

ЗАДАЧИ АНАЛИЗА И ПРОГНОЗА ПРИ УПРАВЛЕНИИ ЦЕЛЕВОЙ АУДИТОРИЕЙ В МАРКЕТИНГЕ¹

Агиева М. Т.²

(Ингушский государственный университет, Назрань)

Бабичева Ю. В.³, Окулист Н. М.⁴, Угольницкий Г. А.⁵

(Южный федеральный университет, Ростов-на-Дону)

Рассмотрены задачи анализа и прогноза на сетях применительно к маркетингу. Приведен краткий обзор моделей влияния и управления в социальных сетях. К указанным задачам относятся выявление сильных подгрупп и спутников, вычисление количественных характеристик сети, определение финальных мнений членов целевой аудитории исходя из их начальных мнений с учетом взаимного влияния. Для решения задачи прогноза разработан оригинальный алгоритм, реализованный на языке программирования R. Алгоритм описан в деталях и охватывает частный случай сильно связного орграфа и общий случай. Проведена оценка сложности этого алгоритма. Другие поставленные задачи могут быть решены стандартными средствами этого языка, что проиллюстрировано на тестовых примерах вместе с результатами авторского алгоритма. Тестовые примеры используют орграфы влияний с 10-15 и более чем со ста вершинами. Вычислены значения различных мер близости для орграфа: центральность по степени, центральность по близости, центральность по посредничеству, дана их интерпретация. Решение задач анализа и прогноза существенно облегчает задачу управления мнениями целевой аудитории в маркетинге. Намечены пути дальнейших исследований по построению и анализу моделей управления мнениями в маркетинге.

Ключевые слова: маркетинг, модели анализа и прогноза на сетях, управление целевой аудиторией.

1. Введение

Модель влияния в социальной группе впервые рассмотрена в работах [20, 24] и более детально изучена в [18]. Впоследствии

¹ Работа выполнена при финансовой поддержке РФФ, проект №17-19-01038.

² Мовлатхан Тугановна Агиева, к.т.н., доцент (agieva25@mail.ru).

³ Юлия Витальевна Бабичева, бакалавр прикладной математики (yulya.babicheva1012@yandex.ru).

⁴ Наталья Менахимовна Окулист, бакалавр прикладной математики (nata-0k@rambler.ru).

⁵ Геннадий Анатольевич Угольницкий, д.ф.-м.н., профессор (ougoln@mail.ru).

рассматривались многочисленные обобщения и уточнения этой модели: случай меняющегося во времени взаимовлияния – в работах [23, 25, 28, 29], условия сходимости мнений – в [14, 16, 19], скорость сходимости – в [18, 23], условия единственности финального мнения – в [23, 25].

Одной из самых подробных монографий по моделированию сетей является [26]; см. также статью [27]. Применение сетевых моделей к анализу политических процессов описано в монографии [3]. Стохастические модели социальных сетей представлены в статье [4]; это направление впоследствии развито в монографии [5], посвященной математическим моделям коллективного поведения. Монография [7] посвящена моделям сетевой экспертизы. Вопросы коммуникации и координации в социальных сетях анализируются в [17]. Моделированию «сарафанного радио» посвящены работы [15, 21, 22]. Анализ относительного влияния вершин сети дан в [30]. Системный подход к сетям представлен в [31]. Сетевые модели социального влияния описаны также в [32, 33].

Подробный анализ моделей влияния представлен в монографии [6]. Авторы предлагают классификацию таких моделей. Приведен анализ динамики мнений для ряда характерных модельных примеров.

В статье [1] приведена постановка задач управления на сетях применительно к маркетингу с соответствующей интерпретацией элементов, в статье [2] – постановка задач анализа целевой маркетинговой аудитории.

Оставшаяся часть статьи организована следующим образом. Во втором разделе приводятся постановки задач анализа и синтеза на сетях и описывается базовый тестовый пример. В третьем разделе описываются ещё два примера большей размерности, даётся подробное описание количественных характеристик сети применительно к маркетингу и результаты вычисления их значений и выявления структуры целевой аудитории для тестовых примеров с помощью средств языка программирования R [8]. В четвёртом разделе представлен оригинальный алгоритм вычисления финальных мнений членов целевой аудитории и его реализация на языке R с расчётами по тестовым

примерам. В Заключении подводятся итоги и намечаются планы дальнейших исследований.

2. Постановки задач и базовый тестовый пример

Базовая модель влияния в социальной сети представляет собой взвешенный ориентированный граф, вершины которого соответствуют членам социальной группы, а дуги описывают их взаимное влияние. Каждой вершине приписывается вещественное значение (мнение члена группы), которое может меняться со временем, а каждой дуге – вещественное число (вес), характеризующее степень влияния одного члена группы на другого (или, что то же самое, степень доверия второго к первому).

Таким образом, в модели $X = \{x_1, \dots, x_n\}$ – множество базовых агентов; $x_i \rightarrow u_i(t)$, $i = 1, \dots, n$, – мнение базового агента; $u^0 = (u_1^0, \dots, u_n^0)$ – вектор начальных мнений базовых агентов; a_{ij} – коэффициент влияния i -го базового агента на мнение j -го базового агента; $A = \|a_{ij}\|$ – матрица влияний (задает множество дуг социально-сетевой модели). Динамика мнений задается уравнением

$$(1) \quad u_j(t+1) = \sum_{i=1}^n a_{ij} u_i(t), \quad j=1, \dots, n;$$

или в векторной форме

$$u(t+1) = A^T u(t), \quad u(0) = u^0, \quad t=0, 1, \dots, T-1.$$

Показано, что все члены каждой i -й сильной подгруппы (невырожденной сильной компоненты сети) приходят к единому финальному мнению, определяемому формулой

$$(2) \quad u_i^\infty = \sum_{k=1}^{n_i} w_k^i u_k^{i0},$$

где w_k^i – компонента стационарного вектора для цепи Маркова с переходной матрицей A^T , n_i – число членов i -й сильной подгруппы. Финальные мнения остальных агентов, не входящих в сильные подгруппы («спутников»), вычисляются как

$$(3) \quad u_j^\infty = \sum_{i=1}^r b_{ji} u_i^\infty,$$

где b_{ji} – вероятность перехода агента j в сильную подгруппу i как эргодическое множество, r – общее число сильных подгрупп.

Кроме того, для заданной сети можно вычислять различные количественные показатели: центральности, престижа и др. [9, 10].

Таким образом, с помощью моделей влияния на сетях можно решать задачи анализа и прогноза. К задачам анализа относятся выделение сильных подгрупп и спутников, а также вычисление количественных характеристик сети. Задачи прогноза позволяют предсказывать динамику мнений членов социальной группы с учетом ее структуры, выявленной при анализе.

Применительно к маркетингу оргграф влияний определяет взаимовлияние членов целевой аудитории некоторой фирмы. Участники целевой аудитории имеют начальные мнения, например, по поводу суммы денег, которую они готовы потратить на приобретение товаров или услуг данной фирмы в течение некоторого периода времени. В силу взаимодействия внутри целевой аудитории эти мнения могут меняться со временем. Выделение сильных подгрупп позволяет определить тех людей (группы, организации), на которых следует оказывать маркетинговое воздействие для получения желаемых финальных мнений (воздействовать на спутников бессмысленно). Количественные характеристики сети дают дополнительное, более детальное описание целевой аудитории.

На данном этапе исследование проводилось для следующего базового тестового примера. Пусть имеется небольшая компания, в которой десять сотрудников и руководитель. У руководителя скоро день рождения, и принято решение собрать деньги ему на подарок. Каждый сотрудник имеет своё мнение по поводу суммы взноса – это начальные мнения. Также в компании имеются социальные связи – некоторые сотрудники дружат между собой и прислушиваются к мнению друг друга. Матрица влияний представлена на следующей странице.

Пусть начальные мнения по поводу суммы подарка в рублях следующие: (500, 900, 500, 400, 400, 500, 800, 500, 600, 400, 100, 200, 200, 300, 500).

Тогда получаем сеть, изображенную на рис. 1.

Решим для этого тестового примера задачи выделения сильных подгрупп и спутников, вычисления характеристик сети и оценки финальных мнений членов группы.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{12} & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{12} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{12} & 0 & 0 & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{3} \end{pmatrix}$$

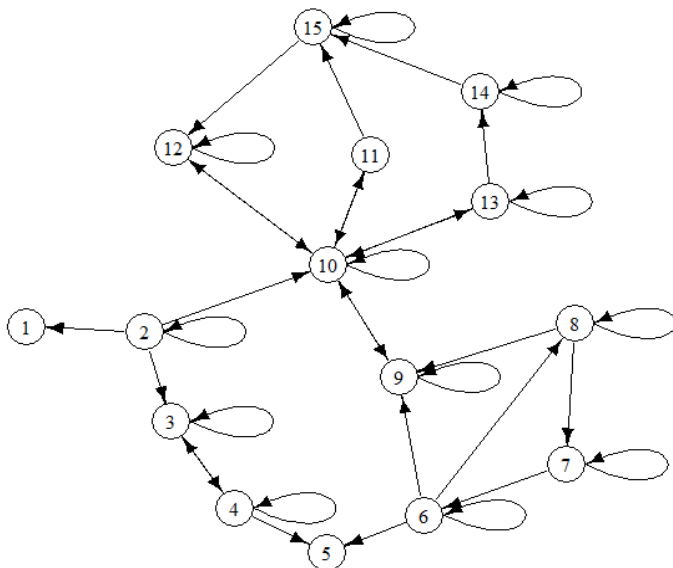


Рис. 1. Граф для тестового примера

3. Задачи анализа на сетях

Под социальной сетью на качественном уровне понимается социальная структура, состоящая из множества агентов (семьи, индивиды, организации) и определенного на нем множества отношений (совокупность связей между агентами). Математически сеть представляет собой ориентированный граф. Техника социометрии (представление социальных групп в виде графа) была впервые предложена в работах Я. Морено [11]. Термин «социальная сеть» был введен в 1954 г. социологом Джеймсом Барнсом [13].

