

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЭНЕРГЕТИЧЕСКИХ СИСТЕМ

Фурсов Дмитрий Викторович
Выпускная квалификационная работа

**Имитационное моделирование информационного
воздействия с использованием средств
массовой коммуникации**

Направление 01.03.02
«Прикладная математика, фундаментальная информатика и
программирование»

Научный руководитель,
кандидат физ.-мат. наук,
доцент
Сwirкин М.В.

Санкт-Петербург
2018

Содержание

Введение.....	3
Постановка задачи.....	6
Обзор литературы.....	8
Информационное воздействие и особенности его ведения.....	9
§1.1 Содержательное описание информационного воздействия.....	9
§1.2 Формализация введённых терминов.....	15
§1.3 Информационно-логическая модель.....	19
Математическая модель информационного воздействия.....	21
§2.1 Вероятностно-статистический анализ и задание пространства параметров.....	21
§2.2 Математическая модель.....	25
§2.3 Имитационное моделирование.....	36
Имитационное моделирование на примере социальной сети «ВКонтакте»...	39
§3.1 Сбор данных и выбор параметров в социальной сети «ВКонтакте»....	39
§3.2 Проведение имитационного моделирования.....	41
§3.3 Анализ результатов.....	44
Заключение.....	45
Список литературы.....	46
Приложение 1.....	48

Введение

С древних времён информация являлась важнейшим источником управления людьми и ресурсами. Владея информацией, можно воздействовать на все сферы человеческой жизни. А в настоящее время, используя средства коммуникации: прежде всего интернет, социальные сети, электронную почту, мессенджеры можно воздействовать на массы людей.

Сегодня резко возрос объем информации, ежедневно поступающей и обрабатываемой человеком. В социальных сетях современный человек проводит большое количество времени. В связи с этим в сетях активно используется и совершенствуется способ распространения информации - таргетинговая реклама. Это рекламный механизм, позволяющий выделить из всей имеющейся аудитории только ту часть, которая удовлетворяет заданным критериям (целевую аудиторию), и показать рекламу именно ей. Поэтому в качестве объекта изучения взята социальная сеть – «ВКонтакте» и её специальный сервис – «Рекламный кабинет». Этот сервис позволяет распространить информацию в новостной ленте у пользователя (рекламная карусель, рекламная запись с кнопкой, реклама в сообществах); продвигать как записи, опубликованные на стене сообщества, так и рекламные записи, созданные в рекламном кабинете (универсальная рекламная запись); создавать объявления со ссылкой на внешний сайт, сообщество или приложение «ВКонтакте» (рекламные объявления на страницах сайта). В данной работе распространение информации будет происходить в новостной ленте у пользователя, а именно – реклама в сообществах. Во всех сервисах используется лишь одна функция: «настройки целевой аудитории», за исключением рекламы в сообществах, в котором предоставляется статистика каждого сообщества. Таким образом, появляется большее количество параметров, которые можно менять, исходя из общедоступной статистики, комбинировать выбор сообществ по нескольким параметрам, в том числе и по стоимости публикации записи в них, что позволит более взвешено тратить

выделенные на рекламу денежные средства, не ухудшая при этом выходные параметры. А если при всём этом известна статистика распространения информации по определенным темам, то можно основываясь на этой статистике выдвигать гипотезу о том, какой ожидается результат и с какой вероятностью он будет достигнут по выбранной тематике и т.д.

Используя вероятностно-статистический анализ и имитационное моделирование, будет построена модель распространения информации. С её помощью любой пользователь социальной сети «ВКонтакте» сможет получать с определенной вероятностью предполагаемый «feedback»¹, задав ряд необходимых входных параметров.

В настоящее время существуют монографии, в которых формализована в заданной предметной области небольшая часть параметров модели и отражены взаимосвязи основных сущностей, которые характеризуют предметную область. Но все подобные исследования носят описательный характер и не используют при составлении модели в достаточной мере аппарата математического анализа.

Актуальность темы очевидна в связи с тем, что сети стали новой ступенью в развитии не только коммуникаций между людьми, но и в сферах маркетинга, рекламы, торговли, экономики, государственного управления. Главное - социальные сети, а точнее их пользователи, учувствуют и оказывают воздействие на разрешение острых социальных, межконфессиональных, политических и экономических проблем человечества. Так, например, в настоящее время сети активно используют террористы для вербовки в свои ряды новых членов. Социальные сети используют также для улучшения различных государственных систем и повышения уровня жизни людей, проводятся различные акции и шествия, о проведении которых люди узнают через них. А данная модель позволит любому пользователю данной социальной сети заранее узнать с

¹«feedback» от англ. – обратная связь.

определенной вероятностью сценарий, который его ожидает, по распространению интересующей его информации.

В представленной работе область исследований – социальная сеть, а в качестве предмета исследования выступает один из нескольких сервисов социальной сети «ВКонтакте» - реклама в сообществах. Под рекламой в сообществах будет пониматься: публикация в данном сообществе записи, содержащей: текстовую информацию, медиа файлы и внешние ссылки. Под информационным воздействием в дальнейшем понимается производство и распространение информации.

В данной работе будет рассматриваться информационное воздействие через социальные сети. Она будет включать анализ предметной области, выделение основных сущностей модели распространения информации, а также будут введены параметры, характеризующие каждую сущность, будет построена информационно-логическая модель, являющаяся основой имитационной модели анализа информационного воздействия, приведён пример имитационного моделирования, его программная реализация в Microsoft Visual Studio 2013 и произведён анализ полученных результатов.

Постановка задачи

Анализ и исследование такого рода воздействий, является основополагающим фактором, обеспечивающим национальную безопасность, потому что информационное воздействие через социальные сети может носить как положительный, так и отрицательный характер.

Задача будет состоять в том, чтобы проанализировать информационное воздействие с использованием социальных сетей, для получения выводов, на которых будет строиться математическая модель.

Анализ предметной области и математическое моделирование являются неотъемлемой составляющей данной работы, она будет проходить в три этапа:

«анализ предметной области» → «построение модели» → «анализ результатов моделирования».

Стоит обратить внимание на то, что при анализе результатов моделирования используются понятия и представления, возникшие в процессе моделирования. Это позволяет «выйти» за пределы возможностей математического моделирования, т.е. получить прогноз развития системы и её свойств, который нельзя получить математическими средствами, с одной стороны, а с другой – позволяет по-новому взглянуть на изучаемую систему с точки зрения построения математической модели. Таким образом, класс систем, которые можно анализировать вышеизложенным способом, расширяется.

Распространение информации в целом представляет собой сложный процесс, для описания которого требуется большое количество параметров и связей, а для анализа такого рода объекта необходим системный подход. В данной работе необходимо осуществить процесс построения имитационной модели распространения информации в социальной сети «ВКонтакте» при использовании специального сервиса «Реклама ВКонтакте», который позволяет анализировать статистику сообществ. Для достижения

поставленной цели необходимо:

- на основе анализа предметной области определить основные сущности, используемые в модели;
- описать параметры, характеризующие каждую из них;
- рассмотреть и описать связи между сущностями;
- построить информационно-логическую модель распространения информации в социальной сети;
- на основе полученной информационно-логической модели построить математическую модель с выделенным пространством параметров;
- по математической модели в выделенном пространстве параметров провести имитационное моделирование;
- на модельном примере показать работоспособность предлагаемого подхода определения вероятностей распространения информации на основе имитационного моделирования.

Таким образом, решается задача моделирования происходящего процесса: составления математической модели и на её основе проведения имитационного моделирования информационного воздействия. Также, используя имитационное моделирование, его результаты будут определяться случайным характером процессов и по этим данным можно получить достаточно устойчивую статистику, которую в дальнейшем можно будет использовать для корректировки входных параметров модели при решении однотипных задач.

Обзор литературы

При написании данной работы была рассмотрена научная литература и статьи на тему информационного воздействия. Был проведен анализ литературы и выделены основные работы по данной теме, такие как Н.В. Белотелов, Ю.И. Бродский, Ю.Н. Павловский «Сложность. Математическое моделирование. Гуманитарный анализ: Исследование исторических, военных, социально-экономических и политических процессов»/ Предисл. Г.Г. Малинецкого. – М.: Книжный дом «ЛИБРОКОМ», 2009ю – 320с. [1], Ю.Н. Павловский, Н.В. Белотелов, Ю.И. Бродский «Имитационное моделирование» : учеб.пособие для студ. высш. учеб. заведений / - М.: Издательский центр «Академия», 2008. – 236с. – (Университетский учебник. Сер. Прикладная математика и информатика) [2], РАСТОРГУЕВ С.П. «Математические модели в информационном противоборстве». Экзистенциальная математика. — М.: АНО ЦСОиП, 2014. — 260 с. [3], Влияние через социальные сети / Под общей ред. Е.Г. Алексеевой [4]. А также интернет ресурсы [5-8]. С использованием данных источников было проведено ознакомление с причинами появления и видами моделей информационного воздействия. В данной работе на основе рассмотренного материала была построена модель ведения информационного воздействия, а именно социальных сетей, она относится к группе имитационных.

В современном мире анализ распространения информации в социальных сетях необходим, так как всё большее количество людей становятся пользователями различных социальных сетей. Выработав универсальный способ распространения информации в пределах одной сети, получим новый скачок в развитии рекламы и маркетинга в этой области. После изучения предметной области были найдены подобные исследования, однако они, не отражают всех взаимосвязей между аспектами информационного воздействия, что еще раз подтверждает новизну данного вопроса и предоставляет возможность для дальнейшего изучения.

