

# ПОСТРОЕНИЕ АГЕНТНЫХ МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ

© Неклюдов Д.Ю.\*

Национальный исследовательский университет «Высшая школа экономики»,  
Москва

Работа посвящена агент-ориентированному моделированию. В ней говорится о том, что для ряда областей можно минимизировать количество данных для построения модели, используя закон по которому данный процесс происходит. А также в работе уделено внимание автоматизированному созданию агентных моделей распространения информации, основываясь на динамики распространения информации в социальной сети.

Данная работа является исследованием, относящимся к области агент-ориентированного моделирования. В работе затрагиваются теория графов, теория распространения эпидемий, теория распространения знаний, стохастическое моделирование и генетические алгоритмы. Одна из современных проблема заключается в том, что для анализа и прогнозирования требуется большое количество данных, и часто чем больше данных (видов данных), тем точнее осуществляется прогноз. И это абсолютно верно для тех областей, законы которых науке неизвестны, но вероятно есть возможность получения качественных результатов и при меньших данных для тех областей, законы которых нам известны или достаточно хорошо структурированы. Вероятно, сами законы уже содержат в себе основную, достаточную информацию, для проведения качественных анализов и построения весьма точных прогнозов. Именно, исходя из этого, в данной работе и берется научная область распространения информации в социальной среде, законы которой задаются самим обществом, а структура данного процесса распространения может быть показанная в виде сети графов.

В данной работе узлы сети графов – это агенты, которые распространяют или получают информацию (взгляд, моду, слухи и т.д.); ребра – это связи, через которые агенты обмениваются информацией (взглядами, модой, слухами и т.д.). Именно это и приводит нас к агент-ориентированному моделированию.

В современной науке особенную остроту приобретает тема агент-ориентированного моделирования, построения искусственных обществ, так как существует постоянное стремление к повышению качества анализа и точности прогнозирования. Особенную критичность к анализу и прогнозированию приобретает наука в вопросах непосредственно связанных с деятельностью людей, с взаимодействием людей, в вопросах управления обществом.

---

\* Магистрант.

Так агент-ориентированное моделирование, основываясь на теории графов и принципах распространения информации, может создавать «искусственные общества», осознание ценности которых растет с каждым годом. «Искусственные общества» хороши тем, что позволяют проводить над ними различные эксперименты без самой возможности нанести случайный вред самому социуму. До того как применять какое-либо изменение можно оценить социальный резонанс на получаемой модели. Имея подобную модель, можно опробовать на ней различные методы информационного воздействия, выбрать наиболее эффективный метод для данной ситуации, следствием всего этого и должно стать повышение качества анализа информационного взаимодействия в социуме, и точности прогнозирования результатов влияния различных методов на реальное общество. Степень разработанности агент-ориентированного моделирования, на текущий момент еще невелика, требуется создание технического инструмента для решения описанной проблемы.

Одна из обоснованных гипотез, решающих поставленную проблему такова, что можно создать инструмент автоматически генерирующего максимально реалистичную агент-ориентированную модель, отражающую передачу данных в какой-либо среде распространения информации по наблюдениям динамики распространения информации. Основываясь на возможности генетического алгоритма находить глобальный экстремум, при моделировании масштаба агентной модели 1:1, учитывая сложность социальных процессов и предполагая практическую невозможность двух социальных сетей иметь одинаковую динамику распространяющих узлов. Можно предположить, что будет получена реалистичная модель, полностью отражающая жизненную ситуацию, либо настолько близкую, что можно пренебречь отличиями от идеальной модели.

Цель работы – Создание максимально реалистичных моделей передачи информации (моды, взглядов, убеждений и т.д.) в социальной среде, путем подбора параметров конкретных социальных сетей по наблюдениям динамики распространения информации.

В настоящее время вычислительные мощности развились до такого состояния, что позволяют моделировать систему практически любой сложности с любым количеством взаимодействующих объектов, без необходимости их агрегирования.

