

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ ПРОБЛЕМ РЕГИСТРАЦИИ ИНФОРМАЦИИ

А.Г. Додонов, Д.В. Ландэ

ЖИВУЧЕСТЬ
ИНФОРМАЦИОННЫХ
СИСТЕМ

КИЕВ НАУКОВА ДУМКА 2011

УДК 004.5
ББК 22.18, 32.81, 60.54
Ж95

А.Г. Додонов, Д.В. Ландэ. Живучесть информационных систем. — К.: Наук. думка, 2011. — 256 с.

Особое место среди задач, получивших актуальность в настоящее время, занимают задачи, связанные с обеспечением живучести информационных систем.

В рамках данной работы информационная система рассматривается как совокупность содержательно связанных элементов информационного пространства. Рассматриваются вопросы живучести, моделирования деструктивных воздействий, восстановления информационных сетевых структур, формулируются актуальные для функционирования информационных систем задачи, описываются математические модели.

Для широкого круга специалистов в области информационных технологий и информационной безопасности, студентов старших курсов, аспирантов.

Особливе місце серед задач, які стали на сьогодні актуальними, належить задачам, пов'язаним із забезпеченням живучості інформаційних систем.

У межах даної праці інформаційна система розглядається як сукупність змістовно зв'язаних елементів інформаційного простору. Розглядаються питання живучості, моделювання деструктивних впливів, відновлення інформаційних мережних структур, формулюються задачі, що є актуальними для функціонування інформаційних систем, наводяться математичні моделі.

Для широкого кола фахівців у галузі інформаційних технологій та інформаційної безпеки, а також для аспірантів і студентів вищої школи.

Рекомендовано к изданию Ученым советом Института проблем регистрации информации НАН Украины (протокол № 16 от 13 декабря 2010 года)

Рецензенты:

А.Е. Литвиненко — доктор технических наук, профессор.

В.В. Мохор — доктор технических наук, профессор.

Научно-издательский отдел физико-математической литературы
Редактор *В.В. Вероцкая*

ISBN 978-966-00-0973-9

© А.Г. Додонов, Д.В. Ландэ, 2011

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	5
ВВЕДЕНИЕ	7
1. ИНФОРМАЦИОННЫЕ СИСТЕМЫ	12
1.1. Свойства информационных систем	12
1.2. Моделирование информационного пространства	15
1.3. Многоагентные системы	18
1.3.1. Искусственные общества	18
1.3.2. Модель «тепловых жуков»	19
1.3.3. Природные вычисления	20
1.3.4. Модель предпочтений групп людей	20
1.4. Имитационные модели	22
1.5. Индивидуум-ориентированные модели	23
1.6. Модели клеточных автоматов	28
2. АНАЛИЗ И ОЦЕНКА ЖИВУЧЕСТИ	38
2.1. Функциональная живучесть	39
2.2. Структурная живучесть	45
2.3. Теоретико-игровой подход	50
2.4. Логико-вероятностные модели	60
2.5. Оценка живучести системы по ее состоянию	66
2.6. Оценка живучести по результатам выполнения задания	71
2.7. Оценка живучести по цели функционирования	73
2.8. Энтропийный подход к оценке живучести	77
3. СЛОЖНЫЕ ИНФОРМАЦИОННЫЕ СЕТИ	83
3.1. Параметры сложных сетей и задачи живучести	84
3.1.1. Параметры узлов сети	86
3.1.2. Общие параметры сети	87
3.1.3. Распределение степеней связности узлов	87
3.1.4. Путь между узлами	88
3.1.5. Коэффициент кластерности	90
3.1.6. Посредничество	91
3.1.7. Эластичность и уязвимость сети	92
3.1.8. Коэффициент элитарности	93
3.1.9. Корреляция степеней связанных вершин	94
3.2. Модель слабых связей	96
3.3. Модель малых миров	97
3.4. Живучесть системы с разветвленной структурой	100
3.5. Моделирование деструктивного воздействия на сети	102

4. КРИТИЧЕСКИЙ УРОВЕНЬ ЖИВУЧЕСТИ СИСТЕМ	104
4.1. Фазовые переходы	105
4.2. Задача теории перколяции	107
4.3. Характеристики перколяционных сетей	108
4.4. Перколяция на случайных сетях	110
4.5. Теория перколяции и моделирование атак на сети	113
5. ИНФОРМАЦИОННЫЕ ПОТОКИ И ВОПРОСЫ ЖИВУЧЕСТИ	115
5.1. Информационные потоки	115
5.2. Моделирование информационных потоков	125
5.2.1. Тематические информационные потоки	126
5.2.2. Традиционные модели информационных потоков	129
5.2.3. Выявление информационных кластеров	142
5.2.4. Эмерджентность информационных систем	145
5.2.5. Синергетический подход	147
5.2.6. Теоретико-игровой подход	151
5.2.7. Экстремальные подходы	156
5.3. Нелинейные динамические модели	162
5.4. Взаимодействие информационных систем	169
5.4.1. Динамика типа «Конкуренция»	172
5.4.2. Динамика типа «Хищничество»	176
5.4.3. Динамика типа «Симбиоз»	179
5.5. Анализ временных рядов	180
5.5.1. Корреляционный анализ	182
5.5.2. Вейвлет-анализ	183
5.5.3. Фрактальный анализ: R/S-анализ	188
5.5.4. Отклонение от линейного тренда	191
5.5.5. Визуализация на основе ΔL -анализа	192
5.5.6. Мультифрактальный анализ	195
6. ЖИВУЧЕСТЬ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ	198
6.1. Понятие «информационные операции»	198
6.2. Информационные операции как социальные процедуры	201
6.3. Информационное влияние	204
6.4. Этапность информационных операций	205
6.5. Особенности моделирования информационных операций	209
6.6. Мониторинг и анализ информационных операций	212
ЗАКЛЮЧЕНИЕ	236
ГЛОССАРИЙ	237
СПИСОК ЛИТЕРАТУРЫ	246

ПРЕДИСЛОВИЕ

Сложным системам присущи различные свойства, но среди них есть такие, которыми нельзя пренебрегать, исследуя их функционирование, прогнозируя развитие, анализируя их взаимодействие с внешней средой. К таким свойствам относится, в частности, живучесть. Под живучестью системы понимают ее способность адаптироваться к новым непредусматриваемым условиям функционирования, противостояния нежелательным влияниям при одновременной реализации основной функции. Живучесть в традиционном понимании — это фундаментальное свойство сложных систем. Биологические, социальные и многие другие системы изначально обладают свойством живучести, что позволяет им сохранять целостность, выполнять свои функции и развиваться вопреки деградации, независимо от наличия неблагоприятных (деструктивных) воздействий со стороны внешней среды [1]. Методы и средства обеспечения живучести применяются при создании сложных искусственных объектов, в частности информационных систем.

Живучие системы способны поддерживать непрерывное выполнение своих основных функций, временно или постоянно отказываясь от выполнения менее важных функций, изменять свою структуру и поведение, находить и выполнять новые функции, необходимые для успешного противостояния неблагоприятным воздействиям, приспосабливаясь к условиям своего функционирования [2]. Механизмы обеспечения живучести, входящие в такие системы, являются их неотъемлемой частью, и эволюция систем обуславливает эволюцию их механизмов обеспечения живучести. Таким образом, развитие систем является фактором развития механизмов обеспечения живучести.

В рамках данной книги информационная система рассматривается как совокупность содержательно связанных в сеть элементов информационного пространства. Соответственно анализируются свойства этой сети, имеющие отношение к живучести, среди которых уязвимость как характеристика нарушения или сохранения связности при деструктивном воздействии на отдельные компоненты или связи.

Информационную систему конечно же нельзя считать ни биологической, ни технической системой, хотя отдельные элементы этих систем необходимы для их существования. Скорее всего, информационные системы можно отнести к системам коммуникационным, на формирование которых существенное влияние оказывает так назы-

ваемый «человеческий фактор», который сложнее всего поддается формализации.

К основным целям функционирования информационных систем относится информирование (например, информирование о важнейших аспектах определенного процесса или события). При этом отдельные информационные системы, взаимодействуя между собой, являются элементами информационного пространства.

Известно, что системы в общем случае разделяются на целенаправленные и нецеленаправленные. Информационные системы могут относиться как к первому, так и ко второму классу. Живучесть искусственно формируемых информационных систем имеет решающее значение, например, в ходе рекламных кампаний и других информационных операций [3].

В данной книге рассматриваются теоретические вопросы живучести информационных систем, моделирования деструктивных воздействий на них, восстановления информационных сетевых структур после нежелательных воздействий. При этом рассматриваются элементы теорий надежности, нелинейной динамики, перколяции, сложных сетей, фазовых переходов и т.п., формулируются актуальные для функционирования информационных систем задачи.

Книга ориентирована на достаточно широкий круг читателей: специалистов в области информационных технологий и информационной безопасности, студентов, аспирантов; хочется верить, что она будет также полезна и аналитикам, которые при решении задач в различных областях хотят учитывать особенности современного сетевого информационного пространства. Надеемся, что книга окажется полезной при подготовке учебных курсов по теоретическим и практическим вопросам живучести информационных систем.

Авторы выражают искреннюю благодарность своим коллегам и соавторам по работам, фрагменты которых обобщены в этой книге: А.А. Снарскому, М.Г. Кузнецовой, Е.С. Горбачик, С.М. Брайчевскому, В.Н. Фурашеву, В.Г. Путятину, а также рецензентам А.Е. Литвиненко и В.В. Мохору за конструктивные обсуждения и замечания.

Александр Додонов,
Дмитрий Ландэ
14 декабря 2010 г.

ВВЕДЕНИЕ

В традиционном понимании информационная система включает инфраструктурные элементы, такие как компьютеры, линии связи и т.п., а также информационную составляющую (библиотеки, СМИ, телевидение, радио, рекламные щиты и многое другое), т.е. представляет собой некоторое содержательно цельное подмножество информационного пространства.

В узких рамках данной работы, в отличие от приведенного выше традиционного подхода, информационная система рассматривается как некоторая смысловая сущность, совокупность информационных блоков, информационный кластер, признаки группировки элементов которого не всегда известны заранее. При этом информационные системы возникают и развиваются в информационном пространстве.

Под информационным пространством (*от англ. — Information space*) будем понимать совокупность элементов, образующих некоторое информационное единство. В данном случае элементами информационного пространства могут быть самые разнообразные информационные ресурсы: документы, обобщающие самые различные виды информации — файлы, электронные письма, веб-страницы, независимо от видов и форматов их представления. Эти информационные ресурсы, в свою очередь, по смысловым признакам могут группироваться в информационные системы.

Приведенное определение информационного пространства является качественным. Конечно же, термин «пространство» в данном случае не совпадает с понятием «пространство» в математике или физике. В рамках данной работы пространство понимается как своеобразный общий контейнер для рассматриваемых информационных объектов. В качестве примеров удачных теоретических моделей информационного пространства можно привести «векторно-пространственную» модель Г. Солтона или модель старения информации Бартона–Кеблер. Модель веб-пространства была построена в конце XX века А. Бредером [4].

Информационные системы в нашем понимании можно трактовать как контентные системы (англ. *content* — содержание), которые, как известно [5], полностью удовлетворяют общему определению систем. Как частные случаи информационных систем можно рассматривать, например, тематические подборки документов, веб-сайты, посвя-

щенные некоторой проблематике, или информационные кластеры (сюжеты) — массивы информационных сообщений, опубликованных на различных веб-сайтах, посвященных одной тематике или одному событию [5]. При таком подходе информационные системы фактически охватывают множества элементарных единиц контента, которые в узких рамках данной работы будем называть «документами» не различая понятий «документ», «сообщение», «публикация» и т.п.

«Среда обитания» информационных систем сегодня вполне репрезентативно представлена веб-пространством, что, однако, не должно ограничивать круг вопросов рассмотрением только этой сети.

Эта книга посвящена такому свойству информационных систем, как живучесть, т.е. возможность реализации цели функционирования в случаях неблагоприятных воздействий (НВ). При неблагоприятных воздействиях на информационную систему может оказаться недостижимой некоторая часть данных, отсутствовать возможность восстановления доступа к этим данным. Механизмы обеспечения живучести при этом могут быть самыми разнообразными. Например, может порождаться новый информационный процесс, реализующий цель функционирования информационной системы. Такая возможность будет характеризовать информационную систему как живучую. В основу технологий подобных динамических функциональных перестроек систем закладываются механизмы адаптации, компенсации, распознавания, реконструкции, реконфигурации и реорганизации. Эти механизмы используются для поддержки доступности, целостности и конфиденциальности информации на уровне, необходимом для реализации с заданным качеством цели функционирования, т.е. живучесть — необходимое качество систем, ориентированных на технологии аналитической деятельности, важнейшим условием для достижения целей функционирования которых является обеспечение точной и наиболее полной информацией пользователей в любое время и там, где эта информация необходима.

Очевидно, что живучесть информационной системы может, с одной стороны, рассматриваться как ее объективное свойство, которое зависит от тематики, аудитории, времени, а с другой — как характеристика, которую хотят придать ей в случае искусственного формирования, например в ходе проведения информационных операций.

Свойство живучести позволяет сложной системе сохраняться целостной в экстремальных для нее условиях, приспособиться к ним,

изменяя поведение, структуру, зачастую и цель функционирования [6]. Живучесть трудно заметить в нормальных условиях функционирования. Это свойство рельефно проявляется только в случаях возникновения нарушений в структуре системы, отказе ее составляющих, функциональных нарушений, целенаправленных деструктивных влияний. В зависимости от класса систем, их сложности, степени организованности, а также от выбранного уровня анализа свойство живучести может оцениваться как устойчивость, надежность, адаптивность, отказоустойчивость.

Устойчивость системы — это ее способность возвращаться в начальное состояние после окончания влияния, которое вывело систему из этого состояния, это активное сохранение системой определенных характеристик безотносительно к тому, играют ли они какую-либо роль в общей системе.

Надежность является комплексным свойством системы, которое состоит в ее способности выполнять (в определенных условиях функционирования) заданные функции, сохраняя свои основные характеристики в определенных границах. Самым распространенным показателем надежности, которые обычно вероятностны по своему характеру, — это вероятности безотказной работы, средняя наработка на отказ, готовность и т.д.

Отказоустойчивость — свойство системы сохранять работоспособность в случае отказ одной или нескольких компонент.

Адаптивность — свойство системы приспосабливаться к изменяемым условиям внутренней и внешней среды путем использования различных механизмов приспособления.

Определенные свойства характеризуют систему с точки зрения возможностей сохранения ее функциональности в случае наличия изменений во внутренней и внешней среде. Если надежность характеризует функциональность системы в четко определенных условиях, стойкость — в условиях незначительных отклонений, отказоустойчивость — в условиях наличия и накопления различных отказов, то живучесть — в любых, в том числе непредусматриваемых, условиях, в частности при умышленных воздействиях деструктивного характера.

Данная книга состоит из шести основных разделов, в первом из которых приведено определение информационной системы, рассматриваются ее свойства, связанные с целями и функциями. Первый раздел также посвящен моделированию информационного пространства

и информационных систем, одному из самых популярных подходов — многоагентному моделированию. Рассматриваются основные модели в рамках данного подхода.

Второй раздел посвящен вопросам анализа живучести, в нем рассматриваются свойства функциональной и структурной живучести, приводятся различные подходы к оценке живучести, в частности теоретико-игровой, энтропийный, базирующийся на логико-вероятностных моделях и т.д.

Современные информационные системы представляют собой сетевые структуры из содержательно связанных компонент — элементов информационного пространства. Третий раздел посвящен современному направлению — теории сложных сетей, в рамках которых оцениваются такие важные с точки зрения живучести параметры сети, как кластерность, посредничество, уязвимость, рассматриваются модели «малых миров», свойства безмасштабных сетей, сетей со слабыми связями.

Вопросы живучести многих сетей связаны с таким свойством, как «протекание», изучаемое в рамках теории перколяции. Порог протекания во многих случаях можно трактовать как «порог живучести» информационной системы. Четвертый раздел посвящен вопросам критического уровня живучести информационных сетевых структур, фазовым переходам, моделированию атак на сети.

В пятом разделе рассматривается широкий круг вопросов, связанных с информационными потоками с точки зрения теории живучести. Дается формальное определение, рассматриваются различные модели информационных потоков, подробно обсуждаются результаты нелинейного моделирования, а также использования современных методов анализа временных рядов, соответствующих информационным потокам.

Шестой раздел посвящена вопросам живучести информационных операций, рассматриваемых как социальные процедуры, инструментом которых выступают информационные системы. Подробно описывается этапность информационных операций, особенности их планирования и моделирования, а также методы мониторинга и анализа. При этом приводится большое количество практических примеров.

Хочется подчеркнуть, что наблюдаемый в настоящее время процесс в области интеллектуализации автоматизированных систем,

перехода от простой обработки данных к процессам поддержки принятия решений потребовал новых подходов. Кроме того, изначальная парадигма информационных систем, сформированная несколько десятилетий тому назад, уже не отвечает реальной ситуации — объемам и динамике информационных потоков, сетевой топологии. Необходим поиск новых принципов, в рамках которых оказалось бы возможным проектирование качественно новых систем обработки больших и динамичных массивов информации — информационных систем. Особое место среди задач, получивших актуальность при этом, занимают задачи, связанные с обеспечением живучести информационных систем.

Цель этой книги — систематически изложить состояние существующих теоретических и технологических возможностей, представить возможные перспективы развития, дать импульс новым идеям в области живучести информационных систем — содержательно связанных кластеров в информационном пространстве.

1. ИНФОРМАЦИОННЫЕ СИСТЕМЫ

1.1. Свойства информационных систем

Для того чтобы сопоставить свойства информационных систем со свойствами других систем, обратимся к классическому определению, в соответствии с которым система — это совокупность объектов и связей между ними, выделенных из среды на определенное время и с определенной целью. Система в широком смысле рассматривается как динамически изменяемая совокупность сильно связанных объектов, обладающая свойствами организации, связности, целостности и членимости.

Соответственно, информационную систему в нашем понимании можно трактовать как контентную систему, совокупность документов, связанных взаимными контекстными ссылками, гиперссылками, цитированием, общей лексикой, фактографией и т.д., выделенных из среды (информационного пространства) на определенное время (время актуальности) с определенной целью или по определенному поводу. Т.е. действительно, информационная система — это совокупность сильно связанных объектов, обладающая свойствами организации, связности, целостности (определяемой тематикой или событием) и членимости (на отдельные документы, их фрагменты).

Рассмотрим свойства систем, связанные с целями и функциями.

1. Синергетичность — однонаправленность действий компонентов усиливает эффективность функционирования системы. В случае информационных систем, в частности, информационных сюжетов, направленность отдельных документов усиливает информационную функцию всей информационной системы.

2. Приоритет интересов системы перед интересами ее компонент (общую тематическую тенденцию определяет вся информационная система, а не отдельные документы как компоненты).

3. Эмерджентность — цели (информационные функции) компонент (отдельных документов) системы не всегда совпадают с целями (функциями) всей информационной системы.

4. Мультипликативность — и позитивные, и негативные эффекты функционирования компонент системы обладают свойством умножения, а не сложения (аналогии — количество информации в документах, информационная энтропия).

5. Целенаправленность информационных систем в случае их искусственного формирования (вместе с тем, существуют и нецеленаправленные системы, в том числе и информационные сюжеты).

6. Альтернативность путей функционирования и развития. Важнейшие документы могут быть актуальными на протяжении длительного времени или отдельные документы по одной тематике, генерируемые в большом количестве, но имеющие небольшой срок актуальности.

Связанные со структурой свойства информационных систем следующие:

1. Целостность — первичность целой информационной системы по отношению к отдельным ее элементам (в общем случае — документам).

2. Неаддитивность — принципиальная несводимость свойств информационной системы к сумме свойств составляющих ее элементов (конечно, если информационная система состоит более чем из одного документа).

3. Структурность — возможна декомпозиция информационной системы на компоненты (документы), установление связей между ними.

4. Иерархичность — компоненты системы (информационные сообщения, документы, пожалуй, кроме элементарных одноаспектных) могут также рассматриваться как подсистемы информационной системы.

Информационные системы, как и отдельные документы, являются частями информационного пространства, и им, соответственно, присущи такие свойства, связанные с внешней средой:

1. Коммуникативность — существование сложной системы коммуникаций информационных системы с информационным пространством, в частности, отдельные документы из информационной системы могут быть связаны не только с другими документами из той же информационной системы, но и с другими частями информационного пространства.

2. Взаимодействие и взаимозависимость информационной системы и информационного пространства.

3. Адаптивность — стремление к состоянию устойчивого равновесия, которое предполагает адаптацию параметров информационной системы на определенных этапах ее жизненного цикла к изменяющимся параметрам внешней среды.

4. Надежность — существование информационной системы при выходе из строя отдельных ее компонент (информационных ресурсов), сохраняемость значений параметров системы в течение определенного периода.

5. Интерактивность — взаимодействие с внешней средой и «ответная» изменяемость информационных систем.

Существует еще ряд системных свойств информационных систем, таких как:

1. Интегративность — наличие системообразующих, системо-сохраняющих факторов.

2. Эквивинальность — способность информационных систем достигать состояний, не зависящих от исходных условий и определяющихся только параметрами системы.

3. Наследственность.

4. Возможность развития.

5. Самоорганизация и т.д.

Для полноценной работы или сохранения минимального набора критически важных функций информирования информационная система должна обладать вполне определенным запасом устойчивости к внешним дестабилизирующим воздействиям из внешней среды (информационного пространства), обусловленным, в свою очередь, влияниями со стороны общества, государства, коммерческих структур и т.д. Как на всю информационную систему, так и на ее отдельные элементы могут оказываться различные дестабилизирующие информационные воздействия, атаки, например, удаление отдельных материалов с веб-сайтов сети Интернет, уничтожение или отключение информационных серверов, публикация документов, которые в определенном направлении искажают исходную информационную систему, или порождение новой информационной системы, которая может снизить актуальность или попросту уничтожить исходную систему.

Понятно, что для полноценной работы и сохранения минимального набора критически важных функций информирования информационная систем должна обладать вполне определенным запасом устойчивости к внешним дестабилизирующим воздействиям. При этом нарушение целостности информационной системы на фоне снижения актуальности ее компонент влечет за собой дезорганизацию, одновременную потерю гибкости — понижение живучести и нарушение целостности, то есть потерю важнейших функций информационных систем (рис. 1) [7].

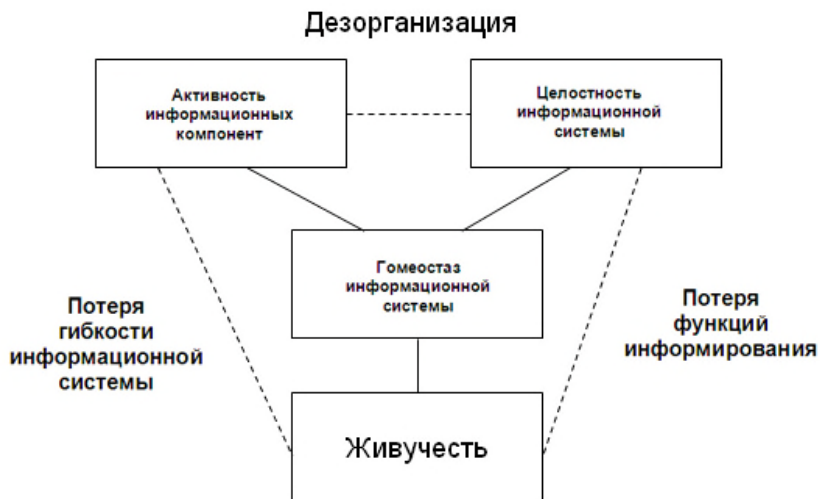


Рис. 1. Модель гомеостаза информационной системы

Задачи живучести информационных систем требуют не только экспертных, качественных решений, но и применения подходов, базирующихся на математическом и компьютерном моделировании. Например, можно выделить задачи моделирования информационных операций, которые неразрывно связаны с так называемым социальным моделированием (англ. — *Social Modeling*).

Несмотря на то, что социальное моделирование является междисциплинарным направлением, именно оно находит в настоящее время широкое применение при решении задач в области информационной безопасности [3].

1.2. Моделирование информационного пространства

Информационное пространство — множество связанных по смыслу элементов (документов), образующих информационные системы — кластеры близких по тематике документов. При этом оно за все время существования сохраняет свои устойчивые закономерности. Многочисленными исследованиями показано, что параметры частотного и рангового распределений документов во многих информационных системах остаются одинаковыми и определяются параметрами,

зависящими от содержания, тематики информации. В связи с этим С.А. Иванов [8] заметил, что «информационное пространство — это документальная среда, в которой формируются кластерные структуры научных публикаций в периодических изданиях, являющиеся фракталами». Информационные системы отражают в информационном пространстве коммуникационные процессы в своей тематической области, появление новых тематик сопровождается возникновением новых фрактальных массивов в информационном пространстве.

Как и многие другие сложные системы, информационное пространство можно представить как коммуникационную среду — в виде системы с комплексом связей информационных источников и преобразователей между собой, влияющих друг на друга в зависимости от уровня восприятия генерируемых и преобразуемых ими отдельных информационных сообщений.

При этом для моделирования источников и преобразователей информации, с одной стороны, вполне подходит классическая теория информации, т.е. математическая теория связи, разработанная Шенноном в 40-х годах XX столетия и существенно дополненная и расширенная в последующие годы работами Н. Винера, В.А. Котельникова и А.Н. Колмогорова. Однако классическая теория информации не учитывает взаимодействия между источниками и преобразователями информации, что, с другой стороны, вполне укладывается в идеологию современной теории сложных систем и реализуется, например, с помощью многоагентного моделирования.

Идея многоагентного (англ. — *Agent-based*) моделирования возникла в середине XX века. В соответствии с ней агент — это некоторая абстрактная сущность, которая обладает активностью, автономным поведением, может принимать решения в соответствии с некоторым набором правил, может взаимодействовать с окружением и другими агентами, а также может сама эволюционировать.

Моделирование на базе применения агентов (или многоагентное моделирование) — это ориентированная на использование компьютеров методология, которая позволяет исследователям проектировать, анализировать и исследовать формальные модели, представленные как некоторые «искусственные миры», населенные агентами, которые взаимодействуют друг с другом на основании наперед заданных относительно простых правил. Многоагентные модели значительно варьируются по своим правилам, однако все они состоят из множеств авто-

номных, близоруких (воспринимающих информацию лишь из некоторой локальной области) единиц. Хотя индивидуальные действия агентов полностью определены на микроуровне, в то время, когда функционирует большое количество таких агентов, их взаимодействия формируют макроуровень, который не всегда предсказуем.

Вместе с тем, сегодня многоагентные модели широко используются для анализа децентрализованных систем, закономерности динамики функционирования которых не изучены в достаточной мере. Многоагентные модели используются для того, чтобы получить представление об общем поведении таких систем, выявить правила функционирования систем, исходя из предположений об индивидуальном поведении их отдельных компонент (агентов).

Цель многоагентного моделирования может быть сформулирована как создание компьютерных микромиров, в которых агенты взаимодействуют, реагируя на условия из своего окружения и осуществляя изменения.

В рамках простейшей модели информационного пространства информационный блок (документ) может генерироваться агентом A_1 , передаваться некоторому множеству агентов A_2, \dots, A_n , которые преобразуют получаемую информацию, искажая ее (случайно или умышленно, в зависимости от информационного влияния других агентов), и передают ее дальше в виде собственных информационных блоков. При этом сходные по информационному содержанию блоки могут группироваться в кластеры — информационные системы. Ниже будет подробно рассмотрена одна из простейших многоагентных моделей, реализующих распространение новостей в информационном пространстве — модель диффузии информации.

В соответствии с определением К. Лангтона [9] моделирование сложных адаптивных многоагентных систем часто основывается на следующих принципах:

- модель состоит из популяции простых агентов;
- не существует единого агента (центра), направляющего остальных агентов;
- каждый агент подробно рассматривает способы, которыми осуществляется простая реакция на локальные изменения в окружении, включая контакты с другими агентами;
- не существует единого правила в системе, которое описывало бы глобальное поведение.

В соответствии с этими принципами любое поведение на уровне выше индивидуального является эмерджентным, порожденным взаимодействиями локальных агентов, т.е. «простые правила могут вызывать сложное поведение и структуры».

1.3. Многоагентные системы

Многоагентные модели по определению являются децентрализованными. При этом сложное глобальное поведение системы является результатом деятельности большого количества агентов, каждый из которых следует простым правилам. Многоагентные модели позволяют исследовать достаточно широкий круг проблем, для которых строгие аналитические методы оказываются неэффективными.

1.3.1. Искусственные общества

Во многих случаях многоагентные модели реализуются в виде так называемых «искусственных обществ» (англ. — *Artificial Societies*) как компьютерных моделей общества реального. Таким образом, искусственное общество является подклассом многоагентных моделей. Агенты искусственного общества ведут себя автономно: принимают решения, действуют самостоятельно и взаимодействуют с другими агентами. В литературе описаны такие области применения агентских моделей, как, например, выбор юрисдикции для проживания, голосование за политические партии, экономическую политику, создание или модификацию юрисдикций, принятие решений об участиях в сообществах и т.д.

Дж. Эпштейн [10] выделил следующие основные характеристики многоагентных моделей:

— неоднородность. Агенты чем-то отличаются друг от друга, что принципиально отличает эти модели от большинства индивидуум-ориентированных моделей;

— автономия — агенты действуют независимо друг от друга;

— пространство, в котором действуют агенты, задано или описано заранее в явном виде;

— наличие локальных взаимодействий;

— ограниченная рациональность, присущая агентам;

— наличие динамики системы, кроме случаев достижения равновесия.

Благодаря взаимодействию между агентами могут моделироваться социальные процессы и процедуры, именно поэтому искусственное общество может рассматриваться как определенный класс ориентированных моделей.

1.3.2. Модель «тепловых жуков»

Моделирование агентов и многоагентных систем без использования современных средств моделирования (например, AnyLogic) является достаточно несложной процедурой. Так, в [11, 12] в качестве примера рассматривается простая модель коллективного поведения, названная «тепловыми жуками» (англ. — *Heat Bugs*) [13]. В дискретной среде, разбитой на клетки, двигаются «жуки», выделяющие тепло, которое распространяется в среде.

Для каждого отдельного жука существует собственная «идеальная» температура среды, в которой он предпочитает находиться, и он имеет сенсор, с помощью которого может определить, в каком направлении температура среды ближе к его «идеальной» температуре. Это позволяет жуку найти то направление, в котором он должен двигаться, чтобы достигнуть клетки с устраивающей его температурой. Среда имеет следующие характеристики:

- тепло распространяется равномерно во всех направлениях со скоростью, пропорциональной разнице температур в соседних клетках;
- тепло «уменьшается» в каждой клетке пропорционально количеству тепла, которым обладает клетка.

«Мир жуков» является динамическим и трудно предсказуемым. Например, даже находясь в клетке с «идеальной» для него температурой, жук нагревает ее, выделяя тепло, поэтому со временем он может направиться в более комфортную для него клетку, поскольку состояние среды изменяется.

Анимация модели «мира тепловых жуков» осуществляется путем обозначения температуры в каждой клетке определенным цветом. В результате моделирования видно, что жуки собираются в группы, согревая друг друга. Модель позволяет проследить поведение каждого конкретного жука, определить его параметры, координаты, идеальную температуру. Эксперименты с этой моделью показывают, как изменения локальных параметров влияют на глобальное поведение и живучесть всей системы.

1.3.3. Природные вычисления

В последнее время в рамках концепции многоагентного моделирования интенсивно разрабатываются так называемые «природные вычисления» (англ. — *Natural Computing*), объединяющее математические методы, в которых заложены принципы принятия решений, аналогичные реализуемым в природе механизмам. Имитация самоорганизации муравьиной колонии (или колонии термитов) составляет основу так называемых муравьиных алгоритмов оптимизации [14], одного из перспективных методов природных вычислений, в которых колония муравьев рассматривается как многоагентная система, где каждый агент функционирует автономно по достаточно простым правилам. В противовес примитивному поведению агентов, поведение всей системы оказывается очень сложным, приближенным к разумному. Муравьиные алгоритмы основаны на имитации самоорганизации социальных насекомых посредством использования динамических механизмов, обеспечивающих достижение системой глобальной цели в результате низкоуровневого взаимодействия ее элементов при условии использования элементами системы только локальной информации, исключая централизованное управление. В настоящее время получены хорошие результаты муравьиной оптимизации таких сложных комбинаторных задач, как задачи коммивояжера, оптимизации маршрутов транспорта, раскраски графа, оптимизации сетевых графиков, календарного планирования, оптимизации процессов в распределенных нестационарных системах, например трафиков в телекоммуникационных сетях.

1.3.4. Модель предпочтений групп людей

В качестве иллюстраций применения мультиагентных систем рассмотрим еще одну модель — Аксельрода и Хаммонда [15]. В соответствии с этой моделью исследовались предпочтения групп людей. При этом сначала предполагалось, что группы различаются только по этнической принадлежности. Однако, построенная модель может также учитывать любые другие типы различия, в которых индивидуальное членство в группе является видимым и устойчивым.

В модели Аксельрода–Хаммонда агент — это индивид. Каждый агент «окрашивается в цвет», который может интерпретироваться как его этническая принадлежность или другой признак членства в группе.

Каждый агент также имеет стратегию, состоящую из двух частей. Первая часть стратегии определяет, сотрудничает ли агент (или нет) с соседом, который имеет тот же цвет. Вторая часть стратегии агента определяет, сотрудничает ли агент с соседом, который имеет отличный от него цвет. Как и во всех многоагентных моделях, вначале устанавливаются правила взаимодействия агента, а затем используется компьютерное моделирование для того, чтобы проследить историю эволюции. Изначальная цель проекта состояла в том, чтобы понять условия, при которых население в результате поставит у власти таких людей, которые будут заботиться лишь о членах собственной группы и отказываться предоставлять помощь членам других групп.

Так как цель модели была эвристической, главный критерий проекта был простым. Описание модели очень простое. Рассматривается система клеточных автоматов — поле 50×50 клеток, каждая из которых имеет 4 соседа — окрестность фон Неймана. В этой модели агенты могут также интерпретироваться как небольшие группы семейств типа деревень с одной этнической принадлежностью. Рассматривается 2000 шагов итерации, в течение которых могут выполняться следующие действия:

1. Иммиграция. Иммигрирующий агент со случайными генами появляется на случайном пустом участке.

Агент имеет три гена:

— признак: один из четырех цветов;

— выбор при встрече агента с тем же цветом: помощь или нет;

— выбор при встрече агента с другим цветом: помощь или нет.

2. Взаимодействие. Каждый агент начинает период с потенциалом воспроизведения (PTR) = 12 %. Каждая смежная пара агентов решает «одношаговую дилемму заключенного»:

— если агент оказывает помощь: PTR уменьшается 1 %;

— если агент получает помощь: PTR увеличивается 3 %.

3. Воспроизводство. В случайном порядке каждый агент с вероятностью PTR выполняет воспроизводство в смежную пустую ячейку, если она доступна, с мутацией/геном = 0,5 %

4. Смерть. Каждый агент имеет 10-процентный шанс умереть.

Один из результатов, полученный по итогам анализа работы модели, показал, что чаще всего способность агентов различать круг лиц с общими интересами и группами ведет к более высокому уровню сотрудничества чем если бы агенты «страдали дальтонизмом».

При применении модели для решения реальных проблем Аксельрод и Хаммонд инициализировали ее пространство, учитывая географические и исторические данные Средней Азии. Для этого использовалась карта распределения этнических групп, которые «привязывались» к карте модели. Модель использовалась в тех местах, где ожидалась этнические конфликты. Так как модель очень проста, не ожидалась высокая степень точности в ее «прогнозах», однако она представляет определенный интерес при идентификации некоторых потенциальных зон конфликтов.

Известна адаптация приведенной модели для исследования определенных аспектов современности, например, динамики роста количества средств массовой информации. Это могло быть сделано путем добавления к модели механизма, позволяющего учитывать влияние на агента и его соседей СМИ различного радиуса действия. С помощью такой модели можно выяснить, станут ли провинциальные СМИ достаточно сильными, чтобы бросить вызов господству общенациональных СМИ.

С помощью предложенной модели можно было также исследовать потенциальные эффекты множественных и кросскультурных связей типа богатых/бедных, городских/сельских, христианских/мусульманских различий.

1.4. Имитационные модели

Имитационные модели, в том числе и многоагентные, являются удобными для постановки вопросов типа «что если?» [16]. Фактически, благодаря тому, что такие модели, как правило, содержат вероятностный элемент, управляя моделированием много раз, часто можно получить полное распределение возникающих ситуаций.

В последнее время в связи с бурным развитием компьютерных технологий важным и перспективным с точки зрения применения на практике математического моделирования является класс так называемых имитационных моделей. Такая модель представляет собой алгоритм, при помощи которого компьютер генерирует наборы данных, описывающих заданные характеристики реальной системы, представляющей интерес. При этом выполняемые машиной операции не имеют никакого отношения к природе и свойствам изучаемой системы. Отметим, что сам по себе факт выяснения возможности имитационного моделирования является немалым достижением современной нау-

ки. Действительно, оказывается, что структура реального процесса в известной мере не зависит от его природы и, так сказать, материальной основы. Числа, получаемые в результате манипулирования другими числами по определенным абстрактным правилам, могут в точности соответствовать числам, описывающим конкретные процессы, происходящие в реальном мире.

Разумеется, при разработке имитационной модели принимаются в расчет свойства исследуемого явления, но на уровне не внутренних механизмов, которые либо не известны, либо слишком сложны для явного использования, а общих характеристик протекания соответствующих процессов.

В плане практического применения имитационные модели хороши тем, что позволяют проводить так называемые машинные эксперименты, целью которых является изучение изменения поведения объекта исследования в зависимости от изменений внутренних параметров или (и) внешних условий. Такие методики дают возможность определять ход развития событий, которые по тем или иным причинам невозможно реализовать в реальной жизни. Например, для анализа распределения потоков беженцев в случае затопления Калифорнии не следует осуществлять подобную катастрофу в действительности. Имитационное моделирование (при наличии удовлетворительных моделей) в подобных случаях позволяет получить данные на вполне приемлемом уровне точности.

Построение имитационных моделей представляет собой достаточно сложную задачу, требующую, помимо знания предметной области, еще и высокого профессионализма в сфере программирования. Однако в случае успеха результаты окупают издержки.

1.5. Индивидуум-ориентированные модели

Одна из разновидностей многоагентного моделирования — индивидуум-ориентированное моделирование (англ. — *Individual-Based Modeling*) является относительно молодым направлением. Первые работы в этой области появились в 80-х годах XX века. Индивидуум-ориентированное моделирование можно рассматривать, с одной стороны, как разновидность многоагентного моделирования, а с другой стороны, как логическое развитие количественной экологии популяций, начало которому было положено в работах Дж. Форрестера (в частности, его моделью «Мировая динамика» [17]), который впервые

построил и рассчитал модели большой размерности с помощью имитационного моделирования с использованием компьютеров.

Индивидуум-ориентированные модели — это многоагентные имитационные модели, в которых интегральные характеристики популяции являются результатом множества локальных взаимодействий агентов (индивидуумов). Модель состоит из описания среды, в которой происходят взаимодействия, и набора индивидуумов, для которых определены правила поведения и характеристические параметры, которые также могут изменяться в ходе эволюции. Индивидуум-ориентированный подход в моделировании предполагает создание имитационных моделей, учитывающих некоторые свойства отдельных индивидуумов и их локального взаимодействия для построения интегральных моделей целых популяций, сформированных из множества индивидуумов. Индивидуум в рамках этих моделей рассматривается как уникальная, дискретная единица, у которой есть некоторый набор характеристик, изменяющихся в течение жизненного цикла. Модели, основанные на данном подходе, строят снизу вверх, начиная с «частей» системы (индивидуумов), описывая в итоге всю популяцию. Целью исследования часто становится понимание того, каким образом свойства системы возникают из взаимодействия между частями [18].

Построение модели на уровне описания отдельного индивидуума обуславливает ряд преимуществ, таких, как прозрачность по отношению к объективным природным механизмам, возможность описывать исследуемый объект с большой степенью детализации, извлекать больше полезной информации из результатов моделирования.

Индивидуум-ориентированное моделирование — это один из разделов теории сложных систем, имеющий целью обеспечить исследователей инструментами моделирования для решения задач, не поддающихся рассмотрению традиционными методами.

Первые индивидуум-ориентированные модели имели отношение исключительно к конкретным задачам из областей биологии и экологии и не носили системного характера. Как будет показано ниже, они вполне логично переносятся на область изучения социальных процессов и процедур.

Безусловно, построение интегральной модели системы на основе приблизительного описания правил поведения отдельного индивидуума может оказаться весьма далеким от реальности, однако в данном случае многое зависит от уровня описания этих правил, свойств

отдельных индивидуумов и предполагаемой динамики популяции. Индивидуум-ориентированное моделирование предоставляет ряд таких преимуществ, как простота описания отдельных особей и их локального взаимодействия, возможность детализации этих описаний в ходе моделирования, а также прозрачность обратной связи «правила – модель – реальность».

Одной из наиболее популярных моделей в рамках индивидуум-ориентированного моделирования является так называемая «сахарная модель» искусственного общества Дж. Эпштейна и Р. Акстеля [19].

Несмотря на свою простоту, эта модель является мощным инструментом анализа информационных систем, в частности, такого свойства, как живучесть. Сегодня эта модель имеет множество модификаций, даже сами создатели нашли вариант описания этой же модели в формализме клеточных автоматов. Приведем описание оригинальной версии «сахарной модели».

Пространством в данной модели является двумерная сетка с равными клетками. В каждый момент времени t существует постоянное конечное число агентов, расположенных в пространстве. В момент времени t каждая клетка (x, y) может содержать агента a ($a_t(x, y) = a$), т.е. агент a находится в клетке (x, y) , или не содержать агента ($a_t(x, y) = \emptyset$). Количество «сахара» в клетке (x, y) в момент времени t составляет определенную величину, зависящую от x и y .

Агент появляется на поле (рождается) с двумя параметрами: зрение (количество клеток в решетке, которое он может видеть) и уровень метаболизма (количество сахара, съедаемое им в единицу времени для выживания). Агент может переносить любое количество сахара. Если у агента нет должного количества сахара для еды, то он умирает. Одновременно со «смертью» одного агента «рождается» другой агент со случайно выбранными параметрами и расположением в пространстве. Таким образом, численность всей популяции агентов — величина постоянная. Правила поведения агентов следующие:

— агент изучает окрестность своего видения (четыре или восемь направлений решетки) и определяет свободную клетку, имеющую наибольшее количество сахара;

— после этого агент перемещается в эту клетку и собирает весь сахар.

На рис. 2. представлены различные этапы эволюции сахарной модели.

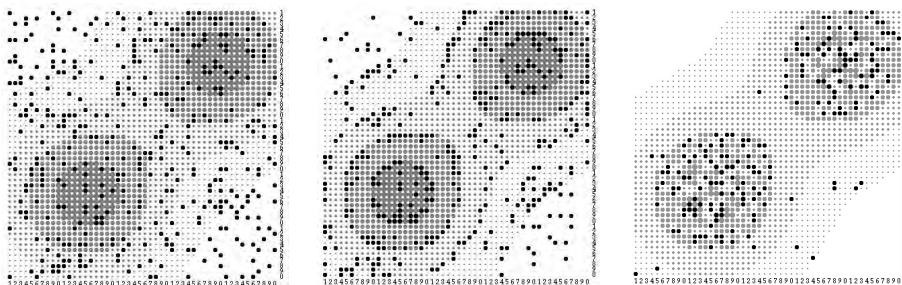


Рис. 2. Эволюция простейшей сахарной модели

На основе сахарной модели авторы получили результаты, вполне соответствующие обычному социальному поведению. Уже на простейшей версии модели без явных взаимодействий между агентами авторы исследуют вопрос о возможностях борьбы с концентрацией населения, в частности, в связи с проблемой загрязнения окружающей среды. Для этой цели в модель вводится загрязнение среды как результат сбора и потребления сахара, т.е. в этом случае каждая клетка содержит сахар и некоторый уровень загрязнения. По новым правилам агент передвигается в свободную клетку, где соотношение (сахар/уровень) загрязнения максимально. Введение этих дополнительных условий приводит к коренным изменениям в эволюции модели: увеличивается количество смертей, а распределение популяции становится более равномерным.

В более поздних модификациях сахарной модели рассматриваются разные виды взаимодействий между агентами, а также другие усложнения. Это позволяет анализировать более широкий круг социальных процессов и процедур.

В рамках «сахарной модели» исследуются следующие вопросы:

- распределение богатства между агентами;
- распределение агентов по возрасту;
- миграция агентов;
- ввод в модель новых свойств модели, например загрязнения среды, и соответствующая модификация правил;
- введение новых свойств агентов, например разделение агентов по полам;

- изменение правил рождения новых агентов;
- введение правил наследования, например, когда богатство умершего агента равномерно распределяется между его потомками;
- введение множества продуктов, например сахара и меда;
- введение правил обмена продуктами между агентами (торговли) и т.д.

Индивидуум-ориентированное моделирование дает возможность описывать следующие свойства моделируемого объекта (особенно важные для моделирования социальных процессов):

- учет пространственных аспектов;
- учет воздействия материальных средств и влияния СМИ;
- учет социальных аспектов и индивидуальных особенностей;
- учет уровня детализации.

Учет пространственных аспектов. Индивидуум-ориентированное моделирование охватывает пространственно-распределенные модели (англ. — *Spatially-explicit Models*), в которых каждый индивидуум ассоциирован с определенным положением в пространстве.

Для полноценного моделирования электоральных полей, как правило, желательна пространственная распределенность, учет перемещений в пространстве. Например, при моделировании возможных фальсификаций выборов по открепительным талонам.

Учет воздействия материальных средств и влияния СМИ. Как известно, электоральные процессы существенно зависят от материальных средств и влияния средств массовой информации, используемых во время избирательных кампаний.

Модели, учитывающие информационные потоки, называются информационными. В этих моделях описание изменения состояния индивидуумов базируется на правилах, основанных на анализе информационных потоков (в том числе, например, и состояние соседних индивидуумов).

Учет социальных аспектов и индивидуальных особенностей. Для моделирования электоральных процессов особенно важно описание всей системы на уровне отдельных особей.

Для электоральных процессов социальные механизмы играют очень важную роль, поэтому электоральные поля — это типичные социальные сети. Если модель направлена на исследование социальных механизмов, требует учета индивидуальных различий и обучения особей, необходимо выбирать данный подкласс моделей.

Как известно, индивидуальная изменчивость — основополагающий принцип эволюции. Вместе с тем учет этого фактора зачастую существенно усложняет модели, поэтому он не учитывается, в частности, в рассматриваемом ниже методе клеточных автоматов.

Уровень детализации. Свойства модели существенно зависят от ее пространственно-временного масштаба. Модели также различаются по количеству рассматриваемых индивидуумов. От масштаба задачи напрямую зависит объем вычислений. Этот факт приходится учитывать при выборе масштаба модели и ее реализации.

Следует отметить, что индивидуум-ориентированные модели требуют большего объема вычислений, чем аналитические модели. Вместе с тем для многих областей, в том числе и исследования информационных процессов и систем, разработка индивидуум-ориентированных моделей оправдана в связи с тем, что:

- данных реальных наблюдений исследуемых параметров зачастую не хватает для идентификации аналитической модели;
- необходим учет пространственных аспектов;
- необходим учет социальных механизмов популяции, индивидуальных различий особи, обучения.

1.6. Модели клеточных автоматов

Многоагентное моделирование охватывает также концепцию клеточных автоматов — математических объектов, представляющих собой дискретную динамическую систему. Впервые клеточные автоматы были введены фон Нейманом в 40-х годах XX столетия как формальный аппарат для исследования сложных распределенных систем (см. [20]).

Простые в реализации модели систем клеточных автоматов зачастую проливают свет на такие проблемы, анализ которых другими методами затруднен.

Клеточные автоматы являются полезными дискретными моделями для разделов теории динамических систем, которые изучают характерные коллективные явления.

Модели систем клеточных автоматов чаще всего реализуются с использованием планарной сетки квадратных ячеек. Различные состояния ячейки представлены на мониторе компьютера различными цветами. Эволюция системы клеточных автоматов демонстрируется как последовательность изменения состояний отдельных ячеек, кото-

рые в соответствии с определенными правилами обновляются на каждом временном такте системы клеточных автоматов. Соответствующие правила учитывают состояние каждой ячейки, а также ее окрестности. Существуют различные типы окрестностей в моделях клеточных автоматов.

Дискретность модели, точнее, возможность представить модель в дискретной форме, в настоящее время относится к существенным преимуществам, поскольку открывает широкие возможности использования компьютерных технологий. Эволюция системы клеточных автоматов может быть представлена как дискретная последовательность шагов, при этом изменения на каждом шаге определяются изменениями на предыдущем. Однако часто непрерывные по своей природе процессы вполне приемлемо аппроксимируются подходящими дискретными конструкциями, но для этого модель, как правило, должна содержать большое число дискретных элементов, а эволюция проходить за большое число тактов.

Главным достоинством клеточных автоматов является их абсолютная совместимость с алгоритмическими методами решения задач. Поэтому в последние годы, в связи с бурным развитием компьютерных технологий, они находят широкое применение в самых разнообразных областях науки и техники.

Клеточные автоматы по своей сути являются пространственно-немобильными дискретными индивидуально-ориентированными моделями. В традиционной системе клеточных автоматов все ячейки равноправны (пространство однородно), в то время как в индивидуум-ориентированной кроме описания ячеек существует понятие индивидуума, который может занимать разные ячейки (или несколько различных индивидуумов могут занимать одну ячейку). Таким образом, в клеточном автомате ячейки изменяют свое состояние синхронно, и цикл моделирования представляет собой перебор ячеек. В индивидуальных моделях цикл может состоять из перебора индивидуумов, т.е. в клеточном автомате моделирование основано на разбиении пространства на однородные участки, в индивидуум-ориентированных моделях описываются сущности, которые изменяют положение в пространстве. Конечно, ячейки в клеточном автомате могут находиться в различных состояниях, и с помощью определения сложных состояний можно моделировать наличие особей в ячейках и их перемещение между ячейками. Но это возможно лишь при существенных ограничениях.

Клеточный автомат представляет собой дискретную динамическую систему, совокупность одинаковых клеток, определенным образом соединенных между собой. Все клетки образуют сеть (решетку) клеточных автоматов. Состояние каждой клетки определяется состоянием клеток, входящих в ее локальную окрестность и называемых ближайшими соседями [21]. Окрестностью конечного автомата с номером j называется множество его ближайших соседей.

Состояние j -го клеточного автомата в момент времени $t + 1$ определяется следующим образом:

$$y_j(t+1) = F(y_j(t), O(j), t),$$

где F — некоторое правило, которое можно выразить, например, языком булевой алгебры. Во многих задачах считается, что сам элемент относится к своим ближайшим соседям, т.е. $y_j \in O(j)$. В этом случае формула упрощается: $y_j(t+1) = F(O(j), t)$. Клеточные автоматы в традиционном понимании удовлетворяют таким правилам:

- изменение значений всех клеток происходит одновременно (единица измерения — такт);
- сеть клеточных автоматов является однородной, т.е. правила изменения состояний для всех клеток одинаковы;
- на клетку могут повлиять лишь клетки из ее локальной окрестности;
- множество состояний клетки конечно.

Теоретически клеточные автоматы могут иметь любую размерность, однако чаще всего рассматривают одномерные и двумерные системы клеточных автоматов.

Рассматриваемая здесь модель является двумерной, поэтому дальнейший формализм будет относиться к этому случаю. В двумерном клеточном автомате решетка реализуется двумерным массивом. Поэтому в этом случае удобно перейти к двум индексам, что вполне корректно для конечных решеток.

В случае двумерной решетки, элементами которой являются квадраты, ближайшими соседями, входящими в окрестность элемента $y_{i,j}$, можно считать или только элементы, расположенные вверх-вниз и влево-вправо от него (так называемая окрестность фон Неймана:

$Y_{i-1,j}, Y_{i,j-1}, Y_{i,j}, Y_{i,j+1}, Y_{i+1,j}$), или добавленные к ним еще и диагональные элементы (окрестность Мура (G. Moore): $Y_{i-1,j-1}, Y_{i-1,j}, Y_{i-1,j+1}, Y_{i,j-1}, Y_{i,j}, Y_{i,j+1}, Y_{i+1,j-1}, Y_{i+1,j}, Y_{i+1,j+1}$). В модели Мура каждая клетка имеет восемь соседей, в расширенной модели — двадцать четыре (рис. 3).

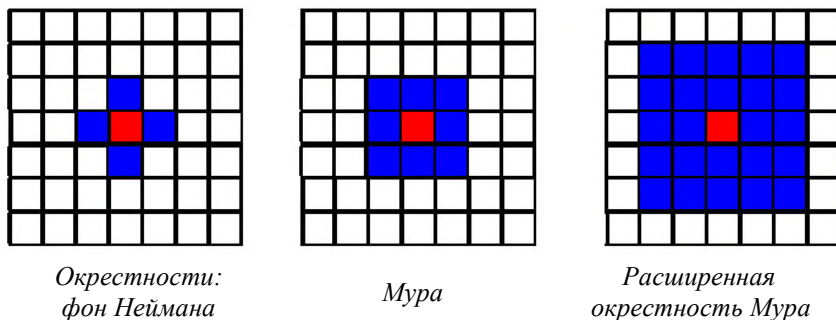


Рис. 3. Окрестности двумерных клеточных автоматов

Для устранения краевых эффектов решетка топологически «сворачивается в тор» (рис. 4), т.е. первая строка считается продолжением последней, а последняя — предшествующей первой. То же самое относится и к столбцам.

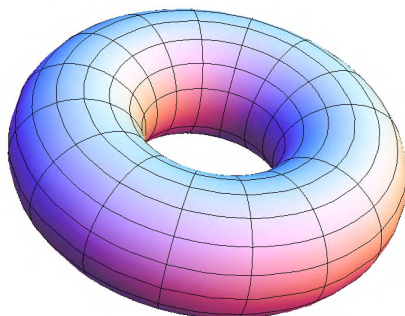


Рис. 4. Прямоугольная область, свернутая в тор

Это позволяет определять общее соотношение значения клетки на шаге $t + 1$ по сравнению с шагом t [22–24]:

$$y_{i,j}(t+1) = F(y_{i-1,j-1}(t), y_{i-1,j}(t), y_{i-1,j+1}(t), \\ y_{i,j-1}(t), y_{i,j}(t), y_{i,j+1}(t), y_{i+1,j-1}(t), y_{i+1,j}(t), y_{i+1,j+1}(t)).$$

С. Вольфрам [21], классифицируя различные клеточные автоматы, выделил те, динамика которых существенно зависит от начального состояния. Подбирая различные начальные состояния, можно получать самые разнообразные конфигурации и типы поведения. Именно к таким системам относится классический пример — игра «Жизнь», изобретенная Дж. Конвеем и известная широкому кругу читателей благодаря публикации в книге М. Гарднера [25].

Среди клеточных автоматов известные такие, динамика которых существенным образом зависит от начального состояния. Подбирая различные начальные состояния, можно получать разнообразные конфигурации и типы обращения. Именно к таким системам относится классический пример — игра «Жизнь», изобретенная Дж. Конвеем.

Правила варианта игры «Жизнь» такие. Клетка находится в одном из двух состояний — живом и неживом (черном и белом). Если в окрестности клетки меньше двух или больше трех черных клеток, то на следующем шаге она закрашивается в белый цвет (умирает). Если соседями клетки являются ровно три черных клетки, то на следующем шаге она закрашивается в черный цвет (рождается).

Поведение систем клеточных автоматов в общем случае может быть сведено к динамике сложных нелинейных систем [26], однако это весьма сложная научная задача для каждой конкретной системы клеточных автоматов.