Информационное воздействие и особенности его ведения

§1.1 Содержательное описание информационного воздействия

Как говорил Н.М. Ротшильд: «Кто владеет информацией, тот владеет миром», но было бы лучше сделать небольшую ремарку: « Кто владеет информацией и правильно её использует, тот владеет миром». Действительно, с этим утверждением трудно не согласиться. На сегодняшний день человек, владеющий информацией, может использовать её и дистанционно воздействовать на сознания других людей, прибегнув к такому инструменту её распространения как интернет.

Интернет - всемирная система объединённых компьютерных сетей для передачи информации и хранения. В этой работе пойдёт речь об информационном воздействии в среде социальных сетей. Понятие «социальная сеть» появилось еще до появления Интернета и соответственно до современных интернет-сетей. Данный термин был введен в 1954 году британским социологом Джеймсом Барнсом. Этот термин стал применяться во второй половине 20 века в европейских странах при исследованиях человеческих отношений и социальных связей и стал общеупотребительным. Спустя некоторое время в социальной сети стали рассматривать не только людей, как представителей общества, но и прочие объекты, которые могут иметь различные социальные связи, например, сайты, фирмы, города, и т.д. Но стоит отметить, что в данной работе мы рассматриваем социальные сети только в рамках сети Интернет и в этих рамках это понятие является довольно новым.

Дадим определение социальным сетям. Социальная сеть - онлайн-сервис, сайт, разработанный с целью собрать людей с общими интересами, возможностью пользователей самостоятельно публиковать и обмениваться информацией о себе и общаться между собой, а также выкладывать

различные медиа файлы. Если упростить определение, то социальная сеть – это набор узлов, например, организации, люди, государственные образования или веб-страницы. Также это отношения или связи между этими узлами. Каждое отношение соединяет несколько узлов. Почтовая сеть, например, является направленной сетью отправителей и получателей. Социальная сеть, организованная с помощью программного обеспечения, обычно является ненаправленной сетью пользователей. Рассмотрим основные типы социальных сетей.

1. Цельные сети – это отношения людей в пределах определённо очерченных границ. Примеры подобных объектов, использующихся в онлайн-исследованиях: люди, работающие в одном офисе, которые связаны локальными сетями, электронная почта, списки рассылки, социальные сети. Цельные сети – наиболее часто используемые объекты в анализе социальных сетей. Однако данные объекты не всегда могут быть использованы. Они могут меняться в зависимости от практических требований исследователя; сбор информации о сети размером с офис не является особенно трудной задачей, но получение действительного списка всех пользователей социальной сети фактически невозможен, поскольку этот список динамичен и будет, меняется во время сбора данных.

2. Эго-сети. Анализ эго-сетей предназначен для определения разницы в размерах, формах и качестве разнообразных эго-сетей. Эго-сети, в основном, это спонтанная выборка пользователей. При этом всегда стремятся к представительности даже временных таких выборок. Каждый элемент выборки в подобном анализе обозначается как «эго», а узлы, связанные с эго, обозначаются как «другие». Мы можем собрать данные или звездообразной сети (т.е. эго-узел и его связи с другими узлами), или полной эго-сети, включающей связи других узлов друг с другом.

Можно обнаружить эго-сеть в социальных сетях интернета. В таком случае это будет список, например, список друзей. Далее, в ходе анализа списка можно будет понять, кто к кому привязан.

3. Неполные сети – это выборка из относительных данных, созданная с помощью метода «снежного кома». Неполные сети – это компромисс между желанием охватить цельную сеть и тем, что некоторые цельные сети слишком большие, чтобы их охватить целиком. Можно начать с нескольких страниц («мини-сеть») или одной-единственной веб-страницы, далее будем искать страницы, связанные с этой мини-сетью, а затем страницы, связанные с этими страницами. Процесс организации выборки заканчивается тогда, когда собрано достаточное количество страниц или когда выборка соответствует определенному критерию (например, когда собраны все страницы с более чем 500 датами), или когда собраны все возможные страницы.

Неполные сети являются практичным решением для анализа массивной совокупности данных, расположенных в сети интернет. Невозможно собрать сведения обо всех пользователях, или блогах, можно выстроить сеть отношений, соединяющую эго-сети нескольких пользователей.

Теперь выделим свойства, присущие социальным сетям:

- сайт есть автоматизированная среда, в рамках которой пользователи могут создавать связи с другими пользователями (социальные связи) или социальные объекты (тематические группы);
- содержание или контент сайта (создается преимущественно его пользователями);
- пользователям доступны функции коммуникации с другими пользователями и социальными объектами;
- пользователи могут получать статическую и динамическую информацию об объектах, которые существуют в данной социальной среде и о социальных

связях между ними;

- разная степень подверженности членов социальной сети влиянию;
- существование косвенного влияния в цепочке социальных контактов;
- существование «лидеров мнений»;
- наличие собственных мнений пользователей;
- изменение мнений под влиянием других членов социальной сети;
- различная значимость мнений (влиятельности, доверия) одних пользователей для других;
- локализация групп («по интересам», с близкими мнениями);
- учет факторов «социальной корреляции»;
- существование порога чувствительности к изменению мнения окружающих;
- существование внешних факторов влияния (реклама, маркетинговые акции) и, соответственно, внешней стороны (средства массовой информации, производители товаров и т.п.);
- воздействие структурных свойств социальных сетей на динамику мнений;
- возможность образования коалиций;
- игровое взаимодействие пользователей;
- наличие лавинообразных эффектов;
- информационное управление в социальных сетях.

Среди функций социальной сети следует выделить следующие:

1. Коммуникационная. В рамках коммуникационной функции люди устанавливают контакты, обмениваются новостями, информацией (медиа файлы, комментарии, сообщения и ссылки на сайты), кооперируются для достижения совместных целей (создание и удержание социальных связей).

2. Информационная. Поток информации имеет двухстороннюю направленность, потому что участники общения выступают как в роли «передатчика», так и в роли получателя информации.

3. Социализирующая. Саморазвитие, самоанализ в системе «групп» и «друзей».

4. Актуализирующая. Самопрезентация пользователей сети.

5. Идентификационная. При создании индивидуального профиля пользователь наполняет его информацией о себе – ФИО, дата рождения, школа, ВУЗ, интересы семейный статус, и т.д. Это позволяет проводить поиск пользователей по заданным параметрам.

6. Функция формирования идентичности. Человек склонен сравнивать себя с людьми, у которых с ним есть большое количество схожих черт. Более того схожие люди положительно оценивают друг друга. Это такой механизм, с помощью которого человек может научиться формулировать своё мнение о других людях или группах.

7. Развлекательная. Социальные сети позволяют обмениваться мультимедийными файлами, более того, стоит отметить роль виджетов – это мини-программы развлекательного характера, которые создаются сторонними лицами для расширения возможностей пользователя, таких как игры, медиа-приложения и прочие.

Выше, даны основные параметры социальных сетей. Они понадобятся для того, чтобы составить информационно-логическую модель. Суть, которой будет заключаться в следующем: она будет давать количественно-вероятностную оценку распространения информации, которая была опубликована в пределах данной социальной сети. Может случиться и так, что опубликованная информация станет распространяться по принципу «снежного кома», в результате, эту новость станут распространять за пределами социальной сети: теле, радио-эфир, газеты и т.д. Однако нас будут интересовать только, те события, реакции людей, «лидеров мнений», которые будут происходить в пределах социальной сети.

Информационное воздействие в социальных сетях может осуществляться через публичные страницы («паблики»), закрытые группы (группа, доступ в которую осуществляет её администратор), известных личностей («лидеров мнений»), блоггеров и отдельный информационный посыл может быть запущен с помощью флэш-моба с определенным хэш-

тэгом. Вопрос выбора распространения зависит от цели информационного посылы и желания быстро распространить его на определенную аудиторию.

В данной работе для анализа выберем третий тип социальной сети. Неполные сети наилучшим образом подходят для эффективного изучения информационного воздействия в выбранной предметной области.

§1.2 Формализация предметной области

Перейдём к формализации предметной области, выделению основных сущностей и определению их взаимосвязей. Данный подход приведёт к построению информационно-логической модели.

Имеется социальная сеть-это набор узлов, например, таких как люди, организации, веб-страницы или государственные образования. Также это отношения или связи между этими узлами. Каждое отношение соединяет несколько узлов. Будет выделена некоторая структура, отражающая картину произведенного воздействия. Затем необходимо определить и выбрать ряд параметров, чтобы на следующем этапе выбрать путь воздействия. Это такие параметры как: основная тема информационного посыла, целевая аудитория на которую производится воздействие, срок распространения информации. Зная вышеуказанные параметры можно выбрать несколько информационных каналов: «лидеры мнений», «паблики», закрытые группы, популярные блоггеры, хэш-тэг. И, как итог, надо выбрать формат информационного посыла: видео, текст или текст с приложенными медиа-файлами. Мы не рассматриваем в данной работе «вирусную рекламу» или новость, которая как снежный ком набирает «лайки, комментарии и репосты», потому как при распространении подобных новостей нет нужды ни в одном из вышеперечисленных параметров. Скорость распространения подобных новостей зависит лишь от того, как хорошо над ней поработали психологи и маркетологи.

Рассмотрим каждую сущность информационно-логической модели более подробно.

1.Выбор типа сети.

На данном этапе, исходя из целей воздействия, мы выбираем тип сети, и, следовательно, дальнейший принцип работы с распространяемой информацией. Рассматривается самая распространённая и популярная в

России и странах СНГ² [9] социальная сеть – «ВКонтакте». Здесь необходимо использовать неполные сети, т.к. исходя из определения, можно сделать довольно большую выборку пользователей и при этом проделав такую процедуру несколько раз, можно получить эффект сравнимый с тем, при котором рассматривали полные сети. Полная сеть – множество объединённых неполных сетей.

