

## **Метод вычислительного интеллекта, заимствующий модель взаимодействия молекул внутри клетки**

**Н.А. Бабосюк, А.А. Каргин**

При автоматизации большинства процессов обработки информации одним из основных инструментов является моделирование. Глубоко и всесторонне изучены детерминированные и стохастические модели оценки и принятия решений [1,2], нечёткие модели [3]. Для всех трёх моделей характерно то, что они позволяют найти численное значение некоторого показателя или оценки в виде средних и среднеквадратичных отклонений для стохастических моделей и результат дефаззификации для нечётких моделей. Определение двух последних моделей базируется либо на статистических данных (функция распределения вероятностей, заданная на множестве значений параметра), либо на экспертных оценках (функция принадлежности, заданная на множестве значений параметра). Ключевым моментом здесь является то, что в наличии имеется информация для определения функции на множестве значений параметра.

Однако, это не всегда возможно. Известен другой класс слабо формализуемых задач, для которых решения строятся на экспертных знаниях, представление которых в виде лингвистических переменных на *предметной шкале* невозможно: они характеризуются более высоким уровнем неточности описания [4]. При решении этого класса задач используются методы искусственного интеллекта, в частности модели, основанные на знаниях и экспертных оценках [5], в том числе нейросетевое моделирование и нейронечеткие модели [6]. Однако, и эти технологии моделирования не позволяют браться за решение задачи, например, такой, как получить оценку влияния вербально выраженной стратегии на экономический показатель в численном виде. Например, устанавливается зависимость количества клиентов предприятия связи [7] или показателей

экологической безопасности [8] от маркетинговой стратегии, выраженной в виде вербального списка – «обеспечить доступные цены для потребителя услуг», «обеспечить конкурентоспособность цен», «формировать тарифы на услуги безотносительно к технологии» и другие. Эксперт способен оценить эту связь только качественно путём вербального описания *степени влияния* в виде *высокая, средняя* или *низкая*. Указанные вербальные оценки могут быть формализованы в виде лингвистических переменных, термы которой заданы на некотором универсальном множестве, например,  $[0,100\%]$ , но непосредственно формализация *степени влияния стратегий* на численный показатель (зависимости количества клиентов от стратегии) на сегодняшний день проблематична. Эти трудности особенно возрастают, когда приходится решать многомерную задачу типа

$$N = f_1(M^T, M^R, M^S, M^Q, O^S, O^R, O^T, P^O, P^{PR}, P^{NEW}, P^{CH}), \quad (1)$$

в которой необходимо учесть связи большого количества факторов с некоторой анализируемой переменной, причём переменная должна быть выражена численно, а знания о влиянии параметров относятся к выше сказанному классу. Выражение (1) приведено из работы [6] и отражает влияние различных маркетинговых ( $M^T, M^R, M^S, M^Q$ ), организационных ( $O^S, O^R, O^T$ ) и производственных ( $P^O, P^{PR}, P^{NEW}, P^{CH}$ ) стратегий и мероприятий на количество привлечённых клиентов. Знания о связях (1), установленные экспертом и выраженные, как было сказано ранее, вербально затем формализуются в виде множества независимых правил типа:

$$\begin{aligned} &\text{ЕСЛИ ДОСТУПНОСТЬ ЦЕНЫ есть } \textit{низкая} \\ &\text{И КАТЕГОРИЯ КЛИЕНТА есть } \textit{эконом} \\ &\text{ТО СТЕПЕНЬ ВЛИЯНИЯ есть } \textit{низкая}. \end{aligned} \quad (2)$$

В такой постановке задачу оценки критерия в (1), по-прежнему, затруднительно решать каким либо из известных традиционных методов, в том числе, путём полного перебора или имитационного моделирования. Известны методы вычислительного интеллекта, построенные на моделях

взаимодействия живых организмов, например, в колонии муравьёв, пчелином рое, стае птиц и другие[9]. В основном, они нацелены на решение задач оптимизации или поиска пути *при наличии явных чётких или нечётких моделей, связывающих целевой параметр с возмущающими факторами.* Поэтому требуется новый подход для решения задач определения численной зависимости некоторого критерия от факторов, о которых отсутствует какая-либо численная информация в виде значений или их оценок, или функций распределения.