Основными преимуществами АОМ считаются возможность смоделировать систему, которая будет максимально приближена к реальности, так как в ней не обязательно применение агрегирования, и при этом для анализа правил поведения реальных агентов требуется меньше сил, чем выявление и описание подобных правил для всей системы сразу. Одним из важных преимуществ агент-ориентированного моделирования является масштабируемость – легкое добавление и удаление агентов из модели. Именно благодаря механизму, заложенному в самом понятии агент-ориентированного моделирования – построение модели, которая максимально реалистично повторяет

реальный исследуемый объект, повторяя не само поведение системы, а повторяя поведение отдельных объектов – агентов системы, следствием этого и будет появление общих эффектов. И так агент-ориентированное моделирование может объяснить причины возникновения таких явлений, как терроризм, войны, всплески изменений на рынке акций и т.д.

Среди ожиданий развития данного направления присутствуют усиление текущих преимуществ. В связи с дальнейшим уменьшением стоимости вычислений будет происходить увеличение масштабов моделирования и как следствие увеличение точности прогнозирования, нахождение различных эффектов, которые могут появиться только на моделях большего размера.

Но, не смотря на уже достаточную проработанность и наличие уже зарекомендовавших себя средств, разработка модели с нуля на любом высокоуровневом языке является целесообразной, если планируется использовать какие-либо нетипичные для данных продуктов функции. В этой связи любой продукт, выполняющий различные сложные функции всегда тем самым ограничивает пользователя в функционале, и даже добавляя возможность дорабатывать что-либо, зачастую ограничивают или усложняют данный процесс.

Агент-ориентированная модель занимает ключевое место в нашей работе, но также стоит отметить и другие важные аспекты исследования и разработки. Для отображения и построения модели используется сеть графов, где агенты – это узлы, а ребра – это связи агентов. Теория графов является одной из наиболее разработанных научных теорий затрагиваемых в данной работе. Поэтому именно она станет основой для агентной модели. Из данной теории будет взято построение сети, параметры сетей. Как уже было указано, используя графы, будет строиться представление модели. Подобная форма вывода данных представляется распространенной и потому понятной.

В данной работе предполагается использование генетического алгоритма для подбора параметров системы с целью приближения ее к реальности, т.е. подбор параметров сети, узлов и т.д. Направление генетических алгоритмов является весьма проработанным, а принцип функционирования подобного алгоритма закладывается эволюцией.

Для того чтобы определить какие параметры должны быть у узлов и их связей и у тех параметров, которые задают систему в целом, необходимо использование принципов распространения информации. Т.е. это агенты, которые распространяют информацию, связи, по которым данная информация передается и собственно информация. Каждый из перечисленных объектов обладает собственными параметрами, определение данных параметров будет произведено в дальнейшем в ходе выполнения работы.

Основываясь на гипотезе данной работы, мы выделим ключевые блоки требуемой системы и их принцип и цели функционирования.

Отталкиваясь от цели работы необходимо построить реалистичную сеть, а значит должен быть блок программы, который способен строить сеть по

заданным параметрам сети. Второй составной частью системы является блок, который по методу генетического алгоритма способен подобрать сеть динамика распространения информации по агентам будет минимально отличаться от заданной динамики. Это и будет нашей целевой функцией, минимизацию которой и будет стремиться выполнить генетический алгоритм. И последний блок, без которого система не будет полноценной это визуальное отображение данной системы и статистики, которая и должна помочь удостовериться в том, что поставленная цель достигнута, далее будем называть это блок вывода.

Перейдем к содержанию каждого блока. Начнем в том же порядке, в котором и были перечислены все блоки.

Построение сети. Это основной блок программы, который строит сеть графов, связывает все узлы сети между собой, а также управляет агентами, просчитывает модель на каждый интервал времени. Все эти действия производятся с учетом параметров сети, агентов и распространяемой информации.

Для построения сети используются параметры: количество узлов сети, среднее количество соединений на узел, степень сети (согласно которой и распределяется количество связей на каждый узел). Необходимо использовать степенной закон распределения, так как он наиболее свойственен крупным сетям. Далее необходимо всем узлам дать различные характеристики, чтобы они могли взаимодействовать. Так как агенты в нашей сети будут распространять информацию, то у них должны быть следующие характеристики: степень убеждения и степень убеждаемости, текущий уровень убеждения (то, на сколько узел убежден в распространяемой информации), переломный момент уровня убеждения (степень убежденности после которой агент сам начинает распространять данную информацию), список связей узла (те узлы сети, с которыми он связан), а также само убеждение агента в каком-либо отношении (это свойственно для моды, т.е. когда агент может периодически изменять свое отношение к вопросу. Необходимо также выделить тех агентов, которые будут убеждены первоначально. Агент будет выбираться из его относительного количества связей, все остальные агенты будут выбираться случайным образом из связей агента. Также в процессе моделирования распространения информации, надо учитывать, что ценность различной информации – различна, и ценность какой-либо информации может меняться от времени, причем закон изменения ценности информации для каждой информации будет различным. В противовес информации выступают все взгляды противоречащие прямо или косвенно. Так как с точки зрения распространяемой информации все, что ей противоречит – плохо, принципиальной разницы происходит это прямо или косвенно нет, а также для упрощения разработки инструмента на текущем этапе. Всю информацию, которая так или иначе противоречит распространению нашей информации (или идеи) будем называть анти информация (или анти идея).

