На правах рукописи

Попков Сергей Игоревич

# МОДЕЛИРОВАНИЕ ПОВЕДЕНИЯ ВЕРОЯТНОСТНЫХ МНОГОАГЕНТНЫХ СИСТЕМ С ДЕЦЕНТРАЛИЗОВАННОЙ АРХИТЕКТУРОЙ

Специальность 05.13.18
Математическое моделирование, численные методы и комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва, 2020

Работа выполнена на кафедре прикладной информатики и мультимедийных технологий факультета информационных технологий Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный психолого-педагогический университет»

**Научный руководитель:** доктор технических наук, профессор
**Куравский Лев Семенович**

**Официальные оппоненты:** **Чечкин Александр Витальевич**, доктор физико-математических наук, профессор
ФГКВОУ ВО «Военная академия Ракетных войск стратегического назначения имени Петра Великого» Министерства обороны РФ,
ФГОБУ ВО «Финансовый университет при Правительстве РФ»,
лауреат Государственной премии СССР,
Заслуженный работник высшей школы, профессор

**Колбин Илья Сергеевич**, кандидат физико-математических наук, научный сотрудник ФИЦ «Информатика и управление» РАН

**Ведущая организация:** ФГУ «Федеральный исследовательский центр «Институт прикладной математики им. М. В. Келдыша» Российской академии наук»

Защита состоится «24» апреля 2020 года в 12 ч. 00 мин. на заседании диссертационного совета Д 212.125.04 Московского авиационного института по адресу: 125993, Москва, А-80, ГСП-3, Волоколамское шоссе, 4.

С диссертацией можно ознакомиться в библиотеке Московского авиационного института по адресу: 125993, Москва, А-80, ГСП-3, Волоколамское шоссе, 4 или на сайте МАИ по ссылке: https://mai.ru/events/defence/index.php?ELEMENT_ID=110809

Автореферат разослан: “__” ____________ 20__ г.

Отзывы в 2-х экземплярах, заверенные печатью, просим отправлять по адресу: 125993, Москва, А-80, ГСП-3, Волоколамское шоссе, 4, Ученый совет МАИ

Ученый секретарь диссертационного совета Д 212.125.04, кандидат физико-математических наук — Рассказова Варвара Андреевна

# Общая характеристика работы

Проблемы исследования коллективного поведения и организации группового управления привлекают внимание исследователей с 30-х годов 20-го века. Их значимость возросла в последние годы в связи с актуальностью задач управления коллективом роботов, включая группы беспилотных летательных аппаратов и других мобильных систем. В настоящее время исследования и разработки в этой области ведутся во многих странах с привлечением большого количества специалистов. Под многоагентными системами далее понимаются системы, образованные совокупностью взаимодействующих интеллектуальных агентов. Класс многоагентных систем с децентрализованной архитектурой, в отличие от централизованной, определяет агентов в рамках системы как равноценных и способных, при необходимости, действовать автономно согласно определенной цели.

Концепции, базовые понятия, принципы построения, перспективы практического применения и другие аспекты создания многоагентных систем рассмотрены в работах Вяткина А.Ю., Смирнова Д.В., Кочетова И.А., Городецкого В.И., Карсаева О.В., Самойлова В.В., Пантелеева А.В., Скавинской Д.В., Серебрякова С.В., Городецкого В.И., Бухвалова О.Л., Скобелева П.О. Однако математический аппарат и программные средства, приемлемые для управления поведением составляющих систему агентов в реальном времени и оперативной оценки ресурсов, необходимых для решения прикладной задачи, к настоящему времени в полном объёме ещё не созданы. Отсутствует пригодное к практическому применению математическое и программное обеспечение оценки уровня подготовки и обучения операторов многоагентных систем на специализированных тренажёрах.

Описанные выше практические проблемы моделирования и анализа коллективного поведения, а также применения группового управления позволяют говорить о необходимости:

- разработки новых методов математического моделирования поведения многоагентных систем, позволяющих эффективно прогнозировать поведение этих систем на основе вероятностных оценок в реальном времени;
- разработки численных методов, учитывающих требования к современным шаблонам проектирования программных продуктов и позволяющих оптимизировать вычислительные процедуры, прогнозирующие поведение многоагентных систем;
- разработки кроссплатформенного комплекса программ для имитационного моделирования, обеспечивающего оценку уровня подготовки оператора сложной многоагентной системы и реализующего принципы адаптивного обучения на специализированных тренажёрах.

**Актуальность темы диссертации** обусловлена необходимостью создания новых подходов к решению задачи группового управления многоагентными системами, обеспечивающих прогнозирование моделируемой ситуации и принятие решений на основе количественных критериев, а также оценку уровня подготовки и обучение операторов, работающих с этими системами.

**Цель работы:** разработка математической модели управления поведением многоагентных систем и реализация на её основе комплекса программ для прогнозирования такого поведения и оценки уровня подготовки и обучения операторов специализированных тренажёров.

Для этого решены **задачи**:

- создания математической модели для исследования поведения частного класса многоагентных систем;
- создания математической модели и метода прогнозирования, обеспечивающих оценку ресурсов, необходимых для решения задачи, на основе количественных критериев;
- проведения вычислительных экспериментов, необходимых для оценки параметров прогнозирования;
- идентификации параметров модели исследуемой прикладной многоагентной системы;
- создания на основе разработанных моделей и методов комплексов программ для прогнозирования поведения исследуемой прикладной многоагентной системы и оценки уровня подготовки и адаптивного обучения её операторов.

**Методологические основы и методы:** для решения поставленных задач использовались модели теории случайных процессов, методы анализа данных и оптимизации, а также численные методы.

**На защиту выносятся следующие научные результаты:**

- математическая модель и алгоритм поведения прикладной многоагентной системы исследуемого класса в реальном времени;
- математическая модель и метод прогнозирования, обеспечивающие оперативную оценку ресурсов, необходимых для решения задачи, на основе количественных критериев;
- адаптивный численный метод идентификации параметров модели прикладной многоагентной системы;
- комплексы программ для прогнозирования поведения прикладной многоагентной системы и оценки уровня подготовки и адаптивного обучения её операторов.

**Научная новизна заключается:**

- в математической модели и алгоритме поведения прикладной многоагентной системы;
- в математической модели и методе прогнозирования, обеспечивающих оперативную оценку ресурсов, необходимых для решения задачи;
- в адаптивном численном методе идентификации параметров прикладной многоагентной системы;
- в концепциях, лежащих в основе разработанных комплексов программ.

**Практическая значимость** диссертационной работы заключается в возможности:

- создания на основе разработанных математических моделей и алгоритмов прикладных многоагентных систем с полностью или частично автоматизированным групповым управлением в реальном времени;
- вычисления на основе разработанных методов прогнозирования лицом, принимающим решения, оценок ресурсов, необходимых для решения поставленной задачи;

— оценок уровня подготовки и адаптивного обучения на специализированных тренажёрах операторов многоагентных систем с использованием созданных комплексов программ.