В статье предлагается новый метод вычислительного интеллекта, формализующий модели взаимодействия молекул внутри клетки. Первое, что заимствует рассматриваемый метод из биологической модели взаимодействия молекул внутри клетки [10] – расчленение процесса решения задачи на отдельные последовательно связанные этапы: продукт взаимодействия элементов предшествующего этапа выступает в качестве элементов, участвующих во взаимодействии на следующем этапе как показано на рис.1. На каждом этапе взаимодействие протекает в «своём» поле, обуславливающим характер взаимодействия. Если возвратиться к примеру, приведенному выше, то решение задачи предлагаемым в статье методом можно расчленить на этапы в соответствии с количеством классов стратегий. На первом этапе рассматривается взаимодействие элементов в поле влияния стратегий маркетинга ( $M^T, M^R, M^S, M^Q$ ), на втором этапе – организационных мероприятий ( $O^S, O^R, O^T$ ), на третьем этапе – взаимодействие в поле производственно-технических мероприятий ( $P^O, P^{PR}, P^{NEW}, P^{CH}$ ) и на четвертом этапе моделируется взаимодействие в поле социально-экономических факторов ( $SCL, REV$ ).

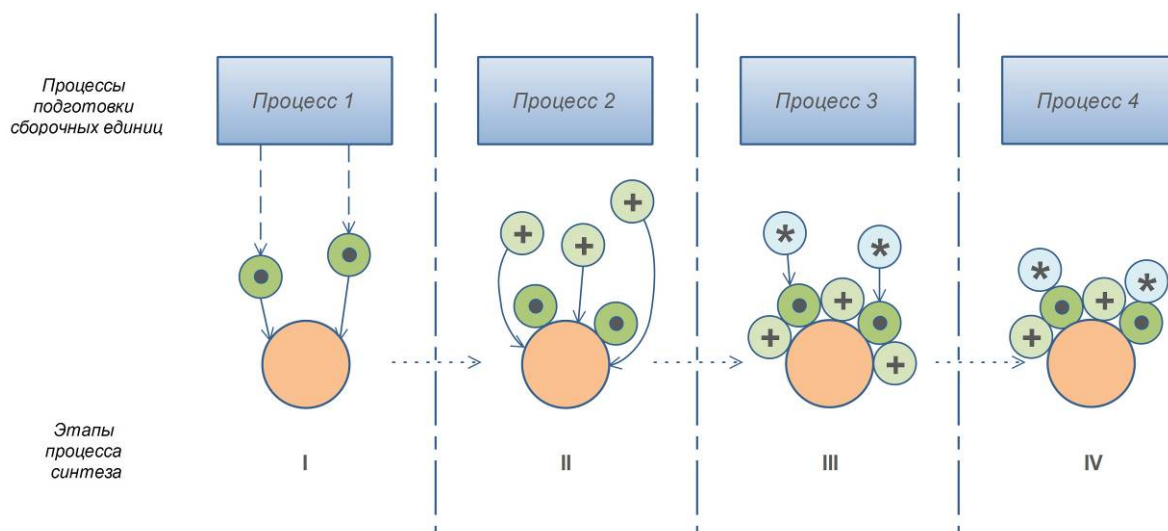


Рис. 1 – Процессы образования новых структур

На каждом этапе взаимодействий в биологической модели рассматривается множество разного типа молекул, находящихся в непрерывном движении [10]. Если взять отдельную молекулу в некотором пространстве клетки, то на произвольный момент времени она характеризуется координатами начала, ориентацией системы координат, в которой описана ее форма, и направлением движения. В биофизике [11] движение молекул моделируется как броуновское. Формальная модель броуновского движения [12] вкратце выглядит следующим образом. Моделируемое пространство (двумерное, трёхмерное или большей размерности) покрывается равномерной сеткой. Элементы могут находиться в узлах сетки. На  $i$ -м шаге моделирования по координатам элемента (где элемент сейчас находится) вычисляется местоположение элемента на следующем  $i+1$ -м шаге (где элемент окажется на следующем шаге моделирования). Принимается равновероятная модель переходов: элемент может перейти с равной вероятностью в один из восьми соседних узлов сетки.

