

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

Н.Ф. Гусарова

**ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ
В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ
ПРОЦЕССАМИ**

Учебное пособие

 УНИВЕРСИТЕТ ИТМО

Санкт-Петербург

2015

Гусарова Н.Ф. **Интеллектуальные системы в управлении социальными процессами.** – СПб: Университет ИТМО, 2015. – 90 с.

Пособие охватывает теоретический материал, необходимый для усвоения программы курса «Интеллектуальные системы в управлении социальными процессами». Рассматриваются вопросы, связанные с базовыми понятиями в управлении социальными процессами, а также с возможностями применения интеллектуальных информационных систем для управления социальными процессами. Особое внимание уделяется такому актуальному феномену современной социальной жизни, как социальные сети, их моделированию и организации воздействия на них с применением интеллектуальных информационных систем.

Пособие адресовано студентам, обучающимся по направлению подготовки 036000 – Интеллектуальные системы в гуманитарной сфере.

Рекомендовано к печати Ученым советом ЕНФ, протокол № 4 от 23 апреля 2015.



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО – становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО, 2015

© Гусарова Н.Ф., 2015

Оглавление

Введение	5
1. Основные понятия в управлении социальными процессами.....	6
1.1. Социальные процессы и их структура	6
1.2. Общесистемные характеристики социальных процессов.....	8
1.2.1. Классификация систем по уровню сложности.....	8
1.2.2. Активные и пассивные системы.....	11
1.2.3. Равновесие, гомеостаз и гомеокинез.....	11
1.2.4. Равновесие в играх.....	14
1.3. Основные понятия в управлении социальными процессами	19
1.4. Базовая модель управления в социальных процессах.....	23
1.5. Управление в малых группах.....	25
1.6. Управление в социальных сетях.....	27
1.6.1 Основные понятия управления в социальных сетях	27
1.6.2. Подходы к моделированию управления в социальных сетях	29
1.6.3. Параметры управления в социальных сетях: мнение, влияние, доверие	31
1.6.4. Эффекты и свойства в реальных социальных сетях.....	33
2. Математические модели социальных сетей	36
2.1. Пороговые модели.....	36
2.2. Модели распространения информации в сети	36
2.3. Модели диффузии инноваций.....	41
3. Модели влияния и управления в социальных сетях	47
3.1. Оценка силы влияния агента.....	47
3.1.1. Определение самых влиятельных агентов в сети.....	47
3.1.2. Обнаружение каскадов распространения влияния.....	48
3.1.3. Индексы влияния.....	48
3.2. Общее знание и его учет в социальных сетях. Рефлексии	49
3.3. Коллективные действия и общественные блага	51
3.3.1. Примеры общественных благ	51
3.3.2. Управление через выбор механизма игры.....	52
3.3.3. Коллективные действия агентов в социальных сетях.....	53
3.3.4. Информационное управление в социальных сетях.....	55
4. Графовые модели социальных сетей	59
4.1. Основная терминология	59
4.2. Сообщества и связанные подгруппы	61
4.2.1. Эффекты степени вершины в сети	61
4.2.2. Структура сообщества.....	62
4.2.3. Достаточные сети.....	64
4.2.4. Small-world модель.....	65
4.2.5. Параметры качества взаимодействия	66

4.3. Модели динамики сети	67
4.4. Методы на основе онтологий.....	68
4.5. Случайные сети и безмасштабные графы.....	69
5. Примеры применения интеллектуальных систем в управлении социальными процессами.....	72
5.1. Примеры прикладных задач.....	72
5.2. Поддержка рекламной кампании в социальной сети	76
5.2. Автоматизированное пополнение онтологии предметной области	81
Литература.....	87

Введение

Управление в социальных системах – одна из важнейших сфер приложения информационных и, в частности, интеллектуальных систем в современном мире.

Для организации любого управления необходимо, в первую очередь, определить и описать объект управления. Специфической особенностью вновь появляющихся типов социальных систем – в первую очередь, социальных сетей – является их принципиально открытый характер, который недостаточно полно учитывает классическая теория управления. С другой стороны, исследования в области социальных сетей, широко представленные в гуманитарных науках, как правило, носят умозрительный или описательный характер, т.е. также не предоставляют возможности построения моделей открытых социальных систем с достаточной степенью формальности.

Междисциплинарный характер и высокая востребованность проблематики отражены в большом числе публикаций научной периодики. Однако, как отмечается в [Давыденко, 2005], ощущается острый дефицит систематического изложения методов и алгоритмов анализа современных социальных систем, пригодных для поддержки управления в них с применением интеллектуальных систем.

Предлагаемое пособие представляет собою краткое введение в проблематику управления в современных социальных системах с применением информационных и, в частности, интеллектуальных систем. Пособие практически полностью построено на материалах научной периодики и монографий. Большое внимание уделено подходам к моделированию социальных систем как объекта управления. Приводятся также примеры постановки и решения практических задач применения интеллектуальных систем при организации управления социальными системами.

1. ОСНОВНЫЕ ПОНЯТИЯ В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ ПРОЦЕССАМИ

1.1. СОЦИАЛЬНЫЕ ПРОЦЕССЫ И ИХ СТРУКТУРА

Все социальные изменения, происходящие в обществе, взаимосвязаны, взаимообусловлены и приводят к направленным сдвигам в состоянии общества, называемым социальными процессами.

Социальный процесс есть результат деятельности людей, в которой достигаются определенные цели в развитии социальных объектов (социальных систем), а в самих объектах происходит последовательная смена их состояний.

Структура социального процесса включает следующие элементы:

1. Субъект социальных действий. Это могут быть личность или социальная общность, организация, которые осуществляют действия для достижения поставленных целей.

2. Объект социальных действий. Это социальная система или социальное явление, на которые направлены социальные действия субъекта с целью изменить их состояние.

3. Условия обитания, в которых осуществляется социальное действие (микро- и макросреда).

4. Целенаправленное социальное действие.

5. Результат социального действия – те изменения, которые произошли в социальном объекте в результате воздействия на него субъекта.

Социология [Социология, 2009; Общая социология, 2009] предлагает классификации социальных процессов по различным основаниям, обобщенные в таблице 1.

Таблица 1.

№	Основание классификации	Типы процессов
1	Формы социальных процессов	направленные однонаправленные разнонаправленные ступенчатые ненаправленные неупорядоченные, циклические, спиральные, стагнация
2	Результаты социальных процессов	морфогенетические трансмутационные репродуктивные трансформационные

3	Осведомленность населения о социальных процессах	явные скрытые «процессы-бумеранги»
4	Движущие силы социальных процессов	эндогенные экзогенные
5	Уровень протекания социальных процессов	глобальный уровень уровень больших социальных групп уровень институтов и организаций уровень межличностных отношений
6	Время протекания социальных процессов	кратковременные долговременные постоянные
7	Сферы проявления социальных процессов	экономические политические экологические воспитательные и т.д.
8	Механизм возникновения	стихийные сознательные естественно-исторические
9	Степень управляемости социальных процессов	управляемые слабо управляемые неуправляемые
10	Роль социальных процессов в обществе	функциональные дисфункциональные
11	Направленность социальных процессов	прогрессивные регрессивные
12	Интенсивность изменений	эволюционные революционные

Дадим некоторые пояснения к таблице 1.

Направленными являются необратимые социальные процессы, в которых каждая последующая стадия отличается от предыдущей и включает ее результат, а более ранняя стадия подготавливает условия для более поздней. Направленные социальные процессы делятся на однонаправленные, разнонаправленные и ступенчатые.

В ненаправленных социальных процессах изменения носят либо чисто случайный, хаотический характер, либо подчиняются определенным повторяющимся или, по крайней мере, сходным моделям, причем каждая последующая стадия идентична или качественно напоминает предыдущие виды. Различают четыре вида ненаправленных социальных процессов: неупорядоченные, циклические, спиральные, стагнация.

Морфогенетические социальные процессы приводят к возникновению новых состояний общества, социальных структур и условий существования социума. Репродуктивные изменения (адаптивные, уравнивающие, поддерживающие) в основном имеют количественный характер. Однако некоторые репродуктивные процессы приводят к качественным изменениям состояния общества – тогда они называются трансформационными.

В зависимости от осознания населением конечных результатов социальных процессов выделяются явные, скрытые и «процессы-бумеранги».

Смысл и последствия явных процессов осознаются населением, в то время как смысл скрытых процессов населением не осознается. «Процессы-бумеранги» – социальные процессы, смысл которых населением осознается, однако их последствия для большинства населения не ясны.

Эндогенные (движимые внутренними причинами) процессы раскрывают потенциальные возможности, свойства и тенденции, заключенные внутри изменяющегося социума. Экзогенные (движимые внешними причинами) процессы являются реактивными, отвечают на вызов, стимул или давление извне.

Социальные процессы считаются кратковременными, если время их протекания меньше времени существования социального объекта. Долговременные процессы происходят на протяжении всего времени существования объекта. Постоянные процессы – глобальные социальные процессы, протекание которых не связано с существованием какого-либо одного социального объекта.

1.2. ОБЩЕСИСТЕМНЫЕ ХАРАКТЕРИСТИКИ СОЦИАЛЬНЫХ ПРОЦЕССОВ

Как следует из предыдущего параграфа, социальный процесс и его компоненты представляют собой сложные системные образования и в связи с этим должны подчиняться общесистемным закономерностям. Рассмотрим общесистемные характеристики социальных процессов, наиболее важные как с точки зрения организации управления в них, так и с точки зрения использования в них интеллектуальных систем.

1.2.1. Классификация систем по уровню сложности

К. Боулдинг предложил классифицировать системы по тому, как они перерабатывают входную информацию.

В соответствии с этой классификацией (рис. 1) все системы располагаются на условной порядковой шкале из 9 уровней.

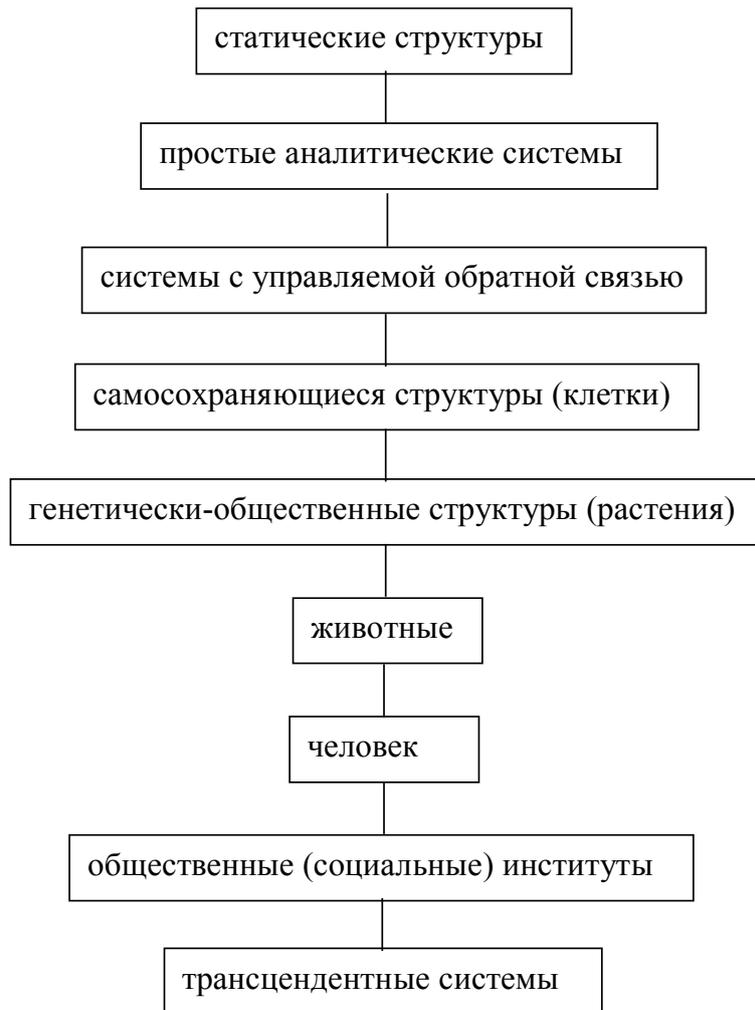


Рис. 1. Классификация систем по К. Боулдингу

Поясним эту классификацию.

1. Уровень статической структуры – это уровень статических систем, существование которых не предопределяется обработкой информации.

Примеры: камень, гора.

2. Уровень простой динамической системы с предопределенными, обязательными движениями.

Пример: часовой механизм.

Хотя эти системы и являются динамическими, их существование также не связано с самостоятельной переработкой информационных потоков.

3. Уровень систем с управляемой обратной связью.

Пример: термостат.

Это простейший из всех уровней систем, существующих в мире, где информационные потоки и их переработка могут влиять на систему.

4. Уровень самосохраняющихся структур.

Пример: одноклеточные организмы.

Это уровень зарождения собственного отношения системы к входящей информации, уровень промежуточный между пассивной и активной реакцией на входную информацию.

5. Уровень генетически-общественных структур.

Пример: растения.

Имеет место специфическая реакция на возмущающую информацию – механизмы приспособляемости и другие реакции на внешние воздействия.

6. Уровень животных.

Особенности этого уровня:

- подвижность, целенаправленное поведение и осведомленность;
- развиты специализированные приемники информации (глаза, уши и т.д.), что приводит к значительному увеличению потока перерабатываемой входной информации;
- имеются развитые нервные системы и, в пределе, мозг, который формирует из воспринимаемой информации основные черты явления – образ;
- затруднено предсказание поведения, так как между воздействием и реакцией на него «вклинивается» образ.
- чем выше организация животного, тем заметнее, что его поведение не является простым ответом на воздействие, а определяется образом, или структурой знания, или окружающей обстановкой в целом.

7. Уровень отдельного человека как системы.

Особенности этого уровня:

- сохраняются почти все особенности, характерные для животных;
- помимо простой осведомленности, возникает самосознание, т.е. выделение себя из окружающего мира: человек не только знает, но и осознает, что знает;
- использование символов, сложный язык.

8. Уровень социальных институтов.

Согласно классификации Боулдинга, этот уровень помещен выше уровня отдельного человека, так как составляет для человека законодательную и контролирующую среду, а также во многом определяет формы организации его существования. Однако многие современные исследования показывают, что хотя бы отчасти поведение человека в социальных институтах подчиняется коллективным механизмам уровня животных или даже коллективов автоматов.

Эта ситуация вполне понятна. Человек сочетает в себе три уровня переработки информации – уровень неживой природы, уровень живой природы и уровень сознания. Какой именно уровень будет доминировать в том или ином проявлении человека, зависит от решаемой задачи.

Примеры: как показано в социальной психологии, человек в толпе, особенно в случае паники, ведет себя по тем же законам, что и животные, т.е. его сознание в этом поведении не участвует.

В то же время здесь возможно включение индивидуального сознания (например, если человек сам становится лидером движения или если он видит осмысленные образцы для подражания – людей, активно сопротивляющихся панике).

9. Трансцендентные системы.

Примеры: мир идей Платона.

О существовании таких систем говорили многие философы, в том числе И. Кант. В рассматриваемой классификации этот уровень появляется вполне логично и должен соответствовать полному отрыву информации от физического носителя.

1.2.2. Активные и пассивные системы

Активность или пассивность систем связывается с наличием у них внутреннего источника энергии и собственного целеполагания.

Согласно [Новиков, Петраков, 1999], активная система – это модель организационной системы, в которой в существенной степени учитывается наличие несовпадающих интересов у субъектов управления (агентов) и их активное поведение, т.е. представление информации управляющему органу (центру) и выбор действий исходя из собственных интересов.

Общим для всех пассивных систем является их предсказуемость с точки зрения управления: у пассивной системы нет свободы выбора своего состояния, собственных целей, средств их достижения, она не может прогнозировать поведение управляющего органа.

Иначе обстоит дело в активных системах. Здесь управляемые субъекты (точнее говоря, хотя бы один субъект) обладают свойством активности, в том числе свободой выбора своего состояния. Кроме того, элементы активной системы обладают собственными интересами и предпочтениями, т.е. осуществляют выбор состояния целенаправленно (в противном случае их поведение можно было бы рассматривать как пассивное).

При организации управления активными системами нужно учитывать проявления активности управляемых субъектов. Чаще всего управляемые субъекты стремятся к выбору таких своих состояний (стратегий), которые являются наилучшими с точки зрения их предпочтений при заданных или прогнозируемых значениях управляющих воздействий. Достаточно часто можно считать, что эти состояния являются одним из вариантов равновесия.

1.2.3. Равновесие, гомеостаз и гомеокинез

Как реализуется состояние равновесия в системах различного уровня сложности?

В простейших линейных системах устойчивое равновесие достигается за счет отрицательной обратной связи, которая удерживает систему в области устойчивости.

Пример: шарик в лунке и шарик на выпуклой поверхности.

С повышением уровня сложности системы механизмов отрицательной обратной связи становится все больше, и между ними уже требуется координация. Возникает понятие «гомеостаз» (от греч. *homoios* – подобный, одинаковый, и. *stásis* – состояние, неподвижность).

Гомеостаз в физиологии – относительное динамическое постоянство состава и свойств внутренней среды и устойчивость основных физиологических функций организма человека, животных и растений.

Термин «гомеостаз» был предложен У. Кенноном в 1929 году для обозначения относительного динамического постоянства внутренней среды и устойчивости основных физиологических функций. В основе такого постоянства лежит свойство живых клеток возвращаться к исходному состоянию после сдвигов в них в результате каких-либо возмущающих воздействий. Подобными свойствами обладают и более сложно организованные структуры – органы, системы и организм в целом.

Примеры:

- Выравнивание артериального давления осуществляется регуляторными механизмами, вступающими в действие по принципу цепных реакций с обратными связями (изменение давления крови воспринимается барорецепторами сосудов, сигнал о нем передается в сосудистые центры, изменение состояния которых ведет к изменению тонуса сосудов и сердечной деятельности; одновременно раздражаются и хеморецепторы сосудов, включающие систему нейро-гуморальной регуляции, и кровяное давление возвращается к норме).
- Сохранение постоянства уровня влажности листьев растений осуществляется путем открывания и закрывания устьиц.

Понятие гомеостаза применимо также к сообществам организмов.

Примеры:

- гомеостаз в биоценозе – сохранение постоянства видового состава и числа особей;
- генетический гомеостаз – способность популяции поддерживать динамическое равновесие генетического состава, что обеспечивает ее максимальную жизнеспособность.

В кибернетике термин «гомеостаз» применяют по отношению к любому саморегулирующемуся механизму.

Графически любое явление гомеостаза можно представить в виде кривой (рис. 2).

Однако, в отличие от вечного двигателя, все реальные системы со временем изменяются и, в конце концов, прекращают свое существование. Строго говоря, все системы достаточно высокого уровня сложности, для

которых уже нельзя пренебречь временным фактором их существования, находятся в состоянии неравновесия: для них существует устойчивое состояние динамического равновесия, к которому они стремятся, но никогда не могут достигнуть.

Для таких систем понятия «гомеостаз» уже недостаточно, и его заменяют на понятие «гомеокинез».

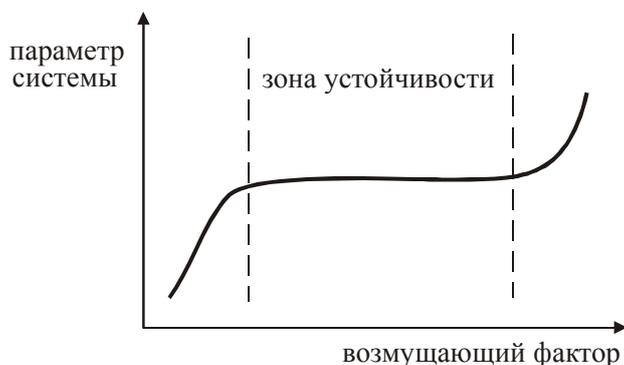


Рис. 2. Кривая гомеостаза

Пока система «жива», она противостоит тенденции перехода в состояние с большей энтропией, для чего постоянно вводит в себя энергию и обрабатывает входную информацию. Этот процесс можно рассматривать как попытки системы достичь состояния равновесия и сохранить его, т.е. пребывать в пределах некоего «гомеокинетического плато» (рис. 3).

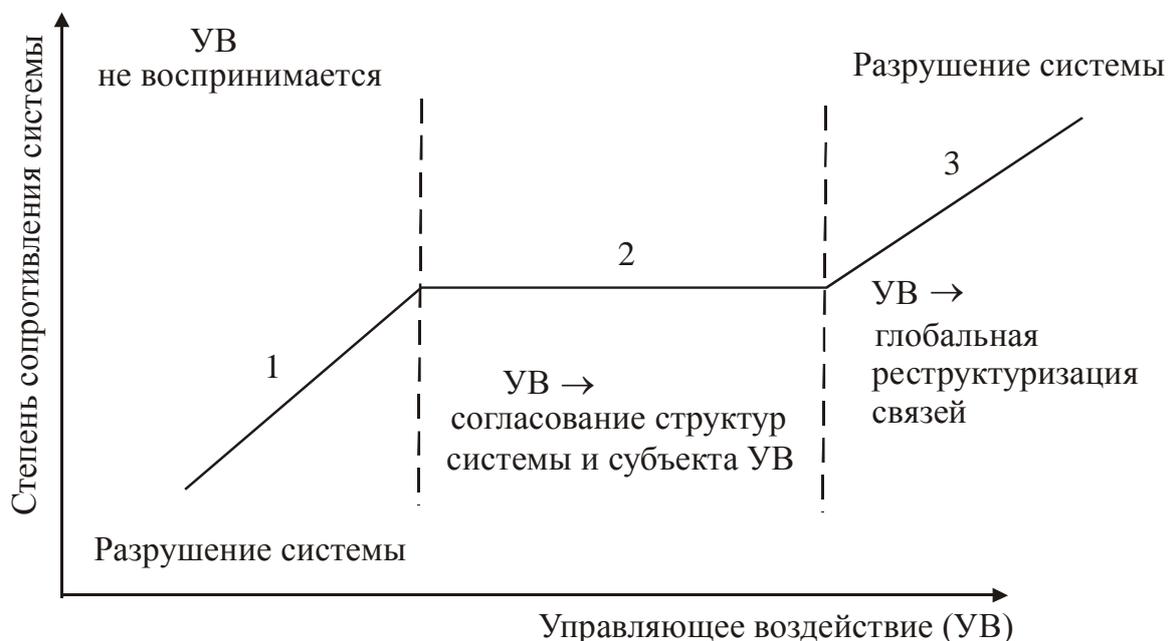


Рис. 3. Кривая гомеокинеза

Рассмотрим применение рис. 3 на примере системы «батальон». Батальон состоит из людей со своими особенностями физиологии и психики,

своими интересами и т.д. Управляющее воздействие (УВ) на него оказывает командование (субъект УВ). Очевидно, что далеко не всегда личные интересы членов батальона совпадают с интересами командования, что вызывает определенное сопротивление УВ.

Если УВ мало или не оказывается вовсе (зона 1), то система не существует как стабильное образование. Это не батальон, а просто группа людей (например, на призывном пункте или сразу после демобилизации).

В диапазоне УВ, соответствующем зоне 2, система работоспособна. Здесь имеет место текущее согласование целей, которые ставит командование, и внутренней активности личного состава.

Если УВ слишком велико (зона 3), то система «перестает быть сама собой». Люди либо полностью теряют инициативу, либо проявляют открытое неповиновение, и батальон в целом становится небоеспособным. Ситуацию можно исправить путем реформирования батальона. В этом случае система перейдет на новое плато, частично или полностью изменив свои свойства и характеристики.

1.2.4. Равновесие в играх

Под игрой понимается процесс, в котором участвуют две или более сторон, ведущих борьбу за реализацию своих интересов. Каждая из сторон имеет свою цель и использует некоторую стратегию, которая может вести к выигрышу или проигрышу – в зависимости от поведения других игроков. Игра по определению происходит в условиях неполной информированности всех игроков.