**Достоверность результатов** работы подтверждается:

- оценками военных экспертов Всероссийского конкурса Министерства обороны РФ 2018 года по поиску в интересах Вооруженных Сил Российской Федерации научно-исследовательских работ граждан Российской Федерации,
- оценками экспертов Всероссийского межотраслевого молодёжного конкурса научно-технических работ и проектов «Молодёжь и будущее авиации и космонавтики»,
- сопоставлением результатов имитационного моделирования с эмпирическими данными,
- вычислительными экспериментами, подтвердившими эффективность разработанного численного метода идентификации параметров прикладной многоагентной системы.

**Апробация**. Работа стала победителем финального этапа Всероссийского межотраслевого молодёжного конкурса научно-технических работ и проектов «Молодёжь и будущее авиации и космонавтики» в 2018 году в номинации "Математические методы в аэрокосмической науке и технике" и заняла II место на Всероссийском конкурсе Министерства обороны РФ 2018 года по поиску в интересах Вооруженных Сил Российской Федерации научно-исследовательских работ граждан Российской Федерации. Теоретические и практические результаты работы были представлены на Всероссийских научных конференциях «Нейрокомпьютеры и их применение» в 2016-2019 годах (отмечены дипломами за лучший научный доклад), Всероссийской выставке научно-технического творчества молодежи «НТТМ-2015» (отмечены дипломом «НТТМ-2015»), а также на научных семинарах в Военной академии РВСН имени Петра Великого и Главном научно-исследовательском испытательном центре робототехники Министерства обороны РФ.

# Основное содержание диссертации

Диссертация содержит 129 страниц основного текста, состоящего из введения, шести глав, заключения и списка использованной литературы.

**Во введении** обоснована актуальность проблемы, сформулированы цель и задачи исследования, дана общая характеристика работы.

**В первой главе** проанализировано современное состояние проблемы создания многоагентных систем, рассмотрены современные подходы к построению соответствующих методов и моделей, включая деревья решений, обучение с подкреплением, генетические алгоритмы и марковские процессы, исследованы достоинства и недостатки применяемых подходов, обоснована необходимость разработки новых концепций построения моделей в данной предметной области, сделан вывод о том, что для эффективного практического применения создаваемых систем необходимо обеспечить:

- наличие целевой функции, учитывающей особенности применения модели;
- недетерминированность поведения агентов;

– быструю оценку ресурсов для решения поставленной задачи с заданным уровнем надежности;
– создание средств обучения (тренажеров) для операторов этих систем.

Особенности практического применения рассматриваемых систем требуют, чтобы каждый интеллектуальный агент обладал:

– умением выполнять задачи в сложном окружении без постоянной поддержки извне (автономностью);
– способностью улучшать качество выполняемой работы на основе приобретенного опыта (адаптивностью);
– способностью к организации деятельности в соответствии с алгоритмом функционирования;
– собственной целевой функцией.

При этом иерархия взаимодействия среди агентов должна отсутствовать - не допускается существование агентов, управляющих всей системой.

Под термином «интеллектуальный агент» (далее – «агент») понимается процесс, получающий информацию в виде данных о совокупности других управляемых процессов и способный влиять через управление этими процессами, способствуя достижению поставленной цели.

**Во второй главе** рассматривается моделирование работы многоагентной системы, которая разработана на основе требований, сформулированных в первой главе.

Агенты $L_k$ $(k = 0, \ldots, K-1)$ перемещаются по плоскому игровому полю, на котором находится одна *подвижная цель* $T$, согласно представленным далее правилам, стараясь поразить цель. Для определения их положения вводится разбиение области поверхности, прилегающей к цели, на ячейки, образованные пересечением $M$ концентрических колец и $N$ секторов, причём цель $T$ находится в центре внутреннего кольца. Относительная система координат игрового поля привязана к перемещаемой цели $T$. Положение агентов определяется с точностью до ячейки $(i, j)$, индексы которой задаются номером кольца $i$ $(i = 0, \ldots, M-1)$ и номером сектора $j$ $(j = 0, \ldots, N-1)$. Вероятность пребывания агента $L_k$ в ячейке $(i, j)$ в момент времени $t$ описывается функцией $p_{k,ij}(t)$.

В дискретные моменты времени, разделённые *интервалом дискретизации* $\Delta t$, агент $L_k$ может атаковать и, возможно, поразить цель $T$, а также быть атакованным и, возможно, поражённым этой целью с определёнными вероятностями, зависящими от его положения.

В дискретные моменты времени все работоспособные агенты получают информацию о том, в каких ячейках они находятся. В зависимости от игровой ситуации, в указанные моменты времени агенты могут получать или не получать информацию о наличии и положении других работоспособных агентов.

Изменение распределения и состава агентов на игровом поле, происшедшее при переходе от одного дискретного момента времени к другому моменту, следующему за ним по порядку с интервалом $\Delta t$, называется *тактом игры*.

В зависимости от возможностей агентов получать релевантную информацию, *условием завершения игры* является или получение информации о поражении цели, или поражение целью всех агентов. Первый исхода игры рассматривается как *победа агентов*, а второй – как *победа цели*.

Обозначив текущий момент времени как $t$, определим следующие этапы эволюции этой системы.

Этап 1. Задание начального пространственного распределения агентов при $t = 0$ (*этап инициализации*).

Этап 2. Получение агентами информации о состоянии и пространственном расположении друг друга в момент времени $t$ (*первый этап, реализующий взаимодействие агентов*).

Этап 3. Расчёт вероятностного распределения всех работоспособных агентов в пространстве в момент времени $t + \Delta t$, обеспечивающего экстремальное значение заданной целевой функции, зависящей от состояния и текущего пространственного распределения всех агентов, путём решения задачи оптимизации (*этап оптимизации и одновременно второй этап, реализующий взаимодействие агентов*).

Этап 4. Перемещение агентов в позиции, соответствующие их новому распределению в пространстве, вычисленному на этапе 3, в течение интервала времени $\Delta t$ (*этап перемещения*).

Этап 5. Получение работоспособными (то есть не пораженными целью и способными выполнять действия, определенные моделью) агентами информации о состоянии и пространственном расположении друг друга в текущий момент времени $t$ (*третий этап, реализующий взаимодействие агентов*).

Этап 6. Согласованная реализация работоспособными агентами попытки поражения цели в случае выполнения заданных условий, наложенных на значения целевой функции (*этап атаки*).

Этап 7. Получение агентами информации о состоянии цели; переход к этапу 2 в случае неудачной попытки поражения цели или завершение игры в случае удачной попытки её поражения или отсутствия работоспособных агентов (*этап контроля*).

Взаимодействие агентов данной системы обеспечивается на этапах 2, 3 и 5. Перемещение каждого агента по ячейкам рассматриваемой области поверхности описывается *марковским случайным процессом с дискретными состояниями и непрерывным временем*, структура которого приведена на рисунке 1. В рамках данной постановки задачи цель полагается неподвижной.