Таким образом, в модели вычислительного интеллекта молекуле ставится в соответствие некоторый элемент, характеризующийся

множеством  $\Omega$  нечётких характеристик признаков  $\tilde{S}_i \in \Omega$  и координатами в пространстве  $\tilde{X}$ .

Второе, что заимствуется из взаимодействия молекул – это два типа взаимодействия: дальнее и ближнее. Вначале рассмотрим **модель дальнего взаимодействия**. Из биофизической модели [10] заимствуется механизм влияния *поля*, создаваемого одиночным элементом, группой элементов или внешними факторами, на движение элементов. В статье развивается модель источника поля, создаваемого одиночным элементом. Степень взаимодействия элементов в поле определяется значениями нечётких характеристик определённого набора свойств элементов. Поле имеет радиус действия и может иметь постоянные или изменяющиеся в пространстве и времени характеристики. По известным характеристикам элемента, имеющих отношение к полю, и координатам элемента-источника поля можно рассчитать характеристики взаимодействия в поле для произвольной точки пространства. Взаимодействие в поле рассматривается попарно (для пары элементов), результат взаимодействия одного элемента с другими, окружающими его, обладает свойством суперпозиции. Взаимодействие пары элементов, в зависимости от класса, которому они принадлежат, может проявляться по-разному: элементы притягиваются друг к другу, отталкиваются один от другого и ведут себя нейтрально. Формально результат взаимодействия двух элементов в поле отражается в изменении распределения вероятностей броуновского движения.

Пусть, после расчётов всех возможных попарных взаимодействий некоторого элемента получено множество из  $N$  данных

$$\{(\varphi_l, \nu_l)\}_{l=1}^N, \quad (3)$$

где  $\varphi_l \in [0, 360]$  – угол между прямой, проходящей через два взаимодействующих элемента, как показано на рис. 2, и направлением север-юг;

$v_l \in [-1,+1]$  – относительная «сила» взаимодействия (положительное значение – притягивание, отрицательное – отталкивание).

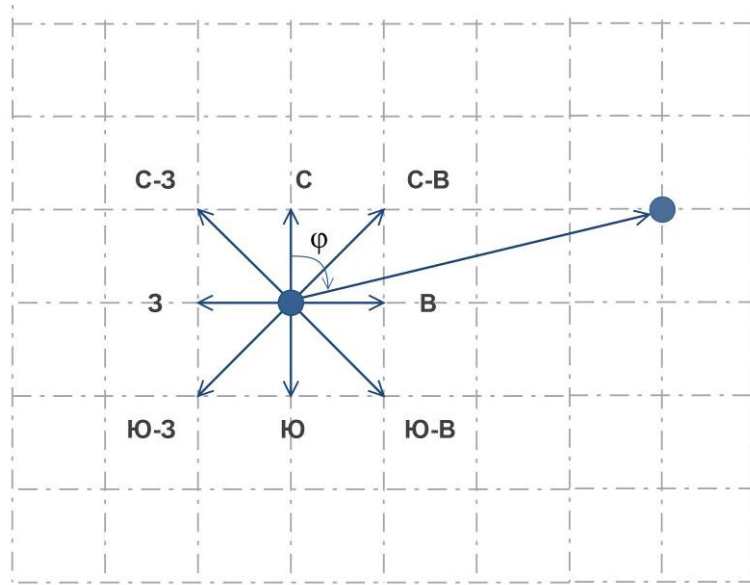


Рис. 2. – Пояснения к изменению распределения вероятностей броуновского движения

Относительная «сила» взаимодействия находится по значениям характеристик дальнего взаимодействия пары элементов. В простейшем случае модель дальнего взаимодействия включает три характеристики:  $r$  – радиус взаимодействия,  $v \in [-1,+1]$  – константная «сила» взаимодействия на протяжении всего радиуса и описатели классов элементов

$$v = \begin{cases} +v, & \text{если } \zeta \leq r \text{ и элементы принадлежат одному классу;} \\ -v, & \text{если } \zeta \leq r \text{ и элементы принадлежат разным классам;} \\ 0, & \text{если } \zeta > r \text{ и в остальных случаях,} \end{cases} \quad (4)$$

где  $\zeta$  – расстояние между взаимодействующими элементами.