В последнее время проводятся многочисленные попытки типизации систем клеточных автоматов. Приведем один из наиболее удачных.

К системам клеточных автоматов первого типа относятся такие системы, которые независимо от начального состояния за конечное число шагов выполняют переход к однородному состоянию — все автоматы оказываются в состоянии покоя.

В процессе эволюции клеточных автоматов второго типа система приходит к локализованным стационарным или периодическим решениям.

Результаты активности системы клеточных автоматов третьего типа являются аperiodическими, т.е. эти системы демонстрируют хаотическое поведение.

Динамика систем клеточных автоматов четвертого типа существенно зависит от начального состояния. Подбирая различные начальные состояния, можно получать самые разнообразные конфигурации.

В литературе имеются многочисленные примеры применения моделей клеточных автоматов для решения прикладных задач при анализе и моделировании социальных процессов и процедур. В работах Т. Брауна [27] рассматривается ряд контекстуальных моделей электорального процесса, в которых предполагается, что избирательные предпочтения индивидуума определяются установками его ближайшего окружения.

В одной из «бинарных» моделей предполагается, что индивид принимает решение голосовать в момент $t + 1$ за одну из двух альтернатив: за республиканцев или демократов в соответствии с правилами простого большинства своего окружения — ближайших соседей. Если из девяти человек пятеро или больше поддерживают демократов, то индивид также голосует за демократов. Если большинство составляют республиканцы, то индивид также разделяет точку зрения большинства. В данном случае клеточный автомат имеет два состояния: 1 — голосование за республиканцев; 0 — голосование за демократов.

Некоторые примеры клеточных автоматов, применяемых в задачах социологии, приведены в [23, 24]. В частности, описывается модель процесса расовой сегрегации при выборе места жительства [28].

В рассматриваемом примере предполагается, что каждая расовая группа предпочитает иметь определенный процент соседей с тем же цветом кожи. Если это условие не выполняется, то семья перебирается в ближайший дом, где процентный состав соседей является приемлемым. В [28] использовалась модель конечных автоматов с простыми правилами и окрестностью Мура. Построенная модель вполне реалистично описала процесс разделения региона на несколько расово-однородных областей.

Подобная модель была предложена в [29], в рамках которой реализуются такие правила взаимодействия существующих двух типов агентов. Агент предпочитает жить в окружении себе подобных, т.е. перемещается в область, где подобных ему агентов больше. В результате формируется структура, похожая на индийские касты.

Клеточные автоматы с успехом применяются и при моделировании процессов распространения новостей, инноваций [30].

В статье Т. Брауна [27] рассматривается модель электорального

процесса. Он считает (с чем вполне солидарны авторы), что избирательные предпочтения индивида определяются установками его ближайшего окружения. В одной из моделей предполагается, что индивид принимает решение голосовать в момент $t + 1$ за республиканцев или демократов в соответствии с правилом простого большинства. Учитываются взгляды индивида и четырех его ближайших соседей в момент t (окрестность фон Неймана). Модель исследовалась на большом временном горизонте — до 20000 тактов. Оказалось, что партийная борьба приводит к очень сложным конфигурациям, существенно зависящим от исходного распределения.

Другая, созданная еще в 1956 г. модель, известная как модель Тибу [31], рассматривает людей, расположенных в конечном числе районов города. В рамках данной модели каждый житель принимает решение остаться в юрисдикции своего проживания или переехать в другой район с целью максимизации значения его функции полезности.

Агент принимает решение на основании информации об юрисдикциях: набор локальных общественных благ и уровень налогообложения. Имеется множество численных результатов, связанных с существованием равновесий модели Тибу и ее оптимальными свойствами.

Для изучения проблем живучести информационных систем как сложных многопараметрических систем, параметры которых еще малоизучены, наиболее подходящей методикой является математическое моделирование. Жизненный цикл информационных систем может описываться, например, моделью диффузии информации [32]. Напомним, что в естественных науках под диффузией понимают взаимное проникновение друг в друга соприкасающихся веществ, вызванное, например, тепловым движением их частиц. Процессы диффузии информации, как и процессы диффузии в физике, достаточно точно моделируются с помощью методов клеточных автоматов.

Клеточные автоматы являются полезными дискретными моделями для исследования динамических систем. Дискретность модели, а точнее, возможность представить модель в дискретной форме, может считаться важным преимуществом, поскольку открывает широкие возможности использования компьютерных технологий.

Модель диффузии информации, которую будем рассматривать в дальнейшем, является двумерной, поэтому вся система клеточных автоматов описывается двумерным массивом. В рамках данной модели, которая относится к распространению новостей в информационном

пространстве, применяются окрестность Мура [20] и вероятностные правила распространения новостей по заданной тематике.

Предполагается, что клетка может быть в одном из трех состояний: 1 — «свежая новость» (клетка окрашивается в черный цвет); 2 — новость, устаревшая, но сохраненная в виде сведений (серая клетка); 3 — клетка не имеет информации, переданной новостным сообщением (клетка белая, информация не дошла или уже забыта). Правила развития инфосюжета следующие:

— изначально все поле состоит из белых клеток за исключением одной — черной, которая первой «приняла» новость;

белая клетка может перекрашиваться только в черный цвет или оставаться белой (она может получать новость или оставаться «в неведении»);

— белая клетка перекрашивается, если выполняется условие: $Cpm > 1$, где p — псевдослучайная величина ($0 < p < 1$), m — количество черных клеток в окрестности, C — константа ($C = 1,5$ при $m = 1$; $C = 1$ при $m \neq 1$);

— если клетка черная, а вокруг нее исключительно черные и серые, то она перекрашивается в серый цвет (новость устаревает, но сохраняется как сведения);

— если клетка серая, а вокруг нее исключительно серые и черные, то она перекрашивается в белый цвет (происходит забывание сведений при их общеизвестности).

Описанная система клеточных автоматов вполне реалистично отражает процесс развития инфосюжета (рис. 5). На поле размером 40×40 (размеры были выбраны исключительно с целью наглядности) состояния системы клеточных автоматов полностью стабилизируются за ограниченное количество тактов, т.е. на практике процесс — сходящийся.

Типичные зависимости количества клеток (последовательности количества однотипных клеток), пребывающих в различных состояниях, в зависимости от шагов итерации приведены на рис. 6.

При анализе приведенных графиков следует обратить внимание на такие особенности: 1 — суммарное количество клеток, пребывающих во всех трех состояниях на каждом шагу итерации постоянно и равно размеру поля; 2 — при стабилизации клеточных автоматов соотношение количества серых, белых и черных клеток приблизительно составляет $0,75 : 0,25 : 0$; существует точка пересечения кривых, определяемых всеми тремя последовательностями на уровне 33 % каждая.

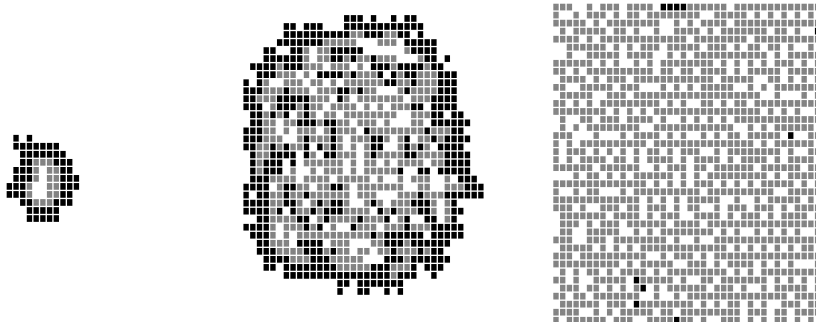


Рис. 5. Состояния эволюции системы клеточных автоматов

Именно черные клетки образуют актуальный инфосюжет, динамика которого представлена на рис. 6.

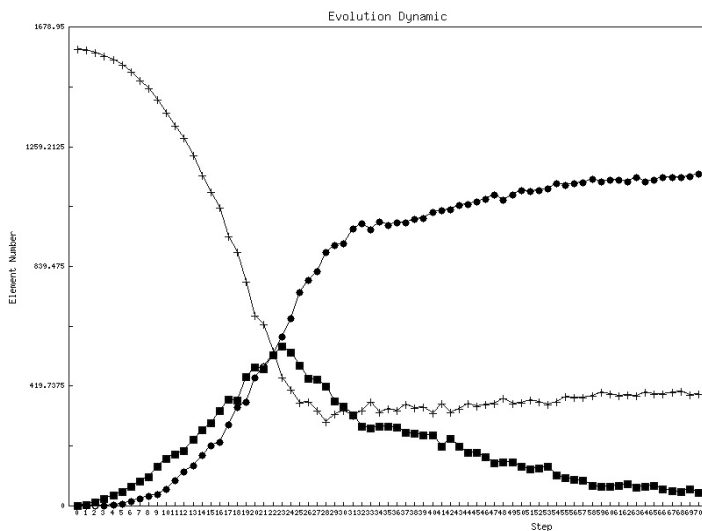


Рис. 6. Распределение клеток в зависимости от такта системы клеточных автоматов: белые клетки — (+); серые клетки — (•); черные клетки — (■)

Полученные в результате аналитического моделирования зависимости количества серых x_g , белых x_w и черных x_b клеток от шага эволюции системы клеточных автоматов, выражаются формулами:

$$x_g = \frac{0,75}{1 + e^{-0,15(t-30)}},$$
$$x_w = 1 - \frac{0,75}{1 + e^{-0,25(t-20)}},$$
$$x_b = 0,75 \left(\frac{1}{1 + e^{-0,25(t-20)}} - \frac{1}{1 + e^{-0,15(t-30)}} \right).$$

Жизненный цикл инфосюжетов так же, как многих других систем, может быть описан с помощью еще двух больших классов моделей: булевых и марковских.

Следует отметить, что зависимость диффузии новостей, полученная в результате моделирования, хорошо согласуется с реальным поведением тематических информационных потоков на интернет-источниках, а на локальных временных промежутках — с традиционными моделями.

2. АНАЛИЗ И ОЦЕНКА ЖИВУЧЕСТИ

Необходимо оценивать способности системы продолжать нормальное функционирование в условиях постоянно действующих деструктивных влияний и противостоять им, адаптировать алгоритмы функционирования к новым условиям и организовывать функциональное восстановление или обеспечить функционирование при постепенном процессе деградации, возможно без потери наиболее значимых «критических» информационных функций; необходим переход от анализа и оценки надежности к анализу и оценке живучести.

Под живучестью мы понимаем способность информационной системы сохранять и восстанавливать выполнение основных функций в заданном объеме и на протяжении заданного времени в случае изменения структуры системы и/или алгоритмов и условий ее функционирования вследствие неблагоприятных воздействий (НВ).

Кроме возможности «внутреннего» восстановления системы после НВ, живучесть системы характеризуется также возможностью воздействия на внешнюю среду, в которой сама система функционирует. Эта возможность особо четко видна как раз в случае информационных систем.

Одним из показателей живучести системы является запас живучести (*d*-живучесть) — критическое количество дефектов, уменьшенное на единицу. Под дефектом будем понимать единицу измерения ущерба, нанесенного информационной системе НВ. Если обозначить через *C* критическое количество дефектов, то показателем *d*-живучести будет $d = C - 1$.

Критическим называют минимальное количество дефектов, возникновение которых приводит к утрате информационной системой своей работоспособности (возможности информационного воздействия).

С другой стороны, запас живучести можно определить как максимальное количество дефектов, которое еще может выдержать система без утраты работоспособности.

Пусть $m_i = i$ -я комбинация дефектов, при которой система не утрачивает своей работоспособности, тогда запас живучести определяется как $m = \max_{(i)} m_i$.

Модели анализа и оценки живучести могут быть статическими и динамическими. В статических моделях задается участок поражения

информационной системы и интенсивность влияния конкретных видов НВ, определяют перечень элементов, которые могут быть поражены НВ (например, страниц веб-сайта), и с помощью логической функции работоспособности находят показатель качества функционирования системы. Динамические модели анализа являются имитационными моделями, которые включают: модель возникновения и развития НВ, модель изменения состояний элементов информационной системы под воздействием НВ и модель функционирования в условиях изменения структуры и значений параметров системы, связанных с НВ.

2.1. Функциональная живучесть

Живучесть систем анализируют и оценивают на различных уровнях проектирования, моделирования и функционирования информационных систем. Во время исследования функциональной живучести могут использоваться теоретико-игровые, вероятностные, графовые, матричные модели.

При исследовании функциональной живучести информационных систем особенности топологии сети межкомпонентных связей учитываются опосредовано. Предполагается, что в информационных системах обеспечивается необходимая связность работоспособных компонент.

При анализе живучести функционирования информационная система характеризуется:

— целью функционирования (информирование, дезинформирование, информационное воздействие и т.п.);

— множеством задач $Q = \{q_1, \dots, q_m\}$, решение которых обеспечивается с ее помощью;

— множеством компонент (информационных ресурсов) $\{S_1, S_2, \dots, S_p\}$, являющихся составными частями системы.

В процессе функционирования информационной системы ее компоненты могут находиться в одном из состояний: работоспособном, не работоспособном, частично работоспособном, т.е. работоспособном, но при частичном снижении (в допустимых пределах) значения каких-либо показателей качества функционирования.

На основе теоретико-игровых моделей исследуют живучесть систем, которые функционируют в условиях целенаправленного влия-

ния противника, внешних и внутренних деструктивных воздействий, когда компенсировать нештатные ситуации, потоки отказов и сбоев можно лишь за счет внутренних резервов системы и воздействия на источник деструктивных воздействий.

Вероятностные, графовые, матричные модели анализа и оценки живучести достаточно разнообразны. В каждом конкретном случае для разных моделей, учитывая различные цели функционирования, а также условия работоспособности системы возможно нахождение количественных оценок живучести. Показатели живучести различных систем можно сравнивать, если цели их функционирования совпадают.

Количественные показатели живучести существенно зависят от параметров, определяющих условия работоспособности информационной системы. Текущий уровень работоспособности определяет количество, качество и содержание функций, которые обобщаются понятием «цель функционирования системы». Для обеспечения цели функционирования системы можно применить одну из стратегий:

— f -стратегию — стратегию обеспечения отказоустойчивости (*fault-tolerance*);

— s -стратегию — стратегию обеспечения живучести (*survivality*).

В процессе формирования f -стратегии необходимо определить множество состояний системы $S^{(f)} = \{s_v^{(f)}\}$, при которых необходимо противодействовать угрозам нарушения работоспособности, задавать варианты распределения функций между работоспособными компонентами информационных системами в состояниях множества $S^{(f)}$.

Стратегия обеспечения отказоустойчивости ориентирована на полную компенсацию предусмотренных функциональных отказов и обеспечения показателей эффективности функционирования систем в этих случаях.

В процессе формирования s -стратегий для каждого состояния множества $S^{(f)}$ необходимо дополнительно наработать решения, относящиеся к функциям системы: суживать или нет множество функций, которые вместе составляют цель функционирования; как это сделать; упрощать или нет алгоритм реализации функций и т.д.

Вариант решения относительно цели функционирования системы в условиях наличия нежелательных влияний может быть одним из таких:

1. Множество функций системы не может быть изменено, должны использоваться все функции, возможно с меньшей эффективностью или с ухудшением качества, т.е. в любом состоянии из $S^{(f)}$ должно выполняться условие:

$$\prod_{i \in I} x(f_i) = 1, \quad x(f_i) = \begin{cases} 1, & \text{если } f_i \text{ выполняется;} \\ 0, & \text{если } f_i \text{ не выполняется.} \end{cases}$$

2. В любом состоянии из $S^{(f)}$ должно выполняться некоторое подмножество функций F^* , которые реализуют цель функционирования информационной системы, т.е.

$$\prod_{f_i \in F^*} x(f_i) = 1.$$

Множество функций F^* зависит от состояния системы и заданных условий к функциональной живучести.

3. В произвольном состоянии из $S^{(f)}$ система должна обеспечивать выполнение хотя бы одной функции из множества F^* , т.е.

$$\sum_{i \in I} x(f_i) \geq 1.$$

Функциональная живучесть информационной системы зависит от заданной заранее цели ее функционирования. Функциональная живучесть различных информационных систем можно сравнивать, если они имеют одинаковые цели функционирования. Оценка живучести одной и той же информационной системы может изменяться в случае изменения цели функционирования. При этом столь же существенное влияние на количественные показатели живучести информационных систем, как цель функционирования, оказывают параметры, определяющие условия их работоспособности.

Выбирая механизмы повышения функциональной живучести конкретной информационной системы, необходимо учесть цель функционирования (множество функций, которые реализует система), структуру связей, особенности функциональных компонент.

Под целью функционирования при этом имеют в виду понятие, которое вводится для живучих и отказоустойчивых информационных систем, но изменение ее предполагается возможным только для систем, обладающих свойством живучести.

Качественная зависимость цели функционирования (числа выполняемых информационных функций) от количества отказов компонент для отказоустойчивых и живучих систем показана соответственно на рис. 7, а и 7, б.

Вместе с тем, существуют живучие системы, для которых рассматриваемая зависимость выражается зависимостью, представляющей собой нечто промежуточное между зависимостями, приведенными на рисунках. Подробно поведение таких систем (систем с фазовыми переходами) будет рассмотрено в главе, посвященной критическому уровню живучести систем.

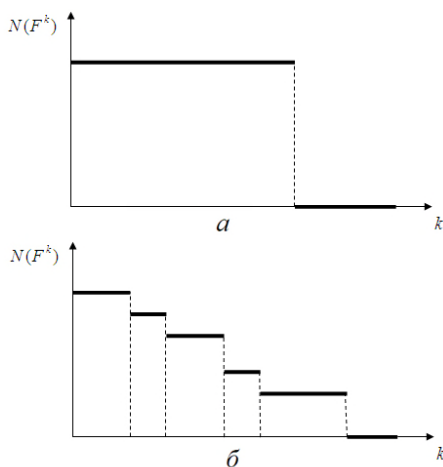


Рис. 7. Зависимость количества выполняемых функций (вертикальная ось) от количества отказов компонент (горизонтальная ось, k)

В процессе анализа и оценки функциональной живучести информационной системы допускается, что можно обеспечить необходимые связи между отдельными функциональными компонентами.

Обозначим множество информационных функций, реализуемых системой, через $F = \bigcup_{i \in I} F_i = \{f_1, f_2, \dots, f_n\}$, причем функциональная

компонента Φ_k потенциально может выполнять множество функций $\varphi_n : \{1, 2, \dots, p\} \rightarrow P(F)$, где $P(F)$ — множество всех подмножеств F . Если $\varphi_n(k) = \{f_{k_1}, f_{k_2}, \dots, f_{k_j}\}, 1 \leq k_r \leq n$, то функциональная компонента Φ_k может выполнять функции $f_{k_1}, f_{k_2}, \dots, f_{k_j}$.

В каждый конкретный момент времени функциональная компонента Φ_k предназначается для выполнения некоторого множества функций, определяемых таким образом: $\varphi_t : \{1, 2, \dots, p\} \rightarrow P(F)$. Если $\varphi_n(k) = \{f_{k_1}, f_{k_2}, \dots, f_{k_j}\}, 1 \leq k_r \leq n$, то функциональная компонента Φ_k может выполнять функции $f_{i_1}, f_{i_2}, \dots, f_{i_j}$. Если $\varphi_t(k) = \emptyset$, то Φ_k неработоспособна.

Каждая функция $f_i \in F$ характеризуется некоторой эффективностью выполнения c_i (например, временем выполнения, объемом аудитории при подаче информации и т.п.). Функция эффективности для системы может быть определена как $\varphi_{ef} : F \times \{1, 2, \dots, p\} \times P(F) \rightarrow C$, где C — некоторое числовое множество; $\varphi_{ef}(f_i, k, \varphi_t(k)) = c_{i_k}$ означает, что функциональная компонента Φ_k предназначена для выполнения функций $\varphi_t(k) = \{f_{i_1}, f_{i_2}, \dots, f_{i_j}\}$, тогда эффективность выполнения $f_i \in \{f_{i_1}, f_{i_2}, \dots, f_{i_j}\}$ равна c_{i_k} .

Условия достижения цели функционирования (выполнения определенного множества функций с заданной эффективностью) можно определить таким образом:

$$\bigcup_{k=1}^p \varphi_n(k) \supseteq F;$$

$$\varphi_t(k) \subseteq \varphi_n(k) \quad \forall k = \overline{1, p};$$

$$\sum_{k=1}^p \varphi_{ef}(f_i, k, \varphi_t(k)) \geq c_i, \quad \forall i = \overline{1, n}.$$

Под функциональным отказом будем понимать невозможность выполнения функциональной компонентой некоторой функции. В этом случае изменяется состояние функциональной компоненты и соответственно функции φ_i . Если это приводит к нарушению условий достижения цели функционирования, то средства обеспечения живучести должны скорректировать функционирование системы таким образом, чтобы выполнялись приведенные выше условия, возможно, за счет изменения функции φ_i . Оптимальность поведения системы можно определить путем выполнения необходимых условий за счет изменения минимального количества φ_i , т.е. изменения количества функциональных компонент, задействованных в процедурах компенсации отказов. Критерием живучести системы (количественной оценкой) при этом может быть количество компенсированных функциональных отказов.

Если компоненты Φ_k функционально однородные, а цель функционирования определяют через заданный уровень производительности системы, который обеспечивается благодаря наличию соответствующего количества работоспособных функциональных компонент:

$$R(\Phi_k, t) \geq R^* = \text{const},$$

где $R(\Phi_k, t)$ — среднее количество работоспособных функциональных компонент в системе в момент времени $t \geq 0$, R^* — минимально допустимое количество работоспособных функциональных компонент, при котором производительность системы не меньше необходимой, то оценкой функциональной живучести может быть функция

$$N(\Phi_k, t) = \overline{\Omega}(\Phi_k, t) / (N\omega),$$

где $\overline{\Omega}(\Phi_k, t)$ — математическое ожидание производительности системы в момент времени $t \geq 0$; $N\omega$ — суммарная производительность всех функциональных компонент.

2.2. Структурная живучесть

Структурная живучесть рассматривается как возможности реконструкции, реорганизации, реконфигурации при НВ, которые позволят создать структуру, обеспечивающую выполнение критического подмножества функций для достижения цели функционирования системы.

При рассмотрении структурной живучести учитывается топология сети межкомпонентной связи и надеждностные характеристики компонент. Задачи, связанные с анализом структурной живучести, можно свести к задачам надежности, связности топологических структур, в зависимости от введения понятия «разрушение».

Анализ структурной живучести требует определения:

— структуры для выполнения цели функционирования системы в некоторый момент времени, когда возникают нежелательные влияния на систему;

— требований к отдельным видам ресурсов системы и их взаимосвязи;

— требований к функциональным возможностям компонент системы;

— особенностей характера нежелательных влияний или их последствий.

Структурную живучесть системы можно оценивать при некоторых допущениях, которые позволяют упростить задачу оценки и свести ее к задаче анализа связности графов, оценки вероятности формирования работоспособной структуры в случае нежелательных влияний и т.п.

При рассмотрении информационной системы необходимо также учитывать количество структур в этой системе, которые могут выполнять критическое подмножество информационных функций. Для подсчета этого количества необходимо выделять такие структуры, т.е. переходя на язык теорий графов и сложных сетей — находить в сетевых структурах множество связанных определенным образом компонент (не обязательно «сильно связанных»), например клик, при учете, что связи формируются за счет смысловых зависимостей. Переход к рассмотрению сетевых моделей должен базироваться также на оценке характеристик сетей, их максимальных и минимальных сечениях, потоках в этих сетях, стоимостях и т.д.

При исследовании структурной живучести с помощью графовых моделей совокупность компонент системы $G(V, R)$ изображают в виде вершин графа $v \in V$, а ребра графа $r \in R$ соответствуют связям между ними. Систему, которая моделируется с помощью графа, считают разрушенной, если в случае удаления вершин или ребра граф будет удовлетворять одному или большему количеству условий:

- граф состоит минимум из двух компонент;
- не существует направленных путей для определенных множеств вершин;
- количество вершин в наибольшей компоненте графа $G(V, R)$ меньше некоторого заранее заданного числа;
- кратчайший путь превышает некоторую заданную величину.

Соответственно система считается живучей, если эти условия не выполняются. Структурную живучесть систем обычно характеризуют различными показателями связности. Расчет таких показателей, как например, вероятность связности при условиях случайного существования ребер графа, на практике ограничивается вычислительной сложностью таких задач. Одновременно, используя пути и разрывы графа, моделирующего систему, можно получить достаточно простые предельные оценки необходимых показателей.

Опишем механизм реконфигурации информационной системы (информационного кластера) при деструктивном воздействии — удалении элемента (документа).

Пусть C_1, C_2, \dots, C_M — различные информационные системы, воздействующие на субъекты (объекты влияния) U_1, U_2, \dots, U_L . Информационная система C_j состоит из элементов $D_{11}, D_{12}, \dots, D_{1j}$.

При этом сила влияния каждого элемента D_{ij} информационной системы C_j на субъекта U_k оценивается значением $f_v(D_{ij}, U_k) = V(i, j, k) \geq 0$.

На рис. 8, *a* изображена схема влияния на субъекты U_1, U_2, U_3 информационной системы C_1 . При этом на субъект U_2 также оказывается ненулевое воздействие со стороны информационной системы C_2 через элемент D_{21} . Следует отметить, что как между элементами из одной информационной системы, так и между отдельными инфор-

мационными системами и их элементами существуют информационные связи (не отображенные на рис. 8).

В результате деструктивного воздействия путем удаления элемента D_{13} информационная система C_1 теряет свое влияние на субъект U_2 (рис. 8, б).

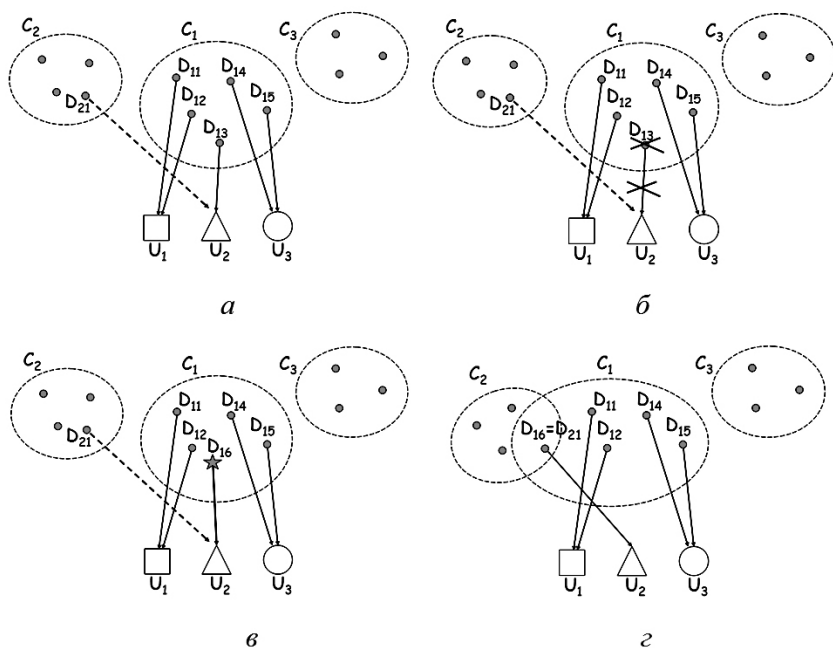


Рис. 8. Восстановление после деструктивного воздействия на информационную систему C_1

Восстановление влияния на субъект U_2 возможно с помощью реконфигурации информационной системы C_1 в трех направлениях:

1) генерации новых документов, воздействующих на U_2 , например, D_{16} (рис. 8, в);

2) модификации существующих сообщений из C_1 , добавление к ним контента, оказывающего влияние на U_2 ;

3) модификации сообщения D_{21} в направлении тематики информационной системы C_1 , и, таким образом, объединения части информационной системы C_2 с C_1 (рис. 8, 2). В реальной жизни этот путь может оказаться затруднительным, например, в том случае, когда информационная система C_2 принадлежит силам, конкурирующим (или находящимися в антагонистических отношениях) с владельцами C_1 .

В любом случае можно оценить стоимость затрат на создание/реконфигурацию информационной системы.

Исходя из того, что общая оценка влияния информационной системы C_i определяется по формуле:

$$W(i) = \sum_{k=1}^L \sum_{j=1}^{|C_i|} f_v(D_{ij}, k),$$

где $|C_i|$ — количество элементов в информационной системе C_i , суммарная стоимость создания этой информационной системы составляет

$$S(i) = \sum_{j=1}^{|C_i|} E(D_{ij}),$$

где $E(D_{ij})$ — затраты на создание элемента D_{ij} информационной системы C_i , то один из способов определения эффективности информационной системы, состоит в оценке выражения

$$Q = \frac{\sum_{k=1}^L \sum_{j=1}^{|C_i|} f_v(D_{ij}, k)}{\sum_{j=1}^{|C_i|} E(D_{ij})}.$$

Важную группу характеристик структурной живучести составляют так называемые «меры живучести». Определяя их, выходят из

допущения, что противник, зная структуру сети, пытается нарушить ее функционирование. Живучесть системы считают высокой, если необходимо разрушить значительное количество узлов и/или ребер, чтобы существенно ухудшить или прервать ее функционирование.

В математической теории графов показатели живучести интерпретируются как количественные меры связности для структуры графа: минимальное сечение, узловая связность, обобщенная связность, длина пути и т.п.

Для проектирования информационных систем задача анализа структурной живучести может формулироваться как задача оценки величины максимального потока, который может передаваться в сети в случае отказов ее элементов и понижения до допустимого уровня качества функционирования.

Например, структурную живучесть сети MPLS (*Multiprotocol Label Switching*) [2] оценивают с помощью такого алгоритма:

1. Рассчитывают общую величину потока в безотказном состоянии для всех классов сервиса;
2. Моделируют различные отказные состояния и рассчитывают вероятности этих состояний;
3. Для каждого отказного состояния рассчитывают: относительное понижение общей пропускной способности; оптимальный поток; максимальный поток; резерв пропускной способности и т.д.

Связанные со структурной живучестью автоматическая реконфигурация и перезагрузка, например, в случае DoS-атак, сегодня де-факто являются обязательными функциями при организации веб- и прокси-серверов, в частности под управлением HTTP-сервера Apache (<http://apache.org/>). В частности, существуют внешние программные модули этого веб-сервера, вносящие изменения в его конфигурацию и обеспечивающие перезагрузку при обнаружении DoS-атак. В настоящее время для другого HTTP-сервера — nginx (<http://sysoev.ru/nginx/>) — подобные средства являются обязательными компонентами. Модуль `ngx_http_limit_req_module` этой системы позволяет ограничить число запросов для заданной сессии или, как частный случай, с одного адреса. Модуль `ngx_http_limit_zone_module` позволяет ограничить число одновременных соединений для заданной сессии или, как частный случай, с одного адреса. В декабре 2009 года nginx использовался на 4 % самых посещаемых сайтов в мире. Реконфигурированием и пе-

резапуском прокси-сервера Squid (<http://www.squid.org/>) занимается специальная программа, постоянно находящаяся в памяти (так называемый демон) `samsdaemon`. Эта программа через заданные промежутки времени проверяет, необходима ли реконфигурация. Например, если пользователь превысил предел своих полномочий, возможно автоматическое внесение его данных в соответствующий конфигурационный файл прокси-сервера. Таким образом, реконфигурирование прокси-сервера Squid происходит путем изменения файла `squid.conf` с последующей подачей сигнала на реконфигурацию.

2.3. Теоретико-игровой подход

В теоретико-игровых моделях живучесть информационных систем можно рассматривать с точки зрения обмена некоторого количества затраченных V ресурсов и некоторого количества потребляемых W ресурсов. В процессе количественной оценки живучести систем, которые описываются моделями такого типа, информационные системы соответствуют продвижению товаров на рынке в экономической среде. При этом вводится и конкретизируется содержание (V, W) -обменов, выполняется оптимизация этих обменов. Живучесть информационных систем можно оценивать по наличию в них жизненно значимых элементов, которые определяются по экспертным оценкам.

Рассмотрим систему A , находящуюся во взаимодействии с системой B .

Каждая система во времени может переводить себя в различные состояния, изменяя свою структуру и поведение. Предположим, что целью \underline{A} системы A является ее стремление достигнуть определенных предпочтительных для нее состояний, целью \underline{B} системы B является такое же стремление достигнуть состояний, предпочтительных для нее. Целесообразность структуры $|A|(|B|)$ и целенаправленность поведения $\bar{A}(\bar{B})$ системы $A(B)$ оценивается эффективностью, с которой система достигает цели $\underline{A}(\underline{B})$. Итак, систему $A(B)$ при фиксации ее цели можно характеризовать двумя факторами:

$$A = (|A|, \bar{A}); \quad [B = (|B|, \bar{B})].$$

Целью системы является наиболее выгодный (V, W) -обмен; т.е. за минимальное количество ресурса V получить возможно наибольшее количество W , которое является функцией структуры и поведения обеих систем:

$$W = W(V, |A|, |B|, \bar{A}, \bar{B}) = W(V, A, B).$$

В результате взаимодействия между собой системы A и B получают следующие (V, W) -обмены:

$$\begin{aligned} \underline{W}_a &= W_a(V_a, A_0, B_0) = \\ &= \max_{\{\bar{A}, |A|\}} \left[(1 + \alpha) W_a(V_a, A, B) - \alpha \max_{\{\bar{B}, |B|\}} W_a(V_a, A, B) \right]; \\ \underline{W}_b &= W_b(V_b, A_0, B_0) = \\ &= \max_{\{\bar{B}, |B|\}} \left[(1 + \alpha) W_b(V_b, A, B) - \alpha \max_{\{\bar{A}, |A|\}} W_b(V_b, A, B) \right]; \end{aligned}$$

где

$$\alpha = \begin{cases} -1 & \text{— в конфликтной ситуации;} \\ 0 & \text{— в индифферентной ситуации.} \end{cases}$$

Здесь A_0 и B_0 — оптимальные системы, т.е. системы, структуры и поведение которых является оптимальным, так как с их помощью можно произвести (V, W) -обмены, близкие к оптимальным.

Для определения своей цели каждой системе необходимо решить, что для нее важнее, получить ли самый выгодный (V, W) -обмен или помешать это сделать другой системе. При этом системы могут варьировать значения (V, W) -обменов в некоторых пределах:

$$\underline{W}_1 \leq W_a \leq \bar{W}_1;$$

$$\underline{W}_2 \leq W_b \leq \bar{W}_2,$$

где \underline{W}_1 и \underline{W}_2 соответствуют максимально агрессивным системам, а \bar{W}_1 и \bar{W}_2 — наиболее осторожным. Если цели систем известны, то имеется вполне определенная ситуация. Если же каждая система или одна из них скрывает свои намерения, то налицо игровая ситуация относительно выбора цели.

Обозначим через \underline{A}_i и \underline{B}_j ($i=1, \dots, n$; $j=1, \dots, m$) цели систем A и B соответственно. Цели \underline{A}_1 и \underline{B}_1 состоят в нанесении максимального ущерба противоположной системе, а цели \underline{A}_n и \underline{B}_m соответствуют крайней осторожности обеих систем. Все остальные цели \underline{A}_i и \underline{B}_j соответствуют промежуточным ситуациям и нумерованы в порядке последовательного перехода от \underline{A}_1 к \underline{A}_n и от \underline{B}_1 к \underline{B}_m . Предполагая, что в ситуации $\{\underline{A}_i, \underline{B}_j\}$ системы A и B получают выигрыш $\underline{W}_a = a_{ij}$ и $\underline{W}_b = b_{ij}$, получаем биматричную игру на определение оптимальной цели с платежными матрицами $\|a_{ij}\|$ и $\|b_{ij}\|$. Если принять, что $a_{ij} \leq a_{sj}$ при $i \leq s$; $b_{ij} \leq b_{sj}$ при $i \leq s$; $a_{is} \leq a_{ij}$ при $j \leq s$; $b_{is} \leq b_{ij}$ при $j \geq s$, то решение игры тривиально и состоит в том, что обеим системам необходимо придерживаться целей \underline{A}_n и \underline{B}_m , т.е. вести себя осторожно и не проявлять агрессивности. Взаимодействие сложных систем становится строго антогонистическим, когда (V, W) -обмен для одной системы положительный, а для другой — отрицательный и $\underline{W}_a = \underline{W}_b$.

Как правило, взаимодействие систем A и B при (V, W) -обмене носит стохастический характер, и поэтому можно говорить лишь о некоторой вероятности $P(V, W)$ достижения каждой системой своей цели. Эта вероятность служит показателем эффективности

поведения системы. Максимальное значение этой вероятности определяется как предельная эффективность.

Системы A и B не могут рассчитывать на получение $W_a > \bar{W}_a$ и $W_b > \bar{W}_b$. Поэтому в пределе

$$\begin{aligned} P_a(V_a, W_a) &\cong 0 \text{ при } W_a > \bar{W}_a; \\ P_b(V_b, W_b) &\cong 0 \text{ при } W_b > \bar{W}_b. \end{aligned}$$

Аналогично значения $W_a < \bar{W}_a$ и $W_b < \bar{W}_b$ системы можно получить с большими вероятностями, т.е. в пределе

$$\begin{aligned} P_a(V_a, W_a) &\cong 1 \text{ при } W_a < \bar{W}_a; \\ P_b(V_b, W_b) &\cong 1 \text{ при } W_b < \bar{W}_b. \end{aligned}$$

Такое асимптотическое поведение вероятности $P(V, W)$ позволяет в конкретной ситуации найти предельный закон (V, W) -обмена, который устанавливает пределы эффективности функционирования системы.

Все элементы A и B разделяются на три класса:

— рабочие (жизненно важные) a - и b -элементы;

— защитные R_a - и R_b -элементы;

— активные (воздействующие на внешнюю среду) c_a - и c_b -элементы.

Приведем пример. Будем считать, что до начала взаимодействия системы A и B имеют некоторые ограниченные «ресурсные запасы» \bar{V}_a и \bar{V}_b , из которых генерируются жизненно важные элементы, причем

$$\bar{V}_a = \{V_{aj}\}; \bar{V}_b = \{V_{bj}\}; \quad (j = 1, \dots, n_a(n_b)),$$

где n_a — количество типов жизненно важных элементов системы A ; n_b — количество типов жизненно важных элементов системы B .

Из $V_{ai}(V_{bj})$ можно воспроизвести $A_i(B_j)$ $a(b)$ -элементов j -го типа ценностей $a_j(b_j)$. Ценности элементов систем определяются в относительных единицах, выбираемых на основе экспертных оценок и в каждом случае по своим правилам.

Защитные и активные элементы каждой системы воспроизводятся (генерируются) жизненно важными элементами.

Начальные ресурсы \bar{V}_a и \bar{V}_b вследствие своей ограниченности могут воспроизвести либо много малоэффективных и невысокой стоимости элементов (например, покупка рекламных ссылок на большом количестве малорейтинговых веб-ресурсов), либо мало высокоэффективных и высокой стоимости элементов (например, покупка рекламных ссылок в престижном каталоге системы Яндекс).

Таким образом, в информационной системе имеются определенные критические соотношения между количеством всех элементов, обеспечивающих максимальную живучесть.

Будем считать, что системы A и B в момент начала взаимодействия имеют:

1) A_j и B_j жизненно важных элементов типа ($j=1, \dots, n_a(n_b)$) и ценностей a_j и b_j соответственно, причем

$$\sum_{j=1}^{n_a} a_j A_j = N_a(0); \quad \sum_{j=1}^{n_b} b_j B_j = N_b(0);$$

2) r_a и r_b типов защитных элементов по α_m и β_m , где

$$\sum_{m=1}^{r_a} \alpha_m = N_{R_a}(0); \quad \sum_{m=1}^{r_b} \beta_m = N_{R_b}(0);$$

3) S_a и S_b типов защитных элементов по V_m^a и V_m^b , причем

$$\sum_{m=1}^{S_a} V_m^a = M_a(0); \quad \sum_{m=1}^{S_b} V_m^b = M_b(0).$$

Взаимодействие систем A и B происходит на некотором отрезке времени и состоит во взаимном обмене определенными порциями активных c -элементов.

Системы A и B в каждый момент времени t определяют свое поведение:

$$\begin{aligned}\bar{A}(t) &= \{ \mu_{\omega}^a(t), \sigma_{\omega}^a(t) \}; \\ \bar{B}(t) &= \{ \mu_{\omega}^b(t), \sigma_{\omega}^b(t) \},\end{aligned}$$

где $\mu_{\omega}^{a(b)}(t)$ и $\sigma_{\omega}^{a(b)}(t)$ — порции $R_a(R_b)$ - и $c_a(c_b)$ -элементов, направленных на защиту и/или уничтожение ω -элементов (ω -элемент — ($\omega = a, b, R, C$)).

С течением времени порции c_a - и c_b -элементов заполняют системы B и A соответственно, и, таким образом, со временем элементы системы редуют, если нет их пополнения.

Для получения количественной оценки взаимодействия систем A и B необходимо конкретизировать смысл (V, W) -обменов, к оптимизации которых стремятся обе системы. Будем считать, что живучесть системы определяется наличием в ней жизненно важных элементов. Другими словами, система $A(B)$ в момент t функционирует нормально, если выполняются условия:

$$\begin{aligned}Q_a(A_1, \dots, A_{n_a}, t) &> Q_a \min(t); \\ Q_b(B_1, \dots, B_{n_b}, t) &> Q_b \min(t),\end{aligned}$$

где Q_a и Q_b — заданные функционалы времени и структур систем A и B соответственно.

В частности, можно принять

$$\begin{aligned}Q_a &= \sum_{j=1}^{n_a} a_j A_j(t); \quad Q_a \min(t) = \sum_{j=1}^{n_a} Q_{aj} a_j A_j(0); \\ Q_b &= \sum_{j=1}^{n_b} b_j B_j(t); \quad Q_b \min(t) = \sum_{j=1}^{n_b} Q_{bj} b_j B_j(0),\end{aligned}$$

где $0 \leq Q_{aj}, Q_{bj} \leq 1$ — параметры, определяемые спецификой систем и требованиями, которые предъявляются к их живучести.

Выбор критериев оптимальности в приведенной форме представляется естественным в тех ситуациях, когда живучесть системы сильно зависит от суммарной ценности жизненно важных элементов и в меньшей — от распределения элементов по их типам.

Параметры Q_{aj} и Q_{bj} характеризуют жизнеспособность систем A и B соответственно. Действительно, если система A в момент t функционирует нормально, то $(1 - Q_{aj})$ -я доля a -элементов j -го типа ($j = 1, \dots, n$) в этот момент находится в рабочем состоянии. При $Q_{aj} = 0$ ($j = 1, \dots, n$) система A является наиболее жизнеспособной, а при $Q_{aj} \rightarrow 1$ жизнеспособность системы A стремится к нулю. Аналогично для системы B .

Если учесть, что процессы (V, W) -обмена каждой системы взаимосвязаны, то ситуацию взаимодействия систем A и B можно рассматривать как обобщенный конфликт, для описания которого необходимо задание функции выигрыша. Вид функции выигрыша каждой системы существенно зависит от степени конфликтности ситуации, т.е. от того, как влияет одна система на выполнимость цели другой системой, а также от уровня, на котором ведется конфликт.

В общем случае, с учетом возможных изменений целей интенсивность конфликта вид функции выигрыша также может изменяться с течением времени. Таким образом, рассматриваемая обобщенная конфликтная ситуация между системами A и B представляет собой игру двух игроков с нулевой суммой со следующими функциями выигрыша:

$$Q_a(T) = \sum_{j=1}^{n_a} a_j A_j(T); \quad Q_b = \sum_{j=1}^{n_b} b_j B_j(T)$$

при фиксации времени окончания конфликта, а при неавтономности критериев оптимальности игра заканчивается в момент времени $t^* = \min(t_a, t_b)$, где t_a, t_b — наименьшие неотрицательные корни уравнений $Q_a = Q_{a \min}$ и $Q_b = Q_{b \min}$.

Исход конфликта существенно зависит от степени важности для каждой системы постановки вопроса: устранить ли противника или выжить самой любыми средствами, т.е. в зависимости от напряженности конфликта выигрыши обеих систем могут изменяться в некоторых пределах:

$$\underline{W}_1 \leq \max_{(A,A)} \min_{(B,B)} \sum_{j=1}^{n_a} a_j A_j(T) \leq \bar{W}_1;$$

$$\underline{W}_2 \leq \max_{(B,B)} \min_{(A,A)} \sum_{j=1}^{n_b} b_j B_j(T) \leq \bar{W}_2.$$

Пусть взаимодействие систем A и B происходит на некотором отрезке времени $0 \leq t \leq T$ в дискретные моменты t_i . Системы A и B в каждый момент времени t определяют свое поведение:

$$\begin{aligned} \bar{A}(t) &= \{ \mu_\omega^a(t), \sigma_\omega^a(t) \}; \\ \bar{B}(t) &= \{ \mu_\omega^b(t), \sigma_\omega^b(t) \}; \end{aligned}$$

где $\mu_\omega^{a(b)}(t)$ и $\sigma_\omega^{a(b)}(t)$ — порции $R_a(R_b)$ - и $c_a(c_b)$ -элементов, направленных на защиту и уничтожение ω -элементов. Обозначим порцию активных элементов: $\sigma_\omega^a = (\sigma_{1\omega}^a, \dots, \sigma_{s_a\omega}^a)$; $\sigma_\omega^b = (\sigma_{1\omega}^b, \dots, \sigma_{s_b\omega}^b)$ и порцию защитных элементов: $\mu_\omega^a = (\mu_{1\omega}^a, \dots, \mu_{s_a\omega}^a)$; $\mu_\omega^b = (\mu_{1\omega}^b, \dots, \mu_{s_b\omega}^b)$.

В этих обозначениях в момент t_i стратегии поведения систем A и B будут иметь вид

$$\begin{aligned} \bar{A}^{(i)} &= \left\{ \|\mu_a^a\|, \|\mu_c^a\|, \|\sigma_b^a\|, \|\sigma_R^a\|, \|\sigma_c^a\| \right\}; \\ \bar{B}^{(i)} &= \left\{ \|\mu_b^b\|, \|\mu_c^b\|, \|\sigma_a^b\|, \|\sigma_R^b\|, \|\sigma_c^b\| \right\}. \end{aligned}$$

Изменение среднего числа выживших до момента t_{i+1} элементов будет описываться следующими соотношениями:

— жизненно важные и защитные для A :

$$A_n(t_{i+1}) = \max \left\{ 0, A_n(t_i) - \sigma_{an}^b(t_i) P_{an}^b(t_i) \right\} \quad (n = 1, \dots, n_a);$$

$$\alpha_j(t_{i+1}) = \max \left\{ 0, \alpha_j(t_i) - \sigma_{R_j}^b(t_i) P_{R_j}^b(t_i) \right\} \quad (j = 1, \dots, r_a).$$

— жизненно важные и защитные для B :

$$B_m(t_{i+1}) = \max \left\{ 0, B_m(t_i) - \sigma_{bm}^a(t_i) P_{bm}^a(t_i) \right\} \quad (m = 1, \dots, m_b);$$

$$\beta_s(t_{i+1}) = \max \left\{ 0, \alpha_s(t_i) - \sigma_{R_s}^a(t_i) P_{R_s}^a(t_i) \right\} \quad (s = 1, \dots, r_b),$$

где вероятность выхода из строя соответствующих активных элементов

$$P_{an}^b(t_i) = (P_{1an}^b, \dots, P_{s_b an}^b); \quad P_{R_j}^b(t_i) = (P_{1R_j}^b, \dots, P_{s_b R_j}^b);$$

$$P_{bm}^a(t_i) = (P_{1bm}^a, \dots, P_{s_a bm}^a); \quad P_{R_s}^a(t_i) = (P_{1R_s}^a, \dots, P_{s_a R_s}^a).$$

При этом, учитывая возможность выхода из строя C -элементов в каждой системе, получаем следующие ограничивающие условия:

$$\sum_{i=0}^{N-1} \left\{ \sigma_{cn}^b(t_i) P_{cn}^b(t_i) + \sum_{j=1}^{n_b} \sigma_{nb_j}^a(t_i) + \sum_{s=1}^{n_b} \sigma_{nR_s}^a(t_i) + \sum_{j=1}^{S_b} \sigma_{nc_j}^a(t_i) \right\} = v_n^a(0), \quad (n = 1, \dots, S_a);$$

$$\sum_{i=0}^{N-1} \left\{ \sigma_{cs}^a(t_i) P_{cs}^a(t_i) + \sum_{j=1}^{n_a} \sigma_{sa_j}^b(t_i) + \sum_{j=1}^{r_a} \sigma_{nR_j}^b(t_i) + \sum_{j=1}^{S_a} \sigma_{sc_j}^b(t_i) \right\} = v_s^b(0), \quad (s = 1, \dots, S_b),$$

где $P_{cm}^b(t_i) = (P_{1cm}^b, \dots, P_{s_b cm}^b); \quad P_{cs}^a(t_i) = (P_{1cs}^a, \dots, P_{s_a cs}^a).$

Приведенные уравнения характеризуют состояние систем A и B в среднем. Решение этих уравнений позволяет найти стратегии и определить условия выживаемости каждой системы.

При построении модели взаимодействия двух систем A и B необходимо учитывать такие факторы, как состояние информированности сторон и возможность превентивности действий. Фактор инфор-

мированности играет большую роль при выборе критерия оптимальности и достаточно сильно влияет на качество оптимальных стратегий.

Предположим, что системы A и B оценивают ситуацию исходя из функций выигрыша Q_a и Q_b . Тогда исходную игру можно заменить эквивалентной ей игрой с функцией выигрыша $Q = Q_a - Q_b$. Система A будет стремиться максимизировать Q , а система B — минимизировать. При этом возможны такие ситуации:

1. Система B информирована о выборе стратегии $\bar{A}(t)$ системой A . Система A такой информации не имеет. Поэтому оптимальная стратегия $\bar{A}_{opt}(t)$ системы A должна определяться из условия

$$\begin{aligned} & \max_{\bar{A}(t)} \min_{\bar{B}(t)} Q[\bar{A}(t), |A(t)|, \bar{B}(t), |B(t)|] = \\ & = \max_{\bar{A}(t)} Q[\bar{A}(t), |A(t)|, \bar{B}^*[t, \bar{A}(t)], |B(t)|] = \\ & = Q[\bar{A}_{opt}(t), |A(t)|, \bar{B}^*[t, \bar{A}_{opt}(t)], |B(t)|] = V_1[|A(t)|, |B(t)|]. \end{aligned}$$

2. Система A информирована о выборе стратегии $\bar{B}(t)$ системой B . Система B такой информации не имеет. В этом случае оптимальная стратегия $\bar{B}_{opt}(t)$ системы B может быть определена из условия

$$\begin{aligned} & \min_{\bar{B}(t)} \max_{\bar{A}(t)} Q[\bar{A}(t), |A(t)|, \bar{B}(t), |B(t)|] = \\ & = \min_{\bar{B}(t)} Q[\bar{A}^*[t, \bar{B}(t)], |\bar{A}(t)|, \bar{B}(t), |B(t)|] = \\ & = Q[\bar{A}^*[t, \bar{B}_{opt}(t)], |A(t)|, \bar{B}^*[t, \bar{A}^*_{opt}(t)], |B(t)|] = V_2[|A(t)|, |B(t)|]. \end{aligned}$$

3. Системы A и B имеют полную информацию о выборе стратегии каждой из них. Тогда оптимальные стратегии $\bar{A}^*_{opt}(t)$ и $\bar{B}^*_{opt}(t)$ систем A и B могут быть найдены из соотношения

$$\begin{aligned}
& \max_{\bar{A}(t)} \min_{\bar{B}(t)} Q[\bar{A}(t), |A(t)|, \bar{B}(t), |B(t)|] = \\
& = \min_{\bar{B}(t)} \max_{\bar{A}(t)} Q[\bar{A}(t), |A(t)|, \bar{B}(t), |B(t)|] = \\
& = Q[\bar{A}^*_{opt}[t, \bar{B}^*_{opt}(t)], |A(t)|, \bar{B}^*[t, \bar{A}^*_{opt}(t)], |B(t)|] = \\
& = V_3[|A(t)|, |B(t)|].
\end{aligned}$$

Из приведенного следует

$$V_1[|A(t)|, |B(t)|] \leq V_3[|A(t)|, |B(t)|] \leq V_2[|A(t)|, |B(t)|].$$

Таким образом, ценность взаимной тактики информированности систем A и B определяется разностью $\Delta V = V_2 - V_1$.

2.4. Логико-вероятностные модели

Для анализа и оценки живучести информационных систем, которые функционируют в условиях НВ, можно воспользоваться логико-вероятностными моделями, в соответствии с которыми допускают, что элементы системы и сама система имеют двузначную логику функционирования, и все события в системе независимы. Описание системы возможно с помощью статической модели, которая не содержит параметра времени среди независимых переменных. Функциональные зависимости между переменными при этом можно полностью представить с помощью функций алгебры логики. Результаты действия влияний также оценивают с помощью двоичной схемы: или сохраняется состояние из множества работоспособных состояний, или работоспособность нарушается (система переходит в состояние из множества неработоспособных состояний).

Допустим, что элементы системы являются точечными объектами, которые соединены между собой линиями связи. Последовательность НВ импульсного типа формирует поток независимых событий. Вторичных последствий НВ не существует, поэтому установленное состояние системы известно непосредственно после НВ.

Рассмотрим систему, которая состоит из K_s объектов, s — но-

мер варианта системы, в которой после одноразового нежелательного влияния на i -й объект возникает ущерб c_i^s .

Объекты нумеруют для каждого варианта системы по мере уменьшения ущерба, т.е. $c_1^s > c_2^s > \dots > c_n^s$. Установим пороговое допустимое значение ущерба A и будем считать, что в случае многократного НВ влиянию поддаются разные объекты и прежде всего объекты, влияние на которые приносит наибольший ущерб.

Ущерб для системы в целом составляет сумму ущербов от влияний на отдельные объекты. Показатель ущербов системы K_s рассчитывается таким образом:

$$K_s^A = \min_{C_s \geq A} K_s; \quad c_s = \sum_{i=1}^{K_s} c_i^s.$$

Если K_s^A превышает пороговое допустимое значение ущерба A , то система становится неработоспособной.

Пусть некоторая информационная система, имеющая базовую структуру S_0 , выполняет задание на протяжении некоторого времени t . После НВ структура системы может измениться, т.е. система будет иметь новую структуру S_i , с которой она может быть работоспособной, иначе говоря, структура принадлежит множеству работоспособных структур:

$$S^P = \{S_i, i = 1, \dots, N_p\}.$$

Или система может быть неработоспособной, в этом случае S_i принадлежит множеству неработоспособных структур, а именно

$$S^{NP} = \{S_i, i = N_{p+1}, \dots, N\}.$$

Таким образом, после НВ может образоваться или работоспособная процедура из множества S^P , или неработоспособная процедура из S^{NP} .

После n -кратного НВ система с соответствующей структурой должна начать выполнять задание, и время его выполнения должно составлять t .

Допустим, что каждый элемент системы может пребывать в одном из трех состояний: e_0 — элемент работоспособен и включен в работу; e_1 — элемент работоспособен, но не включен в работу по некоторым причинам; e_2 — элемент неработоспособен. Связи между элементами определены и стационарны во времени, т.е. в произвольный момент времени состояние элемента можно определить по состоянию работоспособности этого элемента и состояниям других элементов. Признаки работоспособности системы неизменны во времени и позволяют одновременно определять состояние системы по совокупности состояний ее элементов.

Существуют такие основные этапы анализа живучести системы на основе логико-вероятностной модели.