2. Основные параметры.

Основных параметров в модели три: тема информационного посыла, срок распространения информации и целевая аудитория.

Дадим краткую характеристику каждому из параметров.

а) Тема информационного посыла.

Здесь стоит отметить: есть темы, которые актуальны на данный момент и, соответственно, любая информация, связанная с этими темами будет распространяться быстрее, за счёт заинтересованности самих пользователей в её популяризации вне зависимости от возраста, пола и хобби. Есть также темы, которыми интересуются определённые группы лиц. Градация этих групп может быть абсолютно разной. Например: по возрасту, по полу, по сфере деятельности, по увлечению чем-либо и т.д.

б) Срок распространения информации.

Данный параметр зависит от темы информационного посыла и целевой аудитории. Здесь зависимость очевидна: неактуальную информацию для данной возрастной категории будет сложно распространить в короткий срок.

с) Целевая аудитория.

Необходимо определиться с оптимальным форматом того, что мы хотим распространить. Например: люди в возрасте от 12-18 с большей вероятностью обратят внимание на пост, в котором будет немного текста, несколько фотографий и видеофайл.

² Содружество Независимых Государств (СНГ) — международная организация (международный договор), призванная регулировать отношения сотрудничества между государствами, ранее входившими в состав СССР (не всеми). СНГ не является надгосударственным образованием и функционирует на добровольной основе.

3. Способ распространения информации.

В данной модели есть вложенная зависимость, то есть последующий пункт зависит от предыдущего. Здесь речь пойдет о выборе информационного канала, но он необязательно должен быть единственным. Будем рассматривать «сетевых титанов» по посещаемости и разнообразию объединения людей по группам интересов это: «лидеры мнений», «паблики», закрытые группы, популярные блоггеры и особый вид распространения посылы - хэш-тэг. Дадим краткое описание каждому термину в отдельности.

а) «Лидеры мнений»

Это пользователи, которые имеют «вес» среди групп пользователей и могут повлиять на скорость распространения информации. Градация может быть по возрасту, по роду деятельности и увлечению.

б) «Паблики»

Это открытые группы по интересам, которые, как правило, имеют большое количество подписчиков по сравнению с закрытыми группами.

в) Закрытые группы

Это сообщество с ограниченным доступом и, как правило, меньшим количеством подписчиков.

г) Популярные блоггеры

Это распространенное явление в Интернете, в качестве блоггера может выступать как обычный человек, так и широко известная личность.

Блог (англ. blog, от weblog – это интернет-журнал событий, интернет-дневник, онлайн-дневник) - веб-сайт, основное содержимое которого - регулярно добавляемые записи, содержащие текст, изображения или мультимедиа.

д) Хэш-тэг

Суть хэш-тэгов в том, что если помимо вас в данной социальной сети пользователи опубликовали на своей личной странице запись, прокомментировали чью-то медиа-запись, оставили отзыв в «паблике» и т.д. с таким же хэш-тэгом, то вы можете, нажав, например, на #олимпиада2018,

перейти ко всем записям, к любому человеку на страницу, к любому посту, где есть такой же хэш-тэг. Сейчас многие медийные личности, новостные ленты призывают людей публиковать записи с определенными хэш-тэгами, чтобы быстрее распространить информацию.

4. Формат информационного посыла.

Здесь выбирается тот формат. Его выбор зависит от целевой аудитории, темы, информационного канала. Может быть в виде: видео, ссылки на внешний источник, текста или текста с медиа-файлами.

В данном параграфе была рассмотрена и формализована предметная область и выделены основные сущности модели. Далее будет составлена информационно-логическая модель.

§1.3 Информационно-логическая модель

Для проведения имитационного моделирования необходимо построить информационно-логическую модель рассматриваемой предметной области и на её основе построить теоретико-вероятностную математическую модель, связывающую взаимодействие элементов информационно-логической модели.

Иерархия модели:

1. Тип сети
 - a) Полные
 - b) Эго-сети
 - c) Неполные
2. Основные параметры
 - a) Тема
 - b) Срок
 - c) Целевая аудитория
3. Информационные каналы
 - a) «Лидеры мнений»
 - b) «Публики»
 - c) Закрытые сообщества
 - d) Блоггеры
 - e) Хэш-тэг
4. Формат публикуемой информации
 - a) Текст с приложенными медиа-файлами
 - b) Текст
 - c) Видео файлы

Поскольку система является сложной [10,11], все четыре параметра связаны между собой и представляют структуру воздействия на пользователей сети.

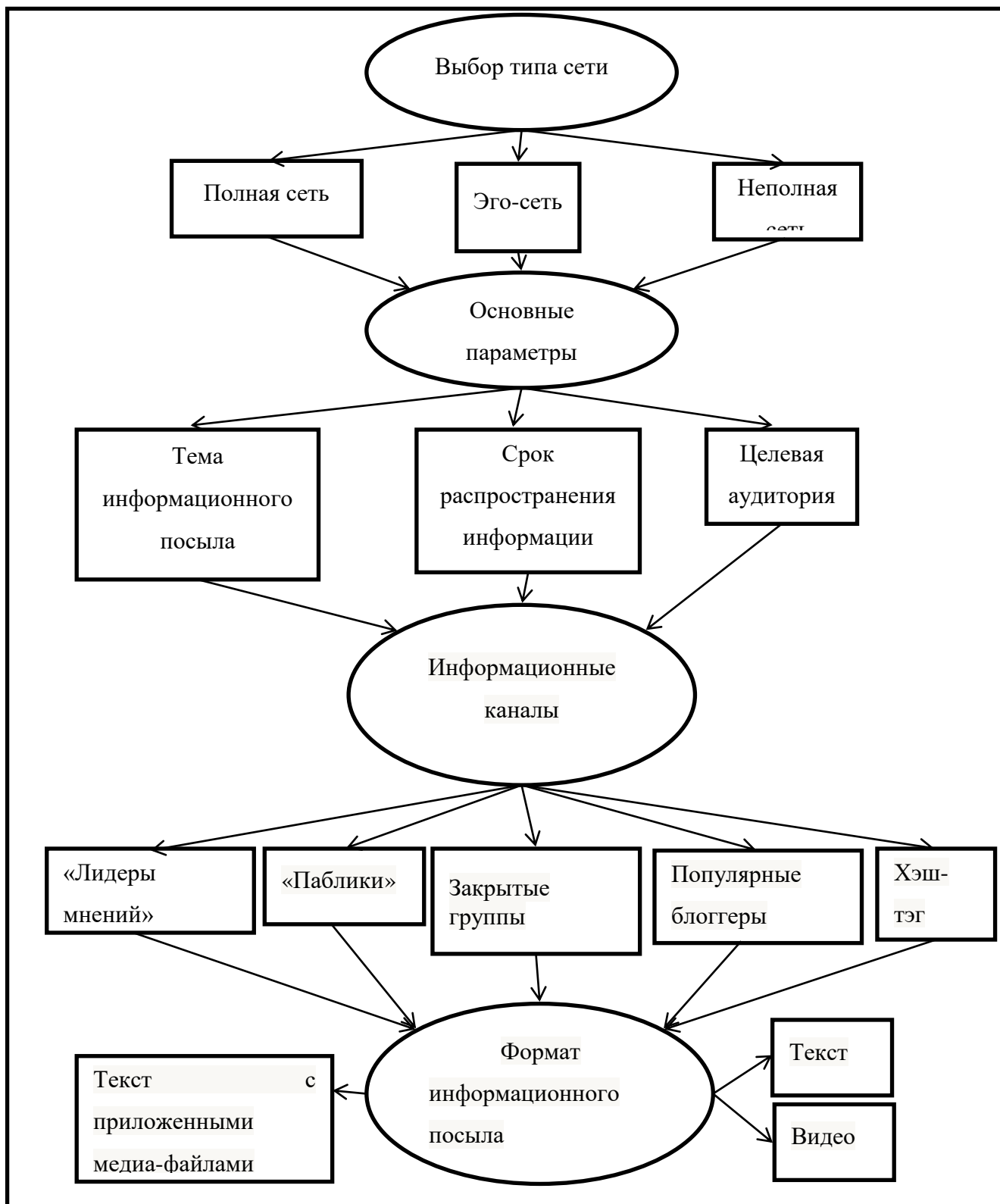


Схема.1 Информационно-логическая модель распространения информации в социальных сетях

Таким образом, определив предметную область, проведя её анализ, выделив и описав основные сущности, была построена информационно-логическая модель (см. схему 1).

Математическая модель информационного воздействия

§2.1 Вероятностно-статистический анализ и задание пространства параметров

Логическая часть модели была дана выше, здесь речь пойдёт о математической её части и вероятностно-статистическом анализе. Выбор такого подхода к построению модели обусловлен тем, что необходимо будет выбирать несколько информационных каналов для распространения информации, просчитывать и оценивать их одновременное использование. Социальная сеть «ВКонтакте» предоставляет возможность уже на первом этапе выбора сообщества отсеять некоторые из них и выбрать наиболее подходящие. Но, после того как будут применены представленные на рисунке 1 фильтры, всё равно останутся несколько сообществ.

Бюджет 15000	Тематика Выберите тематику	Дневной охват, от Введите дневной охват	Участников, от Введите кол-во участников
Регион Выберите страну	Пол Не имеет значения	Возраст Выберите возраст	Охват записи, от Введите охват записи
Диапазон времени размещения 18 Мая 2018		22 : 40	19 Мая 2018

Рис.1 Встроенный фильтр сообществ

И на данном этапе выбора необходимо анализировать предоставленную статистику (см. рис.2, рис.3, рис.4, рис.5). Кроме статистики, некоторые сообщества оставляют комментарии о размещении рекламной записи (см. рис.1).