Согласно принципу суперпозиции по результирующей «силе» взаимодействия по каждому из восьми направлений находится вероятность, например, для направления север по формуле

$$P(c) = \frac{\sum_{l=1}^N \mu_{\tilde{c}}(\varphi_l) \cdot v_l}{\sum_{l=1}^N (\mu_{\tilde{c}}(\varphi_l) + \mu_{\tilde{c-\beta}}(\varphi_l) + \mu_{\tilde{\beta}}(\varphi_l) + \mu_{\tilde{\beta-\gamma}}(\varphi_l) + \mu_{\tilde{\gamma}}(\varphi_l) + \mu_{\tilde{\gamma-\delta}}(\varphi_l) + \mu_{\tilde{\delta}}(\varphi_l)) \cdot v_l}, \quad (5)$$

где  $\mu_c(\varphi_l)$  – степень принадлежности угла  $\varphi_l$  нечёткому ограничению  $C$  (терму север лингвистической переменной НАПРАВЛЕНИЯ ВЗАИМОДЕЙСТВИЯ, показанной на рис. 3.) .

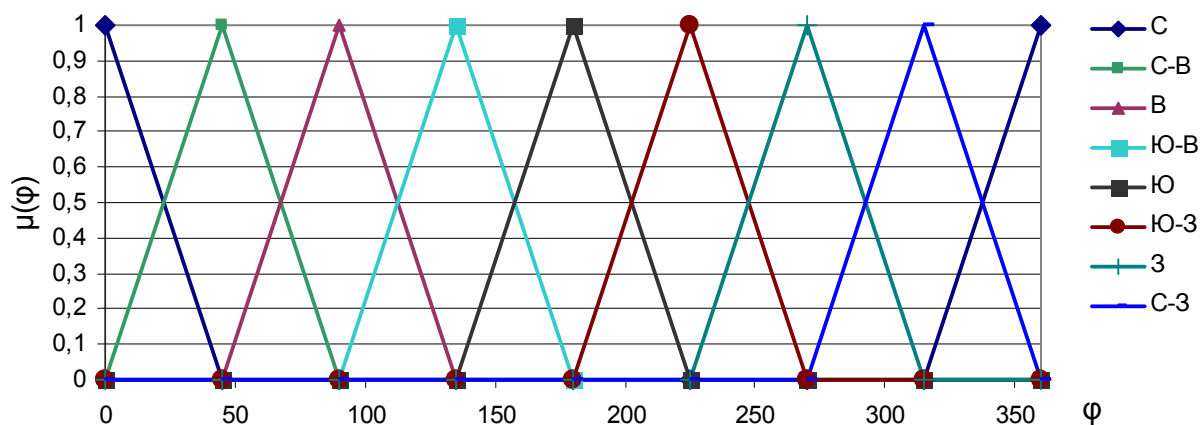


Рис. 3. – Функция принадлежности для лингвистической переменной *НАПРАВЛЕНИЯ ВЗАИМОДЕЙСТВИЯ*

При сближении двух или более элементов, когда они попадают в смежные узлы сетки, моделируется **ближнее взаимодействие**. Ближнее взаимодействие, наблюдаемое в природе, относят к одному из следующих двух видов [10]: 1) все элементы, участвующие во взаимодействии являются равноправными и играют одну роль – объекта; 2) элементы, участвующие во взаимодействии, играют разные роли – есть субъект и объекты.

Для обоих видов взаимодействия результат может быть либо изменение состояния, либо изменение структуры объектов (разрушение старой или появление новой). Результат зависит от характеристик, в том числе и от состояния (системные свойства) элементов, которые взаимодействуют. Если говорить о молекулярном уровне внутри клетки [10], то во взаимодействии фермента с субстратами первый играет роль субъекта, а последние – объектов, во взаимодействии кофермента с ферментом также первый играет роль субъекта, а последний – объекта. Ближнее взаимодействие является избирательным. Все элементы обладают высокой

специфичностью (уникальностью), что обеспечивает избирательность взаимодействия: форма активного центра фермента в точности должна соответствовать форме определённой части молекулы субстрата для того, чтобы началось взаимодействие.