Свойства социальных сетей: 1) наличие собственных мнений агентов; 2) возможность членов социальной сети влиять на мнения друг друга; 3) влиятельность и авторитет одних агентов для других; 4) различный уровень осведомленности агентов; 5) косвенное влияние агентов, например, «друг моего друга – мой друг»; 6) наличие лидеров, кумиров, агентов с максимальным «влиянием», формализация индексов влияния;

7) существование порога чувствительности к изменению мнения окружающих; 8) объединение агентов группы «по интересам»; 9) наличие специфических социальных норм; 10) учет факторов «социальной корреляции»; 11) существование внешних факторов влияния (реклама, маркетинговые акции) и, соответственно, внешних агентов (средства массовой информации, производители товаров), и т.д.

Сам по себе анализ социальных сетей (social network analysis, SNA) представляет собой направление, которое занимается описанием и анализом возникающих в ходе социального взаимодействия и коммуникации связей (сетей) различной плотности и интенсивности. Различные меры SNA используются для представления взаимодействия между участниками, учёта сильных или слабых связей, определения ключевых/центральных игроков и подгрупп в сети.

Для анализа социальной сети существует множество характеристик и метрик. Для их вычисления и представления в работе используется язык R, предназначенный для статистической обработки данных и работы с графикой и представляющий собой свободную программную среду вычислений. Язык R широко применяется как статистическое программное обеспечение для анализа данных и фактически стал стандартом для статистических программ [8].

В частности, задача выделения сильных компонент и спутников может быть решена стандартными встроенными средствами языка R. Для тестового примера результат будет иметь следующий вид (рис. 2).

Граф отображает коллектив, который договаривается о сумме на подарок начальнику. Вершина 10 оказывает сильное влияние на вершины 11–15, которые представляют собой: 10 – аналитик, 11–15 – программисты, т.е. люди из одной финансовой категории, с одинаковыми интересами, которые общаются только между собой и с аналитиком как связующим звеном между ними и всем остальным коллективом. Вершины 6–8 (сильная подгруппа) – отдел маркетинга и продаж. Вершина 9 – менеджер проекта, 2 – отдел внедрения.

Остановимся подробнее на вычислении метрик социальной сети на примере графа на рис. 1.

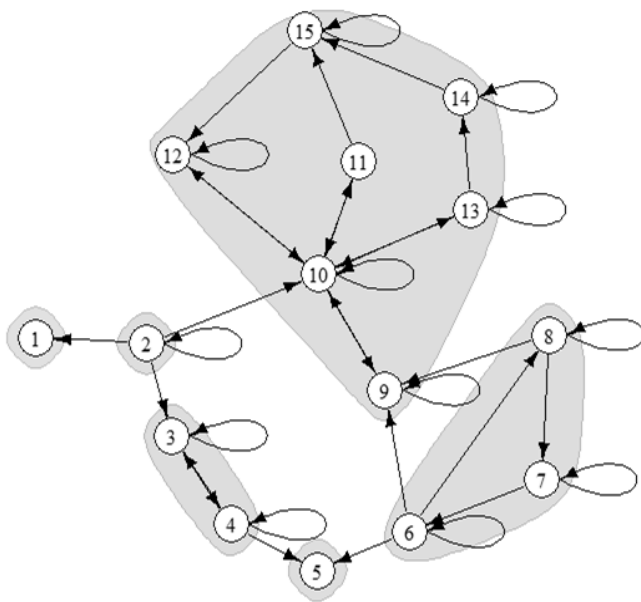


Рис. 2. Сильные компоненты и спутники

Степень (Degree). В теории графов степень вершины графа – это число рёбер, проходящих через вершину. Чем больше связей у данной вершины, тем выше степень.

Близость (Closeness). Эта величина показывает скорость передачи информации от одного узла (вершины) к другим.

Плотность (Density) или коэффициент кластеризации. Плотность измеряет степень, в которой узлы сети соединяются друг с другом [2].

3.1. МЕТРИКИ ЦЕНТРАЛЬНОСТИ (CENTRALITY METRICS)

Метрики центральности – количественная оценка конкретного лица в социальной сети. Мера центральности описывает положение конкретного узла по сравнению с другими узлами.

3.1.1 СТЕПЕНЬ ЦЕНТРАЛЬНОСТИ /ЦЕНТРАЛЬНОСТЬ ПО СТЕПЕНИ (DEGREE CENTRALITY)

Степень центральности является одной из основных метрик, определяет количество связей агента, то есть число зна-

комств в конкретной группе лиц. Узел с более высоким значением степени центральности часто рассматривается как центр и активный субъект в сети [2].

Иногда возникает необходимость сравнивать степень центральности разных графов, тогда нормированную величину вычисляют по формуле

$$(4) \quad D = \frac{\sum_{i=1}^N (D^* - D_i)}{(N-1)(N-2)},$$

где D^* – максимальная степень центральности узла из всех рассчитанных.

Вершины с высоким показателем степени центральности имеют большое влияние на других объектов сети, т.е. они могут быть информаторами либо знаменитостями, быстро распространяющими информацию. Степень центральности иллюстрирует степень общительности индивида или группы. Также различают входящую (in-degree) и исходящую (out-degree) центральность. Входящая указывает число связей, направленных к узлу, а исходящая – число связей, направленных от узла. Например, если сайт имеет множество внешних ссылок на другие сайты, как Яндекс или Google, это – власть, а если много людей размещают ссылки на сайт, это – известность.

Вычислим некоторые макропараметры сети.

Плотность сети: 0,1714286. Значение плотности не очень велико, что свойственно данному коллективу.

Взаимность: 0,4166667. В большинстве случаев связи оказываются не взаимными.

Транзитивность: 0,1666667. Очень низкая транзитивность объясняется тем, что коллеги из разных отделов редко общаются друг с другом.

Найдем степень центральности для примера на рис. 1 (рис. 3).

Степени центральности вершин графа для каждой вершины:

$D_1 = 0,067$; $D_2 = 0,333$; $D_3 = 0,333$; $D_4 = 0,333$; $D_5 = 0,133$;
 $D_6 = 0,400$; $D_7 = 0,267$; $D_8 = 0,333$; $D_9 = 0,400$; $D_{10} = 0,733$;
 $D_{11} = 0,200$; $D_{12} = 0,333$; $D_{13} = 0,333$; $D_{14} = 0,267$; $D_{15} = 0,333$.

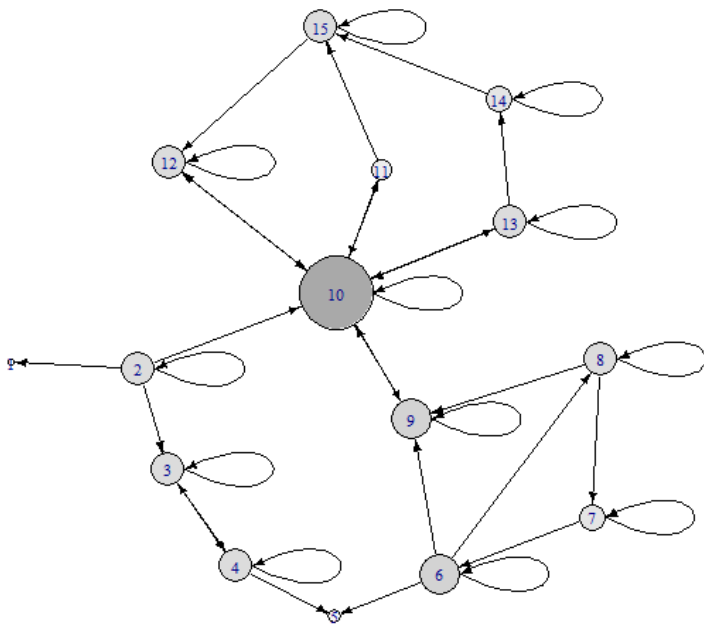


Рис. 3. График центральности по степени

Видим, что самый большой показатель центральности по степени у 10-й вершины, причём она обладает высокой и входящей, и исходящей центральностью, а это значит, что узел – «знаменитость» и «власть». Значит, за таким человеком будет следить много людей, он может быстро распространять информацию, а также быстро ее получать.

Этот узел можно использовать для продвижения продукта.