Описание состояний элементов. Для каждого элемента вводят две логические переменные: x_i — индикатор работоспособности i -го элемента:

$$x_i = \begin{cases} 1, & i\text{-й элемент работоспособен;} \\ 0, & i\text{-й элемент неработоспособен;} \end{cases}$$

y_i — индикатор состояния i -го работоспособного элемента:

$$y_i = \begin{cases} 1, & i\text{-й элемент работает;} \\ 0, & i\text{-й элемент не работает.} \end{cases}$$

Для отображения действия влияния на элементы вводят также индикаторы z_{ij} и $z_i = \bigcup_{(j)} z_{ij}$, где

$$z_{ij} = \begin{cases} 1, & \text{когда влияние } j\text{-го типа} \\ & \text{действует на } i\text{-й элемент;} \\ 0, & \text{когда влияние } j\text{-го типа} \\ & \text{не действует на } i\text{-й элемент.} \end{cases}$$

Далее определяются индикаторы трех состояний элемента, которые затем используются для составления логических зависимостей:

$$u_{i0} = x_i y_i \bar{z}_i;$$

$$u_{i1} = x_i \bar{y}_i \bar{z}_i;$$

$$u_{i2} = \bar{x}_i \vee x_i z_i.$$

Составление логических зависимостей. На основании предварительного анализа динамических моделей информационных процессов с учетом противодействия НВ, реконфигурации и управления можно составить систему логических уравнений, которую можно представить в векторной форме:

$$Y = f_Y(X, Y, Z).$$

Для информационной системы определяющим является состояние относительно небольшой группы самых весомых исходных элементов, однако из-за наличия опосредованных связей работоспособность системы определяется и состоянием всех других элементов.

Функцию работоспособности системы (ФРС) записывают в виде

$$F = \bigwedge_{(i)} f_i(X, Y, Z),$$

где f_i — логическая функция ($i = 1, \dots, N$) — индикатор выполнения i -й функции системы.

Решение системы логических уравнений. Рассмотренная выше система является линейной, которую можно решить, в частности, методами определителей, подстановок и т.п. [33].

Вероятностное описание элементов и внешних влияний. Каждый элемент в вероятностной модели является вероятностью $p_i = P(x_i = 1)$ того, что в заданный произвольный момент времени элемент x_i работоспособен. С появлением влияний $z_{ij} = 1$.

Стойкость i -го элемента к j -му возбуждению можно учитывать с помощью вероятности того, что элемент сохранит работоспособность

в случае появления возбуждения. Кроме того, задаются отдельные вероятности попадания элемента x_i в область действия j -го фактора НВ.

Преобразование ФРС к форме перехода к замещению. Различают формы перехода к частичному и полному замещению. Формы полного замещения: полная дизъюнктивная нормальная форма, дизъюнкция ортогональных неповторных форм, неповторная форма в базисе «И-НЕТ» и т.п. [33] После сведения к одной из этих форм выполняется одношаговое замещение логических переменных и операций на вероятности и арифметические операции.

Запись смешанной формы. Замещение неповторных переменных у преобразованной ФРС является, как правило, не полной, а частичной заменой, в результате которой некоторая логическая переменная и операция замещаются на вероятность и арифметическую операцию. Полученную таким образом ФРС называют смешанной формой, поскольку она содержит одновременно логические переменные и вероятности.

Определение показателей живучести. С помощью многошаговой процедуры замещения логических переменных в смешанных формах, которые составляются для базовой S_0 и структур S_i , находят вероятности работоспособных состояний в момент времени t — $P(t/S_0)$ и $P(t/S_i)$, а также условную функцию живучести $G_i(t)$:

$$G_i(t) = \frac{P(t/S_i)}{P(t/S_0)}.$$

После этого можно найти функцию выживания, безусловную функцию живучести и количество НВ (среднее), приводящее систему в неработоспособное состояние.

Определив функцию живучести как функцию выживания системы после n -кратного НВ (события A_n), усредненную по всем возможным структурам, $P_n(k)$ — усредненную вероятность возникновения структуры S_n после n -кратного НВ, можно записать

$$G(t/A_n) = G(t, n) = \sum_{k=1}^N P_n(k) G_k(t).$$

Безусловная функция живучести является усреднением по всем возможным событиям A_n функции выживания системы:

$$G(t) = \sum_{n=1}^{\infty} P(A_n)G(t / A_n) = \sum_{k=1}^N P(S_k)G_k(t),$$

где вероятность $P(S_k)$ определяется по формуле

$$P(S_k) = \sum_{n=1}^{\infty} P(A_n)P_n(k).$$

После этого можно найти функцию выживания, безусловную функцию живучести и количество НВ (среднее), приводящее систему в неработоспособное состояние.

Определим далее показатель выживания системы после n -кратного НВ (в случае наступления события A_n):

$$R(n) = 1 - Q(n) = P(F = 1 / A_n),$$

где $Q(n)$ — условный закон уязвимости: $Q(n) = P(F = 0 / A_n)$.

Среднее количество НВ, приводящее к потере системой работоспособности, находится по формуле

$$\bar{\omega} = \sum_{n=1}^{\infty} R(n),$$

а средний запас живучести, $\bar{d} = \bar{\omega} - 1$.

В свою очередь, среднее количество НВ, приводящее к невыполнению задания, можно определить так:

$$\bar{\omega}(t) = \sum_{n=1}^{\infty} R(n) = \sum_{n=1}^{\infty} n(G(t, n-1) - G(t, n)) = \sum_{n=1}^{\infty} G(t, n).$$

2.5. Оценка живучести системы по ее состоянию

Рассмотрим информационную систему, которая состоит из N элементов с произвольными связями между собой и функцией работоспособности $F_i = f(X)$, $X = \{X_1, X_2, \dots, X_N\}$. Система пребывает под действием потока независимых точечных НВ:

$$\varphi_{kj} = \frac{1}{N},$$

где φ_{kj} — вероятность того, что k -й элемент попадет в область действия j -го НВ, т.е. предполагается равновероятная уязвимость каждого элемента при НВ.

Оценим живучесть такой системы. Функцию выживания системы $R(n)$ после n -кратного НВ можно представить в виде

$$R(n) = \sum_{X \in X'} P(x / A_n) = P(F = 1 / A_n),$$

где X' — множество векторов X , которые соответствуют работоспособным состояниям системы. Вероятность $P(x / A_n)$ находится по формуле

$$P(x / A_n) = \sum_{\vec{n} \in M_n} P(\vec{n}) P(X / \vec{n}),$$

где $\vec{n} = (n_1, n_2, \dots, n_k)$ — вектор количества НВ, которые приходится на k подсистем; M_n — множество векторов, удовлетворяющих условию $n_1 + n_2 + \dots + n_k = n$. Вероятность $P(\vec{n})$ вычисляется по формуле

$$P(\vec{n}) = \frac{n!}{n_1! n_2! \dots n_k!} \gamma_1^{n_1} \gamma_2^{n_2} \dots \gamma_k^{n_k},$$

где γ_i — вероятность того, что i -я подсистема входит в область действия НВ.

При условии равномерности уязвимости элементов системы приведенные выражения можно упростить. Представим ФРС в виде ортогональной дизъюнктивной нормальной формы: $F = \bigvee_{i=1}^m Q_i$.

Соответственно,

$$R(n) = \sum_{i=1}^m P(Q_i = 1 / A_n).$$

Для случая равновероятного попадания элементов в область действия НВ, исходя из базовой структуры S_0 , определяют все возможные работоспособные структуры $S^p = \{S_i, i = 1, 2, \dots, N_p\}$, а функция выживания принимает вид

$$R(n) = \sum_{j=1}^{N_p} \frac{r_j(n)}{N^n},$$

где $r_j(n)$ — количество случаев, при которых возникает структура S_j после n -кратного НВ; $r_j(n)$ можно определить по формуле

$$r_j(n) = \sum_k L_{nk} B_{kj},$$

где L_{nk} — количество перестановок из n элементов k типов; B_{kj} — количество разных векторов X с k нулями, которые приводят к структуре S_j .

Числа L_{nk} могут быть вычислены непосредственно по формуле

$$L_{nk} = \sum_{i=1}^k C_k^i i^n (-1)^{k+1}.$$

В булевой модели можно предположить, что инфосюжет состоит из n элементов (документов), при этом i -му элементу соответствует бу-

лева переменная x_i , которая может принимать значения $\{0, 1\}$, т.е.

$$x_i = \begin{cases} 1, & \text{если элемент } i \text{ активен,} \\ 0, & \text{иначе.} \end{cases}$$

Принимая во внимание тот факт, что система динамическая, можно зафиксировать значение n как заведомо большое число, превышающее максимально наблюдаемое количество документов в информационных сюжетах. Несуществующим (недостающим) документам можно присвоить нулевые значения x_i .

Состояние информационной системы определяется структурной булевой функцией ее работоспособности (действенности), зависящей от переменных x_1, x_2, \dots, x_n :

$$S(x_1, x_2, \dots, x_n) = \begin{cases} 1, & \text{если система активна,} \\ 0, & \text{иначе.} \end{cases}$$

Если активность элемента информационной системы рассматривать как функцию времени, то состояние i -го документа можно рассматривать как случайный процесс $x_i(t)$, принимающий в произвольные моменты времени $t \geq 0$ значения 0 и 1. Для информационной системы определяется вероятность ее работоспособности по известным правилам [34, 35].

Среди недостатков булевых моделей можно назвать предположение только о двух состояниях компонентов — активности и неактивности. Кроме того, булевы модели не учитывают то, что весьма существенную роль может играть последовательность, в которой отказывают отдельные компоненты. Кроме того, в общем случае характер отказов отдельных компонент зависит от состояния других компонент. Это находится в противоречии с изначально предполагаемой независимостью элементов в булевой модели.

Информационную систему можно описать также марковской моделью. Пусть система имеет m возможных состояний. Обозначим множество состояний через $M = \{z_1, z_2, \dots, z_m\}$. Для любого фикси-

рованного момента времени $t \geq 0$ состояние системы $z_i(t)$ интерпретируется как случайная величина. Заданы множество всех состояний M , вектор распределения начальных вероятностей $p(0)$ и функция переходных вероятностей. Определяется вероятность актуальности информационной системы в заданный момент времени t (работоспособность системы) [34].

Применимость марковских моделей также имеет свои границы. Интенсивности переходов между отдельными состояниями системы могут быть нестационарными, принимаемые при расчете допущения относительно распределения интенсивности отказов могут значительно снизить точность полученных результатов; число состояний системы может быть так велико, что расчет становится практически невозможным.

Проведя оценку надежности компонент системы и получив общие показатели ее надежности, можно оценить ее живучесть на всех этапах их жизненного цикла. Существует несколько подходов к проведению оценки живучести, имеющих общий характер. Живучесть системы можно оценить относительно некоторого стандартного внешнего воздействия либо относительно множества внешних воздействий [36].

Пусть $E = \{e_i\}$ — множество деструктивных воздействий на информационную систему; $\sigma_j(e_i)$ — показатель эффективности (качества) функционирования j -го варианта информационной системы при воздействии e_i ; $H_j(E) = \min_{e_i \in E} \sigma_j(e_i)$ — показатель живучести информационной системы для множества возможных воздействий на нее E . Тогда при целенаправленном формировании информационной системы задача проектирования состоит в том, чтобы из множества вариантов информационных систем Ω найти такой, для которого выполняется

$$H_k(E) = \max_{X_j \in \Omega} \min_{e_i \in E} \sigma_j(e_i).$$

Кроме необходимости сохранения множества функций информационных систем при неблагоприятных для информационной системы воздействиях, часто ставится задача сохранения определенного уровня ее эффективности (актуальности, информативности).

Для количественной оценки живучести существуют многочисленные подходы, наиболее распространенный среди которых заключается в определении соотношения количества функциональных (работоспособных) состояний системы к общему возможному количеству состояний системы, возникающих при деструктивных воздействиях.

В качестве простого примера информационной системы рассмотрим информационный сюжет, состоящий из четырех документов ($n = 4$). Деструктивное воздействие на информационный сюжет — устранение из информационного пространства входящих в него документов. Причем первый из документов считается определяющим — его устранение из информационного пространства фактически ведет к потере информационной функциональности всего информационного сюжета. Остальные три документа считаются равноправными. Устранение любых двух из них ($k = 4$) также ведет к потере функциональности информационного сюжета.

Если обозначить состояние информационного сюжета 4-элементным кортежем, то множество неработоспособных состояний можно представить как объединение двух подмножеств, первое из которых соответствует состояниям с устраненным первым документом, а второе — с актуальным первым документом, но отсутствующими двумя другими.

Мощность первого подмножества составляет $2^{n-1} = 8$, перечислим его компоненты:

(0, 0, 0, 0)

(0, 0, 1, 0)

(0, 0, 0, 1)

(0, 0, 1, 1)

(0, 1, 0, 0)

(0, 1, 1, 0)

(0, 1, 0, 1)

(0, 1, 1, 1).

Мощность второго подмножества составляет $C_n^k + 1 = 4$, его компоненты:

(1, 0, 0, 0)

(1, 0, 1, 0)

(1, 0, 0, 1)

(1, 1, 0, 0).

Мощность всего множества состояний после деструктивного воздействия составляет $2^n - 1 = 15$.

Таким образом, живучесть G инфосюжета составляет

$$G = (15 - 8 - 4) / 15 = 0,2.$$

Рассмотрим случай, когда все состояния инфосюжета после деструктивного воздействия равнозначны, т.е. равновероятны.

Если состояния информационного сюжета не являются равновероятными, то живучесть G информационного сюжета составляет

$$G = \frac{\sum_{i=1}^m p_i - \sum_{i=1}^j p_i^{(0)} - \sum_{i=1}^l p_i^{(1)}}{\sum_{i=1}^m p_i} = 1 - \sum_{i=1}^j p_i^{(0)} - \sum_{i=1}^l p_i^{(1)},$$

где m — мощность всего множества состояний после деструктивного воздействия ($m = 2^n - 1$); j — мощность подмножества состояний с устраненным первым сюжетом; l — мощность подмножества состояний с актуальным первым сюжетом ($l = C_n^k + 1$); p_i — вероятность i -го состояния после деструктивного воздействия; $p_i^{(0)}$ — вероятность i -го состояния после деструктивного воздействия и устранения первого сюжета; $p_i^{(1)}$ — вероятность i -го состояния после деструктивного воздействия и сохранения первого сюжета, но при потере общей актуальности инфосюжета.

Приведенная модель позволяет оперировать с численными результатами, однако является слишком упрощенной. Переходя к более реалистичным подходам, следует отметить, что, характеризуя информационную систему, необходимо обращать внимание на такие ее свойства, как делимость или ослабление информационного сообщения, потока, воздействия на неоднородное информационное пространство, нестационарность этого воздействия, динамические изменения областей самого информационного пространства.

2.6. Оценка живучести по результатам выполнения задания

Для выполнения информационной системой общего задания и достижения системной цели каждый элемент должен функциониро-

вать в соответствии с установленным графиком, обеспечивая выполнение на протяжении времени t заданного набора функций. Пусть вероятность того, что i -й элемент выполняет определенные для него функции, составляет $p_i(t)$. В общем случае $p_i(t)$ может быть вероятностью выполнения задания со сложным режимом работы и сложными ограничениями функционирования.

Оценивая живучесть, будем считать, что все НВ завершились к началу интервала времени $(0, t)$.

Живучесть оценивается следующим образом. На основе логико-вероятностной модели записывается ФРС для базовой структуры. Затем выявляются все другие работоспособные структуры системы, подставляя в ФРС векторы, в которых один, потом два, три и больше элементов заменяются нулями. Если после подстановки логическая функция не становится тождественно равной нулю, то она соответствует одной из работоспособных структур. Одновременно во время таких испытаний определяются коэффициенты B_{ki} . Умножая B_{ki} на матрицу $\|L_{nk}\|$, находим матрицу коэффициентов $\|r_{ni}\|$.

Отношение r_{ni} / N^n дает вероятность $P_n(i)$ того, что после n -кратного НВ возникает структура $S_i \in S^p = \{S_k, k = \overline{1, N_p}\}$. Для каждой из структур S_i строится функция работоспособности $F^{(i)}(X)$, которая представляется в таком виде:

$$F^{(i)}(X) = x_n \left(\bigcup_{j=1}^l x_j^{f_j^{(i)}(X)} \right),$$

где x_n — переменная, которая соответствует полюсу системы; x_j — неповторяющиеся переменные; $f_j^i(X)$ — функции алгебры логики произвольного вида.

Далее переходим к расчету вероятностей:

$$P(F^{(i)}(X) = 1) = p_n \left(1 - \prod_{j=1}^l q_j^{f_j^{(i)}(X)} \right),$$

где $p_n = p(x_n = 1)$, $q_j = 1 - p_j = p(x_j = 0)$.

После этого выполняется последовательное замещение остальных логических переменных и получение функции $P(t/S_i)$. Потом определяется $G_i(t)$ и $G(t, n)$. Если удастся определить вероятность $P(A_n)$, то можно найти безусловную функцию живучести.

2.7. Оценка живучести по цели функционирования

Более общую оценку живучести информационной системы можно построить, исходя из цели ее функционирования, множества задач информирования $Q = \{q_1, q_2, \dots, q_m\}$ и множества компонент (документов). Действительно, любая задача $q_i \in Q$, $i = \overline{1, m}$, характеризуется набором элементарных функций (информирования об отдельных аспектах) $F_i = \{f_{j_1}, f_{j_2}, \dots, f_{j_k}\}$, $1 \leq j_k \leq n$, из которых строятся решения этой задачи.

Обозначим через $F = \bigcup_{i=1}^m F_i$ множество наборов элементарных функций информационной системы. Для каждой задачи информирования q_i задается характеристика эффективности решения. Введем функцию потенциальных возможностей функциональных модулей $\varphi: \{1, 2, \dots, p\} \rightarrow P(F)$, где $P(F)$ — множество всех подмножеств F .

Для характеристики возможных конфигураций информационной системы введем матрицу потенциальных возможностей системы:

$$a_{ij} = \begin{cases} 1, & \text{если } f_i \in \varphi(j), \\ 0, & \text{если } f_i \notin \varphi(j), j = \overline{1, p}, i = \overline{1, n}. \end{cases}$$

Текущую конфигурацию информационной системы будем характеризовать тем, на выполнение каких информационных функций

нацелен каждый модуль. Введем двоичную матрицу B размерности $n \times p$ — матрицу текущей конфигурации системы, такую, что

$$b_{ij} = \begin{cases} 1, & \text{если модуль } S_j \text{ выполняет функцию } f_i, \\ 0, & \text{в противном случае.} \end{cases}$$

Определим функцию эффективности модулей $\varphi_{эф}: I_s \times I_f \times B \rightarrow T$, где $I_s = \{1, 2, \dots, p\}$ — множество индексов модулей; $I_f = \{1, 2, \dots, n\}$ — множество индексов элементарных функций; B — множество матриц конфигураций; T — числовое множество количественных мер эффективности (например, размер аудитории, читающей документы инфосоюжета, и т.п.) Если $\varphi_{эф}(i, j, B) = t_{ij}$, то в конфигурации, определенной матрицей B , модуль S_j выполняет функцию f_i с эффективностью t_{ij} , $\varphi_{эф}(i, j, B) = 0$, если модуль S_j не выполняет функцию f_i .

Для характеристики информационной системы введем понятие характеристического вектора состояния — n -мерного вектора (n — мощность множества элементарных функций системы). Начальной конфигурации информационной системы при условии, что выполняется все множество функций F , будет соответствовать характеристический вектор состояния $(0, 0, \dots, 0)$. Некоторой текущей конфигурации информационной системы будет соответствовать характеристический вектор $(d_1, d_2, \dots, d_i, \dots, d_n)$, где d_i — число «отказов» функции $f_i \in F$. Под «отказом» функции $f_i \in F$ понимается невозможность выполнения функции информирования f_i , т.е. d_i — количество реконфигураций информационной системы из-за «отказа» функции информирования $f_i \in F$.

Решим задачу нахождения множества характеристических векторов состояний информационной системы, в которых реализуется конфигурация, обеспечивающая выполнение цели функционирования.

Мощность этого множества также может служить мерой живучести системы.

Поставленную задачу можно решить в два этапа.

1. Нахождение множества характеристических векторов состояний информационной системы S_f , определяющих состояния, в которых возможен выбор конфигурации, обеспечивающей выполнение множества элементарных функций F . Пусть некоторая начальная информационная система характеризуется матрицей B_0 . Первоначальную конфигурацию можно построить, исходя, например, из следующих предположений: каждый модуль выполняет только одну функцию, и каждая функция выполняется только одним модулем, т.е.

$$\sum_{j=1}^p b_{ij} = 1, \forall i = \overline{1, n},$$

$$\sum_{i=1}^n b_{ij} = 1, \forall j = \overline{1, p}.$$

В качестве критерия оптимизации естественно выбрать

$$\Phi = \sum_{i=1}^n \sum_{j=1}^p b_{ij} \varphi_{эф}(i, j, B_0) \rightarrow \min(\max).$$

В зависимости от конкретного смысла функции эффективности справедлива задача нахождения либо максимума Φ , либо его минимума.

Приведенные выше выражения описывают задачу комбинаторного типа, которую можно решить, например, венгерским методом или с помощью эвристического алгоритма.

Предположим, что возникают «отказы» функций, т.е. изменяется состояние информационного объекта. Новому состоянию, с учетом имевших место отказов, соответствует характеристический вектор состояния $(d_1, d_2, \dots, d_i, \dots, d_n)$. Восстановление осуществляется за счет перераспределения функций между ее модулями. Задачу нахождения новой конфигурации системы можно описать в следующей постановке:

$$\Phi = \sum_{i=1}^n \sum_{j=1}^p b_{ij} \varphi_{\Phi}(i, j, B) = \max(\min),$$

$$\sum_{i=1}^n b_{ij} = 1, \forall j = \overline{1, p},$$

$$\sum_{j=1}^p b_{ij} = 1, \forall i = \overline{1, n},$$

$$\Phi \geq \Phi^* (\Phi \leq \Phi^*).$$

где Φ^* — величина, определяющая минимально (максимально) допустимую эффективность.

Искомое множество S_f включает лишь характеристические векторы, для которых разрешима приведенная задача. На этом этапе решения из множества характеристических векторов состояния информационной системы S_f выделим подмножество S_q , определяющее состояния системы, в которых возможен выбор конфигураций, обеспечивающих выполнение цели функционирования.

В качестве оценки живучести информационной системы можно взять мощность множества S_q . В случае информационных систем на первое место выходит проблема информирования относительно их различных аспектов независимо от наличия или отсутствия неблагоприятных факторов. В связи с этим в качестве количественного критерия оценки живучести целесообразно использовать отношение количества функций, выполняемых системой при наличии определенных неблагоприятных воздействий либо множества таких воздействий, к общему количеству функций системы, с учетом критичности выполняемых и невыполняемых функций. Критичность каждой конкретной функции определяется индивидуально для каждой конкретной информационной системы, исходя из его специфики. Количественный показатель живучести конкретной информационной системы в заданных условиях можно вычислять по формуле: $S = \sum_{i \in \Delta} \alpha_i / \sum_{j \in \Theta} \alpha_j$, где Θ

— множество всех функций информирования; Δ — множество функций информационной системы, выполняемых в заданных условиях

$(\Delta \subseteq \Theta)$; α_n — критичность n -й функции. Таким образом, количественная оценка живучести информационной системы будет изменяться в интервале $[0, 1]$. Живучесть тем выше, чем больше ее количественная оценка.

2.8. Энтропийный подход к оценке живучести

Информационная система может находиться в самых разнообразных допустимых для нее состояниях. Система работоспособна, пока ее структура или организация позволяют избегать деструктивных воздействий либо их локализовать, т.е. живучесть системы зависит от ее поведения в пространстве состояний.

Пусть множество $\{X \langle m \rangle\}$ образует пространство допустимых состояний информационной системы. В общем случае $X \langle m \rangle$ может быть случайной функцией, зависящей от времени: $X \langle m \rangle = X \langle m \rangle(t) = \{x_1(t), \dots, x_m(t)\}^T$.

Обозначим через $F\lambda$ оператор деструктивного воздействия с параметром λ . Если деструктивное воздействие сказывается на информационной системе, находящейся в состоянии $X^j \langle m \rangle(t) \in \{X \langle m \rangle(t)\}_n$, $I = 1, \dots, n$ и, например, $F\lambda(X^j \langle m \rangle(t)) \in \{X \langle m \rangle(t)\}_n$, то можно говорить о том, что система «выжила» после воздействия с параметром λ . Если информационная система так организована, что $\forall \lambda: F\lambda(X^j \langle m \rangle(t)) \in \{X \langle m \rangle(t)\}_n$, то уровень живучести системы очень высок. Таким образом, система имеет тем больший уровень живучести, чем больше мощность множества допустимых состояний (разнообразия) [37].

Эти требования соответствуют закону У. Эшби, согласно которому разнообразие допустимых состояний уменьшает разнообразие недопустимых. Кроме того, что уровень живучести системы зависит также от ее стремления удержаться во множестве допустимых состояний [38], т.е. от ее устойчивости.

Любой объект или систему можно рассматривать как множество, обладающее разнообразием. Изменение этого разнообразия соот-

ветствует изменению состояний системы, т.е. состава ее элементов, структуры или поведения. Таким образом, множеству состояний системы можно сопоставить обладающее эквивалентными свойствами множество вероятностей этих состояний.

Мера разнообразия состояний системы напрямую связана с живучестью. В качестве такой меры разнообразия в настоящее время широко применяется энтропия состояния системы [39–41]. Известно, что энтропия является мерой неопределенности состояния системы, мерой недостатка информации о действительной ее структуре и поведении. В общем случае известно, что распознавание системы имеет наименьшую трудность, если энтропия состояния этой системы минимальна. Кроме того, для максимизации сложности распознавания объекта на фоне среды необходимо максимизировать энтропию состояния объекта до уровня энтропии среды.

Если использовать энтропию в качестве величины, характеризующей уровень живучести системы, то она, учитывая общие свойства энтропийной функции, обладает следующими свойствами:

— энтропийная характеристика позволяет учесть неопределенность (в структуре и поведении) состояний системы и при этом выражается через вероятностные характеристики системы;

— энтропийная характеристика системы зависит от размерности пространства состояний системы, числа и разнообразия элементов системы;

— энтропийная характеристика не зависит от выбора начала координат в пространстве состояний системы.

В некоторых случаях целесообразно измерять живучесть как относительную величину:

$$G = \frac{H}{H_{\max}},$$

где H_{\max} — максимально возможная энтропия системы (далее — рекорд), характеризующая предельное состояние неупорядоченности.

Таким образом, приведенный показатель живучести представляет собой отношение энтропии состояния системы к максимально возможной энтропии этой системы в условиях наложенных на нее ограничений.

Для всех рассмотренных к настоящему времени классов систем H_{\max} существует и единственная, хотя ведутся теоретические работы по поиску таких систем, для которых бы не существовало H_{\max} .

Очевидно, что определение показателя уровня живучести должно производиться по такой схеме:

— определяется способ оценивания значения энтропии H в зависимости от типа системы;

— вычисляется значение H анализируемой системы;

— анализируются ограничения, наложенные на систему;

— в рамках этих ограничений синтезируется система с параметрами, обеспечивающими H_{\max} ;

— определяется значение показателя G .

Следовательно, перед тем как оценить значение показателя живучести, необходимо решить задачу синтеза этой системы по критерию максимума неопределенности в рамках имеющихся ограничений. В некоторых случаях реализация принципа максимума неопределенности обеспечивает наименьшие затраты по сравнению с другими способами действий и тем самым повышает живучесть.

Применение показателя живучести предполагает возможность оценивания живучести, т.е. знание влияния всех составляющих системы на сам показатель живучести и их взаимное влияние друг на друга в процессе функционирования и эволюции системы.

Справедливо утверждение о том, что уровень живучести системы G остается постоянным тогда и только тогда, когда относительное приращение энтропии равно относительному приращению рекорда. В работе [42] показано, что при эволюции системы уровень ее живучести не изменяется, если $H = H_{\max}$. Там также показано, что при реконструкции или развитии системы ее живучесть будет расти тогда и только тогда, когда относительное приращение энтропии системы больше относительного приращения рекорда. Эти два утверждения объединяются в одно: живучесть системы не убывает при эволюции системы, если относительное приращение энтропии системы не меньше относительного приращения рекорда системы:

$$\frac{H'}{H} \geq \frac{H'_{\max}}{H_{\max}}.$$

Таким образом, у системы, обладающей тенденцией к росту уровня живучести, ее энтропия в процессе эволюции всегда стремится к максимально возможной.

Рассмотрим более конкретные свойства показателя живучести различных систем.

Предположим, что система организована так, что $H = \text{const}$ всегда. Следовательно, для выполнения приведенного выше неравенства необходимо, чтобы выполнялось условие $H'_{\max} < 0$.

Таким образом, при $H = \text{const}$ у системы, стремящейся повысить свою живучесть в процессе эволюции, ресурсы должны использоваться на снижение рекорда, его приближение к H .

Пусть для некоторой системы $H = 0$. Множество возможных состояний системы $\{X_{\langle m \rangle}\}_n = \{x_{\langle 1 \rangle}, \dots, x_{\langle m \rangle}\}_n^T$ и множество недопустимых $\{X_{\langle m \rangle}\}_k = \{x_{\langle 1 \rangle}, \dots, x_{\langle m \rangle}\}_k^T$, составляют вместе полную группу событий; в результате можно, как уже отмечалось, сопоставить вероятные меры $P(\{X_{\langle m \rangle}\}_n)$ и $P(\{X_{\langle m \rangle}\}_k)$.

Рекорд системы можно определить как $H_{\max} = -P(\{X_{\langle m \rangle}\}_n) \times \ln P(\{X_{\langle m \rangle}\}_n) - P(\{X_{\langle m \rangle}\}_k) \cdot \ln P(\{X_{\langle m \rangle}\}_k)$.

Необходимо выполнить условие $H'_{\max} < 0$, что приведет к минимизации H_{\max} , так как по предположению $H = 0$. Из основных свойств энтропии [38] следует, что $H_{\max} = 0$ тогда и только тогда, когда все вероятности, кроме одной, равны нулю, а эта единственная равна 1.

Таким образом, при $H = \text{const}$ и $H_{\max} > 0$ живучесть системы со временем будет возрастать, если при эволюции системы будет происходить перераспределение вероятностей всех состояний системы, при котором только одно состояние, причем с минимальной вероятностью, соответствует положению системы, отождествленному с потерей работоспособности.

2. Пусть $H_{\max} = \text{const}$. Для выполнения приведенного выше неравенства необходимо, чтобы $H' > 0$, т.е. при $H_{\max} = \text{const}$ у системы,

стремящейся повысить свою живучесть в процессе эволюции, энтропия должна возрасть, стремиться к рекорду.

3. Если эволюция системы протекает так, что $H' > 0$ и $H'_{\max} > 0$ то, как легко видеть, этот случай является комбинацией двух предыдущих.

4. Если $H' > 0$, поскольку $\lim H = H_{\max}$, уровень живучести которой не убывает в процессе эволюции, должно выполняться условие

$$\frac{H'}{H'_{\max}} > 1.$$

Таким образом, энтропия системы в этом случае должна расти быстрее, чем максимально возможная энтропия.

5. Если система такова, что $H' < 0$, то необходимо, чтобы при этом выполнялось условие $H'_{\max} < 0$, что также приводит к общему условию:

$$\frac{H'}{H'_{\max}} > 1.$$

Таким образом, в данном случае энтропия системы должна убывать медленнее, чем рекорд. Это требование соответствует условию обеспечения энтропийной устойчивости динамических систем, введенному В.Л. Стротановичем.

Рассмотрим систему дифференциальных уравнений, описывающих динамику системы:

$$\frac{dx_j}{dt} = f_j(x_1, \dots, x_m, t), \quad j = 1, \dots, m.$$

Предположим, что функции $f_j(x_1, \dots, x_m, t)$ — непрерывные в некоторой открытой области и являются дифференцируемыми функциями своих аргументов. При случайных начальных условиях частное решение приведенной системы будет представлять собой случайные

функции времени. В соответствии с [39, 42] приведенная система обладает общей монотонной энтропийной устойчивостью, если при произвольном начальном законе распределения координат общая энтропия системы монотонно убывает с течением времени. Известно, что необходимым и достаточным условием общей монотонной энтропийной устойчивости такой системы является выполнение условий

$$\frac{dH}{dt} = \sum_{j=1}^m M \left[\frac{\partial f_j(x_1, \dots, x_m, t)}{\partial x_j} \right] < 0, dt = x \text{ при произвольном начальном}$$

законе распределения. Полученное условие можно рассматривать применительно к оценке живучести системы, описываемой приведенными выше дифференциальными уравнениями.

Предложенный формальный подход подтвердил тот факт, что система, реагирующая на воздействие деструктивных факторов по заранее определенному предписанию, сверхчувствительна к малейшим отклонениям условий функционирования от предусмотренных и не может обладать тем самым высоким уровнем живучести.

Определено, что уровень живучести системы при воздействии деструктивных факторов зависит от мощности множества допустимых состояний системы и разнообразия этого множества и в целом соответствует закону необходимого разнообразия У. Эшби. Исходя из этого, в качестве показателя уровня живучести допустимо использование отношения энтропии состояния системы и максимально возможной энтропии при наложенных на систему ограничениях. При этом живучесть, которая определяется данным показателем, не убывает при эволюции системы, если относительное приращение энтропии системы не меньше относительного приращения рекорда.

Кроме того, в [37] определена тесная связь энтропии с традиционными показателями устойчивости систем. Для использования предложенного подхода и оценивания уровня живучести необходимо оценить энтропию анализируемой системы и в рамках существующих на эту систему ограничений синтезировать систему, обладающую рекордом, и оценить последний.

3. СЛОЖНЫЕ ИНФОРМАЦИОННЫЕ СЕТИ

Информационные системы могут быть представлены как сетевые структуры, так называемые динамические сети [43]. Текущее состояние информационной системы может быть представлено в виде графа $\langle M, L \rangle$, где M — это множество компонент (например, документов) информационной системы, а L — множество ребер, например, связей подобия, цитирования, ссылок и т.д. Свойство живучести напрямую связано с такими свойствами графов, как связность, кластерность, средний кратчайший путь между вершинами и т.п.

В настоящее время наряду с традиционными теориями графов, систем и сетей массового обслуживания активно развивается теория сложных сетей (англ. — *Complex Networks*) [44], в рамках которой предлагаются подходы к решению вычислительно сложных задач, характерных для современных сетей.

Основной причиной актуальности теории сложных сетей являются результаты современных работ по описанию реальных компьютерных, биологических и социальных сетей. Такие сети имеют характеристики, не свойственные сетям с равновероятной связностью узлов, а строятся на основе связанных структур, степенных распределений и узлов-концентраторов.

Представляющие интерес сети чаще всего разрежены — присутствует лишь малая часть возможных ребер, соединяющих отдельные узлы. Поэтому сегодня особую актуальность приобретают методы работы с разреженными матрицами.

Действительно, практически все современные сети можно считать сложными. Например, известная задача синтеза топологии сети допускает комбинаторный подход, опирающийся на представление сети в виде конечного графа без петель и кратных ребер, вершины которого соответствуют узлам сети, а ребра — линиям связи.

Вместе с тем, использование методов перечисления графов для решения задачи топологической оптимизации считается неперспективным, так как необходимо исследовать огромное количество возможных вариантов соединения узлов линиями связи. Например, в сети из 10 узлов существует 2^{45} вариантов размещения линий связи (для 10 узлов теоретически возможно $C_{10}^2 = \frac{10 \cdot 9}{2} = 45$ линий соединений).

Каждая из этих возможных линий связи может реально существовать

— состояние «1», или не существовать — состояние «0», т.е. всего возможностей 2^{45}).

Для меньшего количества узлов (например, $n = 3$) линии связи, могут быть реально посчитаны ($2^{\frac{3 \cdot 2}{2}} = 8$) вариантами (рис. 9).

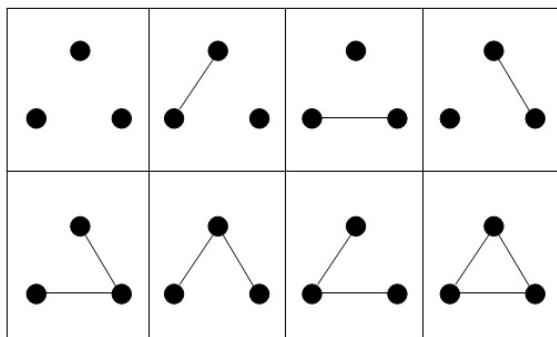


Рис. 9. Варианты размещения линий связи при $n = 3$

Сложные сети обычно рассматриваются в абстрактном пространстве, в котором расположение вершин не имеет значения. Для некоторых реальных типов сетей такое рассмотрение оправдано.

Однако, существует множество систем, в которых расположение компонент весьма важно, поскольку влияет на эволюцию сети. Такие сети называются географическими или пространственными. В географических сетях существование прямого соединения между вершинами может зависеть от многих ограничений, таких как расстояние между ними, географический рельеф, территориальные ограничения и т.д. Модели, предназначенные для представления таких сетей, должны учитывать эти ограничения.

3.1. Параметры сложных сетей и задачи живучести

Теория сложных сетей как область дискретной математики изучает характеристики сетей, учитывая не только их топологию, но и статистические феномены, распределение весов отдельных узлов и ребер, эффекты протекания, просачивания, проводимости в таких сетях тока, жидкости, информации и т.д. Оказалось, что свойства мно-

гих реальных сетей существенно отличаются от свойств классических случайных графов. Изучения таких параметров сложных сетей, как кластерность, посредничество или уязвимость, напрямую относятся к теории живучести, так как именно от этих свойств зависит способность сетей сохранять свою работоспособность при деструктивном воздействии на их отдельные узлы или ребра (связи).

Несмотря на то, что в рассмотрение теории сложных сетей попадают различные сети — электрические, транспортные, информационные, наибольший вклад в развитие этой теории внесли исследования социальных сетей. Термин «социальная сеть» обозначает сосредоточение социальных объектов, которые можно рассматривать как сеть (или граф), узлы которой — объекты, а связи — социальные отношения. Этот термин был введен в 1954 году социологом из «Манчестерской школы» Дж. Барнсом (J. Barnes) в работе «Классы и сборы в норвежском островном приходе». Во второй половине XX столетия понятие «социальная сеть» стало популярным у западных исследователей, при этом в качестве узлов социальных сетей стали рассматривать не только представителей социума, но и другие объекты, которым присущи социальные связи. В теории социальных сетей получило развитие такое направление, как анализ социальных сетей (*Social Network Analysis, SNA*). Сегодня термин «социальная сеть» обозначает понятие, оказавшееся шире своего социального аспекта. Оно включает, например, многие информационные сети, в том числе и веб-пространство или социальные интернет-сети.

В рамках теории сложных сетей рассматривают не только статистические, но и динамические сети, для понимания структуры которых необходимо учитывать принципы их эволюции [45].

В теории сложных сетей выделяют три основных направления: исследование статистических свойств, которые характеризуют поведение сетей; создание модели сетей; предсказание поведения сетей при изменении структурных свойств. В прикладных исследованиях обычно применяют такие типичные для сетевого анализа характеристики, как размер сети, сетевая плотность, степень центральности и т.п.

О «структуре сообщества» в сложной сети можно говорить тогда, когда существует фрагмент сети — группа узлов, которые имеют высокую плотность ребер между собой, притом, что плотность ребер между отдельными фрагментами — низкая. Традиционный метод для выявления структуры сообществ — кластерный анализ. Существуют

десятки приемлемых для этого методов, которые базируются на разных мерах расстояний между узлами, взвешенных путевых индексах между узлами и т.п. В частности, для больших социальных сетей наличие структуры сообществ оказалось неотъемлемым свойством.

К потере живучести информационной системы может привести разрыв связей между ее компонентами, например, при устранении из информационного пространства наиболее весомых компонент, т.е. таких, которые имеют, допустим, наибольший коэффициент посредничества (*betweenness*). Этот коэффициент для конкретного узла сети определяется как сумма по всем парам узлов сети соотношений количества кратчайших путей между ними, проходящими через заданный узел, к общему количеству кратчайших путей между ними.

При анализе сложных сетей, как и в теории графов, исследуются параметры отдельных узлов, параметры сети в целом, сетевые подструктуры.

3.1.1. Параметры узлов сети

Для отдельных узлов выделяют следующие параметры:

- входная степень связности узла — количество ребер графа, которые входят в узел;
- выходная степень связности узла — количество ребер графа, которые выходят из узла;
- расстояние от данного узла до каждого из других;
- среднее расстояние от данного узла до других;
- эксцентricность (*eccentricity*) — наибольшее из геодезических расстояний (минимальных расстояний между узлами) от данного узла к другим;
- посредничество (*betweenness*), показывающее, сколько кратчайших путей проходит через данный узел;
- центральность — общее количество связей данного узла по отношению к другим;
- уязвимость, рассматриваемая как уровень спада производительности сети в случае удаления вершины и всех смежных ей ребер.

Степень связности k_i узла i — это количество ребер, соединенных с этой вершиной.

Соответственно, средняя степень всей сети рассчитывается как среднее всех k_i для всех узлов сети.

Как отмечено выше, в случае ориентированных сетей имеется две разновидности степени связности узла: выходная, соответствующая количеству исходящих из данного узла ребер, и входная, равная количеству входящих в данный узел ребер.

3.1.2. Общие параметры сети

Для расчета индексов сети в целом используют такие параметры: число узлов, число ребер, геодезическое расстояние между узлами, среднее расстояние от одного узла к другим, плотность — отношение количества ребер в сети к возможному максимальному количеству ребер при данном количестве узлов, количество симметричных, транзитивных и циклических триад, диаметр сети — наибольшее геодезическое расстояние в сети, уязвимость, рассчитываемая как максимальная уязвимость всех вершин сети, ассортативность как мера корреляции между степенями узлов и т.д.

Существует несколько актуальных задач исследования сложных сетей с точки зрения живучести, среди которых можно выделить следующие основные:

- определение фрагментов сети (клик, кластеров), в которых узлы связаны между собой сильнее, чем с членами других подобных фрагментов;
- выделение фрагментов сети (компонент связности), которые связаны внутри и не связаны между собой;
- нахождение перемычек, т.е. узлов, при изъятии которых сеть распадается на несвязанные части.

3.1.3. Распределение степеней связности узлов

Важной характеристикой сети является функция распределения степеней узлов $P(k)$, которая определяется как вероятность того, что узел i имеет степень $k_i = k$, т.е. распределение степеней $P(k)$ отражает долю вершин со степенью k .

Для ориентированных сетей существует распределение выходящей полустепени $P^{out}(k^{out})$, и полустепени входной $P^{in}(k^{in})$, а также распределение общей степени $P^{io}(k^{in}, k^{out})$. Последнее задает вероятность нахождения узла с входной полустепенью k^{in} и выходной

полустепеню k^{out} .

Сети, характеризующиеся разными $P(k)$, демонстрируют весьма разное поведение. $P(k)$ в некоторых случаях может быть распределением Пуассона ($P(k) = e^{-m} m^k / k!$, где m — математическое ожидание, экспоненциальным ($P(k) = e^{-k/m}$) или степенным ($P(k) \sim 1/k^\gamma$, $k \neq 0$, $\gamma > 0$).

Важной особенностью многих реальных сетей является распределение степеней узлов $P(k)$ по степенному закону.

Сети со степенным распределением степеней связности узлов называются безмасштабными (*scale-free*). Именно безмасштабные распределения часто наблюдаются в реально существующих сложных сетях. При степенном распределении возможно существование узлов с очень высокой степенью, что практически не наблюдается в сетях с пуассоновым распределением.

3.1.4. Путь между узлами

Если два узла i и j можно соединить с помощью последовательности из m ребер, то такую последовательность называют маршрутом (*walk*) между узлами i и j , а m называют длиной маршрута.

Говорят, что узлы i и j связны, если существует маршрут между ними. Отношение связности транзитивно, т.е. если узел i связан с узлом j , а j связан с k , то i связан с k . При этом маршрут, у которого начало и конец находятся в одном и том же узле, причем все остальные вершины используются ровно один раз, называется циклом.

Расстояние между узлами определяется как длина маршрута от одного узла до другого. Естественно, узлы могут быть соединены прямо или опосредованно. Путем между узлами d_{ij} назовем кратчайшее расстояние между ними. Для всей сети можно ввести понятие среднего пути как среднего по всем парам узлов кратчайшего расстояния между ними:

$$l = \frac{2}{n(n+1)} \sum_{i \geq j} d_{ij},$$

где n — количество узлов; d_{ij} — кратчайшее расстояние между узлами i и j .

Венгерскими математиками П. Эрдешем (P. Erdős) и А. Реньи (A. Rényi) было показано, что среднее расстояние между двумя вершинами в случайном графе растет как логарифм от числа вершин [46, 47].

На практике живучесть сети связи определяют как вероятность наличия пути между любой парой узлов.

Некоторые сети могут оказаться несвязными, т.е. в них найдутся узлы, расстояние между которыми является бесконечным. Соответственно, средний путь может оказаться также равным бесконечности. Для учета таких случаев вводится понятие глобальной эффективности сети как среднего инверсного пути между узлами, рассчитываемое по формуле

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}},$$

где сумма учитывает все пары узлов. Эта характеристика отражает эффективность сети при пересылке информации между узлами (предполагается, что эффективность в пересылке информации между двумя узлами i и j обратно пропорциональна расстоянию между ними).

Обратная величина глобальной эффективности — среднее гармоническое геодезических расстояний:

$$h = \frac{1}{E}.$$

Так как данная формула снимает проблему расхождения при определении среднего пути, то эта характеристика лучше подходит для графов с несколькими компонентами связности.

Эффективное расстояние между двумя узлами в общем случае больше, чем кратчайшее расстояние.

Сети также характеризуются таким параметром, как диаметр или максимальный по длине путь, т.е. путь, равный максимальному значению из всех d_{ij} .

3.1.5. Коэффициент кластерности

Д. Уаттс (D. Watts) и С. Стрoгaтц (S. Strogatz) в 1998 году определили такой параметр сетей, как коэффициент кластерности [48], который соответствует уровню связности узлов в сети. Этот коэффициент характеризует тенденцию к образованию групп взаимосвязанных узлов, так называемых клик (*clique*). Кроме того, для конкретного узла коэффициент кластерности показывает, сколько ближайших соседей данного узла являются также ближайшими соседями друг для друга.

Коэффициент кластерности для отдельного узла сети определяется следующим образом. Пусть из узла выходит k ребер, которые соединяют его с k другими узлами, ближайшими соседями. Если предположить, что все ближайшие соседи соединены непосредственно друг с другом, то количество ребер между ними составляло бы $\frac{1}{2}k(k-1)$, т.е. это число, которое соответствует максимально возможному количеству ребер, которыми могли бы соединяться ближайшие соседи выбранного узла. Отношение реального количества ребер, которые соединяют ближайших соседей данного узла, к максимально возможному (такому, при котором все ближайшие соседи данного узла были бы соединены непосредственно друг с другом) называется коэффициентом кластерности узла $i - C(i)$. Естественно, эта величина не превышает единицы.

Существует еще один способ вычисления коэффициента кластерности (транзитивности), базирующийся на такой формуле:

$$C = \frac{3N_{\Delta}}{N_3},$$

где N_{Δ} — количество 3-циклов в сети, а N_3 — количество связных 3-компонент.

3-цикл определяется при этом как множество трех узлов с ребрами между каждой парой узлов. Связная 3-компонента — множество, состоящее из трех узлов, в котором каждый узел достижим из другого узла, непосредственно или опосредованно. Таким образом, в 3-компоненте центральный узел должен быть инцидентен двум другим. Множитель 3 введен из учета вариантов различных 3-компонент для каж-

дого 3-цикла, этот множитель обеспечивает выполнение неравенства $0 \leq C \leq 1$. Тогда мы получаем

$$N_{\Delta} = \sum_{k>i>j} a_{ij} a_{ik} a_{jk};$$

$$N_3 = \sum_{k>i>j} (a_{ij} a_{ik} + a_{ji} a_{jk} + a_{ki} a_{kj}),$$

где a_{ij} — элементы матрицы смежности A , соответствующей сети, сумма берется по всем компонентам различных узлов i , j и k только один раз.

Коэффициент кластерности может определяться как для каждого узла, так и для всей сети. Соответственно, уровень кластерности всей сети определяется как нормированная по количеству узлов сумма соответствующих коэффициентов отдельных узлов.

Разница между двумя подходами к определению кластерности состоит в том, что, усреднив по вершинам, мы получаем во втором случае одинаковое влияние для каждого треугольника в сети, а в первом случае учитывается равный взнос для каждого узла.

Это приводит к разным значениям коэффициента кластерности, потому что узлы с большими степенями с большей вероятностью входят в состав большего количества треугольников, чем вершины с меньшими степенями.

Рассмотренный ниже феномен «малых миров» непосредственно связан с уровнем кластерности сети.

3.1.6. Посредничество

Значение узла для сети тем больше, чем в большем количестве путей он задействован. Поэтому, полагая, что обмен данными происходит по кратчайшим путям между двумя вершинами, можно измерить количественно значение узла с точки зрения посредничества (*betweenness*), определяемого количеством кратчайших путей, проходящих через узел. Эта характеристика отражает роль данного узла в установлении связей в сети. Узлы с наибольшим посредничеством играют главную роль в установлении связей между другими узлами в сети. Посредничество b_m узла m определяется по формуле

$$b_m = \sum_{i \neq j} \frac{B(i, m, j)}{B(i, j)},$$

где $B(i, j)$ — общее количество кратчайших путей между узлами i и j ; $B(i, m, j)$ — количество кратчайших путей между узлами i и j , проходящих через узел m .

Если учитывать, что кратчайшие пути могут быть неизвестны, и вместо этого для навигации в сети используются поисковые алгоритмы, то посредничество (промежуточная центральность) узла может быть выражена вероятностью его нахождения поисковым алгоритмом.

Уровень преобладания наибольшего посредника в этом случае определяется в соответствии с формулой:

$$CPD = \frac{1}{n-1} \sum_i (B_{\max} - B_i),$$

где B_{\max} — самое большое в сети значение уровня посредничества.

Преобладание центрального узла будет равно 0 для клики и 1 для звезды, в которой центральный узел входит во все пути.

3.1.7. Эластичность и уязвимость сети

Противоположные свойства эластичности и уязвимости сетей относятся к распределению расстояний между узлами при изъятии отдельных узлов. Эластичность сети зависит от ее связности, т.е. существования путей между парами узлов. Если узел будет изъят из сети, типичная длина этих путей увеличится. Если этот процесс продолжать достаточно долго, сеть перестанет быть связной. Р. Альберт (Reka Albert) из университета штата Пенсильвания (США) при исследовании атак на интернет-серверы изучала эффекты, возникающие при изъятии узла из сети, представляющей собой подмножество WWW из 326000 страниц [49].

Среднее расстояние между двумя узлами, как функция от количества изъятых узлов, почти не изменилось при случайном удалении узлов (высокая эластичность). Вместе с тем целенаправленное удаление узлов с наибольшим количеством связей приводит к разрушению

сети. Таким образом, Интернет является высокоэластичной сетью по отношению к случайному отказу узла в сети, но высокочувствительной к намеренной атаке на узлы с высокими степенями связей с другими узлами.

Один из способов найти критичные компоненты сети — поиск самых уязвимых узлов [50]. Если производительность сети связана с ее глобальной эффективностью, уязвимость узла может быть определена как спад производительности в случае удаления узла и всех смежных ему ребер из сети:

$$V_i = \frac{E - E_i}{E},$$

где E — глобальная эффективность исходной сети, а E_i — глобальная эффективность после удаления узла i и всех смежных ему ребер.

Упорядоченное распределение узлов относительно их уязвимостей связано со структурой всей сети. Таким образом, наиболее уязвимый узел занимает наивысшую позицию в сетевой иерархии. Мера уязвимости сети — максимальная уязвимость среди всех ее узлов:

$$V = \max_i V_i.$$

3.1.8. Коэффициент элитарности

В наукометрии на протяжении длительного времени исследуются сети цитирований. Известно, что влиятельные исследователи определенных областей формируют сообщества сетевого типа, выражающиеся, например, в публикации совместных работ. Такая закономерность наблюдается также в других реальных сетях и отражает такую тенденцию, как хорошая связность между узлами-концентраторами. Это явление, известное под названием элитарность (или феномен «клуба богатых» — *rich-club phenomenon*), может быть охарактеризовано коэффициентом элитарности, введенным в работе [51]. Анализ топологии веб, проведенный Ши Жоу (S. Zhou) и Р. Дж. Мондрагоном (R.J. Mondragon) из Лондонского университета, показал, что узлы с большой степенью исходящих гиперссылок имеют больше связей между собой, чем с узлами с малой степенью, тогда как последние

имеют больше связей с узлами с большей степенью, чем между собой. Исследование показало, что 27 % всех соединений имеют место между всего 5 % наибольших узлов, 60 % приходится на соединения других 95 % узлов с 5 % наибольших и только 13 % — это соединения между узлами, которые не входят в лидирующие 5 %.

Элитарность степени k у сети G — это некое множество узлов со степенью, большей k , $\mathfrak{R}(k) = \{v \in N(G) | k_v > k\}$. Коэффициент элитарности степени k выражается следующим образом:

$$\varphi(k) = \frac{1}{|\mathfrak{R}(k)|(|\mathfrak{R}(k)| - 1)} \sum_{i, j \in \mathfrak{R}(k)} a_{ij},$$

где сумма соответствует удвоенному количеству ребер между вершинами в «элите». Эта характеристика подобна коэффициенту кластерности, она определяет долю связей, существующих между узлами со степенью, превышающей k .

3.1.9. Корреляция степеней связанных вершин

Значительное количество структурных и динамических свойств сети определяется с помощью оценки корреляции между степенями соседних узлов. Такая корреляция может быть выражена через совокупное распределение $P(k, k')$, т.е. как вероятность того, что произвольно выбранное ребро соединяет узел степени k с узлом степени k' . Зависимость между степенями вершин можно выразить в терминах условной вероятности того, что произвольно выбранный сосед вершины степени k имеет степень k' [52]:

$$P(k' | k) = \frac{\langle k \rangle P(k, k')}{kP(k)}.$$

При этом $\sum_{k'} P(k' | k) = 1$. В случае неориентированных сетей

$$P(k, k') = P(k', k) \text{ и } k'P(k | k')P(k') = kP(k' | k)P(k).$$

Если сеть ориентированная, то k — это степень предшествующего узла, k' — степень последующего узла, значения k и k' могут быть входными, выходными или полными степенями. В общем случае $P(k, k') = P(k', k)$.

Значения $P(k, k')$ и $P(k | k')$ формально описывают корреляции степеней узлов, однако их сложно вычислять экспериментальным путем, что связано с размером сети и малой мощностью выборки узлов с высокими степенями. Эту проблему можно решить, вычислив среднюю степень ближайших соседей узлов с заданной степенью k по формуле

$$S(k) = \sum_{k'} k' P(k' | k).$$

Показатель корреляции степеней связности позволяет выделить отдельные классы сетей. Если корреляция отсутствует, то $S(k)$ не зависит от значений k , $S(k) = \langle k^2 \rangle / \langle k \rangle$. Если $S(k)$ возрастает при увеличении k , то узлы больших степеней тяготеют к соединениям с узлами больших степеней, и сеть относят к ассортативным (отсюда и феномен «клуба богатых»), тогда как если $S(k)$ — убывающая функция от k , то вершины больших степеней тяготеют к соединениям с вершинами малых степеней, и сеть называют дизассортативной [53].

В работе [53] был подсчитан коэффициент корреляции Пирсона для некоторых реальных и смоделированных сетей. Было обнаружено, что, несмотря на отображение моделями специфических особенностей структуры (степенное распределение степени связности узлов, свойство «малого мира»), большинство из них не воспроизводит ассортативность реальной сети. Например, для модели построения сети со степенным распределением степеней связности узлов Барабаши–Альберта [54] значение $r = 0$. Тем не менее, социальные сети склонны к ассортативности, а биологические и технологические часто дизассортативны.