	ФЕЛОСАФЫ Юмор		49 000 участников	9 000 / 50 000 человек	368 руб.	Исключить
	Кино клуб • Новые фильмы... Кино		370 000	13 000 / 130 000	413 руб.	Исключить
	Факты Наука и техника				1 106 руб.	Исключить
	Леди Ночь Дом и семья				177 руб.	Исключить

Взрослая, платежеспособная аудитория.
Отличное сообщество, подходит для размещения Вашей рекламы, эффект гарантируем!

Рис.2 Статистика сообществ

Охват аудитории

На этом графике учитываются пользователи, просмотревшие записи сообщества на стене или в разделе Мои Новости.

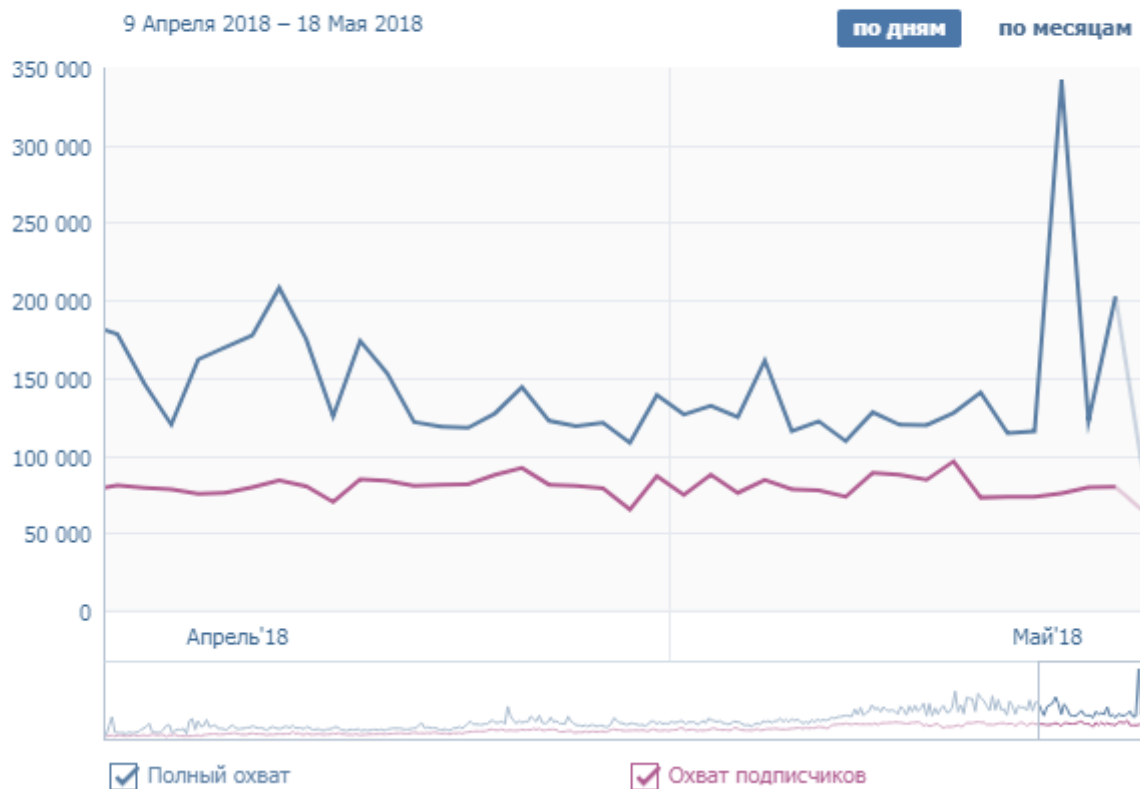


Рис.3 Статистика сообществ (охват аудитории)

Обратная связь

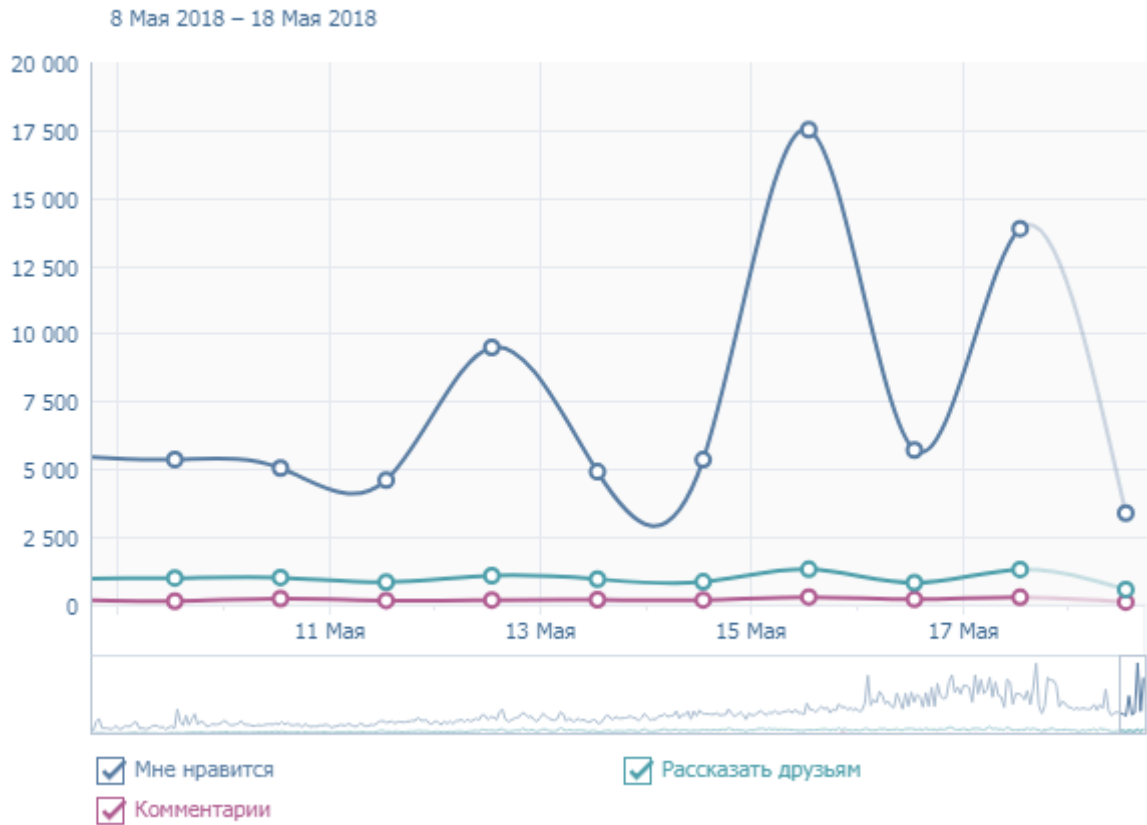


Рис.4 Статистика сообществ (обратная связь)

Пол / Возраст

Показать в виде графиков

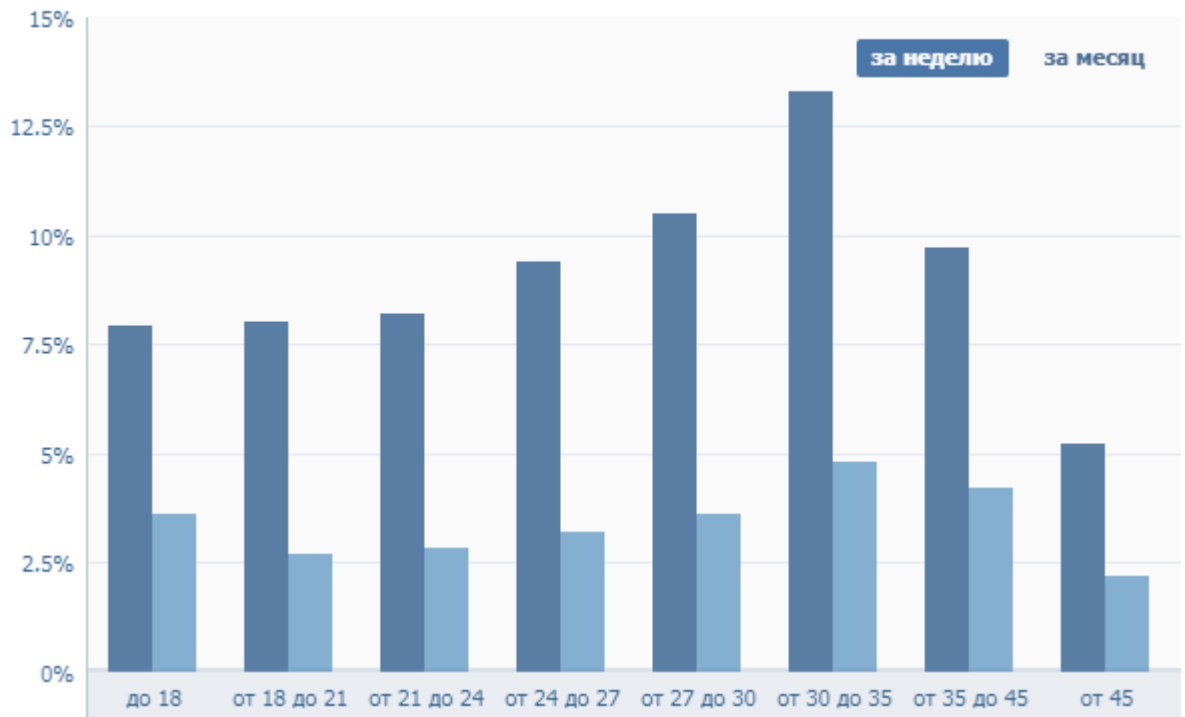


Рис.5 Статистика сообществ (пол/возраст)

Исходя из статистики, находящейся в открытом доступе, выделим следующие статистические показатели социальной сети (рассматриваются «лидеры мнений», блоггеры, «паблики», закрытые группы):

- a) α_1 – количество подписчиков (участников);
- b) α_2 – полный охват пользователей (количество пользователей, посетивших сообщество за месяц в среднем);
- c) β_1 – охват записи данного канала (среднее количество подписчиков, просматривающих запись сообщества);
- d) β_2 – дневной охват (среднесуточное количество пользователей, просматривающих записи сообщества);
- e) пол / возраст аудитории данного канала;
- f) l – количество лайков;
- g) r – количество репостов;
- h) c – количество комментариев.