3.1.2. ЦЕНТРАЛЬНОСТЬ ПО БЛИЗОСТИ (CLOSENESS CENTRALITY)

Данная характеристика задаёт скорость распространения информации по сети. Если узел близок ко всем остальным узлам сети, то центральность по близости будет высокой, и чем дальше находится узел, тем меньше этот показатель [2].

Высоким уровнем центральности по близости обладают те узлы, которые имеют множество связей, при этом их должно

быть достаточно и у его соседей. Следовательно, такие узлы получают возможность доступа к большому количеству участников сети, и могут влиять на них.

Соответственно, мера достигнет своего максимума для заданной сети, когда субъект имеет доступ ко всем остальным вершинам в сети, и своего минимума, когда субъект изолирован. С увеличением сети эта величина будет уменьшаться. Это интуитивно понятно, так как чем больше компания, тем тяжелее поддерживать связь со всеми ее членами.

Рассчитаем значение центральности по близости для вершин нашего графа рис. 1 (рис. 4): $C_1 = 0,005$; $C_2 = 0,015$; $C_3 = 0,005$; $C_4 = 0,005$; $C_5 = 0,005$; $C_6 = 0,012$; $C_7 = 0,011$; $C_8 = 0,012$; $C_9 = 0,008$; $C_{10} = 0,008$; $C_{11} = 0,008$; $C_{12} = 0,008$; $C_{13} = 0,008$; $C_{14} = 0,007$; $C_{15} = 0,007$.

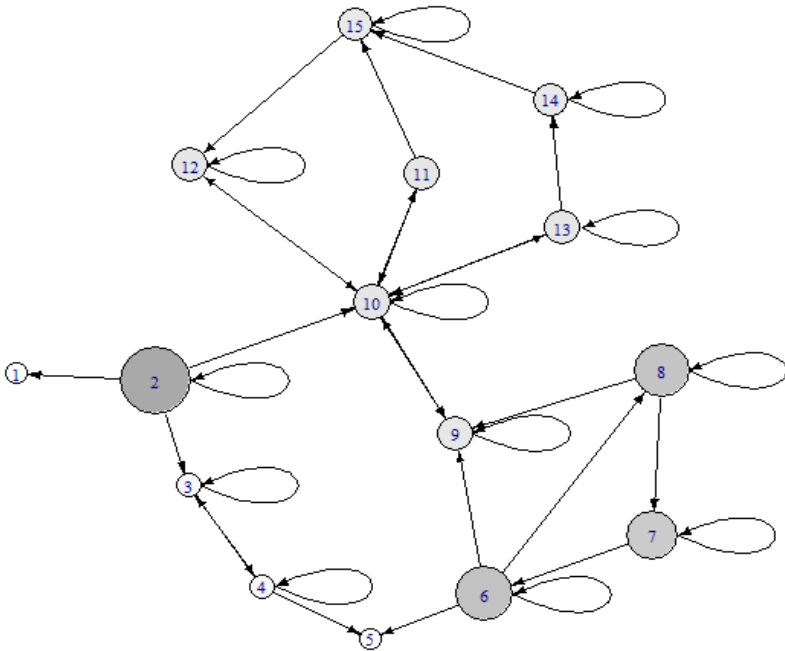


Рис. 4. Центральность по близости

Самый высокий показатель центральности по близости у вершины 2. Это значит, что она может получить быстрый доступ ко всем остальным узлам сети, следовательно, имеет возможность наиболее эффективно распространять информацию по сети.

3.1.3. ПРОМЕЖУТОЧНОСТЬ/ЦЕНТРАЛЬНОСТЬ ПО ПОСРЕДНИЧЕСТВУ (BETWEENNESS CENTRALITY)

Промежуточность позволяет найти те узлы, которые служат мостами для соединения с другими группами или сообществами в сети [2].

Узел с высоким значением промежуточности рассматривается как мощный узел с большим количеством влияния, чаще всего это единственная связь между различными частями сети.

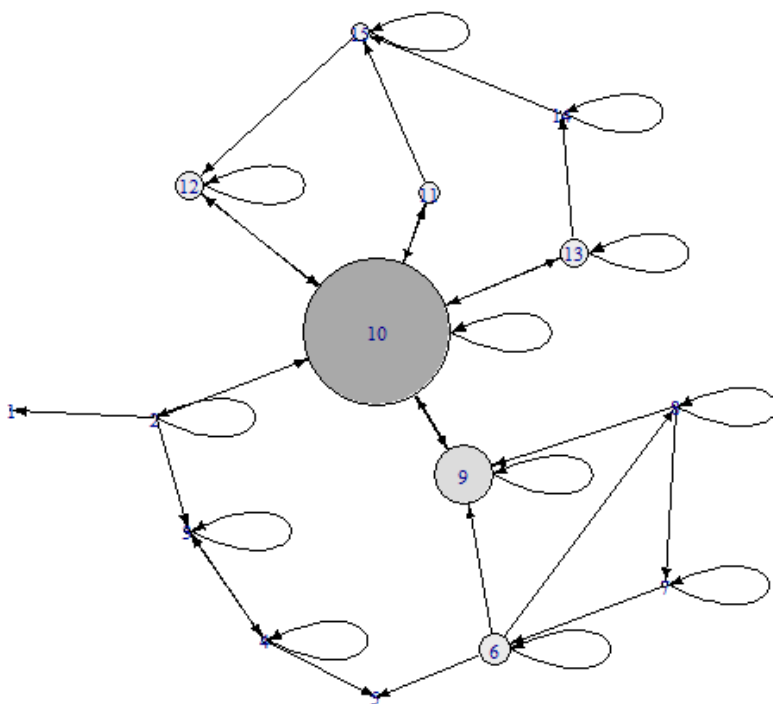


Рис. 5. Центральность по посредничеству

Значения промежуточности для вершин тестового графа:

$V_1 = 0,0$; $V_2 = 0,0$; $V_3 = 2,0$; $V_4 = 2,0$; $V_5 = 0,0$; $V_6 = 10,0$;
 $V_7 = 2,0$; $V_8 = 1,0$; $V_9 = 18,0$; $V_{10} = 44,5$; $V_{11} = 7,0$; $V_{12} = 9,0$;
 $V_{13} = 9,0$; $V_{14} = 1,0$; $V_{15} = 5,5$.

Самый большой показатель промежуточности у вершины 10 (рис. 5) оправдан тем, что она является единственным связующим узлом между двумя частями сети. Данный участник обладает значительным социальным капиталом и системными компетенциями, позволяющими ему контролировать информационные потоки внутри системы совместной сетевой деятельности.

Рассмотрим еще два примера социальных сетей. Первый случай – группа друзей собирается пойти на футбольный матч и договаривается о цене билетов (мнения).

Вершина 1 – успешный бизнесмен; 2 – офисный работник, у которого есть молодой коллега (вершина 8) – тоже фанат футбола, а у него, в свою очередь, компания друзей из 4-х человек (вершины 9, 10, 11, 20). 4 – «душа компании» и блогер, у которого есть свои подписчики (13–16 вершины); 6 – друг детства 4 и 3. Вершина 3 – женатый человек и ему важно мнение его жены (вершина 5), также ему придётся взять выходной у начальника – тоже фаната футбола (вершина 7); 17 – узнал из видео-блога о предстоящем матче и позвал друга, но так как он тоже женат, ему придется согласовать поход на матч со своей женой. Итак, получаем следующий граф (рис. 6).

Каждый из участников данной сети преследует свои интересы, например, успешный бизнесмен хочет большего комфорта, жёны хотят меньших затрат, молодые специалисты хотят попасть на матч, но потратить как можно меньше. Билеты на матч продаются в диапазоне от 100 руб. до 10000 руб. Отсюда вектор начальных мнений имеет вид (10000, 5000, 5000, 5000, 1000, 5000, 8000, 1000, 1000, 1000, 1000, 300, 500, 800, 400, 900, 200, 1000, 1000, 1000).

Базовые макропараметры социальной сети:

Плотность сети: 0,09210526. Значение плотности невелико, так как состоит из людей чаще всего не знакомых друг с другом лично. Взаимность: 0,4705882. Связи чаще всего не взаимны.

Транзитивность: 0,2790698. Низкая транзитивность говорит о том, что друг моего друга не всегда и мой друг тоже.

Самая высокая центральность по степени у вершины 4, что объясняется его увлечением, блогеры быстро распространяют информацию (рис. 7).

Самый высокий показатель центральности по близости у вершин 1, 2 и 4 (рис. 8). Данные вершины так или иначе оказывают сильное влияние на все остальные вершины.

Вершина 4 обладает самым высоким показателем посредничества (рис. 9), так как является связующим узлом между двумя частями сети. Совершенно незнакомые люди узнали о матче, просто посмотрев видео-блог.