Известно, что ассортативные сети менее уязвимы к равновероятным атакам, а дизассортативные менее уязвимы к целенаправленным атакам на узлы-концентраторы. Также, например, синхронизация

состояния компонент сети происходит быстрее в ассортативных сетях. Например, при распространении заразной болезни социальные сети в идеальном случае должны быть ассортативны: при контроле малой доли узлов-концентраторов сеть разбивается на изолированные компоненты связности, что позволяет эффективно контролировать распространение инфекции.

3.2. Модель слабых связей

Существует класс сложных сетей, которым присущи так называемые «слабые» связи. Аналогом слабых социальных связей являются, например, отношения с далекими знакомыми и коллегами. В некоторых случаях эти связи оказываются более эффективными, чем связи «сильные». Так, группой исследователей из Великобритании, США и Венгрии был получен концептуальный вывод в области мобильной связи, заключающийся в том, что «слабые» социальные связи между индивидуумами оказываются наиболее важными для существования социальной сети [55].

Для исследования были проанализированы звонки 4,6 млн абонентов мобильной связи, что составляет около 20 % населения одной европейской страны. Это был первый случай в мировой практике, когда удалось получить и проанализировать такую большую выборку данных, относящихся к межличностной коммуникации.

В социальной сети с 4,6 млн узлов было выявлено 7 млн социальных связей, т.е. взаимных звонков от одного абонента другому и обратно, если обратные звонки были сделаны на протяжении 18 недель. Частота и продолжительность разговоров использовались для того, чтобы определить силу каждой социальной связи.

Было выявлено, что именно слабые социальные связи (один-два обратных звонка на протяжении 18 недель) связывают воедино большую социальную сеть. Если эти связи проигнорировать, то сеть распадется на отдельные фрагменты. Если же не учитывать сильных связей, то связность сети нарушится (рис. 10).

Оказалось, что именно слабые связи являются тем феноменом, который связывает сеть в единое целое. Надо полагать, что данный вывод справедлив и для веб-пространства, хотя исследований в этой области до сих пор не проводилось.

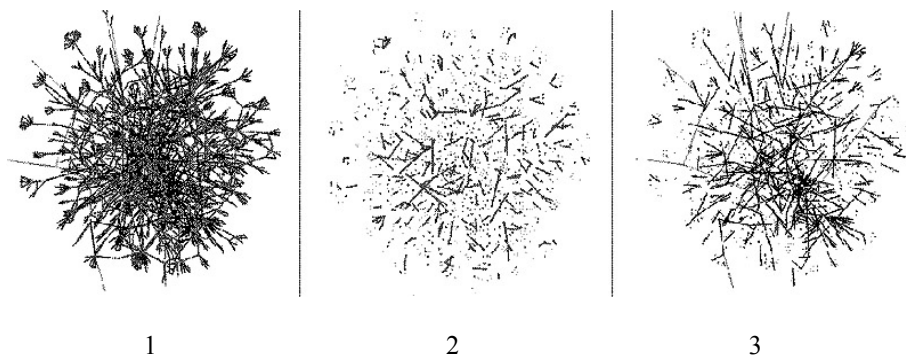


Рис. 10. Структура сети:

1 — полная карта сети социальных коммуникаций; 2 — социальная сеть, из которой изъяты слабые связи; 3 — сеть, из которой изъяты сильные связи: структура сохраняет связность

3.3. Модель малых миров

Несмотря на огромные размеры некоторых сложных сетей, во многих из них (в веб-пространстве, в частности) существует сравнительно короткий путь между двумя любыми узлами — геодезическое расстояние. В 1967 г. психолог С. Милгран в результате проделанных масштабных экспериментов вычислил, что существует цепочка знакомств, в среднем длиной шесть, практически между двумя любыми гражданами США [56].

Д. Уаттс и С. Строгатц обнаружили феномен, характерный для многих реальных сетей, названный эффектом «малых миров» (*Small Worlds*) [57].

Сетевые структуры, соответствующие свойствам малых миров обладают следующими типичными свойствами: малая средняя длина пути относительно диаметра сети (что характерно также для случайных сетей) и большой коэффициент кластеризации (что присуще сетям с регулярной структурой).

При исследовании этого феномена ими была предложена процедура построения наглядной модели сети, которой присущ данный феномен.

Чтобы построить сеть «малого мира», следует начать с регулярной циклической решетки с N вершинами, каждая из которых соединена с k (в частности, $k = 2$) ближайшими соседями в каждом направлении. Для каждой вершины задается $2k$ связей, где $N \gg \log_2(N) \gg 1$. Затем каждое ребро пересоединяется со случайной парой вершин с вероятностью p .

При условии $p = 0$ получается упорядоченная решетка с большим количеством циклов и большими расстояниями, а при условии $p \rightarrow 1$ сеть становится случайным графом с короткими расстояниями и малым количеством циклов. В некоем среднем случае присутствуют и короткие расстояния, и большое количество циклов.

Три состояния этой сети представлены на рис. 11: регулярная сеть, каждый узел которой соединен с четырьмя соседними, та же сеть, у которой некоторые «ближние» связи случайным образом заменены «далекими» (именно в этом случае возникает феномен «малых миров») и случайная сеть, в которой количество подобных замен превысило некоторый порог.

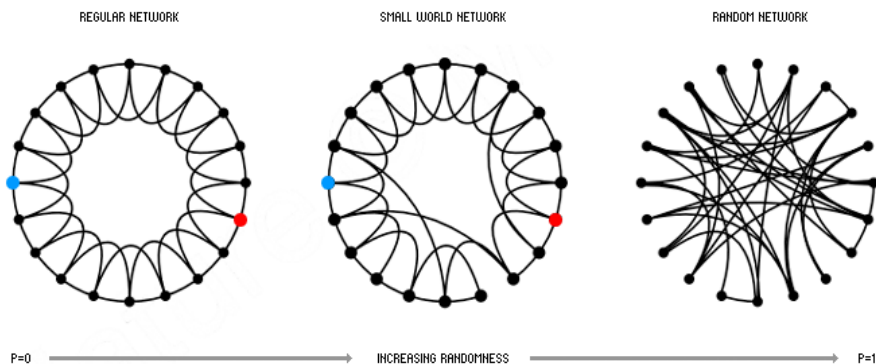


Рис. 11. Модель Уаттса–Строгатца

На рис. 12 приведены графики изменения средней длины пути и коэффициента кластеризации искусственной сети Д. Уаттса и С. Строгатца от вероятности установления «далеких связей» (в полулогарифмической шкале).

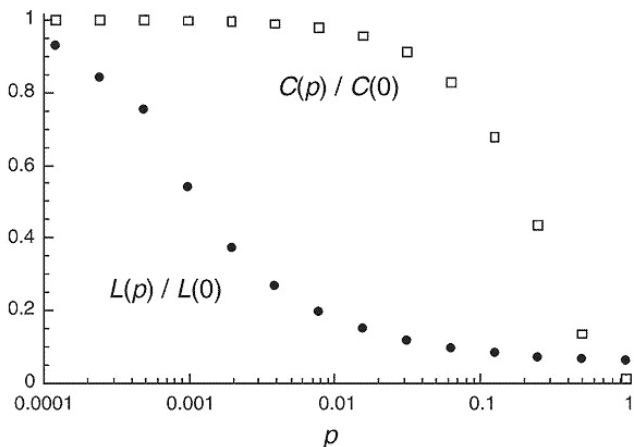


Рис. 12. Динамика изменения длины пути и коэффициента кластерности в модели Уаттса–Строгатца в полулогарифмической шкале (горизонтальная ось — вероятность замены ближних связей далекими)

В реальности оказалось, что именно те сети, узлы которых имеют одновременно некоторое количество локальных и случайных «далеких» связей, демонстрируют одновременно эффект «малого мира» и высокий уровень кластеризации. Веб-пространство также является сетью, для которой также подтвержден феномен малых миров.

Эти исследования дают основания полагать, что зависимость веб-пространства от больших узлов значительно существеннее, чем предполагалось ранее, т.е. она еще более чувствительна к злонамеренным атакам. С концепцией «малых миров» связан также практический подход, называемый «сетевой мобилизацией», которая реализуется над структурой «малых миров». В частности, скорость распространения информации благодаря эффекту «малых миров» в реальных сетях возрастает на порядки по сравнению со случайными сетями, ведь большинство пар узлов реальных сетей соединены короткими путями.

Кроме того, сегодня довольно успешно изучаются масштабируемые, статические, иерархические «малые миры» и другие сети; исследуются их фундаментальные свойства такие, как стойкость к деформациям и перколяция. Недавно было показано, что наибольшую информационную проводимость имеет особый класс сетей, называемых «запутанными» (англ. — *entangled networks*). Они характеризуются

максимальной однородностью, минимальным расстоянием между любыми двумя узлами и очень узким спектром основных статистических параметров. Считается, что запутанные сети могут найти широкое применение в области информационных технологий, в частности, в новых поколениях веб, позволяя существенным образом снизить объемы сетевого трафика.

3.4. Живучесть системы с разветвленной структурой

Для оценки живучести системы с разветвленной структурой необходимо установить показатель качества, по изменению которого определяется степень ухудшения характеристик при наличии действия внешних возбуждающих воздействий. Для разветвленных структур в качестве такого показателя естественно выбрать количество работоспособных веток. Поскольку это количество — случайная величина, то должно быть найдено распределение количества работоспособных веток $\{P(t/S_i)\}$. На векторе функций $P(t/S_i)$ определяют функционал $\varphi(P(t/S_i))$, который можно рассматривать как показатель эффективности, по изменению которого можно оценить живучесть.

Рассмотрим такие варианты показателя эффективности:

— *среднее количество работоспособных ветвей*

$$\varphi_1(P(t/S)) = \bar{N}(S) = \sum_{i=1}^{N_0} P_i(t/S)(N_0 - i),$$

где $P_i(t/S)$ — вероятность того, что в момент времени t неработоспособны i веток; N_0 — количество веток в полностью работоспособной структуре;

— *вероятность работоспособности количества веток, не меньше заданной*

$$\varphi_2(P(t/S)) = \sum_{i=1}^{i_0} P_i(t/S),$$

где i_0 — максимально допустимое количество неработоспособных веток (количество работоспособных веток не меньше, чем $N_0 - i_0$).

Допустим, что многократному воздействию подвергается k -ярусная структура с коэффициентами разветвления на ярусах r_1, r_2, \dots, r_k .

В рамках данной модели предполагается, что в область действия точечных НВ попадают только узлы системы, и нарушения работоспособности различных узлов одного яруса — равновероятные. Поскольку общее количество узлов в системе составляет $1 + r_1 + r_1 r_2 + \dots + r_1 r_2 \dots r_k$, можно записать такое выражение:

$$\gamma_0 + \gamma_1 r_1 + \gamma_2 r_1 r_2 + \dots + \gamma_k r_1 r_2 \dots r_k = 1,$$

где γ_k — вероятность нарушения работоспособности узла k -го ранга после одноразового НВ. Возможны различные гипотезы об уязвимости элементов, которые учитывают особенности узлов различных рангов, в частности, возможна такая гипотеза, что

$$\gamma_i = \gamma = (1 + r_1 + r_1 r_2 + \dots + r_1 r_2 \dots r_k)^{-1}.$$

Тогда, исходя из гипотезы о «равномерности» ярусов, получаем

$$\gamma_i = \frac{1}{(k+1)r_1 r_2 \dots r_i}, \quad i \geq 1; \quad \gamma_0 = \frac{1}{k+1}.$$

В более общем случае может выполняться

$$\gamma_i = a_i \gamma_{i+1}, \quad 0 \leq i \leq k-1, \quad a_i \geq 1.$$

Тогда получаем

$$\gamma_k = \gamma, \quad \gamma_i = a_i a_{i+1} \dots a_{k-1} \gamma, \quad 0 \leq i \leq k-1, \quad \gamma = \left(\sum_{n=0}^k \prod_{i=n}^{k-1} a_i \prod_{j=1}^n r_j \right)^{-1}.$$

3.5. Моделирование деструктивного воздействия на сети

Как и сеть террористов, восстановление которой после деструктивного воздействия описано в [3], информационная система также является динамической системой, восстановление которой после уничтожения лучших «посредников» осуществляется за счет латентных связей с другими компонентами информационного пространства. После того как информационная система разделяется на изолированные фрагменты, она может «использовать» эти связи и быстро восстановить связность (рис. 13).

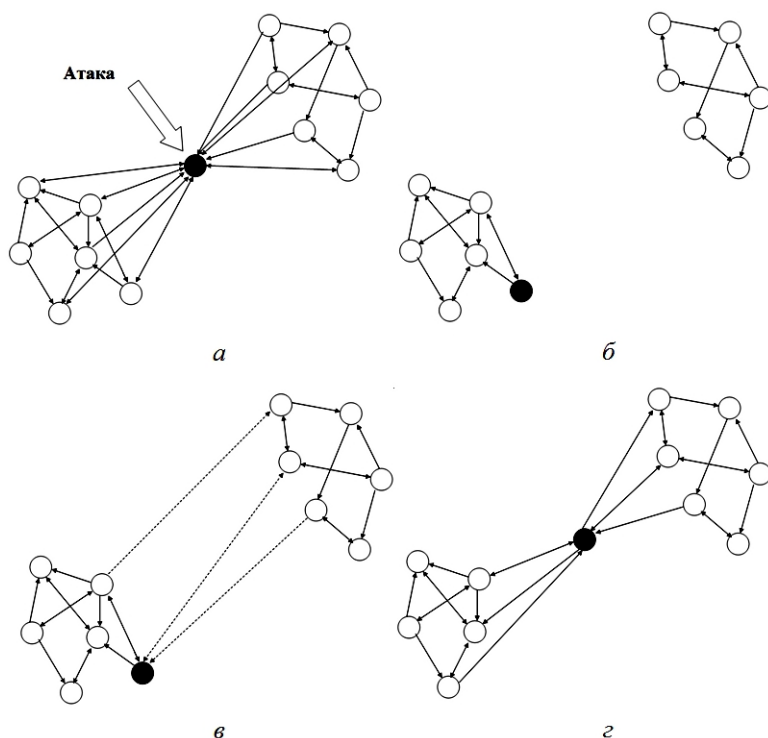


Рис. 13. Восстановление структуры сети путем выбора нового «посредника»: а — атака на сеть; б — несвязная сеть с изъятым посредником; в — возрождение скрытых (латентных) связей; г — связность сети восстановлена

Воссоединение частей сети не состоится, если ни одна из пар уцелевших после деструктивного воздействия компонент не сможет найти скрытые связи между собой (возможно, не прямые, а через другие компоненты информационного пространства). В этом случае влияние разъединения на показатели функционирования информационной системы зависит от того, смогут ли снова разъединенные части получить взаимные связи, недостаток которых наблюдается в этой части информационной системы.

Если часть информационной системы была близкой к самодостаточности, то она продолжает функционировать самостоятельно. В противном случае она прекращает функционирование до тех пор, пока не сформируется новая связь. Если одно из соединений оказывается успешным, то его инициатор становится новым «посредником», который объединяет две части сети.

4. КРИТИЧЕСКИЙ УРОВЕНЬ ЖИВУЧЕСТИ СИСТЕМ

Если в сети слишком мало ребер, т.е. средняя степень ее вершин достаточно мала, то в ней присутствует много изолированных узлов и кластеров из немногочисленных узлов. При добавлении ребер в такую сеть (например, установление новых связей, цитирование и т.п.) небольшие кластеры объединяются в большие.

По достижении некоторого критического уровня, характеризующего порог протекания сети, большая часть вершин будет объединена в кластер, размер которого сопоставим с количеством узлов сети, т.е. вся сеть станет проводящей.

Этот эффект — фазовый переход — описывается в литературе как «появление гигантского кластера» [58].

При исследовании живучести сетевой структуры основной интерес представляет противоположный эффект — переход от гигантского кластера к разреженной сети в результате деструктивных воздействий, выражающихся в удалении элементов сети — ребер (или узлов). При этом функциональный отказ рассматривается как удаление отдельного элемента. В этом случае возникает достаточно точная аналогия с известным в теории перколяции пределом протекания (или перколяционным порогом), который связывается с фазовым переходом.

Существующие математические модели такого поведения сетевых структур (диодных сетей, сетей террористов и т.п.) при деструктивном воздействии приводятся к перколяционной задаче [59–61]. Как правило, в рамках этой задачи рассматривается решетка из N проводящих элементов, после чего рассчитывается проводимость всей решетки при удалении отдельных ее элементов. Порог проводимости рассматривается как аналог порога живучести.

На рис. 14 показан пример типичного поведения подобных моделей. С увеличением количества функциональных отказов k наблюдается спад количества выполняемых функций. В общем случае в точке, которую назовем «критический уровень живучести», происходит резкий спад (фазовый переход II рода).

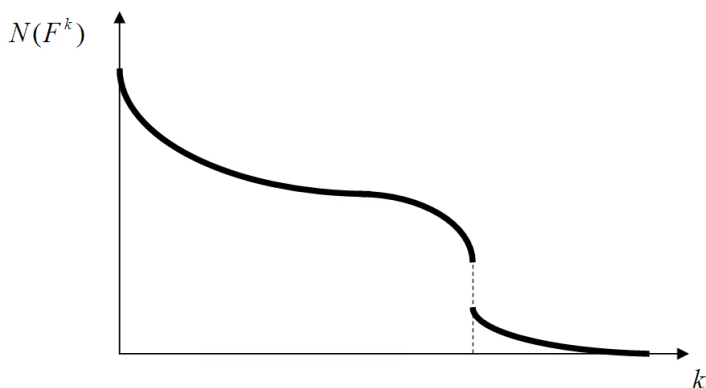


Рис. 14. Зависимость количества выполняемых функций (вертикальная ось) от количества функциональных отказов k (числа элементов, вышедших из строя)

4.1. Фазовые переходы

Понятие фаз встречается во множестве явлений, приведем несколько примеров. Для воды известно три фазы: жидкая, твердая (лед) и газообразная (пар). Каждая из них характеризуется своими значениями параметров. Существенно то, что при изменении внешних условий одна фаза (лед) переходит в другую (жидкость). Еще один объект — ферромагнетик (например, железо или никель). При низких температурах $t < T_c = 360$ °С никель является ферромагнетиком, при снятии внешнего магнитного поля он остается намагниченным. При температуре выше T_c это свойство теряется, при выключении внешнего магнитного поля он переходит в парамагнитное состояние. Здесь, как и в предыдущем примере, четко видно существование разных фаз — парамагнитной и ферромагнитной. При изменении температуры происходит фазовый переход.

В [60] приведен пример из теории перколяции. Случайно вырезая из сетки ребра, в конце концов, когда концентрация оставшихся ребер p станет меньше некоторого значения p_c , по решетке уже нельзя будет пройти «из конца в конец». Таким образом, сетка из фазы «протекания» перейдет в состояние фазы «непротекания».

Из этих примеров ясно, что для каждой из рассмотренных систем существует так называемый параметр порядка, определяющий, в какой из фаз находится система. В ферромагнетизме параметр порядка — намагниченность в нулевом внешнем поле, в теории перколяции — связность сетки.

Фазовые переходы бывают нескольких типов. Фазовый переход I рода характеризуется тем, что в системе может одновременно существовать несколько фаз. Например, при температуре $0\text{ }^{\circ}\text{C}$ лед плавает в воде. Если система находится в термодинамическом равновесии (нет привода и отвода тепла), то лед не тает и не нарастает. Для фазовых переходов II рода существование одновременно нескольких состояний невозможно. Сетка со случайно вырезанными ребрами либо связная, либо нет.

Решающим в создании теории фазовых переходов II рода, начало которой положил Л.Д. Ландау, было введение параметра порядка η как отличительного признака фазы системы. В одной из фаз, например парамагнитной, $\eta = 0$, а в другой, ферромагнитной, $\eta \neq 0$. Для магнитных явлений параметр порядка η — это намагниченность системы.

Для описания фазовых переходов вводится некоторая функция параметров, определяющих состояние системы — $G(\eta, T, \dots)$. В физических системах это энергия Гиббса. В каждом явлении (перколяция, сеть «малых миров» и т.д.) эта функция определяется по-своему. Главное свойство указанной функции, первое предположение Л.Д. Ландау — в состоянии равновесия эта функция принимает минимальное значение:

$$\frac{\partial G}{\partial \eta} = 0, \quad \frac{\partial^2 G}{\partial \eta^2} > 0.$$

В физических системах говорят о термодинамическом равновесии, в теории сложных сетей можно говорить об устойчивости. Заметим, что условие минимальности определяется варьированием параметра порядка.

4.2. Задача теории перколяции

Одной из важных характеристик сложных сетей является возможность протекания по их ребрам тока, жидкости, информации (трафика) и т.п. Впервые задача перколяции (англ. *percolation* — просачивание, протекание) была сформулирована в 1957 г. в работе С.Р. Бродбента (S.R. Broadbent) и Дж.М. Хаммерсли (J.M. Hammersley) [61]. В последствии была развита целая область исследований, названная теорией перколяции, имеющая многочисленные применения на практике. Оказывается, что многие вопросы, которые возникают при анализе живучести информационных сегментов, также непосредственно относятся к теории перколяции.

Перед теорией перколяции стоят многие вопросы, которые выходят за стандартные рамки дискретной математики и теории вероятностей [59].

Самая простая формулировка задачи теории перколяций следующая. Дана решетка из связей, случайная часть которых p — «черная», проводящая, а остальная — «белая», не проводящая поток. Необходимо найти такую минимальную концентрацию p_c «черных» связей, при которой еще есть связный путь по «черным» связям сквозь всю решетку, т.е. такую концентрацию, когда решетка в целом проводит поток.

При $p = 0$ все связи решетки «белые» — решетка не проводит поток. При увеличении концентрации «черных» — проводящих связей, при $p = p_c$ в решетке возникает перколяционный, проникающий кластер из «черных» связей, соединяющий противоположные края сетки. При размере сетки, стремящемся к бесконечности, размер этого кластера также бесконечен, в связи с чем и был введен термин «бесконечный кластер». Другие совокупности соединенных между собой связей конечного размера называются конечными кластерами.

При переходе через порог протекания, т.е. при возникновении бесконечного кластера, свойства системы, характеризующие ее в целом, резко изменяются. Если, например, «черная» связь проводит ток, а «белая» нет, то проводимость всей системы $R \sim 1/G$ вблизи p_c резко уменьшается. Учитывая логарифмический масштаб на рис. 15, ясно, что вблизи порога протекания пропускная способность сети при очень небольшом уменьшении p может резко упасть.

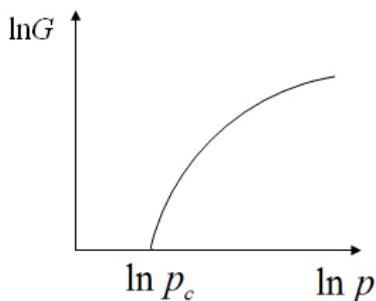


Рис. 15. Проводимость системы вблизи порога протекания (в логарифмическом масштабе)

Казалось бы простая задача определения p_c , а «не поддалась» точным методам теории вероятностей. За редкими исключениями p_c не удастся вычислить аналитически — необходимо численное моделирование.

В настоящее время известно много важных обобщений перколяционной задачи, например, рассматриваются случаи, когда «непроводящие» связи проводят, но намного хуже проводящих; можно говорить о значениях проводимости для разных связей; можно рассматривать однонаправленные «диодные связи» и т.п.

К задачам, решаемым в рамках теории перколяции и анализа сложных сетей, относятся такие, как определение предельного уровня проводимости (пропускной способности), изменения длины пути и его траектории (извилистости, параллельности) при приближении к предельному уровню проводимости, количества узлов, которые необходимо удалить, чтобы нарушить связанность сети.

4.3. Характеристики перколяционных сетей

Существуют важные характеристики ребер сетей, по которым происходит, например, прохождение тока, жидкости, информации (в интересующих нас случаях).

Приведем некоторые из них:

$P_\infty(p)$ — вероятность того, что случайным образом выбранный узел принадлежит бесконечному кластеру, соединяющему две проти-

воположные стороны сети — эта характеристика имеет смысл выше порога протекания, когда существует бесконечный кластер;

$$S(p) = \frac{\sum_s s^2 n_s(p)}{\sum_s s n_s(p)} - \text{средний размер кластера (англ. — mean}$$

cluster size), где n_s — число кластеров из s узлов, приходящихся на один узел решетки — эта характеристика имеет смысл ниже порога протекания, когда все существующие в системе кластеры имеют конечный размер;

$\xi(p)$ — корреляционная длина, характеризующая скорость спада корреляций в решетке. Чтобы ввести ее, необходимо определить парную корреляционную функцию $G(r, p \approx p_c) \equiv G(|r_i - r_j|, p) = \langle g(r_i, r_j) \rangle$, где функция $g(r_i, r_j)$ от радиусов векторов узлов i и j равна единице, если узлы связаны «черными» связями, принадлежащими одному конечному кластеру, и равной нулю во всех остальных случаях, а угловые скобки означают усреднение по всем узлам. При $r \rightarrow \infty$ корреляционная функция $G(r, p)$ убывает экспоненциально, характерным масштабом при этом как раз и является корреляционная длина $G(r, p): \exp[-r / \xi(p)]$.

Отметим, что при $p \rightarrow p_c$ корреляционная длина $\xi(p)$ расходится — это полностью соответствует качественным представлениям. Чем ближе к порогу протекания ($p \geq p_c$), тем меньше «черных» проводящих связей, тем существеннее каждая часть кластера, соединяющая две «бесконечности», обрыв одной из них может сказаться далеко от нее. На самом пороге протекания одна оборванная связь может разрушить весь проводящий путь, называемый бесконечным кластером. Структура бесконечного кластера не простая, на самом пороге протекания он является фрактальным объектом, состоящим из связей, входящих в этот кластер, с размерностью $d_f(d=2) = d - \beta/\nu \approx 1,896$, $d_f(d=3) \approx 2,54$.

Вводят также и другие характеристики бесконечного кластера — остов кластера (англ. *backbone*) — та часть кластера, по которой

происходит перенос, а также отсеченные «мертвые концы» (англ. *dead ends*), и многие другие.

В более развитой и реалистичной теории протекания и «белые» связи считаются проводящими, с проводимостью намного меньшей «черных» (принято говорить, что в $1/h$ раз меньше, где $0 < h \ll 1$). При введении параметра h теория протекания может быть сформулирована в терминах теории фазовых переходов II рода, одном из самых сложных разделов теоретической физики.

Описание проводимости перколяционной сети также может быть сформулировано в терминах теории фазовых переходов II рода. Для определенности будем говорить о так называемой эффективной проводимости σ_e , которая характеризует проводимость системы в целом. Роль параметра порядка в этом случае играет σ_e , а близость к критической точке $\tau = (p - p_c) / p_c$, где p_c — порог протекания. Соответствующее скейлинговое соотношение для эффективной проводимости σ_e имеет вид

$$\sigma_e = h^{\frac{t}{\varphi}} f(\tau / h^{\frac{1}{\varphi}}),$$

$$f(z) \sim \begin{cases} z^t, & z \rightarrow +\infty, \\ \text{const}, & z \rightarrow 0, \\ |z|^{-q}, & z \rightarrow -\infty. \end{cases}$$

В этих формулах t и q — так называемые критические индексы проводимости, $\varphi = t + q$.

Необходимо уточнить, что универсальное скейлинговое поведение σ_e имеет место только вблизи порога протекания, т.е. при $|\tau| \ll 1$.

4.4. Перколяция на случайных сетях

Перколяцию можно рассматривать и на случайных сетях Эрдоса–Реньи, Уаттса–Строгатца, и т.п. В этом случае вместо бесконечного перколяционного кластера говорят о гигантской связанной компоненте (англ. — *giant connected component*).

Для случайной сети Эрдоша–Реньи из N узлов известно, что порог протекания $p_c \approx 1/N$, т.е. что протекание наступает когда средняя степень узла $\langle k \rangle \geq 1$.

М. Ньюманом (M. Newman) и Д. Уаттсом [58] была рассмотрена задача перколяции на сетях типа «малый мир». При этом была использована модификация таких сетей. В отличие от того, как «малый мир» вводился в [57], новые связи, называемые *shortcuts*, набрасывались в начальную сеть дополнительно, все старые связи между соседями не прерывались, оставаясь на своих местах. Как и в стандартной теории протекания, может быть введена характерная длина (в теории протекания она называется корреляционной длиной):

$$\xi = \frac{1}{(\varphi kd)^{1/d}},$$

где φ — вероятность встретить переброшенную связь (*shortcuts*), φN — число этих связей, N — число узлов сети, k — число соседей, d — мерность сети. Смысл этой величины — расстояние между концами разных наброшенных связей.

Для кратчайшего пути l (среднего, конечно) существует скейлинговая функция:

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right),$$

$$f(z \ll 1) \rightarrow \text{const}, \quad f(z \gg 1) \rightarrow \frac{\log z}{z}.$$

Так как $\xi \sim 1/(\varphi k)^{1/d}$, асимптотика скейлинговой функции означает при $z = N/\xi \ll 1$, что $N\varphi^{1/d} \ll 1$, т.е. что число переброшенных связей (*shortcuts*) очень мало. Тогда

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right) \sim N,$$

т.е. сеть представляет собой «большой мир», среднее расстояние между двумя узлами возрастает пропорционально числу узлов.

Если же $z = N / \xi \gg 1$, т.е. $N\varphi^{1/d} \gg 1$, то выполняется

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right) \sim \log\left[N(\varphi kd)^{1/d}\right] \sim \log N,$$

сеть является «малым миром», расстояние между узлами возрастает намного медленнее, как логарифм числа узлов.

Вернемся теперь к вопросу о перколяционных свойствах таких сетей. Для исследования проблем живучести систем особый интерес представляют несколько вопросов, например, что будет происходить, если некоторая доля $q = 1 - p$ узлов выпадет (не будет проводить информацию, ток и др.), чему равно критическое значение $q_c = 1 - p_c$, когда в сети еще существует гигантская связанная компонента (англ. — *giant connected component*), т.е. когда значительное число узлов сети еще соединено между собой? Известно, что порог протекания p_c (Np_c — число неразорванных связей) связан с долей переброшенных связей φ следующим образом:

$$\varphi = \frac{(1 - p_c)^k}{2kp_c \left[1 + kp_c (1 - p_c)^k\right]}.$$

На рис. 16 показаны зависимости величины порога протекания p_c от φ , при которой еще есть гигантская связанная компонента. Как видно, при протекании, чем меньше переброшенная доля связей, тем больше число узлов должно быть переброшено.

Перколяция на безмасштабных сетях с распределением степеней узла $P(k) \sim k^{-\gamma}$ обладает своей спецификой, отличной от перколяции в «малых мирах». Выражение для порога протекания p_c различное для различных диапазонов параметра γ . Например, при $\gamma > 3$:

$$q_c = 1 - p_c = 1 - \frac{1}{\frac{\gamma-2}{\gamma-3} k_0 - 1},$$

где степень узла лежит в диапазоне $k_0 \leq k \leq K_0$.

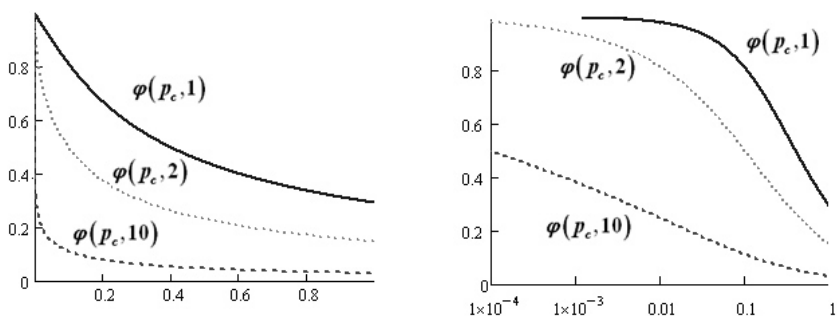


Рис. 16. Зависимость порога протекания p_c от доли переброшенных связей φ : сверху вниз — число соседей один, два и десять. На правой части ось абсцисс — в логарифмическом масштабе

При $k_0 = 0$ выражение для q_c упрощается $q_c = 4 - \gamma$. Таким образом, если доля удаленных узлов больше чем $4 - \gamma$ — гигантский кластер отсутствует, т.е. можно считать, что сеть разрушена. Подробно об этом и о том, что происходит в других диапазонах параметра γ , можно прочитать, например, в [62].

4.5. Теория перколяции и моделирование атак на сети

Приведенные результаты вполне пригодны для моделирования атак на такую глобальные сети. Известно, что узлы сети Интернет связываются линиями связи с различной пропускной способностью. Отдельные сегменты Интернет соединены в точки обмена трафиком, существуют опорные (*backbone*) каналы. Моделирование атак, т.е. разрывов отдельных связей или удаление отдельных, как показывают расчеты, очень немногих узлов может нарушить такую связность и

дает большие шансы террористам. Нахождение «слабых» участков глобальной сети, проектирование резервных узлов и каналов требуют точных расчетов на базе теории перколяции.

Следует отметить, что выше все время речь шла о выводе из строя элементов (узлов или связей) с вероятностью q случайным образом. В тоже время можно, например, вырезать узлы как в стандартной перколяционной задаче, так в других типах сложных сетей, целенаправленным образом, выбирая такие узлы, при вырезании которых сеть разрушается максимально быстро. В сети Интернет такое направленное вырезание узлов (серверов) называется «запланированной атакой». При этом выведение из строя порядка 1 % целенаправленных узлов уменьшает производительность всей сети в два раза [63].

Возможно также обобщение перколяционной задачи на так называемую диодную или направленную (англ. — *directed*) перколяцию. В этой модификации некоторые из хорошо проводящих («черных») связей в двухфазной версии перколяции заменяются на «диодные», пропускающие только в одну сторону.

Отдельные веб-сайты (узлы) и гиперсвязи в веб-пространстве образуют диодную перколяционную сеть, так как в соответствии с протоколом HTTP не предусматривается, чтобы гиперсвязи, ведущие с веб-страниц, вели в обратную сторону. Относительно небольшое количество веб-сайтов (например, каталоги веб-ресурсов) содержит большое количество гиперссылок, в то время как большая часть веб-сайтов содержит крайне мало ссылок на другие веб-ресурсы. Если допустить, что будет произведена атака на веб-сайты, имеющие наибольшее количество ссылок на другие веб-сайты, то в зависимости от мощности такой атаки может быть нарушена связность этой диодной сети. В результате поисковые серверы не смогут охватить некоторую часть веб-пространства, что автоматически переведет последнюю в зону «скрытого» веб [64].

5. ИНФОРМАЦИОННЫЕ ПОТОКИ И ВОПРОСЫ ЖИВУЧЕСТИ

5.1. Информационные потоки

Информационные системы как сетевые структуры в информационном пространстве состоят из отдельных элементов, образующих в динамике своей эволюции (появление, развитие, модификация, уничтожение) информационные потоки. Следовательно, живучесть информационных систем напрямую зависит от свойств информационных потоков.

Для исследования современных информационных потоков в Интернете, т.е. потоков сообщений, которые публикуются на страницах веб-сайтов, в социальных сетях, блогах, и тому подобное, должен применяться принципиально новый инструментарий, потому что классические методы обобщения информационных массивов (классификации, фазового укрупнения, кластерного анализа и тому подобное) не всегда способны адекватно отражать состояние динамической составляющей информационного пространства. В этом случае речь идет не столько об анализе документальных массивов фиксированных размеров, пусть даже очень больших, сколько об обобщении динамического потока гипертекстовых данных.

Конечно, большая часть информации, которая представлена в Интернете, находит своего потребителя. Однако если рассматривать всю совокупность сетевых публикаций как определенную общность по отношению к конкретному пользователю (или группе пользователей), то можно увидеть ряд проблем, связанных с полнотой, релевантностью и оперативностью получения данных. Поиск, фильтрация, сбор информации в Интернете требуют достаточной квалификации персонала и, к сожалению, при этом не могут учитываться все особенности информационной структуры сети и представления в ней данных. Это, в свою очередь, ведет к тому, что единичные выборки информации из веб-пространства не могут считаться репрезентативными.

При этом информационный поток, который «потребляется» конкретным пользователем, носит, как правило, выраженную предметную направленность, которая характеризуется областью его интересов. Поиск и обработка информации в ручном режиме — достаточно трудоемкий, а главное, длительный процесс, который чаще всего не дает

желаемого результата. Решение проблемы на практике возможно путем создания автоматизированных систем сбора, фильтрации и анализа информации, так называемых интеллектуальных посредников между пользователем или корпоративной информационной системой и сетью Интернета. Подобная система должна осуществлять сбор и селекцию информации из Интернет и создавать документальную базу данных, специфицированную предметной областью пользователя, т.е. выполнять функции интеграции информационных потоков. Загрузка информации в базу данных должна сопровождаться ее классификацией и структуризацией. Для последующей информационно-аналитической работы пользователю должны предоставляться эффективные средства навигации, поиска и обобщения информации, которая сохраняется в соответствующей динамической документальной базе данных.

Современный уровень развития информационного пространства обуславливает интерес к подходам, основанным на понимании информации как меры упорядоченности некоторой системы и, соответственно, к статистическим методам ее обработки. Для организации эффективной коммуникации в сетях сегодня приходится постоянно возвращаться к математическим истокам теории информации, понятиям энтропии, теории Шеннона, уравнениям Больцмана и др., широкие перспективы применения мощного аппарата математики и физики в решении теоретико-информационных задач [65].

Для формального описания информационных потоков введем некоторые общие для всего последующего изложения предположения. Дадим определение информационного потока [66], которое корреспондируется с классическим определением из теории информации.

Рассмотрим отрезок (a, τ) действительной оси (оси времени), где $\tau > a$. Допустим, что на этом отрезке времени в соответствии с некоторыми закономерностями в сети публикуется некоторое количество информационных документов — k . На оси времени моменты публикации отдельных документов обозначим как $\tau_1, \tau_2, \dots, \tau_k$ ($a \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq \tau$). Информационным потоком будем называть процесс $N_\alpha(\tau)$, реализация которого характеризуется количеством точек (документов), появившихся в интервале (a, τ) , как функцию правого конца отрезка τ . В соответствии с этим определением реали-

зация информационного потока — необывающая ступенчатая всегда целочисленная функция $N_{\alpha}(\tau)$.

Приведенное определение на локальных временных областях соответствует действительности, но не учитывает такой эффект, как старение информации, которое противоречит «накопительной» способности информационного потока $N_{\alpha}(\tau)$ на больших интервалах времени.

Так, определенный информационный поток учитывает лишь количество информационных сообщений, вне зависимости от их содержания. В общем случае определение содержания, тематики отдельных документов — достаточно субъективный процесс. Для строгого моделирования тематических информационных потоков используют модели, которые различают документы по отдельным словам или словосочетаниям (обычно их называют терминами, от англ. — *Terms*).

Задачи мониторинга информационных потоков большого объема в компьютерных сетях, их адаптивного агрегирования и обобщения осложняются отсутствием типовых методик и решений, неполнотой существующих технологических подходов. В настоящее время исследования по проблемам анализа информационных потоков большого объема в компьютерных сетях чаще всего носят узкоспециализированный характер. Вместе с тем опыт создания и внедрения корпоративных информационных систем свидетельствует о необходимости создания и внедрения документальных информационных хранилищ для обеспечения научных исследований, получения разнообразных аналитических сведений, навигации в документальных информационных потоках больших объемов.

При моделировании этих процессов используются методы нелинейной динамики, теории клеточных автоматов и самоорганизованной критичности. При моделировании информационных потоков изучаются структурные связи между входящими в них массивами документов. Сегодня при этом все чаще применяется фрактальный анализ, подход, базирующийся на свойствах сохранения внутренней структуры массивов документов при изменениях их размеров или масштабов рассмотрения. Теория информации, которая ранее находила свое основное применение в области передачи данных, становится полезной и для анализа текстовых массивов, динамически порождаемых в сетях.

Предусматривается, что новостные сообщения обладают свойством старения, т.е. теряют свою актуальность со временем. Все ин-

формационное пространство можно с достаточной мерой условности разделить на две составляющие — стабильную и динамическую, которые имеют очень разные характеристики своего развития. В частности, процесс старения информации в известной модели Бартона-Кеблера описывается уравнением, состоящим из двух компонент:

$$m(t) = 1 - ae^{-T} - be^{-2T},$$

где $m(t)$ — часть полезной информации в общем потоке через время T ; первое вычитаемое соответствует стабильным ресурсам, а второе — динамическим, новостным. Это уравнение также в полной мере соответствует объемам информации, которые формируются в информационном пространстве по определенным тематикам, которые время от времени возникают и исчезают. Стабильная составляющая информационного пространства содержит информацию «долгосрочного» плана, тогда как динамическая составляющая включает ресурсы, которые постоянно обновляются. Некоторая часть последней составляющей впоследствии «вливается» в стабильную составляющую, однако большая часть «исчезает» из информационного пространства или попадает в сегмент его, так называемой скрытой части, не доступной пользователям, использующим обычные информационно-поисковые системы (ИПС).

В традиционной сетевой информационно-поисковой системе информационное пространство, состоящее из стабильной и динамической частей, которое индексируется с помощью этой ИПС, через некоторое время изменяет свое наполнение: одни документы помещаются в стабильную часть в виде архивов, а другие исчезают. В этом случае пользователь при обращении к ИПС находит релевантные запросу документы из стабильной части, ссылки из динамической части, которые устарели, и ничего не находит из обновленной динамической части.

Для навигации в информации из динамической части информационного пространства требуется применение системы-посредника между пользователем и информационным пространством. Подобный посредник (или агент новостей) может выполнять работу по сбору и селекции информации. Принцип индексирования, которое должно осуществляться этим посредником, немного отличается от индексирования традиционными поисковыми системами: индексируется не весь

контент информационного пространства, а лишь его динамическая часть.

Информация возникает в информационном пространстве не сама по себе. Ее публикуют, размещают на веб-сайтах, страницах социальных сетей, узлах пиринговых сетей и тому подобное. В дальнейшем такие информационные ресурсы будем называть информационными источниками.

Отметим, что система контент-мониторинга InfoStream [67], являющаяся одним из практических «полигонов» этого исследования, потребляет политематический информационный поток более чем из 5000 источников. На рис. 17 приведен фрагмент динамики этого потока (вертикальная ось — количество документов) в разрезе времени (суток) и наиболее продуктивных источников. На этом графике явно выражены недельные колебания объемов информационных потоков.

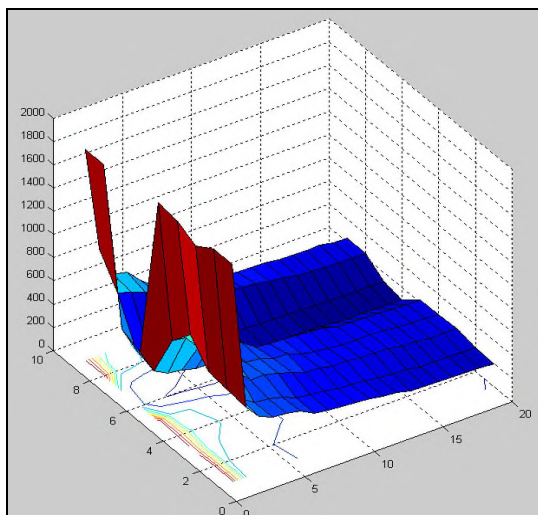


Рис. 17. Динамика публикации информации наиболее рейтинговыми источниками (получено с помощью системы InfoStream)

В настоящее время мощные возможности Интернет порождают проблему оптимизации состава и количества источников, которые могут использоваться корпоративной информационной системой для обеспечения необходимого количества и качества документов, удов-

летворяющих потребностям пользователей. В связи с этим актуальными оказываются вопросы ранжирования и выбора источников новостной информации — веб-сайтов, к которым нужно обеспечить доступ через один интерфейс как в поисковом режиме, так и в режимах пересмотра и аналитического обобщения.

Принципам ранжирования веб-документов посвящено большое количество научных работ и практических разработок. Ссылочное ранжирование веб-сайтов сегодня является отдельным направлением интернет-бизнеса — SEO (*search engine optimization*). Вместе с тем вопросам ранжирования и отбора информационных ресурсов с учетом их контента, объемов и стабильности тематики публикаций уделяется значительно меньшее внимание. Безусловно, основным критерием при выборе источников для таких систем мониторинга новостей является их содержание. Распределение источников по контенту, который соответствует тематическим потребностям корпоративных пользователей, удовлетворяет закон Бредфорда, соответственно, при отборе источников обязательно должно учитываться их ранжирование по степени соответствия тематике.

Однако реализация такого выбора ведет к некоторым сложностям. На практике такое ранжирование осуществляется экспертами путем оценки количества документов, релевантных заранее определенному пакету тематических запросов, которые адресуются к фрагменту базы данных, составленной из документов анализируемого источника. А это неизбежно приводит к определенному субъективизму со всеми вытекающими последствиями.

Поэтому представляется перспективным дополнить традиционный подход более объективными и более строгими методами, которые позволяют оптимизировать процесс формирования информационной базы систем интеграции информационных потоков.

На рис. 18 приведен график распределения (в полулогарифмическом масштабе) количества документов, опубликованных источниками, сканирующимися системой InfoStream, которые ранжируются по параметру — количеству документов, опубликованных источником.

Ниже приведены распределения, которые относятся к массиву документов за марта 2008 года объемом свыше 1,2 млн документов из более чем 2500 источников — открытых веб-сайтов. Центральная часть графика хорошо аппроксимируется прямой, которая свидетельствует о близости представленной зависимости к гиперболической (т.е.

о действии обобщенного закона Ципфа). На рис. 19 представлено общее количество документов, которые охватываются системой мониторинга в зависимости от источников, учитывающихся в ней, также ранжированных по количеству опубликованных документов. Поскольку закон Ципфа допускает аппроксимацию плотности распределения гиперболической зависимостью вида a/x , то функция распределения количества документов

$$f(x) \sim \int \frac{a}{x} dx = a \ln x + C,$$

с хорошим приближением описывается логарифмическим законом.

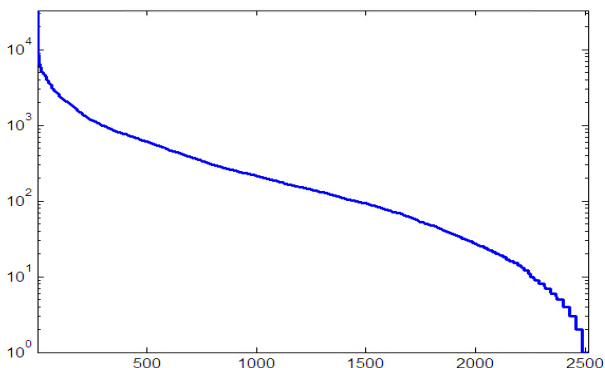


Рис. 18. Распределение количества публикаций (ось ординат) по ранжированному списку источников (ось абсцисс)

Приведенная зависимость позволила построить критерий выбора необходимой части источников для разных корпоративных применений из общего списка, которые удовлетворяют потребностям пользователей.

Как уже отмечалось, цитируемость отдельных документов и веб-сайтов сегодня является одним из основных критериев оценки рангов документов в сетевых поисковых системах (PageRank, HITS, TrustRank, ТИЦ и тому подобное). Идея оценки уровня цитирования позволила реализовать одну из первых моделей динамической части веб-пространства [68]. Отметим, что оценка уровня источника инфор-

мации как «автора» преимущественно по количеству веб-сайтов, из которых на него осуществляются гиперссылки, полностью согласовывается с предложенным Мораном и Лемпелем алгоритмом Salsa.

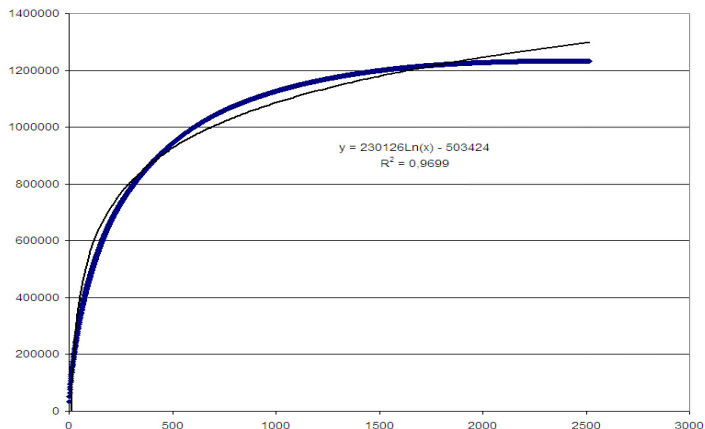


Рис. 19. Зависимость количества публикаций в системе мониторинга (ось ординат) от источников, ранжированных по количеству документов (ось абсцисс)

Если допустить, что все источники делают одинаковый вклад по количеству опубликованных документов, то данная зависимость могла бы быть линейной и выражаться формулой:

$$f_{lin}(n) = n \frac{f_{max}}{N},$$

где f_{max} — максимальный объем охваченных документов; N — общее количество источников; n — номер текущего источника.

Очевидно, что отклонение реальной зависимости от линейной вначале возрастает, а потом уменьшается до нуля. Будем называть количество источников пороговым n_p , если значение реальной зависимости максимально отклоняется от приведенной линейной

$$n_p = \arg \max_n \{f(n) - f_{lin}(n)\}.$$

На рис. 20 представлена зависимость n_p от различных значений N , т.е. когда выбирается N самых производительных источников.

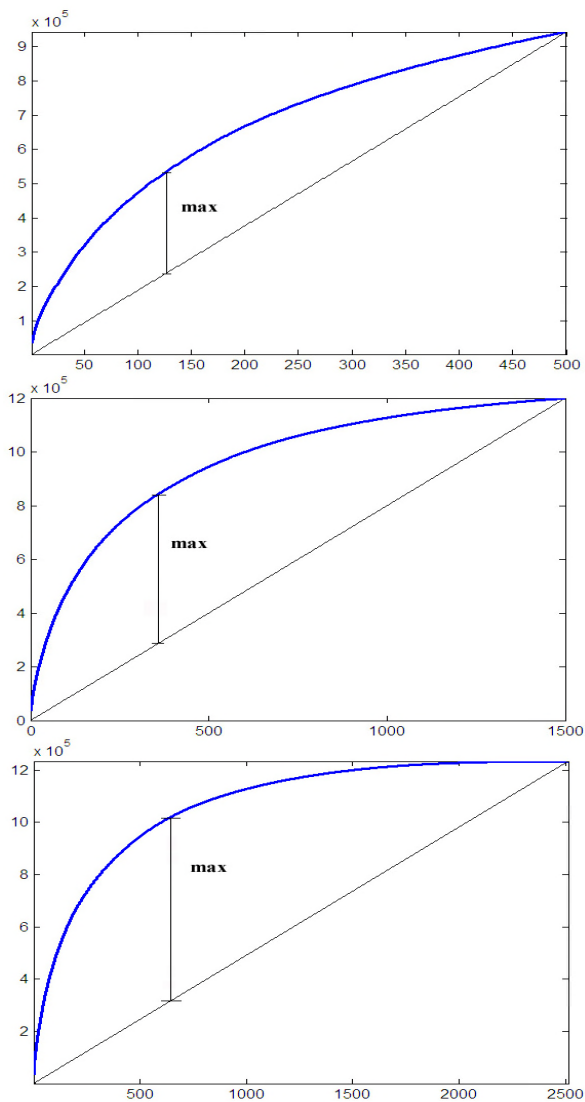


Рис. 20. Количество публикаций в системе мониторинга (ось ординат) при подключении новых наиболее интенсивных источников (ось абсцисс)

Интересно (и полностью соответствует характеру функции $f(n)$), что значения n_p практически линейно зависят от N (рис. 21): $n_p \sim 0,24N$, при этом количество охваченных документов, соответствующих n_p при максимальном количестве источников (2514, рис. 22), достигает 80 процентов от f_{\max} .

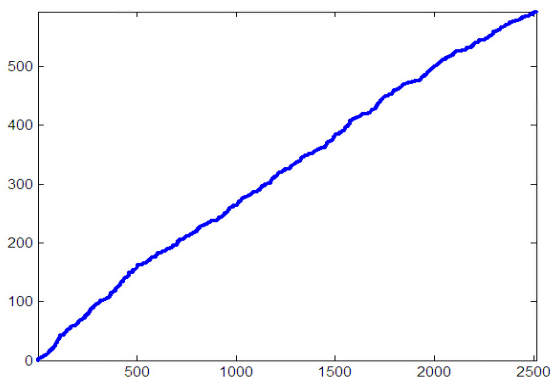


Рис. 21. Зависимость порогового значения (ось ординат) от начального количества источников (ось абсцисс)

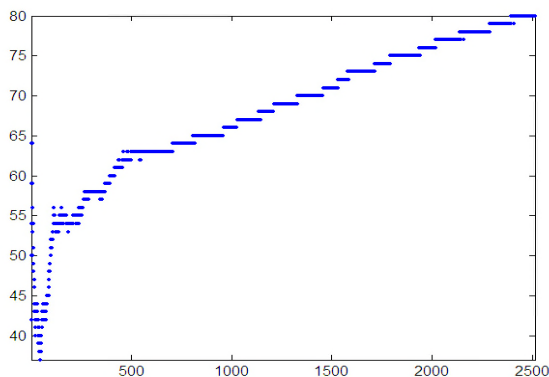


Рис. 22. Зависимость удельного количества документов, охватываемого системой (ось ординат), от изменения начального количества источников (ось абсцисс)

Отметим, что построенная зависимость удовлетворяет принципу Парето: приблизительно 20 % наиболее производительных источников публикуют 80 % документов.

Специальное место в исследовании занимало изучение содержательного дублирования информации. При этом следует отметить, что процент документов, которые дублируются по смыслу, в системе мониторинга InfoStream значительно меньше, чем во всем новостном веб-пространстве. Это объясняется подбором источников для сканирования, в число которых не входят много новостных интеграторов.

Как уже отмечалось, одной из главных особенностей новостной информации является наличие большого количества сообщений, дублирующих друг друга. Кроме того, о событии мирового значения напишут все средства массовой информации (СМИ), причем, скорее всего, на одной из первых страниц. Потребитель (за исключением некоторых специфических направлений аналитических исследований информационного пространства), однако, желает получать по каждому событию одно сообщение.

Поэтому исследование характера и свойств дублирования информации приобретает в современных технологиях исключительно важное значение. В частности, актуальным становится задание отбора наиболее оригинальных источников, которые позволяют (по крайней мере статистически) исключить не только формальное, но и содержательное дублирование информации. Дублирование сообщений на веб-сайтах зависит от разных причин, поэтому проведенные измерения для ранжированного по количеству публикаций списка источников показывают разный уровень, причем информация не носит наглядный характер. Вместе с тем, исследование авторов свидетельствует об устойчивой тенденции: чем более производительный источник информации, тем больше он содержит заимствований из других источников.

5.2. Моделирование информационных потоков

Живучесть информационных систем как тематических информационных потоков определяется их составом, структурой и содержанием (контентом). Анализ динамики тематических информационных потоков, которые генерируются в веб-пространстве, является сегодня одним из наиболее информативных методов исследования актуальности тех или иных тематических направлений [69]. Эта динамика обусловлена факторами, большая часть которых не поддаются точному

анализу. Однако общий характер временной зависимости количества тематических публикаций в Интернете все же допускает построение математических моделей.

В поведении информационных потоков наблюдаются две характерные особенности: во-первых, выразительная тенденция к постоянному возрастанию их объемов, а во-вторых, усложнение динамической структуры. Наблюдения временных зависимостей числа сообщений в сетевых информационных потоках убедительно свидетельствуют о том, что механизмы их генерации и распространения, очевидно, связаны со сложными нелинейными процессами общей сетевой динамики.

Традиционными считаются два класса моделей информационных потоков: линейные и экспоненциальные. Оба класса имеют существенную ограниченность — монотонный характер временной зависимости, т.е. они мало пригодны для изучения реальной динамики сетевых информационных потоков в течение длительных интервалов времени.