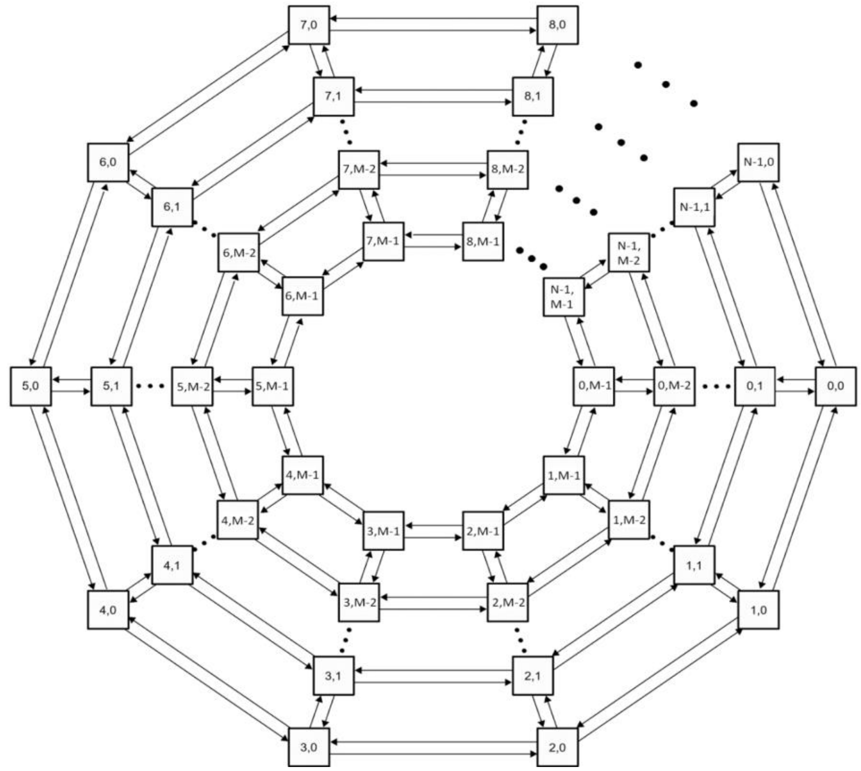


Рисунок 1. Структура связей между состояниями марковского случайного процесса, представляющего перемещения агентов по ячейкам игрового поля.

Пребывание в состоянии марковского процесса соответствует пребыванию в ячейке указанной области поверхности, имеющей те же самые индексы, а переходы, для которых выполняются свойства пуассоновских потоков событий, возможны только между состояниями, соответствующими *смежным ячейкам*, имеющим общую границу, не совпадающую с точкой. Модель этого процесса может быть представлена ориентированным графом. Число переходов между смежными состояниями $X$, попадающих в любой временной интервал τ, начинающийся в момент $t$, распределено согласно *закону Пуассона*:

$$P_{t,\tau}(X = m) = \frac{a(t,\tau)^m}{m!} e^{-a(t,\tau)},$$

где $P_{t,\tau}(X = m)$ – вероятность появления $m$ переходов в течение рассматриваемого интервала, $a(t,\tau)$ – среднее число переходов, попадающих в интервал $\tau$, начинающийся в момент времени $t$. Далее рассматриваются только стационарные потоки переходов, в которых $a(t,\tau) = \eta\tau$, а $\eta = const$ есть интенсивность стационарного потока. Предположение о пуассоновском распределении переходов между смежными состояниями обычно для прикладных задач, поскольку это распределение часто встречается на практике благодаря предельным теоремам для потоков событий.

Поведение каждого агента определяется автономно. Динамика изменения вероятностей пребывания $k$-го агента в состояниях марковских процессов определяется системой *уравнений Колмогорова* в матричной форме:

$\frac{d\boldsymbol{p}_k(t)}{dt} = \boldsymbol{M}_k(\boldsymbol{\lambda}_k)\boldsymbol{p}_k(t)$, где $\boldsymbol{p}_k(t)$ представляет вероятности пребывания $k$-го агента в $n$ состояниях процесса, $\boldsymbol{\lambda}_k$ – множество интенсивностей переходов между смежными состояниями для $k$-го агента, $\boldsymbol{M}_k$ – имеющая порядок $n$ матрица интенсивностей переходов между состояниями для $k$-го агента. Начальные условия: $p_{k,i_0 j_0}(0) = 1$, $\{p_{k,ij}(0) = 0\}_{i \neq i_0,\ j \neq j_0}$, где $(i_0, j_0)$ – индексы ячейки, в которой $k$-й агент находится при $t = 0$.

Расчёт вероятностей $\boldsymbol{p}_k(t)$ для всех агентов выполняется синхронно в дискретные моменты времени с шагом $\Delta t$. Допускается нахождение нескольких агентов в одной и той же ячейке.

Используются следующие обозначения:

- $A$ – поражение цели в случае её атаки;
- $B_k$ – поражение $k$-го агента в случае его атаки целью;
- $D_k$ – атака цели $k$-м агентом;
- $C_k$ – поражение цели в случае её атаки $k$-м агентом;
- $H_{ijk}$ – пребывание $k$-го агента в ячейке $(i, j)$.

Вероятность поражения цели в случае её атаки $k$-м агентом рассчитывается по формуле полной вероятности:

$$p(C_k) = \sum_{i,j} p\left(A|H_{ijk}\right) p\left(H_{ijk}\right)$$

*Вероятность поражения цели $k$-м агентом* при атаке из ячейки $(i, j)$ в момент времени $t$ определяется *картой осуществимостей*, представленной функцией $f_a$:

$$p\left(C_k|H_{ijk}\right) = f_a(i, j, t).$$

Вероятности $p\left(H_{ijk}\right)$ вычисляются путём решения приведённой выше системы уравнений Колмогорова.

*Вероятность поражения целью $k$-го агента* в ячейке $(i, j)$ в момент времени $t$ определяется *картой уязвимостей*, представленной функцией $f_b$:

$$p\left(B_k|H_{ijk}\right) = f_b(i, j, t).$$

Карты осуществимостей и уязвимостей пересчитываются на каждом такте игры, отслеживая перемещение цели, что позволяет учитывать её движение по игровому полю. Распределения вероятностей, которые задают эти карты, при решении многих прикладных задач целесообразно задавать как произведение двух логистических функций, а именно:

$$f_a(i, j, t_*) = c_a \left(\frac{e^{r_{a,d} d_{ij} + q_{a,d}}}{1 + e^{r_{a,d} d_{ij} + q_{a,d}}}\right) \left(\frac{e^{r_{a,h} h_{ij} + q_{a,h}}}{1 + e^{r_{a,h} h_{ij} + q_{a,h}}}\right),$$

$$f_b(i, j, t_*) = c_b \left(\frac{e^{r_{b,d} d_{ij} + q_{b,d}}}{1 + e^{r_{b,d} d_{ij} + q_{b,d}}}\right) \left(\frac{e^{r_{b,h} h_{ij} + q_{b,h}}}{1 + e^{r_{b,h} h_{ij} + q_{b,h}}}\right),$$

где $d_{ij}$ – расстояние между центром ячейки $(i, j)$ и целью, $h_{ij}$ – разность высот между центром ячейки $(i, j)$ и целью; параметры $c_a$, $c_b$, $r_{a,d}$, $q_{a,d}$, $r_{b,d}$, $q_{b,d}$ идентифицируются методом максимального правдоподобия по имеющимся эмпирическим данным так, чтобы обеспечить наибольшую вероятность наблюдаемых удачных и неудачных попыток поражений цели и агентов в контрольной серии опытов.