Необходимость запуска хэш-тэгов является ситуативной составляющей, которая зависит от условий конкретной задачи.

Таким образом, использование имитационного моделирования позволяет получить вероятности воздействия на пользователей социальной сети, а также при дальнейшем исследовании оценивать риски при вложении денежных средств в рекламу в сообществах. Физическое или юридическое лицо, которое захочет распространить информацию в пределах социальной сети «ВКонтакте» сможет заранее знать какой его ожидает результат с определённой вероятностью, а также возможный риск при вложении финансов. Использование данной модели в разы упростит выбор сообществ для распространения информации.

§2.2 Математическая модель

В данном пункте будут определены: математическая модель и реляционная база данных.

Основываясь на вероятностно-статистическом анализе [12], вводим случайные величины $\xi_1, \xi_2, \dots, \xi_n$ и соответствующие вероятности:

- $k = \frac{\beta_1}{\alpha_1}$ - вероятность просмотра записи подписчиками,
- $q = \frac{\beta_2}{\alpha_2}$ - вероятность просмотра записи пользователями,
- $\tau_1 = \frac{l}{\alpha_2}, \tau_2 = \frac{r}{\alpha_2}, \tau_3 = \frac{c}{\alpha_2}$ - вероятности того, что поставят «лайк», репостнут и прокомментируют запись в данном канале.
- Возраст будет разбит на группы: $t = (18, 21, 24, 27, 30, 35, 45)$, то есть $t < t_1$: количество (%) м/ж в этом отрезке, $t_1 < t < t_2$ – количество (%) м/ж в этом отрезке, ... , $t > t_7$: количество (%) м/ж в этом отрезке.

После того, как были определены вероятности, вернёмся к выбору информационных каналов. Выбрав несколько информационных каналов, каждый из которых характеризуется дискретными случайными величинами [13], необходимо:

1. Разбить их значения на интервалы $(\bar{x}_i \pm j * \sigma_j)$, где \bar{x}_i – среднее значение случайной величины, σ_j – её среднеквадратическое отклонение, $i = \overline{1, n}, j = \overline{1, 3}$.
2. Затем для каждой случайной величины будет построена таблица с интервалами и вероятностями попадания в этот интервал.

ξ :

x_i	$(\bar{x}_i - 3 * \sigma_j; \bar{x}_i - 2 * \sigma_j)$	$(\bar{x}_i - 2 * \sigma_j; \bar{x}_i - \sigma_j)$	$(\bar{x}_i - \sigma_j; \bar{x}_i)$	$(\bar{x}_i; \bar{x}_i + \sigma_j)$	$(\bar{x}_i + \sigma_j; \bar{x}_i + 2 * \sigma_j)$	$(\bar{x}_i + 2 * \sigma_j; \bar{x}_i + 3 * \sigma_j)$
p_i	p_1	p_2	p_3	p_4	p_5	p_6

$$\sum_{i=1}^n p_i = 1$$

3. Далее производится розыгрыш случайной величины от 0 до 1 для каждого из выбранных сообществ и определяется в какой интервал она попала.
4. $x_i = s + \sigma_j * (1 - r)$, x_i – значение случайной величины в пределах интервала, который был определён в пункте 3, r – случайная величина от 0 до 1.
5. На каждом шаге поинтервально суммируются значения случайных величин во всех сообществах ($z_i = x_i + y_i + \dots + v_i$).
6. Повторение пунктов 3,4,5 n – раз.
7. Вычисляются интервалы при совместном использовании сообществ, и подсчитывается количество значений, которые оказались в каждом из интервалов.
8. Вычисляется вероятность попадания в каждый интервал при совместном использовании сообществ.
9. Повторение пунктов 3,4,5,6,7,8 m –раз и усреднение результатов.
10. После чего возможны различные комбинации информационных каналов. То есть, предоставляется возможность оценить, например, с определённой вероятностью предполагаемое количество просмотров опубликованной записи среди подписчиков при выборе двух сообществ и т.д.

Данный подход позволит комбинировать сразу несколько информационных каналов и выбрать наиболее подходящий вариант распространения новости. Соответственно, все вышеперечисленные коэффициенты и величины в имитационной модели будут меняться в зависимости от выбора информационного канала, так как у каждого из них будет своя статистика. Поэтому необходимо создать реляционную базу данных для хранения и использования данной статистики.

Для создания используется Microsoft Office Access. База данных должна содержать информацию:

1. Об основных статистических данных о сообществах
2. О стоимости публикаций записи
3. О результатах публикаций

Для эффективного выполнения запросов, необходимо предопределить связи между набором данных. Чтобы построить схему базы данных надо описать свойства полей для каждой таблицы.

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Сообщество	Да	Да	Да	Текст.			
Тематика	Да	Да	Да	Текст.			
Регион			Да	Текст.			
Город			Да	Текст.			

Таблица 2.1 Описание свойств полей таблицы «Данная статистика (территориальное расположение)»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Дата			Да	Дата/время			
Кол-во комментариев			Да	Числов.	Целое		

Таблица 2.2 Описание свойств полей таблицы «Данная статистика с количеством комментариев»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Сообщество		Да	Да	Текст.			
Тематика			Да	Текст.			
Кол-во подписчиков			Да	Числов.	Целое		
Охват записи			Да	Числов.	Целое		
Дневной охват записи			Да	Числов.	Целое		
Кол-во лайков			Да	Числов.	Целое		
Кол-во репостов			Да	Числов.	Целое		
Кол-во комментариев			Да	Числов.	Целое		

Таблица 2.3 Описание свойств полей таблицы «Данная статистика по сообществам информативная таблица»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщения		Да	Да	Числов.	Целое		
Дата			Да	Дата/время			
Кол-во лайков			Да	Числов.	Целое		

Таблица 2.4 Описание свойств полей таблицы «Данная статистика по сообществам с количеством лайков»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщения		Да	Да	Числов.	Целое		
Дата			Да	Дата/время			
Кол-во подписчиков			Да	Числов.	Целое		

Таблица 2.5 Описание свойств полей таблицы «Данная статистика с количеством подписчиков»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Дата			Да	Дата/время			
Кол-во репостов			Да	Числов.	Целое		

Таблица 2.6 Описание свойств полей таблицы «Данная статистика с количеством репостов»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Дата			Да	Дата/время			
Кол-во пользователей			Да	Числов.	Целое		

Таблица 2.7 Описание свойств полей таблицы «Данная статистика с количеством пользователей»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		

Номер сообщества		Да	Да	Числов.	Целое		
Период (неделя, месяц)			Да	Текст.			
до 18			Да	Числов.	С плав. точкой (1)	2	
18-21			Да			2	
21-24			Да			2	
24-27			Да			2	
27-30			Да			2	
30-35			Да			2	
35-45			Да			2	
от 45			Да			2	

Таблица 2.8 Описание свойств полей таблицы «Данная статистика (Пол М) и (Пол Ж)»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Период (неделя, месяц)			Да	Текст.			
до 18			Да			2	
18-21			Да			2	
21-24			Да			2	

24-27			Да	Числов.	С плав. точкой (1)	2	
27-30			Да			2	
30-35			Да			2	
35-45			Да			2	
от 45			Да			2	

Таблица 2.9 Описание свойств полей таблицы «Обратная связь (Пол М) и (Пол Ж)»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Суммарный охват			Да	Числов.	Целое		
Охват подписчиков			Да	Числов.	Целое		
Кол-во переходов по ссылкам			Да	Числов.	Целое		
Кол-во лайков			Да	Числов.	Целое		
Кол-во репостов			Да	Числов.	Целое		
Кол-во комментариев			Да	Числов.	Целое		
Кол-во			Да	Числов.	Целое		

вступивших							
Кол-во вступивших			Да	Числов.	Целое		

Таблица 2.10 Описание свойств полей таблицы «Обратная связь (общая статистика)»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Кол-во подписчиков			Да	Числов.	Целое		
Бюджет			Да	Денежный			
Дата и время публикации			Да	Дата/время			
Дата и время выполнения			Да	Дата/время			

Таблица 2.11 Описание свойств полей таблицы «Обратная связь (с количеством подписчиков)»

Имя поля	Ключевое поле	Уникальное поле	Обязательное поле	Тип данных	Размер данных	Число десятичных знаков	Подпись поля
Номер	Да	Да	Да	Числов.	Целое		
Номер сообщества		Да	Да	Числов.	Целое		
Вероятность просмотра записи подписчиками			Да	Числов.	С плавающей точкой (1)	4	
Вероятность просмотра записи пользователями			Да	Числов.	С плавающей точкой (1)	4	
Вероятность "лайка"			Да	Числов.	С плавающей точкой (1)	4	
Вероятность репоста			Да	Числов.	С плавающей точкой (1)	4	
Вероятность комментария				Числов.	С плавающей точкой (1)	4	

Таблица 2.12 Описание свойств полей таблицы «Вероятности»

После того как были описаны свойства таблиц базы данных, необходимо определить способы связывания данных в таблицах. Схема

данных отображает логические связи таблиц базы данных, которые автоматически используются для объединения записей связанных таблиц (см. схему рис.6).