5.2.1. Тематические информационные потоки

Под тематическим информационным потоком будем понимать последовательность сообщений, соответствующих определенной тематике. Таким образом, информационные системы в нашем понимании также являются тематическими информационными потоками, но в отличие от следующих друг за другом сообщений в простых информационных потоках, информационные системы — это сетевые структуры, учитывающие многочисленные информационные связи.

В узком смысле под тематическим информационным потоком будем понимать количество документов, которые в некотором смысле соответствуют заданной теме. Рассмотрим общую картину динамики тематических информационных потоков, ограничившись механизмами, типичными для динамического сегмента веб-пространства.

Многочисленные факты свидетельствуют о том, что в действительности динамика тематических информационных потоков определяется комплексом внутренних нелинейных механизмов, которые лишь частично коррелируют с объективным окружением. Очевидно, что эта динамика в принципе не может быть объяснена некоторым одним фактором, который полностью отвечает за все разнообразие наблюдаемых эффектов. Именно это обстоятельство и объясняет боль-

шую актуальность проблемы моделирования динамики тематических информационных потоков.

Информационный поток, измеряемый количеством сообщений, является величиной относительно стабильной. Изменяются во времени лишь объемы массивов сообщений, соответствующие той или иной тематике, той или иной информационной системе. Иными словами, возрастание количества публикаций по одной теме при ограниченной способности их генерации (что вполне соответствует действительности) сопровождается уменьшением публикаций на другие темы [69], так что для каждого промежутка времени T имеем:

$$\int_0^T \sum_{i=1}^M n_i(t) dt = NT,$$

где $n_i(t)$ — количество публикаций в единицу времени по теме i ; M — общее количество всех возможных тем, т.е. для локальных временных промежутков можно наблюдать так называемый тематический баланс [69].

Основной интерес в такой формулировке представляет изучение динамики отдельного тематического потока, который описывается плотностью $n_i(t)$.

Теоретически можно допустить, что множества публикаций, ассоциируемых с определенным набором тематик, пересекаются, т. е. существуют публикации, которые могут быть отнесены одновременно к нескольким различным тематикам. В реальности такая политематичность действительно наблюдается, она является эффектом, который необходимо учитывать, но в первом приближении будем считать, что его вклад не искажает общей картины.

Каждая тематика также имеет ряд характерных свойств, которые допускают некоторую классификацию, например, на основе особенностей ее образования и воспроизведения во времени:

— публикации на «разовую» тему, временная зависимость количества которых резко возрастает, выходит на насыщение, а затем убывает и далее асимптотически стремиться к нулю;

— публикации по темам, которые периодически появляются в общем информационном потоке, а затем через некоторое время практически исчезают из него;

— публикации по теме, временная зависимость количества которых колеблется возле некоторого значения и никогда не исчезает полностью.

Таким образом, сообщения могут разделяться на аналогичные категории, причем каждая из них имеет собственную специфику развития во времени.

Еще сложнее выглядит синхронное изменение количества сообщений из нескольких тематических информационных потоков. Их поведение четко напоминает процессы взаимодействия популяций в биоценозе. Так, в ряде случаев увеличение числа публикаций по одной теме сопровождается сокращением числа публикаций по другим темам. Общая динамика в этом случае может описываться системой уравнений, каждое из которых относится к отдельному монотематическому потоку. Подчеркнем, что общие политематические потоки являются стационарными по количеству публикаций, динамика же в основном определяется «конкурентной борьбой» отдельных тематик.

Вместе с тем в практическом плане часто оказывается полностью удовлетворительным упрощенное понимание информационного потока как некоторой зависимой от времени величины $n(t)$, которая описывается уравнением

$$\frac{dn(t)}{dt} = F(n(t), t).$$

В многочисленных литературных источниках описано много разновидностей систем «конкурентной борьбы» для разных модификаций модели в зависимости от целого ряда предположений о реальных условиях протекания процессов. В самом простом виде такие уравнения могут иметь следующий вид:

$$\frac{dm_i(t)}{dt} = p_i \cdot m_i(t) - \sum_{j=1}^{N_m} r_{ij} \cdot m_i(t) \cdot m_j(t),$$

где N_m — количество тематик.

Приведенная система уравнений описывает перераспределение публикаций между тематиками, образующими фиксированный набор. Но

в реальной жизни тематики (сюжеты) появляются и со временем исчезают, поэтому необходимо ввести в эти уравнения соответствующие коррективы. Это можно сделать по-разному, например, определив зависимость коэффициентов p_i и r_{ij} от времени так, чтобы каждый сюжет имел собственный максимум активности на определенном промежутке времени.

5.2.2. Традиционные модели информационных потоков

Линейная модель

В некоторых случаях динамика тематических информационных потоков, выражаемых количеством публикаций за определенный период, их интенсивностью, обусловленной, например, изменением активности тематики (ее повышением или старением), линейна, т.е. количество сообщений в момент времени t можно представить формулой

$$y(t) = y(t_0) + v(t - t_0),$$

где t_0 — стартовое время отсчета; $y(t)$ — количество сообщений к моменту времени t ; v — средняя скорость увеличения (уменьшения) интенсивности тематического информационного потока.

Важные характеристики информационного потока могут быть количественно оценены флуктуацией этого потока — изменением среднеквадратичного отклонения $\sigma(t)$, вычисляемого так:

$$\sigma(t_n) = \sqrt{\frac{1}{n} \sum_{i=0}^n [y(t_i) - (y(t_0) + v(t_i - t_0))]^2}.$$

Если эта величина изменяется пропорционально квадратному корню от времени, то процесс изменения количества публикаций по избранной теме можно считать процессом с независимыми приращениями. При этом связями с предыдущими тематическими публикациями можно пренебречь.

В случае, когда среднеквадратичное отклонение пропорционально некоторой степени от времени: $\sigma(t) \propto t^\mu$ ($1/2 \leq \mu \leq 1$), то чем

большее значение μ , тем выше корреляция между текущими и предыдущими сообщениями в информационном потоке.

Экспоненциальная модель

В некоторых случаях процесс изменения актуальности тематики (увеличения или уменьшения количества тематических сообщений в информационном потоке в единицу времени) аппроксимируется экспоненциальной зависимостью, которая выражается формулой

$$y(t) = y(t_0) \exp[\lambda(t - t_0)],$$

где λ — среднее относительное изменение интенсивности тематического информационного потока.

В реальности актуальность тематики является дискретной величиной, измеряемой в моменты времени t_0, \dots, t_n , которая лишь аппроксимируется приведенной выше зависимостью. В рамках данной модели справедливо соотношение

$$\begin{aligned} y(t_i) &= y(t_0) \exp[\lambda(t_i - t_0)] = \\ &= y(t_0) \exp[\lambda(t_i - t_{i-1} + t_{i-1} - t_0)] = y(t_{i-1}) \exp[\lambda(t_i - t_{i-1})]. \end{aligned}$$

Откуда

$$\frac{y(t_i)}{y(t_{i-1})} = \exp[\lambda(t_i - t_{i-1})].$$

Введем обозначение: $\lambda(t_i)$ — относительное изменение интенсивности тематического информационного потока в момент времени t_i :

$$\lambda(t_i) = \lambda(t_i - t_{i-1}).$$

Прологарифмировав приведенное выше уравнение, получаем

$$\lambda(t_i) = \ln \frac{y(t_i)}{y(t_{i-1})}.$$

Относительное изменение интенсивности в момент времени t_i на практике также часто вычисляется как соотношение

$$\lambda(t_i) = \ln \frac{y(t_i)}{y(t_{i-1})} \approx \frac{y(t_i) - y(t_{i-1})}{y(t_{i-1})}.$$

Изменение флуктуаций величины $\lambda(t_i)$ относительно среднего значения может оцениваться по стандартному отклонению:

$$\sigma(t_n) = \sqrt{\frac{1}{n} \sum_{i=0}^n (\lambda(t_i) - \lambda)^2}.$$

В этом случае, если $\sigma(t)$ изменяется пропорционально квадратному корню от времени, то также можно говорить о процессе с независимыми приращениями — корреляция между отдельными сообщениями несущественна. В случае наличия значительной зависимости сообщений выполняется соотношение: $\sigma(t) \propto t^\mu$, причем значение μ превышает 1/2, но ограничено 1.

Значение μ , которое превышает 1/2, свидетельствует о наличие долгосрочной памяти в информационном потоке. Такой класс процессов получил название автомодельных, для которых предусматривается корреляция между количеством сообщений, публикуемых в разные моменты времени.

Логистическая модель

В отличие от модели Бартона–Кеблера в реальной динамике информационных потоков имеют место процессы как увеличения, так и уменьшения количества документов. Поэтому для построения реалистичной картины, безусловно, необходимо применять более гибкие модели.

В первую очередь, стоит сказать, что документы в информационном потоке во многих отношениях напоминают популяции живых организмов. Они в определенном смысле «рождаются», «умирают» и дают «потомство» (документы, содержащие информацию, ранее появившуюся в других документах). В современной научной литературе понятие популяции часто используется в широком смысле, и потому полностью обосновано введение его и при моделировании информационных потоков.

Во второй половине XX века были достигнуты значительные успехи в построении различных математических моделей динамики популяций, в частности логистическая модель, которая оказалась применимой во многих отраслях науки и техники.

Логистическую модель [66] можно рассматривать как обобщение экспоненциальной модели Мальтуса, предусматривающей пропорциональность скорости возрастания функции $y(t)$ в каждый момент времени ее значению:

$$\frac{dy(t)}{d(t)} = ky(t),$$

где k — некоторый коэффициент.

В реальной жизни, как правило, динамические системы имеют достаточно эффективные обратные связи, позволяющие корректировать характер процессов, происходящих в них, и тем самым удерживать их в определенных рамках. Информационные операции, корректируя эти обратные связи в определенные периоды эволюционного процесса, могут эффективно влиять на характер поведения всей системы.

Наиболее простым обобщением закона Мальтуса, позволяющим уйти от неограниченного возрастания решения, является замена постоянного коэффициента k некоторой функцией времени $k(t)$. Естественно, эта функция должна быть выбрана таким образом, чтобы выполнялись условия:

- решение уравнения имело бы приемлемое поведение;
- структура функции имела бы определенный смысл с точки зрения исследуемого явления.

Главная идея логистической модели заключается в том, что для ограничения скорости роста на функцию $y(t)$ накладывается дополнительное условие, в соответствии с которым ее значением не должно превышать некоторое значение. Для этого выберем $k(t)$ такого вида:

$$k(t) = k \cdot [N - ry(t)],$$

где N — предельное значение, которое функция $y(t)$ не может превысить; r — коэффициент, который описывает негативные для данной тенденции процессы; k — коэффициент пропорциональности. Причем предусматривается, что всегда $n_0 \leq N$. Тогда вместо первого уравнения имеем

$$\begin{cases} \frac{dy(t)}{d(t)} = ky(t)(N - ry(t)), \\ y(t_0) = y_0. \end{cases}$$

Модель, основанная на приведенном выше уравнении, называется логистической. Несмотря на мнимую простоту, подобное обобщение закона Мальтуса никоим образом не является примитивным. Наоборот, оно позволяет явно включить в описание динамики популяций исключительно важную обратную связь. Логистическое уравнение, можно считать феноменологическим: исследователям не обязательно знать, как действуют конкретные механизмы, которые по мере возрастания $y(t)$ снижают скорость ее изменения.

Существует два класса решений логистического уравнения, которые в зависимости от значений коэффициентов описывают возрастание и убывание $y(t)$. Их типичное поведение представлено на рис. 23.

Приведенное выше логистическое уравнение имеет два равновесных решения: $y(t) = 0$ и $y(t) = N$. С формальной точки зрения первое из них неустойчиво, однако на практике это не совсем так. Дело в том, что реальные объемы информационных потоков выражаются дискретными числами, и если в какой-то момент $y(t)$ принимает значение, меньшее единицы, то в дальнейшем возрастать оно уже не

сможет. Поэтому в реальности решение $y(t) = 0$ также можно считать равновесным.

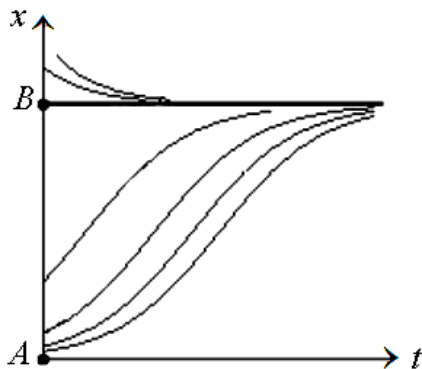


Рис. 23. Обобщенная логистическая модель

Второе решение $y(t) = N$ является равновесным в любом смысле. Действительно, при $y(t) > N$ начинают действовать механизмы убывания зависимости, а при $y(t) < N$, соответственно, возрастания.

Рассмотрим поведение динамики тематического информационного потока, объем которого определяется логистическим уравнением. На рис. 23 изображена результирующая зависимость объемов информационного потока от времени при разных начальных условиях. В точках A и B скорость изменения количества сообщений стремится к нулю: это стационарные состояния. Между A и B скорость позитивна (количество сообщений возрастает), а выше точки B — негативна (количество сообщений убывает).

Модель предусматривает, что впоследствии установится стационарный режим B , который выглядит вполне естественно: больший информационный поток уменьшается, а меньший — увеличивается.

Логистическая модель удовлетворительно описывает многочисленные явления насыщения. Вблизи A , когда объем информационного потока малый, она очень близка к мальтузианской модели. Но при достаточно больших x наблюдается резкое отличие от мальтузианского роста: вместо стремления x к бесконечности, количество публикаций приближается к стационарному значению B .

Рассмотрим, как логистическая модель может применяться во время анализа информационных потоков, а именно определим минимальное начальное количество c сообщений (которое можно, например, выделить для начала некоторой информационной операции). Пусть x — объем тематического информационного потока. На динамику этой величины осуществляется влияние других тематик, уменьшающих ее распространение, которое описывается таким образом: $\dot{x} = x - x^2 - c$.

Вычисления показывают, что поведение системы резко изменяется при некотором критическом значении C (рис. 24).

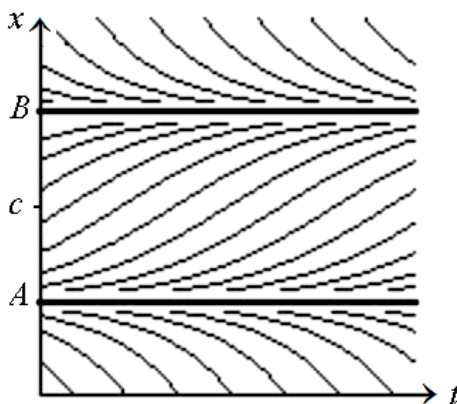


Рис. 24. Две постоянных точки логистической модели

Система имеет два равновесных состояния A и B . Состояние B устойчивое: объем информационного потока в этом случае немного меньше, чем при безконкурентной ситуации. Этот объем возобновляется при малых отклонениях от равновесного значения B .

Состояние A неустойчиво: если по любым причинам объем информационного потока снизится ниже уровня A , то в дальнейшем количество тематических сообщений будет сведено на нет за вполне конечное время.

Очевидно, что при наличии благоприятных внешних условий (при некоторой плотности ресурса) объем информационного потока увеличивается свободно, что способствует логистическому возраста-

нию. В этом случае даже более сложные модели должны давать результаты, подобные приведенным. Вместе с тем это означает, что основные параметры для конкретизации общей модели могут определяться в результате анализа упрощенной логистической модели.

Следовательно, логистическая модель успешно описывает достижение тематическим информационным потоком некоторого равновесного состояния.

Информационную динамику в общем случае можем представить как процесс, обусловленный возникновением и исчезновением отдельных тематик, которые имеют место на фоне общих тенденций информационного пространства. Зафиксируем определенную тематику и допустим, что в момент времени $t = 0$ существует n_0 фоновых публикаций. В силу того, что (в рамках принятой модели) актуальность тематики сохраняется в течение промежутка времени λ , можно рассматривать отдельно две временных области: $0 < t \leq \lambda$ с $D > 0$ и $t > \lambda$ с $D = 0$ (в рамках данной модели $D = \text{const}$ для каждой области — уровень актуальности темы) и, соответственно, функции $u(t)$ и $v(t)$, которые являются решениями для этих областей и «сшиваются» в точке λ :

$$y(t) = \begin{cases} u(t), & 0 < t < \lambda, \\ v(t), & t > \lambda, \\ u(t) = v(t), & t = \lambda. \end{cases}$$

Первой области соответствует процесс возрастания количества публикаций в условиях ненулевой актуальности темы и, возможно, переход к состоянию насыщения.

Реакция медийных средств никогда не бывает мгновенной: всегда существует определенная задержка во времени. Этот аспект учитывается в модели путем введения фактора запаздывания τ .

Соответствующая динамика описывается уравнением, которое после переопределения коэффициентов и их нормировки для функции $u(t)$ можно представить в виде

$$\frac{du(t-\tau)}{dt} = pu(t-\tau)(1-qu(t-\tau)) + Du(t-\tau),$$

$$u(0) = n_0.$$

Подчеркнем, что содержательно величина p определяет нормируемую вероятность появления публикации в единицу времени независимо от актуальности темы. Этот фактор отображает фоновые механизмы генерации информации (типичным примером может быть механическая перепечатка материалов из престижных информационных источников). Величина D характеризует непосредственное влияние актуальности данной темы. Параметр q характеризует уменьшение скорости роста количества публикаций и является величиной, обратной к асимптотическому значению зависимости $u(t)$ при $D = 0$.

Для второй области, описываемой функцией $v(t)$, соответственно, имеем

$$\frac{dv(t-\lambda)}{d(t)} = pv(t-\lambda)(1-qv(t-\lambda)).$$

При этом должно учитываться условие равенства функций $u(t)$ и $v(t)$ в момент $t = \lambda$:

$$v(\lambda) = u(\lambda).$$

Приведенные выше нелинейные дифференциальные уравнения являются вариантами записи уравнения Бернулли:

$$y' = ay^2 + by,$$

которое линеаризуется стандартной заменой $z = 1/y$:

$$z' + bz + a = 0.$$

Общее решение этого уравнения имеет вид

$$z = \frac{1}{\mu(x)} \left[C - a \int \mu(x) dx \right]$$

с интегрирующим множителем

$$\mu(x) = e^{bx}.$$

Переменные C определяются так: для первой области из начальных условий, а для второй — из условия «сшивки». В результате несложных преобразований находим решение для первой области:

$$u(t) = \frac{u_s}{1 + (u_s / n_0 - 1) \exp[-(p + d)(t - \tau)]},$$

где u_s — асимптотическое значение u , значение которого определяет область насыщения:

$$u_s = \frac{p + D}{pq}.$$

Таким образом, модель описывает зависимость, которая имеет S -подобную (логистическую) форму (рис. 25).

Отметим, что решение не зависит от значения n_0 , что свидетельствует о несущественности начальных условий для информационной динамики. Каким бы не было начальное количество публикаций, насыщение будет определяться исключительно параметрами, характеризующими фоновую скорость возрастания количества публикаций, количественную меру актуальности и негативные для процесса факторы.

Кривая, представленная на рис. 25 имеет точку перегиба

$$t_{\text{inf}} = \frac{1}{p + D} \ln(u_s / n_0 - 1) + \tau.$$

Таким образом, для первой области имеем так называемую S -подобную зависимость, а при $t \sim t_{\text{inf}}$ поведение $u(t)$ приближается к

линейной зависимости и соответствует линейной модели.

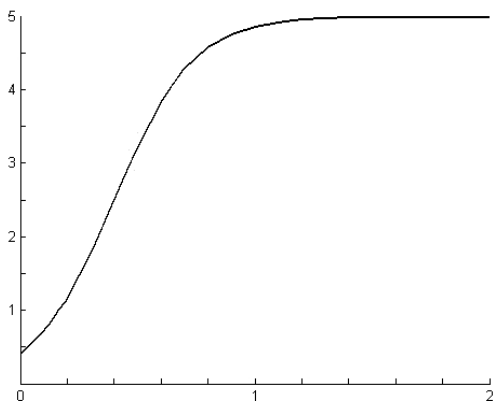


Рис. 25. Функция возрастания

Представим теперь выражение для $u(t)$ следующем виде:

$$u(t) = \frac{u_s}{\exp[(p+D)t] + (u_s/n_0 - 1)\exp[(p+D)\tau]}$$

откуда видно, что при условии

$$t < \frac{1}{p+D} \ln(u_s/n_0 - 1) + \tau = t_{\text{inf}}$$

зависимость $u(t)$ имеет экспоненциальный характер, т.е. для значений t , значительно меньших t_{inf} , модель совпадает с экспоненциальной моделью.

Для второй области, соответственно, имеем (рис. 26)

$$v(t) = \frac{v(\lambda)}{qv(\lambda) + (1 - qv(\lambda))\exp[-p(t - \lambda)]^2},$$

с учетом условия «сшивки»: $v(\lambda) = u(\lambda)$.

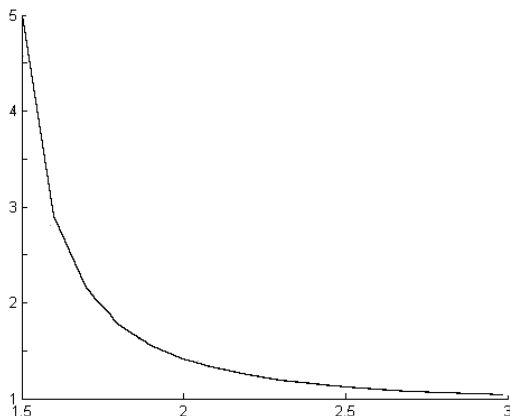


Рис. 26. Функция убывания

Если зависимость $u(t)$ успевает достичь насыщения за интервал времени $t < \lambda$, то приведенное выше уравнение можно упростить, представив его следующим образом:

$$v(t) = \frac{v_s(p + D)}{p + D(1 - \exp[-p(t - \lambda)])},$$

где $v_s = 1/q$ — асимптотическое значение зависимости $v(t)$.

Как и ожидалось, величина v_s также не зависит ни от начального условия, ни от условия «сшивки» с функцией $u(t)$ на границе областей. Как видно, полученная зависимость имеет область насыщения u_s при $t \leq \lambda$ и асимптотику v_s , которая описывает постепенное уменьшение числа публикаций до фонового уровня, т.е. она, по крайней мере, на качественном уровне, согласовывается с общими соображениями о характере информационной динамики, полученными на основе опытных данных. Кроме того, на локальных участках она неплохо аппроксимируется линейной и экспоненциальной моделями. Типичная полная зависимость $y(t)$ приведена на рис. 27.

В случае информационных потоков, которые ассоциируются с конкретными темами, необходимо описывать динамику каждого из

таких потоков отдельно, принимая во внимание то, что возрастание одного из них автоматически приводит к уменьшению других и наоборот. Поэтому ограничение на количество сообщений по всем тематикам распространяется и на совокупность всех монотематических потоков. В случае изучения общего информационного потока наблюдается явление «перетекания» публикаций из одних тематик, теряющих актуальность, в другие.

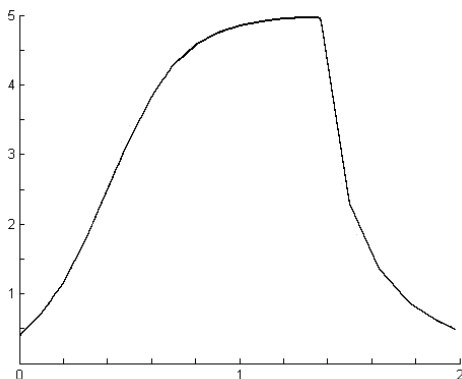


Рис. 27. Обобщенный график динамики тематического информационного потока

Общая динамика должна описываться системой уравнений, каждое из которых принадлежит к отдельному монотематическому потоку. Подчеркнем, что общие политематические потоки являются стационарными по количеству публикаций, динамика же в основном определяется «конкурентной борьбой» отдельных тематик.

Приведенную выше систему уравнений «конкурентной борьбы» в рамках обобщенной логистической модели можно представить так:

$$\frac{dy_i(t)}{dt} = (p_i + D_i(t, \lambda_i)) \cdot \left(y_i(t) - \sum_j r_{ij} \cdot y_i(t) y_j(t) \right).$$

В этих соотношениях коэффициенты p_i и D_i имеют тот же смысл, что и ранее, а λ_i — точки, в которых соответствующие D_i достигают максимальных значений.

5.2.3. Выявление информационных кластеров

На практике при поиске новостной информации всегда возникает задача выявления информационных кластеров (сюжетов), состоящих из отдельных документов, и их ранжирования по некоторым признакам, что должно обеспечить, не только выявление самой важной темы, но и «вверное» многоаспектное освещение всех наиболее значимых аспектов инфосюжетов. Эта задача является решаемой во многих системах с использованием различных подходов и алгоритмов. При этом неизменной остается технологическая цепочка: построение семантической сети из информационных сообщений, кластеризация — выявление наиболее взаимосвязанных групп, т.е. инфосюжетов, «взвешивание» (оценка важности, актуальности) и наглядная визуализация самых весомых из них [66].

При выделении сюжетных цепочек для определения попарной текстуальной близости отдельных документов, как правило, используются алгоритмы выявления подобных документов, ставшие уже традиционными в поисковых системах. Так, матрица попарной близости документов обрабатывается алгоритмами кластеризации, такими как *LSA/LSI*, *k-means*, суффиксных деревьев и т.д. Выделенные классы документов и представляют собой информационные сюжеты.

Для предъявления пользователям информационные сюжеты должны быть ранжированы. Основные факторы, влияющие на ранжирование по важности, — оперативность информации и размер сюжетной цепочки. Под оперативностью понимается некоторая функция от времени публикации всех документов в информационном сюжете, а размер отражает общий интерес к конкретной теме. Во всех этих подходах центральная задача состоит в отождествлении документов, относящихся к одному сюжету, и в выделении «непересекающихся» сюжетов. На рис. 28 представлен типовой подход к выявлению инфосюжетов. Последнее по времени генерации информационное сообщение (документ) сравнивается с предыдущими, оценивается уровень их подобия. Если уровень подобия превышает некоторый порог, то анализируемый документ считается принадлежащим информационному сюжету, к которому относится ранее сгенерированный документ. Если подобных документов нет, фиксируется новый сюжет, состоящий на текущий момент из одного документа.

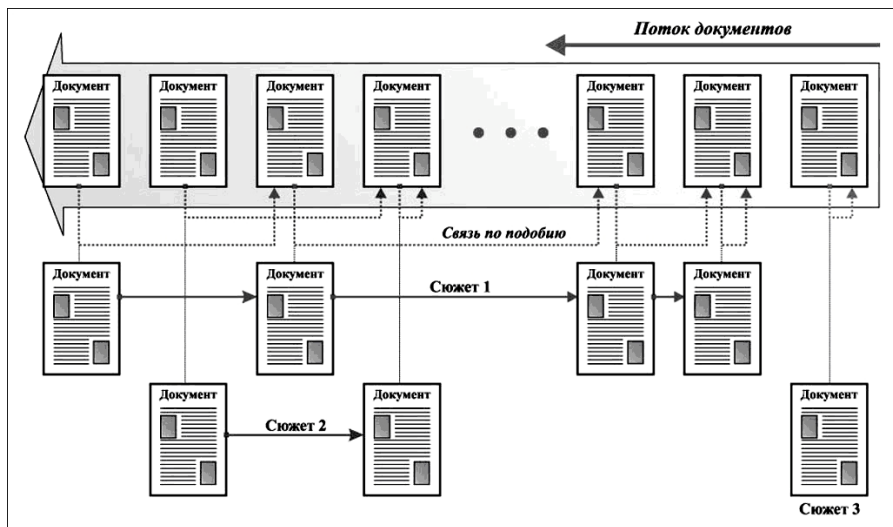


Рис. 28. Типовая схема выявления информационных сюжетов

Для результирующего отображения каждого отдельно взятого инфосюжета используются отобранные по текстуальной близости документы из различных источников, отсортированные в хронологическом порядке. При этом сюжеты могут представлять собой дайджесты, интегрирующие общие места документов по теме, а также уникальную информацию, содержащуюся в отдельных документах. Иными словами, реферирование сюжета в этом случае сводится не к свертыванию информации, а к построению расширенной версии, по сравнению с любым документом из сюжетной цепочки.

Например, в системе Яндекс.Новости (<http://news.yandex.ru>) для выделения информационного сюжета строится матрица попарной близости документов, которая обрабатывается алгоритмом кластеризации с эмпирически подобранными параметрами (в частности, радиусом метрики близости). Для того чтобы увеличить связность крупных сюжетов, дополнительно используют кластеризацию второго уровня, обеспечивающую сбор атомарных кластеров в более крупные. Все сообщения в результатах поиска сгруппированы на сайте, при этом ранжирование инфосюжетов построено на стандартных для Яндекса принципах ранжирования выдачи. Оно основано на числе и ранге новостей внутри новостных сюжетов, причем ранг одной новости

определяется как ее «свежесть» с учетом приоритетов удовлетворения критериев поиска.

В системе InfoStream (<http://infostream.ua>) тематическая близость документов определяется на основе нормированных последовательностей наиболее весомых терминов, входящих в каждый документ [70].

Последовательности подобных (с определенным коэффициентом взаимной близости, превышающим некоторый установленный эмпирически уровень) документов образуют цепочки. При этом каждый документ попадает в какую-нибудь цепочку, в крайнем случае, состоящую только из него самого. Затем цепочки оцениваются по длине и оперативности, после чего пользователю предъявляется определенное количество самых важных тематических инфосюжетов. Для репрезентации сюжетной цепочки, заголовки документов также оцениваются относительно ключевых слов, соответствующих сюжету, а затем из всех заголовков выбираются наиболее «весомые» для отображения (рис. 29).

Обзор основных сюжетов

киргизия; документов - 3000, сюжетов - 500

В виде графа (Java) Распечатать

- 1. Брата президента Киргизии объявили в розыск**
 Жаныш Бакиев. Фото с сайта post.su. Временное правительство Киргизии объявило в розыск младшего брата президента Курманбека Бакиева, который возглавляет Службу государственной охраны президента, сообщает "Интерфакс-Казахстан". Заместитель главы временного правительства Азимбек Бекназаров в эфире национального телевидения заявил, что вся вина за погибших 7 апреля в Бишкеке лежит на главе Службы государственной охраны президента Бакиева
 Сюжет полностью (1106)

2010.04.08 15:09 Россия признала смену власти в Киргизии Гruzия Online
 1106

2010.04.09 15:15 В Киргизии за расстрел митингующих разыскивают брата президента Бакиева BantM-фо.ru
- 2. Курманбек Бакиев отказался сложить с себя полномочия президента Киргизии**
 Президент Киргизии Курманбек Бакиев возложил ответственность за происходящие в Киргизии события на оппозицию. В киргизское информагентство "24" поступило сегодня его заявление, в котором он заявляет об отказе сложения с себя полномочий главы государства. "В результате безответственных действий лидеров оппозиции наша страна понесла ничем не оправданную, невосполнимую утрату: погибли ни в чем не
 Сюжет полностью (361)

2010.04.08 15:11 Оппозиция в Кыргызстане стала преемником правительства в Бишкеке. Что будет с "Манасом"? ("Christian Science Monitor", США) RT.KORR
 361

2010.04.09 15:14 "Ярлык на княжение!" "Россия и соотечественники"
- 3. В Киргизии объявлен траур по погибшим**
 GroBansicht des Bildes mit der Bildunterschrift: Столица Киргизии Бишкек В Киргизии объявлен траур по погибшим в ходе недавних событий. В столице Бишкеке в ночь на пятницу произошли столкновения между милицией и мародерами, к утру ситуация нормализовалась. В Киргизии 9 и 10 апреля объявлены траурными днями в память о погибших в ходе событий 6-7 апреля.
 Сюжет полностью (116)

2010.04.08 15:10 10 апреля объявят в Киргизии днем траура Lenta.Ru
 116

2010.04.09 14:42 Бишкек после погрома: мародеры бесчинствуют. ФОТО MlNews.com.ua
- 4. Делегация временного правительства Киргизии проведет встречи в Москве**
 Бишкек 9 апреля. ИНТЕРФАКС - Делегация временного правительства Киргизии вылетела в Москву для проведения переговоров, сообщил "Интерфаксу" источник в правительстве республики. При этом собеседник агентства не уточнил, какие встречи запланированы в Москве и каков их уровень. Делегацию возглавляет заместитель главы временного правительства Киргизии по вопросам экономики Алмазбек Атамбаев.
 Сюжет полностью (61)

2010.04.08 15:30 "Эйр Астана" приостанавливает воздушное сообщение Today.kz
 61

2010.04.09 15:07 В Москву вылетела делегация временного правительства Киргизии Говорит Москва

Рис. 29. Пример отображения инфосюжетов в системе InfoStream

5.2.4. Эмерджентность информационных систем

В сложных системах (а современные информационные системы, без сомнения, являются таковыми) среди многих других характеристик наиболее четко проявляется целостность, т.е. наличие таких свойств, которые не присущие ни одному элементу (в рассматриваемом случае — документу), составляющему систему, взятому отдельно. Свойство, которое называют «эмерджентностью», — результат возникновения между элементами системы особенных синергических связей. Под термином «эмерджентность», введенном Ф. Льюисом понимается то, что в системах целое является чаще всего большим, чем сумма частей [71], т.е. на каждом уровне сложности возникают новые, часто непредвиденные качества, которые не свойственны составным частям.

Так, если в качестве системы рассматривать часы — прибор, который показывает текущее время, то ни одна из его деталей не сможет показывать время. Она не может показывать даже «часть времени». Свойство показывать время появляется у всех деталей вместе, причем после того, как они будут определенным образом собраны в единый комплекс и, тем самым, вступят друг с другом в определенное взаимодействие.

Эмерджентность информационной системы позволяет не ограничиваться изучением ее элементов и связей между ними, а выполнять целостный анализ всей системы. До конца XX века при анализе сложных, в том числе и социальных систем, в основном использовался редукционистский подход, который объяснял множество свойств сложных систем свойствами их элементов — «атомов» или «молекул». В результате развития системного анализа, появления науки о сложности технологического прорыва в вычислительных возможностях ситуация резко изменилась. В настоящий момент развитие получили такие направления, как теории хаоса, сложных сетей, нелинейных и самоорганизующихся систем. Оказалось, что многие свойства сложных систем не могут быть выведены из предварительно определенного набора динамических уравнений. Наоборот, уравнения могут быть получены только вследствие численного моделирования.

Вместе с тем очевидно, что невозможно разработать и применять на практике некоторую универсальную методику моделирования информационных систем. В основном это связано со слабой формализацией многих понятий и факторов, в первую очередь субъективных.

В каждом отдельном случае приходится доверять информированности и интуиции аналитиков, которые профессионально занимаются вопросами анализа информационных процессов.

С объективными факторами дело обстоит иначе. Они полностью поддаются анализу на статистическом уровне и допускают количественные оценки, которые могут использоваться для построения обоснованных прогнозов. Современные методы прикладной статистики, анализа временных рядов включают большой арсенал детально проработанных и апробированных методов. Однако статистика позволяет описывать лишь формальные аспекты изучаемых явлений, оставляя «за бортом» аспекты содержательные. Поэтому существует необходимость расширения набора инструментальных средств, которые используются при анализе и моделировании информационных потоков. Одним из наиболее перспективных направлений в этом плане является математическое моделирование. Сегодня математическое моделирование широко применяется во многих отраслях науки и техники, тогда как моделирование информационных систем остается открытой проблемой.

В области информационных систем перспективно моделирование, обусловленное некоторыми реалистическими правилами поведения отдельных элементов (документами, тематиками), которые уточняются некоторыми параметрами, изменяемыми при моделировании. В этом случае большую ценность приобретает также и обратная задача — по реальному поведению некоторой зависимости оценить величину параметров модели. Знание общего поведения устойчивых решений позволяет прогнозировать развитие информационных систем даже в том случае, когда не существует точного представления о конкретных механизмах, которые определяют их динамику, причем такого рода прогнозы могут оказаться точнее, чем полученные традиционными экспертными методами. Если решения оказываются неустойчивыми, то из этого также может быть получена важная информация о системе, позволяющая в отдельных случаях прогнозировать, в какую сторону направлена динамика отдельных информационных систем.

Попытки моделирования информационных потоков осуществлялись уже давно, но они тормозились вычислительными трудностями, особенно в случае необходимости описания динамики систем с обратными связями. Сегодня существует достаточное количество возможностей для эффективной компьютерной обработки данных, что позволяет, с одной стороны, подготавливать наборы входных пара-

метров на основании анализа результатов статистических исследований, а с другой — решать формализованные задачи с достаточной степенью точности и в допустимое время. Все это дает основания полагать, что в ближайшее время математическое моделирование станет основным инструментальным средством анализа и управления информационными потоками.

Одним из применений концепции эмерджентности к моделированию сегодня является многоагентагентное моделирование (см. гл. 2). Многоагентные модели широко применяются для анализа децентрализованных систем, закономерности динамики функционирования которых не изучены в достаточной мере. Эти модели используются для изучения общего поведения сложных систем, выявления правил их функционирования с учетом предположений об индивидуальном поведении ее отдельных компонентов.

5.2.5. Синергетический подход

Под влиянием внешней среды информационные системы могут переходить к непредсказуемому поведению — хаосу.

Неупорядоченное, непрогнозируемое, случайное поведение системы связывают с недетерминированным хаосом, при котором невозможно вывести закономерности определения будущего состояния системы, зная ее предыдущее состояние. Сегодня все большее внимание ученых обращается к детерминированному хаосу, который порождается не случайным поведением большого количества элементов системы, а внутренней сущностью нелинейных процессов. Поведение информационных систем в полной мере соответствует определению детерминированного хаоса. Для сложных систем, которыми, безусловно, являются информационные системы, уравнения, описывающие их поведение оказываются настолько сложными, что не могут решаться аналитическими методами. Поэтому их исследование обычно проводится средствами компьютерного моделирования.

При решении нелинейных задач состояние системы и степень ее организованности изображают с помощью так называемого фазового пространства, координатами в котором являются параметры, характеризующие систему. Например, для описания систем в механике как координаты фазового пространства используют положение отдельных точек и их скоростей. В этом случае детерминированный хаос отображается непрерывной траекторией, которая порой может постепен-

но заполнять все фазовое пространство (любую малую окрестность точки в фазовом пространстве будет пересекать множество фазовых траекторий). Это свойство детерминированного хаоса приводит к понятию фракталов, фрактальной размерности, например, хаусдорфова размерность траектории, плотно покрывающей плоскость, не может быть целым числом. Дробная размерность — это один из основных признаков фракталов.

Ключевыми понятиями синергетики являются «бифуркации» и «аттракторы». Под точкой бифуркации обычно понимают состояние системы, после которого допустимо некоторое множество вариантов ее развития. Та траектория, или то множество траекторий, по которым возможно развитие системы после точки бифуркации, и которые отличаются от других относительной устойчивостью, называются аттракторами. Другими словами, аттрактор будто бы притягивает к себе множество траекторий, возможных после точки бифуркации. Свойства точек бифуркации и аттракторов изучаются в теории сложных систем, где устанавливаются закономерности развития таких систем, переходы от хаоса к порядку и наоборот.

Действительно, бифуркация может привести к хаосу. Опишем абстрактный, но довольно убедительный пример, каскад бифуркаций М. Фейгенбаума (M. Feigenbaum), один из типичных сценариев перехода от простого периодического режима к сложному аperiodическому при бесконечном удвоении периода [72]. Последовательность Фейгенбаума имеет самоподобную, фрактальную структуру — увеличение какой-нибудь области демонстрирует сходство выделенного участка со всей структурой.

Фейгенбаум анализировал в основном логистическое уравнение $X_{n+1} = CX_n - CX_n^2$ (C — внешний параметр), откуда вывел, что при некоторых ограничениях во всех подобных уравнениях происходит переход от равновесного состояния к хаосу.

Логистическое уравнение, которое, как известно, имеет два устойчивых решения, обычно трактуется как условие популяционной динамики и допускает следующую трактовку: предполагается, что изолировано существует популяция особей удельной численностью X_n . Через год появляется потомство удельной численностью X_{n+1} . Возрастание популяции описывается первым членом правой части уравнения (CX_n), где коэффициент C определяет скорость возраста-

ния и является определяющим параметром. Убывание численности популяции (за счет перенаселенности, недостатка пищи и т.п.) определяется нелинейным членом (CX_n^2). Результаты расчетов (рис. 30) показывают, что:

- при $C < 1$ популяция с увеличением n вымирает;
- в области $1 < C < 3$ численность популяции приближается к постоянному значению $X_0 = 1 - 1/C$, которая является областью стационарных, фиксированных решений. При значении $C = 3$ точка бифуркации становится отталкивающей фиксированной точкой;
- в диапазоне $3 < C < 3,57$ начинают появляться бифуркации — разветвление каждой кривой на две (действительно, логистическое уравнение имеет два устойчивых класса решений). Численность популяции колеблется между двумя значениями, которые лежат в этих областях. Сначала популяция резко возрастает, на следующий год возникает перенаселенность и через год численность снова уменьшается;
- при $C > 3,57$ происходит перекрытие разных областей, и поведение системы становится на вид хаотичным.

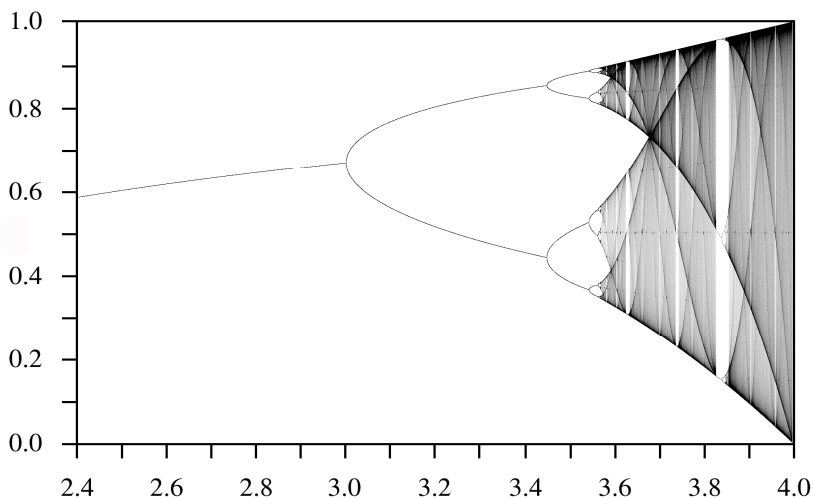


Рис. 30. Каскад бифуркаций (последовательность Фегенбаума): ось абсцисс — значение параметра C , ось ординат — значение X_n

Таким образом, заключительным состоянием эволюционирующих физических систем является состояние динамического хаоса.

С использованием теории бифуркаций можно прогнозировать характер поведения, проявляющееся при переходе системы в качественно другое состояние, а также область существования системы и оценить ее устойчивость.

Бифуркации очень часто возникают при переходе системы от состояния видимой стабильности и равновесия к хаосу. На основе хаоса не только нельзя построить или уточнить прогноз, но и, соответственно, проверить его. Однако это не должно указывать на неверность теории хаоса, подтвержденной как в математических расчетах, так и в жизни. В настоящее время еще не существует математически точного применения теории хаоса для исследований информационных потоков, вместе с тем эта теория уже сегодня способна предусмотреть переходы моделей систем, представленных в аналитическом виде, в хаотичное состояние. Очевидно, это действительно одно из самых перспективных направлений прикладных исследований информационных процессов.

В синергетике строго доказывается, что никакими внешними влияниями невозможно «навязать» системе желаемое для кого-нибудь поведение — можно лишь наилучшим образом выбрать подходящую траекторию из потенциально заложенных [73].

Поэтому при планировании и моделировании информационных систем одна из основных задач — нахождение точек бифуркации информационных процессов и формирование флюктуаций, обуславливающих выбор требуемой траектории эволюции (аттрактор).

Как известно, в математике катастрофами называются скачкообразные изменения, возникающие как ответ системы на плавное изменение внешних условий. Информационные системы могут вызвать процессы, которые лучше всего описываются в рамках теории катастроф: «близ точек бифуркации в системах можно наблюдать значительные флюктуации. Такие процессы будто колеблются перед выбором одного из нескольких путей эволюции. Небольшая флюктуация может служить началом эволюции в полностью новом направлении, которое резко изменит все поведение макроскопической системы» [74]. Это объясняет, почему так тяжело препятствовать катастрофе, когда ее признаки уже заметны: скорость ее приближения в это время беспредельно возрастает [75].

Если рассматривать общество как сложную систему, то информационные системы можно рассматривать в качестве средств воздействия на нее, как было показано выше для выбора определенных путей развития. Таким образом, модели информационных систем являются частью более общих социальных моделей.

5.2.6. Теоретико-игровой подход

Характерной особенностью многих информационных систем является то обстоятельство, что их компоненты (акторы) находятся в состоянии конфликта интересов, и при этом действуют в условиях отсутствия полной информации о намерениях друг друга. В частности, при анализе информационных процессов практически всегда приходится анализировать конфликтные ситуации, в которых сталкиваются интересы двух или более конкурирующих сторон, преследующих различные цели. Математической теорией, которая посвящена изучению конфликтных ситуаций, является теория игр. Таким образом, представляется вполне естественным попытаться применить к изучению информационных систем теорию игр. В обобщенной игре (в качестве игры, например, может рассматриваться некоторая совокупность информационных операций противоборствующих сил) могут сталкиваться интересы двух или нескольких противников. При этом игроки могут образовывать коалиции, в этом случае игра — коалиционная.

Структура любой игры описывается тремя блоками:

- 1) допустимые множества ходов или стратегий участников;
- 2) цели участников;
- 3) тип поведения и информированности участников.

В теории игр игры классифицируются как кооперативные (коалиционные) и некооперативные.

В кооперативных играх участники могут объединяться в группы, беря на себя некоторые обязательства перед другими игроками и координируя свои действия. Этим они отличаются от некооперативных игр, в которых каждый обязан играть за себя.

Из двух типов игр, некооперативные описывают ситуации в мельчайших деталях и выдают более точные результаты. Кооперативные рассматривают процесс игры в целом.

Ниже остановимся на некооперативных играх. Формально некооперативной игрой называется тройка $\Gamma = \langle I, X_i, H_i \rangle$, где I —

множество участников игры; X_i — множество стратегий участника, $i \in I$; H_i — функция выигрыша участника, $i \in I$, определенная на множестве ситуаций (конкретных реализациях стратегий всех участников игры), отображающая его на множество действительных чисел.

Некооперативная игра предполагает следующий порядок разыгрывания:

1. Игроки, одновременно и независимо друг от друга, выбирают из множеств X_i свои стратегии. Вектор стратегий $x = (x_1, x_2, \dots, x_n)$ всех игроков представляет собой ситуацию в игре.

2. Каждый участник получает выигрыш, определяемый значением функции $H_i(x)$; на этом взаимодействие между игроками прекращается.

Анализ игры заключается в умении прогнозировать решение игры — множество возможных ходов и их результатов. Важными понятиями в теории игр являются также оптимальная стратегия, цена игры, средний выигрыш. В частности, стратегии P^* первого и Q^* второго игроков называются оптимальными, а число V — ценой игры, если для любых стратегий P первого и Q второго игроков выполняются неравенства [76]:

$$M(P, Q^*) \leq V \leq M(P^*, Q),$$

где $M(P, Q)$ — математическое ожидание выигрыша первого игрока, выбравшего стратегию P , при условии, что вторым выбрана стратегия Q .

Во многих задачах из теории игр неопределенность вызвана не противодействием противника, а недостаточной осведомленностью игрока об условиях, в которых действуют стороны, например, об информационных операциях против него, информационных воздействиях. Такие игры принято называть «играми с природой», при решении которых используют так называемые матрицы рисков. В рамках данного подхода действия интеллектуального агента, принимающего решения (англ. — *intelligent rational decision-maker*), определяются его информированностью о состоянии окружающей среды и о представлениях оппонентов. Элементом информационного воздействия при

этом могут быть как переданные агенту сведения об окружающей среде, так и о представлениях оппонентов.

Информация, передаваемая агенту для информационного воздействия, может представлять собой [77]:

- «сухие» факты;
- логически обоснованные выводы, аналитические суждения, опирающиеся на определенный набор фактов;
- эмоционально окрашенные утверждения.

В качестве данных агенту может передаваться также прогноз, зависящий от неопределенного параметра и действий самого агента. Каждый агент на основании «активного прогноза» может «восстановить» информацию об окружающей среде и использовать ее при принятии решений (например, при вычислении равновесных действий).

При решении задач в условиях неопределенности, когда вероятности отдельных частных исходов неизвестны, возникают трудности при математическом моделировании. Поэтому теория принятия решений, в частности, рекомендует применять подход, базирующийся на известной теореме Байеса. Стратегия оптимизации в таких случаях строится на основе байесовской теории принятия решений. При этом принятая в теории игр функция потерь рассматривается как обобщение вероятности ошибки. Соответственно, предполагается выбирать решение, минимизирующее функцию потерь. Байесовский подход к оценке вероятностных связей играет решающую роль в теории принятия решений в условиях неопределенности последствий этих решений или в условиях противодействия со стороны природы, или конкуренции. В этих условиях ключевой является стратегия управления, основанная на апостериорной (послеопытной) вероятности события. Обязательное условие корректности такого подхода — постоянное обучение системы. Стратегия управления вначале строится на базе определенных представлений о вероятности событий, а по мере функционирования системы реализуется коррекция управления — использование накапливаемого опыта путем перерасчета вариантов стратегий с учетом изменившихся значений вероятностей.

Отметим, что применение теории игр имеет два различных аспекта: во-первых, она может использоваться в целях оптимизации механизмов принятия решений противоборствующими сторонами, а во-вторых — для выработки принципов их организации. В частности, во втором случае крайне актуальным становится вопрос об устойчивости

игры (которой мы описываем электоральный процесс) в смысле Нэша. Формально равновесие по Нэшу определяется следующим образом [78]. В случаях, когда свои ожидания о поведении партнера каждый игрок строит по прошлому опыту подобных игр, устойчивое в каком-то смысле решение игры называют равновесием этой популяции. Тогда особое значение приобретает равновесие по Нэшу — профиль стратегий, от которого никому не выгодно отклоняться, если партнеры не отклоняются, т.е. игра называется устойчивой в смысле Нэша, если ни один из игроков не может увеличить свой выигрыш только в результате своих собственных действий. Нэшевское равновесие (*NE*) — точка, из которой ни одному игроку нет пользы уходить при текущих ходах партнеров, а строгое Нэшевское равновесие (*SNE*) — точка, из которой невыгодно уходить. Когда каждый игрок $i \in I$ выбирает стратегию x_i из вектора стратегий $x = (x_1, x_2, \dots, x_n)$, игрок i получает выигрыш $H_i(x)$. При этом выигрыш i -го участника игры зависит от всего профиля стратегий: не только от стратегии, выбранной самим игроком i , но и от чужих стратегий. Вектор стратегий x^* является равновесием по Нэшу, если изменение своей стратегии не выгодно ни одному игроку, т.е. для любого i справедливо условие:

$$H_i(x^*) \geq H_i(x_i, x_{-i}^*).$$

Здесь x_i, x_{-i}^* — вектор, составленный из всех координат вектора x^* кроме i -й, которой соответствует значение x_i .

Игра может иметь равновесие Нэша в чистых или в смешанных стратегиях (т.е. при выборе чистой стратегии стохастически с фиксированной частотой). Нэш доказал, что если разрешить смешанные стратегии, то в каждой игре n игроков будет хотя бы одно равновесие Нэша.

При анализе социальных процессов часто рассматриваются ситуации несимметричных условий для разных игроков. В таких случаях имеет смысл рассмотреть равновесие Штакельберга [79], которое в отличие от симметричных условий предполагает различные принципы формирования ожиданий разных игроков.

Первый игрок (лидер) ориентируется на оптимальные ответы партнеров, зная их предпочтения, а остальные играют, как в случае Нэшевского равновесия, лишь реагируя на его ход и на ходы друг

друга. Равновесие Штакельберга может возникать, например, когда один из игроков делает свой выбор раньше других и знает их цели. Или когда он один, а однотипных ведомых игроков достаточно много, чтобы каждый из них не мог просчитывать общие последствия своего хода.

Рассмотрим частный случай модели Штакельберга — борьбу двух информационных систем за электоральные предпочтения людей. Пусть существует две партии, одна из которых «лидер», а другая — «преследователь». Пусть затраты на избирательную кампанию являются линейной функцией общего количества электората Q :

$$P(Q) = a - bQ.$$

Предположим также, что издержки каждой из двух партий (реклама, информационные системы, локальные информационные операции и т.п.) на одного сторонника постоянны и равны соответственно c_1 и c_2 . Тогда условная «прибыль» первой партии будет определяться формулой

$$\Pi_1 = P(Q_1 + Q_2) \times Q_1 - c_1 Q_1,$$

а количество голосов второй

$$\Pi_2 = P(Q_1 + Q_2) \times Q_2 - c_2 Q_2.$$

В соответствии с моделью Штакельберга, первая партия — лидер — на первом шаге добивается количества своих сторонников Q_1 . После этого вторая партия — преследователь — анализируя действия лидера, добивается количества сторонников Q_2 . Цель обеих партий — максимизация условной прибыли.

Равновесие Нэша в этой игре определяется методом обратной индукции. Рассмотрим предпоследний этап игры — ход второй партии. На этом этапе вторая партия знает оптимальное количество сторонников первой партии Q_1^* . Тогда задача определения оптимального количества своих сторонников Q_2^* сводится к решению задачи нахождения точки максимума функции прибыли второй партии. Максими-

зируя функцию Π_2 по переменной Q_2 , при заданном Q_1 , находим, что оптимальное количество сторонников второй партии составляет (будем считать, что $c_1 = c_2 = c$):

$$Q_2^* = \frac{(a - bQ_1^* - c)}{2b}.$$

Это наилучший ответ партии-преследователя на выбор лидером значения Q_1^* . Партия-лидер может максимизировать свою функцию прибыли, учитывая вид функции Q_2^* . Точка максимума функции Π_1 по переменной Q_1 при подстановке Q_2^* определяется так:

$$Q_1^* = \frac{(a - c)}{2b}.$$

Подставляя это в выражение для Q_2^* , получаем

$$Q_2^* = \frac{(a - c)}{4b}.$$

Таким образом, при равновесии партия-лидер должна приобрести в два раза больше сторонников, чем партия-преследователь.

Следует иметь в виду, что модели теории игр в меньшей степени, чем многие другие из рассмотренных, могут использоваться для точных расчетов и прогнозов. Скорее здесь можно говорить о хорошо обоснованной методологии, которая может существенно повысить эффективность действий участников социальных процессов. Эти модели, по сути, представляют собой наборы рекомендаций, которые дают заметные преимущества тем, кто их использует.

5.2.7. Экстремальные подходы

Экстремальные подходы в моделировании поведения сложных систем находят широкое применение в естественных науках, в последнее время наибольшее развитие получили на стыке экологии и

биологии [80], где они с успехом используются для изучения популяционной динамики — развития отдельных популяций.

Исследования, проводимые в этой области, почти без технических изменений могут применяться для изучения человеческих сообществ, социальных процессов, динамики электоральных популяций, в частности, под влиянием информационных операций.

В соответствии с экстремальными подходами к моделированию, реализуются лишь те состояния систем, которые соответствуют экстремумам некоторой целевой функции (описываемой уравнениями) при определенных граничных условиях. Самым тонким вопросом при этом являются принципы составления уравнений, которые в случае исследования информационных операций (как впрочем и в других областях) базируются на опыте экспертов, аналогиях, неполных эмпирических закономерностях.