Перемещения между смежными ячейками выполняются со скоростью $\vec{v} = (v_\lambda, v_\mu)$, имеющей следующие компоненты:

$v_\lambda = \Delta l_\lambda / \tau_{\lambda*}$, $v_\mu = \Delta l_\mu / \tau_{\mu*}$, где $\Delta l_\lambda$ и $\Delta l_\mu$, соответственно, есть расстояния между центрами смежных ячеек в радиальном и трансверсальном направлении, а $\tau_{\lambda*}$ и $\tau_{\mu*}$ – периоды времени, затрачиваемые на преодоление указанных расстояний.

Вероятность поражения цели при коллективной атаке $p(A)$ определяется по формуле сложения вероятностей. События $C_i$ и $C_j$ полагаются независимыми при $i \neq j$. Допускается расположение нескольких агентов в одной и той же ячейке.

Реализация этапов 2, 5 и 7 эволюции системы обеспечивается имеющимися техническими средствами и лежит вне математической постановки задачи. Задача, решаемая на этапах 1, 3, 4 и 6, формулируется следующим образом.

**Даны:**

(1) структура связей между состояниями марковского случайного процесса, представляющего перемещения агентов по ячейкам игрового поля;
(2) распределение всех работоспособных агентов по состояниям данного марковского случайного процесса в момент времени $t$;
(3) функция $f_a(i,j,t)$, определяющая вероятность поражения цели при атаке из ячейки $(i,j)$ в момент времени $t$;
(4) функция $f_b(i,j,t)$, определяющая вероятность поражения целью агента в ячейке $(i,j)$ в момент времени $t$;
(5) пороговые вероятности индивидуальной $p_t$ и коллективной $p_{tn}$ атаки цели;
(6) пороговая вероятность завершения игры $p_{tmax}$;
(7) наибольшая допустимая вероятность поражения агентов $p_B$;
(8) наибольшая допустимая скорость перемещения агентов по игровому полю $v_{max}$;
(9) число агентов $B$, одновременно атакуемых целью.

**Найти:** распределение всех работоспособных агентов по состояниям заданного марковского случайного процесса в момент времени $t + \Delta t$, обеспечивающее наибольшее значение вероятности поражения цели

– при коллективной атаке в случае наличия информации о положении работоспособных агентов или
– при индивидуальной атаке в случае её отсутствия

при условии

– выполнения всех ограничений, заданных параметрами (5)-(7) условия задачи,
– оценки вероятностей поражения цели и агентов с помощью функций $f_a(i,j,t)$ и $f_b(i,j,t)$и
– недетерминированного перемещения агентов по состояниям марковского процесса.

Для решения данной задачи разработан специальный алгоритм. Следует отметить, что при его выполнении не решается математическая задача управления в классической постановке, однако численное решение задачи оптимизации является одним из основных компонентов.

Эволюция системы определяется следующим алгоритмом.

- Шаг **1**. Задать начальные условия задачи.
- Шаг **2**. Для текущего расположения агентов в текущий момент времени $t_*$ определить не более $B$ агентов, имеющих наибольшие вероятности поражения целью в соответствии с «картой уязвимостей» $f_b(i,j)$ и случайным образом, соразмерно этим вероятностям, удалить с игрового поля часть указанных агентов; проверить критерии индивидуальной ($p(\mathrm{C}_k) \geq p_t$) или коллективной ($p(\mathrm{C}_1 + \cdots + \mathrm{C}_n) \geq p_{tn}$) атаки цели агентами; если хотя бы один из них выполнен, то атаковать цель.

- Шаг **3**. Если в момент времени $t_*$ выполнено хотя бы одно из условий завершения игры (исход, приводящий к поражению всех агентов либо цели), завершить игру (перейти к шагу 6).
- Шаг **4**. Выполнить идентификацию значений свободных параметров марковских процессов $\{\boldsymbol{\lambda_k}\}_{k=0,\ldots,K-1}$ при условии $|\vec{v}_k| \le v_{max}$ . Если агенты имеют возможность получать релевантную информацию друг о друге, то максимизируемой *целевой функцией игры*, определяющей результаты этой идентификации, является вычисленная с учётом всех агентов групповая вероятность поражения цели $p(\mathrm{C}_1 + \cdots + \mathrm{C}_n)$ в момент времени $t_* + \Delta t$; в противном случае задача идентификации решается для каждого агента автономно, причём в качестве максимизируемых целевых функций игры используются индивидуальные вероятности поражения цели $p(\mathrm{C}_k)$ в тот же момент времени $t_* + \Delta t$.
- Шаг **5**. Выбрать одну из ячеек игрового поля, смежных по отношению к ячейке, в которой находится $k$-й агент в момент времени $t_*$. Использовать «*метод рулетки*» с вероятностями выбора объектов, пропорциональными прогнозируемым байесовским оценкам

$p(\widetilde{H}_{ijk}\,|C_k) = \frac{p(C_k|\widetilde{H}_{ijk}\,)p(\widetilde{H}_{ijk}\,)}{p(C_k)} = \frac{p(\widetilde{H}_{ijk}\,)}{p(C_k)}$, где вероятности пребывания $k$-го агента в момент времени $t_* + \Delta t$ в ячейке $(i,j)$, смежной по отношению к ячейке, в которой этот агент находился в момент времени $t_*$ $p(\widetilde{H}_{ijk})$, рассчитываются для момента времени $t_* + \Delta t$ как результат выполнения предыдущего шага алгоритма. Переместить в неё этого агента со скоростью $\vec{v}_k$, случайные компоненты которой рассчитываются, используя идентифицированные интенсивности переходов между состояниями, если выполнены следующие ограничения: $|\vec{v}_k| \le v_{max}$, $p(B_k) \le p_B$; Если не выполнено ограничение $|\vec{v}_k| \le v_{max}$, то перемещение происходит со скоростью $|\vec{v}_k| = v_{max}$ ; если не выполнено ограничение $p(B_k) \le p_B$, то перемещение не происходит.

Интервал дискретизации $\Delta t$ определяется на данном такте игры как наибольшее время, необходимое для перемещения между центрами смежных ячеек: $\Delta t = \max\limits_{k\in\{0,\ldots,K-1\}}(\Delta l_k/|\vec{v}_k|)$, где $\Delta l_k = \sqrt{{\Delta l_{\lambda,k}}^2 + {\Delta l_{\mu,k}}^2}$. Переходы агентов между состояниями синхронизируются по единому (для всех этих объектов) интервалу $\Delta t$.

Перейти к следующему по порядку дискретному моменту времени $t_* + \Delta t$, рассматривая его далее как текущее время; перейти к шагу 2.

- Шаг **6**. Завершить игру.