Игра описывается тройкой

$$\langle I, S_i, u_i \rangle,$$

где $I = \{1, \dots, N\}$ – конечное множество игроков, S_i – множество действий, доступных i -му игроку, $u_i: S \rightarrow R$ – множество функций выплат. Если множество стратегий конечно, то множество исходов игры можно выразить N -мерной матрицей, в которой каждому набору стратегий соответствует определенная выплата.

Стратегия S доминируема, если существует другая стратегия S' , которая не хуже S в каждой точке, при любых возможных комбинациях стратегий других агентов. И наоборот, стратегия S является доминантной, если всякая другая стратегия S' ею доминируется. Это означает, что игроку достаточно выбрать доминантную стратегию, все равно никакая другая ни при каком исходе ничего лучшего не даст.

Равновесие в доминантных стратегиях для игры $\langle I, S_i, u_i \rangle$ – это такой набор (вектор) стратегий $s^* \in S$, что для всякого агента i стратегия $(s_i)^*$ является доминантной.

Такое равновесие является самым устойчивым из всех. Однако оно достижимо далеко не всегда. Имеется много типов игр, в которых для каж-

дой стратегии s_i игрока i существует вектор стратегий других игроков s_{-i} , при котором игроку i было бы выгодно сменить s_i на другую.

Что же делать участвующим в игре агентам, когда доминантных стратегий у них нет и не предвидится? Более того, в играх возникает много парадоксов, т.е. не всегда то, что выгодно каждому по отдельности, выгодно всем вместе. Тогда приходится учитывать не только свои собственные стратегии, но и стратегии других агентов. Этот учет приведет к другим понятиям равновесия.

Рассмотрим пример – дилемма заключенного.

Двоим заключенным предлагают признаться в преступлении и «заложить» своего сообщника. Доказательств у обвинения нет, поэтому, если оба промолчат, то получают небольшой срок (0,5 года) за другие, более мелкие преступления; если оба признаются, то обоим за примерное поведение дадут по 2 года; если же один признается, а другой – нет, то признавшегося за сотрудничество отпустят, а упорствующий получит полный срок – 10 лет. Информировать друг друга заключенные не могут. Что делать каждому из них?

Запишем матрицу возможных стратегий:

	Промолчать	Сознаться
Промолчать	(0,5; 0,5)	(10; 0)
Сознаться	(0; 10)	(2; 2)

Легко видеть, что, вне зависимости от выбора первого заключенного, второму в любом случае выгоднее признаться.

Этот же пример в интерпретации экономики: две фирмы производят один и тот же продукт, других фирм на рынке нет. Нужно ли каждой из них рекламироваться (= сознаться) или обеим не рекламироваться (= промолчать)?

Таким образом, в общей постановке игровой задачи каждый агент решает свою локальную задачу, пытаясь максимизировать свою собственную прибыль. В результате система приходит в какое-то равновесное состояние. В каждом взаимодействии могут существовать различные виды равновесий, в том числе: равновесие доминирующих стратегий, равновесие по Нэшу, равновесие по Штакельбергу, равновесие по Парето.

Доминирующая стратегия, которая уже рассматривалась выше – такой план действий, который обеспечивает участнику максимальную полезность вне зависимости от действий другого участника. Соответственно, равновесием доминирующих стратегий будет пересечение доминирующих стратегий обоих участников игры.

Равновесие по Нэшу – ситуация, в которой ни один из игроков не может увеличить свой выигрыш в одностороннем порядке, меняя свой план действий. Равновесие по Нэшу – ситуация, в которой стратегия каждого из игроков является лучшим ответом на действия другого игрока. Иными

словами, это равновесие обеспечивает игрока максимумом полезности в зависимости от действий другого игрока.

Равновесие по Штакельбергу – ситуация, когда ни один из игроков не может увеличить свой выигрыш в одностороннем порядке, а решения принимаются сначала одним игроком и становятся известными второму игроку. Равновесие по Штакельбергу возникает тогда, когда существует временной лаг в принятии решений участниками игры: один из них принимает решения, уже зная, как поступил другой. Таким образом, равновесие по Штакельбергу соответствует максимуму полезности игроков в условиях неодновременности принятия ими решений. В отличие от равновесия доминирующих стратегий и равновесия по Нэшу, этот вид равновесия существует всегда.

Равновесие по Парето – ситуация, когда нельзя улучшить положение ни одного из игроков, не ухудшая при этом положения другого и не снижая суммарного выигрыша игроков. Равновесие по Парето существует при условии, что нельзя увеличить полезность обоих игроков одновременно.

Рассмотрим отдельные типы равновесия более подробно.

Равновесие по Нэшу легко понять на основе дилеммы заключенного.

Каждому игроку неизвестно, какую стратегию выберет противник, поэтому ему самому целесообразнее выбрать стратегию, рассчитанную на самое неблагоприятное для него поведение противника (принцип «гарантированного результата»). Действуя осторожно и считая конкурента сильным, этот игрок выберет для каждой своей стратегии минимально возможный выигрыш. Таким образом, он из всех минимально выигрышных стратегий выберет такую, которая обеспечит ему максимальный из всех минимальных выигрышей (принцип «максимин»).

Его противник рассуждает так же. Он найдет для себя наибольшие проигрыши во всех стратегиях этого игрока, а потом из этих максимальных проигрышей выберет минимальный (принцип «минимакс»).

При равенстве максимина минимаксу решения игроков будут устойчивыми, а игра будет иметь равновесие. Устойчивость (равновесие) решений (стратегий) заключается в том, что обоим участникам игры будет невыгодно отходить от выбранных стратегий. Когда же максимин не равен минимаксу, то решения (стратегии) обоих игроков, если они хотя бы в какой-то мере угадали выбор стратегии противника, будут неустойчивыми, неравновесными.

Значит, равновесие Нэша – результат, в котором стратегия каждого из игроков является лучшей среди других стратегий, принятых остальными участниками игры. Это определение основывается на том, что каждый из игроков изменением собственной роли не может достичь наибольшей выгоды (максимизации функции полезности), если другие участники твердо придерживаются собственной линии поведения.

Равновесие по Нэшу – фундаментальное понятие теории игр, но оно применимо не всегда. Для реализации равновесия по Нэшу требуется выполнение нескольких условий:

- каждый агент знает структуру игры полностью;
- каждый агент знает, что другие ее тоже знают (принцип рефлексии – см. § 3.2);
- каждый агент знает, что все действуют рационально;
- каждый агент знает, что все выберут одно и то же равновесие по Нэшу (а их может быть несколько).

Поэтому свою «формулу равновесия» Дж.-Ф. Нэш усилил показателем оптимального объема информации. Он вывел его из анализа ситуаций с полным информированием игрока о своих противниках и с неполным информированием о них. Переведя этот постулат с математического языка на язык экономической жизни, ученый ввел как важный информационный элемент знания об условиях «внешней среды» неуправляемые переменные рыночных отношений.

Здесь появляется следующий тип равновесия – ***равновесие по Байесу-Нэшу***, которое предъявляет более естественные требования к информированности агентов:

агент не знает наверняка, каковы функции полезности других игроков (т.е. чего они хотят), но он знает выплаты для каждого возможного типа и имеет априорное распределение этих выплат на типах для каждого из других агентов. В этом случае стратегия агента оптимальна по распределению типов других агентов, хотя в каждом частном случае игры агент может выбрать и неоптимальное поведение.

Равновесие по доминантным стратегиям – еще более сильное равновесие. Примером такого равновесия являются тендеры.

Тендеры проводятся по схеме аукциона Викри: участники подают свои заявки в конвертах, в момент проведения аукциона их вскрывают, и объект продается тому, кто предложил самую высокую цену. Можно показать, что в таком аукционе каждому участнику выгодно не скрывать свою цену, потому что это будет его доминантная стратегия независимо от других.

Пример 2 – в дилемме заключенного для каждого из них «сознаться» – доминантная стратегия, и в результате они получают срок 2 года каждый, а не по 0,5 года, независимо от стратегии другого.

Равновесие по Парето учитывает суммарную эффективность игры в целом. Здесь признается право на все изменения, которые не приносят никому дополнительного вреда. Равновесие по Парето – ситуация, когда нельзя улучшить положение ни одного из игроков, не ухудшая при этом положения другого и не снижая суммарного выигрыша игроков. Оптимальность по Парето означает, что если кому-то стало лучше, чем в предлагаемой функцией социального выбора варианте, то кому-то другому обя-

зательно стало хуже, т.е. нельзя монотонно улучшить состояние сразу всех агентов.

Пример. В России считается справедливым, что пенсионеры должны иметь возможность передвигаться достаточно свободно по территории страны, используя общественный транспорт (в том числе в рамках одного населенного пункта), поэтому долгое время он оставался де-факто бесплатным для пенсионеров. Но многие пенсионеры (например, живущие в сельской местности) не пользовались транспортом вообще. А другие (преимущественно в крупных городах) использовали общественный транспорт для поездок в дешевый магазин на другом конце города, где они могли купить хлеб на 10 коп. дешевле, при этом стоимость их проезда для государства могла превышать 50 рублей. Как видно, бесплатный проезд был, с одной стороны, несправедлив к сельским пенсионерам, а с другой – крайне неэффективен для городских (экономия 10 копеек обходится в 10–50 рублей). Поэтому была предложена концепция монетизации льгот, по которой всем пенсионерам выдается одинаковая сумма денег, которой должно хватить на покупку карточки для проезда в общественном транспорте в большинстве городов (за исключением 10–15 самых крупных и дорогих). Таким образом, пенсионеры, получив средства на руки, могут потратить их с большей пользой.

Равновесие по Парето реализуется только при ряде условий:

- имеется большое число потребителей;
- никто не может влиять на цену, то есть никто не имеет рыночной власти;
- товар на рынке является нормальным;
- нет экстерналий (внешних эффектов) в экономике;
- предпочтения потребителей являются выпуклыми.

Примеры экстерналий:

- положительная экстерналия со стороны производителя – взаимодействие расположенных рядом пасеки и яблоневого сада: пчелы способствуют повышению урожая яблок, а яблони – увеличению сбора меда, при этом их хозяева не вступают между собой ни в какие рыночные отношения. При этом оптимально было бы увеличить количество данного товара, но сделать его платным для потребителей. Такого рода экстерналия часто ассоциируется с «эффектом безбилетника», то есть когда потребитель не платит за пользование товаром или услугой при условии, что производитель инвестировал в их производство. Создателем положительного внешнего эффекта можно назвать, к примеру, жителя дома, установившего освещение в своем подъезде в личных интересах, что при этом принесло пользу соседям по дому.
- отрицательная экстерналия – загрязнение окружающей среды, когда увеличение прибыли предприятия в результате наращивания производства оборачивается ущербом окружающей среде, от чего страдают не-

которые фирмы и общество в целом. Государство может установить различные виды налогов, пошлины для «вредителей». При увеличении налогов фирмы сокращают предложение своего товара, при этом повышая на него цену.

На рынке может существовать несколько точек Парето-равновесия. Если условия выполнены, то любое существующее парето-оптимальное состояние можно реализовать рыночными методами. То есть, если из множества парето-оптимальных состояний рынка нам по какой-то причине (например, мы считаем его наиболее справедливым) нравится какое-то конкретное, то, манипулируя исключительно доходами (количеством имущества во времени) индивидов и фирм, мы можем достичь принятия ими таких решений, которые в равновесии будут давать эту самую справедливость. Таким образом, экономическая наука при помощи этой теоремы разделяет проблемы справедливого распределения и эффективности.

Заметим, что сейчас во всем мире реализуется справедливость 20/80: 80% результата получают благодаря 20% усилий; с другой стороны, 80% усилий принесут лишь 20% результата. Эту пропорцию сформировал Парето исходя из исследований уровня благосостояния итальянских семей. В его время 20% семей принадлежало около 80% всех доходов.

Сейчас уместнее было бы говорить, что 90% всех доходов приходится на 10% всех семей. Те же самые пропорции справедливы и по отношению к любой профессии: 10% актеров, спортсменов, художников, инвесторов, бизнесменов и др. зарабатывают 90% денег. Важно понять, что соотношения могут быть очень разными, но общий принцип закона Парето справедлив.

Принцип Парето тесно связан с экономической эффективностью. Экономическая эффективность, согласно Парето – это такая ситуация рынка, когда ни один из его участников не может улучшить свое положение, не ухудшив положения других его участников. В качестве наглядного примера экономической эффективности по Парето можно указать на расширяющуюся из года в год пропасть между богатыми и бедными. Богатство одних увеличивается в основном за счет других. Другое название для такой ситуации - оптимум Парето.

1.3. ОСНОВНЫЕ ПОНЯТИЯ В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ ПРОЦЕССАМИ

Словарь [БСЭ] определяет управление следующим образом:

Управление – функция организованных систем различной природы (биологических, социальных, технических), обеспечивающая сохранение их определенной структуры, поддержание режима деятельности, реализацию программы, цели деятельности.

Социальное управление – воздействие на общество с целью его упорядочения, сохранения качественной специфики, совершенствования и развития – есть непереносимое, внутренне присущее свойство любого общества, вытекающее из его системной природы, общественного характера труда, необходимости общения людей в процессе труда и жизни, обмена продуктами их материальной и духовной деятельности.

Понятие «социальное управление» рассматривается в узком и широком значении слова [Иванов, 2001]. В узком значении социальное управление определяется как процесс воздействия на социальные процессы для достижения поставленных целей. В широком смысле слова социальное управление рассматривается как сфера деятельности людей.

Необходимость социального управления обусловлена фактом разделения труда в группах людей, в коллективах, в масштабе государства, а также необходимостью кооперации при совместной деятельности.

Ключевое понятие управления – воздействие. Оно осуществляется в любых системах, в том числе в технических и биологических системах. В социальных системах оно может быть понято как воздействие субъекта управления на объект управления с целью перевода его в новое желательное состояние. Однако воздействие в социальных системах может быть разным по объему, содержанию, направленности, силе власти. В зависимости от этого меняется содержание самого понятия управления.

Важнейшим элементом управления является целеполагание.

Цель управления – желаемое, возможное и необходимое состояние системы, которое должно быть достигнуто.

Целеполагание – процесс обоснования и формирования целей развития управляемого объекта на основе анализа общественных потребностей в продукции, услугах, качестве социальных связей, исходя из реальных возможностей их наиболее полного удовлетворения.

Механизмы управления в социальных системах. В обществе сложились два типа механизма управления: стихийный и сознательный. При стихийном механизме упорядочивающее, управляющее воздействие на систему является усредненным результатом столкновения и перекрещивания различных, нередко противоречащих друг другу сил, массы случайных единичных актов; это воздействие автоматически по своей природе и не требует вмешательства людей. Таков, например, рынок – основной регулятор капиталистической экономики, главная управляющая сила производства и определяемой им всей системы общественных отношений.

Наряду со стихийными факторами на любой ступени развития общества действуют сознательные факторы управления, в том числе формируются специфические общественные институты – субъекты управления, т.е. система организаций, осуществляющих целенаправленное воздействие на общество.

Сознательные факторы управления в ходе истории претерпели глубокие изменения – от управления посредством сложившихся и передаваемых из поколения в поколение традиций, обычаев в первобытном обществе до управления обществом на научной основе в условиях социализма.

Факторы воздействия в социальном управлении. Управляемая социальная система является активной (см. § 1.2.2) Поэтому управление в социальных системах тем эффективнее, чем в большей мере поставленные цели соответствуют собственному целеполаганию системы или ее членов, в том числе затрагивают ценности, убеждения, мировоззрение, коренные жизненные интересы (экономические, социальные, политические, духовно-культурные) каждого члена социальной группы.

Источником активности, как известно, являются стимулы. Поэтому, создавая ту или иную систему стимулирования, т.е. представления каких-либо жизненных гарантий по результатам активных действий в соответствии с поставленной целью, можно эффективно управлять социальными процессами, социальным поведением.

Стимул содержит в себе и причину, и цель активности в их нерасчлененном виде, а цель не всегда может быть осознанной. Напротив, мотив активно включает сферу сознания, идеальный образ, что делает саму активность более целесообразной, эффективной. Мотивация к деятельности – это качественно иной способ стимулирования активности, который включает все сущностные силы личности: ценности, идеалы, мировоззрение, направленность человека, «Я-концепцию» и т.п.

Все перечисленные (а также и другие) аспекты могут быть факторами воздействия в социальных системах. Тем самым открывается много возможностей для организации управления в социальных системах.

Современное социальное управление опирается на главные ресурсы: творческий потенциал личности, сильную социальную организацию, интеллектуальную собственность и ее составляющую – информацию.

Типы управленческих отношений. Управленческие отношения отличаются исключительной сложностью, представляя собой единство объективного и субъективного, и реализуются в деятельности субъектов управления.

Организационные отношения – необходимый элемент механизма управления. Они имеют свою внутреннюю структуру, в которой выделяют отношения централизма и самостоятельности, субординации и координации, ответственности, соревнования и др. Они подразделяются на вертикальные и горизонтальные, формальные и неформальные.

Помимо отношений субординации типа «руководители – подчиненные», существуют отношения координации, или отношения между находящимися на одном иерархическом уровне членами организации, направленными на взаимное согласование действий, исходя из совместных задач руководства или исполнения.

Формальное построение организации по подразделениям и должностям создает логическую структуру для организационной деятельности. При реализации управления в активных системах важно учитывать, помимо формальной структуры организации, свойства конкретных людей, а также соотносить формальную структуру с неформальной.

Момент неформальных отношений в организаторской деятельности нередко не учитывается, руководители делают акцент на собственную волю или близкое окружение. Особо актуален момент неформальных коллективных отношений в больших социальных организациях, в обществе в целом, где морально-психологические и нравственные регуляторы имеют все большее значение.

Трактовки понятия «интеллектуальные системы управления». С точки зрения классической социологии [Иванов, 2001] под интеллектуальными системами управления, понимаются системы, состоящие из творческих индивидуальностей, мотивированных к управленческой деятельности, способных достигать наилучших результатов в управлении общественными делами, обеспечивать упорядоченность и высокий уровень организации социальных систем.

Однако применительно к специфике нашего курса мы будем использовать трактовку понятия «интеллектуальные системы», принятую в ИТ [Рыбина, 2010; Лисьев, 2011], а именно:

интеллектуальная система (ИС, англ. intelligent system) – это техническая или программная система, способная решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы.

Интеллектуальные системы изучаются группой наук, объединяемых под названием «искусственный интеллект».

В технологиях принятия решений выделяются:

- интеллектуальная система – информационно-вычислительная система с интеллектуальной поддержкой, решающая задачи без участия человека – лица, принимающего решение (ЛПР);
- интеллектуализированная система – информационно-вычислительная система с интеллектуальной поддержкой, в которой оператор присутствует.

Функции управления. Важнейшей частью любого процесса управления являются его функции:

1. *Планирование* – функция управления, с помощью которой определяются цели деятельности, необходимые для этого средства, а также разрабатываются методы, наиболее эффективные в конкретных условиях. Начальным элементом планирования является составление прогнозов, показывающих возможные направления будущего развития объекта, рассматриваемого в тесном взаимодействии со средой.

2. *Организация* – функция управления, задачей которой является формирование структуры объекта, а также обеспечение всем необходимым для его нормальной работы – персоналом, материалами, оборудованием, зданиями, денежными средствами и т.д. В любом плане всегда имеется стадия организации, т.е. создания реальных условий для достижения запланированных целей.

3. *Мотивация* – функция управления, имеющая целью активизировать работающих и побудить их активно трудиться для выполнения целей, поставленных в планах. Для этого осуществляется материальное и моральное стимулирование работающих, обогащается само содержание труда и создаются условия для проявления творческого потенциала работников и их развития.

4. *Контроль* – функция управления, задача которой состоит в количественной и качественной оценке и учете результатов работы. Главные инструменты выполнения этой функции – наблюдение, проверка всех сторон деятельности, учет и анализ. Среди функций управления контроль выступает элементом обратной связи, так как по его данным производится корректировка ранее принятых решений, планов, а также норм и нормативов.

5. *Координация* – центральная функция управления, обеспечивающая его бесперебойность и непрерывность. Главная задача координации – достижение согласованности в работе всех звеньев системы путем установления рациональных связей (коммуникаций) между ними. Характер этих связей может быть самым различным, так как зависит от координируемых процессов. Наиболее часто используются отчеты, интервью, собрания, компьютерная связь, средства радио- и телевидения, документы. С помощью этих и других форм связи устанавливается взаимодействие между подсистемами, осуществляется маневрирование ресурсами, обеспечивается единство и согласование всех функций управления, а также действий руководителей.

1.4. БАЗОВАЯ МОДЕЛЬ УПРАВЛЕНИЯ В СОЦИАЛЬНЫХ ПРОЦЕССАХ

В теории автоматического управления при постановке задачи управления выделяется управляющий орган и управляемая система (объект управления) (рис. 4). Состояние управляемой системы зависит от трех факторов:

- внешних воздействий;
- воздействий со стороны управляющего органа;
- действий самой управляемой системы, если она является активной.

Задача управляющего органа состоит в том, чтобы осуществить такие управляющие воздействия, чтобы с учетом информации о внешних воздействиях обеспечить требуемое состояние управляемой системы.

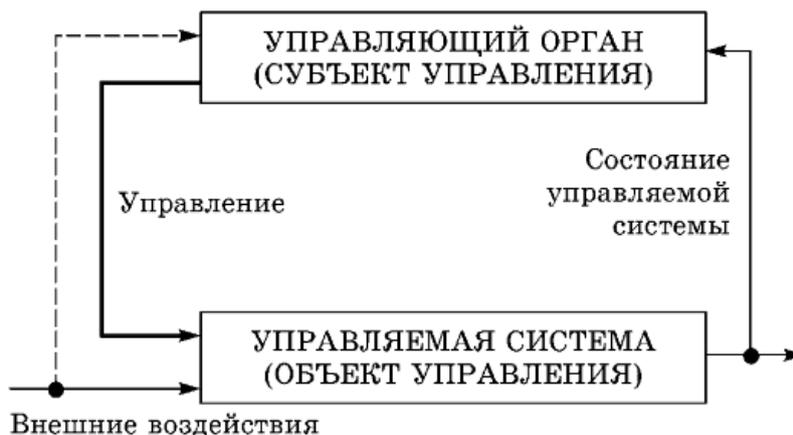


Рис. 4. Структура системы управления

Чтобы вся схема рис. 4 заработала, необходимо иметь модель управляемой системы. В целом можно сказать, что эффективность всей системы управления в решающей степени зависит от того, насколько адекватно выбранная модель управляемой системы отражает ее реальное поведение. Поэтому большая часть дальнейшего изложения посвящена различным вариантам моделирования социальных систем как объектов управления. Однако уже сейчас можно дать некоторые обобщенные классификации задач управления, наиболее применимые для управления в социальных системах.

Классификация по принципу «на что мы влияем»:

- управление составом (набором элементов, из которых состоит управляемая система);
- управление структурой (набором или интенсивностью связей между элементами);
- институциональное управление (ограничениями или нормами деятельности элементов системы);
- мотивационное управление (предпочтениями элементов системы);
- информационное управление (информированностью, т.е. тем объемом информации, которым элементы системы обладают на момент принятия решения).

Классификация по принципу «что мы делим / вокруг чего мы объединяемся»:

- в малых группах – различные варианты моделирования межличностного взаимодействия (см. § 1.5);
- в социальных сетях (см. § 1.6)
 - структурный подход,
 - ресурсный подход,
 - нормативный подход,
 - динамический подход.

1.5. УПРАВЛЕНИЕ В МАЛЫХ ГРУППАХ

При организации управления в малых социальных группах можно выделить две цели, на первый взгляд, противоречащие друг другу. С одной стороны, хочется максимально сохранить творческий потенциал группы, т.е. индивидуальную активность членов группы, иначе они будут действовать как роботы. С другой стороны, хочется обеспечить управляемость группы, т.е. выполнение распоряжений руководства. Эти противоречивые цели можно совместить тогда, когда группа работает в условиях характерного для нее состояния равновесия (см. пример с батальоном в § 1.2.3).