При моделировании информационных систем могут применяться подходы, основывающиеся на логистических уравнениях возрастания популяций [81], получаемых в результате решения оптимизационных задач, принципы стационарного состояния открытых систем [82], максимального разнообразия популяции [83], максимальной обобщенной энтропии [84], максимума мальтузианского параметра [85] и многие другие.

Остановимся подробнее на некоторых из существующих подходов.

Принцип выживания

При исследовании динамики информационных систем в качестве критерия оптимальности можно использовать принцип выживания (сохранения компонент информационной системы и их функционального значения), применяя математический аппарат, предложенный в [86].

Предполагается, что динамику информационной системы адекватно описывает система уравнений, в качестве параметров которых выступают некоторые социальные условия, а также структурно-функциональные параметры всех информационных систем. Выделяют s -ю информационную систему и некоторый структурный или функциональный параметр α_{s_k} этой системы. Можно сделать предположение о том, что информационная система состоит из двух подсистем, различающихся значением некоторого фенотипического параметра (ха-

раактеристиками, присущими компонентам на определенной стадии развития). Пусть $x_s^{(1)}$, $x_s^{(2)}$, $\alpha_{s_k}^{(1)}$, $\alpha_{s_k}^{(2)}$ — количества компонент и функциональные параметры двух подсистем.

Моделирование системы, в которую внесены соответствующие изменения, учитывающие различия данного функционального параметра у компонент s -й популяции, дает возможность анализировать асимптотические свойства размера подсистем. Один из возможных вариантов поведения — вытеснение второй подсистемой первой, когда параметр $\alpha_{s_k}^{(1)}$ имеет преимущество по сравнению с $\alpha_{s_k}^{(2)}$ в заданной информационной ситуации, т.е.:

$$\lim_{t \rightarrow \infty} x_s^{(1)} > 0, \lim_{t \rightarrow \infty} x_s^{(2)} = 0.$$

Оптимальным значением параметра $\alpha_{s_k}^{(1)}$ с точки зрения вытеснения второй подсистемой является такое $\alpha_{s_k}^*$, при котором для любого отличного от этого значения параметра $\alpha_{s_k}^{(2)} \neq \alpha_{s_k}^*$ выполняются приведенные выше условия для любых начальных состояний системы. В некоторых случаях информационная система может не обладать таким оптимальным значением параметра, т.е. система может стабильно существовать при любом значении параметра α_{s_k} , принадлежащего области, соответствующей условию стабильного существования информационной системы, даже при значении, не равном оптимальному. Оптимальное же значение устанавливается в результате конкуренции особей с различными значениями рассматриваемого функционального параметра. Именно вследствие этой конкуренции компоненты систем с неоптимальными значениями параметра $\alpha_{s_k} \neq \alpha_{s_k}^*$ «покидают» информационную систему [87].

Используя критерий отбора, в рассматриваемом случае необходимо учитывать ограничения, вытекающие из информационных закономерностей процесса. Для дальнейшего моделирования в качестве критерия оптимальности может выступать простейшее требование максимума относительной скорости увеличения размера информационной системы:

$$k = \frac{d \ln x}{dt} = \max .$$

Такой критерий может быть применен для определения оптимальных значений структурно-функциональных параметров, если относительная скорость возрастания количества компонент информационной системы представлена в виде функции этих параметров.

Принцип максимальной неожиданности

Один из путей исследования динамики биологических популяций — изучение «дарвиновских систем», описывающих динамику естественного отбора. Рассмотрим, как он может применяться для моделирования информационных систем. В работе Е.В. Евдокимова [88] приведен способ описания дарвиновских систем (ДС) по Эйгену. Такие системы являются открытыми, состоящими из самокопирующихся с небольшим количеством ошибок единиц различных видов, использующих для своего размножения свободную энергию поступающих извне питательных компонент. В случае информационных систем в качестве такого внешнего энергетического воздействия можно рассматривать информационные воздействия, в том числе и информационные операции. При данном подходе ограничением может являться постоянство суммарной численности элементов системы (размера информационной системы). Для описания ДС используют дифференциальное уравнение

$$\dot{x}_i = x_i(A_i Q_i - \Delta_i) + \sum_{j \neq i}^w u_{ij} x_j - F_i ,$$

которое упрощается до следующего выражения:

$$\dot{y}_i = y_i(\mu_i(s) - D),$$

где $i, j = 1, 2, \dots, w$ ($w = \text{const}$) — количество популяций в системе; $s = \{s^1, s^2, \dots, s^m\}$ — концентрации «питательных компонентов» (объемы воздействий); $\mu_i(s)$ — удельная скорость увеличения i -й популяции; D — скорость протока в системе. В зависимости от наложен-

ных ограничений различают ДС с постоянной организацией, у которых сумма $\sum_{i=1}^w y_i$ и концентрация s постоянны, и ДС с постоянным протоком, характеризующиеся условием $D = \text{const}$.

Для разрешения проблемы неполноты приведенных выше уравнений и труднодоступности информации на микроуровне предлагается использовать постулат [88], заключающийся в том, что «процесс эволюции ДС протекает наименее неожиданным способом» (принцип минимальной неожиданности протекания эволюции).

В качестве целевой функции используется функция неожиданности эволюции ДС с постоянной организацией:

$$I(P_i(t)/P_{i0}) = P_i(t) \log(P_i(t)/P_{i0}),$$

где $P_i(t) = P_\mu(\mu = \mu_i, t)$ — вероятность того, что случайно выбранная в момент t популяция i ($i = 1, 2, \dots, w$) имеет мальтузианский параметр μ_i , $P_{i0} = P_i(t)|_{t=0} = y_i / \sum_{k=1}^w y_{k0}$, значения y_{k0} задаются экспериментально.

В этом случае вариационная задача формулируется следующим образом:

$$\begin{cases} \sum_{i=1}^w I(P_i(t)/P_{i0}) \rightarrow \min; \\ \delta[I(P_i(t)/P_{i0})] = 0. \end{cases}$$

Решение было получено методом неопределенных множителей Лагранжа:

$$P_i(t) = \frac{P_{i0} e^{\mu_i t}}{\sum_{k=1}^w P_{k0} e^{\mu_k t}},$$

причем доказано, что оно соответствует решению приведенной выше системы уравнений для \dot{y}_i . Придавая множителю Лагранжа

$\lambda_0 = \log \sum_{i=1}^w P_{i0} e^{\mu_i t}$ информационный смысл, можно получить «основную теорему естественного отбора» Фишера:

$$\frac{d\hat{\mu}}{dt} = \sigma_{\mu}^2.$$

Кроме того, доказано, что множитель λ_0 пропорционален «энергопотреблению» популяции.

Таким образом, исходя из эвристического принципа минимальной неожиданности протекания эволюционного процесса, получены результаты, описывающие динамику отбора в дарвиновских системах, которые полностью идентичны уравнениям, выводимым из кинетики размножения и конкуренции, а множители Лагранжа, используемые для решения вариационной задачи, вполне осмыслены и имеют прогностическую ценность.

Принцип максимума параметра Мальтуса

Пусть информационное пространство состоит из w информационных систем. Информационное пространство может быть описано численностями составляющих его информационных систем x_i . Пусть

$x = \sum_{i=1}^w x_i$ — суммарное количество элементов информационного про-

странства. Допустим, что в течение некоторого локального интервала времени i -я информационная система характеризуется мальтузианским параметром $\mu_i(t)$ из уравнения $dx_i/dt = \mu_i x_i$. Пусть $p_i = x_i/x$ — относительная доля i -й информационной системы в информационном пространстве. Тогда набор $p = \{p_1, p_2, \dots, p_w\}$ назовем структурой информационного пространства; величину $\hat{\mu} = (\mu, p)$ — средним мальтузианским параметром ($(\mu, p) = \mu_1 p_1 + \mu_2 p_2 + \dots + \mu_w p_w$), а динамика развития размера информационного пространства описывается уравнением $dx/dt = \hat{\mu}x$.

В основе данного подхода постулируется принцип максимума среднего мальтузианского параметра, т.е. то, что сообщество взаимодействующих информационных систем эволюционирует таким образом, что его средний мальтузианский параметр всегда возрастает, достигая в устойчивом равновесии своего максимума. В работе [87] структуры разделяются на вероятные и маловероятные, а также определены условия, при которых в процессе адаптации из сообщества элиминируются все популяции, кроме одной или ни одна из популяций не покидает сообщества (в нашем случае информационного пространства).

5.3. Нелинейные динамические модели

Известно, что если состояние системы не изменяется во времени, ее называют статической, в противном случае — динамической. Понятно, что в плане изучения информационных систем основной интерес представляют именно динамические системы: в первую очередь нас интересуют происходящие изменения, а статические системы никаких изменений не порождают. В свою очередь, динамические системы делятся на два класса: линейные и нелинейные. Линейными называют системы, характеристики которых зависят от изменения состояний этих систем. И наоборот, характеристики нелинейных систем зависят от таких изменений.

Очень часто нелинейные социальные системы проявляют себя, прежде всего, несоразмерностью отклика на внешнее воздействие. Хорошо известно, что такие системы могут поразительно легко и без последствий переносить тяжелые потрясения, и вместе с тем мгновенно «пойти в разнос» от малозначимого события или влияния. Именно информационные операции (инструментом и объектом которых являются информационные системы) могут рассматриваться как подобное влияние или воздействие.

Модели, используемые применительно к нелинейным системам, также называют нелинейными, так как модель — это тоже система, и она, естественно, может быть нелинейной. Нелинейность модели может быть формально выражена в структуре используемых уравнений, а их решение в некоторых случаях может быть вполне осуществимой задачей, по меньшей мере, в числовом виде.

Основной акцент при построении нелинейных конкурентных моделей информационных процессов в настоящее время делается на анализе принципиальных внутренних взаимодействий динамических систем на основании логистических моделей. Моделирование динамики развития на основе дифференциальных логистических уравнений широко используется для моделирования самых разнообразных как естественных, так и информационных процессов.

Естественно, перед применением математических моделей необходимо обосновать их адекватность. Для этого используются хорошо известные методики, в частности, ретроспективный анализ.

Чаще всего для моделирования сложных систем применяют дифференциальные уравнения, описывающие динамику изменения состояний таких систем. Как правило, это система уравнений первого порядка, имеющая вид [89]

$$dx_i / dt = f_i(X, a),$$

где $X = (X_1, \dots, X_n)$ — вектор переменных, характеризующих состояние социальной системы; a — вектор параметров системы; t — время.

Решения приведенной системы уравнений обычно представляют в виде траекторий в фазовом пространстве. Если фиксировать значения всех параметров, т.е. выбрать точку в параметрическом пространстве, то решения приведенной системы дифференциальных уравнений будут зависеть только от начальных условий. Однако для качественного подхода важны не столько частные решения, сколько по возможности наиболее полное описание поведения системы во всем динамическом пространстве [90].

Эта общая картина будет преимущественно зависеть от значений, к которым стремиться решения при $t \rightarrow \infty$ или при $t \rightarrow -\infty$. Наиболее важные для рассматриваемой предметной области асимптотические решения такого типа — стационарные точки и предельные циклы. При этом наибольший интерес представляет специальный вид устойчивости системы: устойчивость по отношению к изменениям параметров системы. Система, общий динамический характер которой не изменяется при малых изменениях параметров называется грубой или «жесткой» (англ. — *hard*). В противном случае системы называются «мягкими».

Первым шагом при исследовании приведенной выше системы дифференциальных уравнений является определение стационарных точек, т.е. решение системы уравнений:

$$f_i(X, a) = 0.$$

Вторым шагом исследований является определение характера особых точек. Для этого переходят к новым переменным — отклонениям от координат стационарной точки:

$$u_i = x_i - x_i^0.$$

Вблизи стационарной точки правые части исходных уравнений системы можно разложить в ряд Тейлора:

$$\dot{u}_i = \left. \frac{\partial f_i}{\partial x_j} \right|_{x_m^0} u_j + \left. \frac{\partial^2 f_i}{\partial x_k \partial x_l} \right|_{x_m^0} u_k u_l + \dots$$

Поскольку вблизи x_i^0 выполняется: $u_i \ll 1$, то во многих случаях можно ограничиться исследованием линейной системы:

$$\dot{u}_i = \left. \frac{\partial f_i}{\partial x_j} \right|_{x_m^0} u_j = a_{ij} u_j.$$

В качественной теории это исследование сводится к определению собственных значений матрицы коэффициентов a_{ij} , т.е. к решению уравнения

$$|a_{ij} - \delta_{ij} \lambda| = 0.$$

В случае, когда все собственные значения λ_i различны и имеют отличные от нуля действительные части ($\text{Re } \lambda_i$), существует «грубая»

стационарная точка. Если все $\operatorname{Re} \lambda_i$ не равны нулю, то используются следующие теоремы:

1. Если все $\operatorname{Re} \lambda_i < 0$, то стационарная точка асимптотически устойчива.

2. Если хотя бы одно $\operatorname{Re} \lambda_i > 0$, то стационарная точка неустойчива.

Вопрос об устойчивости стационарной точки можно решить также на основании критерия Гурвица, без непосредственного вычисления собственных значений. Для этого характеристическое уравнение переписывается в виде

$$a_0 \lambda^n + b_0 \lambda^{n-1} + a_1 \lambda^{n-2} + b_1 \lambda^{n-3} + \dots = 0,$$

где $a_0 = 1$.

Матрицей Гурвица называется матрица n -го порядка:

$$H = \begin{vmatrix} b_0 & b_1 & b_2 \dots b_{n-1} \\ a_0 & a_1 & a_2 \dots a_{n-1} \\ 0 & b_0 & b_1 \dots b_{n-2} \\ 0 & a_0 & a_1 \dots a_{n-2} \\ 0 & 0 & b_0 \dots b_{n-3} \\ \vdots & \vdots & \vdots \end{vmatrix}.$$

Миноры матрицы H (от первого до n -го порядка) называются определителями Гурвица. Критерий Гурвица заключается в следующем: для того чтобы все $\operatorname{Re} \lambda_i < 0$, необходимым и достаточным условием является положительность всех определителей Гурвица. В результате общее исследование характера особых точек исходной системы может быть основано на приведении ее матрицы к жордановой форме.

В простейшем случае для системы второго порядка исходное уравнение может быть записано в виде

$$\lambda^2 - \sigma \lambda + \Delta = 0,$$

где $\sigma = a_{11} + a_{22}$, $\Delta = a_{11}a_{22} - a_{12}a_{21}$.

Тогда

$$\lambda_{1,2} = \frac{\sigma}{2} \pm \frac{1}{2} \sqrt{\sigma^2 - 4\Delta}.$$

В данном простейшем примере ограничимся рассмотрением лишь грубых стационарных точек. Если ни один из параметров σ , Δ не равен нулю, то качественная картина фазового пространства в окрестности стационарной точки зависит только от линейных членов и стационарная точка является грубой. Существуют три грубые положения равновесия:

1. Стационарная точка типа «узел» (рис. 31, а): собственные значения λ_i действительные, одного знака, $0 < \Delta < \frac{\sigma^2}{4}$, $\sigma \neq 0$. При $\sigma < 0$ узел устойчивый, при $\sigma > 0$ — неустойчивый.

2. Стационарная точка типа «седла» (рис. 31, б): собственные значения λ_i действительные, разных знаков, $\Delta < 0$, $\sigma \neq 0$.

3. Стационарная точка типа «фокус» (рис. 31, в): собственные значения λ_i комплексные, $\Delta > \frac{\sigma^2}{4}$, $\sigma \neq 0$. При $\sigma < 0$ — устойчивый фокус, движение около стационарной точки носит характер затухающих колебаний. При $\sigma > 0$ фокус неустойчивый, возникают колебания нарастающей амплитуды.

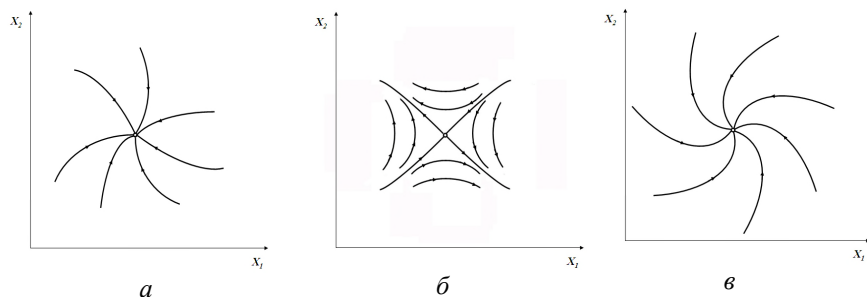


Рис. 31. Фазовые портреты грубых особых точек

Анализ фазовых траекторий позволяет сделать заключение о характере эволюции системы, определять области ее детерминированного поведения и области бифуркаций (т.е. области значений параметров, при которых возникает неустойчивость и происходит изменение вида решений уравнения, описывающего поведение системы).

Сложные системы часто имеют несколько устойчивых состояний (аттракторов), в одном из которых они рано или поздно оказываются. В этих случаях пути эволюции не дискретны: возможен лишь определенный набор путей, соответствующих аттракторам (рис. 32). При этом переходы от одного аттрактора к другому не могут произойти самопроизвольно, для этого необходимо изменение внешних условий или свойств системы. В частности, в социальных системах именно для этого применяются информационные операции.

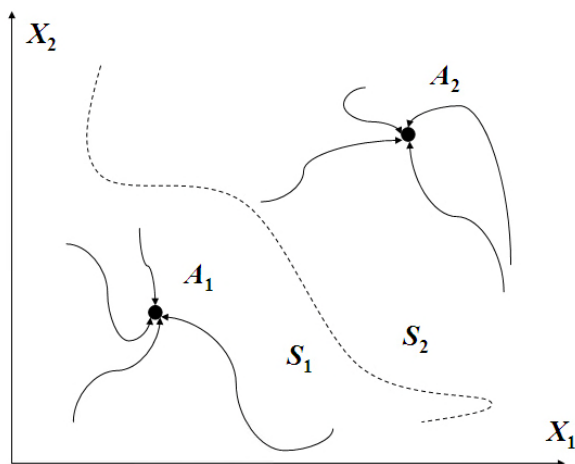


Рис. 32. Пример фрагмента фазового пространства динамической системы (плоскостной срез по двум координатам X_1 и X_2 (A_1 и A_2 — аттракторы, S_1 и S_2 — соответствующие области притяжения аттракторов))

Из синергетики известно, что переход системы из одного состояния в другое проходит фазу хаоса при ослаблении структурирующих процессов. В информационных системах хаос обеспечивает систему первоначальным набором вариантов дальнейшего развития — аттракторами. В периоды хаоса, кризисов, когда возникают возмож-

ности многовариантного развития, информационные операции приобретают решающую роль. В эти периоды социальные и информационные системы наиболее уязвимы относительно информационных воздействий, которые могут сыграть решающую роль в дальнейшем развитии событий [89], выборе того или иного аттрактора.

Таким образом, можно констатировать, что на разных стадиях эволюции информационные системы обладают различной чувствительностью и уязвимостью к внешним воздействиям — информационным операциям, которые становятся наиболее опасными в периоды хаоса — кризисов. В этих случаях даже не слишком интенсивные информационные воздействия могут задать направление развития системы и повлиять на характер ее последующей эволюции. Безусловно, как и для эффективного проведения информационных операций, необходимо знать структуру имеющихся аттракторов, а также владеть набором приемов перевода системы от одного аттрактора в другой.

Рассмотрим конкретные нелинейные модели динамики информационных систем, начиная с простейшей модели роста ($\dot{x} = kx$), которая была предложена Мальтусом для расчета динамики численности населения Земли. Эта модель ведет, как было показано выше, к экспоненциальному возрастанию населения x во времени; она может быть применима, например, к развитию информационных систем на начальном этапе. Решением приведенного уравнения, как известно, является экспонента.

Очевидно, ни один реальный процесс — ни физический, ни информационный, не может развиваться по экспоненциальному закону неограниченное время. Действительно, начиная с некоторого момента, зависимость слишком быстро стремится к бесконечности, которая в природе, по понятным причинам, не реализуется. Следовательно, приходится признать, что рано или поздно, и скорее рано, чем поздно, произойдет некоторая катастрофа, которая изменит характер зависимости и вернет ее в рамки допустимого интервала значений.

В более или менее стабильных системах всегда присутствует элемент самосогласованности, в силу чего на значительных промежутках времени зависимость динамики их развития описывается более сложными уравнениями, содержащими обратные связи. Поэтому характер зависимости со временем изменяется, причем изменения происходят не тривиальным образом. В результате возникают следующие типичные случаи:

— зависимость достигает насыщения, и система переходит в статическое (или, возможно, гомеостатическое) состояние;

— зависимость имеет локальный максимум, за которым следует убывание (в том числе и до нуля);

— устанавливается колебательный режим (обычно затухающий, но возможен и автоколебательный).

В реальной жизни, как правило, динамические системы обладают достаточно эффективными обратными связями, позволяющими корректировать характер происходящих в них процессов и тем самым удерживать их в определенных рамках. Информационные операции, корректируя эти обратные связи в определенные периоды эволюционного процесса, могут достаточно эффективно повлиять на характер поведения всей системы.

Как и в случае информационных потоков, используется широко применяемое обобщение закона Мальтуса, известное как логистическая модель.

5.4. Взаимодействие информационных систем

Логистическое уравнение описывает динамику одной информационной системы, взаимодействующей только с окружающим информационным пространством. В теории популяционной динамики разработана классификация различных форм такого взаимодействия популяций [91–93], в нашем случае — информационных систем.

К основным информационным системам относятся следующие:

— нейтрализм (отсутствие прямого воздействия популяций друг на друга);

— конкуренция (взаимное подавление популяций);

— амменсализм (одностороннее подавление одной популяции);

— хищничество (уничтожение особями одной популяции особей другой);

— симбиоз (продуктивное сосуществование популяций).

В динамике взаимодействующих популяций выделяются две категории воздействий, отличающиеся временным характером:

— фазовые (однократные);

— параметрические (постоянные).

Логистическая модель позволяет вполне удовлетворительно описывать и динамику M взаимодействующих между собой популя-

ций. В общем случае это осуществляется с помощью уже приведенной для случая информационных потоков системы уравнений, но с несколько иным смыслом параметров:

$$\frac{dn_i(t)}{dt} = n_i(t) \left[p_i - \sum_{j=1}^M q_{ij} n_j(t) \right],$$

$$n_i(0) = n_{0i}.$$

Тип процесса, описываемого этой системой уравнений, определяется значением и знаком коэффициентов p_i и q_{ij} . Следует также учитывать, что в каждом уравнении диагональные члены $n_i(t)n_j(t)$ описывают внутривидовое взаимодействие, а перекрестные $n_i(t)n_j(t)$ — межвидовое.

Иными словами, диагональные члены описывают влияние на популяцию внешней среды, в том числе исчерпание доступных ресурсов, а перекрестные — воздействие одной популяции на другую (положительные значения соответствуют благоприятному влиянию, отрицательные — неблагоприятному). Коэффициенты p_i имеют смысл скоростей увеличения соответствующих популяций при отсутствии взаимодействия.

Важным моментом является также поведение популяции при заданных значениях параметров и в отсутствие взаимодействия (например, ее рост ограничен сам по себе).

Приведенная выше система уравнений в принципе может описывать широкий спектр зависимостей, и это, в определенном смысле, является проблемой, так как при желании из ее решений можно «вытащить» все что угодно. Поэтому работа с ней требует взвешенного и ответственного отношения.

Однако решения, характеризующие реальные процессы, обычно относятся к одному из следующих режимов:

- стационарному;
- автоколебательному;
- квазистохастическому.

Как правило, эти режимы в полной мере проявляют себя на достаточно больших (не обязательно бесконечных) промежутках времени. Но переходные процессы, которые предшествуют установлению определенного режима, исключительно полиморфны, их поведение может во многом определять последующую динамику. Именно эти процессы могут рассматриваться как основной объект информационных операций в случае планирования социальных процедур.

Ниже предлагаются результаты моделирования частотных характеристик некоторых информационных процессов в рамках логистической модели. Несомненным преимуществом этой модели является то, что она совмещает в себе простоту исходных формулировок с гибкостью в постановке заданий.

Данное выше описание динамики популяций в рамках логистической модели сначала было сформулировано для биологических систем, однако в настоящее время распространено и на другие области исследований, в том числе на информационные процессы.

Приведенная система логистических уравнений позволяет описывать динамику любого числа популяций, как взаимодействующих между собой, так и находящиеся в изоляции. Но для понимания основных ее закономерностей бывает достаточно ограничиться небольшим их числом. Даже изучение поведения двух взаимодействующих популяций позволяет проследить общие закономерности их динамики, по крайней мере, качественные. Практика показывает, что для разумных прогнозов этого вполне достаточно.

Ниже исследуется динамика трех информационных систем, из которых две считаются основными в том смысле, что их поведение должно иллюстрировать интересующие нас моменты, а третья (дополнительная) введена для демонстрации роли, которую в данном случае может играть общеинформационный контекст. При этом изучаются три основных варианта взаимодействия информационных систем: конкуренция, хищничество и симбиоз.

Поскольку аналитические решения приведенной выше системы уравнений в случаях, когда они могут быть построены, оказываются громоздкими и плохо поддающимися анализу, изначально применяются численные методы, тем более, что графическая форма представления результатов в данном случае наиболее удобна и наглядна. Так как нас интересует качественное поведение зависимостей, приводимые ниже результаты будут представлены в условных единицах.

5.4.1. Динамика типа «Конкуренция»

Конкуренция представляет собой форму взаимодействия популяций, при которой они взаимно подавляют друг друга (в рассматриваемой упрощенной модели ввиду ограниченности общей ресурсной базы). Главной особенностью конкуренции является то, что конкурирующие популяции не оказывают друг на друга непосредственного воздействия. Взаимодействие осуществляется опосредованно — путем вытеснения друг друга из области ограниченных ресурсов. При этом возможно полное подавление одной из популяций, в результате чего она исчезает. Именно конкурентные отношения представляют реальную опасность для информационных систем, и именно они наиболее характерны для основных участников информационных процессов, находящихся в информационном противостоянии.

В зависимости от условий, в которых находятся взаимодействующие информационные системы, и значений определяющих динамику параметров, возможны как различные равновесные состояния системы, так и механизмы их достижения.

Исследования устойчивости систем, популяций в частности к внешним воздействиям проводится на основе моделирования конкурентной борьбы, которой посвящено большое количество работ. В частности, в [94] приведена система нелинейных дифференциальных уравнений, описывающих изменение соотношения сил противоборствующих сторон в результате конкурентной борьбы.

В этой модели было показано, что характер эволюции и самоорганизации систем в решающей мере зависит от таких обстоятельств (условий):

- стационарности функционирования системы;
- замкнутости или открытости системы;
- ресурсной обеспеченности системы.

Ниже приведены наиболее характерные, на наш взгляд, случаи динамики информационных систем при конкурентных отношениях. Конкуренции соответствует приведенная выше система логистических уравнений с положительными значениями коэффициентов q_{ij} .

Равновесное сосуществование без внешних влияний

Будем считать, что взаимодействие между двумя основными силами, с одной стороны, и третьей силой, с другой, в основном сводит-

ся к взаимному ограничению ресурсной базы. Главная конкуренция имеет место между основными силами.

При достаточно малых значениях коэффициентов q_{ij} , описывающих влияние одной конкурирующей информационной системы на другую; при достаточно больших значениях скоростей их роста, каждая система достигает равновесного состояния и стабилизируется в нем. В зависимости от значений других параметров и начального размера информационная система в процессе достижения этого состояния может как увеличиваться, так и уменьшаться.

Как видим из рис. 33, обе основные информационные системы достигают равновесного значения, однако при этом график одной из них возрастает, а другой убывает.

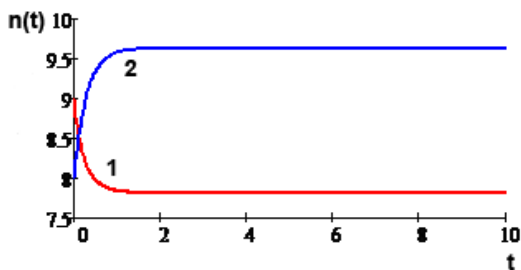


Рис. 33. Равновесное сосуществование информационных систем 1 и 2 в собственных нишах

Может показаться, что приведенные результаты очевидны, но, если учесть, что речь идет о конкурентных отношениях, это отнюдь не так. Оказывается, что обе конкурирующие информационные системы могут оказаться в равновесном состоянии, из которого они сами по себе не смогут выйти ни при каких условиях.

Полное подавление одной информационной системы другой

При значительном увеличении q_{ij} численность одной из информационных систем сокращается до нуля (рис. 34). Если значения коэффициентов q_{12} и q_{21} близки, то ситуация становится неустойчивой в том смысле, что будет подавлена информационная система или нет — зависит от малых отклонений в значениях других параметров.

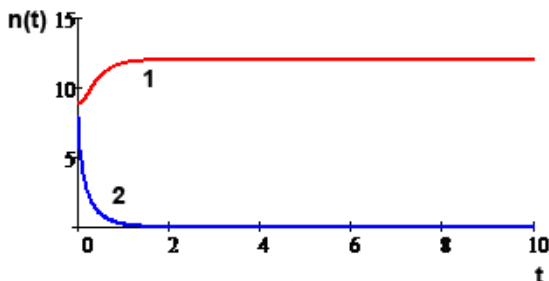


Рис. 34. Полное подавление одной информационной системы 2 информационной системой 1 (здесь t — время, $n(t)$ — размер информационной системы)

Подавление одной из конкурирующих информационных систем другой — сценарий, который чаще всего воспринимается как естественный и закономерный. Поэтому приведенные результаты как таковые не представляют особого интереса. Интересно другое: при выбранных условиях победившая информационная система не исчерпала «освободившиеся ресурсы». Ее увеличение незначительно по сравнению с потерями проигравшей информационной системы.

Таким образом, победа в конкуренции не означает автоматической поддержки со стороны информационных ресурсов конкурирующей стороны.

Равновесное сосуществование за счет третьей силы

В предыдущих случаях пренебрегали непосредственным влиянием третьей силы на основные. Однако, если третья сила тем или иным способом «подпитывает» одну из основных сил, т.е. является информационной операцией первой информационной системы относительно второй, то их динамика может кардинально измениться (рис. 35). Например, возможен сценарий сосуществования.

Равновесное сосуществование конкурирующих основных информационных систем, возникающее за счет положительного воздействия на одну из них дополнительной информационной системы, отражает такую крайне важную ситуацию, как влияние на взаимодействующие информационные системы нового контекста, которое может быть как положительным, так и отрицательным. И в том и ином случае, такое влияние способно эффективно компенсировать сильные и

слабые стороны основных конкурентов, приводя к, казалось бы, неожиданным сценариям.

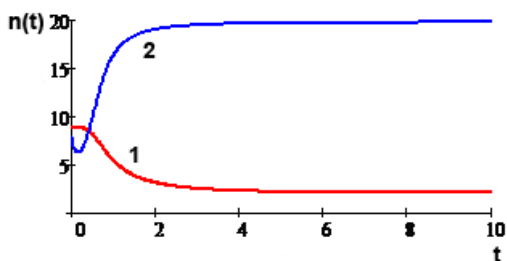


Рис. 35. Равновесное сосуществование информационных систем 1 и 2 за счет воздействия третьей силы

Одним из ключевых вопросов является устойчивость решений по отношению к численным значениям параметров. В ряде случаев поведение, в том числе и качественное, решений приведенной выше системы уравнений очень сильно зависит от них. Проиллюстрируем это на примере зависимости решений от скорости роста (рис. 36).

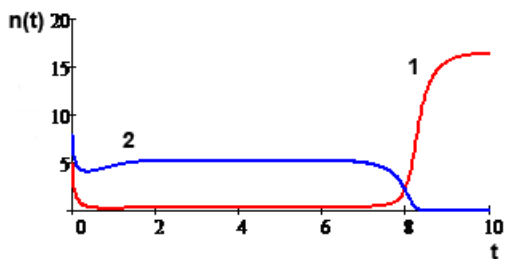


Рис. 36. Зависимость размера информационных систем 1 и 2 от скорости их роста

Представленные графики различаются незначительными изменениями коэффициента p_2 . Как видим, в данном случае поведение приведенных кривых различаются не только количественно, но и качественно. Более наглядно это показано на фазовых портретах

(рис. 37), соответствующих информационным системам, поведение которых отражено на рис. 36.

Таким образом, динамика информационных систем может существенно зависеть от малых изменений скорости их роста. Это означает, что информационная система может победить в конкуренции из-за малого преимущества в этой характеристике.

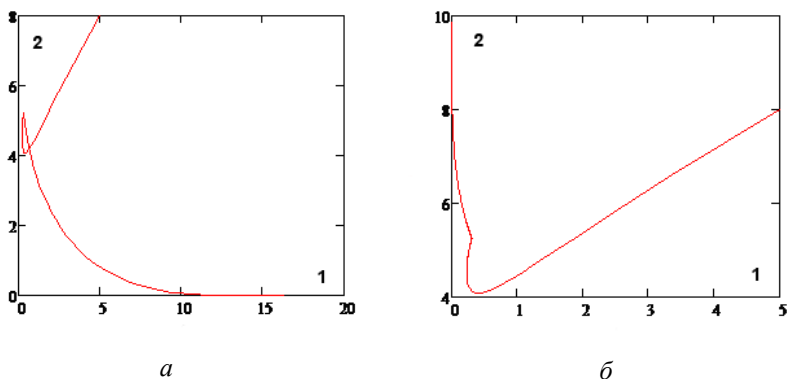


Рис. 37. Фазовые портреты зависимостей размера информационных систем 1 (а) и 2 (б) от скорости их роста

Отметим нетривиальное поведение первой зависимости: вначале доминирует вторая информационная система, а первая — подавлена, но потом они меняются местами. Таким образом, при конкуренции двух информационных систем в присутствии третьей возможны ситуации, когда та из них, которая вначале была бесспорным лидером, со временем вытесняется на периферию.

5.4.2. Динамика типа «Хищничество»

Динамика поглощения

Хищничество часто рассматривают как разновидность конкуренции, при которой компоненты одной информационной системы непосредственно поглощают компоненты другой. В отличие от случая обычной конкуренции, здесь взаимоотношения информационных систем асимметричны: одна играет роль хищника, другая — жертвы, причем их нельзя поменять местами.

Отношения хищник–жертва сами по себе никогда не приводят к подавлению одной информационной системы другой: хищники не могут воспроизводить свою популяцию без жертв — своей пищи (как в буквальном, так и переносном значении). Сокращение популяции жертв, в свою очередь, вызывает сокращение популяции хищников, так как часть из них остается без средств к существованию. Но сокращение числа хищников приводит к снижению внешнего воздействия на жертвы, и они начинают восстанавливаться. Поэтому для данного случая характерны разнообразные колебательные режимы.

Естественно, влияние третьей силы может существенно изменить типичную картину. Этот вид взаимодействия информационных систем является наиболее сложным и, вместе с тем, наиболее интересным. Главная причина состоит в том, что здесь основную роль играют механизмы второй группы.

Действительно, в случае взаимодействия биологических видов сокращение численности жертв, начиная с некоторого момента, обуславливает сокращение численности хищников, так как тем самым сокращается их ресурсная база (доступные объемы пищи). В информационных системах на уровне их размеров ничего похожего не происходит. Наоборот, компоненты одной информационной системы, «съеденные» другой, механически увеличивают численность хищников, причем без каких-либо ограничений. Информационная система, поглотившая остальные, становится на определенное время информационным монополистом. Поэтому механизмы первой группы сами по себе в данном случае могут приводить лишь к тривиальным эффектам механического расширения одной силы за счет других.

С механизмами второй группы ситуация намного сложнее. Часто вес одной информационной системы основан на эксплуатации ресурсов другой. Например, пропаганда конкурирующих сил может быть построена на противоположностях, которые исключают друг друга, но одновременно и придают друг другу смысл. В таких случаях для поддержания активности необходим противник, с которым имело бы смысл полемизировать. Реальный «информационный вес» определяется превосходством в этой полемике. Еще один пример — использование «образа врага».

Также возможна систематическая эксплуатация одной информационной системой содержательных наработок, созданных другой.

При таких сценариях вполне можно говорить об отношениях хищник–жертва между информационными системами. Проанализируем несколько ситуаций, характерных для отношений хищничества. Хищничеству соответствует приведенная выше система уравнений с отрицательными значениями коэффициентов p_i и q_{ij} для хищника и положительными — для жертвы.

Воздействие информационной операции

Если третья сила (информационная система) оказывает на основные информационные системы слабое воздействие (не участвует в потреблении соответствующих ресурсов), то имеем обычную картину хищничества (рис. 38).

Происходят типичные колебания размеров обеих информационных систем, причем затухающие. В достаточно больших интервалах времени обе зависимости стремятся к состоянию некоторого равновесия, но амплитуда колебаний никогда не становится равной нулю.

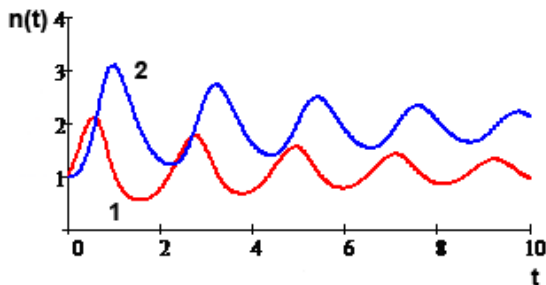


Рис. 38. Слабое воздействие третьей силы на информационные системы 1 и 2

В зависимости от значений параметров кривые могут иметь различные соотношения фаз, что иногда приводит к интересным эффектам, являющихся темой отдельного исследования.

Из соображений наглядности допустим равенство начальных численностей обеих основных информационных систем. Колебательный режим, характерный для отношений—жертва, в поведении динамики информационных систем действительно наблюдался в целом ряде аналитических исследований.

Следующие две зависимости показывают нам, как на динамику хищника и жертвы может влиять третья сила (рис. 39). Приведенные решения отличаются друг от друга значением коэффициента q_{13} . Он равен -1 в первом случае (положительное воздействие на первую из основных информационных систем) и 1 во втором (отрицательное воздействие на нее).

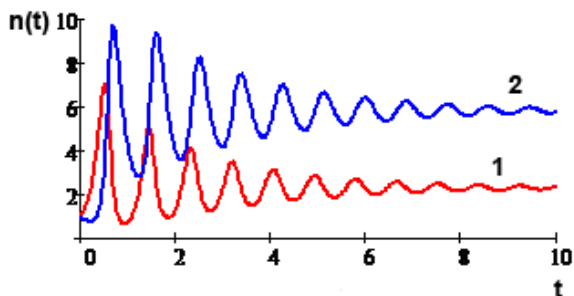


Рис. 39. Влияние третьей силы на информационные системы 1 и 2

При положительном воздействии третьей силы также наблюдаются существенные изменения характера колебательных процессов, в частности, возрастает их частота, но при этом сокращается время затухания (внешнее стабилизирующее влияние). Важно также то обстоятельство, что информационные системы могут многократно меняться местами по своим размерам. С точки зрения постороннего наблюдателя такой процесс может выглядеть странным и тревожным, однако он быстро нормализуется и переходит в (квази)равновесное состояние.

5.4.3. Динамика типа «Симбиоз»

Симбиоз информационных систем возникает в том случае, когда по каким-либо причинам они или не мешают друг другу, или друг друга поддерживают. В таком случае чаще всего отсутствует необходимость проведения дополнительных информационных операций, а вмешательство третьей силы оказывается неэффективным.

В ряде случаев популяции (информационные системы в отношениях симбиоза) оказывают друг на друга положительное влияние,

помогая выжить в жесткой борьбе с другими системами. Однако не следует думать, что симбиоз всегда представляет собой мирное и благодостное сосуществование. В действительности отношения могут быть жесткими и даже антогонистическими. Просто такие информационные системы могут не иметь реальной возможности воздействовать друг на друга посредством информационных механизмов.

Симбиоз, пожалуй, наиболее интересен как в теоретическом, так и в прикладном плане. Действительно, динамика информационных систем, находящихся в отношениях симбиоза, мало чем отличается от динамики невзаимодействующих систем. Однако важен сам факт, что взаимодействующие системы могут находиться в таком состоянии. Тем не менее, для полноты изложения он включен в общую картину.

Симбиоз предполагает, что для всех участвующих в нем информационных систем коэффициенты p_i положительные, а q_{ij} отрицательные.

Отметим также, что информационные системы, между которыми действуют отношения симбиоза, быстро и практически одновременно достигают своих равновесных состояний. Таким образом, этот вид взаимодействия систем можно назвать наиболее статичным.

5.5. Анализ временных рядов

Базой для успешного моделирования информационных процессов и прогнозирования их результатов является учет взаимосвязи событий с информационным пространством, в частности, с его наиболее динамичной и современной частью — информационными ресурсами веб-пространства. Задача изучения свойств информационного пространства и информационных систем — многоплановая, предполагает активное использование методов анализа сложных зависимостей, временных рядов, позволяющих глубже понять специфику той или иной предметной области.

Рассмотрим возможности современных аналитических средств на примере исследования тематических информационных потоков веб-публикаций, собранных из сети Интернет системой InfoStream [95].

В качестве запроса к системе мониторинга веб-ресурсов InfoStream, выражающего тематику исследуемого информационного массива, связанную с развитием кризисных явлений в Украине в 2008 году [96], было выбрано выражение:

(парламентс~криз)|(политич~криз)|(финасов~криз)|(економич~криз)

Исследовались информационные потоки, поступающие из более тысячи украинских сетевых информационных ресурсов, среди которых лидерами по количеству релевантных запросу публикаций были такие авторитетные источники, как Укринформ, УНИАН, РБК-Украина, Радио Свобода, Корреспондент.net, Главред и т.п. Ретроспективный период исследования составлял весь 2008 год, т.е. 366 дней. За этот период системой InfoStream было охвачено свыше 12 млн сетевых документов. В результате поиска по запросу, который учитывал все основные аспекты кризисных явлений, найдено 57245 релевантных документов. На основе обработки этих данных получены достаточно полные картины экспериментальных данных — временные ряды за заданный период. На рис. 40 приведен график изменения количества соответствующих тематических публикаций по дням 2008 года.

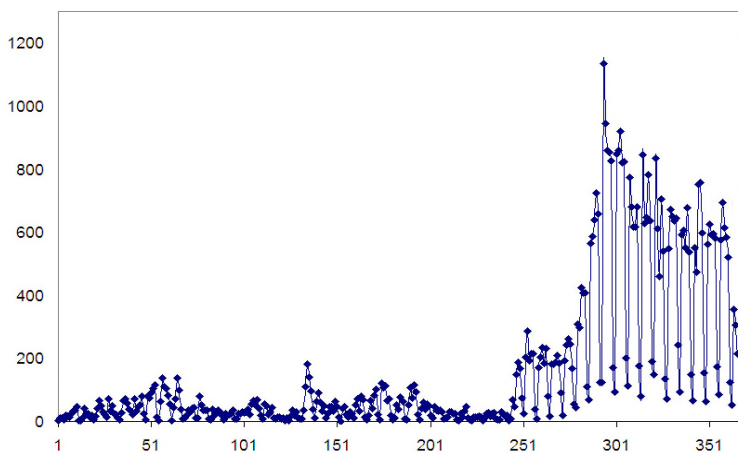


Рис. 40. Динамика количества тематических публикаций по дням 2008 года (всего 57245 публикаций)

Приведенный график позволяет увидеть недельные колебания в объемах публикаций (по выходным дням, например, в сети публикуется значительно меньше документов, чем в будние). На графике также можно видеть, что приблизительно на 250-й день года общее количество

чество сообщений по кризисной проблематике резко увеличилось (усилился парламентский кризис).

5.5.1. Корреляционный анализ

Один из основных методов современного анализа рядов измерений — корреляционный анализ. Остановимся более подробно на формализме корреляционного анализа.

Если обозначить X_t член ряда измерений (например, количество электронных сообщений, поступивших в день t , $t = 1, \dots, N$), то функция автокорреляции для этого ряда X при «окне измерения» в k дней определяется так:

$$F(k) = \frac{1}{N-k} \sum_{t=1}^{N-k} (X_{k+t} - m)(X_t - m),$$

где m — среднее значение ряда X . Коэффициенты автокорреляции для рядов измерений X длиной N при окне измерения, равном k , рассчитываются по формуле

$$R(k) = \frac{F(k)}{\sigma^2},$$

где $F(k)$ — функция автокорреляции; σ^2 — дисперсия.

Важное свойство функции автокорреляции — возможность выявления гармонических составляющих, а также самоподобия исходного процесса. Известно, что эта функция обладает тем свойством, что, если существует скрытая периодическая составляющая в ряде измерений, то ее значение асимптотически стремится к квадрату среднего значения этого ряда. Кроме того, функция автокорреляции периодического ряда также есть периодической, содержит основную частоту и гармоники. Если ряд измерений X является суммой некоторой содержательной составляющей N и синусоидального сигнала S , то функция автокорреляции ряда X включает в себя явным образом выраженную периодическую составляющую [97].

Графическое представление коэффициента автокорреляции для ряда наблюдений, соответствующих динамике рассмотренного выше

тематического информационного потока веб-публикаций, свидетельствует о неизменности корреляционных свойств по дням недели (рис. 41), а тренд — о возможном самоподобии исходного временного ряда.

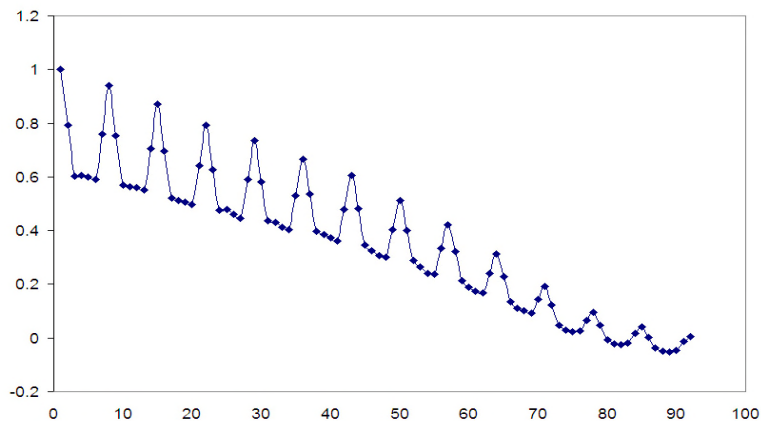


Рис. 41. Коэффициенты автокорреляции ряда наблюдений $R(k)$ (ось ординат) в зависимости от k (ось абсцисс)

5.5.2. Вейвлет-анализ

Основа вейвлет-анализа [98, 99] — вейвлет-преобразование, представляющее собой особый тип линейного преобразования, базисные функции которого (вейвлеты) имеют специфические свойства. Вейвлетом (малой волной) называется некоторая функция, сосредоточенная в небольшой окрестности некоторой точки и резко убывающая к нулю по мере удаления от нее как во временной, так и в частотной области. Существуют разнообразные вейвлеты, имеющие разные свойства. Вместе с тем все вейвлеты имеют вид коротких волновых пакетов с нулевым интегральным значением, локализованных на временной оси, являющихся инвариантными к сдвигу и масштабированию.

К любому вейвлету можно применить две операции:

- сдвиг, т.е. перемещение области его локализации во времени;
- масштабирование (растяжение или сжатие).

Главная идея вейвлет-преобразования заключается в том, что нестационарный временной ряд разделяется на отдельные промежутки (так называемые окна наблюдения), и на каждом из них выполня-

ется вычисление скалярного произведения (величины, которая характеризует степень близости двух закономерностей) исследуемых данных с разными сдвигами некоторого вейвлета на разных масштабах. Вейвлет-преобразование генерирует набор коэффициентов, с помощью которых представляется исходный ряд. Они являются функциями двух переменных: времени и частоты, поэтому образуют поверхность в трехмерном пространстве. Эти коэффициенты, показывая насколько поведение процесса в данной точке аналогично вейвлету на данном масштабе. Чем ближе вид анализируемой зависимости в окрестности данной точки к виду вейвлета, тем большее абсолютное значение имеет соответствующий коэффициент. Отрицательные коэффициенты показывают, что зависимость похожа на «зеркальное отражение» вейвлета. Использование этих операций, с учетом свойства локальности вейвлета в частотно-временной области, позволяет анализировать данные на разных масштабах и точно определять места их характерных особенностей во времени.

Технология использования вейвлетов позволяет обнаруживать единичные и нерегулярные «всплески», резкие изменения значений количественных показателей в разные периоды времени, в частности объемов тематических публикаций в веб-пространстве. При этом могут обнаружиться моменты возникновения циклов, а также моменты, когда за периодами регулярной динамики следуют хаотические колебания.

Рассматриваемый временной ряд может аппроксимироваться кривой, которая, в свою очередь, может быть представленная в виде суммы гармонических колебаний разной частоты и амплитуды. При этом колебания, которые имеют низкую частоту, отвечают за медленные, плавные, крупномасштабные изменения значений исходного ряда, а высокочастотные — за короткие, мелкомасштабные изменения. Чем сильнее изменяется описываемая данной закономерностью величина при данном масштабе, тем большую амплитуду имеет составляющая соответствующей частоты. Таким образом, исследуемый временной ряд можно рассматривать в частотно-временной области, т.е. в области исследований закономерности, описывающей процесс в зависимости как от времени, так и от частоты.

Непрерывное вейвлет-преобразование для функции $f(t)$ строится с помощью непрерывных масштабных преобразований и переносов выбранного вейвлета $\psi(t)$ с произвольными значениями масштабного коэффициента a и параметра сдвига b :

$$W(a, b) = (f(t), \psi(t)) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt.$$

Полученные коэффициенты представляются в графическом виде как карта коэффициентов преобразования, или скейлограмма. На скейлограмме по одной оси откладываются сдвиг вейвлета (ось времени), а по другой — масштабы (ось масштабов), после чего точки полученной схемы раскрашиваются в зависимости от значений соответствующих коэффициентов (чем больше коэффициент, тем ярче цвета изображения). На скейлограмме видны все характерные особенности исходного ряда: масштаб и интенсивность периодических изменений, направление и значение трендов, наличие, расположение и продолжительность локальных особенностей.

Например, известно, что комбинация нескольких разных колебаний может иметь настолько сложную форму, которая не позволяет аналитику выявить их. Периодические изменения, которые происходят для значений коэффициентов вейвлет-преобразования на некотором непрерывном множестве частот, выглядят как цепочка «холмов», имеющих вершины, расположенные в точках (по оси времени), где эти изменения достигают наибольших значений.

Другим важным показателем является выраженная тенденция динамики временного ряда (тренд) вне зависимости от периодических колебаний. Наличие тренда может быть неочевидным при простом рассмотрении временного ряда, например, если тренд объединяется с периодическими колебаниями. Тренд отражается на скейлограмме как плавное изменение яркости вдоль оси времени одновременно на всех масштабах. Если тренд возрастающий, то яркость будет увеличиваться, если убывающий — уменьшаться.

Еще одним важным фактором, который необходимо учитывать при анализе временных рядов, являются локальные особенности, т.е. возможные резкие, скачкообразные изменения характеристик исходного ряда. Локальные особенности представленные на скейлограмме вейвлет-преобразования как линии резкого перепада яркости, которые исходят из точки, соответствуют времени возникновения скачка. Локальные особенности могут иметь как случайный, так и систематический характер, при этом «маскируя» периодические зависимости или краткосрочный тренд. Анализ локальных особенностей позволяет вос-

становить информацию о динамике исходного процесса и в некоторых случаях прогнозировать подобные ситуации.

Таким образом, каждый из основных факторов динамики имеет свое, характерное отражение на скейлограмме, причем вся аналитическая информация представляется в наглядном и удобном для изучения виде. Благодаря наглядности представления результатов в виде скейлограммы, иногда достаточно одного взгляда, чтобы увидеть на ней наиболее существенные факторы [100, 101].

На рис. 42 представлена скейлограмма — результат непрерывного вейвлет-анализа (вейвлет Гаусса) временного ряда, соответствующего рассматриваемому выше процессу.

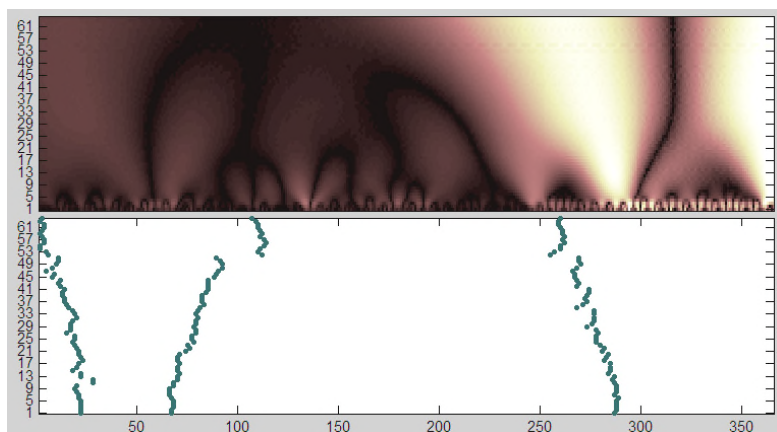


Рис. 42. Результат вейвлет-анализа (непрерывное вейвлет-преобразование):
сверху — вейвлет-скейлограмма;
снизу — линии локальных максимумов (скелетон)

Приведенный пример показывает, что вейвлет-анализ позволяет обнаруживать не только очевидные аномалии в исследуемом ряде, но и критические значения, которые скрыты за относительно небольшими абсолютными значениями элементов ряда. Например, на скелетоне наибольшие значения отмечены не только в 250-й день, но показаны и неявные экстремумы (25-й и 75-й дни).

Безусловно, финансово-экономические факторы имеют непосредственное влияние на общественные процессы. На рис. 43 представлена динамика изменения курса продажи наличного доллара США

в банках Украины в течение 2008 года. На рис. 44 приведен пример применения вейвлет-анализа к этому ряду. Сдвиг последней линии локальных трендов на этой скейлограмме в сравнении с предыдущий (см. рис. 42) свидетельствует о том, что наличный курс является производной от общекризисных явлений.

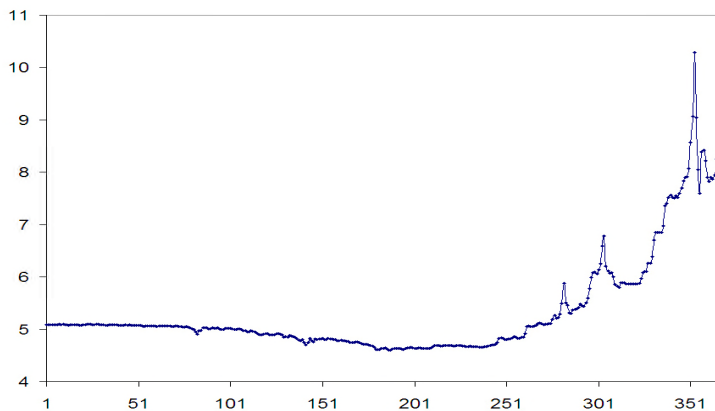


Рис. 43. Динамика изменения курса наличного доллара США в гривнах в течение 2008 г.

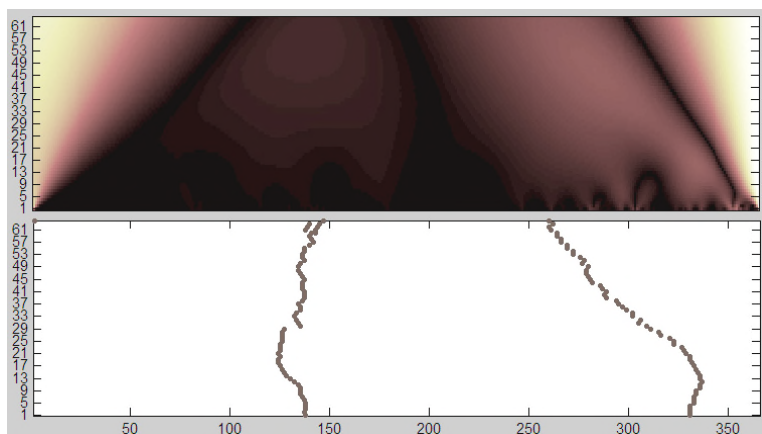


Рис. 44. Результат вейвлет-анализа ряда, соответствующего курсу наличного доллара США: сверху — вейвлет-скейлограмма; снизу — линии локальных максимумов (скелетон)

5.5.3. Фрактальный анализ: R/S -анализ

Теория фракталов [102] широко применяется как подход к исследованию рядов наблюдений, позволяющий получать важные характеристики соответствующих процессов, не вдаваясь в подробный анализ их внутренней структуры.