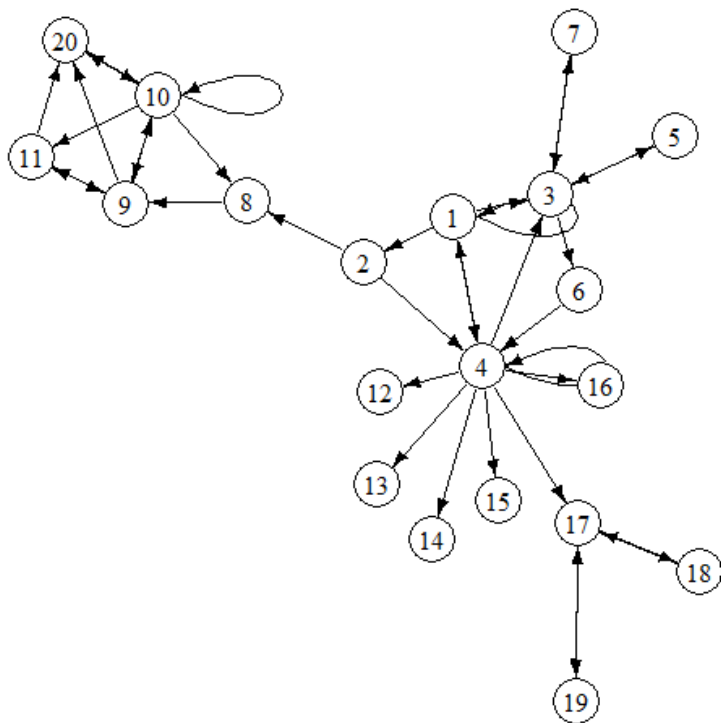


Рис. 6. Футбольный матч

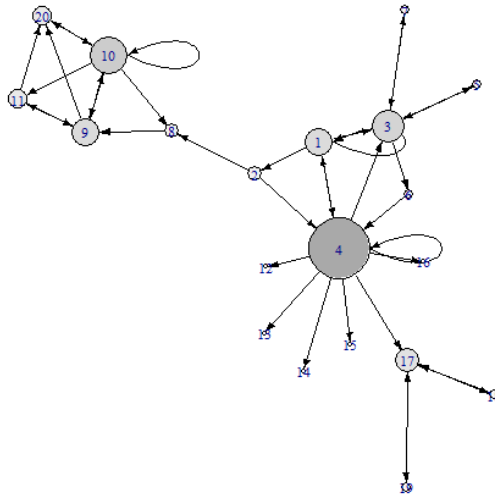


Рис. 7. График центральности по степени

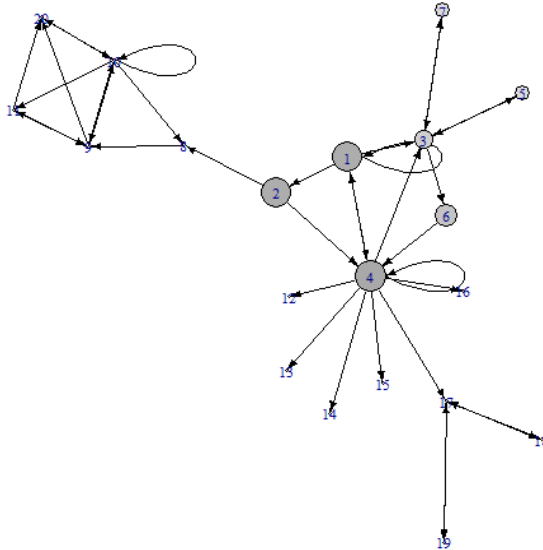


Рис. 8. График центральности по близости

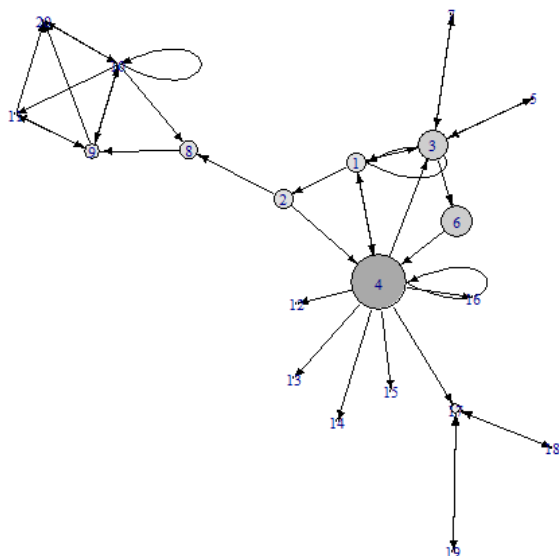


Рис. 9. График центральности по посредничеству

Наконец, вычислим сильные компоненты данной сети и их спутников (рис. 10).

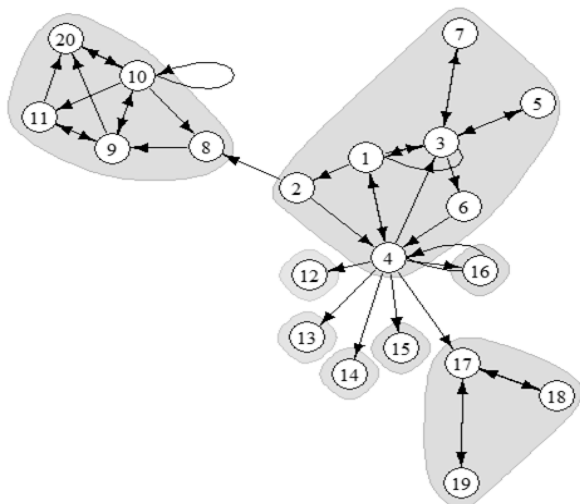


Рис. 10. Сильные компоненты и спутники

Рассмотрим следующий пример, иллюстрирующий довольно распространенный вид деятельности в социальных сетях, а именно набирающее популярность блогерство. Вершины графа представляют пять подруг в декретном отпуске, ведущих блоги о красоте, здоровье и своих детях (вершины 1–5), а также их подписчики, которые обсуждают сумму на покупку коляски (мнения). Так как данный граф содержит несколько сотен тысяч вершин, его сложно изобразить. Возьмем для наглядности граф со 100 вершинами (рис. 11) и вычислим характеристики для такой социальной сети.

Очевидно, что в такой сети сильной подгруппой будут являться сами блогеры, а их подписчики – спутниками (рис. 12).

Плотность сети: 0,03464646. Взаимность: 0,03292181. Транзитивность: 0,09532888.

Очень низкие показатели характеристик объясняются структурой сети и типом взаимосвязи ее объектов (рис. 13).

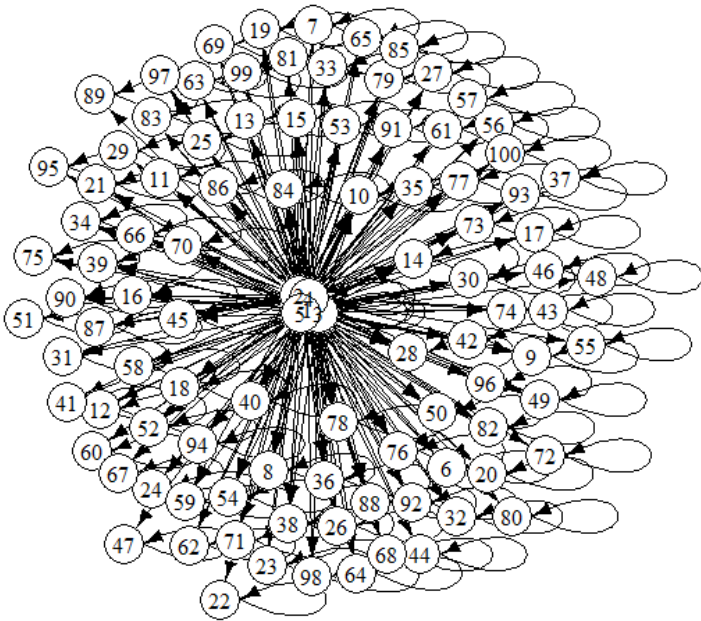


Рис. 11. Социальный граф блогеров и их подписчиков

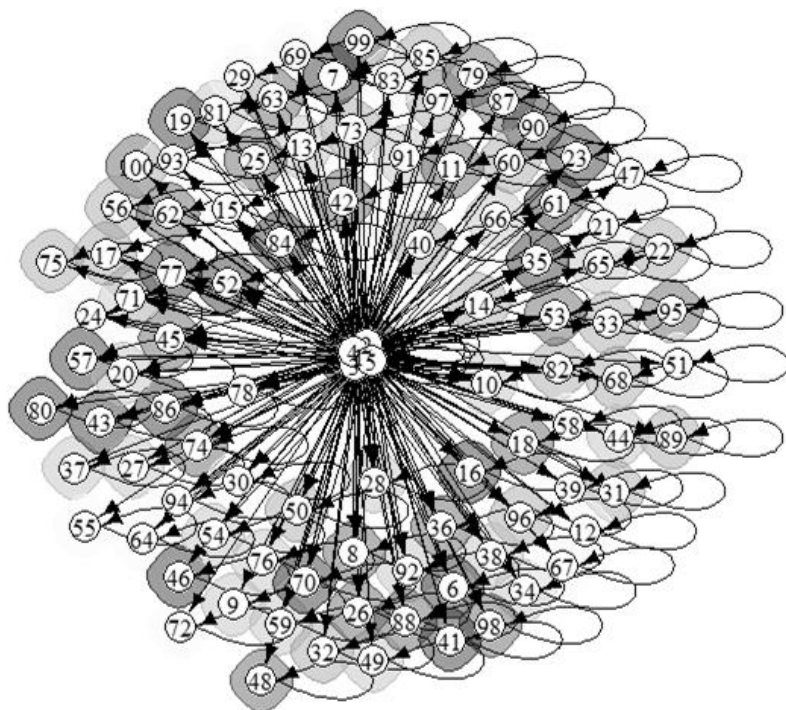


Рис. 12. Сильные компоненты и спутники

Аналогичная ситуация с микропараметрами сети. Блогеры обладают властью над своими подписчиками, которые прислушиваются к их мнению. Блоги служат хорошим инструментом для рекламы.