В модели вычислительного интеллекта такая избирательность представлена как ситуационное взаимодействие по правилу

*ЕСЛИ* значения конкретного подмножества характеристик элемента-объекта нечётко соответствуют значениям такого же подмножества характеристик элемента-субъекта

*ТО* реакция (характеристики реакции зависят от нечёткого соответствия).

На основании аппарата нечёткого фактора уверенности [13] введен формализм нечёткого соответствия  $\rho(S_i, S_j)$  двух структур, представленных нечёткими характеристиками  $S_i$  и  $S_j$ , а на этой основе формализована нечёткая характеристика реакции

$$\underset{\sim}{R}\{\underset{\sim}{[R]}_t = F(\underset{\sim}{[\Theta}^{x_1}]_t, \underset{\sim}{[\Theta}^{x_2}]_t, \dots, \underset{\sim}{[\Theta}^{x_n}]_t, \{\rho(\underset{\sim}{S}_i, \hat{\underset{\sim}{S}}_i), \forall i \in \Omega\}, \quad (6)$$

где  $\underset{\sim}{[\Theta}^{x_j}]_t$  – нечёткая характеристика состояния  $x_j$  элемента-субъекта;  $\hat{\underset{\sim}{S}}_i$  – прототип  $i$ -го свойства элемента-субъекта;  $\underset{\sim}{S}_i$  – нечёткая характеристика  $i$ -го свойства элемента-объекта, участвующего во взаимодействии;  $\Omega$  – множество свойств, определяющих индивидуальное избирательное взаимодействие элемента-субъекта с элементом-объектом.

Следующая компонента модели, заимствованная у клетки, – управляемое взаимодействие. Реакции, протекаемые с участием ферментов, могут быть управляемыми ко-факторами – молекулами, присутствие которых влияет на проявление каталитической активности ферментов. В статье используется модель управляемого ближнего взаимодействия, формализованная следующим выражением [10]:



$$\begin{aligned} \{[R]_{\tilde{t}} = F([\Theta^{x_2}]_{\tilde{t}}, \rho([S^{sub}]_{\tilde{t}}, [S^{ob}]_{\tilde{t}})), \\ \{[\Theta^{x_2}]_{\tilde{t}+1} = T_1([\Theta^{x_1}]_{\tilde{t}}, \rho(\hat{In}, [In]_{\tilde{t}})). \end{aligned} \quad (7)$$

Рассматривается случай, когда элемент-субъект имеет два состояния  $x_1$  и  $x_2$ , прототип  $S^{sub}$ , специфицирующий взаимодействие с объектом, определён на одном свойстве, а, значит, сопоставляется с нечёткой характеристикой этого одного свойства  $S^{ob}$  объекта. В (7) первое выражение устанавливает зависимость между нечёткой характеристикой процесса взаимодействия (протекания реакции  $R$  с целью получения нового элемента-комплекса) и двумя нечёткими характеристиками: состояниями  $x_2$  и близости характеристик взаимодействующих субъекта  $S^{sub}$  и объекта  $S^{ob}$ . Второе выражение в (7) связывает состояния элемента-субъекта с нечёткой характеристикой управляемого входа. Чем ближе ситуация, соответствующая прототипу управляемого воздействия  $\rho(\hat{In}, [In]_{\tilde{t}})$ , тем больше уверенность в том, что элемент-субъект находится в состоянии  $x_2$ , готовом к взаимодействию с элементом-объектом.

Независимо от типа модели ближнего взаимодействия результат взаимодействия – образование новой структуры – определяется по характеристике «сила внутренней связи образованного комплекса»  $\alpha_R$ , параметр уверенности нечёткого фактора уверенности  $R$  в (6) и (7).