Когда модель построена, то ее нужно просчитать с заданным шагом, что и создаст из него искусственное общество. Одна итерация искусственного общества состоит из просчета ценности информации и анти информации, и просчета всех значений убежденности в идее. Когда просчитаны их значения для следующего хода, то для каждого агента модели по всем его связям вычисляются значения и из них происходит изменение текущего веса узла, по итогам решается если ее значение выше порога – то она распространяет идею, если ниже, то распространяет анти идею. Таким образом, у нас создается любое требуемое количество моделей, каждая из которых зависит от состояния модели на прошедший ход.

Построение модели это основа системы, но цель заключается в автоматическом построении модели и для этого нужен блок, который способен автоматически подобрать параметры. Поэтому генетический алгоритм будет заниматься тем, что подбирать параметры для сети, которые позволяют модели быть как можно реалистичнее.

Наш генетический алгоритм (далее в этой главе алгоритм) состоит из нескольких составных частей и его работа регулируется рядом параметров. Механизм работы алгоритма следующий:

1. Для работы алгоритма требуются следующие параметры: вероятность мутации, вероятность скрещивания, количество членов поколения, процент селекции, изменяемые параметры и границы их изменений, пользовательская статистика динамики распространения информации, условия остановки работы алгоритма.

2. Пользователь вводит статистику динамики распространения информации, интересующего процесса, с заданным шагом.

3. Далее указывает для тех параметров сети, которые ему заранее известны – информацию по ним.

4. Указывает алгоритму, какие параметры, соответственно, необходимо изменять, а также для каждого параметра – допустимые рамки изменения и запускает работу алгоритма.

5. После запуска алгоритм формирует общество из параметров, значения которых выбраны случайно из рамок заданных пользователем.

6. Когда общество моделей со случайными параметрами готово, то по каждой из них строится модель. То далее все модели необходимо оценить относительно целевой функции минимизации. Получаются значения динамики распространения ровно за то же самое количество ходов, которое было введено пользователем в динамике распространения информации. Вычитается разность на каждом шаге динамики полученной модели из реальной динамики. Суммируются все отклонения, и получается значение функции, которую алгоритму необходимо минимизировать, т.е. он будет стремиться построить такую модель, отклонения которой от целевой динамики распространения будут минимальны.

7. Когда для каждой модели получено значение целевой функции, алгоритм сортирует их в зависимости от значения целевой функции. Далее переходит к созданию нового поколения, которое должно содержать потомков лучших, чем в предыдущем поколении. Для этого выбираются два родителя случайным образом, в зависимости от минимальности полученного целевого значения функции (т.е. чем меньше значение целевой функции, тем больше вероятность выбора параметров именно этой модели, для передачи моделям следующего поколения).