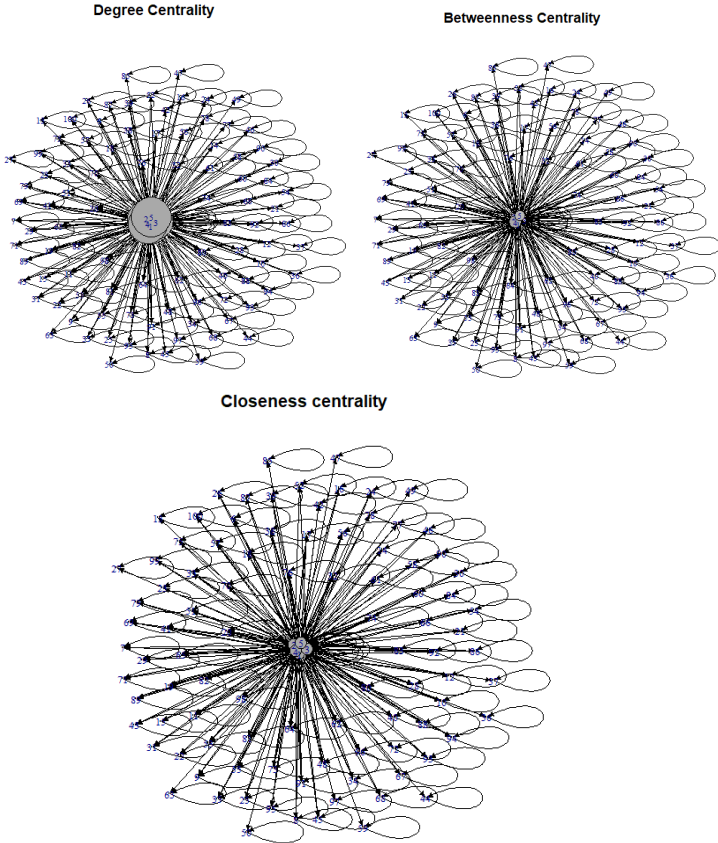


Рис. 13. Центральность

4. Задачи прогноза на сетях

Для решения задачи прогноза разработан и программно реализован с помощью языка R оригинальный алгоритм. Он базируется на таких понятиях, как эргодическое множество, поглощающие состояния и спутники [6, 12].

Находим обращение графа P . Сначала необходимо выяснить, сильно связный граф или нет. Если да, то финальное мнение группы будет единым. Тогда необходимо решить систему

$$\begin{aligned}
 a_1 \cdot w_1 + b_1 \cdot w_2 + \dots + m_1 \cdot w_n &= w_1, \\
 a_2 \cdot w_1 + b_2 \cdot w_2 + \dots + m_2 \cdot w_n &= w_2, \\
 &\dots \\
 a_{n-1} \cdot w_1 + b_{n-1} \cdot w_2 + \dots + m_{n-1} \cdot w_n &= w_{n-1}, \\
 w_1 + w_2 + \dots + w_n &= 1,
 \end{aligned}$$

где

$$P = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ c_1 & c_2 & \dots & c_n \\ & \dots & & \\ m_1 & m_2 & \dots & m_n \end{pmatrix} \text{ — обращение орграфа влияний.}$$

В данной системе линейных алгебраических уравнений (СЛАУ) неизвестными служат элементы матрицы W . Перенос всех неизвестных в правую часть равносильна операции в два шага:

1. $P^T - E$, P^T — транспонированная матрица, которая является исходным орграфом влияний, а E — единичная матрица.
2. Замена последней строки полученной матрицы на единицы.

Таким образом реализуется нахождение матрицы W : из исходной матрицы вычитается единичная и последняя строка заменяется на единицы. Получаем матрицу A в СЛАУ $AX = b$. В векторе b все элементы равны 0, кроме последнего, который равен 1. Далее решаем СЛАУ с помощью встроенных методов языка R и получаем матрицу W .

Если граф не сильно связный, то поиск финальных мнений делится на а) поиск сильных подгрупп (эргодических множеств) и б) поиск спутников:

- а. Для поиска эргодических множеств будем сначала искать сильные компоненты. Если из сильной компоненты не выходит никакая дуга в другие вершины, то множество эргодическое. Также заметим, что если множество содержит только одну вершину, то такое состояние называется поглощающим.

б. Вершины, которые не входят в эргодические множества – переходные, иначе говоря – спутники.

3. Для каждого из эргодических множеств можно найти финальное мнение, используя алгоритм для сильно связного графа.

4. Далее строим канонический вид обращения исходной матрицы

	S_1	S_2	...	S_r	T
S_1	1				0
S_2		1			
...			1		
S_r				1	
T	R				Q

где S – сильные подгруппы (в том числе поглощающие состояния), T – спутники.

В матрице R строки – это номера вершин, которые являются *переходными* состояниями, а столбцы – это *поглощающие* состояния и *эргодические* множества. В матрице Q и строки, и столбцы – это номера вершин, которые являются *переходными* состояниями.

Если первоначально цепь Маркова находится в переходном состоянии $y_{r+1,j}$, то вероятность ее перехода в эргодическое множество S_i равна значению элемента $b_{ji} \in B = NR$.

Фундаментальная матрица N поглощающей цепи Маркова находится по формуле $N = (E - Q)^{-1}$, где E – единичная матрица. Формула для B имеет вид $B = NR$.

Финальное мнение *поглощающих* состояний остается равным их начальному мнению.

Осталось найти финальное мнение *переходных* состояний

$$x_{r+1,j}^\infty = \sum_{i=1}^r b_{ji} \left(\sum_{k=1}^{n_i} w_k^{(i)} x_{ik}(0) \right).$$

Вернемся к тестовому примеру из раздела 2. Разберём подробно алгоритм на примере с сотрудниками. Сотрудникам можно дать неделю на размышления, но секретарю нужно знать сразу, в какой ценовой категории искать подарок. Поэтому важно спрогнозировать финальное мнение каждого сотрудника.

Сначала необходимо найти обращение матрицы влияний:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{12} & \frac{1}{2} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Для реализации общего случая понадобится вычислять сильные компоненты орграфа. Для этого в языке R есть специальный инструмент, который имеет три составляющие: показывает общее количество компонент, число вершин в каждой компоненте, а также к какой компоненте принадлежит каждая вершина. Выполняя поиск сильных компонент, получим

```
$membership
[1] 5 6 4 4 2 3 3 3 1 1 1 1 1 1 1
$size
[1] 7 1 3 2 1 1
$no
[1] 6
```

Membership: Первая вершина принадлежит 5-й компоненте, вторая компонента принадлежит 6-й компоненте, третья и четвёртая принадлежат 4-й компоненте, пятая – 2-й компоненте, шестая, седьмая и восьмая принадлежат 3-й, с девятой по пятнадцатую – 1-й компоненте.

Csize: вторая, пятая и шестая компоненты содержат по одной вершине, четвёртая – две вершины, третья компонента имеет три вершины, а первая – семь вершин.

No: всего 6 компонент.

Для дальнейшей реализации понадобится каждую компоненту выделять в отдельный вектор. Так как все компоненты могут иметь различную длину вектора, поступим следующим образом. Создадим матрицу (назовём её G), где каждая строка обозначает свою компоненту. В начале каждой строки идут номера вершин, которые принадлежат данной компоненте, далее идут нули. Вершины обязательно должны идти в начале, чтобы в дальнейшем при обходах строк в матрице идти не до конца строки, а сразу по значащим вершинам. Номера вершин в матрице уникальны и не могут повторяться в другой строке. Количество строк в такой матрице возьмем равным значению No (количество компонент).

Для примера выше матрица с векторами будет выглядеть следующим образом:

$$G = \begin{pmatrix} 9 & 10 & 11 & 12 & 13 & 14 & 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 7 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Также понадобится вектор для поглощающих состояний, который будет содержать номера вершин.

По определению поглощающее состояние – это эргодическое множество с одним состоянием. Можно сказать, что это та вершина, из которой не выходит ни одной дуги в другие вершины, есть только петля. В матрице P (транспонированной к ис-

ходной) собраны те вершины, где на диагонали стоят единицы. Заметим, что не обязательно такие состояния должны быть.

В нашем примере одно поглощающее состояние – вершина 2, так как только элемент $(2, 2)$ имеет значение 1.

Далее необходимы матрица, где будут собираться отдельно эргодические множества, и вектор, который будет собирать *только* переходные состояния, т.е. поглощающие состояния собираются отдельно (выше был описан алгоритм). Напомним, что не любая сильная компонента соответствует эргодическому множеству.