Для прогноза состояния системы в случае простых вариантов развития игры получены аналитические выражения для вероятностей поражения цели в случае её атаки $k$-м агентом и поражения $k$-го агента в случае его атаки целью.

Назовём поведение $k$-го агента *простейшим*, если

– он перемещается к цели, сохраняя постоянную ненулевую интенсивность переходов между смежными кольцами в направлении цели и нулевую интенсивность переходов между смежными кольцами в противоположном направлении,

– вероятность поражения цели агентом из $i$-го кольца есть $p\left(A \middle| E_{k,i}\right) = if_{a*}$, где $f_{a*} = const$,

– вероятность поражения целью агента из $i$-го кольца есть $p\left(B_k \middle| E_{k,i}\right) = if_{b*}$, где $f_{b*} = const$.

Объединим все ячейки каждого кольца игрового поля в укрупнённые состояния и рассмотрим марковский процесс с постоянной интенсивностью переходов, представляющий случайные переходы от одного кольца к другому в случае простейшего поведения. В этом случае интенсивность переходов между смежными кольцами равна сумме интенсивностей переходов между соответствующими парами состояний по секторам (в рассматриваемом случае эти интенсивности не зависят от номеров колец):

$$\lambda_k = \lambda_{k,i} = \sum_{j=0}^{N-1} \lambda_{k,i,j,>}, i = 0, \dots, M-2.$$

В случае указанного процесса с постоянной интенсивностью $\lambda_k$ вероятность попадания в $i$-е состояние в течение временного интервала $\tau$ есть вероятность совершения $i$ переходов между состояниями за указанное время:

$$p_\tau\left(E_{k,i}\right) = P_{0,\tau}(X = i) = \frac{(\lambda_k \tau)^i}{i!} e^{-\lambda_k \tau}.$$

Множество $\left\{E_{k,i}\right\}_{i=0,\dots,M-1}$ есть полная группа несовместных событий для $k$-го агента. Вероятность поражения цели этим агентом в течение временного интервала $\tau$ определяется по формуле полной вероятности:

$$p_\tau(A) = \sum_{i=0}^{M-1} p\left(A \middle| E_{k,i}\right) p_\tau(E_{k,i}).$$

Полагая события $A$ и $B_k$ независимыми, получены следующие выражения для событий, имеющих место в течение временного интервала $\tau$:

$$p_\tau(A|\bar{B}_k) = p(A)p(\bar{B}_k) = f_{a*}R(\lambda_k \tau, M)\left(1 - f_{b*}R(\lambda_k \tau, M)\right),$$
$$p_\tau(B_k|\bar{A}) = p(B_k)p(\bar{A}) = f_{b*}R(\lambda_k \tau, M)\left(1 - f_{a*}R(\lambda_k \tau, M)\right),$$

где

$$R(\lambda_k \tau, M) = \sum_{i=0}^{M-1} \frac{i(\lambda_k \tau)^i}{i!} e^{-\lambda_k \tau}.$$

Таким образом, доказаны следующие утверждения.

Утверждение 1. Вероятность поражения цели $k$-м агентом в течение временного интервала $\tau$ в случае его простейшего поведения и сохранения непоражённым есть $f_{a*}R(\lambda_k \tau, M)\left(1 - f_{b*}R(\lambda_k \tau, M)\right)$.

Утверждение 2. Вероятность поражения целью $k$-го агента в течение временного интервала $\tau$ в случае его простейшего поведения и сохранения цели непоражённой есть $f_{b*}R(\lambda_k \tau, M)\left(1 - f_{a*}R(\lambda_k \tau, M)\right)$.

Очевидно, что в общем случае, когда поведение агентов простейшим не является, задача поддаётся решению только с помощью численных методов.

**В третьей главе** рассматриваются методы оценки макропараметров модели, позволяющих осуществлять прогноз развития событий во время работы моделируемой системы, с целью определения стратегии действий лица, принимающего решение.

Используя рассмотренный подход, можно путём вычислительных экспериментов оценить, сколько агентов необходимо, чтобы поразить цель, а также исследовать зависимость исхода от числа агентов и распределений вероятностей поражения целей и агентов. Однако, учитывая стохастический характер игры, её исход практически остаётся в бóльшей или мéньшей степени неопределённым, и однократное моделирование хода игры не может служить основанием для прогноза.

Чтобы обеспечить прогнозирование результатов игры при решении различных практических задач, опираясь на общие характеристики её начальных условий, следует:

- определить удобные для интерпретации и практического контроля макропараметры,
- вычислить для каждого сочетания таких параметров путём имитационного моделирования достаточно представительный ансамбль реализаций хода игры и
- определить статистические характеристики различных её исходов.

Для рассматриваемой игры в качестве таких макропараметров удобно использовать:

- $T_{win}$ – время до выигрыша в тактах,
- $L_{lost}$ – относительное количество поражённых агентов в конце игры,
- $N_r$ – номер ближайшего к цели кольца, в котором располагается агент,
- $P_a$ – количество агентов, располагающихся на игровом поле.

С целью обеспечить независимость прогнозируемого игры результата от конкретного вида игрового поля, первый и четвёртый из указанных параметров рассматриваются в отношении к числу колец. Значения макропараметров $N_r$ и $P_a$ в момент времени $t = i\Delta t$ будем обозначать как $N_{ri}$ и $P_{ai}$, соответственно.

Проведя достаточное число вычислительных экспериментов и их статистический анализ, результаты такого прогнозирования могут быть представлены в виде зависимостей $T_{win,\gamma} = f_T(N_{r0}, P_{a0})$ и $L_{lost,\gamma} = f_L(N_{r0}, P_{a0})$, где $T_{win,\gamma}$и $L_{lost,\gamma}$ – $\gamma$ -квантили эмпирических распределений величин $T_{win}$ и $L_{lost}$, $N_{r0}$ и $P_{a0}$ – значения макропараметров в начальный момент времени ($t = 0$), которые графически удобно изображать на плоскости в виде поверхностей уровня.

Знание зависимостей $T_{win,\gamma} = f_T(N_{r0}, P_{a0})$ и $L_{lost,\gamma} = f_L(N_{r0}, P_{a0})$ позволяет решить *обратную задачу*.

<u>Обратная задача</u>. Найти начальные значения макропараметров $N_{r0}$ и $P_{a0}$, обеспечивающие выполнение неравенств $T_{win} \le T_*$ и $L_{lost} \le L_*$ с вероятностью $p \ge p_*$.

Для решения этой задачи следует использовать имеющиеся представления зависимостей $T_{win,p*} = f_T(N_{r0}, P_{a0})$ и $L_{lost,p*} = f_L(N_{r0}, P_{a0})$, с помощью которых в области значений макропараметров $N_r$ и $P_a$ определяется пересечение множеств, заданных неравенствами $T_{win} \le T_*$ и $L_{lost} \le L_*$, что и является искомым результатом.