Если  $\alpha_R > \varepsilon$ ,  $\varepsilon > 0$ , то ближнее взаимодействие считается завершённым успешно (образовалась новая структура в виде элемента-комплекса) и этот комплекс «аннигилирует» в следующее пространство (удаляется из пространства текущего этапа моделирования и появляется в пространстве следующего этапа).

Если  $\alpha_R < \varepsilon$ ,  $\varepsilon > 0$ , то считается, что взаимодействие не произошло и элементы, принимающие участие во взаимодействии случайным образом размещаются в текущем пространстве.

Графически для наглядности избирательного характера ближнего взаимодействия элементы изображаются [6] в виде «гребёнки», зубцы которой символизируют свойства, соответствие которых требуется для ближнего взаимодействия – образования новой структуры комплекса, как показано на рис. 4.

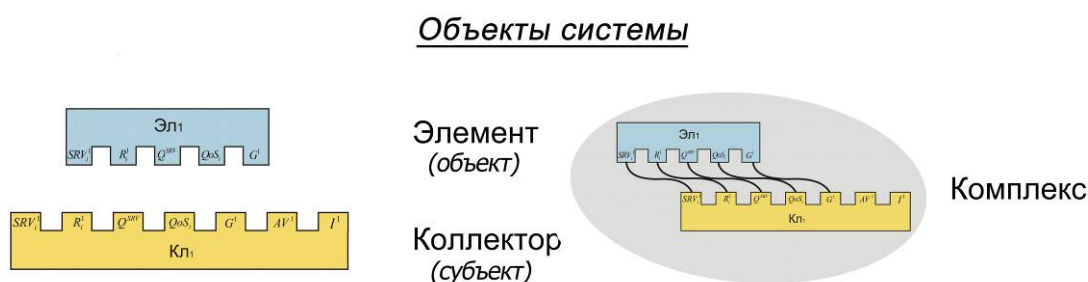


Рис.4. – Схематическое представление компонентов взаимодействия

Для решения прикладной задачи вычислительным методом требуется настройка параметров модели на предметную область. Для каждого этапа настраиваются параметры модели взаимодействия, а именно, элементы взаимодействия, характеристики дальнего и ближнего взаимодействия, начальные условия моделирования (количества элементов каждого типа, структура пространства моделирования, начальные распределения элементов в пространстве и другие).

Так при моделировании предметной области, затронутые в начале статьи, создаются следующие виды структур:

- Элемент – клиент в поле маркетинговой деятельности.
- Коллектор – предприятие в поле маркетинговой деятельности;
- Комплекс – структура, объединяющая коллектор-предприятие и один элемент-клиент.

Элементу-клиенту приписаны следующие признаки: категория клиента (эконом (э), оптималист (о), премиум (п)); перечень основных услуг

(широкополосный доступ: кабельная связь (кс), цифровая абонентская линия DSL (цл), WiMAX (wm), Metro Wi-fi (mw), RadioEthernet (re), спутниковая связь (сс); коммутируемый доступ (кд)); перечень дополнительных услуг (регистрация доменных имен (ри), выделение IP-адресов (ва), электронная почта (эп), решения для бизнеса (рб): разработка проекта, установка оборудования и программного обеспечения, инсталляция технического решения); граничные тарифы на соответствующие услуги.

Для каждой категории элемент-клиента на основании обработки статистических данных известен его портрет в виде:

- наборов нечётких факторов уверенности:

$$\text{категория клиента} = \{\Theta_{\sim c}, \Theta_{\sim kn}, \Theta_{\sim d}\}; \quad (8)$$

$$\text{основные услуги} = \{\Theta_{\sim kc}, \Theta_{\sim цл}, \Theta_{\sim wm}, \Theta_{\sim re}, \Theta_{\sim cc}, \Theta_{\sim kd}\}; \quad (9)$$

$$\text{дополнительные услуги} = \{\Theta_{\sim ри}, \Theta_{\sim ва}, \Theta_{\sim эп}, \Theta_{\sim рб}\}. \quad (10)$$

- наборов нечётких множеств, заданных на шкале тарифов:

$$\text{граничные тарифы} = \{T_{\sim kc}, T_{\sim цл}, T_{\sim wm}, T_{\sim re}, T_{\sim cc}, T_{\sim kd}\} \quad (11)$$

Признаки (8)-(11) используются алгоритмом ближнего взаимодействия. Расчеты характеристик дальнего взаимодействия осуществлены на основе лингвистической переменной «СТЕПЕНЬ ПРИВЛЕКАТЕЛЬНОСТИ», значение которой находится путем нечеткого вывода по правилам базы знаний экспертов, представленных в виде (2).