Чтобы проверить, что множество эргодическое, будем использовать следующий алгоритм. Возьмём матрицу G , в которую собраны компоненты. Напомним, что каждой строке соответствует отдельная компонента. Будем проверять, сколько вершин находится в каждой строке. Напомним, что в начале каждой строки идут номера вершин, которые принадлежат данной компоненте, далее идут нули. Таким образом, можно подсчитать число вершин, просто сосчитав количество положительных элементов.

Если в строке находится только одна вершина, то она подозрительна на занесение в вектор с *переходными* состояниями. Необходимо проверить ещё, что данная отдельная вершина не была занесена в вектор с *поглощающими* состояниями. Если нет, то эта вершина заносится в вектор, содержащий переходные состояния.

Если в строке матрицы G находится более одной вершины, то теперь необходимо проверить, что данная сильная компонента является эргодическим множеством. Для этого проверим, что ни из одной вершины сильной компоненты не выходит дуга в другие вершины. Для каждой сильной компоненты создается отдельный вектор с вершинами, не входящими в эту сильную компоненту. Далее проверяем, что в обращении исходной матрицы нет элементов, равных нулю, у которых номер строки — это вершина из компоненты, а номер столбца – это вершина, не входящая в сильную компоненту. Если всё это выполняется, то множество эргодическое и можно сразу вычислить финальное мнение группы (т.е. вершин, входящих в эргодическое множество), используя алгоритм для сильно связного графа.

Если множество оказалось не эргодическим, то заносим вершины в вектор с *переходными* вершинами. Заметим, что не нужно проверять, что эти вершины не являются поглощающими состояниями.

Весь алгоритм по поиску эргодических, переходных вершин и поглощающих состояний можно представить в виде структурной схемы (рис. 14 и рис. 15)

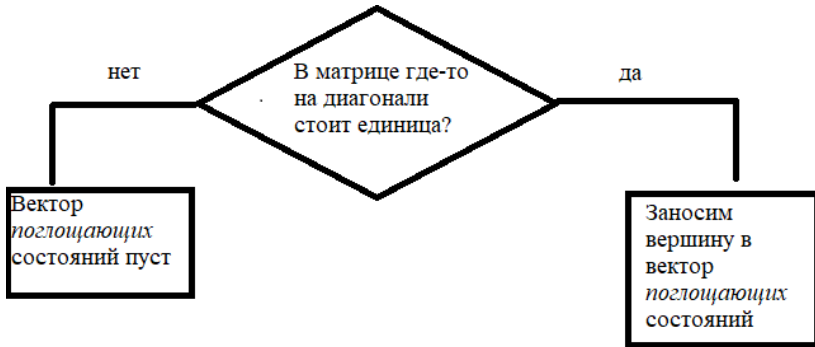


Рис. 14. Этап 1

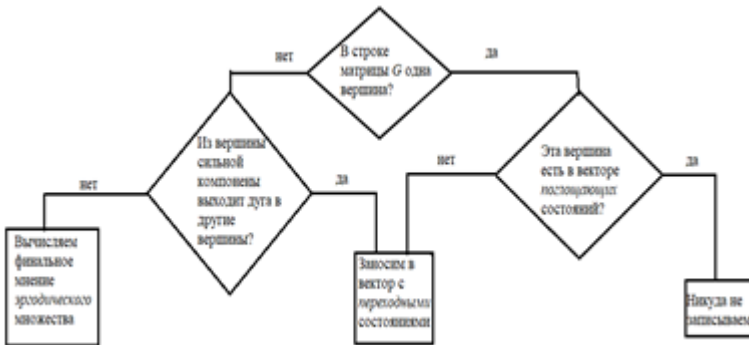


Рис. 15. Этап 2

Проиллюстрируем вышесказанное на тестовом примере из раздела 2.

Вершины 5, 1 и 2 подозрительны на внесение в вектор с *переходными* состояниями. Однако вершина 2 уже попала в вектор

поглощающих состояний. Значит, пока вектор переходных состояний состоит из вершин 5 и 1.

Рассмотрим теперь строки, где больше одной вершины. Это первая строка с вершинами 9–15, четвёртая строка с вершинами 6, 7 и 8, а также пятая строка с вершинами 3 и 4.

Проверим, что в компоненте с вершинами 6, 7 и 8 нет ни одной дуги в другие вершины. Вектор с другими вершинами *Not_Point_in_D* содержит вершины 1, 2, 3, 4, 5, 9 и 10. Пройдя всю матрицу *P*, можно заметить, что нет ни одного элемента с координатами $P[D[i], \text{Not_Point_in_D}[j]] = 0$, т.е. нет ни одной дуги, выходящей из вершин сильной компоненты в другие вершины. Следовательно, это *эргодическое* множество.

Если проделать то же самое с сильной компонентой с вершинами 3, 4, то можно увидеть, что есть дуга из 3 в 2, где 2 не входит в данную сильную компоненту. Следовательно, это не эргодическое множество и вершины 3, 4 записываются в вектор с *переходными* состояниями.

То же видно и у первой сильной компоненты с вершинами 9, 10, 11, 12, 13, 14, 15. Из 9 вершины есть дуги в вершины 6 и 8, а также из 10 выходит дуга во 2 вершину. Следовательно, это не эргодическое множество и вершины 9, 10, 11, 12, 13, 14, 15 тоже записываются в вектор с *переходными* состояниями.

Итак, вектор с *поглощающими* состояниями состоит из единственной вершины 2; вектор с *переходными* состояниями состоит из вершин 1, 3, 4, 5, 9, 10, 11, 12, 13, 14, 15. *Эргодическое* множество здесь одно и состоит из вершин 6, 7 и 8.

Вычислим финальное мнение группы с вершинами 6, 7 и 8. Для этого используем алгоритм для сильно связного графа. Необходимо выписать матрицу *P* – обращение орграфа влияний. Всегда будем получать СЛАОУ вида

$$a_1 \cdot w_1 + b_1 \cdot w_2 + \dots + m_1 \cdot w_n = w_1,$$

$$a_2 \cdot w_1 + b_2 \cdot w_2 + \dots + m_2 \cdot w_n = w_2,$$

...

$$a_{n-1} \cdot w_1 + b_{n-1} \cdot w_2 + \dots + m_{n-1} \cdot w_n = w_{n-1},$$

$$w_1 + w_2 + \dots + w_n = 1,$$

где

$$P = \begin{pmatrix} a_1 & a_2 & a_n \\ b_1 & b_2 & b_n \\ c_1 & c_2 & \dots & c_n \\ & \dots & & \\ m_1 & m_2 & m_n \end{pmatrix}.$$

В данной СЛАУ неизвестные включены в матрицу W . Перенос всех неизвестных в правую часть равносильна операции в два шага:

1. $P^T - E$, P^T – транспонированная матрица, которая является исходным оргграфом влияний, а E – единичная матрица;

2. Замена последней строки полученной матрицы на единичную строку.

Таким образом, будем реализовать нахождение матрицы W : из исходной матрицы вычитаем единичную и последнюю строку заменяем единичной. Получим матрицу A в СЛАУ $AX = b$. Вектор b – это вектор, в котором все элементы равны 0, кроме последнего, который равен 1. Далее решаем СЛАУ с помощью встроенных методов языка R и получаем матрицу W .

Получается следующая матрица:

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Транспонируем эту матрицу, чтобы получить исходную. В данном случае

$$T = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Вычитаем из неё единичную матрицу и заменяем последнюю строку на единичную:

$$L = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Решаем полученное СЛАУ $LW = b$ и получаем

$$W = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}.$$

Осталось вычислить общее финальное мнение группы:

$$\begin{aligned} x^\infty &= \sum_{i=1}^n w_i x_i(0) = \\ &= \frac{1}{3} x_6(0) + \frac{1}{3} x_7(0) + \frac{1}{3} x_8(0) = \frac{1}{3} 500 + \frac{1}{3} 800 + \frac{1}{3} 600. \end{aligned}$$

Далее перейдём к построению матриц R и Q .

Матрица R – матрица, у которой строки – это номера вершин, которые являются *переходными* состояниями, а столбцы – это *поглощающие* состояния и *эргодические* множества.

Матрица Q – матрица, у которой и строки, и столбцы – это номера вершин, которые являются *переходными* состояниями.

Начнём с матрицы Q . Схематично она будет выглядеть как

$$Q = \begin{pmatrix} A[a,a] & A[a,b] & \cdots & A[a,c] \\ A[b,a] & A[b,b] & \cdots & A[b,c] \\ & \dots & & \\ A[c,a] & A[c,b] & \cdots & A[c,c] \end{pmatrix}.$$

где P – обращение исходной матрицы; a, b, \dots, c – *переходные* состояния.

В примере, рассмотренном выше, матрица Q имеет вид

$$\begin{pmatrix}
 9 & \left(\begin{array}{cccccccccccc}
 \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \frac{1}{12} & \frac{1}{2} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 11 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\
 12 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
 13 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 \\
 14 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\
 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\
 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 \\
 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 \\
 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right) \\
 (9 & 10 & 11 & 12 & 13 & 14 & 15 & 5 & 3 & 4 & 1)
 \end{pmatrix}$$

где слева и внизу указаны номера вершин.