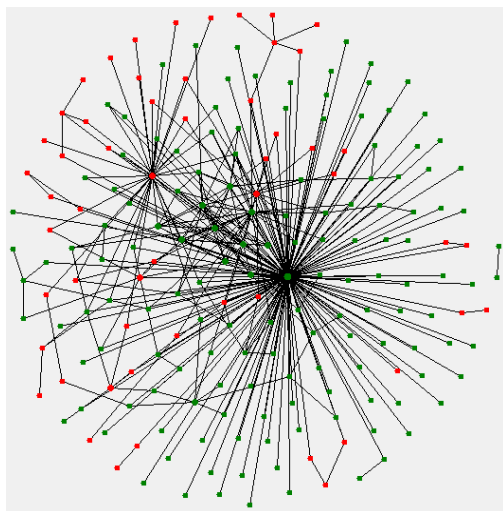
8. Для формирования нового поколения происходит следующий процесс: выбираются два родителя по принципу указанному в пункте 6. Далее из них формируются два потомка, либо прямым копированием, либо через операцию скрещивания, когда первому потомку берется одна часть параметров от одного родителя и другая от другого, а для второго потомка берутся противоположные части от одного и другого родителей. Скрещивание позволяет быстро закрепить лучшие параметры и распространить их на всех потомков. Потом производятся мутации над обоими потомками. Мутации двух вариантов, либо процентное отклонение от текущего значения всех параметров или выбор новых случайных значений для некоторых случайных параметров. За счет этих мутаций главным образом и достигается двойной эффект, одновременно покрытие всей области значений (выпадение новых случайных значений) вместе с плавным обхождением области текущих лучших полученных параметров для моделей. Именно мутация гарантирует нам в неопределенно долгом периоде работы алгоритма достижимость абсолютного совпадения динамик распространения информации модели и реальных значений.

9. Далее в работе алгоритма повторяются шаги 5-7, с единственным отличием в 6 пункте, что часть значений лучших членов предыдущего поколения переносится в новое поколение и далее сортировка и продолжение 6 пункта. Это делается для селекции лучшего значения. Так как от работы алгоритма необходимо получить именно одно значение, которое максимально близко к реальности, то селективность позволит быстрее достичь лучшего значения, хотя это несколько ограничит алгоритм в гибкости поиска лучшего решения, то есть несколько лучших моделей обладая наибольшими вероятностями на выбор в качестве родителей, заставят процесс создания нового поколения строить поколение, основываясь практически только на лучших полученных значениях.

10. И так алгоритм постоянно выполняет действия 5-8. Пока не сработает одно из условий завершения. Остановка алгоритма производится по достижению желаемой точности совпадения, прошествии определенного времени, количеству поколений или по требованию.

Мы описали два блока нашей системы, остался третий заключительный, блок вывода. Он занимается отображением и анализом получаемых моделей. В концептуальную версию необходимо встроить два основных средств-

ва, это визуальное представление работы модели и информационное табло, на которое будет выводиться информация в текстовом представлении. Необходимой к текстовому выводу информацией является ход модели, количество убежденных агентов и их процент от общего количества узлов, средняя убежденность, веса идеи и анти идеи текущего хода, среднее количество связей текущих убежденных узлов. Она должна позволить получить быстрое представление о ситуации в модели на определенный ход. Так же у пользователя есть доступ, для того, чтобы редактировать и просматривать значения всех параметров каждого агента. Визуальное отображение встречает несколько трудностей, это, во-первых, большое количество агентов (а система рассчитана именно на большие социальные сети), значит, требуется некоторое время на построение визуального представления. Во-вторых, учитывая то, как зададутся связи в сети нам заранее неизвестно, то и форма построения должна быть динамической, чтобы подходить под любую сеть. В качестве механизма построения и вывода будем выводить все узлы по кругу, отображая все связи между узлами, связей может быть достаточно, много, что приведет к тому, что будет невозможно разобрать связи между узлами. Для того, чтобы они лучше отображались, а также было лучше видно какие именно убеждаются узлы необходим механизм, который минимизирует расстояние связей между узлами, что заставит связанные узлы быть ближе друг к другу, что должно улучшить визуальное представление изображения, данный процесс условно назван оптимизация отображения. В итоге получается изображение подобное рис. 1.



*Рис. 1.* Представление построенной модели распространения информации, после оптимизации отображения

Концептуально описаны все блоки и общий принцип функционирования системы.

В заключение хочется сказать, что разработка методики и инструмента продолжается, но уже по текущим полученным результатам, инструмент показывает положительные результаты. Достигается 6 % точность построенных моделей при относительной случайно заданной динамики распространения информации. А значит, цель работы уже можно считать достигнутой, согласно полученной методике, возможно, строить агентные модели, для которых автоматически подбираются параметры сети и параметры агентов.

### **Список литературы:**

1. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. – 2010.
2. Макаров. Искусственные общества и будущее общественных наук. – 2009.
3. Макаров // Экономика и управление. – 2009. – С. 13-26.
4. Newman. The structure and function of complex networks. – 2003.
5. Дороговцев. Evolution of Networks. – 2003.