При моделировании взаимодействия с целью определения рациональных значений управляющих переменных стратегий, могут варьироваться все указанные признаки. Кроме этого, к исходным варьируемым данным при моделировании относятся численные значения тарифов и объёма дисконта. На основании этих входных данных путём нечёткого вывода по алгоритму Мамдани [3,13] по правилам типа (2) с использованием определений нечётких множеств, являющихся признаками коллектора и элемента, находится численное значение привлекательности

предприятия. На основании этого значения вычисляются параметры  $V_i$  и  $r$  дальнего взаимодействия в формулах (3)-(5).

Метод вычислительного интеллекта, заимствующий модель взаимодействия молекул в клетке апробирован в качестве инструментария в задачах бизнес-планирования, в частности, составление бизнес-плана предприятия связи: ЧАО «ДОРИС» и ПАО «Укртелеком». С помощью метода проводилось исследование влияния управляемых параметров стратегий развития предприятия (маркетинговой, производственной, организационной, экономико-социальной деятельности) на численные экономические показатели. На рис.5 приведен пример одного из результатов исследований – влияние таких параметров как тарифы, скидки, услуги предприятия стратегий «Обеспечения доступности цен», «Обеспечение конкурентоспособных тарифов» и «Формирование тарифов с учетом уровней дополнительного сервиса» маркетинговой деятельности. Нечеткая формализация рассуждений, использующих вербальные описания в виде «Категория клиента», «Доступность цены» и других позволила найти значение лингвистической переменной «СТЕПЕНЬ ПРИВЛЕКАТЕЛЬНОСТИ» на основании которой найден радиус дальнего взаимодействия. Из рисунка явно видна степень влияния отдельных маркетинговых решений опосредованных через радиус дальнего взаимодействия на количество привлеченных новых клиентов предприятия.