Таким образом, для организации эффективного управления в малых группах целесообразно, взяв за базу кривую гомеостатизации (см. рис. 3), углубить структурирование зоны 2 применительно к экономическим, интеллектуальным и, в особенности, организационным особенностям группы. Это приводит к различным типологиям (стилям) управления.

Стили управления (рис. 5, 6) описываются двумя факторами, отложенными по осям координат. Уточним смысловую характеристику координатных осей.

Сильная ориентация на людей означает, что для руководителя важен микроклимат в коллективе (группе), насколько сотрудники заинтересованы заданием, поручают ли они удовольствие от порученной им работы. Слабая ориентация на людей, в свою очередь, не означает, что руководителя вообще не волнуют нужды его подчиненных; она означает, что в данной ситуации (при решении данной задачи) руководитель непосредственно не решает проблемы создания благоприятного микроклимата, его контакты с подчиненными минимальны.

Ориентация на задачу означает степень структурированности задания (руководитель лишь обозначает проблему или четко указывает, что, как и в какие сроки должно быть сделано). Другими словами, слабая ориентация на задачу вовсе не означает, что руководителя не заботит конечный результат.

Исторически первой была разработана типология К. Левина (рис. 5). Согласно типологии Левина, наилучшим является демократический стиль управления.

Однако идеологии системного анализа лучше отвечает типология управления Херси–Бланшара (рис. 6). Рассматривая организацию как открытую систему, активно взаимодействующую с внешней средой, Херси и Бланшар разработали ситуационную теорию управления. Ее идея состоит в том, что не существует идеального управления, пригодного на все случаи жизни. Его эффективность зависит от адекватности действий руководителя в конкретной ситуации.

Как можно более развернуто охарактеризовать стили руководства по Херси-Бланшару?

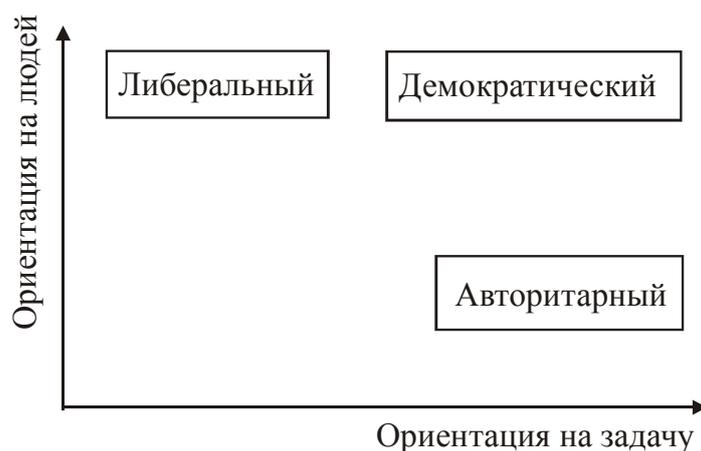


Рис. 5. Стили управления по К. Левину

Предписание. Руководитель точно ставит задачу перед подчиненным, при этом ему не важно, насколько его взгляды разделяются в коллективе, он не принимает предложения подчиненных и не объясняет им происходящих процессов. Работа персонала строится в соответствии с четко определенными инструкциями, строго контролируется руководителем; поощряется четкое следование инструкциям и четкое выполнение задания.

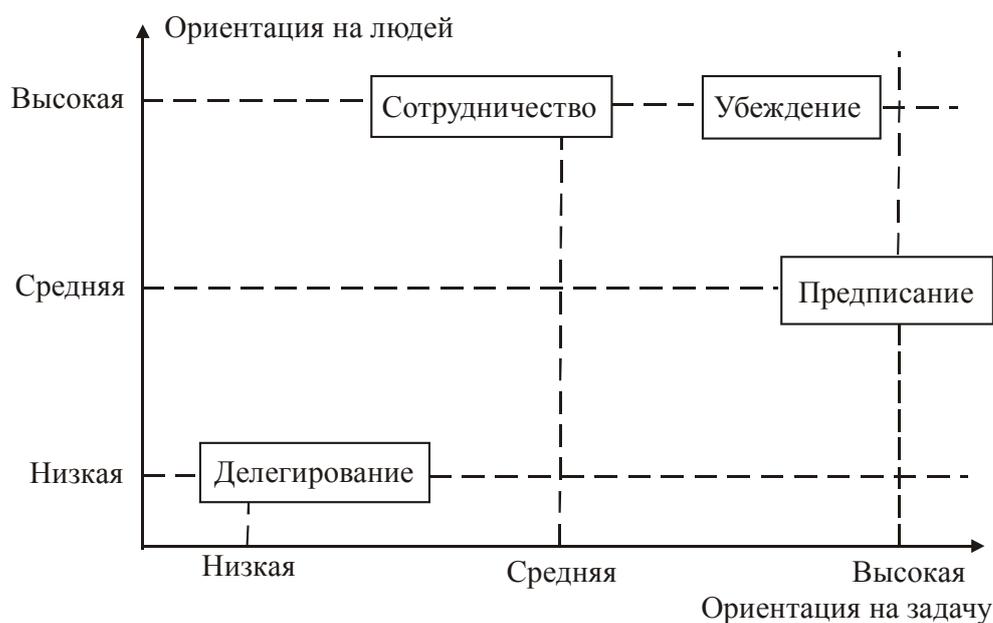


Рис. 6. Стили управления по Херси–Бланшару

Убеждение. Руководитель рекламирует свою позицию и идеи, посредством убеждения стараясь превратить своих сотрудников в союзников. Возникающие идеи могут обсуждаться совместно с подчиненными. Идеи группы принимаются, но при этом руководитель сам контролирует и направляет работу.

Сотрудничество. Руководитель на равных с подчиненными участвует в формировании задачи. Инициатива в определении тактики и ее реализа-

ции принадлежит группе, а задача руководителя – постоянно поддерживать инициативу и заинтересованность подчиненных, наравне с ними участвовать в работе и не слишком контролировать ее.

Делегирование. Руководитель делегирует право принятия решения группе (или подчиненному), однако при необходимости готов оказать помощь. Группа самостоятельно определяет задачи, вырабатывает тактику и выполняет работу. Процесс обсуждения проблем может происходить и без руководителя. Руководитель лишь наблюдает за ходом выполнения работы (не всегда напрямую), но ответственность за конечный результат лежит на нем.

Важно подчеркнуть, что руководитель должен владеть разными стилями руководства в зависимости от обстановки. При выборе конкретного стиля руководства нужно учитывать, по меньшей мере, три фактора:

1. Ситуация (спокойная, стрессовая, неопределенная). В ситуации дефицита времени оправдан авторитарный стиль (крайний пример – военные действия).

2. Задача (уровень ее структурированности). К решению сложных проблем необходимо привлекать экспертов, организовывать дискуссии, и здесь необходим демократический стиль.

3. Группа (ее особенности по полу, возрасту, времени существования). Для сплоченного коллектива, заинтересованного в решении задачи, адекватен демократический стиль; в пределе, в творческих коллективах и при решении творческих задач – либеральный стиль (ситуация типа мозгового штурма и т.д.).

Эффективный руководитель должен иметь в своем арсенале все стили и типологии и использовать их в соответствии с конкретной ситуацией.

1.6. УПРАВЛЕНИЕ В СОЦИАЛЬНЫХ СЕТЯХ

1.6.1 Основные понятия управления в социальных сетях

Конкретизируем общие понятия, предложенные в предыдущих параграфах, применительно к такому важному на сегодняшний день объекту управления, как социальные сети.

В современном обществе все пронизано «сетями» (networks) социальных отношений – устойчивыми системами связей и контактов между индивидами, которые невозможно вписать в рамки традиционной дихотомии «рынок–иерархия», т.е. стихийный или сознательный механизмы управления. Эти сети неформальных отношений позволяют находить работу, обмениваться информацией, разрешать большинство всех проблем и конфликтов, минуя судей и адвокатов. Т.е. экономические институты через социальные связи становятся социальными конструкциями.

Под социальной сетью [Губанов, 2010] на качественном уровне понимается социальная структура, состоящая из множества агентов (субъектов,

индивидуальных или коллективных – например, индивидов, семей, групп, организаций) и определенного на нем множества отношений (совокупности связей между агентами – например, знакомства, дружбы, сотрудничества, коммуникации).

С чем связана привлекательность социальных сетей? Они предоставляют своим членам следующие возможности:

- получение информации (в том числе обнаружение ресурсов) от других членов социальной сети;
- верификация идей через участие во взаимодействиях в социальной сети;
- социальная выгода от контактов (сопричастность, самоидентификация, социальное отождествление, социальное принятие и др.);
- рекреация (отдых, времяпрепровождение).

Формально социальная сеть представляет собой граф $G(N, E)$, в котором $N = \{1, 2, \dots, n\}$ – конечное множество вершин (агентов), а E – множество ребер, отражающих взаимодействие агентов.

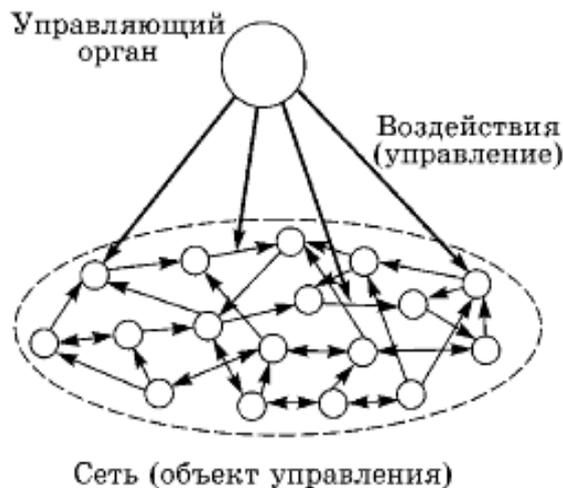


Рис. 7. Социальная сеть как объект управления

Интерпретация социальной сети как объекта управления поясняется рис. 7.

Если предполагать, что вершины графа «пассивны», т.е. не обладают собственными предпочтениями и информированностью, управление может заключаться в целенаправленном воздействии на следующие компоненты сети:

1. состав управляемой системы (тогда управление может заключаться в удалении или добавлении вершин);
2. структура управляемой системы (тогда управление может заключаться в удалении или добавлении дуг);

3. значения параметров, соответствующих вершинам графа (значения состояний элементов сети) или его дуг (значения параметров, отражающих силу связи между элементами сети).

Заметим, что большинство теоретических моделей управления в социальных сетях ограничиваются моделью 3, хотя модели типа 1 и 2 являются весьма перспективными.

1.6.2. Подходы к моделированию управления в социальных сетях

Эти подходы, по существу, конкретизируют применительно к социальным сетям приведенную в § 1.4 классификацию по принципу «что мы делим / вокруг чего мы объединяемся».

Структурный подход. В структурном подходе все участники сети рассматриваются как вершины графа, которые влияют на конфигурацию ребер и других участников сети. Основное внимание уделяется геометрической форме сети и интенсивности взаимодействий (весу ребер), поэтому исследуются такие характеристики, как взаимное расположение вершин, центральность, транзитивность взаимодействий.

При структурном анализе изучается поведение вершин в процессе кластеризации и типичных временных характеристик социальных сетей – например, как меняется структура сети в процессе роста или как меняются поведение и распределение связанных компонентов графа.

Большое значение придается определению сообществ в социальных сетях. Например, задачу определения регионов сети, внутри которых происходит активное взаимодействие участников, алгоритмически можно отнести к задаче о разделении графов: необходимо разделить сеть на плотные регионы на основе поведения связей между вершинами.

Ресурсный подход. Ресурсный подход рассматривает возможности участников по привлечению индивидуальных и сетевых ресурсов для достижения определенных целей и дифференцирует участников, находящихся в идентичных структурных позициях социальной сети, по их ресурсам. В качестве индивидуальных ресурсов могут выступать знания, престиж, богатство, раса, пол. Под сетевыми ресурсами понимаются влияние, статус, объем и характер информации. Основным показателем, определяющим различия в ресурсах участников сети, является сила структурной позиции участника.

Важная задача данного направления – анализ содержания социальных сетей, т.е. сетевого контента. Использование контента сети помогает значительно улучшить качество выводов при анализе социальных сетей, например, в задачах кластеризации и классификации.

Нормативный подход. Нормативное направление изучает уровень доверия между участниками, а также нормы, правила и санкции, влияющие на поведение участников в социальной сети и процессы их взаимодей-

ствий. В этом случае анализируются социальные роли, которые связаны с данным ребром сети, например, отношения руководителя и подчиненного, дружеские или родственные связи. Так как в основе социальных сетей лежит взаимодействие между различными участниками, естественно предположить, что это взаимодействие оказывает влияние на участников в терминах их поведения.

Основные вопросы этого направления: как моделировать влияние на основе информации об участниках; как моделировать распространение влияния; кто из участников наиболее влиятелен в процессе распространения.

Перечислим некоторые типичные задачи, решаемые в рамках нормативного подхода.

В социальных сетях содержится большое количество личной информации об участниках, например, интересы, друзья, демография и др. Это может привести к несанкционированному распространению личной информации в сетях. В решении такого типа задач полезно применять модели на основе механизмов конфиденциальности.

Так как для эффективности и устойчивости социальной сети важны функциональные роли ее участников, социальная сеть может быть инструментом для выявления экспертов в конкретной области. Помимо экспертов, при анализе социальных сетей представляют интерес так называемые брокеры (лидеры) – люди, которые играют роль посредников в социальной сети, связывая между собой группы людей, налаживая связи между специалистами и тем самым открывая им доступ к информации.

Динамический подход. Динамический подход – направление в изучении социальных сетей, в котором объектами исследований являются изменения в сетевой структуре с течением времени: появляются новые участники, некоторые участники прекращают взаимодействие, возникают новые связи, некоторые связи устаревают, так как участники перестают взаимодействовать. Это приводит к изменениям в структуре социальных сетей в целом и в отдельных сообществах.

Здесь рассматриваются следующие вопросы: согласно каким законам происходят долгосрочные изменения между крупными сообществами в социальных сетях, существуют ли какие-либо стационарные конфигурации социальной сети, как развиваются сами сообщества во времени, какие изменения могут происходить, как можно отследить и представить их.

Важной задачей является прогноз формирования связей в социальных сетях. В большинстве приложений для анализа социальных сетей связи считаются динамическими и могут изменяться с течением времени. В процесс прогнозирования связей могут быть вовлечены как структура сети, так и информация об особенностях различных вершин

1.6.3. Параметры управления в социальных сетях: мнение, влияние, доверие

Поведение агента социальной сети (рис. 8) может зависеть от различных факторов:

- индивидуального (определяется внутренним целеполаганием агента и не зависит от внешнего влияния);
- социального (определяется взаимовлияниями с другими агентами сети);
- административного (определяется воздействием управляющего органа – центра)

Основными параметрами управления, т.е. предметом управления, в социальной сети являются: мнение агента, влияние агента / доверие к агенту, репутация агента.

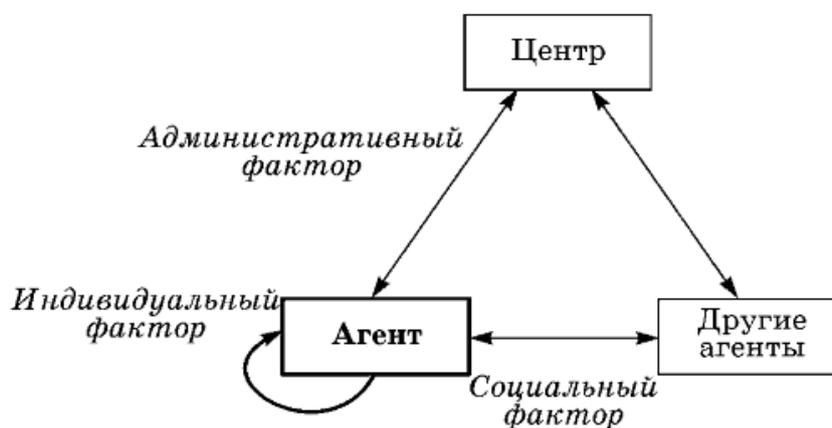


Рис. 8. Факторы, влияющие на поведение агента в социальной сети

Мнение. Выражая свое мнение, т.е. высказывая свое суждение (оценку) по какому-то вопросу, агент получает ту или иную социальную выгоду, в том числе сопричастность, самоидентификацию, социальное отождествление, социальное принятие и др.

Мнение агента может быть выражено в различной форме (рис. 9):

- в личной переписке – текст;
- в блогах – комментарий;
- при помощи механизмов, предоставляемых разработчиками сети (факты, опросы, оценки в баллах и пр.).

Влияние и доверие. Влияние – процесс и результат изменения индивидом (субъектом влияния) поведения другого субъекта (индивидуального или коллективного объекта влияния), его установок, намерений, представлений и оценок (а также основывающихся на них действий) в ходе взаимо-

действия с ним. Другими словами, влияние – способность воздействовать на чьи-либо представления или действия.

Различают направленное и ненаправленное влияние. Направленное (целенаправленное) влияние использует в качестве механизмов воздействия на объект влияния убеждение и внушение. При этом субъект влияния ставит перед собой задачу добиться определенных результатов (например, выбора определенных действий) от объекта влияния. При ненаправленном (нецеленаправленном) влиянии такая задача не ставится; при этом субъект влияния может даже не подозревать о том влиянии, которое он оказывает на объект влияния.

В социальной сети агенты часто не имеют достаточной для принятия решений информации или не могут самостоятельно обработать ее, поэтому их решения могут основываться на наблюдаемых ими решениях или представлениях других агентов. В этом состоит механизм социального влияния.



Рис. 9. Формы выражения мнений агентов социальной сети

Социальное влияние реализуется в двух процессах:

- коммуникация (в ходе общения, обмена опытом и информацией, обсуждения тех или иных вопросов с авторитетными для агента соседями он приходит к определенным представлениям, установкам, мнениям);
- сравнение (в поисках социальной идентичности и социального одобрения агент принимает представления и действия, ожидаемые от него другими агентами в данной ситуации).

При коммуникативном подходе к влиянию агенты могут прийти к сходным представлениям, но не обязательно к сходному поведению. При сравнении же агент обычно косвенным образом копирует поведение.

Очевидно, поведение агента определяется не только представлениями, но и ограничениями, с которыми он сталкивается. Поэтому агенты со схожими представлениями могут вести себя по-разному, и наоборот, агенты с разными представлениями могут вести себя одинаково.

Таким образом, феномен влияния агента в сети связан с его информированностью и позволяет организовать информационное управление. Здесь выделяется три вложенных уровня управления:

- модель информационного влияния описывает зависимость поведения субъекта от его информированности; т.е. здесь фактором управления являются информационные воздействия на него;
- модель информационного управления описывает, какими должны быть информационные воздействия на управляемого субъекта, чтобы добиться от него желаемого поведения;
- модель информационного противоборства позволяет описать взаимодействие нескольких субъектов с несовпадающими интересами, одновременно воздействующих на один и тот же управляемый субъект.

Доверие в сети – это степень взаимного влияния агентов друг на друга. Управляя взаимным доверием агентов, субъект управления может добиваться требуемых результатов.

Репутация. В традиционном смысле репутация – создавшееся общее мнение о достоинствах или недостатках кого-либо, общественная оценка. Применительно к социальным сетям репутация рассматривается в двух смыслах:

- как ожидаемая другими агентами норма деятельности данного агента (т.е. какого поведения от него ожидают остальные);
- как весомость мнения агента, определяемая успешностью его предыдущей деятельности или успешностью его предсказаний (решений) в будущем.

Репутация агента возрастает, если мнение или действия агента совпадают с ожиданиями остальных членов сети или с тем, что впоследствии считается нормой (например, приводит к эффективной деятельности). Наоборот, репутация снижается при нарушении принятых в сети норм поведения, при принятии неэффективных решений.

1.6.4. Эффекты и свойства в реальных социальных сетях

Сложные комбинации факторов и параметров управления, существующие в конкретных социальных сетях, могут приводить к разнообразным эффектам, имеющим место в сетях:

- изменение мнений агента под влиянием других членов социальной сети;
- различная значимость мнений (влиятельности, доверия) одних агентов для других агентов;
- различная степень подверженности агентов влиянию (конформизм, устойчивость мнений);
- существование косвенного влияния в цепочке социальных контактов, которое уменьшается с увеличением «расстояния».
- существование «лидеров мнений» (агентов с максимальным «влиянием»), формализация индексов влияния;
- существование порога чувствительности к изменению мнения окружающих;
- локализация групп («по интересам», с близкими мнениями) внутри сети;
- наличие специфических для конкретной сети социальных норм;
- наличие факторов «социальной корреляции» (общих для групп агентов);
- существование (обычно менее значимых) внешних факторов влияния (реклама, маркетинговые акции) и соответственно внешних агентов (средства массовой информации, производители товаров и т.п.);
- наличие стадий – характерных этапов динамики мнений членов социальной сети (например, процесса диффузии инноваций);
- лавинообразные эффекты (каскады), формализация условий их возникновения и свойств распространения;
- воздействие структурных свойств социальных сетей на динамику мнений:
 - чем больше у агента связей, тем больше у него возможностей через свое окружение повлиять на всю сеть, но больше и уязвимость к чужому влиянию;
 - эффект кластеризации и сильной связи (чем выше плотность связей активных агентов-соседей, тем больше вероятность активизации связанного с ними агента);
 - эффект локальной промежуточности и слабой связи (чем больше промежуточное значение агента, тем больше его значение в распространении мнения/информации из одной части сети в другую (роль информационного брокера), но меньше его влияние на агента-соседа);
- возможность образования группировок, коалиций;
- неполная и/или асимметричная информированность агентов;
- нетривиальная взаимная информированность (рефлексия) агентов;
- игровое взаимодействие агентов;
- оптимизация информационных воздействий;

Умея описывать свойства сети в зависимости от ее параметров и выделив управляемые переменные (т.е. параметры, которые подвергаются целенаправленному управлению со стороны управляющего органа), можно ставить и решать задачи управления в социальных сетях. Именно этот факт является основой для применения интеллектуальных систем в управлении социальными системами.

2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СОЦИАЛЬНЫХ СЕТЕЙ

Как было показано в § 1.6.3, механизмы влияния играют решающую роль в организации социального управления. К настоящему времени предложены разнообразные подходы к моделированию этих механизмов, которые влекут за собой разные модели управления (см. § 1.4). Достаточно большой список этих подходов приводится в [Губанов, 2010]. Необходимо отметить, что этот список постоянно растет, и для подбора модели в конкретной ситуации целесообразно пользоваться статейным материалом.

2.1. ПОРОГОВЫЕ МОДЕЛИ

Модель с линейным порогом. Агент моделируется как узел социальной сети (вершина графа). Агент может находиться в активном и неактивном состояниях, причем возможен переход только из неактивного состояния в активное (обратный переход не допускается).

Агент i испытывает влияние a_{ij} каждого своего j -го соседа в сети так, что выполняется условие

$$\sum_{j \text{ активный узел-сосед } i} a_{ij} \leq 1,$$

при этом агент становится активным в зависимости от выбранного им порога $0 \leq \varphi_i \leq 1$. Таким образом, условие активации имеет вид

$$\sum_{j \text{ активный узел-сосед } i} a_{ij} \geq \varphi_i.$$

Значение порога φ_i может быть различным для каждого агента.

Модель независимых каскадов рассматривает сеть как систему взаимодействующих частиц. Узел сети (агент) определяется так же, как и в модели с линейным порогом.

Когда агент i становится активным, он получает шанс активировать на следующем (и только на следующем) шаге каждого из своих соседей j с вероятностью p_{ji} (причем агента j могут пытаться независимо активировать и другие агенты). В литературе показано, что модели с линейным порогом и модели независимых каскадов в определенном смысле эквивалентны.

2. 2. МОДЕЛИ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ В СЕТИ

Модели распространения основываются на метафоре распространения какого-то ресурса по сети.

Модели просачивания и заражения, используемые в различных приложениях – от моделей эпидемий до исследования нефтяных месторождений, представляют собой популярный способ изучения распространения информации (инноваций). В англоязычной литературе модель полу-

чила название SIR – по первым буквам основных стадий описываемого процесса: susceptible (восприимчив), infected & infectious (заражен и заразен), recovered/removed (выздоровел или умер).