Анализ самоподобия временных рядов может рассматриваться как технология, предназначенная для осуществления аналитических исследований с элементами прогнозирования, пригодная к экстраполяции полученных зависимостей.

Важнейшей характеристикой рядов, имеющих хаотичное поведение, является фрактальная размерность, которая во многих случаях может вычисляться с помощью так называемого R/S -анализа. Точнее говоря, вычисляется не сама фрактальная размерность, а показатель Херста, который связан с ней простым соотношением. R/S -анализ базируется на анализе нормированного разброса — отношения разброса значений исследуемого ряда R к среднеквадратичному отклонению S [103].

Для изучения фрактальных характеристик временных рядов $F(n)$, $n=1, \dots, N$, составленных из количества принадлежащих им сообщений, изучалось значение R/S , где R — так называемый размах:

$$R(N) = \max_{1 \leq n \leq N} X(n, N) - \min_{1 \leq n \leq N} X(n, N),$$

а S — стандартное отклонение:

$$S = \sqrt{\frac{1}{N} \sum_{n=1}^N (F(n) - \langle F \rangle_N)^2};$$

$$\langle F \rangle_N = \frac{1}{N} \sum_{n=1}^N F(n);$$

$$X(n, N) = \sum_{i=1}^n (F(i) - \langle F \rangle_N).$$

Для достаточно широкого класса рядов зависимость R/S имеет показательный тренд, т.е. можно говорить о соотношении

$$R/S = \left(\frac{N}{2}\right)^H,$$

где H — показатель Херста, который при определенных дополнительных условиях связан с хаусдорфовой (фрактальной) размерностью D полой формулой: $D + H = 2$.

Главное условие, при котором показатель Херста связан с фрактальной размерностью в соответствии с приведенной формулой, определено Е. Федером: «... рассматривают клетки, размеры которых малы по сравнению как с длительностью процесса, так и с диапазоном изменения функции; поэтому соотношение справедливо, когда структура кривой, описывающая фрактальную функцию, исследуется с высоким разрешением, т.е. в локальном пределе». Еще одним важным условием является самоаффинность функции. Не вдаваясь в подробности, отметим, что, например, для информационных потоков это свойство интерпретируется как самоподобие, возникающее в результате процессов их формирования. При этом указанными свойствами обладают не все информационные потоки, а лишь те, которые характеризуются достаточной мощностью и итеративностью при формировании.

На рис. 45 представлены значения R/S для ряда количества публикаций по дням 2008 года, которая соответствует приведенному выше запросу. Очевидно, характер нормированного размаха резко изменяется в окрестности 250 дня года приблизительно тогда, когда прозвучали первые серьезные заявления на высшем уровне о финансово-экономическом кризисе, т.е. имеем фактически два разных ряда — с 1 по 250 и с 251 по 366. Как можно видеть, кривая нормированного размаха для второго ряда (рис. 46) удовлетворительно аппроксимируется прямой в двойном логарифмическом масштабе. Наклон этой прямой соответствует показателю Херста.

Численные значения H характеризуют разные типы корреляционной динамики (персистентности). При $H = 0,5$ наблюдается некоррелированное поведение значений ряда, а значения $0,5 < H < 1$ соответствуют уровню автокорреляции ряда. Как видим, показатель Херста для исследуемого информационного потока на рис. 46 соответствует значению $\sim 0,89$, что подтверждает предположение о самопо-

добии и итеративности процессов в информационном пространстве. Републикации, цитирование, гипертекстовая и контекстная ссылки и т.п. порождают самоподобие, наличие высокого уровня статистической корреляции в информационных потоках на продолжительных временных интервалах.

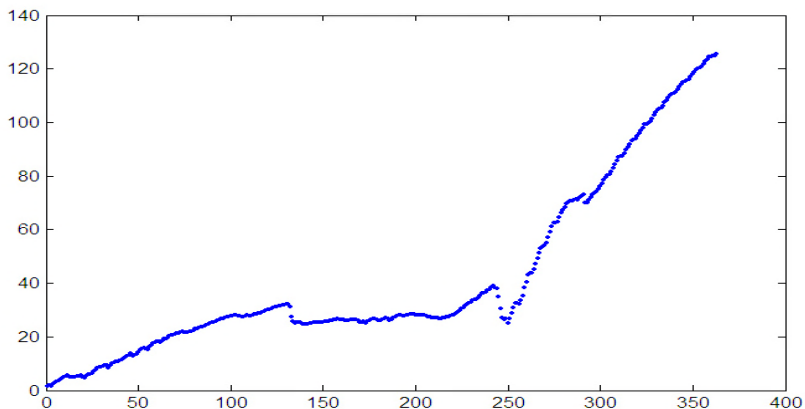


Рис. 45. Показатель нормированного разброса (ось ординат) для всего периода наблюдений исследуемого ряда (ось абсцисс)

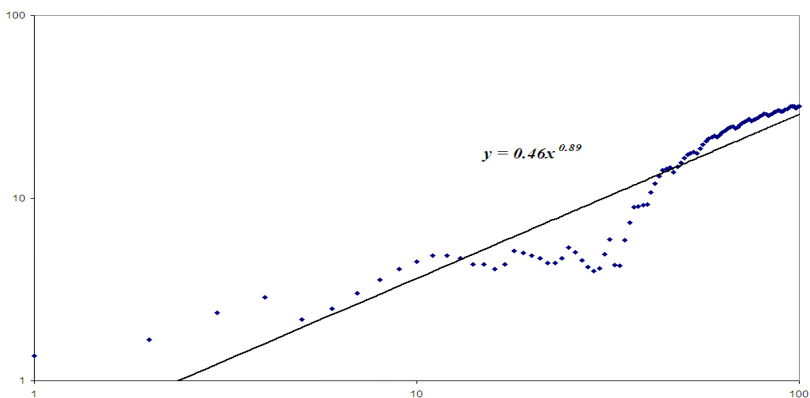


Рис. 46. Показатель нормированного разброса в логарифмической шкале за последние 120 дней года

5.5.4. Отклонение от линейного тренда

Метод DFA (Detrended fluctuation analysis) также чаще всего употребляется для выявления статистического самоподобия сигналов [104].

Этот метод является вариантом дисперсионного анализа одномерных случайных блужданий и позволяет исследовать эффекты продолжительных корреляций в рядах, которые исследуются. В рамках алгоритма DFA анализируется среднеквадратичная ошибка линейной аппроксимации в зависимости от размера участка аппроксимации (окна наблюдения). Пусть есть ряд измерений x_t , $t \in 1, \dots, N$. Обозначим

среднее значение этого ряда измерений: $\langle x \rangle = \frac{1}{N} \sum_{k=1}^N x_k$. Из исходного

ряда строится ряд накопления:

$$X_t = \sum_{k=1}^t (x_k - \langle x \rangle).$$

Потом ряд X_t разделяется на временные окна длиной L , строится линейная аппроксимация ($L_{j,L}$) по значениям $X_{k,j,L}$ с $X_{j,L}$ внутри каждого окна (в свою очередь, $X_{j,L}$ — подмножество X_t , $j = 1, \dots, J$, $J = N/L$ — количество окон наблюдения) и рассчитывается отклонение точек ряда накопления от линейной аппроксимации:

$$E(j, L) = \sqrt{\frac{1}{L} \sum_{k=1}^L (X_{k,j,L} - L_{k,j,L})^2} = \sqrt{\frac{1}{L} \sum_{k=1}^L |\Delta_{k,j,L}|^2},$$

где $L_{k,j,L}$ — значение локальной линейной аппроксимации в точке $t = (j-1)L + k$.

Здесь $|\Delta_{k,j,L}|$ — абсолютное отклонение элемента $X_{k,j,L}$ от локальной линейной аппроксимации.

Далее вычисляется среднее значение

$$F(L) = \frac{1}{J} \sum_{j=1}^J E(j, L),$$

после чего, в случае $F(L) \propto L^\alpha$, где α — некоторая константа, делаются выводы о наличии статистического самоподобия и характер поведения исследуемого ряда измерений.

Этот метод был применен к ряду значений количества публикаций, полученных за представленным выше запросом. На рис. 47 представлена зависимость среднеквадратичной ошибки аппроксимации от длины участков аппроксимации в двойном логарифмическом масштабе.

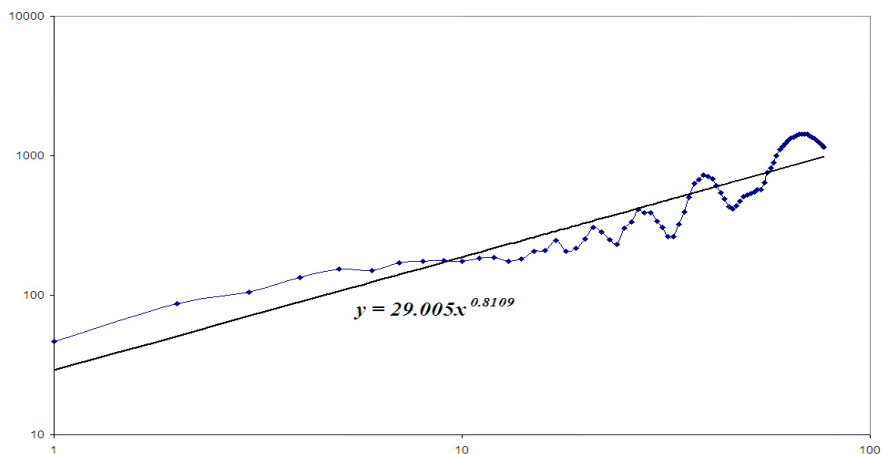


Рис. 47. Зависимость среднеквадратичной ошибки линейной аппроксимации D от длины окна наблюдения k

Близость зависимости $D(k)$ к линейной еще раз подтверждает наличие локального скейлинга во втором полугодии 2008 года.

5.5.5. Визуализация на основе ΔL -анализа

В целях визуализации и анализа временных рядов, связанных с публикациями в информационном пространстве сети Интернет, разработан новый метод дисперсионного анализа, предназначенный для

анализа и визуализации состояния временных рядов интенсивности публикаций по определенной тематике [105].

Задачи выявления и визуализации трендов, выявление гармонических составляющих, трендов, локальных особенностей временных рядов, фильтрации шума сегодня решаются методами фрактального, вейвлет- и фурье-анализа.

Как и в методе DFA, рассмотрим поведения отклонения точек ряда накопления от линейной аппроксимации (но в этом случае абсолютное значение) $|\Delta_{k,j,L}|$. Построение соответствующих диаграмм значений $|\Delta_{k,j,L}|$, которые зависят фактически от двух параметров — L и $t = (j-1)L + k$, названо ΔL -методом визуализации. Такая визуализация в виде «рельефной» диаграммы представляет определенный интерес для изучения особенностей процессов, которые соответствуют исходным рядам измерений.

ΔL -метод оказывается довольно эффективным для выявления гармонических составляющих исследуемого ряда. На рис. 48 показана ΔL -диаграмма ряда, который соответствует синусоиде ($y(i) = \sin(i\pi / 7), i = 1, \dots, 366$). Применение ΔL -метода к ряду, составленному из количества публикаций, собранных системой InfoStream из Интернета без учета тематического деления, имеет явным образом выраженную гармоническую составляющую (общее количество публикаций зависит от дня недели), что можно видеть на рис. 49. Кроме того, на этой диаграмме заметны отклонения от общей динамики объемов публикаций в праздничные дни.

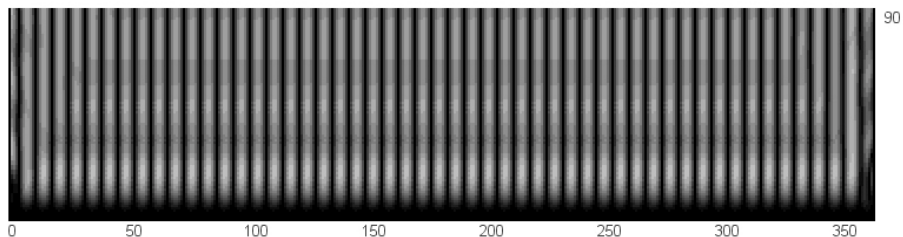


Рис. 48. ΔL -диаграмма синусоиды

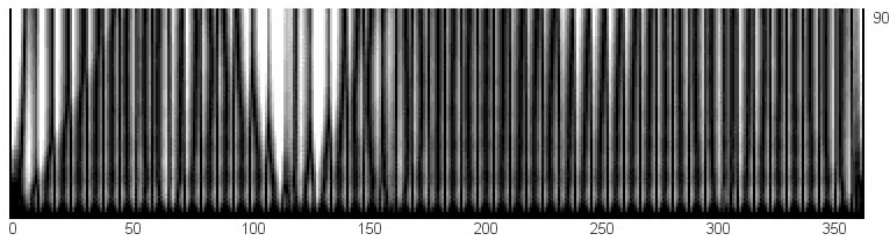


Рис. 49. ΔL -диаграмма ряда количества публикаций, собираемых ежесуточно системой InfoStream в 2008 году

«Рельефные диаграммы», получаемые в результате ΔL -метода (более светлые тона соответствуют большим значениям $|\Delta_{k,j,L}|$), напоминают скейлограммы, полученные вследствие непрерывных вейвлет-преобразований. Следует обратить внимание на то, что темные полосы в центре многих областей светлого закрашивания свидетельствуют о «стабилизации» больших значений рассмотренного ряда на высоком уровне.

ΔL -метод применяется для реальных временных рядов, например тех, которые отражают интенсивность публикаций данной тематики в Интернете. На рис. 50 приведена ΔL -диаграмма для рассмотренного выше временного ряда из количества публикаций сообщений через сутки по выбранной тематике в сети Интернет на протяжении года.

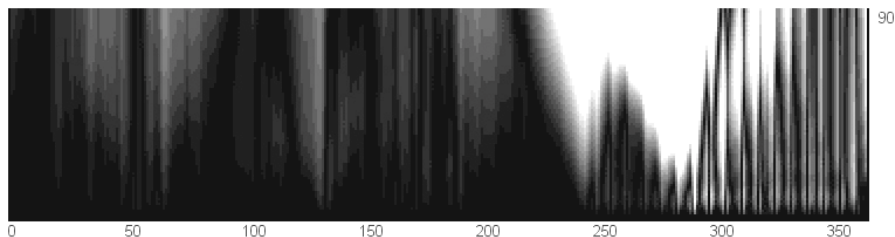


Рис. 50. ΔL -диаграмма временного ряда интенсивности тематических публикаций (ось абсцисс — дни года, ось ординат — величина окна измерений)

На рис. 51 приведена ΔL -диаграмма наличного курса доллара в гривнах на протяжении 2008 года. Еще нагляднее, чем в случае применения вейвлет-анализа, можно убедиться в том, что наиболее значительные отклонения на диаграмме в этом случае наступают с некоторой временной задержкой по сравнению с диаграммой публикациями по кризисной тематике.

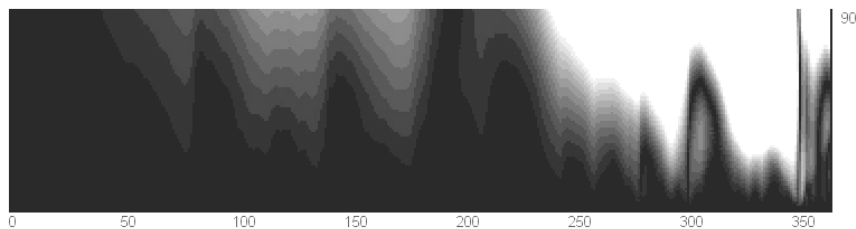


Рис. 51. ΔL -диаграмма временного ряда значений наличных курсов доллара США в гривнах (ось абсцисс — дни года, ось ординат — величина окна измерений)

Предложенный метод визуализации абсолютных отклонений ΔL -диаграмм, как и метод вейвлет-преобразований, позволяет (и как показано на примере — не хуже) обнаруживать единичные и нерегулярные «всплески», резкие изменения значений количественных показателей в разные периоды времени.

Следует отметить, что метод вейвлет-преобразований может применяться с использованием разнообразных вейвлетов. В случае применения ΔL -метода не нужно решать сложную задачу выбора и обоснования применения соответствующего вейвлета; в отличие от методов фрактального анализа предложенный подход не требует значительных объемов точек ряда измерений. Этот метод достаточно простой в программной реализации и базируется на такой мощной теоретической основе, как DFA, и является достаточно эффективным при анализе временных рядов в таких областях, как экономика и социология.

5.5.6. Мультифрактальный анализ

Наиболее общее описание природы самоподобных объектов дает теория мультифракталов, которая позволяет охватывать беско-

нечную иерархию размерностей, и позволяет отличать однородные объекты от неоднородных. Концепция мультифрактального формализма [106–108] представляет собой эффективный инструмент для изучения и количественного описания широкого многообразия сложных систем.

В соответствии с этим формализмом носителем мультифрактальной меры является множество L — объединение фрактальных подмножеств L_α , т.е. мультифрактал можно понимать как какое-то объединение разных однородных фрактальных подмножеств L_α исходного множества L , каждое из которых имеет свое собственное значение фрактальной размерности.

Для характеристики мультифрактального множества используют так называемую функцию мультифрактального спектра $f(\alpha)$ (спектр сингулярностей мультифрактала), к которой полностью подходит термин «фрактальная размерность». Величина $f(\alpha)$ равняется хаусдорфовой размерности однородного фрактального подмножества L_α исходного множества L , которое дает доминирующий вклад в некоторую статистическую сумму.

Приведенные на рис. 52 зависимости относятся к анализу числового ряда интенсивностей сообщений, которые отражают проблематику использования антивирусного программного обеспечения (динамика публикаций в Интернет-новостях сообщений по заданной тематике), а также ряда, полученного по уточненной тематике (первоначальный запрос был расширен словом «тороянский»).

Путем соответствующих расчетов было показано, что ряды, соответствующие динамике появления публикаций, в рассмотренных случаях имеют мультифрактальную природу. Вместе с тем соответствующие зависимости $f(\alpha)$, которые соответствуют исследуемым рядам (рис. 53), имеют разные параметры кривизны. Этот факт, характерный для соотношений тематики и подтематики, свидетельствует, с одной стороны, о том, что ряд, который соответствует подтематике менее стабилен, чем ряд, соответствующий всей тематике, а с другой — о том, что рассмотренная подтематика не является репрезентативной для анализа потока публикаций по общей тематике.

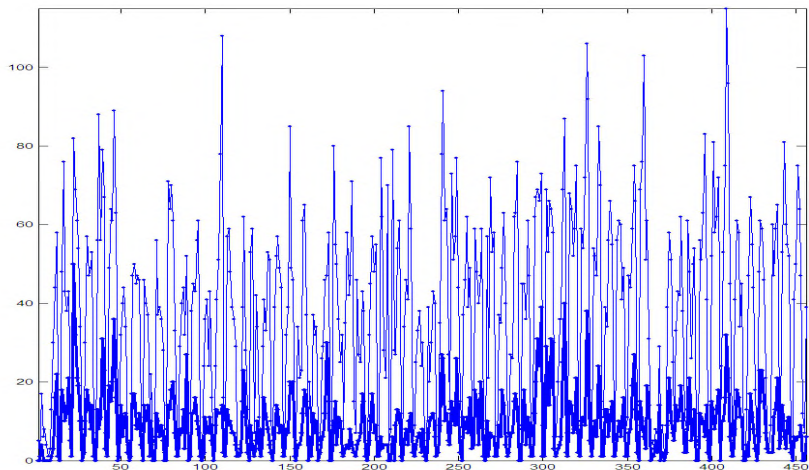


Рис. 52. Диаграммы интенсивности публикаций по основной (тонкая соединяющая линия) и уточненной тематике (жирная линия): ось абсцисс — порядковые номера дней, ось ординат — количество публикаций

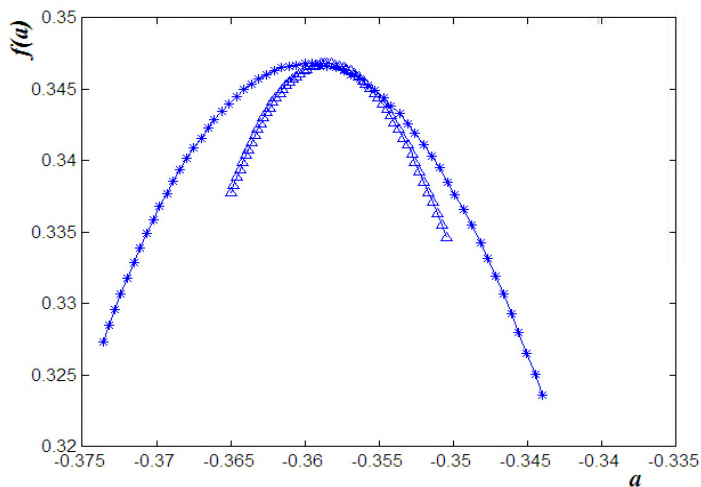


Рис. 53. Сравнение мультифрактальных спектров исследуемых рядов по основной тематике (Δ) и уточненного ($*$) от индекса сингулярности α

6. ЖИВУЧЕСТЬ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ

6.1. Понятие «информационные операции»

В последние годы благодаря многочисленным документам и публикациям Министерства обороны США стал популярен термин «информационные операции», прежде всего потому, что информационные технологии играют постоянно увеличивающуюся роль в военных операциях. При этом информационные операции определяются как «акции, направленные на воздействие на информацию и информационные системы противника и защиту собственной информации и информационных систем» [109]. Информационные операции рассматриваются как объединение основных возможностей радиоэлектронной войны, компьютерных сетевых операций, психологических операций, военных действий и операций по обеспечению безопасности с целью воздействовать, разрушать, исказить информацию, необходимую для принятия противником решений, а также защищать собственную информацию.

Информационные операции охватывают целый комплекс процессов, проводимых в самых разных областях. При этом необходимо отметить, что информационные операции — существенная и традиционная составляющая боевых операций. Несмотря на то, что формальное определение в документах Департамента обороны США ориентировано на военные аспекты информационных операций, оно вполне применимо практически для любой области жизни.

Ниже будут рассматриваться такие информационные операции, которые реализуются с помощью информационных систем. Живучесть этих информационных систем во многом определяет живучесть информационных операций, которые реализуются в виде информационных воздействий на сознание людей.

Информация является отражением вложенного в нее смысла, поэтому сегодня информация превратилась из абстрактного термина в объект, цель и средство информационных операций, стала критическим понятием в проблематике безопасности. Бывший министр обороны США Уильям Коэн 18 марта 1999 г. заявил, что «способность армии использовать информацию, чтобы доминировать в будущих сражениях, даст США новый ключ к победам в течение многих лет, если не в течение нескольких поколений» [110].

При моделировании и проведении информационных операций необходимо учитывать значение ценности информации для лиц, принимающих решение. Ценность информации включает ее своевременность, точность и «аналитичность». С практической точки зрения ценность информации также может быть определена как ее значимость или применимость, пригодность к использованию. Под применимостью информации понимается обеспечение доступа лиц, принимающих решения, к готовой к использованию информации. Стандарт ISO 9241 (ISO — Международная Организация по Стандартизации) определяет применимость в терминах эффективности и удовлетворения потребностей указанного набора пользователей для решения указанного набора задач в специфическом окружении. На практике большая часть полезной информации поступает лицам, принимающим решения от информационно-аналитических систем, обеспечивающих ориентацию в ситуации и поддержку при принятии решений. Согласно полевому уставу военного ведомства США «Информационные операции» (FM 100-6), «ориентация в ситуации означает комбинацию ясного представления о диспозиции своих и вражеских сил с оценкой ситуации и намерений со стороны командования».

Информационные операции осуществляются в некоторой социальной среде, соответственно, для успешного их проведения необходимо адаптироваться к этой среде, преодолеть определенный барьер не очень сильного внимания к информационному воздействию. Этот барьер возникает благодаря так называемой иммунной системе среды, которая может не пропустить информационные воздействия, если она достаточно мощная и/или уже научилась защищаться от подобных воздействий. К подготовительным действиям для проведения информационных операций может относиться создание «иммунодефицита» социальной среды путем воздействия через информационное пространство, например, с помощью материалов в СМИ. Очень часто информационные воздействия используют механизмы «вирусного маркетинга», например, в виде слухов, когда сенсационно поданная дезинформация распространяется с огромной скоростью. Именно иммунная система оказывает противодействие подобным информационным операциям. Очень часто с иммунной системой общества отождествляют государство, призванное обеспечивать безопасность этого общества, т.е. при наличии сильного государственного аппарата вероятность успеха антиобщественных информационных операций существ-

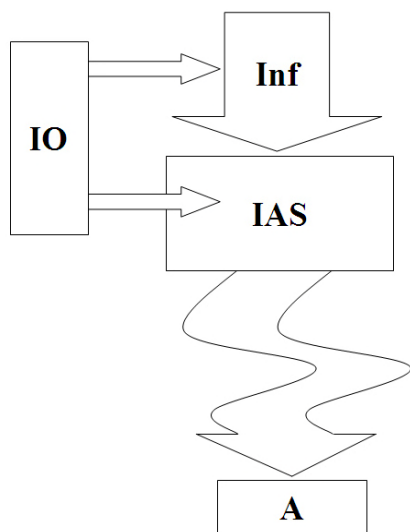
венно снижается. Читатель прекрасно знает, как происходило противодействие подобным информационным процессам в тоталитарных государствах. В демократическом обществе, естественно, тоталитарные методы не применимы. В этом случае иммунитет достигается за счет «обучения», т.е. демократическое общество должно пройти через многие информационные атаки, воздействия, влияния стереотипов, чтобы выработать необходимый иммунитет.

Уровень готовности к проведению информационных операций сегодня считается ключевым фактором успеха проведения любой социальной процедуры, кампании.

Особой целью при проведении информационных операций являются информационно-аналитические системы субъекта воздействия. Оказывая влияния на такие системы, можно добиться того, что принимающие решение лица из лагеря противника примут неадекватные выводы, и требуемый социальный процесс изменит траекторию в необходимом оказывающей влияние стороне направлении (рис. 54).

В данном случае к непосредственным информационным воздействиям может быть отнесено размещение в информационном пространстве документов, компрометирующих противоположную сторону,

реклама (в том числе скрытая) своих преимуществ, искаженные данные о внешней среде, искаженная информация о намерениях и т.д.



*Рис. 54. Воздействие на информационно-аналитическую систему противника:
 Inf — информационное пространство;
 IAS — информационно-аналитическая система;
 А — абонент системы (лицо, принимающее решение);
 IO — информационные воздействия*

При этом информационным операциям присущи такие основные особенности:

— информационные операции — это междисциплинарный набор методов и технологий в таких областях, как информатика, социология, психология, международные отношения, коммуникации, военная наука;

— до сих пор не существует стандартов проведения информационных операций;

— в развитии технологий информационных операций заинтересованы не только оборонные ведомства, но и многие правительственные и коммерческие организации;

— задача формирования научного подхода к информационным операциям является насущной и актуальной.

При проведении информационных операций существенно выявление содержания (знаний), вкладываемого в информацию, с учетом самых разнообразных аспектов — социальных, политических, религиозных, исторических, экономических, психологических, ментальных, культурных, присущих различным слоям общества. Поэтому в настоящее время имеет смысл рассматривать информационные операции шире, как операции, базирующиеся на знаниях (*англ.* — *Knowledge Operations*) [111].

6.2. Информационные операции как социальные процедуры

Социальные процедуры и процессы, как правило, сложно оценивать и моделировать, так как их результаты относятся к психологическим и социологическим, а не физическим. Именно этот факт также определяет проблематичность прогнозирования результатов моделирования информационных операций. Кроме того, экспериментирование с информационными воздействиями в рамках информационных операций более сложны и опасны, чем при моделировании физических процессов. Действия для достижения эффективности влияния на процессы принятия решения противником иногда необходимо предпринимать в течение длительного времени, прежде чем они вступят в силу.

Одна из основных компонент информационных операций — социальное влияние, охватывающее все многообразие процессов влияния. Существенные изменения в убеждениях или отношении людей к

некоторой проблеме или явлению, как ожидается, будут вести к изменению в поведении, связанном с этой проблемой.

В 1948 году Лассвел (Lasswell) [112] разработал модель трансмиссии коммуникаций, состоящую из пяти компонент:

— источник — персона, которая влияет или убеждает другие персоны;

— сообщение — с помощью чего источник пробует убедить цель;

— цель — человек, на которого источник пробует влиять;

— канал — метод доставки сообщений;

— воздействие — реакция цели на сообщение.

Хотя Лассвел прежде всего интересовался массовой коммуникацией, его модель передачи информации может применяться в межличностной коммуникации типа циркулярных моделей Шеннона–Вивера (Shannon–Weaver) и Осгуда–Шрамма (Osgood–Schramm), которые включают петли обратной связи в процессе коммуникаций, утверждая, что коммуникация является циркулярным, а не линейным процессом [113, 114].

Моделирование объективных факторов социального влияния требует междисциплинарных подходов, имеющих отношение к информатике, маркетингу, политологии, социальной психологии. Самые известные модели формирования общественного мнения и социального влияния базируются на теории Латэйна (Latane) динамического социального воздействия [115, 116], развитой такими авторами, как Nowak, Szamrej, Latane [117], Lewenstein, Nowak и Latane [118], Kasperski и Holyst [119], Sobkowicz [120, 121].

Пытаясь обосновать механизм социального влияния сообщений Латэйна [115] подчеркнул важность трех признаков отношений источника и цели:

— сила — социальная сила, вероятность или уровень влияния на индивидуумов;

— непосредственность — физическое или психологическое расстояние между индивидуумами;

— число источников — количество источников, стремящихся к цели.

Современное состояние моделирования информационных операций характеризуется рядом открытых проблем, основные из которых относятся к пониманию понятий информационного влияния и воздействия.

Универсальными характеристиками объектов являются его состояние и возможность воздействия на другие объекты. Реализация возможности воздействия требует определенных условий, которые принято называть его влиянием. При этом объект, который может осуществлять свою волю, называют субъектом, а управлением принято называть воздействие по отношению к объекту воздействия, применяемое с определенной целью.

Когда индивидуум является целью влияния одного или более источников, динамическая социальная теория воздействия утверждает, что уровень социального влияния на индивидуума может быть представлен следующим уравнением, являющимся основой так называемой индивидуум-ориентированной модели [122]:

$$I_i = -S_i\beta - \sum_{j=1, j \neq i}^N \frac{S_j O_j O_i}{d_{i,j}^\alpha},$$

где I_i — величина (количество) социального давления, оказываемого на индивида i , ($-\infty < I_i < \infty$). O_i представляет мнение индивидуума (± 1) по актуальному вопросу, где $+1$ и -1 представляют поддержку или возражение относительно данного вопроса, соответственно. S_i представляет силу индивида i или влияние ($S_i > 0$), β — сопротивление индивидуума к изменениям ($\beta > 0$), $d_{i,j}^\alpha$ — расстояние между индивидуумами i и j ($d_{i,j}^\alpha \geq 1$), α — показатель сокращения расстояния ($\alpha \geq 2$), N — общее количество агентов (индивидуумов, составляющих сообщество). Значение β , тенденция сохранять собственное мнение или сопротивляться изменению определяет то, что индивидуумы в рамках модели могут требовать больших или меньших объемов социального давления для изменения их мнения. Большие уровни значения α соответствуют эффекту возрастания расстояния между источником и целью, что влияет на объем социального давления на цель.

На основе введенных терминов формулируется понятие «информационного поля объекта» [123], описываются его характеристики. Это дает возможность определить информационное воздействие

как воздействие на информационное поле объекта. Исследуя информационные поля объектов и субъектов социальных систем, можно определить информационные влияния и управления. При этом информация может рассматриваться и как объект, и как средство воздействия. Использование информации как средства воздействия требует в процессе управления осуществить подготовку данных, производство соответствующей информации, а лишь затем реализовывать созданную информацию в виде воздействия (влияния).

6.3. Информационное влияние

Одним из основных методов ведения информационных операций является информационное влияние, оказываемое с целью информационного управления. Под информационным управлением в данном случае понимается механизм управления, когда управляющее воздействие носит неявный, косвенный информационный характер и объекту управления дается определенная информационная картина, под влиянием которой он формирует линию своего поведения. Таким образом, информационное управление — это способ воздействия, побуждающий людей к упорядоченному поведению, выполнению требуемых действий.

В соответствии с [123, 124] процесс информационного влияния одного объекта на другие целесообразно декомпозировать на следующие этапы:

- генерация источником влияния данных, информационных элементов и информационных совокупностей;
- передача информации источником влияния;
- прием информации реципиентом;
- генерация совокупности данных, информационных элементов и новых совокупностей объекта влияния;
- соответствующие активные действия объекта влияния.

Информационные воздействия на элементы систем можно классифицировать по таким признакам, как источники возникновения, длительность воздействия, природа возникновения и т.п.

Для выбора конкретных способов реализации информационного управления необходимо конкретизировать задачи, решаемые с помощью информационного воздействия, провести анализ процесса формирования информационных операций и выработать критерии их

оценки. Информационное управление рассматривают как процесс, охватывающий такие три взаимосвязанные направления:

- управление обменом данными между реальным миром и виртуальным миром субъекта влияния;
- управление виртуальным миром субъектов влияния, механизмами принятия решений;
- управление процессом преобразования решений в действия субъектом влияния в реальном мире.

Информационное воздействие может быть двух основных видов:

1. Изменение в требуемую сторону данных, которые использует информационно-аналитическая система объекта воздействия при принятии решений.

2. Непосредственное влияние на процесс принятия решения объекта воздействия, например, на процедуры принятия решения или отдельные лица, принимающие решения.

Важнейшее значение для проведения информационных операций имеет окружающая среда, состояние объектов информационного воздействия, их взаимное влияние. В частности, если в качестве объектов информационных операций выбирается некоторое электоральное поле, то важно учитывать все электоральные популяции, входящее в это поле, которые представляют сторонников (или противников) тех или иных политических сил. Несмотря на то, что в дальнейшем будут рассматриваться и некоторые модели, в которых в явном виде постулируется однородность среды, в общем случае по отношению к информационным операциям окружающая среда может состоять из областей:

- доминирующего восприятия;
- повышенной чувствительности;
- индифферентности к соответствующим информационным воздействиям.

6.4. Этапность информационных операций

Остановимся отдельно на этапности информационных операций. Очевидно, не существует единственного «стандартного» плана проведения как наступательных, так и оборонительных информационных операций. Можно лишь рассмотреть примерную, полученную путем обобщения некоторых уже реализованных информационных операций последовательность действий при их осуществлении.

На практике информационная операция как процесс информационного воздействия на массовое сознание как правило реализуется следующим образом: в результате предварительной разведки вырабатывается план следующего этапа — оперативного управления и намечаются соответствующие мероприятия оперативной разведки, которые являются приближенной моделью решения, после чего реализуется оперативное управление противником. На этапе оперативной разведки определяется уровень отклонения первоначальной модели от реальности, и если оно незначительно, то реализуется первоначальный план. В противном случае строится новый план оперативного управления и управления противником. Далее цикл повторяется до тех пор, пока оперативная разведка не подтвердит модель. При этом окончательное решение принимается с определенным оперативным риском.

Таким образом, процесс информационного воздействия охватывает такие основные этапы [125] (рис. 55), как предварительная разведка (PI) (выявление текущей обстановки, состояние противника (Op), определение тенденций управления), управление противником (M) (информационное воздействие на противника с целью передачи ему сведений соответствующих нашему замыслу), оперативная разведка (OI) (проверка результатов рефлексивного управления), оперативное управление (OM) — действия наших собственных сил по достижению требуемой цели.

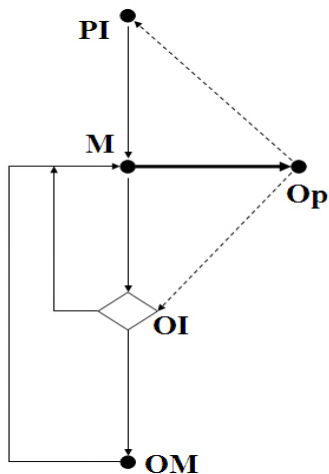


Рис. 55. Основные этапы информационных операций

При планировании или моделировании социальных процессов, в частности информационных операций, всегда необходимо учитывать, что общее поведение социальных систем невозможно определить, оперируя исключительно рафинированными математическими моделями. Это главным об-

разом обусловлено тем, что такие процессы в большой степени зависят от социально-психологических факторов.

Различают два основных типа информационных операций — наступательные и оборонительные. Однако, на практике, большая часть информационных операций является смешанной. Кроме того, большинство процедур информационных операций относятся одновременно к наступательным и оборонительным. Каждый из типов информационных операций, включая приведенные выше основные этапы, подразумевает некоторые особенности и уточнения.

Особенностью наступательных информационных операций (информационных атак) является то, что объекты воздействия таких операций определены и планирование основывается на достаточно точной информации об этих объектах. Информационная атака чаще всего требует нахождения или создания информационного повода (для оборонительных информационных операций поводом может являться сама информационная атака противника), раскрутка этого повода, т.е. пропаганда (в отличие от мер контрпропаганды при оборонительных информационных операциях), а также необходимость принятия мер по препятствию информационному противодействию.

Таким образом, план типовой информационной операции включает совпадающие на верхнем уровне для информационных операций обеих типов такие этапы, как оценка, планирование, исполнение и завершающая фаза. Приведем более детальный перечень компонент информационных операций.

В наступательных информационных операциях можно выделить такие основные фазы:

1. Оценка необходимости проведения операции:
 - 1) определение цели, прогноз достижимости, степени влияния;
 - 2) сбор информации.
2. Планирование.
3. Исполнение информационных воздействия:
 - 1) нахождение или создание информационного повода;
 - 2) раскрутка информационного повода (пропаганда);
 - 3) оперативная разведка;
 - 4) оценка воздействия;
 - 5) препятствие информационному противодействию;
 - 6) корректировка информационного воздействия.
4. Завершающая фаза:

1) анализ эффективности;
2) использование позитивных результатов информационного воздействия;

3) противодействие отрицательным результатам.

Типовая оборонительная информационная информация охватывает такие основные этапы:

1. Оценка:

1) анализ возможных уязвимостей (целей);

2) сбор информации о возможных операциях;

3) определение возможных «заказчиков» информационных воздействий:

— *определение сфер общих интересов объекта и потенциальных «заказчиков»;*

— *ранжирование потенциальных заказчиков по их интересам.*

2. Планирование:

1) стратегическое планирование оборонительной операции (явное или неявное):

— *определение критериев информационных воздействий;*

— *моделирование информационных воздействий с учетом: связей объекта; динамики воздействия; «особых» (критичных) точек воздействия;*

— *прогнозирование следующих шагов;*

— *расчет последствий.*

2) тактическое планирование контропераций.

3. Исполнение — отражение информационного воздействия:

1) выявление и «сглаживание» информационного повода;

2) контрпропаганда;

3) оперативная разведка;

4) оценка информационной среды;

5) корректировка информационного противодействия.

4. Завершающая фаза:

1) анализ эффективности;

2) использование позитивных результатов информационного воздействия;

3) противодействие отрицательным результатам.

Оперативное управление информационными операциями с использованием информационно-аналитических систем можно проиллюстрировать с помощью диаграммы, представленной на рис. 56.

В соответствии с приведенной диаграммой информация из реального мира (R) поступает в информационное пространство, в частности, в средства массовой информации (I) либо непосредственно экспертам (E), также через средства массовой информации. От экспертов или непосредственно из информационного пространства (например, с помощью средств контент-мониторинга) информация поступает в информационно-аналитическую систему (IAS). Информационно-аналитическая система передает лицам, принимающим решения (P), данные, которые определяют меры информационного воздействия на информационное пространство и непосредственно на объекты реального мира (людей, окружающую среду, компьютерные системы и т.д.).

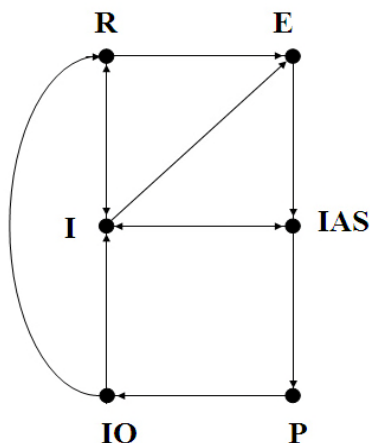


Рис. 56. Диаграмма оперативного управления с использованием информационно-аналитических систем

6.5. Особенности моделирования информационных операций

Моделирование можно рассматривать как один из способов решения проблем, возникающих в реальном мире, в частности, при планировании и проведении информационных операций. Чаще всего моделирование применяется в случаях, если эксперименты с реальными объектами невозможны либо слишком затратные. Моделирование охватывает отображение реальной проблемы в мир абстракции, изучение, анализ и оптимизацию модели и отображение оптимального решения обратно в реальный мир.

При моделировании существует два альтернативных подхода — аналитическое и имитационное моделирование. Идеальные аналитические модели допускают строгое аналитическое решение, или, по меньшей мере, постановку, например в виде систем дифференциаль-

ных уравнений. Однако, аналитические решения не всегда достижимы. Поэтому, особенно в последнее время, и особенно при решении задач из области социальной динамики все чаще применяются методы имитационного моделирования (*англ.* — *Simulation Modeling*). Имитационное моделирование представляет собой более мощное и практически незаменимое средство анализа социальных процедур. Имитационную модель можно рассматривать как множество правил, определяющих будущее состояние системы на основании текущего. При этом процесс моделирования заключается в наблюдении эволюции данным правилам системы во времени, и, соответственно, оценки адекватности модели, когда это возможно.

Наиболее перспективным направлением моделирования информационных операций является математическое описание самоорганизации среды восприятия и распространения информации с учетом сложившихся в текущий момент условий. Самоорганизующиеся среды, для которых отсутствует центральный механизм управления, а развитие идет за счет множества локальных взаимодействий, изучаются теорией сложных систем. Эта теория охватывает такие отрасли знаний, как нелинейная физика, термодинамика неравновесных процессов, теория динамических систем. Взаимодействия между отдельными элементами сложных систем определяют возникновение сложного поведения при отсутствии централизованного управления. Для исследований подобного поведения применяются самые современные методы, охватывающиеся междисциплинарной основой современной методологии — концепцией сложности. В настоящее время к теоретическим и технологическим основам этой концепции относятся теории детерминированного хаоса и сложных сетей, синергетика, фрактальный и волновой (вейвлет) анализ, многоагентное моделирование, теория самоорганизованной критичности (изучающей динамическое развитие до критического состояния, характеризуемого сильными пространственно-временными флуктуациями, без внешнего управления [126]), теория перколяции (*англ.* — *Percolation* — *протекание*) и т.п.

Моделирование социальных процедур (информационные операции, безусловно, относятся к таковым) предполагает проведение вычислительных экспериментов, так как чаще всего возникают существенные ограничения, затрудняющие проведение «полевых» натуральных экспериментов.

При моделировании информационных операций вычислительный эксперимент позволяет сократить операции по уточнению ограничений, подбору исходных данных, выбору правил функционирования компонент модели и т.д. В этом случае появляется возможность учета случаев, трудно реализуемых на практике, используя реальные данные лишь для идентификации параметров математической модели. Вместе с тем математическое моделирование имеет свои ограничения, реальный мир оказывается сложным для моделирования с достаточным уровнем детализации и точности, т.е. более или менее достоверные математические модели настолько сложны и многопараметричны, что не поддаются анализу и оценкам точными методами.

Отработать математические модели при планировании информационных операций можно лишь в процессе моделирования конкретных процедур, постоянно сопоставляя их с реальностью.

Выраженная цель методологии оценки информационных операций состоит в том, чтобы обеспечить своевременный и точный анализ возможных несоответствий между запланированной операцией и фактическим воздействием. Когда обнаруживаются существенные различия, которые влияют на вероятности успеха операции, аналитическая система должна сообщать об этом лицам, принимающим решения, для того, чтобы откорректировать текущие планы и решения. Вместе с тем, при планировании информационных операций нельзя действовать методом проб и ошибок, поэтому необходимо развивать методы, позволяющие обобщать ретроспективные данные, и на их основе проверять адекватность моделей.

В основу успешных моделей информационных операций закладываются синергетические подходы. Действительно, общество является сложной системой, каждая компонента которой характеризуется множеством признаков, имеет множество степеней свободы. При этом важным свойством этой системы является самоорганизация, которая является результатом взаимодействия таких компонент, как случайность, многократность, положительная и отрицательная обратная связь [127].

Особенностью математического моделирования информационных операций следует считать сравнительную простоту интерпретации получаемых результатов. Такие понятия, как «численность электората», «политический вес» и т.д., воспринимаются на интуитивном уровне даже без знакомства с точными (насколько они тут возможны) оп-

ределениями. А это позволяет делать подобный анализ актуальных ситуаций предметом широкого обсуждения.

В силу того, что некоторые решения являются неустойчивыми по отношению к своим параметрам, значения таких параметров необходимо определять с высокой точностью. Для этого требуется комплекс методик, основанных не только на обработке больших объемов статистических данных, но и на разносторонних социологических исследованиях.

В настоящее время реалистичной выглядит постановка задачи, состоящая в использовании математических моделей для прогнозирования возможных сценариев динамики социальных процессов на качественном уровне. В такой формулировке моделирование динамики занимает как бы промежуточный уровень между тем, что изложено здесь, и точным прогнозированием. И все же потребуются выбор значений параметров, которые бы в некотором разумном приближении соответствовали изучаемой ситуации, причем в большинстве случаев продуктивным оказывается использование относительных величин. Так, конечно, не получить достоверных данных о будущем развитии событий, но, скорее всего, можно составить более или менее адекватную картину того, что и как может произойти. А это уже не мало.

Для достижения успеха при этом отдельные информационные воздействия необходимо рассматривать как части единой информационной операции, точно так же, как артобстрел или авиационные атаки можно рассматривать как согласованные части военной операции.

6.6. Мониторинг и анализ информационных операций

Современное информационное пространство представляет собой уникальную возможность получения любой информации по выбранному вопросу, при условии наличия соответствующего инструментария, применение которого позволяет анализировать взаимосвязь возможных событий или событий, которые уже происходят, с информационной активностью определенного круга источников информации. С другой стороны, при ретроспективном анализе любого процесса или явления интерес представляют определенные характеристики его развития, а именно:

— количественная динамика, присущая процессу или явлению, например, количество событий в единицу времени, или количество сообщений, имеющих отношение к нему;

— определение критических, пороговых точек, которые соответствуют количественной динамике явления;

— определение проявлений в критических точках, например, выявления основных сюжетов публикаций в СМИ относительно выбранного процесса или явления;

— после выявления основных проявлений явления в критических точках, эти проявления ранжируются, и исследуется динамика развития отдельных определенных проявлений до и после определенных критических точек;

— осуществляется статистический, корреляционный и фрактальный анализ общей динамики и динамики отдельных проявлений, на основе которых осуществляются попытки прогнозирования развития явления и отдельных его проявлений.

Для исследования взаимосвязи реальных событий и публикаций о них в сети Интернет авторами использовалась система InfoStream, обеспечивающая интеграцию и мониторинг сетевых информационных ресурсов.

Количество веб-публикаций в день по какой-либо теме, а особенно изменения (динамика) этой величины порой позволяют даже небольшим специалистам в предметной области делать более-менее точные выводы.

Получить данные подобной динамики можно, например, ежедневно заходя на сайты интеграторов новостей (news.yandex.ru, webground.su, uaport.net). Конечно, в лучшем положении пользователи профессиональных систем мониторинга типа Интегрум или InfoStream. Именно на основе последней системы получена удивительная статистика по количеству веб-публикаций по тематике эпидемий гриппа в разные периоды.

К примеру, на рис. 57 показана динамика публикаций в RUNet по запросу «птичий грипп» за период с середины 2005 года до конца первого полугодия 2008 года, полученная с помощью системы InfoStream.

Безусловно, мелкие колебания количества веб-публикаций, связанные с недельной цикличностью, можно сглаживать, но все равно на рис. 57 видны три большие пиковые области, с максимумами, прихо-

дящимися на декабрь-январь в течение трех лет. Видно, как из года в год тема (даже в критические сезоны) теряла свою актуальность. Итак, наблюдается периодичность, снижение актуальности, колоколообразная форма динамики в критические сезоны.

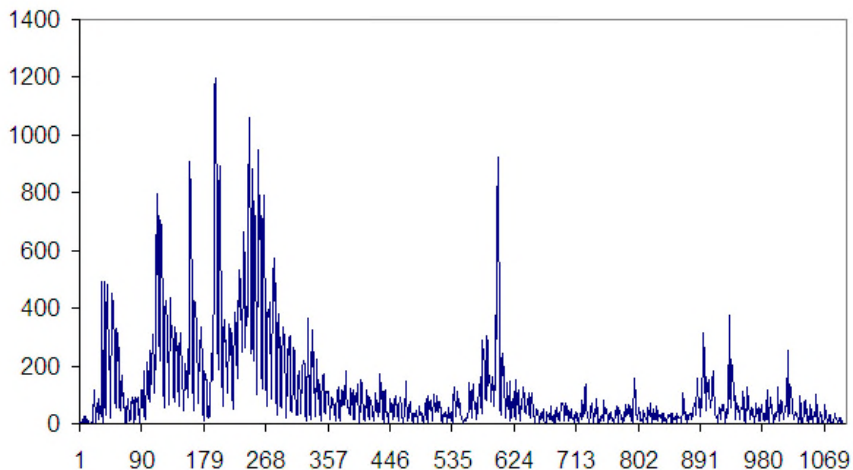


Рис. 57. Динамика веб-публикаций по запросу «птичий грипп»

Совсем иначе выглядит динамика веб-публикаций по запросу «свиной грипп» (рис. 58). Данные получены с момента появления в RUNete первых сообщений по этой теме (второй квартал 2009 года) по октябрь 2010 года.

Можно видеть два внезапно возникающих пика в апреле 2009 года и конец октября 2010 года. Затем количество веб-публикаций резко уменьшается — практически по гиперболической зависимости. Первый пик связан с первыми проявлениям в мире вируса А/Н1N1 и выделениями во всем мире (прежде всего в США) огромных средств на борьбу с ним, другой — с осенними проявлениями во всех странах мира, но прежде всего в России и Украине. Абсолютные пиковые значения более чем в 3 раза превысили пиковые значения птичьего гриппа.

Официальная статистика нам уже рассказала, что с последней эпидемией удалось успешно справиться, смертность от А/Н1N1 оказалась более низкой, чем от обычного сезонного гриппа. Налицо

вспышки активности веб-публикаций, очень быстро рассасывающиеся. Очевидно, сообщения о свином гриппе вначале носили сенсационный характер, но затем не было естественной информационной подпитки. Если бы в средние века был Интернет, то информация о волнах чумы, наверное, имела именно такой же характер... Но последствия тогда были иными. Действительно, очень удобная платформа для сторонников идеи конспирологии и всемирных заговоров.

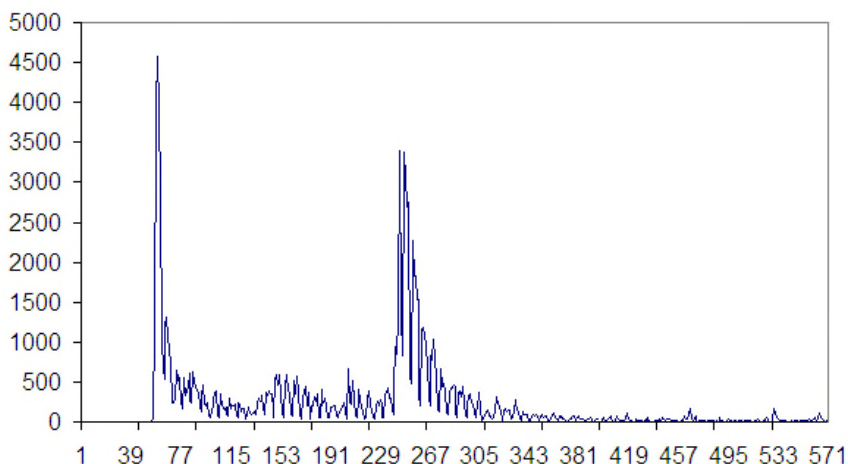


Рис. 58. Динамика веб-публикаций по запросу «свиной грипп»

Как видим, отсутствие естественной информационной подпитки выражается в резком всплеске количества публикаций, а затем также в достаточно резком (гиперболическом или даже экспоненциальном) спаде. Зачастую информационные операции сопровождаются именно таким поведением динамики веб-публикаций.

Приведем пример одной такой информационной операции, проводимой в конце октября 2009 года оператором мобильной связи Tele2 из Латвии. Речь шла о мистификации по поводу падения метеорита в городе Мазсалац. График динамики публикаций по этой теме, полученный с помощью системы InfoStream, приведен на рис. 59.

Понятия в динамике :
♦ Мазсалац Латви метеор

Дата	Всего	Понятие
2009.10.23	61306	0
2009.10.24	17344	0
2009.10.25	14091	43
2009.10.26	61382	703
2009.10.27	62047	366
2009.10.28	62247	117
2009.10.29	63346	18
2009.10.30	62463	14
2009.10.31	18744	3
2009.11.01	14821	0
2009.11.02	62255	6
2009.11.03	60960	1
2009.11.04	44777	2
2009.11.05	63156	7
2009.11.06	60712	6
2009.11.07	17861	0

Рис. 59. Динамика веб-публикаций по тематике появления метеорита в Мазсалаце

Первые публикации 25 октября были посвящены самому факту появления «метеорита» и попыткам объяснить его космическое происхождение (рис. 60).

♦ (Мазсалац Латви метеор) & (2009.10.25)

Найдено документов - 43, страница 1 из 3

Статистика слов: Добавить канал

♦ **МАЗСАЛАЦ** - 2280, **ЛАТВИ** - 103447, **МЕТЕОР** - 20838, **2009.10.25** - 14096.

- Эксперты считают, что в Латвии мог упасть не метеорит, а спутник**
 Правда.Py 2009.10.25 23:59
 РИА Новости Представители Государственной пожарно-спасательной службы Латвии полагают, что упавший на территорию города Мазсалац метеорит мог вполне оказаться спутником или его деталью. Этой информацией с журналистами поделилась помощник начальника ГПСС Инга Ветере, добавив, что в районе падения неопознанного объекта может быть радиационное загрязнение.
 Похожие документы - Оригинал
- В Латвии упал метеорит**
 Zvezda.com.ua 2009.10.25 23:54
 ВВ районе латвийского города Мазсалаца сегодня упал метеорит, образовавший кратер шириной 20 метров и глубиной 5 метров, никто не пострадал, передает телеканал "Вести".
 Похожие документы - Оригинал
- Упавший в Латвии с неба объект может быть фрагментом спутника**
 Газета.Ru 2009.10.25 23:51
 Специалисты Государственной пожарно-спасательной службы (ГПСС) Латвии не исключают, что упавший на территории страны в воскресенье объект - искусственный спутник Земли или его фрагмент, а не метеорит, сообщила помощник начальника ГПСС Инга Ветере. Ранее сообщалось, что в районе городка Мазсалаца в Валмиерском районе Латвии упал метеорит.
 Похожие документы - Оригинал
- Оцепление вокруг упавшего в Латвии "небесного тела" расширено из-за угрозы радиации**
 Trend.az 2009.10.25 23:50
 Пожарные и полиция Латвии расширили зону оцепления вокруг воронки, образовавшейся в результате падения с неба

Информационный портрет

Уточнить запрос

Рубрики (29)

Языки (1)

Страны источников (17)

Источники (50)

AND NOT

РИА Новости *

Новости@MAIL.RU

Газета.Ru

Mixnews.lv

Ves.lv

Trend.az

Date.BS

Телеграф.lv

DELFI

Радио "Эхо Москвы"

Взгляд.ру

Gorod.lv

Рис. 60. Публикации в начале изучаемого информационного сюжета

По-видимому, вопреки ожиданиям мистификаторов, наибольший отклик в веб-пространстве получило их разоблачение, информационная операция вышла из под контроля своих авторов (рис. 61).