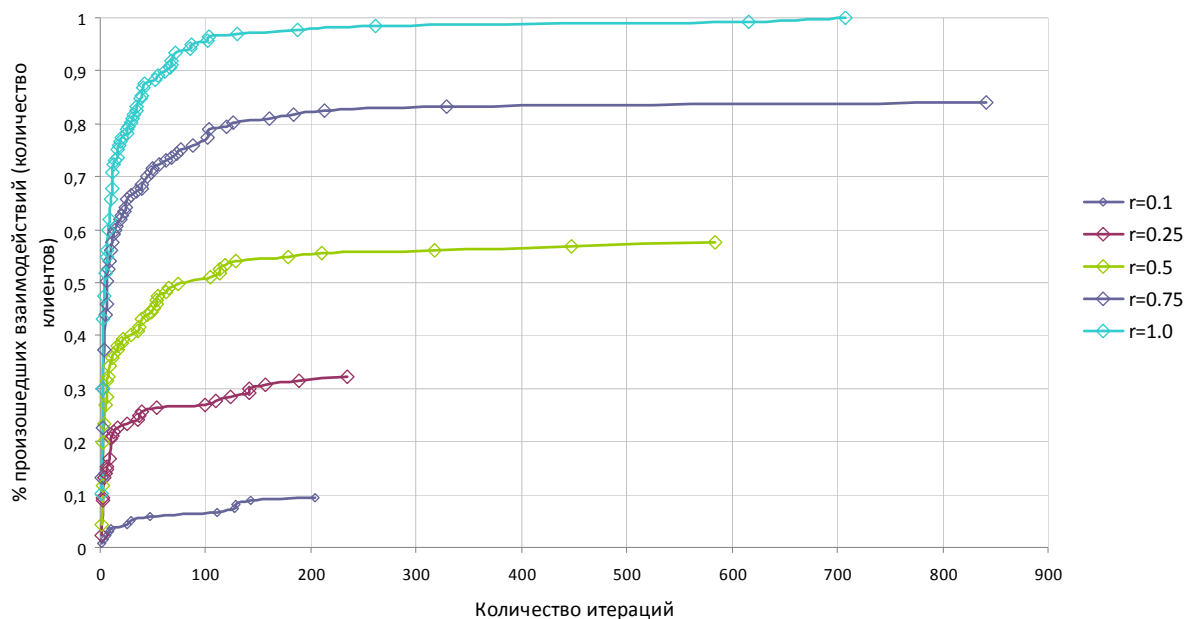


Рис. 5. – Количество произошедших взаимодействий для различных радиусов  $r$ , количества элементов 100, количества коллекторов 10000

### Литература:

1. Медведев Н.А. Пучкова Т.И. Бизнес-планирование [Текст] / Н.А. Медведев, Т.И. Пучкова – М.: Изд-во МГУЛ, 1998. – 147 с.
2. Бабосюк А.Л. Ворхалев П.Ю., Петров Л.П. Исследования возможностей применения математико-статистических методов моделирования для прогнозирования тарифных доходов предприятия связи [Текст] // журнал «Зв'язок», 2002. – №2, – С.45-47.
3. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH [Текст] / А.В. Леоненков – СПб: БХВ-Петербург, 2003. – 736 с.: ил.
4. Коваленко И.И. Метод экспертного оценивания сценариев: [Текст] / И.И. Коваленко, А.В. Швец. – Николаев: Изд-во ЧГУ им. Петра Могилы, 2012. – 156 с.
5. Проталинский О.М., Ажмухамедов И.М. Системный анализ и моделирование слабо структурированных и плохо формализуемых процессов в социотехнических системах [Электронный ресурс] // «Инженерный вестник Дона», 2012, №3. – Режим доступа:

<http://ivdon.ru/magazine/archive/n3y2012/916> (доступ свободный) – Загл. с экрана. – Яз. рус.

6. Терехов В.А. Ефимов Д.В., Тюкин И.Ю. Нейросетевые системы управления [Текст] / В.А. Терехов, Д.В. Ефимов, И.Ю. Тюкин – М.: Высш. шк., 2002., – 183 с.

7. Каргин А.А., Бабосюк Н.А. Об использовании в задачах бизнес-планирования метода вычислительного интеллекта, основанного на модели взаимодействия молекул внутри клетки. [Текст] // Вестник Херсонского национального технического университета, 2013. – №1(46), – С 64-66.

8. Самсонова М.Л. Учет экологических факторов при разработке инновационного бизнес-плана [Электронный ресурс] // «Инженерный вестник Дона», 2009, №1. – Режим доступа: <http://www.ivdon.ru/magazine/archive/n4p2y2012/1424> (доступ свободный) – Загл. с экрана. – Яз. рус.

9. Yan Meng, Yaochu Jin. Bio-Inspired Self-Organizing Robotic Systems [Текст] / Yan Meng, Yaochu Jin – Springer-Verlag Berlin And Heidelberg GmbH Co. Kg., Germany, 2011, – 287с.

10. Фаллер Д.М., Шилде Д. Молекулярная биология клетки. Руководство для врачей. Пер. с англ. [Текст] / Д.М. Фаллер, Д.Шилде – М.: Издательство БИНОМ. – 2011. – 256 с., ил.

11. Джаксон М.Б. Молекулярная и клеточная биофизика. [Текст] / М.Б.Джаксон, – М. : Мир; БИНОМ. Лаборатория знаний, 2009. – 551с.: ил.

12. Kramers, H.A. Brownian motion in a field of force. [Текст] // Physica, 1940, – №7, – С.284-304.

13. Каргин А.А. Введение в интеллектуальные машины. Книга 1. Интеллектуальные регуляторы [Текст]: Монография / А.А.Каргин. – Донецк: Норд-Пресс, ДонНУ, 2010. – 526 с.

14. Пегат А. Нечеткое моделирование и управление (Адаптивные и интеллектуальные системы) Пер.с англ. [Текст] / А. Пегат – М. : БИНОМ. Лаборатория знаний, 2009. – 798 с. :ил.