В терминах указанных макропараметров, вероятностная динамика игры описывается *марковским случайным процессом с дискретными состояниями и дискретным временем* (рисунок 2). Каждое состояние этого процесса соответствует сочетанию диапазонов значений макропараметров $N_r$ и $P_a$ ($q$ диапазонов для $N_r$ и $l$ диапазонов для $P_a$). Интерпретация этих макропараметров делает допустимыми переходы только между смежными состояниями, как в общем виде показано на рис. 3. Число используемых диапазонов определяет точность прогнозирования.

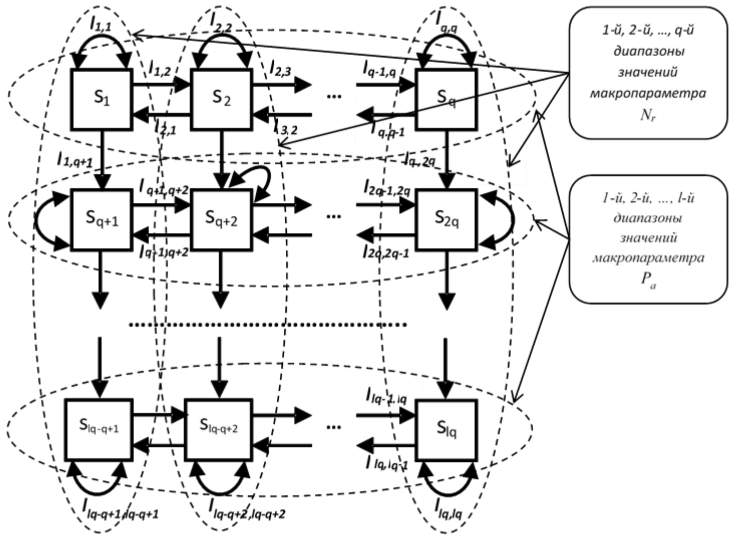


Рисунок 2. Марковский случайный процесс с дискретными состояниями и дискретным временем, представляющий вероятностную динамику игры в терминах макропараметров.

Следует заметить, что, как правило, не все вероятности переходов между состояниями, представленные на рис. 3, являются ненулевыми. В частности, поскольку агенты не восстанавливаются после поражения, переходы в сторону увеличения их численности имеют нулевую вероятность.

Распределения вероятностей пребывания в состояниях в смежные моменты времени $i\Delta t$ и $(i+1)\Delta t$ связаны матричным уравнением:

$$\boldsymbol{p}_{L,i+1} = \boldsymbol{L}\,\boldsymbol{p}_{L,i},$$

где $\boldsymbol{p}_{L,i}$ и $\boldsymbol{p}_{L,i+1}$ представляют вероятности пребывания указанного марковского процесса с дискретными состояниями и дискретным временем в определённых выше $lq$ состояниях в смежные моменты времени $i\Delta t$ и $(i+1)\Delta t$, соответственно; $\boldsymbol{L} = [l_{ij}]$ – матрица вероятностей переходов между этими состояниями, имеющая порядок $lq$. В начальный момент времени компонент $\boldsymbol{p}_{L,0}$, соответствующий исходному сочетанию диапазонов значений макропараметров $N_r$ и $P_a$, равен единице, остальные компоненты этого вектора равны нулю.

Используя указанное матричное уравнение, можно решить *прямую задачу*.

<u>Прямая задача</u>. Задано распределение вероятностей $\boldsymbol{p}_{L,0}$ пребывания в состояниях марковского процесса, представленного на рис. 3, в начальный момент времени. Найти распределение вероятностей $\boldsymbol{p}_{L,k}$ в состояниях данного процесса в момент времени $k\Delta t$, где $k$ – натуральное число.

Искомое распределение тривиально вычисляется с помощью матричной формулы:

$$\boldsymbol{p}_{L,k} = \boldsymbol{L}^k \boldsymbol{p}_{L,0}\ .$$

Моделирование поведения марковских процессов позволяет вычислять ансамбли их реализаций, используя которые можно идентифицировать процесс, представленный на рис. 2, вычислив выборочные оценки вероятностей переходов $l_{ij}$ между состояниями рассмотренного процесса с дискретным временем, а именно:

$$l_{ij} = \frac{F_{ij}}{\sum_{k=1}^{lq} F_{kj}},$$

где $F_{ij}$ – наблюдаемая частота переходов из состояния $j$ в состояние $i$.

Определённый выше процесс с дискретным временем позволяет исследовать общие закономерности динамики игры без привязки к конкретным распределениям агентов по игровому полю и прогнозировать, как изменяются со временем вероятностные распределения значений макропараметров. Переход к описанию хода игры в терминах макропараметров приводит к определённой потере информации о деталях игры, однако позволяет перейти от воспроизведения конкретных её реализаций к исследованию общих закономерностей, выраженных оценками распределения вероятностей и представляющих практический интерес.

**В четвертой главе** описывается программная реализация вероятностной модели поведения прикладной многоагентной системы. Исследованы языки и парадигмы программирования, на основе проведенного анализа выбраны конкретные языки для реализации алгоритма. Программа позволяет сохранять, загружать и модифицировать данные модели, а также осуществлять ее генерацию и следить за ходом изменения моделируемой игровой ситуации с указанным шагом. Организовано взаимодействие двух процессов, написанных на разных языках, в рамках одного исполняемого файла приложения для гарантии переносимости программной версии модели и использования всех доступных преимуществ выбранных языков программирования.

При программной реализации марковский процесс представляется последовательностью связанных между собой блоков с одним входом (X) и выходом (Y), а также значением вероятности, характеризующим сам блок (P). Вход и выход подсоединяются к следующему блоку в последовательности; последний блок, в случае нециклического соединения блоков, имеет значения входа и выхода равными 0.

Нумерация колец и секторов начинается с 0. Кольца нумеруются от самого удаленного от центра кольца, сектора - по часовой стрелке от сектора, наиболее близкого к нулевому градусу.

Структура для хранения марковской модели состоит из двоичного флага, определяющего цикличность, а также из массива вещественных чисел, задающего значения X, Y и P для всех блоков модели.

Совокупность групп является частью структуры, задающей свойства агента в рамках моделируемой игровой ситуации. Также общая структура агента определяется координатами его местоположения (парой целочисленных переменных $P_X$ и $P_Y$), заданной скоростью и двоичным флагом выхода из строя. Дополнительно вводится двоичный флаг, определяющий первый ход агента, чтобы предотвратить избыточный перерасчет значений вероятностей.

Игровое поле задается тремя полями: структурой, описывающей функцию связи между агентами; словарем, хранящим местоположение агентов; массивом вещественных чисел, задающим карты осуществимостей и уязвимостей. Для реализации было решено использовать такие технологии, которые бы в явном виде поддерживали следующие принципы: поддерживаемость, простоту, современность, высокую производительность, универсальность, функциональную полноту, кроссплатформенность, независимость, компилируемость и естественную удобочитаемость. В результате исследований более 20 языков программирования было решено осуществить программную реализацию модели на языках C++, Python и Golang. Компилятор Golang был расширен самостоятельно написанной надстройкой, позволяющей осуществлять кроссплатформенную компиляцию динамических библиотек для осуществления межъязыкового взаимодействия.