Классический пример применения этой модели – модель распространения эпидемии. Содержательно она описывает цикл заболевания:

- первоначально человек восприимчив к заболеванию;
- если он входит в контакт с инфицированным, то заражается с некоторой вероятностью β ;
- впоследствии через некоторый период времени человек становится здоровым, приобретая иммунитет, или умирает;
- иммунитет со временем снижается, и человек снова становится восприимчивым к болезни.

Таким образом, модель описывает переход $S \rightarrow I \rightarrow R$.

Рассмотрим применение этой модели к распространению информации в социальной сети. Блоггер (человек, который ведет блог, т.е. сетевой дневник) может прочитать блог друга (восприимчив), посвященный некоторой теме, а затем может и сам написать об этой теме (инфицирован) и позже вернуться к ней (восприимчив).

Построим математическую модель процесса $S \rightarrow I \rightarrow R$. Введем следующие обозначения:

$S(t)$ – численность подгруппы людей, еще не инфицированных или восприимчивых к болезни в момент времени t ;

$I(t)$ – численность подгруппы инфицированных людей;

$R(t)$ – численность подгруппы выздоровевших людей.

Пусть

$$N = \text{const} = S(t) + I(t) + R(t).$$

Опишем динамику процесса.

$$dS(t) / dt = -\beta N S(t) / N I(t) = -\beta S(t) I(t),$$

т. е. каждый из инфицированных в единицу времени, контактируя с восприимчивыми к болезни, заражает их с вероятностью β ;

$$dR(t) / dt = g I(t),$$

т.е. инфицированные выздоравливают через средний период времени $1/g$; соответственно

$$dI(t) / dt = \beta S(t) I(t) - g I(t).$$

Ключевым показателем модели является «эпидемический порог» λ_c – критическая вероятность заражения соседа, при превышении которой «инфекция» распространяется по всей сети. Эпидемический порог зависит от свойств графа сети, например, от числа вершин, распределения связей, коэффициента кластеризации. Поэтому распространение инфекции сильно зависит от выбранной модели представления графа сети.

Если социальную сеть представить случайным графом, то инфекция с вероятностью заражения выше порога экспоненциально быстро размножается:

$$\lambda = \beta / g > \lambda c;$$

а инфекция с вероятностью заражения ниже порога экспоненциально быстро «вымирает».

Более реалистичной моделью социальной сети является безмасштабный граф (см. § 4.5), в котором некоторые вершины связаны с тысячами и даже миллионами других вершин, в большинстве своем имеющих всего по несколько связей (т. е. отсутствует характерный масштаб). В таком графе распределение количества связей узлов описывается степенным законом. В безмасштабных сетях эпидемический порог отсутствует – эпидемия охватит всю сеть, если возникнет инфекция.

Однако в блогосфере многие обсуждаемые темы могут распространяться без возникновения эпидемий, поэтому порог все же отличен от нуля. Следовательно, нужно или выбирать более адекватную модель сетей со степенным распределением (учесть более «тонкие» свойства таких сетей, например, наличие эффекта кластеризации) или модифицировать модель передачи инфекции (ослаблять вероятность заражения с увеличением «дистанции от инициатора»).

Лавинообразные модели. При определенных условиях модель просачивания и заражения описывает лавинообразные процессы в сетях. К лавинообразным можно отнести широкий круг явлений в природе и обществе, в том числе: процессы горения и взрыва, размножение вирусов или накопление продуктов распада в живом организме, социальные конфликты с митинговым характером протекания, валютные и биржевые паники, ажиотажный спрос на те или иные товары, распространение технологических и управленческих новшеств, в том числе информационных систем и технологий, информационные воздействия на индивидуальных и коллективных субъектов.

Их характерной чертой является лавинообразный, по типу цепной реакции, характер распространения и развития процессов и, как следствие, наличие внутренних или внешних связей, характеризующихся большим, чаще всего экспоненциальным, изменением одного параметра при небольшом изменении другого. Это, как правило, быстропротекающие процессы.

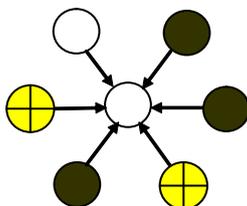


Рис. 10. Воздействие контактирующих на здорового члена группы:

В таких процессах нужно учитывать, что на здорового члена группы могут воздействовать одновременно несколько контактирующих комму-

никаторов, в том числе здоровые (белый кружок), паникующие (заразители) (черный кружок) и оздоровители (кружок с крестиком).

Тогда модель $S \rightarrow I \rightarrow R$ переходит в систему уравнений:

$$\begin{cases} \frac{dS}{dt} = -[p_{озл} + p_{нзл}] \cdot S \\ \frac{dI}{dt} = p_{озл} \cdot S - p_{леч} \cdot I \\ \frac{dR}{dt} = [p_{нзл} \cdot S + p_{леч} \cdot I] \end{cases}$$

где S, I, R – количество восприимчивых к панике, охваченных паникой и невосприимчивых (иммунизированных); $p(\cdot)$ – соответствующие вероятности. Значения вероятностей зависят от ряда параметров сети, в том числе:

- дифференцированности группы по экономическим, социально-психологическим и ролевым характеристикам субъектов;
- степени и характера взаимовоздействия коммуникаторов;
- наличия конструктивных и деструктивных источников информации, в том числе СМИ.

Модель Изинга (намагничивания). В своем исходном варианте модель описывает процесс намагничивания материала. В этом процессе учитывается взаимодействие только ближайших атомов-соседей в кристаллической решетке, причем энергия взаимодействия

$$E_{ij} = -J s_i s_j,$$

где s – спин атома, равный ± 1 ; J – константа обменного взаимодействия. Полная энергия E может быть найдена суммированием попарных энергий взаимодействия по всей решетке:

$$E(s) = -J \sum_{i \sim j} s_i s_j ;$$

при наличии внешнего поля h

$$E(s) = -J \sum_{i \sim j} s_i s_j + h \sum_i s_i .$$

Для ферромагнетика $J > 0$, и энергия минимальна для спинов, направленных в одну сторону. При температуре ниже критической большая часть спинов атомов ориентирована одинаково, при более высокой температуре ориентация спинов становится случайной.

С помощью этой модели можно описать конформность или независимость в большой социальной группе. Для распространения этой модели на поведение социальной группы делается предположение, что влияние ближайших соседей является определяющим, а аналогом температуры является готовность группы мыслить творчески, готовность принять новые идеи. Внешним полем для социальной группы является влияние «авторитета».

Модели на основе клеточных автоматов. Для описания процессов распространения информации в социальной сети можно использовать модельное представление клеточных автоматов.

Клеточный автомат состоит из набора объектов (в данном случае агентов), обычно образующих регулярную решетку. Состояние отдельно взятого агента в каждый дискретный момент времени характеризуется некоторой переменной. Рассматриваемые состояния объекта синхронно изменяются через дискретные интервалы времени в соответствии с неизменными локальными вероятностными правилами, которые могут зависеть от состояния переменных, описывающих ближайших соседних агентов в окрестности данного агента, а также, возможно, от состояния самого агента.

Каждый агент в большой сети относится к одной персональной сети, агенты в которой связаны сильными (стабильными и постоянными) связями. Агент также имеет слабые связи с агентами из других персональных сетей. Вероятность того, что информированный агент повлияет по сильной связи на неинформированного агента (т. е. последний станет информированным), в данный период времени равна β_s , а по слабой – β_w ($\beta_s > \beta_w$). Также неинформированные агенты в данный момент времени с вероятностью α становятся информированными благодаря рекламе и другим маркетинговым приемам.

Эмпирические данные показывают, что влияние, достигаемое посредством эффекта «из уст в уста», существенно выше, чем влияние внешнего маркетинга, т.е. $\beta_s > \beta_w > \alpha$.

В [Губанов, 2010] представлены результаты имитационного эксперимента с такой моделью. Задавались следующие параметры: размер каждой персональной сети, число слабых связей для каждого агента, вероятности β_s , β_w и α . В результате были выявлены следующие закономерности:

- хотя вероятность распространения по слабым связям ниже, т.е. $\beta_s > \beta_w$, но влияние слабых связей на скорость распространения информации, по крайней мере, такое же, как и сильных связей.
- в начальной фазе большее влияние в информировании агентов имеет реклама, но в дальнейшем ее роль незначительна;
- в следующей фазе информация распространяется в персональных сетях благодаря сильным связям; по мере того, как информированных агентов в таких сетях становится больше, эффект сильных связей ослабляется, и возрастает роль слабых связей в активации новых сетей.
- при увеличении размера персональной сети роль сильных связей увеличивается, а слабых – уменьшается. При увеличении количества слабых связей эффект от сильных связей снижается, а от слабых – увеличивается. При усилении рекламы эффект от сильных связей немного увеличивается, а от слабых – уменьшается.

2. 3. МОДЕЛИ ДИФФУЗИИ ИННОВАЦИЙ

Модель диффузии инноваций традиционно применяется для описания распространения нововведений в крупномасштабной сети. Исследователи в этой области предпринимают попытку объяснить, какие условия увеличивают или соответственно уменьшают вероятность принятия нововведения членами социальной системы, с какой скоростью нововведения распространяются в социальной системе.

Динамика процесса распространения изменений традиционно моделируется S-образной (логистической) кривой (рис. 11). Отметим, что такая кривая характеризует большой класс процессов распространения, в том числе любой инфекционный процесс, процесс научения, диффузию инноваций.

S-образная функция содержит три фазы развития: первая – формирование базы развития (медленный рост), вторая – резкий рост, третья – насыщение (медленный рост). Одним из главных факторов, определяющих скорость процессов диффузии, является межличностное общение между сторонниками данной инновации и теми, кто еще колеблется или вообще ничего не слышал о предлагаемом нововведении.

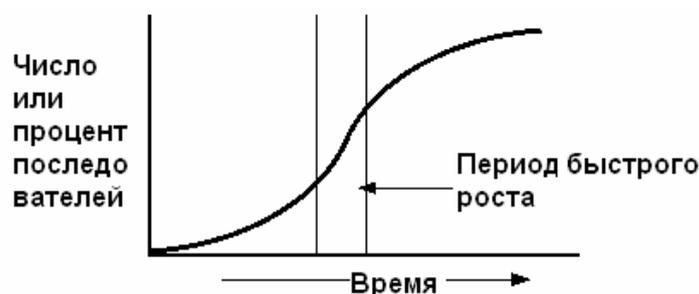


Рис. 11. Логистическая кривая

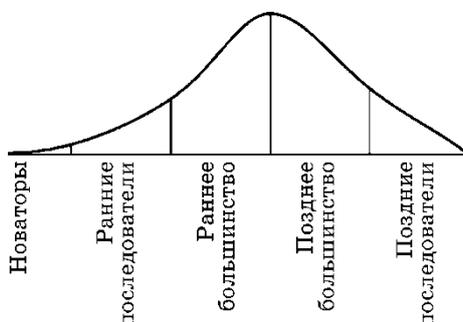


Рис. 12. Кривая стадий

На логистической кривой (рис. 11) различают стадии (рис. 12):

- новаторы (innovators, начинающие первыми воспринимать и использовать нововведение),
- ранние последователи (early adopters, начинающие воспринимать и использовать нововведение вскоре после его появления),
- раннее большинство (early majority, воспринимающие нововведение после новаторов и ранних последователей, но раньше большинства других агентов), позднее большинство (late majority, воспринимающие нововведение после широкого его распространения)
- поздние последователи (late adopters, воспринимают последними).

Новаторов можно охарактеризовать как нонконформистов и «оригиналов», ранних последователей – как агентов, легко поддающихся социальному нормативному и информационному влиянию (или имеющих «нюх» на перспективное), поздних последователей – как трудно поддающихся влиянию и устойчивых агентов в сети.

Рассмотрим основные понятия модели диффузии инноваций:

- диффузия – процесс, посредством которого нововведение распространяется по коммуникационным каналам во времени и в пространстве среди членов социальной системы;
- нововведение – идея, мнение, технология (метод), продукт или любой другой объект, воспринимаемый агентом как новый;
- коммуникация – процесс, посредством которого участники создают и обмениваются информацией друг с другом для достижения взаимного понимания и трансляции нововведений.

Нововведения привносятся в социальную систему агентами изменений (gatekeepers), а затем постепенно принимаются многими агентами, которые передают информацию о нововведении друг другу. Межличностные контакты агентов и средства массовой информации (коммуникационные источники) предоставляют информацию о нововведении (по коммуникационным каналам) и влияют на установки, диспозиции, представления и, в конечном итоге, на решения агентов о принятии нововведения. Так нововведение распространяется через социальную систему, при этом важную роль играет природа социальной системы.

Процесс принятия нововведения агентом проходит через следующие стадии:

- а) знание (агент ознакомлен с новшеством, но не имеет о нем полной информации);
- б) убеждение (агент испытывает интерес, формирует благоприятное или неблагоприятное отношение (установку) к нововведению, осуществляет поиск дополнительной информации);
- в) решение (агент мысленно взвешивает преимущества и недостатки принятия нововведения в текущей или ожидаемой ситуации, решает, стоит ли использовать его (принять или отклонить));
- г) апробация/выполнение (агент использует нововведение);

д) подтверждение (агент оценивает результаты и принимает решение о дальнейшем использовании).

Решение агента о принятии нововведения принимается на основе анализа затрат и выгод. Однако часто возникает неопределенность при принятии решения, которая влияет на скорость распространения инноваций в целом и обусловлена свойствами нововведения, которые оцениваются конкретным агентом:

- 1) относительные преимущества нововведения перед имеющимися аналогами, которые зачастую выражаются в экономических или социальных категориях (прибыльность, экономичность, затраты и т.п.);
- 2) совместимость нововведения – степень соответствия инновации существующей системе ценностей (определяется культурными нормами социальной системы), прошлому общественному и/или индивидуальному опыту и потребностям потенциального последователя;
- 3) сложность нововведения – степень простоты и легкости для понимания, использования или приспособления к нововведению; предполагается, что сложность инновации негативно связана с ее принятием;
- 4) простота апробации/использования – возможность апробации нововведения в ограниченных масштабах;
- 5) коммуникационная наблюдаемость – степень, с которой нововведение и его результаты могут быть замечены и оценены другими агентами.

Эмпирически установлено, что зачастую небольшие изменения в состояниях вершин сетей могут привести к каскадным (лавинообразным) изменениям (локальным, затрагивающим окружение инициатора, и глобальным, ограниченным только размером всей сети). Это – эффект «из уст в уста», или по-русски – сарафанного радио. Этот эффект уже частично рассматривался ранее.

Как же можно моделировать диффузию инноваций? Для этого предложены различные специализированные модели, примеры которых приводятся ниже.

Модель сетевой автокорреляции. Диффузия инноваций рассматривается как детерминированный дискретно-временной линейный процесс, в котором установки (attitude) агента изменяются под влиянием других агентов:

$$y_{t+1} = W y_t,$$

где y – вектор установок, изменяющихся во времени t ; W – матрица влияний (w_{ij} – значение влияния j -го агента на i -го). Если обозначить через $A = \|a_{ij}\|$ матрицу смежности в социальной сети, тогда матрица влияния может определяться, например, следующим образом:

$$w_{ij} = a_{ij} / (a_{ij} + 1);$$

или

$$w_{ij} = r_j a_{ij} / (r_i + \sum_k r_k a_{ik}),$$

т.е. влияние j -го агента определяется числом имеющихся у него ресурсов r_j .

Модель подражательного поведения. Описана модель подражательного поведения, в которой у каждого из агентов возможно одно из двух действий (так называемый бинарный выбор). Каждый агент характеризуется априорной вероятностью выбора того или иного действия, склонностью прислушиваться к мнению других агентов (какие действия планируют выбрать они) и матрицей влияний. Тогда апостериорная вероятность выбора агентом определенного действия вычисляется аналитически по формуле полной вероятности (фактически – в терминах марковских моделей – рассматриваются два момента времени). Это дает возможность исследовать различные случаи принятия агентами решений под влиянием окружения, хорошо интерпретируемые содержательно.

R^* -модели социального влияния. Для анализа связей и отношений между агентами в социальной сети можно учитывать информацию о вероятностном распределении структурных характеристик (конфигураций или подграфов) в генеральной совокупности сетей с заданными свойствами (например, с заданными вершинами). В этом случае можно проверять гипотезы о распределениях, значимости параметров модели, пригодности модели для описания данных.

Для моделирования доступна информация о наблюдаемой социальной сети, рассматривающейся как реализация из генеральной совокупности сетей. Предполагается, что в сетях присутствуют конфигурации выбора ($i \rightarrow j$) и взаимности ($i \leftrightarrow j$). Можно также предположить, что существуют более сложные конфигурации: транзитивность, экспансия, посредничество и т.п. Если допустить, что сеть достаточно однородна, то можно выделить различные изоморфные классы конфигураций и по значениям двух параметров оценить, насколько часто встречаются эти классы в сети.

Модели «диффузии инноваций», связанные с формированием общественного мнения (т. е. само общественное мнение является нововведением – инновацией). Наибольшее распространение получила двухступенчатая модель, в которой средствами массовой информации сначала формируются мнения так называемых «лидеров мнений» (имеющих статус хорошо информированных, уважаемых или просто характеризующихся большим количеством связей), а затем посредством лидеров формируются мнения «обычных» агентов.

Однако против такой простой модели имеются определенные возражения. Во-первых, не только лидеры влияют на обычных агентов, но и обычные агенты влияют на лидеров; влияние может передаваться более чем на два шага. Более того, многие математические модели не требуют введения в явном виде предположения о наличии лидеров мнений или каких-то «особенных» индивидов для формирования S-образной кривой диффузии инноваций.

В статье [Watts, 2007] подробно рассматривается роль лидеров в распространении нововведений (насколько изменения мнений таких лидеров приводит к крупным каскадным изменениям мнений в сети), причем в качестве базовой используется простейшая модель линейного порога:

$$P[\text{принять альтернативу СВ}] = \begin{cases} 1, & \text{если } r_i \geq f_i \\ 0, & \text{если } r_i < f_i \end{cases}$$

Здесь f_i – порог, r_i – доля агентов, выбравших альтернативу СВ. Как оказалось, в большинстве случаев лидеры лишь умеренно «важнее» обычных агентов: фактически к возникновению больших каскадов приводит влияние одних легко поддающихся влиянию агентов на других, столь же легко поддающихся влиянию. Этот эффект авторы объясняют известным феноменом подражания: вероятность того, что i -й агент предпочтет альтернативу СВ (вместо альтернативы СА), увеличивается с числом других агентов, выбравших СВ.

В этой же статье предлагается критерий выявления лидеров. Предполагается, что i -ый агент в популяции размером N влияет на n_i других выбираемых случайно агентов. Число n_i берется из распределения влияния $p(n)$ (среднее $n_{\text{avg}} \ll N$) и означает влияние i -го агента на n_i других относительно данной проблемы. Если считать, что в сети влияния все агенты могут (прямо или косвенно) влиять друг на друга, то можно определить лидеров мнений как агентов, входящих в верхний дециль распределения влияния $p(n)$.

В [Watts, 2007] рассматривается динамика влияния. В начальной стадии агенты не активны (имеют состояние 0), за исключением одного случайно выбранного, так называемого активного инициатора i (лидера мнений), имеющего состояние 1. Этот инициатор может активировать соседей далее по цепочке, иницируя каскад. Если большое число ранних последователей – агентов, непосредственно связанных в рамках сети с инициатором, связано между собой, то может возникнуть глобальный каскад, хотя в целом такие последователи могут составлять небольшую часть всей популяции.

Какие же факторы влияют на появление каскада? На этот вопрос авторы [Watts, 2007] отвечают путем имитационных экспериментов.

1. Размер каскадов, генерируемых одиночными инициаторами, сильно зависит от «средней плотности» сети n_{avg} : если это значение мало, то многие агенты уязвимы, но сеть недостаточно плотна для распространения, и, в конечном итоге, активируется только небольшая часть сети; если же значение n_{avg} велико, то сеть сильно связана, но для активации агентам требуется большое число уже активированных соседей, т.е. небольшое число инициаторов не приведет к образованию глобального каскада. Только средний интервал – так называемое «окно каскадов» –

может привести к образованию глобальных каскадов. В этом промежутке и лидеры, и обычные агенты могут инициировать каскады.

2. Способность агента инициировать каскад зависит скорее от глобальной структуры сети, нежели от персональной степени влияния агента. Если в сети в принципе могут возникать каскады, то любой агент может их инициировать, если нет, то никто. Данное утверждение не зависит от значения порога f , так как последнее просто одинаково сдвинет «окно каскада» и для лидеров, и для обычных агентов.
3. Каскады, инициируемые лидерами, ненамного больше каскадов, инициируемых обычными агентами, за исключением узких границ «окна каскадов», в пределах которых лидеры существенно значимее, чем обычные агенты. С другой стороны, лидеры могут оказать ключевую роль в инициировании глобальных каскадов в качестве образующих критическую массу ранних последователей.
4. Если сеть имеет низкую плотность (n_{avg} примерно равно нижней границе «окна каскадов»), то ранние последователи в среднем более влиятельны (т.е. $n_i > n_{avg}$), но если сеть имеет высокую плотность (n_{avg} у верхней границы «окна каскадов»), то ранние последователи в среднем менее влиятельны (т.е. $n_i < n_{avg}$). Объясняется это тем, что агенты с высоким влиянием (у которых велико n_i) менее уязвимы, но при активации потенциально способны активировать больше других агентов. В то же время эксперименты показывают, что, хотя ранние последователи являются более влиятельными, чем в среднем агенты всей сети, они не являются лидерами мнений (не всегда достаточно влиятельны для генерации глобальных каскадов).
5. В реальных сетях существует определенная локальная структура. Как правило, «знакомые» больше влияют друг на друга, и агенты имеют множественные, обычно перекрывающиеся группы знакомств. Таким образом, популяция из N агентов делится на M групп размером g . В среднем каждая группа случайно связана с m_{avg} группами. Каждый агент i -ой группы с вероятностью p связан с каждым агентом в своей группе и с вероятностью q связан с каждым агентом из m_i соседних групп. Оказалось, что у таких сетей «окно каскадов» шире, чем у рассматриваемых ранее случайных сетей. Однако введение групповой структуры снижает значимость лидеров мнений и ранних последователей, за исключением сети с низкой плотностью.
6. Интересно знать, что произойдет, если на лидеров можно повлиять так же легко, как и на остальных агентов. В работе [Watts, 2007] рассматривается каноническая модель SIR, в которой в одном взаимодействии независимо от других взаимодействий агент активизируется («инфицируется») с вероятностью b и дезактивируется («выздоровливает») со скоростью g , т.е. наиболее влиятельные агенты более легко поддаются влиянию. Как оказалось, в этом случае нет верхней границы «окна кас-

кадов»: чем выше плотность сети (возрастает уязвимость всех агентов), тем большего размера возникают каскады.

3. МОДЕЛИ ВЛИЯНИЯ И УПРАВЛЕНИЯ В СОЦИАЛЬНЫХ СЕТЯХ

3.1. ОЦЕНКА СИЛЫ ВЛИЯНИЯ АГЕНТА

3.1.1. Определение самых влиятельных агентов в сети

Если рассматривать социальную сеть как множество агентов – потенциальных потребителей некоторого товара или услуги или множество потенциальных последователей новой технологии (инновации, нововведения), то, с точки зрения продавца последних, ценность (полезность) агента в социальной сети зависит не только от него самого (например, непосредственной ожидаемой прибыли от продажи товара или технологии именно ему), но и от его влияния на других агентов (т. е. важна конфигурация и состояние сети – совокупность мнений потенциальных потребителей относительно «товара»). Поэтому часто возникает потребность в выявлении небольшого числа агентов (проблема максимизации влияния), которым, например, предоставляются льготы, способствующие распространению нововведения по всей сети.

