой среды исполнения Avian+MPI был реализован программный каркас, в котором выделенный (первый по счёту) узел кластера выдаёт задания

(так называемые стадии расчёта) и собирает результаты со всех участвующих в расчёте узлов (включая его самого). Одному шагу моделирования (году, в терминологии текущей модели) может соответствовать несколько стадий, так как при межузловом межагентном взаимодействии есть вероятность возникновения потребности в согласовании определённых свойств у агентов (например, согласование желаемого количества детей партнёрами).

Для формата ALP, используемого средой AnyLogic для хранения описания моделей, был разработан конвертер, в котором для каждой секции входного файла реализован обработчик, собирающий существенную информацию о модели и генерирующий Java-классы. Все эти обработчики используются конвертером с опорой на программную специфику модели (то есть набор правил трансляции). Например, инициализация поля главного класса модели *Population* обрабатывается так, чтобы получить количество агентов из окружения. Это позволяет задавать значение моделируемого параметра в виде опций командной строки.

Эмуляция необходимых функций AnyLogic (взаимодействия, генераторы случайных чисел, использующие вероятностные распределения, чтение и запись excel-файлов) была реализована в виде отдельного программного пакета RT (RunTime). C++ код для работы с MPI и низкоуровневыми примитивами размещается в отдельном программном модуле, производит инициализацию и вызывает входную точку модели, предоставляя некоторые примитивы, а также полезные средства для отладки.

**Технология получения и запуска параллельной версии.** Исходное, разработанное в среде AnyLogic описание модели представляет собой XML-файл с расширением. ALP (XML-файл в стандарте AnyLogic), в котором содержится описание агентных сущностей модели, их параметры, а также правила пересчёта состояния агентов во время эволюции. Кроме ALP-описания к модели прилагается файл с данными в формате Excel, в котором указаны численные параметры, используемые как на стадии формирования стартового состояния системы, так и при пересчёте соответствующих популяций.

Процесс создания распараллеленного суперкомпьютерного варианта программы включает три стадии. На первой входной ALP-файл с описанием модели считывается конвертором, который реконструирует объектное представление для всех описанных сущностей и правил и, проведя необходимую их обработку, формирует набор программных модулей (Java-классов общим объёмом в несколько тысяч строк), содержащих всю существенную информацию. Переменные агентов преобразуются в поля классов, а правила реакции агентов на события — в соответствующие методы.



На второй стадии входной Excel-файл со стартовыми значениями параметров модели преобразуется в исходный текст на языке Java (также объёмом в несколько тысяч строк), чтобы обеспечить оперативный доступ к параметрам модели на каждом узле суперкомпьютера. Иными словами, внешний набор параметров модели становится составной частью исполняемого модуля для суперкомпьютера.

Последняя стадия предполагает компиляцию сформированных программных модулей совместно с разработанным кодом эмуляции используемых функций среды AnyLogic и их компоновку в машинный код для целевого суперкомпьютера. Исполняемый модуль является полностью автономным и при запуске принимает лишь несколько ключевых параметров в командной строке. Например, запуск модели с миллионом агентов для 20 лет на 12 вычислительных узлах производится следующей командой:

```
$ mpirun -np 12 -x Population = 1000000 rt.Main
```

Выходные данные при этом собираются с вычислительных узлов в процессе счёта, ключевые характеристики популяции и общее потраченное на расчёт время сохраняются в журнале, а также печатаются по окончании запуска:

```
[0]: totalNumberPeople = 990221
```

```
[0]: * * * Total stages time: 21.173s * * *
```

**Сравнительные результаты скорости работы моделей и перспективы развития разработанной технологии.** Для первых экспериментов с параллельной версией упрощённой демографической модели использовался фрагмент лабораторного вычислительного кластера с 12 счётными ядрами и общим объёмом оперативной памяти 96 Гбайт. При такой конфигурации в оперативной памяти без труда помещаются 60 млн. агентов, что позволяет провести моделирование динамики роста населения в масштабах небольшой страны.

Параллельная версия также испытывалась на многоядерном процессоре. Результаты замеров времени работы оригинальной и параллельной версии модели представлены в таблицах 1, 2 и 3, из которых видно, что скорость работы и вместимость по числу агентов суперкомпьютерной версии значительно выше соответствующих показателей оригинальной модели. Также видно, что при повышении локальности межагентных взаимодействий (увеличения числа кварталов и соответствующего уменьшения их площади) производительность счёта повышается. Таким образом, для моделирования динамики развития населения России или Европы достаточно небольшого Blade-сервера, при этом время одного расчёта не превысит нескольких минут. Для проведения более масштабных экспериментов с использованием разработанной модели потребуются уже более вместительный и производительный суперкомпьютер.

Главным положительным моментом предлагаемого нами подхода к распараллеливанию моделей, разработанных в среде AnyLogic, является автоматизация создания их суперкомпьютерных версий. Это значительно упрощает разработку, так как в большинстве случаев после несущественной модификации исходной модели не требуется доработки правил преобразования в исполняемый модуль для суперкомпьютера.