**В пятой главе** рассматривается численный метод оптимизации, применяемый для идентификации параметров марковских процессов. В основу программно-алгоритмической реализации метода оптимизации положен шаблон проектирования «декоратор». Параметры, существенно влияющие на качество и время работы основного алгоритма, но не относящиеся к числу оптимизируемых основным методом значимых параметров, будем называть «параметрами основного алгоритма». Алгоритм-надстройку, осуществляющую модификацию метапараметров основного алгоритма с целью увеличения критерия эффективности, будем называть «внешний алгоритм оптимизации».

Эффективность основного алгоритма за одну итерацию можно рассматривать как отношение значения максимизируемого критерия ко времени, затраченному на работу алгоритма, при условии, что исходные значимые параметры остаются неизменными между итерациями (сохраняют свои изначальные значения в качестве входных параметров основного алгоритма). Тогда, адаптивно изменяя параметры основного алгоритма оптимизации, оказывающие наибольшее влияние на рост его эффективности, можно повысить эффективность основного математического метода в целом.

Эмпирически было определено, что такими метапараметрами являются:

- M – число варьируемых значимых параметров,
- $\Delta$ – начальная поправка для значимых параметров.

Представим «сетку» - двумерное пространство, в котором по одной оси с фиксированным шагом $h_1$ задается количество единовременно изменяемых параметров от 1 до n, а по другой – также, с фиксированным шагом $h_2$, определяется начальный сдвиг искомых параметров (от $h_2$ до 1). Полученное пространство позволяет определить значение критерия эффективности метода в каждой ячейке «сетки» - в зависимости от конкретного значения модифицируемых параметров. В разработанном численном методе задача поиска в рамках заданного пространства параметров выполняется по следующему принципу: выбирается случайная ячейка сетки, для которой, вместе со смежными ячейками, выполняется процедура оптимизации с соответствующими значениями параметров основного алгоритма. Фиксируется значение критерия для каждой пары параметров и затраченное на оптимизацию время. Отношение значения критерия к затраченному времени определяет эффективность внешнего алгоритма оптимизации для заданной ячейки. Предпринимается попытка перехода из заданной ячейки в смежную с максимальным значением эффективности при условии, что ячейка не была посещена ранее. Если такой переход невозможен, вновь выбирается случайная ячейка, и процесс повторяется. Поиск завершается, если превышено суммарное пороговое время, затраченное на выполнение процедур оптимизации поиска, или посещены все ячейки. В итоге, среди посещенных выбирается ячейка с максимальным значением эффективности и соответствующие этой ячейке значения параметров основного алгоритма оптимизации задаются как оптимальные.

Данный метод может быть применен для улучшения представленного метода оптимизации и может быть расширен для n-мерного пространства параметров за счет увеличения числа шагов, обеспечивающих поиск вокруг текущей точки. В частности, при переходе от произвольно выбранной начальной точки в пространстве «сетки» параметров к конечной точке критерий эффективности был увеличен с 10,84 до 13,69 (на 26%). Представленный математический метод может повысить эффективность работы произвольного алгоритма оптимизации за счет адаптивного изменения его параметров.

**В шестой главе** показано практическое применение многоагентной системы на примере разработанного программного комплекса «Тренажер». Этот комплекс программ включает в себя разработанный и представленный в пятой главе численный метод, набор программ по созданию новых и редактированию уже построенных моделей многоагентных систем с целью моделирования конкретной ситуации, заданной параметрами, а также систему оценки макропараметров и осуществления процедур прогнозирования. Дополнительно включена объединяющая программа, предоставляющая пользовательский интерфейс для совместного использования всех составляющих программного комплекса в роли тренажера, предоставляющего средства для адаптивного обучения и определения текущего уровня подготовки операторов сложных систем.

Были рассмотрены аналогичные системы, работа которых могла бы обеспечить описанный процесс обучения. Для некоторых из них характерно наличие математической модели, описывающей адаптивное обучение, однако ограничения реализации в рамках конкретной предметной области не позволяют организовать процесс обучения для оператора сложных систем в общем виде. Для других систем не предусмотрено наличие специализированной математической модели для оценки уровня обучения по заданным критериям, несмотря на адаптацию к предметной области сложных систем. С учетом недостатков исследованных подходов к решению поставленной выше задачи был разработан и реализован программный комплекс – виртуальный адаптивный тренажер.

Принцип работы тренажера основан на марковской модели гибели и размножения, представляющей переходы испытуемого от уровня 0 до уровня M-1, где M – количество доступных тренажеру уровней. Уровень позволяет условно оценить степень компетентности оператора. Предполагается, что при выполнении поставленной задачи значение текущего уровня для оператора рано или поздно (в бесконечности) сойдется к значению, соответствующему приближенной оценке его способностей.

Подбор уровней мог бы осуществляться эмпирически; однако, разработанный ранее механизм прогнозирования благоприятной стратегической ситуации с помощью макропараметров позволяет значительно упростить данную процедуру. Число используемых уровней определяет точность прогнозирования компетентности оператора. При необходимости, уровни можно редактировать вручную с помощью специального редактора, входящего в состав комплекса программ.

Программа, обеспечивающая работу тренажера как целостного комплекса, представляет собой кросс-платформенное приложение с графическим кросс-браузерным интерфейсом на базе языка разметки гипертекста с поддержкой стандарта HTML5, что обеспечивает его переносимость и удобство интеграции в любую современную вычислительную систему. Фрагменты схемы тренажера в нотации BPM, описывающей процесс генерации пула уровней на основе разработанной модели тренажера, представлены на рисунках 3-4.

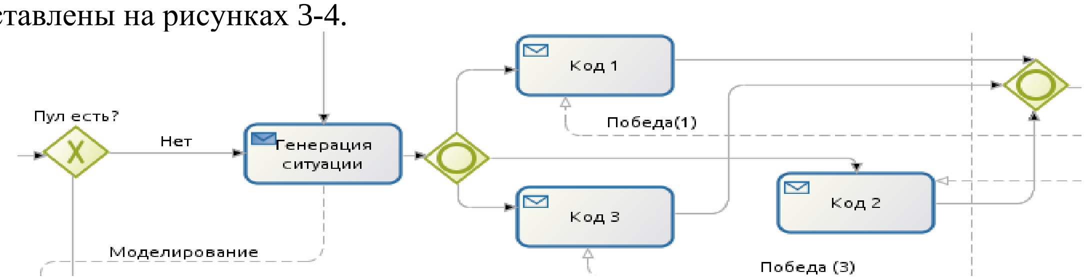


Рисунок 3. Фрагмент генератора пула уровней.


Рисунок 4. Формирование пространства модели (симуляция и отклик).

Тренажер позволяет избежать затрат, связанных с эксплуатацией и сопутствующим износом или выходом из строя реального дорогостоящего оборудования. Тренажер обеспечивает механизм по контролю за уровнем компетентности оператора в реальном времени с помощью макропараметров. Реализованная комплексная схема данных поддерживает последующее расширение текущих возможностей модели через добавление новых (точная система повреждений агента, более гибкая модель подавления сигнала и т.п.). Кроссплатформенная программная реализация тренажера представлена на рисунке 5.