Она отражает связи, которые введены между объектами. Таким образом, можно наблюдать как будет выполняться запрос.

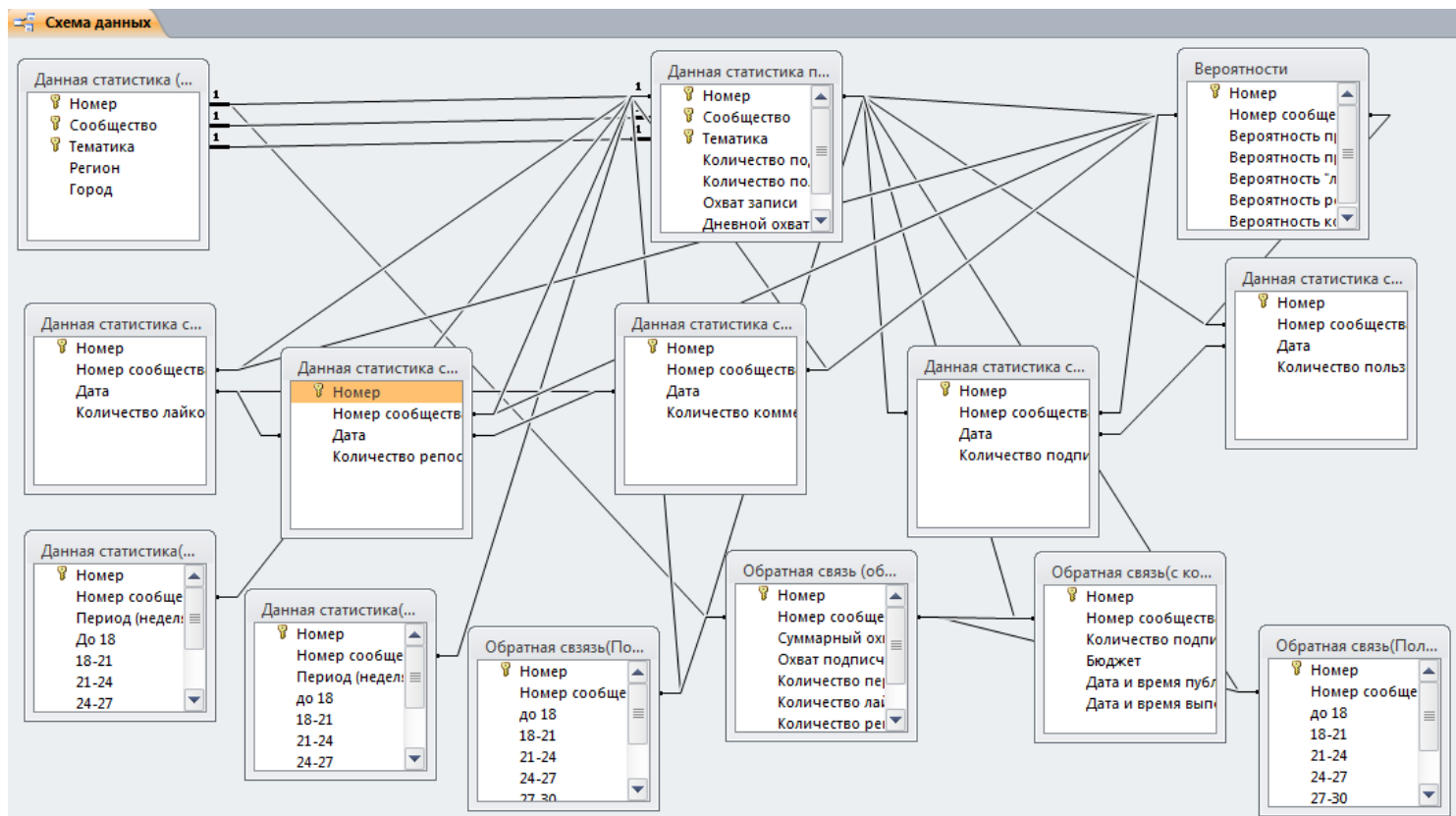


Рис.6 Схема данных

Таким образом, была сделана реляционная база данных для хранения и обработки статистических данных и составлена математическая модель.

§2.3 Имитационное моделирование

Имитационное моделирование - это совокупность методов алгоритмизации функционирования объектов исследований, программной реализации алгоритмических описаний, организации, планирования и выполнения на ЭВМ вычислительных экспериментов с математическими моделями, имитирующими функционирование сложных систем в течение заданного периода. Под алгоритмизацией функционирования сложных систем понимается пооперационное описание работы всех ее функциональных подсистем отдельных модулей с уровнем детализации, соответствующим комплексу требований к модели.

Термин "имитационное моделирование" означает, что мы имеем дело с такими математическими моделями, с помощью которых нельзя заранее вычислить или предсказать поведение системы, а для предсказания поведения системы необходим вычислительный эксперимент (имитация) на математической модели при заданных исходных данных.

Имитационные модели обладают следующими достоинствами:

1. Возможностью описания поведения основных элементов сложных систем с высоким уровнем детализации;
2. Отсутствием ограничений на состояние внешней среды рассматриваемой сложной системы;
3. Возможностью исследования поведения сложной системы в пространстве параметров, характеризующих функционирование данной системы;
4. Возможностью решения задач по оптимальному функционированию и управлению поведением сложных систем.

Одним из видов имитационного моделирования является статистическое имитационное моделирование, позволяющее воспроизводить на ЭВМ функционирование сложных случайных процессов. При исследовании сложных систем, подверженных случайным возмущениям

используются вероятностные аналитические модели и вероятностные имитационные модели. В вероятностных аналитических моделях влияние случайных факторов учитывается с помощью задания вероятностных характеристик случайных процессов (законы распределения вероятностей, спектральные плотности или корреляционные функции). При этом построение вероятностных аналитических моделей представляет собой сложную вычислительную задачу.

В вероятностном имитационном моделировании оперируют не с характеристиками случайных процессов, а с конкретными случайными числовыми значениями параметров. При этом результаты, полученные при воспроизведении на имитационной модели рассматриваемого процесса, являются случайными реализациями. Поэтому для нахождения объективных и устойчивых характеристик процесса требуется его многократное воспроизведение, с последующей статистической обработкой полученных данных. Именно поэтому исследование сложных процессов и систем, подверженных случайным возмущениям, с помощью имитационного моделирования принято называть статистическим моделированием.

Таким образом, под статистическим моделированием будем понимать алгоритм, с помощью которого имитируют поведение сложной системы, подверженной случайным возмущениям, имитируя при этом взаимодействие элементов системы, носящих вероятностный характер.

При реализации на ЭВМ статистического имитационного моделирования возникает задача получения на ЭВМ случайных числовых последовательностей с заданными вероятностными характеристиками. Численный метод, решающий задачу генерирования последовательности случайных чисел с заданными законами распределения получил название "метода статистических испытаний" или "метода Монте-Карло".

Статистическое моделирование - это способ изучения сложных процессов и систем, подверженных случайным возмущениям, с помощью имитационных моделей. Метод Монте-Карло – это численный метод,

моделирующий на ЭВМ псевдослучайные числовые последовательности с заданными вероятностными характеристиками. Методика статистического моделирования состоит из следующих этапов:

1. Моделирование на ЭВМ псевдослучайных последовательностей с заданными законами распределения вероятностей (метод Монте-Карло), имитирующих на ЭВМ случайные значения параметров при каждом испытании;

2. Использование полученных числовых последовательностей в имитационных математических моделях;

3. Статистическая обработка результатов моделирования.

Итак, опишем процесс выбора сообществ.

Алгоритм

1. Применение встроенного сервиса «Реклама ВКонтакте», описанного в главе 2 параграф 2.1.
2. Выбор сообществ будет производиться в зависимости от целей задачи по соответствующим статистическим показателям.
3. Выделение случайных величин.
4. Составление таблиц распределения для каждой из случайных величин.
5. Применение итерационного алгоритма, описанного в параграфе 2.2.
6. Публикация информации в выбранных сообществах.
7. Анализ результатов.

Данный подход позволит изменять пространство и значения параметров, тем самым для одной задачи может получиться несколько различных подходов к её решению. Это позволит моделировать несколько сценариев распространения информации, смотреть какой подходит больше или же комбинировать их для увеличения «прибыли» или уменьшения затрат. Изменив параметр в одном решении, мы можем вычислить, как это повлияет на другое решение, какой в итоге будет результат и с какой вероятностью его можно достичь.

Имитационное моделирование на примере социальной сети «ВКонтакте»

§3.1 Сбор данных и выбор параметров в социальной сети «ВКонтакте»

В данном пункте речь пойдёт о сборе данных и выборе параметров социальной сети с использованием базы данных для решения поставленной задачи.