Данный подход является расширяемым в части используемого целевого языка и программно-аппаратной платформы. Кроме уже успешно опробованных платформ исполнения Avain и ADEVS можно разработать и более низкоуровневые средства для ускорения фазы пересчёта состояния агентов, а в перспективе рассмотреть вопрос об использовании таких аппаратных ускорите-

**Таблица 1.** Оригинальная версия, 1 вычислительное ядро (8 гигабайт оперативной памяти)

Агенты	Секунды
50000	1.8
100000	6
200000	28
400000	92
800000	360
1600000	1485

**Таблица 2.** Параллельная версия, 4 вычислительных ядра

Ядра	Кварталы	Агенты	Секунды
1	12	1000000	210
4	12	1000000	50
4	40	1000000	30
4	40	2000000	75
4	20	4000000	344

**Таблица 3.** Фрагмент вычислительного кластера (12 счётных ядер Core i7, 96 гигабайт оперативной памяти)

Ядра	Кварталы	Агенты	Секунды
12	12	1000000	21
12	24	1000000	12
12	24	10000000	516
12	60	10000000	303
12	300	10000000	132
12	300	30000000	585
12	300	50000000	1371
12	300	60000000	1833

лей, как Xeon Phi и NVidia CUDA. Использование более эффективных низкоуровневых средств предпочтительнее также в целях ускорения работы некоторых примитивов. Как выяснилось в процессе разработки, на итоговую скорость работы модели весьма существенное влияние оказывает эффективность используемых генераторов случайных чисел. Поэтому их оптимизация является приоритетной задачей. Перспективным подходом здесь может быть генерация случайных чисел с нужными параметрами используемых вероятностных распределений в отдельных счётных потоках.

Применённая технология межузловое взаимодействия при помощи технологии активных сообщений позволяет в случае необходимости легко реализовать как интерактивное моделирование, так и интерактивную визуализацию процесса моделирования в масштабе расчётного времени. Однако это возможно лишь в случае, если суперкомпьютер будет доступен в монопольном режиме, например, может использоваться компактный персональный суперкомпьютер.

Основным вопросом, который сохраняет актуальность, является вопрос о максимально достижимой эффективности распараллеливания в случае массового взаимодействия агентов, находящихся на разных вычислительных узлах суперкомпьютера. Достаточно очевидно, что если каждый агент активно взаимодействует со всеми остальными агентами, производительность будет низкой ввиду объёмного межузлового трафика. Тем не менее даже при таких неблагоприятных условиях суперкомпьютер позволяет существенно ускорить моделирование, когда модель запускается или многократно (для набора статистики), или с разными параметрами. В другом крайнем случае, когда почти все взаимодействия локализованы по географическому расположению агентов, эффективность распараллеливания будет хорошей. Таким образом, эффективность параллельной версии напрямую зависит от той доли межагентных взаимодействий, которая потребует пересылки больших объёмов данных между вычислительными узлами.

\* \* \*

В заключение подчеркнём несколько моментов, связанных с использованием разработанных и распараллеленных мультиагентных моделей, имитирующих явление демографического перехода.

Во-первых, хотя обычно различия в репродуктивном поведении людей рассматриваются в связи с междустрановыми сопоставлениями, они могут наблюдаться и внутри одной страны, если её население в этом отношении неоднородно.

Во-вторых, выбранный подход (перенос фокуса моделирования демографических процессов на микроуровень) позволяет адекватно воспроиз-

вести в модели особенности репродуктивного поведения людей с учётом их индивидуальных различий.

В-третьих, результаты проведённых на модели экспериментов соответствуют наблюдаемым в реальной жизни процессам, таким как снижение общей численности и старение населения.

В-четвёртых, за счёт имитации различий в репродуктивном поведении людей из разных социальных групп, а именно, различий в выборе времени вступления в брак, а также в среднем числе рождаемых каждой женщиной детей, модели адекватно демонстрируют изменения в социальной структуре населения.

Ввиду этих особенностей представленные модели могут быть использованы при изучении динамики численности и структуры населения тех или иных регионов.

Работа выполнена при поддержке Российского научного фонда, проект № 14-18-01968.

## ЛИТЕРАТУРА

1. Макаров В.Л., Бахтизин А.Р., Васенин В.А. и др. Средства суперкомпьютерных систем для работы с агент-ориентированными моделями // Программная инженерия. 2011. № 3. С. 2–14.
2. Agent-Based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behavior / Ed. by F.C. Billari, A. Prskawetz. Heidelberg: Springer, 2003.
3. Diaz B.A. Agent-Based Models of Social Interaction and Demographic Behaviour (Ph.D. Thesis). Wien: Technische Universität, 2010.
4. Population Dynamics and Projection Methods: Understanding Population Trends and Processes / Ed. by J. Stillwell, M. Clarke. N. Y.: Springer, 2011.
5. Wu B.M., Birkin M.H. Agent-Based Extensions to a Spatial Microsimulation Model of Demographic Change // Agent-Based Models of Geographical Systems / Ed. by A.J. Heppenstall et al. N. Y.: Springer, 2012. P. 347–360.
6. Вишневский А.Г. Воспроизводство населения и общество: История, современность, взгляд в будущее. М.: Финансы и статистика, 1982.
7. Родители и дети, мужчины и женщины в семье и обществе / Под науч. ред. Т.М. Малевой, О.В. Синявской. М.: НИСП, 2007.
8. Шубат О.М., Багирова А.П. Прогнозирование вторых рождений у российских женщин: социологостатистический подход // Проблемы прогнозирования. 2014. № 3. С. 131–140.
9. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М.: Едиториал УРСС, 2002.
10. Бахметова Г.Ш. Методы демографического прогнозирования. М.: Финансы и статистика, 1982.
11. Шахотко Л.П., Терещенко С.М. Компьютерное решение задачи построения демографических прогнозов // Вопросы статистики. 1999. № 10. С. 57–65.
12. Тоффли Т., Марголис Н. Машины клеточных автоматов. М.: Мир, 1991.