К проблеме определения k самых влиятельных агентов в социальной сети приводит задача организации вирусного маркетинга. Например, авторы [Domingos, 2002] моделируют рынок как социальную сеть агентов, ценность каждого из которых определяется не только непосредственной ожидаемой прибылью от продажи, но и ожидаемой прибылью от продаж другим агентам, на которых повлияет данный, от продаж агентам, на которых они могут повлиять, и т.д. Таким образом, возникает понятие сетевой ценности агента-потребителя.

В [Domingos, 2002] ставится задача определения оптимальных маркетинговых действий $MA = \{MA_1, \dots, MA_n\}$ (MA_i может быть как булевой переменной: 1 – наличие скидки, 0 – ее отсутствие для i -го агента; так и непрерывной – размер скидки) для множества n агентов с предикатом $X_i = 1$, если агент i купил товар, и $X_i = 0$ иначе. Товар описывается следующим множеством атрибутов $Y = \{Y_1, \dots, Y_m\}$. У каждого агента i существует множество соседей N_i , которые прямо влияют на X_i , определяя тем самым сеть агентов. В свою очередь, i -ый агент влияет на своих соседей. Задается также стоимость c маркетинга в расчете на одного агента, выручка rv_1 от продажи товара агенту, если для него была проведена маркетинговая акция, и выручка rv_0 от продажи продукта агенту, если маркетинговая акция не была проведена. В этих условиях можно рассчитать ожидаемое повышение прибыли от маркетинговых акций для выбранных агентов ELP.

Тогда для нахождения k самых влиятельных узлов в социальной сети следует найти такое MA , которое максимизирует ELP. Эта задача является

NP-полной, т.е. требует полного перебора, и для ее решения предложен ряд эвристик.

Другой подход к выявлению наиболее влиятельных агентов предлагают авторы [Kempe, 2003]. Здесь использованы две базовых модели распространения нововведений:

- линейная пороговая модель (см. § 2.1);
- модель независимых каскадов. Имеется начальное множество активных агентов A_0 . В некоторый момент времени новый активный агент получает шанс активировать своих соседей с вероятностью p_{vw} , последние в случае успеха активируются на следующем шаге, и так до тех пор, пока возможны новые активации.

Тогда влияние $\sigma(A)$ множества агентов A можно определить как ожидаемое число активных агентов при завершении процесса распространения нововведений, инициированных агентами из множества A . Эта задача также является NP-полной, т.е. требует полного перебора, и для ее решения предложен ряд эвристик.

3.1.2. Обнаружение каскадов распространения влияния

Часто (например, в области информационной безопасности) необходимо как можно раньше обнаружить каскады распространения в социальной сети. Для этого отслеживаются состояния небольшой части узлов социальной сети. Проблема состоит в том, как определить это множество узлов (так называемых сенсоров) A . Выигрыш зависит от минимального времени обнаружения, числа обнаруженных каскадов, числа «зараженных» узлов, а затраты зависят от свойств выбранных узлов.

В работе [Leskovec, 2003] социальная сеть моделируется графом $G(N, E)$, задан бюджет B для сенсоров и доступны данные о распространении каскада по сети (для каждого каскада, инициированного в узле i , известно время $T(i, u)$, за которое он дойдет до узла u). Выбирается подмножество агентов A для максимизации ожидаемого выигрыша:

$$\max R(A) = \sum_i P(i) R_i(T(i, A)),$$

где $T(i, A)$ – минимальное время обнаружения одним из сенсоров из A каскада i ; P – вероятностное распределение каскадов (по «типам» – узлам возникновения); $R_i(T(i, A))$ – выигрыш от обнаружения каскада i в момент времени $T(i, A)$; затраты $c(A) < B$. Задача решается методами линейного программирования.

3.1.3. Индексы влияния

Количественной характеристикой влиятельности агента является индекс «силы» или индекс влиятельности агента.

В работе [71] изучается следующая модель влияния в социальной сети. Перед агентами группы стоит необходимость принять или отвергнуть некоторое предложение (например, проголосовать за некоторого кандидата на выборах). Предполагается, что агент изначально имеет предрасположенность к некоторому решению (да – «+1» или нет – «-1»), однако из-за влияния других агентов может принять иное итоговое решение.

Иногда для оценки воздействия (decisional power) k -го агента на решение всей группы используется индекс Хеде-Баккера (Hoede-Bakker)

$$HB_k(B, Gd) = \frac{1}{2^{n-1}} \sum_{i_{i_k=+1}} gd(Bi),$$

где N – множество из n агентов; $i \in I$ – вектор предрасположенностей / намерений этих агентов; а $B: I \rightarrow I$ – функция влияния, которая отображает вектор намерений в итоговый вектор решений; $gd: B(I) \rightarrow \{+1, -1\}$ – функция группового решения. Содержательно индекс Хеде-Баккера означает меру соответствия решений группы намерениям агента.

Функция влияния может быть основана на принципе большинства (когда агент принимает решение на основе намерения, поддерживаемого большинством агентов), либо может быть основана на ключевой роли экспертов (агенты поступают согласно известному намерению признанного «гуру»/эксперта) и т.п.

В теории принятия решений на сегодняшний день исследовано множество моделей определения индексов влияния, в том числе учитывающих и другие механизмы влияния. Например, это индексы влияния агентов, определяемые количеством и топологией сильных (непосредственных) и слабых (косвенных) связей, в которых участвует каждый агент. Многим таким индексам соответствуют марковские модели влияния. Более того, как оказалось, задача определения относительного влияния агентов в социальной сети чрезвычайно близка к так называемой задаче ранжирования интернет-страниц (PageRank problem).

Дальнейшее развитие приложений перечисленных моделей влияния к социальным сетям представляется перспективным направлением исследований.

3.2. ОБЩЕЕ ЗНАНИЕ И ЕГО УЧЕТ В СОЦИАЛЬНЫХ СЕТЯХ. РЕ-ФЛЕКСИИ

Основная концепция. Рассмотрим агента, входящего в некоторую социальную сеть. Агент информирован о текущей ситуационной обстановке (действиях и представлениях других агентов, параметрах среды – так называемом состоянии природы и т.п.). Предрасположенность к тем или иным представлениям и ситуационная обстановка (например, действия других агентов) приводят к формированию новых или модификации ста-

рых представлений. В соответствии с этими представлениями и установленной целью агент принимает решение и выполняет действие. Результаты действий приводят к изменению как самой ситуационной обстановки, так и внутренних ценностей, установок и представлений.

Представления n -го порядка (рефлексии). Взаимные представления. В ситуации принятия решения агент перед действием пытается предсказать поведение других агентов. Поэтому важными оказываются также представления агента о представлениях других агентов и т.д.

Пример представлений второго порядка (второго ранга рефлексии): агенту А известно, что агенту В известно, что С известно p .

Другие агенты соответственно могут иметь свои представления разных порядков.

В [Губанов, 2010] выделены следующие подходы к определению понятия взаимного представления: итеративный подход, рефлексивный подход, «мы-представление».

Согласно **итеративному подходу**, в группе G существует взаимное представление о p тогда и только тогда, когда факт p является общим знанием. Это значит: всем агентам из G известно p ; всем известно, что всем известно p , и так далее до бесконечности. Данный подход гипотетически предполагает наличие у агентов сильных когнитивных возможностей и рациональности поведения, что далеко не всегда выполняется. Отметим, что проблема определения максимальных целесообразных рангов рефлексии в зарубежной литературе получила название «level problem». Часто, но далеко не всегда, для успешных действий достаточно представлений второго порядка.

Рассмотрим пример – выполнение совместного действия двумя агентами. Высказывания базового порядка:

- 1) Агент А выполнит свою часть X , $p(A)$;
- 2) Агент В выполнит свою часть X , $p(B)$.

Предположим:

- i) А верит, что 1) и 2);
- ii) В верит, что 1) и 2).

Логически доказывается, что уровень рефлексии $n = 2$ в данном примере необходим и достаточен для выполнения совместного действия. При этом необходимо наличие у агентов представлений следующего порядка:

- iii) А предполагает, что i) и ii);
- iv) В предполагает, что i) и ii).

Согласно **рефлексивному подходу**, в группе G существует взаимное представление о p тогда и только тогда, когда все в группе предполагают p и все предполагают, что в группе G существует взаимное представление о p . Данное представление соответствует второму уровню рефлексии.

«Мы-представление» агента о p таково: (а) каждый агент имеет данное представление (требуется, поскольку не может быть общей позиции

без всех участников) и предполагает, что (b) все в группе имеют такое представление (предоставляет социальную причину для того, чтобы принять и иметь такую позицию), а также каждый агент предполагает, что (c) все предполагают, что есть взаимное представление в смысле (b) (усиливает причину, делая ее межсубъектной).

Общее «мы-представление» предполагает наличие у каждого агента группы «мы-представлений», т. е. соответствует третьему уровню рефлексии.

С помощью взаимных представлений можно пытаться объяснить коллективное мышление и коллективное действие.

3.3. КОЛЛЕКТИВНЫЕ ДЕЙСТВИЯ И ОБЩЕСТВЕННЫЕ БЛАГА

3.3.1. Примеры общественных благ

В коллективном действии важны следующие факторы: информированность, коммуникация и координация. Теория коллективного действия объясняет широкий круг явлений (общественные движения, электоральное поведение, членство в группах по интересам), связанных с достижением общественных благ посредством согласованного совместного участия двух или более людей. Теория также рассматривает влияние внешних факторов на поведение людей в данной группе.

Как известно, общественное благо – это благо, характеризующееся:

- 1) неисключаемостью из потребления, т. е. невозможно исключить из числа потребителей общественного блага тех, кто не платил за него;
- 2) отсутствием конкуренции при потреблении блага: потребление блага одним субъектом не ведет к сокращению потребления этого блага другими людьми.

Примеры чистого общественного блага – национальная оборона, мосты, общественное мнение, выборы, открытая информационная база данных, система коммуникации и т.д.

Поэтому каждый участник может не нести затраты ради достижения общественного блага (не участвовать в коллективном действии) и при этом получить выгоды без участия в издержках. Т.е. возникает затрудненность осуществления взаимовыгодных коллективных действий.

В этом аспекте известно много феноменов:

- «проблема безбилетника», широко известная в современной микроэкономической теории: лично безбилетник имеет выгоду, при этом, если число людей велико, то увеличиваются общественные затраты на выявление «безбилетников» и наложение санкций.
- трагедия общин. Существует сельская община, у которой есть только одно доступное пастбище. На нем все члены общины могут пасти

скот сколько угодно. Выпас скота уменьшает количество травы, растущей на нем и, соответственно, выгоды от скотоводства.

Каждый член общины может увеличить число своего скота, увеличить свой собственный доход, при этом плодородие пастбища сократится незначительно. Однако если все члены общины сделают то же самое, пастбище станет уже намного хуже. Если же член общины уменьшит свой выпас, плодородие поля увеличится, но его личный выигрыш от этого будет намного меньше, чем потерянный доход.

Получается, что каждому члену общины выгодно только увеличивать использование пастбища.

- Прокси-сервер. Похожая ситуация наблюдается с доступом в Интернет через плохо настроенный прокси-сервер. Когда пропускная способность канала делится поровну между соединениями, пользователь может увеличить отведенную ему полосу пропускания, увеличивая количество соединений (например, пользуясь большим количеством одновременно открытых окон браузера или менеджером зачек). Если это делает одновременно большое количество пользователей, пропускная способность каждого отдельного соединения уменьшается настолько, что выход в Интернет становится затруднительным. Для решения этой проблемы в прокси-серверах есть интеллектуальные средства разделения полосы пропускания между пользователями.

Решение во всех случаях заключается в создании некоторого общественного механизма (в лице государства, старейшины общины или администратора прокси-сервера), который регламентирует использование общего ресурса и не дает никому излишне его эксплуатировать. Однако возникает новая проблема – как сделать этот механизм наиболее эффективным? Решения во многом основаны на теории игр и понятии равновесия в играх.

3.3.2. Управление через выбор механизма игры

Чего мы хотим как управленцы? – построить такой механизм игры, который бы реализовал:

- эффективный социальный выбор;
- удачное (как правило, максимально сильное) равновесие для каждого агента;
- социальную справедливость;
- «правдивый» выбор, при котором агентам было бы выгодно сообщать свои настоящие стратегии. Напомним, что примером такого механизма является аукцион Викри (см. § 1.2.4).

Примирить все эти критерии – задача построения управления. Единственного решения (а часто и решения вообще) эта задача не имеет, но ис-

следования позволили выявить пути, чтобы побудить агентов к приложению усилий по созданию общественного блага:

- стимулировать их материальными поощрениями,
- оказывать влияние в рамках тех или иных схем социального влияния.

Ситуации с коллективным действием, в которых все стороны могут получить взаимную выгоду, если примут взаимно согласованные решения (проблема координации), часто моделируются координационными играми. Координационные игры – класс игр с множественными чистыми равновесиями Нэша, в которых игроки выбирают одинаковые или согласованные стратегии.

Какова здесь специфика ИТ? Развитие информационно-телекоммуникационных технологий в коллективных действиях во много раз снижает затраты на коммуникацию и координацию, а также иногда освобождает от необходимости построения формальной структуры. При этом потенциально должны существовать латентные группы, т.е. сообщества с общими групповыми интересами в коллективном благе, которые еще не построили организационную структуру для решения коммуникативных и организационных задач, но со структурами лидерства, центрами, где аккумулируются ресурсы и принимаются решения.

Если, например, рассматривать общественно полезную информацию как общественное благо, то при целенаправленном создании информационной базы и построении сообщества на начальном этапе все же требуется координация участников, и возникает «проблема безбилетника»; при нецеленаправленном создании такой базы, когда участники могут и не знать других участников, самостоятельно размещая информацию на общественно доступных ресурсах (форумах, интернет-страницах), возникает по большей части не проблема участия, а проблема доверия.

3.3.3. Коллективные действия агентов в социальных сетях

Структура социальной сети оказывает сильное воздействие на решения агентов о принятии участия в коллективном действии. Ключевое значение здесь имеют социальные связи:

- социальные связи могут обеспечить эффективный локальный социальный контроль для стимулирования участия в коллективном действии (в силу давления со стороны своих соседей, доверия к ним, социального одобрения, необходимости сохранения положительных отношений и соответствия ожиданиям, эмоциональной привязанности, сохранения своей репутации, отождествления себя с соседями и т.п.).
- социальные связи обеспечивают агента информацией о намерениях и действиях других агентов в сети и формируют его (неполные) представления, на основе которых агент принимает свои решения.

- в пределах социальных связей агенты могут прикладывать совместные усилия по созданию локального общественного блага и пользоваться им.

Агенты, соединенные связями, могут прикладывать совместные усилия по созданию общественных благ и пользованию ими. Это может привести к специализации в сети в обеспечении общественных благ. Доказано, что существует равновесие, в котором агенты вносят определенный вклад (усилия), а другие пользуются этим. Такая специализация может принести пользу обществу в целом, если вкладчики («специалисты») связаны со многими агентами в сети. Новые связи в сети увеличивают доступность общественного блага, но уменьшают индивидуальные стимулы для приложения усилий (увеличения вклада). Следовательно, общее благосостояние всей сети выше в неполных сетях. В этом смысле для будущих исследований перспективно изучение динамики процессов формирования социальной сети. Для этого используются игровые модели, аналогичные рассмотренным в § 1.2.4 (может реализоваться равновесие по Нэшу, по Парето).

В статье [Janku, 2002] рассматривается взаимосвязь между механизмами социального контроля, свойствами социальных сетей и коллективным действием (collective action), предпринимаемым для обеспечения общественного блага всего сообщества агентов. Показано, что ключевыми факторами, влияющими на решения агентов сети в рамках конфликта частных и общественных интересов, являются осуществляемые посредством межличностных связей в социальной сети различные виды социального контроля:

- поведенческое подтверждение (behavioral confirmation) – следование агента социальным ожиданиям,
- социальные стимулы (social selective incentives) – дополнительные персональные блага, предоставляемые данному агенту другими агентами.

Принятие коллективного решения моделируется некооперативной игрой с равновесием по Нэшу. В игре используются следующие параметры. На агента по связям оказывают влияние:

- поведенческие стимулы от соседей, принявших аналогичное решение (аддитивный стимул b_1 и пропорциональный стимул b_2)
- социальные стимулы (s от каждого соседа).

При этом каждый агент принимает решение об участии ($s_i = 1$) или неучастии ($s_i = 0$) в коллективном действии. Участие требует общих затрат c и приносит дополнительное благо a для каждого агента сети, но есть и «безбилетники».

Можно показать, что нет ни одного агента, для которого неучастие при участии всех соседей было бы выгодно. Содержательно это означает, что усиление стимулов, так же как и высокая минимальная степень вершин

сети, приводит к повышению вероятности успешности коллективного действия.

В том случае, когда возможно несколько равновесий Нэша, выбирается равновесие, обеспечивающее больший выигрыш для каждого агента, нежели какое-либо другое. Если возможны как равновесие Нэша полного участия агентов, так и равновесие Нэша полного неучастия, и число агентов превосходит некий порог, то равновесие Нэша с полным участием агентов Парето-доминирует. Другими словами, чем меньше «безбилетников» в равновесии Нэша с частичным участием, тем реже равновесие полного участия будет доминировать, так как шансы «безбилетников» получить больший выигрыш увеличиваются.

3.3.4. Информационное управление в социальных сетях

Как уже говорилось в § 1.6.3, при рассмотрении моделей, учитывающих информированность агентов, традиционно выделяют три вложенных класса задач: моделирование информационного влияния, информационного управления и информационного противоборства (рис. 13).



Рис. 13. Соотношение уровней информационного управления

Модель информационного влияния дает возможность исследовать зависимость поведения субъекта от его информированности и, следовательно, от информационных воздействий. Имея модель информационного влияния, можно ставить и решать задачу информационного управления – какими должны быть информационные воздействия (с точки зрения управляющего субъекта), чтобы добиться требуемого поведения от управляемого субъекта. И, наконец, умея решать задачу информационного управления, можно моделировать информационное противоборство – взаимодействие нескольких субъектов, обладающих несовпадающими интересами и осуществляющих информационные воздействия на один и тот же управляемый субъект.

В работе [Губанов, 2009] рассматривается информационное влияние агентов на формирование мнений друг друга в социальных сетях. Структура сети описывается с помощью введенных понятий: сообщество (множество агентов, которые не подвергаются влиянию агентов вне него), группа (сообщество агентов, в котором каждый агент влияет или подвергается влиянию каждого другого агента группы прямо или косвенно) и спутник (агент, не оказывающий влияния ни на одну из групп). В работе показано, что, если в каждой группе хотя бы один агент хоть сколько-нибудь доверяет своему мнению, то в конечном итоге мнения спутников определяются мнением групп, а внутри групп мнения агентов сходятся и равны.

В такой социальной сети представляется вполне естественным рассмотрение задачи информационного управления: как изменить мнение небольшого множества ключевых агентов в сети таким образом, что в результате распространения изменения мнений сформируются требуемые мнения участников сети. Из этой задачи вытекает, в соответствии с рис. 13, и теоретико-игровая задача информационного противоборства нескольких игроков в сети.

Модель информационного влияния в социальной сети. Агенты, входящие в социальную сеть, описываются множеством N . Агенты в сети влияют друг на друга, и степень влияния задается матрицей t прямого влияния размерности $n \times n$, где $t_{ij} \geq 0$ обозначает степень доверия i -го агента j -му агенту (влияния j -го на i -го). Такая модель широко применяется не только для моделирования социальных групп, но и в теории многоагентных систем.

Если i -й агент доверяет j -му, а j -й доверяет k -му, то это означает, что k -й агент косвенно влияет на i -го. Продолжая эту цепь, можно получить ответ на вопрос, кто в итоге формирует мнение в социальной сети.

У каждого агента в начальный момент времени имеется мнение по определенному вопросу. Мнение всех агентов сети отражает вектор-столбец мнений \mathbf{b} размерностью n .

Агенты в социальной сети взаимодействуют, обмениваясь мнениями. Этот обмен приводит к тому, что мнение каждого агента меняется в соответствии с мнениями агентов, которым данный агент доверяет. Первое измененное мнение агентов равно произведению матрицы непосредственного доверия на вектор начальных мнений: $\mathbf{b}^{(1)} = t\mathbf{b}$. В конечном итоге их мнения сходятся к результирующему (итоговому) мнению $\mathbf{V} = \lim \mathbf{b}^{(n)}$, т.е. итоговое мнение

$$\mathbf{V} = T\mathbf{b},$$

где $T = \lim t^{(n)}$. Таким образом, выявлены следующие свойства информационного влияния в социальной сети:

- в каждой из групп сети итоговые мнения агентов совпадают (этот результат соответствует выводам социальных психологов);

- итоговые мнения спутников полностью определяются мнением одной или нескольких групп.

Задача информационного управления. Имея «основное уравнение» $\mathbf{B} = T\mathbf{b}$, связывающее начальные и итоговые мнения агентов, можно ставить и решать задачу управления – воздействия на агентов социальной сети с целью формирования требуемого влияния, а управляющее (информационное) воздействие заключается в изменении центром начальных мнений агентов путем «добавления» вектора управлений u .

Содержательно управление заключается в изменении начального мнения i -го агента с b_i на $b_i + u_i$, $u_i \in U_i$, $i \in N$. Тогда итоговые мнения \mathbf{B}_u будут определяться следующим уравнением:

$$\mathbf{B}_u = T(\mathbf{b} + u),$$

т.е. результирующее мнение агента является суммой его «невозмущенного» результирующего мнения \mathbf{b} и изменений, вызванных управляющими воздействиями u .

Пусть целевая функция центра (критерий эффективности управления) $F(\mathbf{B}_u, u) = H(\mathbf{B}_u) - c(u)$, где $H(\times)$ – «доход» центра, зависящий от итоговых мнений агентов; $c(\times)$ – затраты на осуществление управляющих воздействий. Заметим, что от предпочтений центра, зависящих от действий агентов, можно перейти к его предпочтениям, зависящим от информированности или мнений агентов, так как с учетом рефлексии агентов (см. § 3.2) можно считать, что действия агентов определяются их информированностью. Тогда задача управления будет заключаться в выборе допустимого вектора управлений, максимизирующего критерий эффективности:

$$\Phi(\mathbf{B}_u, u) \rightarrow \max_{u \in U}.$$

Поскольку, как уже было показано выше, воздействовать на мнения спутников не имеет смысла, можно на основе только матрицы доверия сказать, на каких агентов должно быть нацелено информационное воздействие. Для этого можно проанализировать «индексы влияния».

Другие задачи, решаемые в рамках информационного управления:

- выработка оптимальной последовательности информационных воздействий,
- определение уровня управляемости конкретной социальной группы,
- обратная задача определения множества управляющих воздействий

Теоретико-игровая модель информационного противоборства. В соответствии с рис. 13, имея решение задачи информационного управления, можно переходить к решению задач информационного противоборства. Эта задача ставится и решается в рамках игровых моделей.

Пусть существует множество игроков («игроки» могут совпадать с агентами или их группами), имеющих возможность влиять на начальные мнения агентов и заинтересованных в формировании определенных итоговых мнений агентов. Опишем возникающую между игроками игру. Будем

считать, что воздействия игроков на мнение каждого из агентов аддитивны. Считая, что игроки выбирают свои действия однократно, одновременно и независимо, получим игру в нормальной форме, определяемую заданием соответственно множества игроков, их множеств допустимых действий и целевых функций. Имея игру в нормальной форме, можно исследовать ее равновесия, определять на ней кооперативные, повторяющиеся и другие виды игр. Ряд примеров анализа таких игр приведен в [Губанов, 2009].

В целом можно сделать вывод, что модели влияния в социальных сетях пока еще только становятся самостоятельной дисциплиной, представляя собой синтетический «сплав» теории графов, теории игр, социальной психологии, социологической теории малых групп, теории марковских цепей, теории синтеза механизмов (*mechanism design*), теории многоагентных систем и других научных направлений.