Необходимо провести анализ статистических данных для соответствующих параметров. В самом общем случае отбор сообществ может происходить как по одному, так и по нескольким статистическим показателям:

- Регион
- Город
- Количество подписчиков (как общее количество, так и подписчиков просмотревших записи сообщества на стене или в разделе Новости в динамике);
- Количество пользователей (как среднее количество в месяц, так и пользователи, просмотревшие записи сообщества на стене или в разделе Новости в динамике);
- Охват записи
- Дневной охват
- Бюджет
- Количество «лайков», репостов, комментариев (в динамике)
- Процентное соотношение пол\возраст (в динамике)
- Тематика

Используя общедоступную статистику, можно выделить периоды активности пользователей по дням недели, месяцам, временам года, определить их активность\неактивность, выбрать сообщества с

необходимой целевой аудиторией (пол\возраст) и тематикой и т.д.

Задача заключается в том, чтобы распространить информацию о появлении нового спортивного питания в Санкт-Петербурге и Москве с бюджетом до 500 рублей включительно и как можно больше людей увидели запись об этом спортивном питании.

Для этого было выбрано два сообщества, отбор которых по 5 показателям: регион, город, тематика, бюджет, охват записи (см. рис.7).

Номер	Сообщества	Количество подписчиков	Количество пользователей	Бюджет	Регион	Город	Охват записи	Тематика
	Рубрика спорт	100000	295580	199,00р.	Россия	Санкт-Петерб	4500	Спорт
2	Новости спорт	110000	253303	159,00р.	Россия	Москва	7000	Спорт
*	(№)							

Рис.7 Запрос по базе данных

Выбор этих сообществ обоснован следующими факторами (см. рис.8):

- Территориальное расположение
- Ограничение в бюджете
- Тематика распространяемой информации и сообществ
- Значение охвата записи

Бюджет	Регион	Город	Охват записи	Тематика
Данная статистика по	Данная статистика по	Данная статистика по	Данная статистика по	Данная статистика по
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
< = 250	"Россия"	"Москва"	> = 4000	"Спорт"
< = 500		"Санкт-Петербург"		

Рис.8 Ограничения в запросе по базе данных

Исходя условий задачи и с использованием базы данных, был выполнен запрос и выбраны два сообщества по соответствующим показателям.

§3.2 Проведение имитационного моделирования

В предыдущем параграфе была сформулирована задача, выбрано два сообщества и дискретная случайная величина: охват записи. Составим для них таблицы [14].

$$\xi_1: \bar{x}_1 = 4500, \sigma_1 = 500$$

x_i	(3000,3500)	(3500,4000)	(4000,4500)	(4500,5000)	(5000,5500)	(5500,6000)
p_i	0,05	0,15	0,35	0,35	0,08	0,02

Таблица 3.1 Таблица распределения случайной величины для первого сообщества

$$\xi_2: \bar{x}_2 = 7000, \sigma_2 = 700$$

y_i	(4900,5600)	(5600,6300)	(6300,7000)	(7000,7700)	(7700,8400)	(8400,9100)
q_i	0,05	0,15	0,4	0,3	0,07	0,03

Таблица 3.2 Таблица распределения случайной величины для второго сообщества

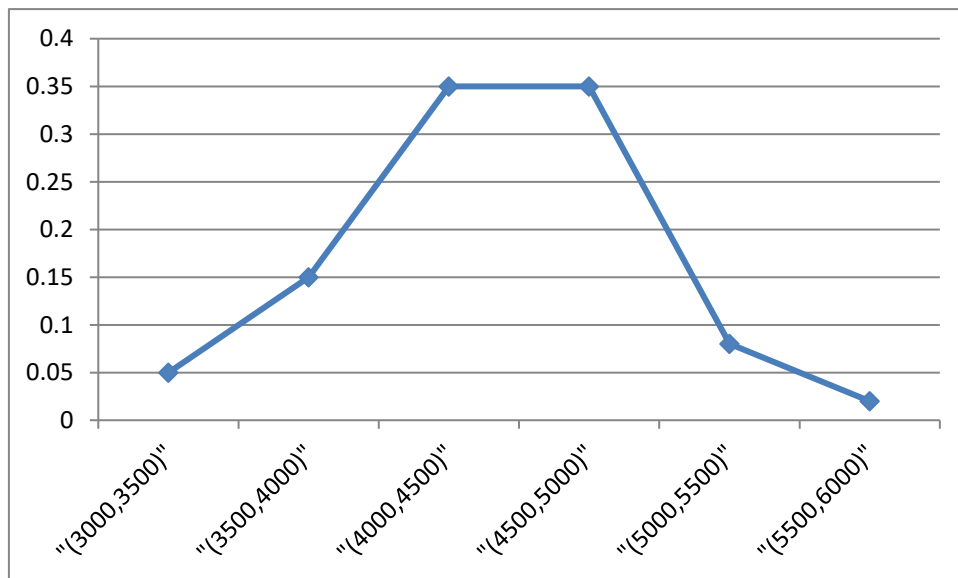


График 3.1 График распределения вероятностей для первого сообщества

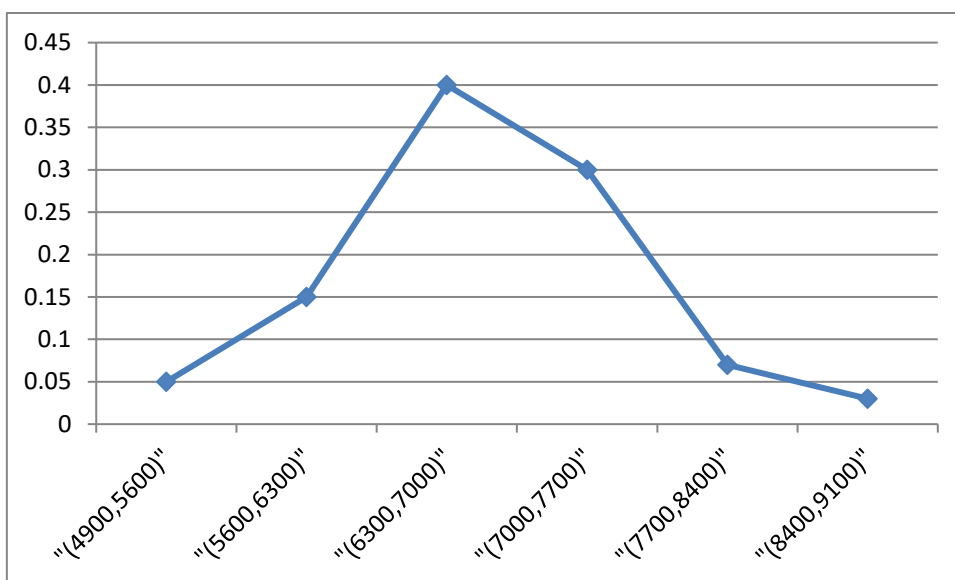


График 3.2 График распределения вероятностей для второго сообщества

Будем исходить из того, что \bar{x}_i – это среднее значение случайной величины, с математической точки зрения и среднее количество подписчиков, просматривающих записи сообщества за определенный период времени, с точки зрения действительного смысла.

Данная задача была решена программно и реализована в Microsoft Visual Studio 2013 (см. приложение 1).

Было произведено 15 циклов по 100000 испытаний. Результат работы программы представлен на рис.9.

```
0.01554 from (7900;9100)
0.173712 from (9100;10300)
0.459195 from (10300;11500)
0.296752 from (11500;12700)
0.0521127 from (12700;13900)
0.00269067 from (13900;15100)
Press any key to continue . . .
```

Рис.9 Результаты имитационного моделирования

Данная программная реализация (см. приложение 1), является решением конкретной поставленной задачи. При дальнейшем исследовании планируется выработать общий подход к решению аналогичных задач.

Проведя имитационное моделирование, становится очевидно, что с наибольшей вероятностью P значение случайной величины будет принадлежать промежутку (10300; 11500). Исходя из результатов моделирования можно выдвинуть гипотезу о том, что охват записи будет принадлежать именно этому интервалу. Условный заказчик согласился на публикацию информации в этих сообществах, потому что они территориально находятся в Москве и Санкт-Петербурге, тематика «спорт» и бюджет не превышает 500 рублей. Формат публикации записи: текст с приложенными медиа файлами и ссылкой на сайт продукции. Стоит отметить, что имитация рассматриваемого процесса также заключается в том, что если, допустим, изменить формат публикуемой новости, то изменится и количество просмотров подписчиками. Таким образом, это означает, что в данной модели нет тривиальных зависимостей, а рассматриваемая система является сложной.

§3.3 Анализ результатов

После того, как запись была опубликована в двух сообществах, появилась статистика. Получили следующие результаты:

1. Охват подписчиков/Суммарный охват (Количество пользователей, увидевших опубликованную запись) = $7224 \setminus 7451$ ($4243 \setminus 4268$).
2. Обратная связь - количество отметок «Мне нравится» = 8 (2), репостов = 1 (0), комментариев = 0 (0).
3. Переходы по ссылкам - суммарное количество человек, которые перешли по любой ссылке в записи = 37 (34).
4. Пол (м/ж %) = $99 \setminus 1$ ($96 \setminus 4$).

Суммарно охват записи (подписчиков) равен $7224 + 4243 = 11467$.

Гипотеза заключалась в том, что с вероятностью $P = 0,459195$ значение случайной величины будет принадлежать промежутку (10300; 11500).

Таким образом, результаты численного эксперимента и выдвинутая в предыдущем параграфе гипотеза подтвердились.

Поскольку задача является многокритериальной, приведённый пример иллюстрирует один из возможных подходов и сценариев распространения новости в рамках поставленной задачи. И даже в рамках рассмотренного примера можно менять сообщества, случайные величины и результаты будут различные. В качестве критерия оценивания проделанной работы можно ввести как критерий качества, так и критерий эффективности, что позволит определять правильность выбора сообществ(а) для распространения информации.