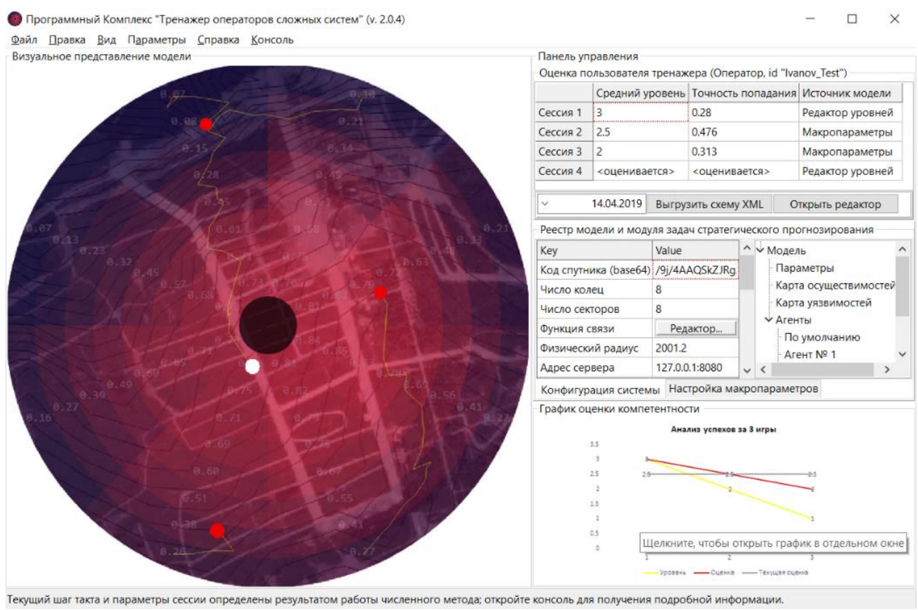


Рисунок 5. Интерфейс программного комплекса.

# Основные результаты работы, выносимые на защиту

1. Математическая модель и алгоритм поведения прикладной многоагентной системы. Модель представлена марковским процессом с дискретными состояниями и непрерывным временем и обеспечивает коллективное и автономное поведение агентов, а также их недетерминированное перемещение по рассматриваемой области поверхности [1, 6]. Основные результаты приведены в источниках [1-11]. К особенностям модели и ее программной реализации [12] относятся:

- способность обучаться без сохранения предыстории действий;
- способность быстро приспосабливаться к изменениям значений параметров, характеризующих свойства моделируемой системы;
- кроссплатформенность, переносимость и расширяемость программной реализации [2, 10];
- способность модели удовлетворять представленным требованиям на программно-аппаратных вычислительных системах с ограниченным количеством ресурсов без принципиальных потерь в скорости принятия решений [4].

2. Математическая модель и метод прогнозирования [9, 11], обеспечивающие оперативную оценку ресурсов, необходимых для решения задачи, на основе количественных критериев [7]. Особенностями модели являются:
   - применение макропараметров, вычисленных, используя эмпирические данные, и определяющих состояние системы и качественные характеристики моделируемой ситуации на заданном такте времени;
   - комплекс программ, обеспечивающий вычисления в автоматическом режиме и позволяющий наглядно представить конечную информацию стратегического характера для лица, принимающего решение.
3. Адаптивный численный метод оптимизации, используемый при идентификации параметров прикладной многоагентной системы [3]. Эффективность метода подтверждена серией вычислительных экспериментов.
4. Комплексы программ для прогнозирования поведения прикладной многоагентной системы и оценки уровня подготовки и адаптивного обучения операторов, работающих с этой системой [5, 8, 10].
5. Аналитические выражения для вероятности поражения цели агентом в случае его простейшего поведения [6].

**Публикации в журналах из перечня ВАК**


1. Куравский Л.С., Попков С.И. Вероятностная модель поведения прикладной многоагентной системы. – Нейрокомпьютеры: разработка, применение, № 9, 2016, с. 22-34.
2. Попков С.И. Программная реализация вероятностной модели поведения прикладной многоагентной системы. – Нейрокомпьютеры: разработка, применение, № 9, 2016, с. 35-44.
3. Попков С.И. Метод внешней оптимизации для идентификации марковских процессов. – Информационные технологии, № 10, 2018, с. 633-641.
4. Попков С.И. Программная реализация межъязыкового взаимодействия на базе динамических библиотек. – Нейрокомпьютеры: разработка, применение, № 3, 2018, с. 39-49.
5. Попков С.И. Применение и разработка тренажера для автоматизированных БЛА и робототехнических комплексов на базе вероятностной модели поведения прикладной многоагентной системы. – Нейрокомпьютеры: разработка, применение, № 5, 2019, с. 5-17.

**Публикации в журналах, индексируемых в международных библиографических базах данных**

6. Kuravsky L.S., Popkov S.I. and Artemenkov S.L. An applied multi-agent system within the framework of a player-centered probabilistic computer game. - International Journal of Modeling, Simulation, and Scientific Computing, Vol. 9, No. 1 (2018), 17 pp, DOI: 10.1142/S1793962317500635 [SCOPUS].
7. Kuravsky L.S., Popkov S.I. Forecasting macro parameters representing the behavior of an applied multi-agent system. - International Journal of Modeling, Simulation, and Scientific Computing, 2018, Vol. 9, No. 6 (2018), 15 pp, DOI: 10.1142/S1793962318500526 [SCOPUS].

**Публикации в сборниках научных трудов**

8. Kuravsky L.S., Popkov S.I. and Artemenkov S.L. Applied multi-agent system to study behavior of operators of complex technical systems. – In: Proc. First World Congress on Condition Monitoring 2017 (WCCM 2017) The International Society for Condition Monitoring (ISCM), British Institute of Non-Destructive Testing (BINDT). 2017 [SCOPUS].
9. Kuravsky L.S., Popkov S.I. Forecasting behavior of a stochastic multi-agent system. - In: Proc. 15th International Conference on Condition Monitoring & Machinery Failure Prevention Technologies, Nottingham, UK, September 2018 [SCOPUS].
10. Попков С.И. Программная реализация вероятностной модели поведения прикладной многоагентной системы (тезисы). XV Всероссийская научная конференция «Нейрокомпьютеры и их применение». МГППУ, 2017. – 2 с.
11. Куравский Л.С., Попков С.И. Представление общих закономерностей поведения многоагентной системы с помощью ее макропараметров (тезисы). XVI Всероссийская научная конференция «Нейрокомпьютеры и их применение». МГППУ, 2018. – 4 с.

**Свидетельство о государственной регистрации программы**

12. Попков С.И., Куравский Л.С. Свидетельство о государственной регистрации программы для ЭВМ №2017618950 "Программа для моделирования стохастического поведения прикладной многоагентной системы (St#MAS)” /Правообладатели Попков С.И., Куравский Л.С. (Россия). — Заявка №2017615896; Заяв. 20.06.2017; Зарегистр. 11.08.2017.—(РОСПАТЕНТ).