4. ГРАФОВЫЕ МОДЕЛИ СОЦИАЛЬНЫХ СЕТЕЙ

4.1. ОСНОВНАЯ ТЕРМИНОЛОГИЯ

Любую социальную сеть можно математически представить в виде графа $G=(V, E)$, где V – множество вершин графа, E – множество ребер графа. В графе социальной сети вершинами являются участники, а ребра означают наличие отношений между ними. Отношения могут быть как направленными, так и ненаправленными. Как правило, рассматривают два основных типа отношений: дружба (люди знакомы друг с другом) и интересы (есть общие интересы, люди входят в одну группу по интересам).

Типы графовых моделей. Выделяют три типа графовых моделей – обычные, стохастические блоковые и вероятностные.

- Обычные графовые модели задаются матрицей связности A размера $N \times N$.
- Стохастические блоковые модели задаются матрицей A размера $N \times N$, где N – число групп (блоков) участников. Элемент $a_{ij} \in [0, 1]$ показывает плотность связей между участниками сети, принадлежащими к группе v_i , и участниками, принадлежащими к группе v_j . При этом граф не содержит дополнительных ребер и вершин, соответствующих связям участников внутри одной группы.
- Вероятностные графовые модели задаются матрицей A размера $N \times N$, где N – число участников сети. Элемент $a_{ij} \in [0, 1]$ показывает вероятность взаимодействия участника v_i и участника v_j в течение определенного периода времени.

Локальные свойства графовых моделей. В прикладных исследованиях обычно применяют такие характеристики, как размер сети, сетевая плотность, степень и плотность центральности и эквивалентность.

Коэффициент плотности определяется как отношение числа ребер в анализируемом графе к числу ребер в полном графе с тем же числом вершин (полный граф – это граф, в котором все вершины соединены между собой).

Число путей заданной длины (путь – последовательность вершин, связанных между собой) – минимальное число ребер, удаление которых разбивает граф на несколько частей.

Центральность – мера близости к центру графа – используется для определения относительной важности (веса) вершин графа (т.е. насколько влиятельным в рамках конкретной сети является ее участник). Центральность можно определить разными способами, поэтому существуют различные меры центральности:

- центральность по степени определяется как количество связей, инцидентных вершине; при этом входящие связи характеризуют популярность человека, выходящие – его общительность;

- центральность по близости является показателем того, насколько быстро распространяется информация в сети от одного участника к остальным, т.е. насколько близок рассматриваемый участник ко всем остальным участникам сети. В качестве меры расстояния между двумя участниками используется кратчайший путь по графу $d_G(v, t)$. Так, непосредственные друзья участника находятся на расстоянии 1, друзья друзей – на расстоянии 2, друзья друзей друзей – на расстоянии 3 и т.д. Далее берется сумма всех расстояний и нормируется. Полученная величина называется удаленностью вершины v от других вершин. Близость определяется как величина, обратная удаленности:

$$C_c(v) = \frac{N-1}{\sum_{t \in V \setminus v} d_G(v, t)}.$$

- центральность по посредничеству – характеристика участника, показывающая его важность при распространении информации. Она рассчитывается как число кратчайших путей между всеми парами участников, проходящих через рассматриваемого участника. Недостатком центральности по посредничеству является ее вычислительная сложность.
- центральность по собственному вектору демонстрирует зависимость между центральностью участника и центральностями его друзей. Пусть центральность рассматриваемого участника – x_v , а центральность его непосредственных друзей (соседних вершин) – x_j, x_k, x_l и т.д. Центральность по собственному вектору определяется как $x_v = (x_j + x_k + x_l + \dots) / \lambda$, где λ – собственный вектор матрицы смежности A . Собственный вектор, соответствующий самому большому собственному значению, как раз образован центральностями соответствующих участников сети. Таким образом, чем больше у участника друзей и чем они центральнее, тем больше его центральность. Верно и обратное: чем больше центральность участника, тем больше центральность его друзей.
- центральность можно вычислить при помощи алгоритма ссылочного ранжирования (PageRank), который используется в поисковой системе Google. В основу положен принцип «важности» веб-страницы: чем больше ссылок на страницу, тем она «важнее». Кроме того, вес самой страницы определяется весом ссылки передаваемой на нее страницы. Таким образом, PageRank – это метод вычисления веса страницы путем подсчета важности ссылок на нее, т. е. вершина, ссылающаяся на другую вершину с большим весом, сама получает большой вес.

Уровень доверия. Алгоритм вычисления уровня доверия (TrustRank) изначально был создан для отделения информативных веб-страниц от спама. Для контрольной выборки эксперты вручную оценивают степень доверия небольшого количества сайтов, которые можно считать надежны-

ми. Эти сайты принимаются за эталон. Далее в основу алгоритма положено утверждение, что хорошие сайты редко ссылаются на плохие, а вот плохие очень часто ссылаются на хорошие. TrustRank – величина, которая дает оценку того, можно ли доверять конкретному сайту, считая, что он не содержит спама. Чем больше ссылок на сайте, тем меньше доверия передается по каждой такой ссылке. Степень доверия сайту (TrustRank) убывает с увеличением расстояния между ним и первоначальной выборкой.

Сила структурной позиции участника является основным показателем, определяющим различия в ресурсах участников сети. В теории сетевого обмена для измерения данной характеристики вводится индекс *GPI* (Genuine Progress Indicator) силы участника

$$GPI_i = \sum_{k=1}^{g-1} (-1)^{k-1} P[i]_k,$$

где $P[i]_k$ – число непересекающихся путей длины k , проходящих через вершину v_i . Сила участника v_i по сравнению с силой участника v_j определяется как

$$GPI_{ij} = GPI_i - GPI_j.$$

В заключение параграфа отметим, что, согласно [Давыденко, 2005], применение большинства упомянутых характеристик (размер сети, сетевая плотность, степень и плотность центральности, эквивалентность) ограничивается двумя принципиальными проблемами – сбора и систематизации первичной информации. Например, расчеты размера сети и сетевого ранга теряют смысл в больших сетях, а сетевую плотность и центральность, оказывается, затруднительно вычислить в больших сетях. В то же время эти характеристики достаточно информативны, например, центральность позволяет уловить такую важную для социальных сетей характеристику, как посредничество.

4.2. СООБЩЕСТВА И СВЯЗАННЫЕ ПОДГРУППЫ

4.2.1. Эффекты степени вершины в сети

Степень вершины в сети – это число ребер, соединенных с заданной вершиной. Определим p_k как часть вершин в сети, которые имеют степень k . В случайном большом графе каждое ребро присутствует или отсутствует с равной вероятностью, и распределения степени – биномиальное или пуассоновское. Далекие от распределения Пуассона распределения степени вершин в большинстве сетей искажены со скосом вправо – распределения имеют длинную правую хвостовую часть значений. Однако в большинстве реальных сетей распределения степени вершин имеют длинную правую хвостовую часть, которая подчиняется степенному закону:

$$p_k \sim k^{-\alpha} \text{ при } \alpha = \text{const.}$$

Распределение степени при удалении вершин определяет такое свойство сети, как эластичность сети. Большинство сетей основано на их связности, т.е. существовании путей между парами вершины. Если удалять некоторые вершины, то весь граф в конечном итоге разрушится. Если вершина удалена из сети, типичная длина этих путей увеличивается, и в конечном счете пары вершины станут разъединенными. Имеется ряд способов удаления вершин, и различные сети показывают вариацию степени эластичности сети к этим способам. Можно случайно удалять вершину из сети или иметь цель удаления определенного класса вершин (например, с самыми высокими степенями).

Рассмотрим пример. Для анализа атак на Internet-серверы были исследованы эффекты удаления вершин сети WWW объемом 326000 страниц со степенным распределением степени. Среднее расстояние вершина–вершина почти не изменилось при случайном удалении вершин; в то же время целенаправленное удаление вершин с наиболее высокими степенями приводит к разрушению сети. Таким образом, Internet является высоко эластичной сетью относительно случайного отказа вершины в сети, но высокочувствителен к преднамеренной атаке на вершины с наиболее высокими степенями.

4.2.2. Структура сообщества

Социальные сети имеют «структуру сообщества», т.е. группы вершин, имеющие высокую плотность ребер между ними, с более низкой плотностью ребер между группами. Оказалось, что для социальных сетей структура сообществ является характеристикой конкретной сети.

Простейший случай связанной группы – это сообщество, где каждый участник связан с каждым, и в данную группу не могут быть включены другие участники сети, поскольку они не имеют связей со всеми членами сообщества. Однако это не обязательно так. Личности не обязательно знают тех, с кем они разделяют группу, но имеются вероятность p знакомства, которая равна нулю для тех, кто не разделяет группу.

Если анализировать процессы распространения информации в графах, можно дать другое определение сообщества: множество участников, путь между двумя любыми участниками которого не содержит более одной промежуточной вершины. В результате информация от одного участника к другому в связанной группе передается с минимальными искажениями.

Как выявить структуру сообщества в сети?

(1) Традиционный метод для извлечения структуры сообществ из сети – кластерный анализ. В этом методе приписывается «сила связи» парам вершин в сети, представляющей интерес. Каждой из $0,5n(n-1)$ возможных пар сети из n вершин назначена такая «сила» (не только соединенным ребрами). Затем, начиная с вершин без ребер между любыми из них, прибавляются ребра в порядке уменьшения силы связи. Когда все ребра добавле-

ны, все вершины соединены со всеми другими – получается требуемая кластеризация.

В социальных сетях кластеризация рассматривается в контексте моделей блоков, которые разделяют сети на блоки согласно какому-либо принципу. Две вершины сети структурно эквивалентны, если они имеют одинаковых соседей. Точная структурная эквивалентность – редкость, но приближительная эквивалентность может использоваться как основание для кластеризации.

Для социальных сетей применимы стандартные алгоритмы кластеризации – как статистические (например, алгоритм k -средних), так и иерархические. Преимуществом иерархических методов является возможность представления результата кластеризации в виде дендрограммы, т.е. на выходе будет получено не просто разбиение графа на группы, а иерархия групп и подгрупп в графе. Кластерный анализ также можно проводить не снизу вверх, а сверху вниз, т.е. сначала вся сеть рассматривается как одна группа, а на каждой итерации происходит последовательное отделение по одной связи.

Основная сложность методов кластеризации – подобрать подходящую меру расстояния (кратчайшего пути между вершинами) или меру сходства (similarity). Это могут быть различные взвешенные меры расстояния вершина–вершина, размеры максимального потока и взвешенных путевых индексов между вершинами. Наиболее часто применяются меры сходства, использующие косинусный коэффициент и коэффициент Жаккара.

Рассмотрим модель, предложенную в [Newman, 2003]. Математически эта модель может быть представлена как процесс просачивания (см. § 2.2) с вероятностью p на сети, сформированной проекцией на личности двудольного графа личностей и групп.

Пусть сеть разделена на группы, и каждый индивидум может принадлежать любому числу групп. Личности не обязательно знают тех, с кем они разделяют группу, но имеются вероятность p знакомства, которая равна нулю для тех, кто не разделяет группу. В дополнение к параметру p характеристиками модели с распределениями вероятностей являются: r_m – вероятность, что личность принадлежит m группам; s_n – вероятность, что группа содержит n личностей.

В простейшем случае каждая личность принадлежит точно одной группе, и размеры группы имеют распределение Пуассона. Тогда $r = p$.

Более реалистичский пример – число групп, к которым принадлежат личности также изменяется согласно распределению Пуассона. Тогда $r = p / (1 + \mu + n + n \cdot m \cdot p)$, где n , m – средние значения двух распределений.

Еще один вариант выявления структуры сети – вычисление индекса кластеризации. В реальной сети присутствие связей между вершинами А и В, и между В и С часто приводит к связи между А и С. В топологических терминах это означает, что в сети существует высокая плотность треуголь-

ников ABC, и кластеризация может быть определена измерением этой плотности:

$C^{(1)} = 3 \times (\text{число треугольников на графе}) / (\text{число связных троек вершине})$, где «связная тройка» – вершина, соединенная непосредственно с неупорядоченной парой других вершин.

Фактически C измеряет часть троек, которые имеют третье ребро для формирования треугольника. Множитель "3" используется для учета того, что каждый треугольник входит в три тройки, поэтому $0 \leq C \leq 1$.

Альтернативное определение индекса кластеризации – следующее:

$$C_i = \frac{\text{число треугольников, соединенных с вершиной } i}{\text{число троек, центрированных на вершине } i}.$$

Для вершин со степенью 0 или 1 $C_i = 0$.

Индекс кластеризации сети в целом рассчитывается усреднением $C^{(1)}$ по всей сети.

(2) Связанные группы также могут быть выделены с помощью многомерного шкалирования или факторного анализа матрицы связей графа.

(3) Игровой подход. Социальная сеть рассматривается как коммуникационная, посредством которой агенты сообщают друг другу о своей готовности принять участие в коллективном действии. Каждый агент информирован о готовности только своих ближайших соседей и на основе этого локального знания принимает решение об участии, используя правило принятия решений «я приму участие, если примешь участие ты» (механизм координации). Иначе говоря, рассматривается координационная игра с неполной информированностью.

4.2.3. Достаточные сети

Если все агенты имеют достаточно «оптимистичные» представления, то все агенты примут участие в коллективном действии независимо от структуры сети. Однако представляет интерес то, какими свойствами должны обладать коммуникационные сети, чтобы вся группа агентов приняла участие в коллективном действии независимо от их априорных представлений. Такие сети были названы достаточными.

Достаточная сеть – сеть, для которой существует такое равновесие, в котором все агенты выбирают участие в коллективном действии (напомним, что такое возможно, только если все они готовы/желают участвовать) независимо от их априорных представлений.

Существуют достаточные сети, называемые минимальными, в которых отсутствуют избыточные коммуникационные связи. Такая минимально достаточная сеть принимает форму иерархии клик.

Кликкой в неориентированном графе $G = (V, E)$ называется подмножество вершин $C \subseteq V$ такое, что для любых двух вершин в C существует

ребро, их соединяющее. Клика – подмножество агентов, в которой каждый агент непосредственно информирует каждого другого агента.

Любая минимально достаточная сеть разбивает агентов на клики, и между кликами существует коммуникационная связь только в одном направлении. Минимально достаточная сеть распределяет агентов в социально-сетевой иерархии ролей, определенной кликами и их отношением: «ведущие» и «последователи».

В целом минимальная структура коммуникаций между агентами, обеспечивающая требуемые свойства социальной сети, – это остоное дерево в орграфе коммуникаций. Заметим, что эта структура является общей для множества моделей влияния в социальных сетях, а именно – для моделей сетевой автокорреляции, подражательного поведения, координации, марковских и др. моделей.

Структура сети должна быть не просто известна агентам, но быть общим знанием для агентов. Важную роль играют клики, в которых «локально» общее знание (ограниченная частью агентов структура) возникает естественно. Информация о готовности «течет» от «ведущих» клик по цепочке. Агенты знают готовность других агентов, но не их действия. Минимальные достаточные сети – неотъемлемые для игры структуры, интерпретируемые как иерархические социальные роли.

Интуитивно понятно, что в сети с большим количеством сильных связей сразу формируются небольшие клики из-за «транзитивности»: друзья моих друзей часто оказываются и моими друзьями. Поэтому в них формирование общего знания на локальном уровне происходит быстрее. И если значения порогов достаточно низки, то есть шансы, что группа, связанная сильной связью, станет ведущей кликой. Однако если пороги высоки, то локальное общее знание в маленьких кликах останется бесполезным, а большее значение приобретут слабые связи в силу того, что они пересекают общество быстро, ускоряют коммуникацию и усиливают распространение знания, создавая предпосылки для коллективного действия.

4.2.4. Small-world модель

Известен эксперимент С. Милграма, в которых письма проходили от лица к лицу и достигали назначенной цели за малое количество шагов – порядка шести. Этот эксперимент получил название «феномен малого мира» (Small-world), или «теория шести рукопожатий» (Six degrees of separation). Гипотеза заключается в том, что каждый человек знаком с любым другим жителем планеты через цепочку общих знакомых, в среднем состоящую из шести человек. Пока что это утверждение не было опровергнуто. Наоборот, в качестве доказательства правильности гипотезы выдвигается наблюдение, что диаметр большинства сетей относительно невелик.

Результат С. Милграма – одна из первых демонстраций small-world эффекта: большинство пар вершин в сети соединены коротким путем. Small-world эффект важен для оценки динамики процессов в сетях. Например, на сетях с эффектом small-world распространение информации будет более быстрым.

Small-world модели могут быть построены на решетках любой топологии, но наилучший анализируемый случай – одномерный. Если мы берем одномерную решетку с n вершинами с периодическими условиями (кольцо) и соединим каждую вершину с ее соседями, удаленными на k или менее ребер решетки, мы получаем систему с L_k ребрами. Small-world модель создается взятием малой части ребер в графе и «заменой путей» в нем. Процедура «замены путей» заключается в просмотре каждого ребра по очереди и перемещении одного конца этого ребра с вероятностью p к новому положению, выбранному случайно на решетке; при этом двойные ребра не создаются.

В пределе, при $p \rightarrow 0$, модель есть «large-world». При больших значениях p модель становится случайным графом. Переход от «large-world» режима к «small-world» в этой модели возможен или увеличением p , или увеличением размера системы n . Выявление эффекта «small-world» на многомерной сети направлено на выявление эффекта «слабых связей».

4.2.5. Параметры качества взаимодействия

Выявление логики взаимодействий участников сети в блоковых моделях позволяет выявлять сходство принципов взаимоотношений участников в различных социальных сетях. Для этого используются различные методы.

Ролевые алгебры. Здесь определяются матрицы симпатии и антипатии:

$$LIKE = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, INLIKE = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Теперь можно анализировать комбинации взаимодействий участников сети, перемножая соответствующие матрицы.

Анализ диад и триад. Диады – это набор из двух участников сети (вершин) и всех взаимодействий (ребер) между ними. Диада для каждого типа взаимодействий может находиться в одном из четырех состояний: нет связи между участниками, связь направлена от первого участника ко второму, связь направлена от второго участника к первому, взаимные связи участников. Анализ диад помогает установить вероятность наличия ребра между ними, степень зависимости от свойств участников, определить условия и направления передачи информации и т.д.

Для триад (три взаимодействующих участника) дополнительно исследуются вопросы транзитивности взаимодействий. Транзитивность – это

выполнение условий вида «если есть взаимодействие между v_1 и v_2 , а также между v_2 и v_3 , то имеет место взаимодействие между v_1 и v_3 ». Важной характеристикой, описывающей локальные связи участников и часто используемой при анализе диад и триад, является сбалансированность. Сбалансированность – это отсутствие ситуаций типа «положительное взаимодействие (дружба, партнерство) между v_1 и v_2 , а также между v_1 и v_3 , но негативное взаимодействие (вражда, соперничество) между v_2 и v_3 . Предполагается, что сбалансированные сети психологически более комфортны для участников и более устойчивы по сравнению с несбалансированными.

4.3. МОДЕЛИ ДИНАМИКИ СЕТИ

В этих моделях сети растут дополнением вершин и ребер; процессы роста приводят к изменениям структурных характеристик сети.

Рассмотрим некоторые примеры.

В моделях «**триадического замыкания**» ребра добавляются между парами вершин, которые имеют третью вершину как общего соседа, образуя треугольник, что приводит к увеличению транзитивности в сети.

Д. Прайс исследовал сеть цитирований между научными работами и нашел, что и входные, и выходные степени имеют степенные распределения. Ш. Саймон показал, что степенные законы возникают, когда «богатые богатеют», т.е. когда увеличивающееся количество растет пропорционально самому количеству. Прайс назвал это кумулятивным преимуществом. В случае кумулятивного процесса вероятность прикрепления одного из новых ребер к старой вершине пропорциональна $k+1$, где k – входная степень старой вершины (см. также § 1.2.4).

Модели эволюции графа. Согласно этой модели, когда новая вершина добавляется к сети, происходит выбор вершин, к которым можно осуществить присоединение, при помощи присоединяющего правила предпочтений. Также выбор вершины может осуществляться случайным образом или «копированием» некоторых ее внешних ссылок.

Расположение ребер играет важную роль в эволюции сетей. Например, если со временем плотность сети увеличивается, то количество ребер увеличивается линейно с ростом количества вершин. В итоге плотность сети меняется по степенному закону. В то же время описано еще одно наблюдение: диаметр сети часто уменьшается с течением времени, что противоречит общепринятому мнению о том, что меры расстояний должны медленно увеличиваться в зависимости от количества вершин.

Модель «закрытого треугольника» [Leskovec, 2005] утверждает, что новые вершины, добавленные к сети, имеют тенденцию к закрытию треугольника. Если считать, что связи, возникающие между участниками, образуют треугольник, то «открытый» треугольник возникает, когда два участника могут быть связаны друг с другом только посредством третьего,

т.е. одна из трех связей пропущена. Когда добавляется третья связь, получается «закрытый» треугольник.

Модель «лесных пожаров» является в некотором смысле обобщением модели закрытого треугольника. Новая вершина присоединяется к существующей путем выбора подграфа, содержащего эту вершину и связывающего ее со всеми вершинами этого подграфа. Процесс начинается в выбранной вершине и напоминает распространение пожара через все вершины сети.

Для анализа эволюции сетей используются различные алгоритмы, базирующиеся на динамическом программировании, полном переборе, максимальном соответствии и «жадных» эвристиках. Основное внимание уделяется определению приблизительных кластеров пользователей и их временным изменениям.

4.4. МЕТОДЫ НА ОСНОВЕ ОНТОЛОГИЙ

Для удобства регистрации пользователей в социальных сетях был разработан формат FOAF (Friend of a friend). Файл (foaf.rdf) имеет «двухуровневую рекурсию»: на первом уровне идет описание пользователя, т.е. указание имени, фото, блога, электронного адреса и т.д., а на втором уровне идет описание тех людей, которых знает этот пользователь (foaf:knows), в таком же формате, как было на первом уровне. Формат RDF (подобие микроформатов) позволяет «вживить» FOAF в текст HTML.

Смысл этого в том, что можно создать распределенную социальную сеть, которая будет распознаваться компьютерами. RDF – это машиночитаемый и машино-понимаемый формат. Есть небольшие программы (в том числе и online), которые позволяют создать foaf.rdf файлы. Зарегистрировавшись, например на yandex.ru или в ЖЖ, можно также получить автоматически такой файл. Yandex сделал также свое расширение для foaf.

Для семантического веба RDF хорош тем, что можно выполнять специальные SPARQL-запросы. Для этого надо будет собирать эти файлы в единое RDF-хранилище (RDF-story), тогда получилась бы семантическая поисковая машина для поиска людей. Однако подобные проекты (типа FOAF search engine) так и не получили большого развития. Существует также концепция Linked Data – идея соединить все RDF данные.

Концепция FOAF файлов поддерживается всеми крупными социальными сетями, которые генерируют для каждого своего пользователя такой файл и становятся, таким образом, открытыми социальными сетями. Тем самым формируется онтология FOAF (Friend of a friend) для определения участников социальной сети и контента, который они добавляют в сеть. Эта онтология описывает людей, их активность и отношения к другим людям и объектам. Описание социальных связей между людьми в FOAF основывается на транзитивности доверия.

Концепция FOAF используется также в других сетевых онтологиях, например, SIOC initiative (Semantically-Interlinked Online Communities), Online Presence Ontology, SemSNI (Semantic Social Network Interactions) онтологии для анализа социальных сетей SemSNA.

При помощи этих онтологий можно оценить параметры социальных сетей (диаметр, количество участников, среднюю длину пути и др.). Можно также проанализировать виды элементов сети – люди, объекты (музыка, фото, видео, сообщения) – и их взаимодействия (знает, сообщает, комментирует и т.д.). В рамках семантического анализа сети удастся вычислить параметры подграфов по разным типам семантических связей (семья, мне нравится, дружба) и типам взаимодействий (комментирует, создает сообщение и др.).

4.5. СЛУЧАЙНЫЕ СЕТИ И БЕЗМАСШТАБНЫЕ ГРАФЫ

Безмасштабная сеть или масштабно-инвариантная сеть (см. подробное описание и библиографию в [Берновский, 2012]) – граф, в котором степени вершин распределены по степенному закону, т.е. доля вершин со степенью k примерно или асимптотически пропорциональна $k^{-\gamma}$.