Перейдём к построению матрицы R . Её схема выглядит как

$$R = \begin{pmatrix}
 P[a,u] & \cdots & P[a,v] & P[a,d] + \dots + P[a,f] & \cdots & P[a,g] + P[a,z] \\
 P[b,u] & \cdots & P[b,v] & P[b,d] + \dots + P[b,f] & \cdots & P[b,g] + P[b,z] \\
 \cdots & & & & & \\
 P[c,u] & \cdots & P[c,v] & P[c,d] + \dots + P[c,f] & \cdots & P[c,g] + P[c,z]
 \end{pmatrix},$$

где P – обращение исходной матрицы; a, b, \dots, c – переходные состояния; u, \dots, v – поглощающие состояния; d, \dots, f – вершины, входящие в первое эргодическое множество; g, \dots, z – вершины, входящие в n -е эргодическое множество.

В нашем примере матрица R выглядит следующим образом:

$$\begin{array}{c}
 \left(\begin{array}{c}
 9 \\
 10 \\
 11 \\
 12 \\
 13 \\
 14 \\
 15 \\
 5 \\
 3 \\
 4 \\
 1
 \end{array} \right) \left(\begin{array}{c}
 \frac{1}{4} + \frac{1}{4} \quad 0 \\
 0 \quad \frac{1}{6} \\
 0 \quad 0 \\
 0 \quad 0 \\
 0 \quad 0 \\
 0 \quad 0 \\
 \frac{2}{3} \quad 0 \\
 0 \quad \frac{1}{4} \\
 0 \quad 0 \\
 0 \quad 0 \\
 0 \quad 1
 \end{array} \right) \\
 (S \quad 2)
 \end{array}$$

Слева – вершины переходных состояний (строки), а внизу S – сильная компонента, она включает в себя вершины 6, 7 и 8. Вершина 2 – поглощающее состояние. Как можно заметить, элемент $[9, S]$ получен путём сложения элементов обращения исходной матрицы: просуммировали элементы $P[9, 6] + P[9, 7] + P[9, 8]$.

После того как получены матрицы R и Q , можно перейти к нахождению матриц N и B . Формула для N : $N = (E - Q)^{-1}$, где E – единичная матрица. Инструменты языка R позволяют быстро найти матрицу N через функцию *solve*. Формула для B : $B = NR$.

Для тестового примера

$$N = \begin{pmatrix} 1,5 & 1,5 & 0,146 & 0,167 & 0,292 & 0,833 & 0,0625 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 0,437 & 0,5 & 0,875 & 0,25 & 0,1875 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 1,437 & 0,5 & 0,875 & 0,25 & 0,1875 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 0,604 & 1,833 & 1,208 & 0,917 & 0,6875 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 0,4375 & 0,5 & 2,875 & 0,25 & 0,1875 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 0,4375 & 0,5 & 2,875 & 4,25 & 0,1875 & 0 & 0 & 0 & 0 \\ 0,5 & 4,5 & 0,9375 & 0,5 & 1,875 & 2,25 & 1,6875 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1,3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{2}{3} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & \frac{3}{4} & \frac{3}{4} & \frac{3}{4} & \frac{3}{4} & \frac{3}{4} & \frac{1}{3} & 1 & 1 & 1 \end{pmatrix}.$$

Осталось найти финальное мнение *поглощающих* состояний и финальное мнение *переходных* состояний.

Финальное мнение поглощающих состояний равно своему первоначальному мнению: $x_2^\infty = 900$.

Финальное мнение *переходных* состояний

$$x_{r+1, j}^\infty = \sum_{i=1}^r b_{ji} \left(\sum_{k=1}^{n_i} w_k^{(i)} x_{ik}(0) \right),$$

$$x_9^\infty = \frac{3}{4} \cdot S + \frac{1}{4} \cdot x_2^\infty = 450 + 225 = 675,$$

$$x_{10}^\infty = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^\infty = 150 + 675 = 825,$$

$$x_{11}^\infty = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^\infty = 825,$$

$$x_{12}^\infty = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^\infty = 825,$$

$$x_{13}^\infty = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^\infty = 825,$$

$$x_{14}^{\infty} = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^{\infty} = 825,$$

$$x_{15}^{\infty} = \frac{1}{4} \cdot S + \frac{3}{4} \cdot x_2^{\infty} = 825,$$

$$x_5^{\infty} = \frac{2}{3} \cdot S + \frac{1}{3} \cdot x_2^{\infty} = \frac{2}{3} \cdot 600 + \frac{1}{3} \cdot 900 = 700,$$

$$x_3^{\infty} = 0 \cdot S + 1 \cdot x_2^{\infty} = 900,$$

$$x_4^{\infty} = 0 \cdot S + 1 \cdot x_2^{\infty} = 900,$$

$$x_1^{\infty} = 0 \cdot S + 1 \cdot x_2^{\infty} = 900.$$

Итого получаем:

$$x_1^{\infty} = 900; x_2^{\infty} = 900; x_3^{\infty} = 900; x_4^{\infty} = 900; x_5^{\infty} = 700; x_6^{\infty} = 600; \\ x_7^{\infty} = 600; x_8^{\infty} = 600; x_9^{\infty} = 675; x_{10}^{\infty} = 825; x_{11}^{\infty} = 825; x_{12}^{\infty} = 825; \\ x_{13}^{\infty} = 825; x_{14}^{\infty} = 825; x_{15}^{\infty} = 825.$$

Таким образом, полностью реализован алгоритм по поиску финальных мнений. Найдены финальные мнения всех групп состояний – *поглощающих*, *переходных* и *эргодических* множеств.

По результатам тестового примера можно сделать вывод, что «влиятельными» членами данной компании (руководителя в счёт не берём) являются сотрудники 2, 6, 7 и 8. Несмотря на то, что группа из сотрудников 9–15 достаточно большая и тесно общается, в данном случае их конечные мнения всё равно формируются благодаря влиятельным сотрудникам, так как работники 10 и 9 являются своеобразным мостом между их группой и влиятельными членами компании. Соответственно, сотрудники 2, 6, 7, 8 косвенно влияют на 11–15, формируя их мнение.

Таким образом, так или иначе все коллеги прислушиваются к мнению «влиятельных» работников, изменяя своё собственное, в то время как влиятельные сотрудники прислушиваются или только к мнению друг друга – как в случае с работниками 6, 7, 8 – или вообще своего мнения не меняют, как в случае с работником 2. Поэтому если секретарь захочет поменять мнение всего коллектива (например увеличить сумму подарка), то достаточно будет повлиять на мнения сотрудников 2, 6, 7 и 8, а не всех пятнадцати человек.

Теперь введём данные второго примера из раздела 2 (футбольный матч) в программу на языке R и интерпретируем полученные результаты. При векторе начальных мнений (10000, 5000, 5000, 5000, 1000, 5000, 8000, 1000, 1000, 1000, 1000, 300, 500, 800, 400, 900, 200, 1000, 1000, 1000) получаем следующие результаты. Вершины 1–7 образуют эргодическое множество, их финальное мнение равно 6876,079. Внутри этой группы люди влияют друг на друга, хотя не все члены между собой обязаны быть знакомы лично (например, жена работника 3 и начальник этого же работника). Но на эту группу никто извне не влияет. Неженатые и не имеющие друзей – фанатов футбола подписчики блогера 4 находятся под полным его влиянием и также готовы потратить на билет 6876,079 рублей. Вершины 17–19 образуют сильную компоненту, однако не являются эргодическим множеством. На семьянина 17 оказывает влияние блогер 4. Несмотря на то, что 17 должен посоветоваться с женой перед покупкой билета, можно видеть, что жена всецело доверяет мужу. Соответственно 17–19 также готовы потратить 6876,079 рублей. Вершины 8–11, 20 также образуют сильную компоненту, но на них извне влияет вершина 2. В их случае 8 готов потратить 7410,387 рублей, 9-й – 7392,495 рублей, 10-й – 7650,439, 11-й – 7451,822 и 20-й – 7520,356 рублей.

Можно сделать вывод, что блогер имеет очень большое влияние не только на своих непосредственных друзей и подписчиков, но и на их жён и друзей.

Программа на R хорошо работает и на третьем примере, показывая итоговые финальные мнения у каждой вершины графа. Подруги-блогеры формируют мнение внутри своей группы и оказывают огромное влияние на своих непосредственных подписчиц, а также на их подруг, которые не являются подписчицами. Можно сказать, что если производитель колясок захочет дать свою рекламу, то лучше обратиться напрямую к этим блогерам, нежели воздействовать непосредственно на покупательниц.

Хотя отдельные действия алгоритма, такие как транспонирование матриц и решение СЛАУ, решаются стандартными инструментами языка R, в совокупности реализация алгоритма

оригинальна, поскольку до сих пор ещё не была написана программа, вычисляющая финальные мнения.