Заключение

В выпускной квалификационной работе рассмотрена задача имитационного моделирования информационного воздействия. На основе анализа предметной области выделены основные сущности модели распространения информации, а также описаны параметры, характеризующие каждую сущность, построена информационно-логическая модель, являющаяся основой имитационной модели анализа информационного воздействия, создана реляционная база данных, для хранения и обработки статистических данных, построена математическая модель и приведён пример иллюстрирующий один из возможных подходов к решению практической задачи, который программно реализован на языке программирования C++.

Дальнейшее развитие работы может осуществляться в нескольких направлениях:

- Определение основных типов задач: максимизации прибыли и минимизации потерь.
- Ввод в систему дополнительных параметров, характеризующих сообщества, что позволит расширить круг выполняемых задач и улучшить результаты имитационного моделирования.
- Разработка общих подходов к решению однотипных задач.
- Программно с использованием метода Монте-Карло реализация эти решения.
- Обращение к социальной сети «ВКонтакте» для сотрудничества (возможность получать расширенную статистику)[15].

Ранее уже проводились исследования на эту тему, однако исследователями не было формализовано всех аспектов в совокупности и не было рассмотрено взаимодействие сущностей между собой.

Рассматриваемая тема требует дополнительных исследований, так как она является достаточно новой и существует ещё много нерешённых задач.

Список литературы

1. Н.В. Белотелов, Ю.И. Бродский, Ю.Н. Павловский «Сложность. Математическое моделирование. Гуманитарный анализ: Исследование исторических, военных, социально-экономических и политических процессов»/ Предисл. Г.Г. Малинецкого. – М.: Книжный дом «ЛИБРОКОМ», 2009ю – 320с
2. Ю.Н. Павловский, Н.В. Белотелов, Ю.И. Бродский «Имитационное моделирование» : учеб.пособие для студ. высш. учеб. заведений / - М.: Издательский центр «Академия», 2008. – 236с. – (Университетский учебник. Сер. Прикладная математика и информатика)
3. РАСТОРГУЕВ С.П. «Математические модели в информационном противоборстве». Экзистенциальная математика. — М.: АНО ЦСОиП, 2014. — 260 с.
4. Влияние через социальные сети / Под общей ред. Е.Г. Алексеевой. М.: Фонд «ФОКУС-МЕДИА», 2010. – 200 с.
5. Журнал «International Journal of Simulation and Process Modelling» (IJSPM). <http://www.inderscience.com/browse/index.php?journalID=100/>
6. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 2, no. 11, 2014. <http://injoit.org/index.php/j1/article/viewFile/149/118/>
7. Журнал «SCS M&S Magazine» Издатель – международное общество имитационного моделирования (The Society for Modeling and Simulation International, SCS). <http://www.scs.org/msmagazine/>
8. Журнал «Simulation Modelling Practice and Theory» (SIMPRA). http://www.elsevier.com/wps/find/journaldescription.cws_home/622330/description/
9. Wikipedia. https://ru.wikipedia.org/wiki/Содружество_Независимых_Государств/
10. КЖЮР «Культурная жизнь юга России» №3 (54), 2014. <http://kjur.kguki.com/content/cms/files/35228.pdf/>

11. Пост наука. <https://postnauka.ru/longreads/20259/>
12. Поллард Дж. Справочник по вычислительным методам статистики.
Под ред. Е.М. Четыркина. – М.: Финансы и статистика, 1982. – 344 с.
13. Ивченко Г.И., Медведев Ю.И. Математическая статистика: Учеб.
Пособие для вузов. – М.: Высшая школа, 1984. – 248 с.
14. Вентцель Е.С. Теория вероятностей.– М.: Государственное
издательство физико-математической литературы, 1962. – 564 с.
15. Социальная сеть «ВКонтакте». <https://vk.com/>

Приложение 1

В данном приложении код программы имитационной модели информационного воздействия в среде разработки Microsoft Visual Studio 2013.

```
#include <iostream>
#include <cmath>
#include <ctime>
using namespace std;
int main(){
    srand(time(NULL));
    double m1[7], h1[7],m2[7], h2[7]; //массивы данных с границами
интервалов значений случайных величин и вероятность попадания в
интервалы (для каждого сообщества)
    double r1 = 0, r2 = 0, q = 0, w = 0; //случайные величины от 0 до 1 (r1,r2)
и значения охвата записи в 1-ом и 2- ом сообществе соответственно
    const int n = 100000; //количество экспериментов
    const int m = 15; //количество повторений экспериментов
    double z[n]; //массив для хранения экспериментальных данных
    double countf[6]; //массив-счетчик количества попадания в интервалы
при совестном использовании двух сообществ
    double sum[6];
    for (int i = 0; i < 6; i++){
        countf[i] = 0;
    }
    for (int i = 0; i < 6; i++){
        sum[i] = 0;
    }
    double prob[6],prob1[6]; //массив с вероятностями попадания в
интервалы при совестном использовании двух сообществ
    for (int i = 0; i < 6; i++){
        prob[i] = 0;
    }
    for (int i = 0; i < 6; i++){
        prob1[i] = 0;
    }
    h1[0] = 3000;
    h1[1] = 3500;
    h1[2] = 4000;
    h1[3] = 4500;
    h1[4] = 5000;
    h1[5] = 5500;
```



```

h1[6] = 6000;
h2[0] = 4900;
h2[1] = 5600;
h2[2] = 6300;
h2[3] = 7000;
h2[4] = 7700;
h2[5] = 8400;
h2[6] = 9100;
double sig1 = h1[3] * 0.1, sig2 = h2[3] * 0.1;//среднеквадратическое
отклонение для первого и второго сообщества соответственно
double h3[7];
for (int i = 0; i < 7; i++){
    h3[i] = h1[i] + h2[i];
}
m1[0] = 0;//массив вероятностей попадания в интервал для 1-го
сообщества
m1[1] = 0.05;
m1[2] = 0.2;
m1[3] = 0.55;
m1[4] = 0.9;
m1[5] = 0.98;
m1[6] = 1;
m2[0] = 0;//массив вероятностей попадания в интервал для 2-го
сообщества
m2[1] = 0.05;
m2[2] = 0.2;
m2[3] = 0.6;
m2[4] = 0.9;
m2[5] = 0.97;
m2[6] = 1;
for (int j = 0; j < m; j++){ //количество повторений экспериментов
    for (int i = 0; i < n; i++){ //количество экспериментов
        r1 = 0.0001 * (rand() % 10001);
        r2 = 0.0001 * (rand() % 10001);
        if (m1[0] <= r1 && r1 <= m1[1]){//проверка попадания в
интервал
            q = h1[0] + sig1*(1 - 0.0001 * (rand() % 10001));//если
попали в интервал, то значение случайной величины рассчитывается по
формуле: q=нижняя граница+сигма*(1-случайное число(от 0 до 1))
        }
        if (m1[1] <= r1 && r1 <= m1[2]){
            q = h1[1] + sig1*(1 - 0.0001 * (rand() % 10001));
        }
        if (m1[2] <= r1 && r1 <= m1[3]){

```

```

        q = h1[2] + sig1*(1 - 0.0001 * (rand() % 10001));
    }
    if (m1[3] <= r1 && r1 <= m1[4]){
        q = h1[3] + sig1*(1 - 0.0001 * (rand() % 10001));
    }
    if (m1[4] <= r1 && r1 <= m1[5]){
        q = h1[4] + sig1*(1 - 0.0001 * (rand() % 10001));
    }
    if (m1[5] <= r1 && r1 <= m1[6]){
        q = h1[5] + sig1*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[0] <= r2 && r2 <= m2[1]){
        w = h2[0] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[1] <= r2 && r2 <= m2[2]){
        w = h2[1] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[2] <= r2 && r2 <= m2[3]){
        w = h2[2] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[3] <= r2 && r2 <= m2[4]){
        w = h2[3] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[4] <= r2 && r2 <= m2[5]){
        w = h2[4] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    if (m2[5] <= r2 && r2 <= m2[6]){
        w = h2[5] + sig2*(1 - 0.0001 * (rand() % 10001));
    }
    z[i] = q + w;
    q = 0; w = 0; r1 = 0; r2 = 0;
}
for (int i = 0; i < n; i++){
    if (h3[0] <= z[i] && z[i] <= h3[1]){//проверка попадания в
интервал при использовании двух сообществ
        countf[0]++;
    }
    if (h3[1] <= z[i] && z[i] <= h3[2]){
        countf[1]++;
    }
    if (h3[2] <= z[i] && z[i] <= h3[3]){
        countf[2]++;
    }
    if (h3[3] <= z[i] && z[i] <= h3[4]){

```

```

        countf[3]++;
    }
    if (h3[4] <= z[i] && z[i] <= h3[5]){
        countf[4]++;
    }
    if (h3[5] <= z[i] && z[i] <= h3[6]){
        countf[5]++;
    }
}
for (int i = 0; i < 6; i++){
    prob[i] = countf[i] / n;
}
for (int i = 0; i < 6; i++){
    sum[i] +=prob[i];
}
for (int i = 0; i < 6; i++){
    countf[i] = 0;
}
for (int i = 0; i < 6; i++){
    prob[i] = 0;
}
}
for (int i = 0; i < 6; i++){//цикл в котором усредняются результаты 1,5
МИЛЛИОНОВ ИСПЫТАНИЙ
    prob1[i] = sum[i]/m;
}
for (int i = 0; i < 6; i++){
    cout << prob1[i] << " " << "from " << " (" << h3[i] << ";" << h3[i +
1] << ")" << endl;
}
system("pause");
return 0;
}

```