Теория случайных графов стала интенсивно развиваться с конца 50-х годов прошлого века после публикации статей Эрдеша–Реньи об эволюции случайных графов. В этой модели все ребра появляются случайно и независимо друг от друга с одинаковой вероятностью p . Под эволюцией понимается изменение свойств графов с ростом вероятности p . Оказалось, что при некоторых значениях p происходит так называемый фазовый переход в эволюции, т.е. свойства графа кардинально меняются. А именно, если

$p = \frac{c \ln n}{n}$, то при $c < 1$ случайный граф почти наверняка несвязен, а при $c > 1$ – почти наверняка связан.

Однако в начале 2000-х годов выяснилось, что модель Эрдеша–Реньи плохо описывает реальные графы, возникающие в различных областях, в частности, графы таких социальных сетей, как Facebook, Twitter и т.п. Барабаши и Альберт заметили следующие закономерности в веб-графе (графе, вершинами которого являются сайты, а ребра соответствуют ссылкам):

- веб-граф разрежен (на n вершинах у него mn ребер, $m \in N$)
- он подчиняется феномену «малого мира» (его диаметр $\approx 5-7$)
- он подчиняется степенному закону:

$$\frac{|\{v : \deg(v) = d\}|}{n} = \frac{c}{d^\lambda}; \quad c = const, \quad \lambda \sim 2.1,$$

где d – диаметр графа, $\deg(v)$ – степень вершины v в графе.

На основании своих наблюдений авторы ввели понятие предпочтительного присоединения. Рассмотрим процесс генерации графа. На n -м

шаге мы добавляем новую вершину n с m ребрами, инцидентными ей, причем вероятность ребра к вершине i пропорциональна степени вершины i (см. рис. 14):

$$P(\text{ребро из } n \text{ в } i) = \frac{\text{deg}(i)}{\sum_j \text{deg}(j)}$$

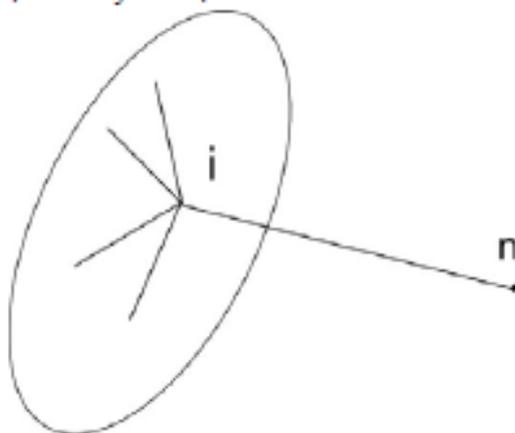


Рис. 14. Добавление новой вершины на n -м шаге

Тем самым мы получили имитационную модель, которая позволяет предсказать многое в поведении реальных сетей. Главная проблема – нужно уметь имитировать показатель степени λ . Для этого были последовательно разработаны следующие группы моделей.

- модель Боллобаша-Риордана (степень равна 3) и ее обобщения (степень не меньше 2). Buckley, Osthus и другие исследователи предложили модификацию модели Барабаша-Альберт, в которой вершины обладают «изначальной привлекательностью»: вероятность того, что старая вершина будет выбрана соседом новой вершины, пропорциональна ее входящей степени плюс константе, т.е. «изначальной привлекательности», am , где m – число ребер, входящих в новую вершину. Если $a = 1$, получаем модель Барабаша-Альберт, так как там используется полная степень, а каждая исходящая степень равна m . Buckley и Osthus обобщили результат для произвольной привлекательности a . Однако и здесь есть проблема: модель говорит о том, что все клики в графе имеют одинаковую мощность, а это не так. Клики большей мощности в сети встречаются, что связано с действием спамеров и агентств по раскрутке сайтов (групп в социальных сетях). Спам в модели Боллобаша-Риордана не учитывается. Эта проблема решена в так называемой модели копирования.
- модели, в которых показатель степенной зависимости может задаваться произвольно, что позволяет изучать эффекты фазовых переходов при его изменении. К числу таких моделей можно отнести модель Чунг-Лу и модель Лучака-Янсона.

Эти модели позволяют моделировать разные типы клик в реальных графах.

– самый многочисленный класс – модели, в которых характеристики и свойства графов определяются эмпирически. Про такие модели не доказано никаких содержательных результатов. Примерами таких моделей могут являться модель Forest Fire и ряд других.

Насколько эти модели соответствуют реальности? Посмотрим, как можно получить ответ на этот вопрос.

Большие социальные сети (например, Facebook и MySpace) не публикуют свой веб-граф, поэтому на текущий момент его характеристики приходится получать экспериментально. Способ для этого – создание популярных приложений в социальной сети (с числом пользователей не менее нескольких миллионов) и последующий сбор данных их пользователей. Например, сотрудники центра RUBINET написали три приложения для Facebook с общим числом пользователей около 8 миллионов. Анонимизированные данные, полученные с помощью этих приложений, доступны online. Проанализировав различные характеристики пользователей (их географическое расположение, связи друг с другом и другие) и их зависимость от типа приложения, исследователи пришли к следующим выводам:

- граф отношений (interaction graph) пользователей обладает свойством «малого мира»;
- пользователи неигровых приложений образуют более плотные сообщества, чем пользователи игр. При этом популярность приложений характеризуется степенным законом с экспоненциальным угасанием.

Другой путь – создание генераторов синтетических веб-графов, способных воспроизводить выбранные модели. Существуют свободно распространяемые генераторы. Качество работы генератора определяется путем сравнения некоторых параметров (например, размера плотных подграфов) получаемого графа с их теоретическими оценками. Наиболее полный отчет приведен авторами генератора DIGG.

5. Примеры применения интеллектуальных систем в управлении социальными процессами

5.1. Примеры прикладных задач

Задачи управления процессами, происходящими в социальных сетях, составляют исключительно интересную и важную область управления социальными процессами в целом. Круг этих задач постоянно расширяется, и в настоящем пособии представлены только некоторые примеры, в которых в качестве средств поддержки решения могут использоваться интеллектуальные системы.

В качестве базы для решения этих задач выбирается один из подходов, рассмотренных в § 1.6.2 – структурный, ресурсный, нормативный или динамический. Напомним их основные концепции.

В структурном подходе все участники сети рассматриваются как вершины графа, которые влияют на конфигурацию ребер и других участников сети. Исследуются такие характеристики сети, как взаимное расположение вершин, центральность, транзитивность взаимодействий.

Ресурсный подход рассматривает возможности участников по привлечению индивидуальных и сетевых ресурсов для достижения определенных целей. В качестве индивидуальных ресурсов могут выступать знания, престиж, богатство, раса, пол, в качестве сетевых ресурсов – влияние, статус, информация, капитал.

Нормативное направление изучает уровень доверия между участниками, а также нормы, правила и санкции, которые влияют на поведение участников в социальной сети и процессы их взаимодействий. Комбинация индивидуальных и сетевых ресурсов участника с нормами и правилами, действующими в данной социальной сети, образует его «сетевой капитал» – преимущества, которые участник может получить в произвольный момент времени для достижения некоторой цели.

Динамический подход исследует изменения в сетевой структуре с течением времени: по каким причинам исчезают и появляются ребра сети, как сеть изменяет свою структуру при внешних воздействиях, существуют ли какие-либо стационарные (равновесные) конфигурации социальной сети и др.

Перечислим некоторые характерные постановки задач сетевого анализа, в инструментарий решения которых входят интеллектуальные системы [Давыденко, 2005; Батура, 2012; Charu, 2011].

Статистический анализ социальных сетей организуется по-разному в зависимости от масштаба сети. Многие из вопросов, которые могли быть заданы при исследовании малых сетей, не имеют смысла в больших сетях. Например, вопрос «Какая вершина сети является наиболее критической в

связности сети, если она будет удалена?» не имеет значения в больших сетях. С другой стороны, вопрос «Какой процент вершин должен быть удален, чтобы существенно влиять на связность сети?» имеет реальное значение для очень больших сетей.

Как правило, статистический анализ фокусируется на крупномасштабных сетях. Исследуются «типичные» социальные сети, и изучается вопрос, как будет выглядеть сеть, если ее увеличить. Изучается взаимное поведение вершин сети исходя из предположения, что у большинства вершин имеется мало связей, возникают ли при этом «ядра» (скопления) или степени вершин распределяются более равномерно. Изучается поведение вершин при кластеризации.

Исследуется поведение типичных временных характеристик социальных сетей. Например, как меняется структура сети в процессе роста или как меняется поведение и распределение связанных компонентов графа. Со временем к сети добавляются новые сущности, но, несмотря на это, некоторые свойства графа могут сохраняться.

Определение сообществ в социальных сетях. Этот вопрос является наиболее важным в анализе социальных сетей. Цель – попытаться определить регионы сети, внутри которых происходит активное взаимодействие участников.

Алгоритмически эту задачу можно отнести к задаче о разделении графов. Нужно разделить сеть на плотные регионы на основе поведения связей между вершинами. Компьютерные социальные сети являются высокодинамичными, что приводит к затруднениям с точки зрения выявления сообществ. Однако в некоторых случаях удастся интегрировать информационное содержимое сети (контент) в процесс определения сообществ. Тогда контент является вспомогательным средством для выявления групп участников с похожими интересами.

Анализ контента социальных сетей. Можно выделить четыре вида анализа контента сети.

Анализ медиаданных. Для обнаружения полезных бизнес-приложений можно анализировать социальные медиасети. Техники анализа данных предоставляют исследователям и специалистам инструменты для анализа больших, комбинированных, постоянно меняющихся медиаданных.

Анализ текстовой информации в социальных сетях. В вершинах социальной сети содержится много текстовой информации в различных формах, например, ссылки на посты (сообщения), блоги или статьи с новостями. Иногда пользователи могут отмечать друг друга, что тоже является формой текстовой информации в виде ссылок. Использование контента сети может сильно улучшить качество выводов при анализе социальных сетей, например, в задачах кластеризации и классификации.

Интеграция данных, поступающих с датчиков, и социальных сетей. Многие современные сотовые телефоны поддерживают возможность вза-

имодействия пользователей друг с другом динамически в режиме реального времени в зависимости от их местоположения и статуса. Подобные приложения также приводят к образованию потоков массивов в режиме реального времени. Их применяют для того, чтобы получить информацию о человеке или совокупности свойств объектов, которые отслеживаются. Поскольку информация о местоположении пользователя является личной, естественно, возникает ряд проблем с точки зрения обработки, исследуются методы интеграции данных, поступающих с датчиков, и данных в социальных сетях.

Анализ мультимедийного контента сети. Существует много сайтов по обмену средствами массовой информации, например, Flickr и YouTube, которые обеспечивают возможность совместного использования этой информации. Такие средства массовой информации общего пользования часто сочетаются с взаимодействием пользователей – размещением тегов и комментариев в различных изображениях. Поэтому подобные сети могут служить источником для широкого спектра приложений в процессе извлечения и анализа данных.

Расстановка тегов. Большинство взаимодействий между пользователями происходит в форме тегирования (расстановки тегов, отметок), в которых пользователи прикрепляют описания различных объектов в социальной сети, такие как картинки, текст, видео или другая мультимедийная информация. В рамках данного подхода изучают свойства потоков тегов, моделей тегирования, семантику тегов, рекомендации по использованию тегов, визуализацию тегов, приложения для расстановки тегов, интеграцию различных систем тегирования и проблем, связанных с использованием тегов. Интересным вопросом является, например, почему люди расставляют теги, что влияет на выбор при тегировании, как промоделировать процесс подобной разметки, разновидности тегов, как создаются теги и как выбрать правильные теги для рекомендации.

Случайные блуждания и их применение в социальных сетях. Классификация – один из наиболее известных методов в веб-поиске. Например, можно упомянуть алгоритм ссылочного ранжирования (PageRank) для приписывания веса веб-документам. Его основной принцип может применяться для поиска и классификации сущностей и участников в социальной сети. Этот алгоритм использует подход случайного блуждания для того, чтобы оценить вероятность посещения той или иной вершины. Естественно, что вершины, которые лучше расположены со структурной точки зрения, имеют более высокий вес, а значит, являются более важными. Методы случайного блуждания могут быть также полезны для объединения участников в группы относительно наиболее влиятельных участников.

Классификация вершин в социальных сетях. Некоторые вершины удобно снабжать пометками, чтобы их отличительные особенности и

структурную информацию можно было распространить на всю сеть. Например, в маркетинговых исследованиях определенные вершины могут обозначать заинтересованность участников сети в конкретном продукте, и было бы желательно применить характерные особенности этих вершин для изучения других участников на предмет заинтересованности этим продуктом. Для этих целей, кроме того, можно использовать информацию о контенте и структуре социальной сети. Другой пример: когда об одной из двух связанных вершин получены некоторые сведения, для второй эти сведения с большой долей вероятности тоже будут верны. Вот почему структуру связей можно применять для распространения меток среди вершин. Содержимое сети и структурные особенности в дальнейшем могут пригодиться для подтверждения качества полученной классификации.

Анализ социального влияния. Так как в основе социальных сетей лежит взаимодействие между различными участниками, естественно предположить, что это взаимодействие оказывает влияние на участников в терминах их поведения. Классическим примером является применение «вирусного маркетинга» для распространения сообщения между взаимосвязанными участниками через всю сеть. Вопросы этого направления:

- как моделировать влияние на основе информации об участниках;
- как моделировать распространение влияния;
- кто из участников является наиболее влиятельным в процессе распространения?

Напомним, что эти вопросы рассматривались в § 2.2 и § 3.1.

Конфиденциальность в социальных сетях. В социальных сетях содержится большое количество личной информации об участниках, например, интересы, информация о дружбе, демографическая информация и др. Это может привести к несанкционированному распространению личной информации в сетях. В решении такого типа задач полезно применять модели на основе механизмов конфиденциальности.

Обнаружение экспертов в сетях. Социальная сеть может являться инструментом для выявления экспертов в конкретной области. Часто в реальности эксперты образуют сеть, которая соответствует социальной сети или организационной структуре компании. Многие сложные задачи требуют коллективного решения нескольких экспертов. В подобных случаях получается, что более эффективно можно достигнуть общей цели, когда специалисты сотрудничают друг с другом. Важной задачей данного направления является обнаружение групп специалистов в определенных узких областях.

Эволюция в динамических социальных сетях. С течением времени в социальных сетях появляются новые участники, некоторые участники прекращают взаимодействие, возникают новые связи, некоторые связи устаревают, так как участники перестают взаимодействовать. Это приво-

дит к изменениям в структуре социальных сетей в целом и в отдельных сообществах. При этом возникает следующие вопросы:

- согласно каким законам происходят долгосрочные изменения между крупными сообществами в социальных сетях;
- как развиваются сами сообщества во времени;
- какие изменения могут происходить, как можно отследить и представить их.

Прогноз формирования связей в социальных сетях. Для извлечения интересующей информации из социальной сети полезны исследования, направленные на определение и предсказание возможных связей между вершинами в будущем. В большинстве приложений для анализа социальных сетей связи считаются динамическими и могут сильно изменяться с течением времени. В процесс прогнозирования связей может быть вовлечена как структура сети, так и информация об особенностях различных вершин. Для решения таких задач предлагается строить разнообразные структурные и реляционные модели.

Визуализация социальных сетей. Социальные сети становятся крупнее и имеют все более сложную структуру. Визуализация помогает естественным образом свести воедино информацию о сетях и сделать ее более доступной для понимания. Визуализация помогает аналитикам в описании социальных сетей. Целью этого направления является также поиск ответа на вопрос, как различные модели могут быть использованы для изучения различных аспектов сетей, таких как структура и семантика. Важным является создание алгоритмов, сочетающих методы анализа и визуализации, чтобы улучшить понимание структуры и динамики сети.

5.2. Поддержка рекламной кампании в социальной сети

Как было сказано выше, социальные сети предоставляют значительные возможности для проведения маркетинговых компаний. В данном параграфе описываются шаги, необходимые для проведения подобной компании среди пользователей сети, интересующихся фотографией.

Для создания рекламного объявления в социальной сети «ВКонтакте» необходимо задать критерий отбора целевой аудитории – пользователей, которым будет отображаться объявление. Социальная сеть «ВКонтакте» предоставляет следующие параметры:

- пол;
- семейное положение;
- возраст;
- страна;
- города;
- районы;
- станции метро;

- улицы;
- учебные заведения;
- год окончания школы;
- год окончания вуза;
- категории групп;
- группы;
- должности;
- религиозные взгляды;
- интересы.

Таким образом, устанавливая значение какого-либо параметра из показанных выше, мы сужаем аудиторию. Задачей эксперимента является сужение аудитории потенциальных клиентов контекстного объявления до множества людей, интересующихся выбранной предметной областью, таким образом достигнув повышения эффективности кампании.

Исходя из анализа имеющихся параметров, можно сказать, что наиболее значимыми для описания интересов человека факторами являются его интересы, группы и категории групп (так как обычно человек вступает в те группы, тематиками которых он заинтересован). Данные параметры представляют собой поля для ввода текста. После создания объявления с соответствующими значениями параметров социальная сеть будет транслировать данное объявление подходящей под описание целевой аудитории.

Одна из имеющихся в данной области проблем заключается в определении релевантного для предметной области списка ключевых слов. Можно поступить несколькими способами:

1. купить ключевые слова;
2. извлечь ключевые слова из данных какого-либо интернет-портала.

Будем использовать метод 2. Наиболее подходящим методом для выделения интересов пользователей социальной сети «ВКонтакте» являлось бы выполнение следующего алгоритма:

1. Получение интересов всех пользователей «ВКонтакте»
2. Выбор по некоторым ключевым словам пользователей, имеющих интересы, связанные с заданной предметной областью
3. Проведение анализа имеющихся данных и выявление встречающихся наиболее часто.

Но выполнению этих действий препятствует следующий пункт правил:

«5.3.9. использовать без специального на то разрешения Администрации Сайта автоматизированные скрипты (программы) для сбора информации на Сайте и(или) взаимодействия с Сайтом и его сервисами;».

Также для анализа страниц социальной сети «ВКонтакте» пришлось бы создавать обработчики HTML-данных (так как информация на сайте представлена в виде HTML), а также политики обхода временного ограни-

чения на доступ к страницам. Ввиду наличия этих факторов было принято решение проводить сходные действия в другой социальной сети.

Как отмечалось в § 4.4, большинство социальных сетей, в том числе сеть «ВКонтакте», экспортируют свои данные в формате FOAF.

FOAF (англ. Friend of a Friend) – проект по созданию модели машинно-читаемых домашних страниц и социальных сетей. Ядром проекта является спецификация, которая определяет некоторые выражения, используемые в высказываниях о ком-либо: например, имя, пол и другие характеристики. Чтобы сослаться на эти данные, используется идентификатор, включающий уникальные свойства друга (например, SHA1 – сумма от E-Mail адреса, Jabber ID, или URI домашней страницы, веблога).

Основанный на RDF, определенный с помощью Web Ontology Language и разработанный для легкой расширяемости, FOAF позволяет распределять данные между различными компьютерными окружениями.

Resource Description Framework (RDF) – это модель для представления данных, в особенности метаданных. RDF представляет утверждения о ресурсах в виде, пригодном для машинной обработки. RDF является частью концепции семантической паутины.

Ресурсом в RDF может быть любая сущность – как информационная (например, веб-сайт или изображение), так и неинформационная (например, человек, город или некое абстрактное понятие). Утверждение, высказываемое о ресурсе, имеет вид «субъект – предикат – объект» и называется триплетом. Утверждение «небо голубого цвета» в RDF-терминологии можно представить следующим образом: субъект – «небо», предикат – «имеет цвет», объект – «голубой». Для обозначения субъектов, предикатов и объектов в RDF используются URI. Множество RDF-утверждений образует ориентированный граф, в котором вершинами являются субъекты и объекты, а ребра помечены предикатами.

RDF сам по себе является не форматом файла, а только лишь абстрактной моделью. Для записи и передачи RDF используется несколько форматов, в том числе:

- RDF/XML – запись в виде XML-документа;
- RDFa (англ. RDF in attributes) – запись внутри атрибутов произвольного HTML- или XHTML-документа;
- N-Triples, Turtle, N3 – компактные формы записи утверждений.

Для обработки данных, представленных в RDF, могут использоваться языки запросов: SPARQL (стандарт W3C), RQL, RDQL.

Крупной социальной сетью, экспортирующей свои данные в формат FOAF, является социальная сеть «LiveJournal», <http://livejournal.com>. LiveJournal занимает 9-е место в списке сайтов, упорядоченных по популярности для Российской Федерации.

Экспорт данных осуществляется следующим образом: к URL адресу блога или сообщества (например, http://community.livejournal.com/blog_medvedevA) необходимо добавить строку «data/foaf», например:

http://community.livejournal.com/blog_medvedev/data/foaf.

По запросу по данному URL адресу клиенту будет выдан FOAF файл, содержащий описание блога\сообщества.

Совершая запросы к данным в этом формате, можно получить всю информацию о блоге, ссылки на профайлы друзей и др. в заранее известном формате, что сильно облегчает обход и сбор информации. Для нашего исследования нам интересны теги, описывающие интересы пользователя.

Вследствие того, что сеть «Живой Журнал» представляет возможность поиска по интересам, был выбран следующий алгоритм действия:

1. Поиск по интересу «фото» всех сообществ
2. Получение списка интересов всех этих сообществ
3. Проведение анализа и выяснение, какие интересы встречаются наиболее часто (для создания рекламной кампании).

Пункт 1 был выполнен полуавтоматическим методом: сначала был проведен поиск всех сообществ (было выдано 412 результатов), после чего выделение непосредственно имени сообщества, являющегося частью URL адреса, было проведено при помощи поиска и замены с помощью регулярных выражений в текстовом редакторе NotePad++. Получение списка всех интересов сообществ было выполнено при помощи создания программы на языке программирования PHP.

```
<?php
//ini_set ('display_errors', 1);
$data = file('data.txt'); require_once 'XML/FOAF/Parser.php';
foreach($data as $z)
{
    $z = chop($z); //echo "$z:\n"; $handle = file_get_contents('http://community.livejournal.com/' . $z . '/data/foaf');
    $parser = new XML_FOAF_Parser;
    $parser->parseFromMem($handle);
    $sarr = $parser->toArray();
    $k = 0;
    //echo "elements: " . (count($sarr) - 2) . "\n";
    foreach ($sarr["dc"] ['title'] as $j => $i)
    {
        if($k == 0){$k++; continue;}
        if($j == 'http://community.livejournal.com/' . $z . '/profile'){continue;}
    }
}
```

```

echo "$i\n";
}

```

?>

Суть программы заключается в чтении файла со списком сообществ, организации http-запроса к данным каждого из них, получении списка интересов и записи их в файл. Обход FOAF-данных осуществляется при помощи плагина к PHP, содержащего уже реализованный анализ FOAF (http://pear.php.net/package/XML_FOAF). Результатом работы программы послужил список из 14000 ключевых слов, из которых при помощи скрипта на PHP были выбраны наиболее частые.

```

<?php
$h =
file('results.txt');
$arr = array(); fo-
reach($h as $i)

{

$i = chop($i);

if ( !isset($arr[$i]))

{
$arr[$i]
= 1; } else
{

$arr[$i]++;

}

}

arsort($arr) ; foreach($arr
as $i => $j )

{
echo "$i\t$j\n";
}
?>

```

Наиболее частотные слова из выбранного списка представлены ниже:

фото	412
фотография	207
photo	104
искусство	81
музыка	79
foto	76
природа	72
photography	68

фотографии	66
дизайн	64
любовь	57
кино	57
art	51
люди	47
солнце	47
реклама	47
креатив	43
canon	43
nikon	43
журналистика	43
путешествия	41
красота	41
секс	41
портрет	38
стихи	38
фотошоп	38
photoshop	38
творчество	36
фотографировать	36
литература	34
Moscow	34
фотоаппараты	32
лето	32

Таким образом, был проведено экспериментальное исследование социальной сети «ВКонтакте» для поддержки планирования рекламной кампании услуг фотографа. С использованием программного обеспечения, анализирующего социальную сеть, было проведено извлечение ключевых слов из социальной сети LiveJournal, служащих для сужения целевой аудитории.