Наиболее сложной частью алгоритма является выполнение циклов. Самая сложная в данной реализации – это вложенный цикл (глубина равна 2, т.е. внутренний и внешний цикл). Оценивая сложность алгоритма, необходимо использовать только ту часть, что растёт больше всего. В итоге получим, что сложность алгоритма вычисляется как $O(n^2)$, где n – число вершин графа. Такая сложность вполне приемлема, соответственно программу можно применять и для графов с куда большим количеством вершин.

5. Заключение

В статье рассмотрены задачи анализа и прогноза на социальных сетях применительно к решению задачи управления целевой аудиторией в маркетинге. Предложен оригинальный алгоритм вычисления финальных мнений членов целевой аудитории для частного случая сильно связанного орграфа влияний и общего случая произвольного орграфа. Этот алгоритм и другие методы анализа реализованы на языке программирования R и апробированы на тестовых примерах.

В дальнейшем предполагается опробовать предложенную методику исследования целевой аудитории на реальных предприятиях, а также рассмотреть задачи управления мнениями, использующие выявленную на этапе анализа структуру целевой аудитории.

Литература

1. АГИЕВА М.Т. *Модели управления на социальных сетях в маркетинге* // Инженерный вестник Дона. – 2018. – №1. – URL: ivdon.ru/ru/magazine/archive/n1y2018/4670.
2. АГИЕВА М.Т. *Задачи анализа на социальных сетях в маркетинге* // Инженерный вестник Дона. – 2018. – №2. – URL: ivdon.ru/ru/magazine/archive/N2y2018/4889.

3. АЛЕСКЕРОВ Ф.Т., БЛАГОВЕЩЕНСКИЙ Н.Ю., САТАРОВ Г.А. и др. *Влияние и структурная устойчивость в Российском парламенте (1905–1917 и 1993–2005 гг.)*. – М.: Физматлит, 2007. – 312 с.
4. БРЕЕР В.В. *Стохастические модели социальных сетей // Управление большими системами*. – 2009. – Вып. 27. – С. 169–204.
5. БРЕЕР В.В., НОВИКОВ Д.А., РОГАТКИН А.Д. *Управление толпой. Математические модели порогового коллективного поведения*. – М.: ЛЕНАНД, 2016. – 168 с.
6. ГУБАНОВ Д.А., НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Социальные сети: модели информационного влияния, управления и противоборства*. – М.: Изд-во физико-математической литературы, 2010. – 228 с.
7. ГУБАНОВ Д.А., КОРГИН Н.А., НОВИКОВ Д.А., РАЙКОВ А.Н. *Сетевая экспертиза*. – М.: Эгвес, 2011. – 166 с.
8. КАБАКОВ Р.И. *R в действии. Анализ и визуализация данных в программе R*. – М.: ДМК Пресс, 2016. – 588 с.
9. КУЗНЕЦОВ А.Л. *Об анализе социальных сетей и их метриках // Материалы VIII Международной студенческой электронной научной конференции «Студенческий научный форум»*. – URL: www.scienceforum.ru/2017/2830/31138 (дата обращения: 14.03.2018).
10. ЛИФИНЦЕВ Д.В. *Оценка связей индивида в микросоциуме на основе методов анализа социальных сетей // Вестник Балтийского федерального университета им. И. Канта. Серия: Филология, педагогика, психология*. – 2013. – №5. – URL: <https://cyberleninka.ru/article/n/otsenka-svyazey-individual-v-mikrosotsiume-na-osnove-metodov-analiza-sotsialnyh-setey> (дата обращения: 14.03.2018).
11. МОРЕНО Я.Л. *СОЦИОМЕТРИЯ: Экспериментальный метод и наука об обществе*. – М.: Академический Проект, 2001. – 384 с.
12. РОБЕРТС Ф. *Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам*. – М.: Наука, 1986. – 496 с.
13. BARNES J.A. *Classes and Committees in a Norwegian Island Parish // Human Relations*. – 1954. – No. 7. – P. 39–58.

14. BERGER R.J. *A Necessary and Sufficient Conditions for Reaching a Consensus using De Groot's method* // J. of American Statistical Association. – 1981. – Vol. 76. – P. 415–419.
15. BUTTLE F.A. *Word-of-Mouth: Understanding and Managing Referral Marketing* // J. of Strategic Marketing. – 1998. – Vol. 6. – P. 241–254.
16. CHATTERJEE S., SENETA E. *Toward Consensus: Some Convergence Theorems on Repeated Averaging* // J. of Applied Probability. – 1977. – No. 14. – P. 159–164.
17. CHWE M.S. *Communication and Coordination in Social Networks* // Review of Economic Studies. – 2000. – Vol. 67 – No. 1– P. 1–16.
18. DE GROOT M.H. *Reaching a Consensus* // J. of American Statistical Association. – 1974. – Vol. 69. – P. 118–121.
19. DE MARZO P., VAYANOS D., ZWIEBEL J. *Persuasion Bias, Social Influence and Unidimensional Opinions* // Quaterly J. of Economics. – 2003. – No. 118(3) . – P. 909–968.
20. FRENCH J.R. *A formal theory of social power* // The Psychological Review. – 1956. – No. 63. – P. 181–194.
21. GODES D., MAYZLIN D. *Using Online Conversations to Study Word of Mouth Communication* // Marketing Science. – 2004. – No. 23. – P. 545–560.
22. GOLDENBERG J., LIBAI B., MULLER E. *Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth* // Marketing Letters. – 2001. – No. 2. – P. 11–34.
23. GOLUB B., JACKSON M. *Naive Learning in Social Networks and the Wisdom of Crowds* // American Economic J.: Microeconomics. – 2010. – No. 2(1). – P. 112–149.
24. HARARY F. *A Criterion for Unanimity in French's Theory of Social Power* // Studies in Social Power. – Michigan: Institute of Sociological Research, 1959. – P.168–182.
25. HEGSELMAN R., KRAUSE U. *Opinion Dynamics and Bounded Confidence Models: Analysis and Simulation* // J. of Artificial Societies and Social Simulation. – 2002. – Vol.5 – No. 3– P. 226–229.

26. JACKSON M. *Social and Economic Networks*. – Princeton University Press, 2008. – 504 p.
27. JACKSON M., WOLINSKY A. *A strategic model of social and economic networks* // J. Economic Theory. – 1996. – Vol. 71(1). – P. 44–74.
28. KRAUSE U. *A Discrete Nonlinear and Non-autonomous Model of Consensus Formation* / Communications in Difference Equations. – Amsterdam: Gordon and Breach Publishers, 2000. – P. 227–236.
29. LORENZ J. *A Stabilization Theorem for Dynamics of Continuous Opinions* // Physica A. – 2005. – Vol. 355. – P. 217–223.
30. MASUDA N., KAWAMURA Y., KORI H. *Analysis of relative influence of nodes in directed networks* // Physical Review. – 2009. – E80, 046114.
31. NEWMAN M. *The Structure and Function of Complex Networks* // SIAM Review. – 2003. – Vol. 45(2). – P. 167–256.
32. ROBINS G., PATTISON P., ELLIOT P. *Network Models for Social Influence Processes* // Psychometrica. – 2001. – No. 66(2). – P. 161–190.
33. WATTS D. *The "New" Science of Networks* // Annual Review of Sociology. – 2004. – Vol. 30. – P. 243–270.

ANALYSIS AND FORECASTING PROBLEMS IN THE CONTROL OF TARGET AUDIENCE IN MARKETING

Movlatkhan Agieva, Ingush State University, Nazran, Cand. Sc., Associate Professor (agieva25@mail.ru).

Julia Babicheva, Southern Federal University, Rostov-on-Don, Bachelor of Science (yulya.babicheva@yandex.ru).

Natalya Okulist, Southern Federal University, Rostov-on-Don, Bachelor of Science (nata-0k@rambler.ru).

Guennady Ougolnitsky, Southern Federal University, Rostov-on-Don, Doctor of Sc., Professor (ougoln@mail.ru).

Abstract: Some networks analysis and forecasting models are considered in the paper relative to marketing. A brief survey of the control and influence models on social networks is presented. Namely, the problems include determination of the strong subgroups and satellites, calculation of quantitative characteristics of the network, determination of the final opinions of the members of the target audience

based on their initial opinions and mutual interactions. For the solution of the forecasting problem an original algorithm is developed and implemented by means of the programming language R. The algorithm is described in details and considers a specific case of a strongly connected digraph and the general case. The complexity of the algorithm is estimated. Other named problems can be solved by embedded functions of the language that is illustrated by test examples together with the results of the authors' algorithm. The model examples use influence digraphs with 10-15 and more than 100 vertices. Such closeness measures as degree centrality, closeness centrality, and betweenness centrality are calculated for the model example, their interpretation is given. Solution of the analysis and forecasting problems essentially facilitates the problem of control of the opinions of target audience in marketing. The directions of future research in building and analysis of the models of opinion control in marketing are outlined.

Keywords: marketing, network analysis and forecasting models, control of target audience.

УДК 519:301

ББК 60.54; 32.81

DOI: <https://doi.org/10.25728/ubs.2019.79.2>

*Статья представлена к публикации
членом редакционной коллегии Я.И. Квинто.*

Поступила в редакцию 06.10.2018.

Опубликована 31.05.2019.