5.2. Автоматизированное пополнение онтологии предметной области

Согласно [Gruber, 1993], под онтологиями понимают систему явной концептуализации предметной области, т.е. формального представления предметной области.

Общим для всех формализаций является выделение множества объектов (концептов, понятий), алфавита отношений, правил установления отношений и аксиом, задающих правила вывода на множестве отношений.

Выделяются фундаментальные онтологии, которые описывают предметную область максимально полно, безотносительно к приложениям и

обычно с максимальной степенью формализации, и прикладные онтологии, которые также называются «легкими» онтологиями и которые формализуются настолько, насколько это необходимо для приложения.

Применение онтологий в управлении социальными системами. Приведем некоторые примеры.

Во многих дисциплинах сейчас разрабатываются стандартные онтологии, которые могут использоваться экспертами по предметным областям для совместного использования и аннотирования информации в своей области. Например, в области медицины созданы большие стандартные, структурированные словари, такие как *snomed* и семантическая сеть *Unified Medical Language System*.

Появляются обширные общецелевые онтологии. Например, Программа ООН по развитию (*the United Nations Development Program*) и компания *Dun & Bradstreet* объединили усилия для разработки онтологии *UNSPSC*, которая предоставляет терминологию товаров и услуг (<http://www.unspsc.org/>).

О создании своих онтологий объявили практически все большие компании, занятые обработкой текстов и семантическим анализом – *ABBYY* в проекте *NLC*, *Яндекс* и *Google* в области семантического анализа.

Помимо онтологий контента, появляются онтологии для организаций. Например, онтологическая модель вуза используется для повышения эффективности автоматизации управления в вузе.

Процесс создания онтологий. Строго говоря, универсального подхода к созданию онтологий, который бы привел к однозначно успешному результату, не существует. Процесс создания онтологий обычно является итеративным, т.е. сначала создается черновой набросок, а затем по мере необходимости происходит возврат для определения деталей, и так продолжается до тех пор, пока онтология не будет отражать концепцию предметной области с определенной степенью.

Практически создание онтологий включает следующие шаги:

1. Определение классов в онтологии;
2. Организация классов в некоторую иерархию (базовый класс – подкласс);
3. Определение слотов и их допустимых значений;
4. Заполнение значений слотов для экземпляров классов.

При правильном моделировании онтология может быть представлена предложениями, где, вероятнее всего, в качестве объектов выступают существительные, а в качестве отношений – глаголы.

На практике обычно используется следующий подход: экспертным путем строится каркас онтологии, т.е. базовая структура отношений, а затем организуется ее автоматизированное пополнение.

«Каркас» онтологии должен отражать иерархическую структуру предметной области, т.е. содержать в первую очередь связи иерархическо-

го типа (a kind of, a part of). Построение целесообразно выполнять в ручном режиме на основе содержательного анализа текста.

Подходы к автоматизированному пополнению онтологий. Общей методики для этого не существует. Практически все алгоритмы автоматизированного пополнения онтологий являются эвристическими, с ними нужно знакомиться по литературным источникам. Ниже приводятся некоторые примеры идей, лежащих в основе таких алгоритмов.

- Используется подход машинного обучения. Выполняется обработка естественного языка (natural language processing, NLP) с классификатором типа «наивный Байес», обучаемым по документам с заданными ключевыми фразами.
- Используется глоссарий (Glossary-based approaches). Онтология пополняется по ключевым словам, которые задает автор.
- Эвристика KIP. Основная задача – выделить многословные понятия. Сравниваются вес ключевой фразы (получается из NLP) и вес ключевого слова в ключевой фразе-кандидате. Веса назначаются по встречаемости в глоссарии предметной области
- Эвристика TextOntoEx. Использование семантических паттернов. TextOntoEx не выявляет новые отношения, но он выявляет примеры уже известных отношений в текущем тексте. Для этого строятся паттерны с учетом семантики предметной области.
- Выделение ключевых концептов на основе графового подхода. Каждый термин рассматривается как узел в графе; их значимость – частота связи с ближайшими узлами. Тогда концепт – это кластер связанных терминов. Проблема – трудно интерпретировать получаемую визуальную репрезентацию связи всех слов в тексте.

Практический пример автоматизированного пополнения онтологии [Chi, 2014]. Цель онтологии – помочь анализу рисков в строительной индустрии (ЖНА). Сюда входят:

- (1) работы с самыми высокими показателями ранений или заболеваний;
- (2) работы, которые могут привести к тяжелым повреждениям или потере трудоспособности;
- (3) работы, в которых простая человеческая ошибка может привести к несчастному случаю с тяжелыми последствиями.

Использованы следующие текстовые ресурсы:

CPWR – база строительных решений. Содержит информацию о активностях, задачах и рисках, которые описывают различные рабочие сценарии в строительном бизнесе. Содержит 15 типов активности (например, покрытие кровлей) и 16 типов рисков (например, падение с высоты).

NIOSH – правительственная программа, которая, с целью формирования превентивных стратегий, идентифицирует случаи со смертельным исходом, особенно в ситуациях, которые включали высокие риски. Доклады NIOSH служат тестовыми данными в задаче классификации текстов (ТС)

базы CPWR. Для этого 243 доклада NIOSH размечены экспертами по активностям и рискам, определяемым в CPWR.

OSHA – стандарт промышленной безопасности, в нем все строительные работы разделены на категории, и для каждой записаны основные правила безопасности.

При этом CPWR и NIOSH могут иметь похожую структуру документов и перекрывающиеся темы; однако они используют разные фразы и разную терминологию для описания одних и тех же концептов. Добавление стандартов OSHA позволяет решить эту проблему.

Все три текстовых ресурса объединяются в онтологию, как показано на рис. 15.

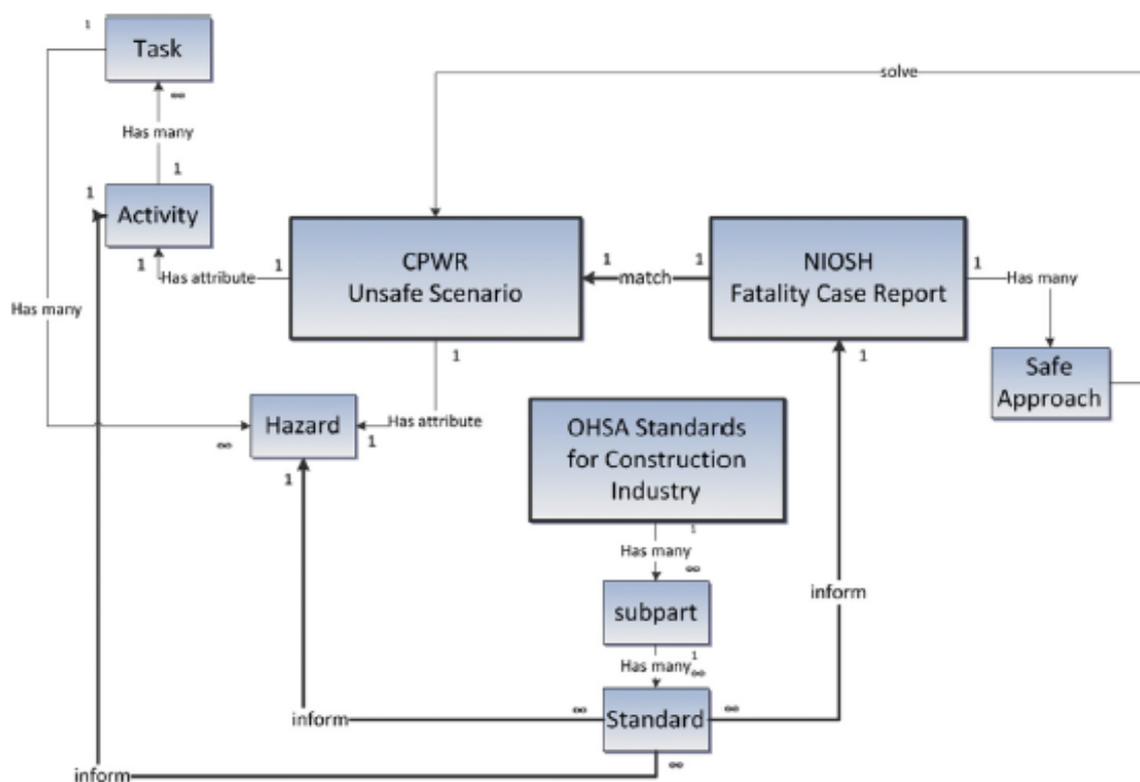


Рис. 15. Структура онтологии рисков в строительной индустрии

Как все это работает при выполнении анализа рисков в строительной индустрии (ЖНА)?

Шаг 1. Описывается небезопасный сценарий в терминах базы CPWR (например, выбирается активность «строительная сталь» и риск «падение с высоты»).

Шаг 2. Такой случай (активность «строительная сталь» и риск «падение с высоты») отыскивается в докладах NIOSH, и там есть рекомендации по безопасности. Например, в докладе «Ironworker Dies Following an 18-foot Fall from Structural Steel Framework» есть 4 таких рекомендации.

Шаг 3. В CPWR и NIOSH для этого случая имеется общее выражение «падение с высоты». В стандарте OSHA отыскиваются кореллирующие этому

выражению стандарты, которые дополняют контент сценариев, чтобы сделать его более точным.

Как пополняется построенная онтология?

При этом решается задача классификации вновь поступающих из Интернета текстов (ТС) – разложение текстов на predetermined категории. Используется простейшая методика – по частотному вектору терминов текста (TF-IDF model)

1 – Получение структуры классификации. Activities разделяются на классы (согласно CPWR и NIOSH) и группируются (при их содержательном совпадении). Производится выделение наиболее важных активностей (по количеству реальных случаев, которые им соответствуют). Остается 9 типов активностей и 4 типа рисков.

2 – Расширение полученной онтологии (фиг. 16).

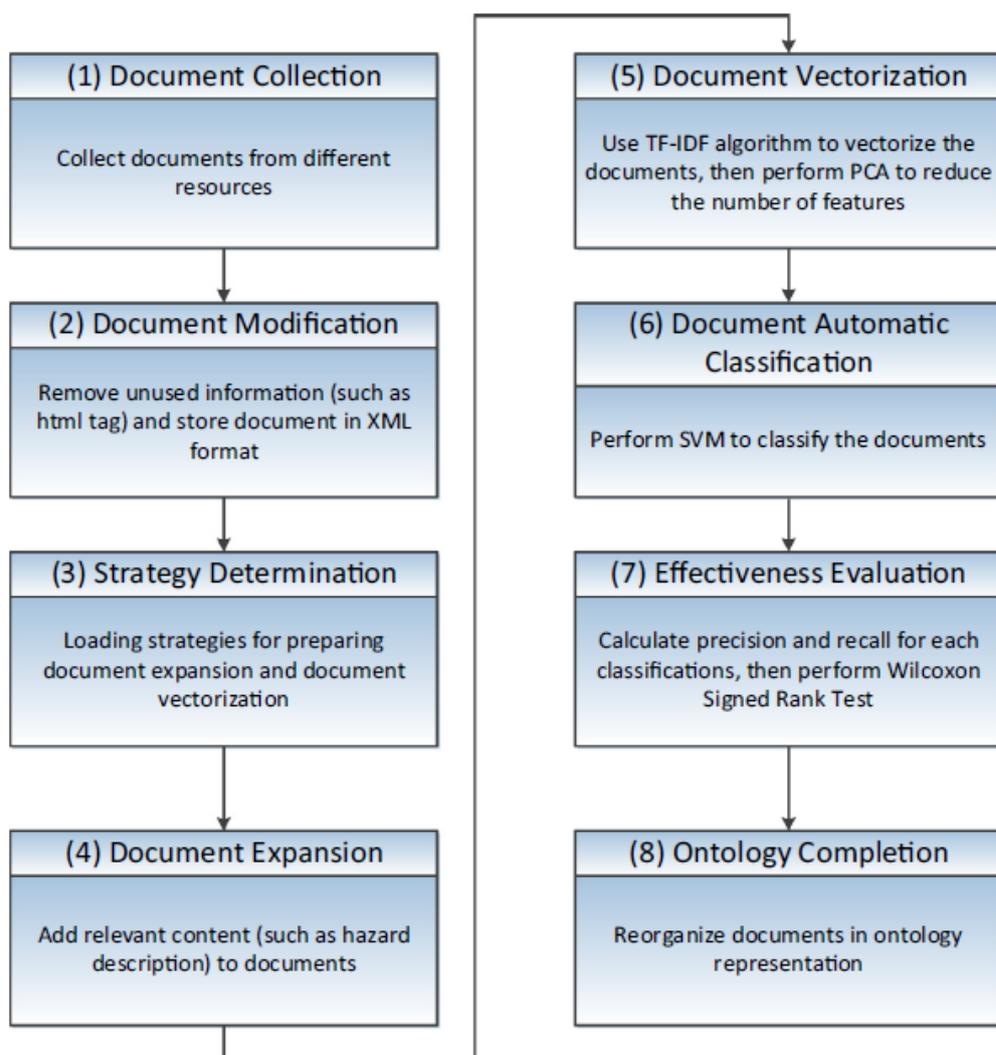


Рис. 16. Процедура пополнения онтологии

Взаимодействие программных продуктов, используемых для пополнения онтологии, представлены на рис. 17. Специфика – модули 2. Document vectorization module и 3. Machine learning module. Текст векторизуется (по схеме TF-IDF), затем посредством SVD-разложения вектор сокращается до 25-75-100 терминов. Затем на этих векторах запускается машинное обучение (машина опорных векторов, SVM), уточняются параметры для классификации текстов.

Результирующая онтология (рис. 18) служит для наиболее точного подбора примеров под классифицируемые случаи.

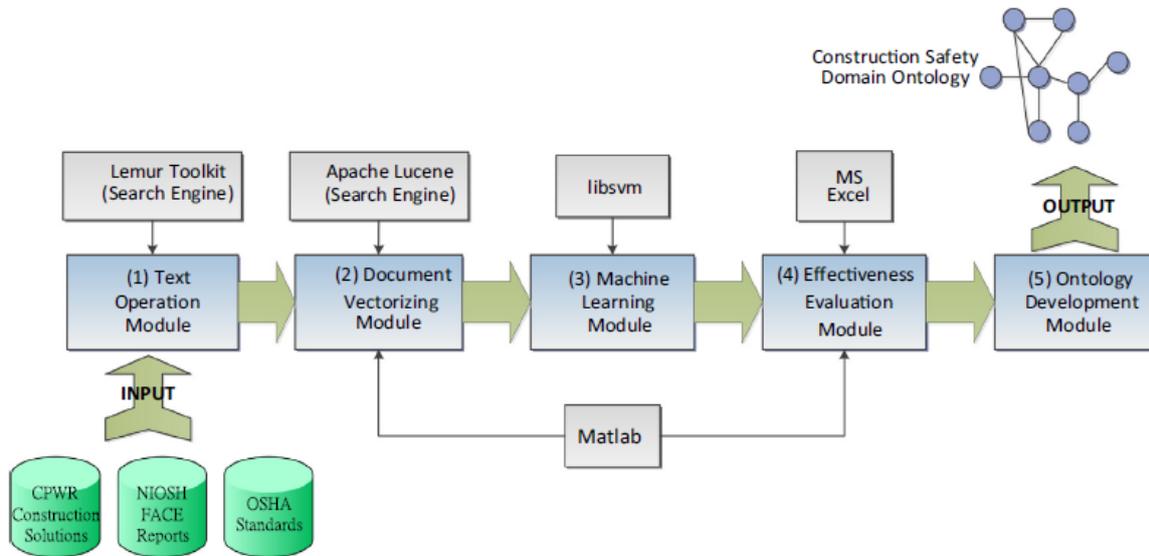


Рис. 17. Взаимодействие программных продуктов для пополнения онтологии

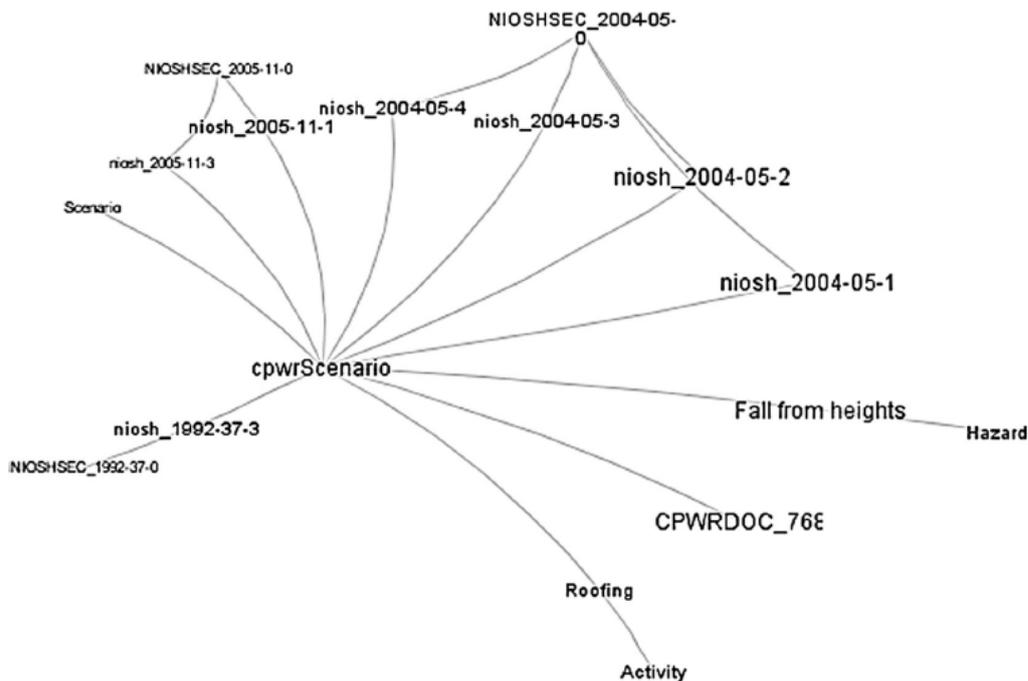


Рис. 18. Фрагмент онтологии строительных рисков

Литература

- [Батура, 2012] Батура Т.В. Методы анализа компьютерных социальных сетей // Вестник НГУ. Серия: Информационные технологии. 2012. Том 10. Вып. 4. С. 13–28.
- [Берновский, 2012] Берновский М.М., Кузюрин Н.Н. Случайные графы, модели и генераторы безмасштабных графов // Труды Института системного программирования РАН. 2012. Том 22. С. 419–432.
- [БСЭ] Большая советская энциклопедия. Электронный ресурс. – Режим доступа: <http://dic.academic.ru/dic.nsf/bse/143029>
- [Губанов, 2009] Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели информационного влияния и информационного управления в социальных сетях // Проблемы управления. 2009. № 5. С. 28–35.
- [Губанов, 2010] Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. – М.: Изд. физ.-мат. литературы, 2010. – 228 с.
- [Давыденко, 2005] Давыденко В.А., Ромашкина Г.Ф., Чуканов С.Н. Моделирование социальных сетей // Вестник Тюменского государственного университета. № 1, 2005. С. 68–79.
- [Иванов, 2001] Иванов В.Н. Основы социального управления. Учебное пособие / А. Г. Гладышев, В. Н. Иванов, В. И. Патрушев и др. Под ред. В. Н. Иванова. — М.: Высш. шк., 2001. — 271 с.
- [Лисьев, 2011] Лисьев Г.А., Попова И.В. Технологии поддержки принятия решений. — 2-е издание, стереотипное. — М.: Издательство «ФЛИНТА», 2011. — 133 с. — ISBN 978-5-9765-1300-6.
- [Новиков, Петраков, 1999] Новиков Д.А., Петраков С.Н. Курс теории активных систем М.: СИНТЕГ, 1999. – 104 с.
- [Общая социология, 2009] Общая социология: Учебное пособие / Под ред. проф. А.Г. Эфендиева. – М.: ИНФРА-М, 2009. – 654 с.
- [Рыбина, 2010] Рыбина Г.В. Основы построения интеллектуальных систем. — М.: Финансы и статистика; ИНФРА-М, 2010. — 432 с. — ISBN 978-5-279-03412-3.
- [Социология, 2009] Социология: Учебник для вузов / Под ред. проф. В.Н. Лавриненко. – М.: ЮНИТИ-ДАНА, 2009. – 448 с.
- [Charu, 2012] Charu C. Aggarwal. Social Network Data Analytics. 2011. 520 p.
- [Chi, 2014] Nai-Wen Chi, Ken-Yu Lin, Shang-Hsien Hsieh Using ontology-based text classification to assist Job Hazard Analysis // Advanced Engineering Informatics (2014)
- [Domingos, 2002] Domingos P., Richardson M. Mining the Network Value of Customers / Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining. 2002. P. 57-66.

- [Janky, 2002] Janky B., Takacs K. Social Control, Participation in Collective Action and Network Stability. – HUNNET Working Paper, 2002. Электронный ресурс. – Режим доступа: www.socialnetwork.hu
- [Kempe, 2003] Kempe D., Kleinberg J., Tardos E. Maximizing the Spread of Influence through a Social Network / Proceedings of the 9-th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2003. P. 137–146.
- [Leskovec, 2003] Leskovec J., Krause A., Guestrin C., Faloutsos C., Vanbriesen J., Glance N. Cost-effective Outbreak Detection in Networks / Proceedings of the 13-th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2007. P. 420-429.
- [Leskovec, 2005] Leskovec J., Kleinberg J., Faloutsos C. Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations // Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining (KDD). N. Y., 2005. P. 177–187.
- [Newman, 2003] Newman M.E.J. The structure and function of complex networks. // SIAM Review. 2003. Vol. 45. pp. 167–256.
- [Watts, 2007] Watts D., Dodds P. Influentials, Networks, and Public Opinion Formation // Journal of Consumer Research. 2007. № 34. P. 441–458.

Миссия университета – генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

КАФЕДРА ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ В ГУМАНИТАРНОЙ СФЕРЕ

Кафедра интеллектуальных технологий в гуманитарной сфере была организована в 1998 году и при образовании получила название «кафедра технологий профессионального обучения». Тогда же кафедру возглавил профессор Потеев Михаил Иванович, и с 2002 года кафедра стала выпускающей. В 2012 году кафедра была переименована в соответствие с основным направлением деятельности.

Центральной идеей образовательных программ, реализуемых кафедрой, является участие студентов в выполнении работ, связанных с возможными направлениями будущей деятельности, и с задачами, решаемыми университетом. Научные исследования, проводимые на кафедре, связаны с интеллектуальным анализом данных, математическим моделированием и проектированием информационных систем. В этих областях много интересных, сложных и нерешенных задач. На старших курсах студенты имеют возможность выбрать то направление в рамках профиля, которое им наиболее интересно.

Существующие международные программы

Сотрудниками кафедры интеллектуальных технологий в гуманитарной сфере ведутся переговоры о внедрении программ двойных/совместных дипломов и реализации международных программ академического студенческого обмена с вузами стран Шанхайской организации сотрудничества, а также Чехии, Англии и Финляндии.

Компании, в которых осуществляется производственная и преддипломная практика, а также компании, трудоустраивающие выпускников

Доктор Web, Государственный Русский музей, ООО «ИНТЕРФОРУМ», Ростелеком, Интерзет, ООО «Интермедиа», «ВКонтакте» и др.

http://www.ifmo.ru/ru/viewdepartment/13/kafedra_intellektualnyh_tehnologiy_v_gumanitarnoy_sfere.htm#ixzz3byzWK4D7

Гусарова Наталия Федоровна

**ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ
В УПРАВЛЕНИИ СОЦИАЛЬНЫМИ ПРОЦЕССАМИ**

Учебное пособие

В авторской редакции

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Н.Ф. Гусарова

Подписано к печати 08.06.2015

Заказ № 3311

Тираж 100

Отпечатано на ризографе

**Редакционно-издательский отдел
Университета ИТМО
197101, Санкт-Петербург, Кронверкский пр., 49**