7. **Упавший в Латвии метеорит был размером с кулак, считают ученые**
ФАКТNEWS 2009.10.26 23:10
Метеорит, который в воскресенье вечером упал в Латвии, скорее всего, состоял из железа | delfi.lv Неуставленный "небесный объект", упавший в воскресенье вечером на территории латвийского города Мазсалаца Валмиерского района, по размеру был не больше кулака, сказал в понедельник специалист Центра среды, геологии и метеорологии балтийской республики Улдис Нупле, который ночью побывал на месте ЧП.
Похожие документы - Оригинал
8. **Латышский "метеорит" - это фальшивка. В искусственной воронке взорвали селитру, заподозрили эксперты**
ФАКТNEWS 2009.10.26 23:10
Упавший в латвийском Мазсалаце "небесный объект" может оказаться фальшивкой. Как передает "Интерфакс", к такому выводу пришел геолог Зиельмевидземской биосферической резервации Дайнис Озолс, прибывший утром на место падения | Первый канал Упавший в латвийском Мазсалаце "небесный объект" может оказаться фальшивкой.
Похожие документы - Оригинал
9. **Тайна латвийского метеорита раскрыта - Репортаж**
Петербург-Пятый канал 2009.10.26 23:02
В поле у латвийского городка Мазсалаца сегодня нашли неопознанный объект. Местный фермер растрюбил на весь мир: этот огненный шар упал с неба, это метеорит или даже военный спутник.
Похожие документы - Оригинал
10. **Латвийский метеорит оказался пиар-акцией оператора мобильной связи**
Завтра.com.ua 2009.10.26 22:54
Сообщение о падении в Латвии метеорита, ставшее одной из главных новостей местных и мировых СМИ, оказалось шуткой. В инсценировке падения "метеорита" признался оператор мобильной связи Tele2.
Похожие документы - Оригинал

Рис. 61. Публикации в «пиковый» период жизни информационного сюжета

И, наконец, хвост информационного сюжета полностью посвящен мерам наказания мистификаторов (рис. 62).

6. **Инсценировка падения метеорита в Латвии не является преступлением**
Политсовет 2009.11.26 16:55
Алена Клименко Сегодня, 26 ноября, помощник начальника Управления полиции Видземского района Латвии Илзе Унгуре заявила, что правоохранители не будут возбуждать уголовный процесс в отношении оператора мобильной связи TELE2, который в октябре инсценировал падение метеорита в районе города Мазсалаца. По словам Унгуре, в действительности компания не обнаружила состава преступления. Вместе с тем, представитель полиции не уточнила, возместил ли оператор все.
Похожие документы - Оригинал
7. **Автор шутки с падением "метеорита" в Латвии не будет подвергнут наказанию**
Penki.lt 2009.11.26 16:51
Автор шутки с падением 26 октября "метеорита" в латвийском городе Мазсалаца оператор мобильной связи "Tele2" не будет подвергнут уголовному наказанию, поскольку в его действиях полиция не усматривает состава преступления.
Похожие документы - Оригинал
8. **Заказчика "падения метеорита" в Латвии оштрафуют на 50 долларов**
Газета "Известия" 2009.11.26 16:38
Полиция не нашла в действиях автор шутки с падением 26 октября "метеорита" в латвийском городе Мазсалаца - оператора мобильной связи Tele2 - ничего противозаконного, а потому он не будет подвергнут уголовному наказанию.
Похожие документы - Оригинал

Рис. 62. Завершение жизненного цикла информационного сюжета

Тематика следующего исследуемого информационного потока определялась запросом к системе InfoStream относительно «военных действий» в информационном пространстве страны:

(інформац~воїн & украї) | (інформац~війн & украї).

Документы, релевантные приведенному запросу, могут быть представлены на двух языках (украинском и русском), содержать словосочетания типа «Информационная война», или «Информационной войне», а также содержать в себе название нашей страны. Приведенные запросы соответствуют понятию «информационные войны», которое чаще всего применяются в веб-среде как функциональный синоним «информационных операций». Ретроспективный период исследования представлял весь 2008 год. В результате поиска по наиболее широким запросам было найдено 6196 документов. На основе обработки этих данных была получена полная картина экспериментальных данных — временные ряды за заданный период. На рис. 63 приведен график количества публикаций по запросу по дням 2008 года.

Представленный график учитывает недельные колебания (по выходным дням, например, в сети Интернет публикуется значительно меньше документов, чем в будни).

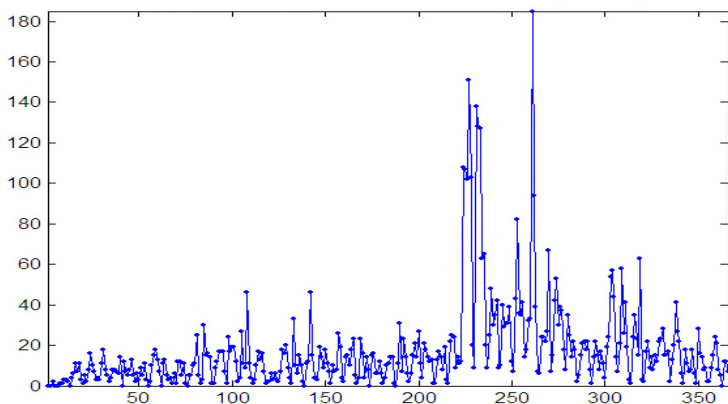


Рис. 63. Динамика количества публикаций по запросу по дням 2008 года (вместе 6196 публикаций)

Для более наглядного отображения тенденций подобные графики сглаживаются методом «скользящего среднего» с окном наблюдения в 7 суток. На рис. 64 приведен сглаженный график, соответствующий приведенной выше динамике. В частности, можно видеть, что приблизительно в районе 220-го дня года общее количество сообщений по тематике информационных войн резко увеличилась.

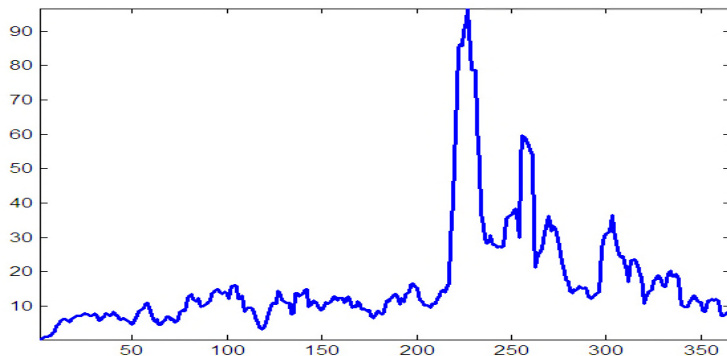


Рис. 64. Сглаженный график количества публикаций по дням 2008 года

На рис. 65–67 приведены обобщенные системой InfoStream содержания сообщений, которые относятся к некоторым пиковым значениям интенсивности публикаций в течение 2008 года.

Задача изучения статистических свойств сетевых документальных массивов [65] является многоплановой, допускает активное использование современных методов, в том числе и методов фрактального анализа [128–130], позволяющих глубже понять специфику предметной области.

(*) Обзор основных сюжетов
 ((информац-воин & украи) | (информац-війн & украї) & (2008.04.17) :
 документов - 46, сюжетов - 7

1. **Противостояние Ющенко и Тимошенко ослабляет позиции Украины**
 Председатель парламентского комитета по вопросам национальной безопасности и обороны Анатолий Гриценко считает, что противостояние между Президентом и правительством ослабляет позиции страны на международном уровне... Как передает корреспондент УНИАН, об этом он заявил журналистам в Киеве. "Противостояние между Президентом и правительством перешло в открыто враждебную фазу. Сюжет полностью (28)

2008.04.17 00:14 Борьба Президента и главы правительства за власть разрушает Украину УРА-Информ 28
 2008.04.17 19:33 А.Гриценко (НУ): Противостояние между Ющенко Тимошенко перешло в "войну на уничтожение" Укрпарлнформ

2. **Война неизбежна**
 Если секретариат президента не прекратит блокировать работу правительства, БЮТ обещает объявить ему информационную войну. Но в силу того что коалиция, малой частью которой руководит Виктор Балага, не может быть изменена, единственным возможным выходом из клинча президента и премьера остается отставка Кабмина и перереформирование коалиции. Вчерашнее заявление премьера Юлии Тимошенко в Страсбурге о том, что Украина Сюжет полностью (11)

2008.04.17 08:02 Взгляд: Коалиция треснула Корреспондент.net 11
 2008.04.17 12:29 Война неизбежна Экономические известия

Рис. 65. Фрагмент основных сюжетов за 17 апреля 2008 года (доминирующая тематика — выступление премьер-министра Украины в Страсбурге)

Обзор основных сюжетов	
(информац-войн & украи) (информац-войн & украи) & (2008.08.14) ;	
документов - 151, сюжетов - 29	
<p>1. Балтийские пособия кавказской трагедии</p> <p>"Южная Осетия - это очень небольшой регион, и я считаю, что ее воссоединение с Грузией - вопрос нескольких месяцев". Это заявление Михаила Саакашвили примечательно тем, что оно сделано не в начале августа 2008-го, а четыре года назад, в интервью эстонской газете "Постимесз" накануне визита президента Грузии в Латвию, Литву и Эстонию. Сюжет полностью (42)</p>	<p>2008.08.14 00:00 Волки и овцы Российская газета 42 2008.08.14 21:01 Кавказский Садам "2000" Еженедельник</p>
<p>2. "Нас никто не остановит при возвращении на базу в Севастополь"</p> <p>"Новая газета" (Россия) Конфликт вокруг Черноморского флота. Наш собственный корреспондент - из Крыма Во вторник 12 августа президент Украины Виктор Ющенко вновь стоял на Майдане - на площади Свободы в центре Тбилиси, куда он прибыл с коллегами из Польши, Эстонии, Литвы и Латвии. Сюжет полностью (36)</p>	<p>2008.08.14 00:19 Исследование: Украина проиграла информационную войну России в освещении конфликта в Грузии "Завтра" 36 2008.08.14 21:22 Новые информационные войны Газета "День"</p>
<p>3. Российские хакеры-"миротворцы" напали на украинский портал</p> <p>Портал delo.ua подвергся хакерской атаке из-за сегодняшних публикаций в газете "Дело". Об этом говорится в пресс-релизе киевского издательства "Экономика", поступившем в "Обком": "Сегодня в 11 часов утра по киевскому времени украинский портал delo.ua подвергся DDoS атаке с десятка тысяч компьютеров, расположенных по всему миру. Сюжет полностью (9)</p>	<p>2008.08.14 16:09 Сайт delo.ua атакували через матеріали про російсько-грузинську інформаційну війну Телекo 9 2008.08.14 21:34 Русские хакеры атаковали украинский сайт обозреватель</p>

Рис. 66. Фрагмент основных сюжетов за 14 августа 2008 года (доминирующая тематика — российско-грузинский военный конфликт)

Обзор основных сюжетов	
(информац-войн & украи) (информац-войн & украи) & (2008.11.14) ;	
документов - 63, сюжетов - 12	
<p>1. Украина могла бы эффективнее защищаться от России в информационной войне</p> <p>Украинское министерство иностранных дел возмущено действиями посла России в Украине Виктора Черномырдина. Он выступил "соорганизатором откровенно провокационной антиукраинской акции", во время которой планировался показ российского фильма "Искусство предательства". В фильме, который так хотело показать украинцам российское посольство, речь идет о российско-грузинском конфликте и якобы участии в нем украинцев. Сюжет полностью (34)</p>	<p>2008.11.14 01:09 Посольство России требует от Киева объяснений за "грубость" Росбалт Украина 34 2008.11.14 21:41 Россия ведет информационную войну относительно Украины Vlasti.net</p>
<p>2. Росія проводить інформаційну війну щодо України - експерт</p> <p>Українське міністерство закордонних справ обурене діями посла Росії в Україні Віктора Черномірдіна. Він виступив "співорганізатором відверто провокаційної антиукраїнської акції", під час якої планувався показ російського фільму "Мистецтво зради". У фільмі, який так хотіло показати українцям російське посольство, йдеться про російсько-грузинський конфлікт та нібито участь у ньому українців. Сюжет повністю (11)</p>	<p>2008.11.14 10:03 Росія веде проти України інформаційну війну Газе по-українськи 11 2008.11.14 19:02 Росія веде проти України інформаційну війну Рух</p>

Рис. 67. Фрагмент основных сюжетов за 14 ноября 2008 года (доминирующая тематика — российский фильм «Искусство предательства»)

На рис. 68 представлено соотношения R/S для ряда количества публикаций по дням 2008 года, соответствующее приведенному выше запросу. Наклон этой прямой соответствует показателю Херста.

Как можно видеть, значение Херста для исследуемых информационных потоков соответствует $\sim 0,81$, что подтверждает предположение о самоподобии и итеративности процессов в информационном пространстве. Это означает, что резонансные публикации многократно дублируются, пересказываются, обсуждаются. Это также означает, что общая информационная напряженность остается на большом уровне, как только исчезает «шлейф» одного сюжета по тематике информационных операций, ему на смену возникает новый сюжет, чаще всего, как показывают тенденции, более интенсивный.

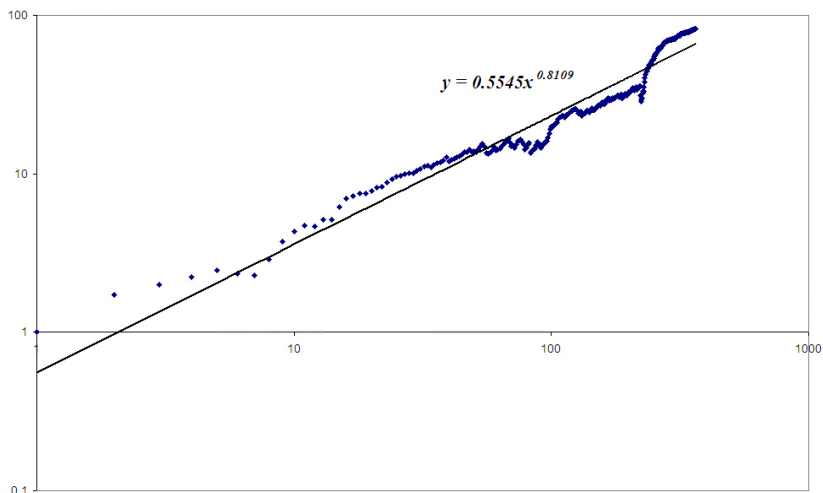


Рис. 68. Показатель нормированного размаха в логарифмической шкале для всего периода наблюдений

К наиболее распространенным инструментальным средствам математического моделирования и оценки рядов наблюдений относится также вейвлет-анализ [100, 101]. Каждый из основных факторов динамики исходного процесса имеет свое, характерное отражение на скейлограмме, при этом вся аналитическая информация представляется в наглядном и удобном для изучения виде. На рис. 69 приведенная скейлограмма — результат непрерывного вейвлет-анализа (вейвлет

Хаара) временного ряда, который соответствует исследуемому процессу.

На скелетоне для большинства частот отмечены не только 220-й день, но и неявные экстремумы (105-й, 130-й, 200-й дни и т.п.) Основное наблюдение: публикации по тематике информационных войн вызывают лавину републикаций и домыслов, которые влияют на общественную мысль и, в конечном счете, на информационную безопасность как бизнеса, так и государства.

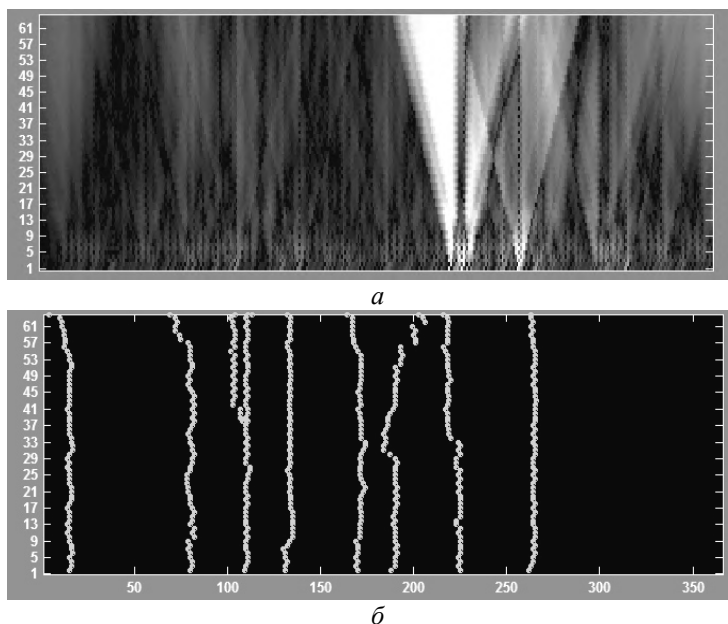


Рис. 69. Результат вейвлет-анализа (непрерывное вейвлет-преобразование): а — вейвлет-скейлограмма; б — линии локальных максимумов (скелетон)

Обычная сетевая информационная атака в веб-среде сегодня производится следующим образом: как правило, создается и некоторое время функционирует веб-сайт (назовем его «первоисточником»), при этом он публикует вполне корректную информацию. В час X на его странице появляется документ, обычно компромат на объект атаки, достоверный либо сфальсифицированный. Затем происходит так называемая «отмывка информации». Документ перепечатаывают ин-

тернет-издания двух типов — заинтересованные в атаке и те, кому попросту не хватает информации для заполнения своего информационного поля. В случае претензий все перепечатавающие издания ссылаются на «первоисточник» и в крайнем случае по просьбе/требованию объекта атаки удаляют со своих веб-сайтов информацию. Первоисточник при необходимости также снимает информацию либо вообще ликвидируется (после чего оказывается, что он зарегистрирован в Интернет на несуществующее лицо). Вместе с тем информация уже разошлась, задача первоисточника выполнена, атака стартовала.

У всех на слуху информационная кампания, направленная против «Проминвестбанка», начавшаяся в конце сентября 2008 г. С помощью системы контент-мониторинга InfoStream (<http://infostream.ua>), сканирующей все основные информационные веб-сайты Украины в режиме реального времени, была определена динамика публикаций на веб-сайтах сообщений, в которых упоминался «Проминвестбанк» за три месяца — сентябрь, октябрь и ноябрь (рис. 70). Эта динамика свидетельствует о небольшом количестве публикаций за первую половину сентября, однако затем пошел ряд публикаций, компрометирующих председателя правления В. Матвиенко, что вызвало относительно небольшой резонанс.

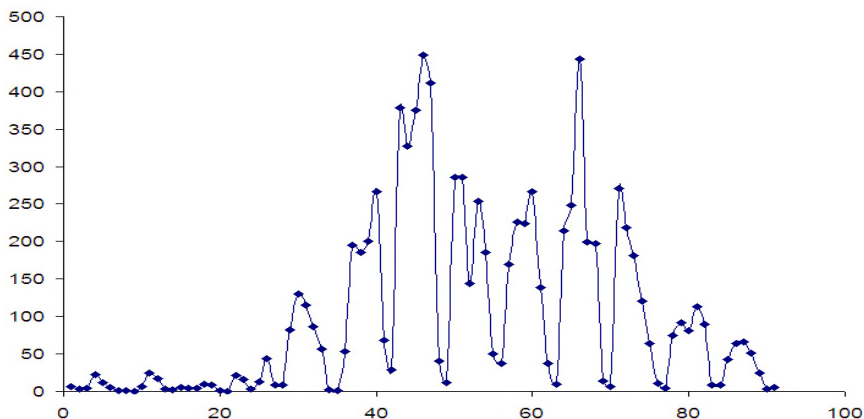
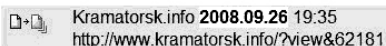


Рис. 70. Динамика публикаций по теме «Проминвестбанк» за три месяца 2008 г.

Как оказалось впоследствии, эти публикации были лишь «арт-подготовкой». 26 сентября появились первые сообщения о возможном банкротстве банка (рис. 71), количество которых вполне соответствовало лавинообразному процессу, ограниченному лишь числом веб-сайтов, способных публиковать подобную информацию. Впрочем, этот процесс вышел на стабильно-средний уровень к декабрю 2008 г.

 Kramatorsk.info 2008.09.26 19:35
<http://www.kramatorsk.info/?view&62181>

В Донбассе вошла в активную фазу атака на Проминвестбанк. ПИБ заявляет, что это атака из-за рубежа

Сегодня в Донецкой области население организовано вышло к проходным **Проминвестбанка**.

Вести о том, что народные массы Донбасса штурмуют отделения ПИБа в Донецке, Авдеевке, Волновахе и пр. населенных пунктах Донецкой области, приходят в "Обком" с середины дня.

Никто из опрошенных нами экспертов не может пока сказать что-либо конкретное по данному поводу, кроме банальных констатаций: ПИБ - серьезный банк, он кредитует промышленный сектор Украины, Донецкое облотделение ПИБа - одно из крупнейших, борьба за него началась еще в середине 90-х годов... Ну а баннеры на киевских дорогах против нынешнего (неизменного) руководства ПИБа во главе с г-ном Матвиенко видели многие автомобилисты и пассажиры столичного транспорта.

"Обком" пока не готов сказать что-то определенное по поводу паники, которая охватила сегодня трудовой Донбасс - хотя сведения для определенных умозаключений, в принципе, имеются. Вместо этого мы предлагаем внимаю вкладчиков сообщение, поступившее от пресс-службы ПИБа:

"**Проминвестбанк** заявляет о стабильной работе, несмотря на дезинформацию в ряде СМИ о якобы приближающемся банкротстве банка.

Проминвестбанк, по оценкам зарубежных экспертов, стабильный банк и занимает в Украине второе место по надежности.

Массовая газетная атака на **Проминвестбанк** организована рейдерскими (бандитскими) группировками зарубежных агентов с участием высокопоставленных чиновников крупных государственных структур, которые по Конституции должны защищать отечественные предприятия и банки. Ложь, шантаж, направленные против банка, преследуют цель вынудить его к продаже иностранцам за комиссионное вознаграждение... Заявляем: банк не продается... **Проминвестбанк** останется украинским!", - говорится в сообщении.

Служба информационной поддержки **Проминвестбанка** также сообщает, что, несмотря на беспокойство вкладчиков, вызванное негативными публикациями о банке, все обязательства перед клиентами и вкладчиками выполняются, а структурные подразделения банка работают в нормальном режиме.

"Обком"

Рис. 71. Одно из первых тревожных сообщений

Нельзя утверждать, что лишь информационная атака через Интернет привела банк к печальному состоянию, однако именно первые тревожные сообщения подорвали доверие многих вкладчиков, заставили их массово забирать свои сбережения из банка. 30 сентября поя-

вилось сообщение, что для спасения банка НБУ решило выделить «Проминвестбанку» 5 млрд гривен рефинансирования, а 5 декабря появилось сообщение, что у «Проминвестбанка» появился новый владелец (рис. 72). После этого объемы публикаций о «Проминвестбанке» существенно сократились, что свидетельствует не столько об его оздоровлении, сколько о системном кризисе банковской системы Украины, «уронившему» многие другие кредитные и банковские учреждения.

Активная база данных: Система интеграции интернет-ресурсов

Главная Помощь Кабинет Источники Статистика Новости проекта

Вход Выход

InfoStream Online

(проминвестбанк)&(2008.11.05)

Период: Другой Убрать дубли Морфология

От: 200809 До: 200811

Найти Динамика Дайджест

Очистить События Сюжеты

Риск запросов Примеры

↓ (проминвестбанк) & (2008.11.05)

Найдено документов - 443, страница 1 из 30

Статистика слов:

ПРОМИНВЕСТБАНК - 26463, 2008.11.05 - 58209

Добавить канал

- Матвиенко поделился "Проминвестбанком" с братьями Клюевыми**
Політбайні 2008.11.05 22:08
40% акций "Проминвестбанка" досталось братьям Клюевым. Эту информацию газете "Сегодня" подтвердил источник в руководстве Партии Регионов. По словам информатора, денег за это братья Клюевы не заплатили. Похожие документы - Оригинал
- ВЧЕРА ОФИЦИАЛЬНЫЙ И РЫНОЧНЫЙ КУРСЫ ДОЛЛАРА ПОЧТИ СРАВНЯЛИСЬ, И В ОБМЕННИКАХ АМЕРИКАНСКУЮ ВАЛЮТУ ПРОДАВАЛИ ПО 6,85 ГРИВНИ**
Газета "Факты и комментарии" 2008.11.05 21:30
А "Проминвестбанк" обрел новых владельцев Роберт ВАСИЛЬ "ФАКТЫ" Национальный банк в среду продолжил свою деятельность по укреплению курса гривны на наличном и межбанковском рынке и параллельно по ослаблению официального курса американской валюты. Курс доллара, установленный Нацбанком, вчера вырос приблизительно на 3,5 копейки и достиг значения 5,82611 гривны за доллар. Похожие документы - Оригинал
- "Проминвестбанк" сменил собственника**
Газета "День" 2008.11.05 21:22
Правительство утверждает, что не будет расходовать деньги налогоплательщиков на рекапитализацию банка Наталья БИЛОУСОВА, "День" Вчера об этом официально сообщил на своем сайте Национальный банк Украины (НБУ). Похожие документы - Оригинал
- Кабинет утвердил правила для капитализации банков**
Укр.Информ 2008.11.05 21:13
Кабинет министров Украины утвердил порядок участия государства в капитализации банков. Соответствующее постановление от 4 ноября 2008 г. N 960 размещено на сайте правительства. Похожие документы - Оригинал
- Дмитрий Фирташ покупает почти 90% банка "Надра"**
Время.info 2008.11.05 20:17
Вчера, 4 октября, украинский предприниматель Дмитрий Фирташ, совладелец швейцарского газового трейдера RosUkrEnergy, подписал предварительное соглашение о покупке 86,7% акций банка "Надра". Похожие документы - Оригинал
- У Проминвестбанка поменялся владелец (5.11.2008 18:00)**
NTV (рус) 2008.11.05 19:45
new У Вас есть видео, которое Вы хотите показать всему миру? Вам сюда В Проминвестбанка изменился владелец. Факт продажи акций банка подтвердили в НБУ. СТБ Похожие документы - Оригинал
- Население будет покупать доллары по официальному курсу**
АМІ Новости:Украина 2008.11.05 19:45
Национальный банк Украины своим постановлением N353 от 5 ноября обязал коммерческие банки продавать населению наличные доллары по курсу не выше официального, сообщают "Українські новини". Похожие документы - Оригинал

Информационный портрет	
Уточнить запрос	
Языки (1)	
Страны источников (1)	
Источники (1)	
Размер (1)	
Цифровая насыщенность (1)	
География (1)	
Компании (1)	
Слова (12)	
Классификатор-навигатор	
ОБКОМ	
ВОЛНОВАХА	
АТАКА	
ФАЗА	
РУБЕЖ	
ВКЛАДЧИК	
ДОНБАСС	
ГОТ	
ВОЛНОВАХА	
АТАКА	
ФАЗА	
РУБЕЖ	
ВКЛАДЧИК	
ДОНБАСС	
УМОЗАКЛЮЧЕНИЕ	
ВОЛНОВАХА	
АТАКА	
ФАЗА	
РУБЕЖ	
ВКЛАДЧИК	
ДОНБАСС	
ПРОХОДНАЯ	
ВОЛНОВАХА	
АТАКА	
ФАЗА	
РУБЕЖ	

Рис. 72. Сообщения, завершившие экстремальную динамику интенсивности публикаций по теме «Проминвестбанк»

Буквально через неделю после описанных выше событий в Украине произошла еще одна публичная знаковая информационная атака, в этот раз на рынке страхования. Это была настоящая информационная операция против НАСК «Оранта». В этом случае первоисточником компромата оказался не веб-сайт, а информационное сообщение, разосланное электронной почтой тысячам пользователей Интернет. В результате применения специальных технических приемов оно разошлось с обозначением адреса пресс-службы объекта атаки. И так, 10 декабря 2008 года в районе 11:30 в виде спама было разослано информационное сообщение, в котором говорилось о том, что страховая компания «Оранта» заявляет о банкротстве. По предварительным данным, информация разлетелась по 1000 адресам, естественно, данные попали к конкурентам и в СМИ. В сообщении говорилось, что компания с 31 декабря 2008 года прекращает выполнять взятые перед клиентами обязательства.

В связи со случившимся НАСК «Оранта» обратилась в правоохранительные органы с просьбой расследовать данный инцидент и наказать виновных. Произошедшее с «Орантой» очень напоминало ситуацию с «Проминвестбанком», с этим согласились многочисленные эксперты. Ведь как банковский бизнес, так и страховой основываются на доверии клиентов, которое легче всего подрывается именно информационными атаками. По словам Олега Спилки, председателя наблюдательного совета НАСК «Оранта», «Это мероприятие готовилось целенаправленно для того, чтобы дискредитировать страховую компанию и подорвать ее репутацию». Не вдаваясь в детали возможных целей атаки (смена владельцев, борьба за блокирующий пакет акций, уничтожение компании и т.п.), с помощью ретроспективного анализа проследим за динамикой публикаций в сети Интернет, в которых упоминалась НАСК «Оранта».

На рис. 73 приведена посуточная динамика количества соответствующих публикаций. На этой диаграмме, кроме всего прочего, отчетливо виден спад интенсивности публикаций по данной теме в начале декабря 2008 г., что вполне можно воспринимать как некоторое «затишье перед бурей».

Для анализа временных рядов в рамках исследования авторов применялся ΔL -метод. На рис. 74 представлена скейлограмма динамики рассматриваемого процесса с помощью метода (ΔL -метода) за второе полугодие 2008 года. Несмотря на отдельные пики в 16 и 55

день квартала, все же наибольший интерес представляет экстремум, приходящийся именно на 10–12 декабря.

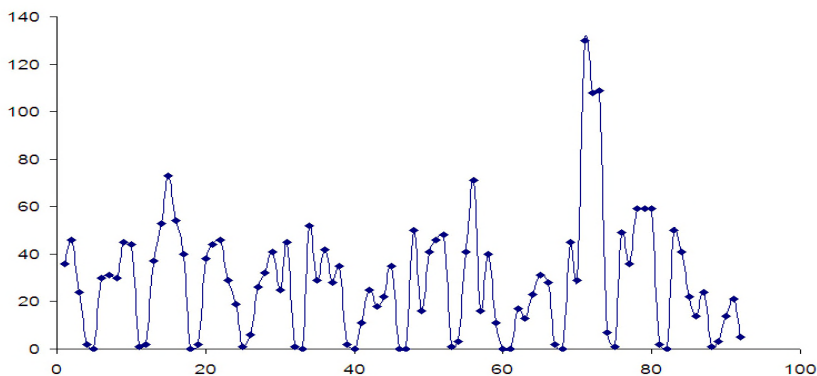


Рис. 73. Интенсивность публикаций в Интернет по теме «Оранта»

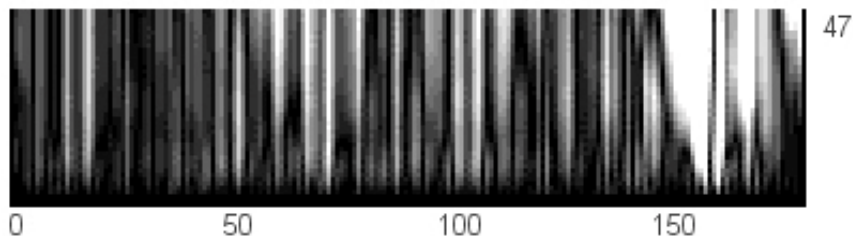


Рис. 74. ΔL -диаграмма ряда публикаций по теме «Оранта»

Более детальная статистика публикаций по теме «Оранта» за декабрь 2008 года получена через интерфейс пользователя системы контент-мониторинга InfoStream (рис. 75).

Проследим за ходом информационной операции, рассматривая сообщения, публикуемые в разные промежутки времени.

На рис. 76 приведен список публикаций по теме «Оранта» в течение первых часов атаки. По словам Олега Спилки, в течение двух часов с начала атаки все почтовые серверы НАСК «Оранта» были выведены из строя, поэтому опровержение в сети задержалось.

6. Живучесть информационных операций

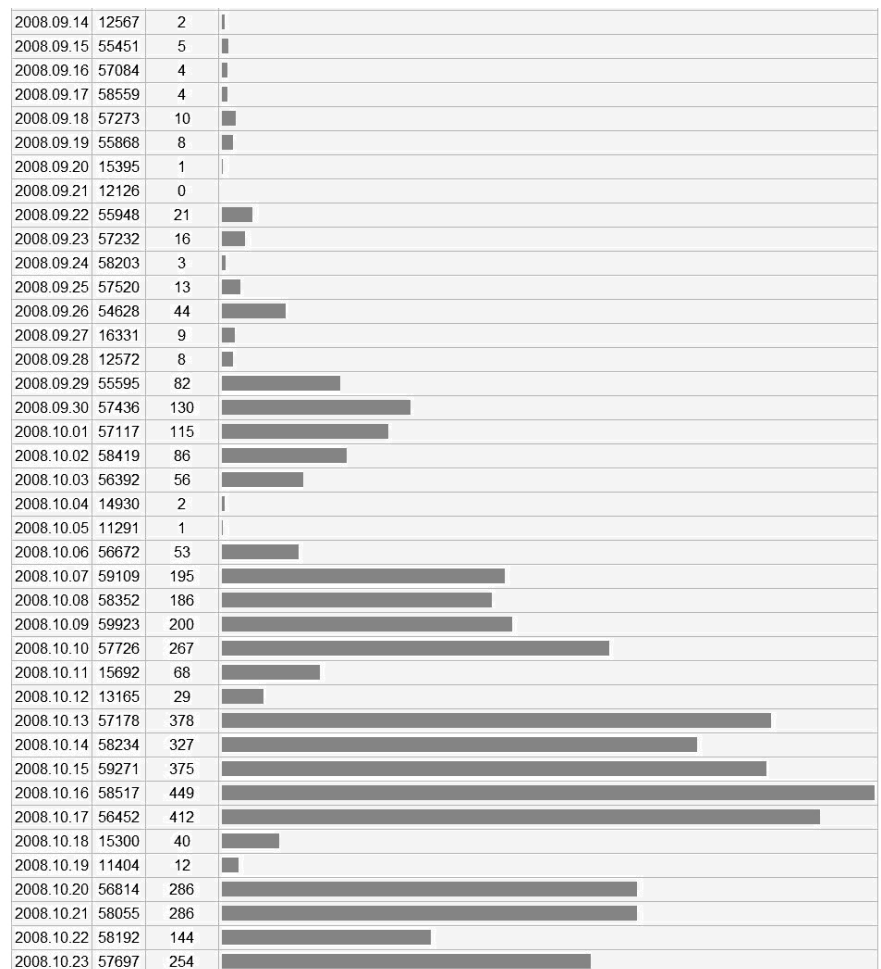


Рис. 75. Детальная диаграмма интенсивности публикаций по теме «Оранта»

В 12:31 на сайте «Экономические новости» появляется странное «обновленное» сообщение с парадоксальным последним предложением (рис. 77).

Далее руководство НАСК «Оранта» опубликовало в Интернете первые опровержения, не спеша обвинять конкурентов в происшед-

Рассмотренные практические примеры позволили выработать некоторую общую методiku проведения оборонительной информационной операции с использованием системы контент-мониторинга веб-ресурсов. Допустим, объектом агрессивной информационной операции является компания «АБВ». Предлагается такие 12 шагов противодействия:

1. Сбор информации с публикациями в «чужих» (не имеющих отношения к «АБВ», неафелированных) СМИ о компании.

2. Построение графика — динамики появления сообщений о компании «АБВ» в сетевых СМИ.

3. Анализ динамики с ретроспективой в 6–12 месяцев с помощью методов анализа временных рядов. После этого анализируется контент публикаций в пороговых точках, определяются моменты, длительность, периодичность воздействия, привязка моментов воздействия к другим событиям из области интереса объекта.

4. Определение источников, публикующих наибольшее количество негатива (публикаций с отрицательной тональностью) о компании «АБВ».

5. Определение «первоисточников» публикаций в СМИ — тех источников, которые первыми опубликовали негативную информацию.

6. Определение вероятных «заказчиков» — владельцев или лиц, влияющих на издательскую политику отдельных СМИ.

7. Определение сфер общих интересов компании «АБВ» и потенциальных «заказчиков» (путем выявления общих информационных характеристик — пересечений «информационных портретов» системы InfoStream, строящихся для объекта и «заказчика»), ранжирование потенциальных «заказчиков» по их интересам.

8. Определение критериев информационных воздействий на основе самых рейтинговых интересов.

9. Моделирование информационных воздействий, для чего находят связи «заказчика» — наиболее связанные с ним персоны и организации, анализируется динамика воздействия со стороны заказчика и строится прогноз этой динамики, анализируется контент публикаций в пороговых точках кривой динамики — определяются критичные точки воздействия.

10. Прогнозируются дальнейшие шаги воздействия путем анализа аналогичной динамики публикаций для других компаний в ретроспективной базе данных системы InfoStream.

11. С учетом реалий и публикаций из ретроспективной базы данных оцениваются вероятные последствия.

12. Организуется информационное (и не только) противодействие. Примеры публикаций в контексте противодействия находятся в ретроспективной базе данных.

Именно системы контент-мониторинга лучше всего подходят для оперативного анализа информационной обстановки по трем причинам. Во-первых, они обеспечивают оперативность, которую не могут обеспечить поисковые системы (время индексации сетевого контента даже лучшими из них составляет от нескольких суток до нескольких недель); во-вторых, специализированные системы контент-мониторинга обеспечивают полноту как в плане источников, так и представления материалов источников, в то время как обычные агрегаторы новостей не всегда обеспечивают необходимую полноту; и, в-третьих, системы контент-мониторинга содержат необходимые аналитические средства, которые могут предоставить пользователю информацию об интенсивности публикаций по заданной тематике в необходимый период времени.

В плане профилактики информационных операций следует внимательно следить за динамикой публикаций о целевой компании, если есть возможность, с учетом тональности этих публикаций, пользоваться доступными аналитическими средствами, например, вейвлет-анализом. При этом следует ориентироваться на возможные модели информационных атак, например, если эта модель охватывает фазы: «фоновые публикации» — «затишье» — «артподготовка» — «затишье» — «атака» (рис. 79), то уже по первым трем компонентам можно с большой вероятностью предсказать грядущие события.

Приведенный выше план, очевидно, является идеальным, ориентированным исключительно на данные контент-мониторинга веб-ресурсов.

Вместе с тем, найденная авторами закономерность нашла свое продолжение и математическое обоснование [131]. Приведенные выше фазы информационных атак, отображаемые в виде локальных и глобальных максимумов на временном ряду, позволили построить шаблон, к которому были применены полиномы Кунченко [132].

В приведенной модели экстремумы соответствуют фазам «попытки» и «атаки». При этом второй максимум будет больше первого. Разница между этими максимумами может быть очень существенной,

фазе «атаки» будет соответствовать интервал с наиболее высокой плотностью публикаций сообщений на протяжении дня.



Рис. 79. Типовое поведение рядов интенсивности тематических публикаций

Выделенные образцы поведения рядов интенсивностей тематических публикаций могут рассматриваться как шаблоны (образцы) функциональной зависимости. Эти шаблоны можно взять в качестве единого базисного элемента некоторого линейного пространства, т.е. в качестве порождающего элемента e для моделирования с помощью полиномов Кунченко.

Тогда как линейную комбинацию линейно-независимых преобразований $f_1(e), f_2(e), \dots, f_n(e)$ соответствующего порождающего элемента можно построить полином P_n приближения n -го порядка к части выходного сигнала $f_s(e)$:

$$P_n = \sum_{\substack{k=0, \\ k \neq s}}^n c_k f_k(e),$$

где коэффициенты c_k определяются из условия обеспечения минимума расстояния между строящимся полиномом и сигналом. Как показано в [133], при этом

$$c_0 = \frac{\langle f_s(e), f_0(e) \rangle - \sum_{\substack{k=1, \\ k \neq s}}^n c_k \langle f_k(e), f_0(e) \rangle}{\langle f_0(e), f_0(e) \rangle},$$

а другие коэффициенты c_k определяются как решение системы линейных уравнений:

$$\sum_{\substack{k=1, \\ k \neq s}}^n c_k F_{i,k} = F_{i,s}, \quad i = 1, \dots, n, \quad i \neq s,$$

где центрированные коррелянты $F_{i,k}$ также рассчитываются с помощью соответствующих преобразований:

$$F_{i,k} = \langle f_i(e), f_k(e) \rangle - \frac{\langle f_i(e), f_0(e) \rangle \cdot \langle f_k(e), f_0(e) \rangle}{\langle f_0(e), f_0(e) \rangle}.$$

Числовой характеристикой, которую можно использовать в критериях качества сопоставления сигнала с выделенным шаблоном, т. е. как меру приближения полинома Кунченко P_n к сигналу $f_s(e)$, можно считать коэффициент эффективности d_n :

$$d_n = \frac{\sum_{\substack{k=1, \\ k \neq s}}^n c_k \langle f_k(e), f_s(e) \rangle}{\langle f_s(e), f_s(e) \rangle}.$$

Рассмотренный метод распознавания определенных образцов с помощью построения пространства с порождающим элементом и поиска коэффициентов соответствующего полинома Кунченко может быть использован в любой проблемной области, в которой можно априори во временном ряду выделить определенные характерные шаблоны.

Таким образом, построив типовые модели поведения рядов интенсивности тематических публикаций во время проведения информационных операций и сопоставив шаблоны, полученные на их основе, можно использовать метод на основе полиномов Кунченко для определения (и предупреждения) возможной информационной атаки. В работе [131] также было обосновано перспективность использования указанного подхода при исследовании моделей в области статистики и социологии.

Естественно, на практике ориентация лишь на единственный тип источников и математических моделей может привести к дефициту информации, необходимой для принятия решений, неточностям, а порой — к дезинформированности. Лишь применение комплексных систем, базирующихся на использовании многочисленных источников, баз данных, математических моделей, наряду с приведенными выше возможностями систем контент-мониторинга может гарантировать эффективную информационную поддержку при противодействии информационным операциям.

ЗАКЛЮЧЕНИЕ

Понятие живучести информационной системы подразумевает ее способность своевременно выполнять свои функции (информирования, влияния, воздействия) в условиях действия дестабилизирующих факторов. В случае информационных систем такими факторами могут выступать устранение отдельных элементов из информационного пространства, потеря их актуальности, доступности документов. Привлечение внимания аудитории к другим темам, порождение и развитие новых информационных сюжетов также может снизить актуальность текущего информационного воздействия. Вместе с тем, с практической точки зрения, происхождение деструктивного информационного процесса играет значительно меньшую роль, чем его последствия.

Информационные системы могут быть как целенаправленными, так и нецеленаправленными. Именно целенаправленные системы могут порождаться в ходе активных рекламных кампаний в качестве отвлекающих факторов, информирования и других элементов информационных операций [3]. При этом живучесть, проявляющаяся как способность целенаправленных информационных систем выполнять свои основные информационные функции на заданном интервале времени без отказов, определяет порог устойчивости, за которым без восстановления компонент и функций информационная система может потерять актуальность и возможность информационного воздействия. Вследствие этого и многих других факторов живучесть информационных систем имеет важнейшее значение для информационной безопасности.

ГЛОССАРИЙ

Адаптивность — свойство системы приспосабливаться к изменяемым условиям внутренней и внешней среды путем использования различных механизмов приспособления.

Ассортативность (от англ. — *assortative mating*) — в теории сложных сетей термин, обозначающий преимущественное соединение узлов сети с большой степенью связности друг с другом (так называемый «клуб богатых»).

Аттрактор (от англ. *attract* — *привлекать, притягивать*) — компактное подмножество фазового пространства динамической системы, все траектории из некоторой окрестности которого стремятся к нему при времени, стремящемся к бесконечности. Так, наиболее простыми вариантами аттрактора являются притягивающая неподвижная точка (к примеру, в задаче о маятнике с трением о воздух) и периодическая траектория (пример — самовозбуждающиеся колебания в контуре с положительной обратной связью).

Вейвлеты (от англ. — *Wavelet*) — математические функции, позволяющие анализировать разные частотные компоненты данных. Вейвлеты локальные во времени и по частоте, все функции одного семейства вейвлетов выводятся из одной с помощью ее сдвигов и растяжений по оси времени. Все вейвлет-преобразования любой функции рассматривают ее в терминах колебаний, локализованных по времени и частоте.

Детерминированный хаос — явление, при котором поведение нелинейной системы выглядит случайным, несмотря на то, что она определяется детерминистическими законами. Причиной появления детерминированного хаоса является неустойчивость системы относительно начальных условий и параметров: изменение начального условия приводит к существенным изменениям динамики системы.

Живучесть — способность системы адаптироваться к новым изменившимся и, как правило, непредвиденным ситуациям, противостоять вредным воздействиям, выполняя при этом свою целевую функцию, за счет соответствующего изменения структуры и поведения системы. Свойство живучести позволяет сложной системе сохраняться целостной в экстремальных для нее условиях, приспособиться к ним, изменяя поведение, структуру, зачастую и цель функционирования. В зависимости от класса систем, их сложности, степени орга-

низованности, а также от выбранного уровня анализа свойство живучести может оцениваться как устойчивость, надежность, адаптивность, отказоустойчивость.

Имитационное моделирование — метод исследования, при котором исследуемая система замещается моделью, которая с достаточной точностью описывает реальную систему. С этой моделью осуществляются эксперименты с целью получения информации о реальной системе. Имитационное моделирование — частный случай математического моделирования. Существует класс объектов, для которых по разным причинам не разработаны аналитические модели, или не разработаны методы решения относительно полученной модели. В этом случае математическая модель замещается имитатором или имитационной моделью — логико-математическим описанием объекта.

Индивидуум-ориентированное моделирование (*от англ. — Individual-based Modeling*) — один из методов компьютерного моделирования сложных систем — популяций. Индивидуум-ориентированная модель описывает механизмы взаимодействия индивидуумов (программных агентов) с окружающей средой и внутрипопуляционные взаимодействия на уровне отдельных программных агентов.

Информационное воздействие — воздействие на массовое сознание аналогичное тому, как психологическое воздействие влияет на сознание индивидуальное. Информационное воздействие как процесс управления — это возбуждение (торможение) в управляемой системе таких процессов, которые стимулируют желательный для воздействующей стороны выбор. Этот способ воздействия на противника не предполагает прямого выведения из строя части элементов его системы, но представляет собой передачу противнику такой информации, которая натолкнет его на выбор необходимого для воздействующей стороны решения.

Информационная война (*от англ. — Information War*) — комплекс мероприятий по информационному воздействию на массовое сознание для изменения поведения людей и навязывания им целей, которые не входят в число их интересов, а также защита от подобных воздействий. Целенаправленные действия, предпринятые для достижения информационного превосходства путем нанесения ущерба информации, информационным процессам и информационным системам противника при одновременной защите собственной информации, информационных процессов и информационных систем [133].

Информационные операции (от англ. — *Information Operations, Info Ops, IO*) — информационное воздействие на массовое сознание (как на враждебное, так и на дружеское), воздействие на информацию, доступную противнику и необходимую ему для принятия решений, а также на информационно-аналитические системы противника, в том числе действия, направленные на физическое поражение информационно-аналитических систем, вывод из строя средств компьютерно-телекоммуникационной инфраструктуры.

Информационное пространство (от англ. — *Information Space*) — 1) совокупность информационных ресурсов, технологий их сопровождения и использования, информационных и телекоммуникационных систем, которые образуют информационную инфраструктуру; 2) множество связанных по смыслу элементов (документов), образующих информационные системы — кластеры близких по тематике документов.

Информационные системы — в контексте данной работы — документальные или контентные системы (от англ. — *content — содержание*) — совокупность содержательно связанных в сеть элементов информационного пространства. Как частные случаи информационных систем можно рассматривать, например, тематические подборки документов, веб-сайты, посвященные некоторой проблематике, или информационные кластеры (сюжеты) — массивы информационных сообщений, опубликованных на различных веб-сайтах, посвященных одной тематике или одному событию.

Информационное управление — процесс выработки и применения управляющих воздействий, которые носят неявный, косвенный, информационный характер. Объекту управления предоставляется определенная информационная картина, ориентируясь на которую, он выбирает линию своего поведения, как ему представляется, самостоятельно, т.е. информационное управление — это способ воздействия, побуждающий людей к упорядоченному поведению, выполнению требуемых от них действий.

Искусственное общество — компьютерная модель общества, состоящего из агентов (программных аналогов индивидуумов), действующих по определенным правилам, в том числе взаимодействующих между собой.

Клеточный автомат — набор клеток, образующих некоторую периодическую решетку с заданными правилами перехода, определяю-

щими состояние клетки в следующий момент времени через состояние клеток, находящимися от нее на расстоянии не больше некоторого, в текущий момент времени. Как правило, рассматриваются автоматы, где состояние определяется самой клеткой и ближайшими соседями.

Контент (*от* англ. — *Content*) — содержательная часть информационных ресурсов; любое содержательное наполнение информационных ресурсов (например, веб-сайтов) — тексты, графика, мультимедиа.

Контент-анализ — метод получения выводов путем анализа содержания текстовой информации. Чаще всего реализуется как систематическая числовая обработка, оценка и интерпретация формы и содержания информационного источника.

Контент-мониторинг — систематическое, непрерывное во времени сканирование и контент-анализ информационных ресурсов.

Коэффициент кластерности — величина, которая соответствует уровню связности узлов в сети. Эта величина показывает, сколько ближайших соседей данного узла являются ближайшими соседями друг для друга и равна отношению реального количества ребер, которые соединяют ближайших соседей данного узла, к максимально возможному.

Латентно-семантический анализ (*от* англ. — *Latent Semantic Analysis, LSA*) — теория и метод для выдержки контекстно-зависимых значений слов с помощью статистической обработки больших наборов текстовых данных. Латентно-семантический анализ основывается на идее, что совокупность всех контекстов, в которых встречается и не встречается терм, задает множественное число ограничений, которые в значительной степени позволяют определить сходство смысловых значений термов между собой. Как начальную информацию LSA использует матрицу «термы/документы», которая содержит весовые значения термов в документах.

Логистическое уравнение — уравнение, которое изначально появилось при рассмотрении модели роста численности населения. Исходные предположения для вывода уравнения при рассмотрении популяционной динамики выглядят следующим образом: скорость размножения популяции пропорциональна ее текущей численности при прочих равных условиях; скорость размножения популяции пропорциональна количеству доступных ресурсов при прочих равных ус-

ловиях. Второй член уравнения отражает конкуренцию за ресурсы, которая ограничивает рост популяции. Точным решением логистического уравнения является логистическая функция, S-образная кривая.

Математическое моделирование — процесс построения и изучения математических моделей — математических представлений о реальности.

Многоагентное моделирование (*от англ. — Multiagent Modeling или Agent-based Modeling*) — моделирование на базе (основе) применения агентов. Компьютерные модели, в которых атомарными элементами являются агенты.

Моделирование — исследование объектов познания на их моделях; построение и изучение моделей реально существующих объектов, процессов или явлений с целью получения объяснений этих явлений, а также для предсказания явлений, которые интересуют исследователя. Моделирование можно рассматривать как процесс создания, применения, использования модели. Главные функции модели — упрощение получения информации о свойствах объекта, передача информации и знаний, управление объектами и процессами, их оптимизация, прогнозирование, диагностика.

Надежность — комплексное свойство системы, заключающееся в ее способности выполнять (в определенных условиях функционирования) заданные функции, сохраняя свои основные характеристики в определенных границах. Самые распространенные из показателей надежности, которые обычно вероятностны по своему характеру, — это вероятности безотказной работы, средняя наработка на отказ, готовность и т.д.

Обратная связь — воздействие результатов функционирования системы на характер этого функционирования.

Отказоустойчивость — свойство системы сохранять работоспособность в случае отказ одной или нескольких компонент.

Поведение системы — действие системы во времени.

Помехоустойчивость — свойство систем противостоять действию помех (возмущающих воздействий).

Природные вычисления (*от англ. — Natural Computing*) — направление в рамках концепции многоагентного моделирования, объединяющее математические методы, в которых заложены принципы принятия решений, аналогичные реализуемым в природе механизмам. Например, имитация самоорганизации муравьиной колонии (или

колонии термитов) составляет основу так называемых муравьиных алгоритмов оптимизации, одного из перспективных методов природных вычислений.

Психологическое воздействие — воздействие на психику отдельных людей, которое может осуществляться различными средствами: информационными, военными, экономическими, политическими. Психологическое воздействие подразделяется на следующие виды: информационно-психологическое, психогенное, психоаналитическое, нейролингвистическое, психотронное, психотропное.

Редукционизм (от лат. — *Reductio*) — методологический принцип, согласно которому сложные явления могут быть полностью объяснены с помощью законов, свойственных явлениям более простым (например, социологические явления объясняются биологическими или экономическими законами).

Синергетика (от греч. συν — совместно и εργος — действующий) — междисциплинарное направление научных исследований, задачей которого является изучение природных явлений и процессов на основе принципов самоорганизации систем (состоящих из подсистем); «...наука, занимающаяся изучением процессов самоорганизации и возникновения, поддержания, устойчивости и распада структур самой различной природы...».

Система (от др.-греч. σύστημα — сочетание) — множество взаимосвязанных элементов, обособленное от среды и взаимодействующее с ней как целое.

Система клеточных автоматов — совокупность математических объектов, которые представляют собой однородную сетку, каждая клетка которой (клеточный автомат) может находиться в одном из возможных состояний. Состояния клеток синхронно обновляются на каждом шаге моделирования в соответствии с определенными правилами перехода; в общем случае таких правил может быть бесчисленное количество, которое соответствует количеству подмножеств счетного множества.

Системный эффект — несводимость системы к сумме свойств ее компонент.

Скейлинг (от англ. *Scaling*) — масштабная инвариантность, самоподобие. Это свойство применяется, в частности, для представления функции двух переменных как функции одной.

Скейлограмма — карта коэффициентов преобразования.

Скелетон — линии локальных экстремумов поверхности скейлограммы.

Сложная система — система, состоящая из множества взаимодействующих составляющих (подсистем), вследствие чего сложная система приобретает новые свойства, которые отсутствуют на подсистемном уровне и не могут быть сведены к свойствам подсистемного уровня.

Сложная сеть (*от англ. — Complex Network*) — сеть (граф) с нетривиальными топологическими особенностями, не свойственными простым сетям, таким как решетки или случайные графы. Изучение сложных сетей — область научного исследования, которая появилась в результате исследования сетей реального мира, таких как компьютерные и социальные сети.

Структура (*от лат. — structura — строение*) — относительно устойчивая фиксация связей между элементами системы. В естествознании структура — внутреннее устройство чего-либо, скрытое внешней формой предмета. Структура связана с упорядоченностью отношений, связывающих элементы системы. Структуры бывают простые, сложные, иерархические.

Текстовый корпус (*от англ. — Text corpus*) — массив текстов, собранных в соответствии с определенными принципами, размеченными по определенному стандарту и обеспеченными специализированной поисковой системой. В некоторых случаях текстовым корпусом первого порядка называют произвольные массивы, объединенных каким-то общим признаком. Разработкой, созданием и использованием текстовых (лингвистических) корпусов занимается специальный раздел языкознания — корпусная лингвистика.

Тематический информационный поток — последовательность сообщений, соответствующих определенному тематическому запросу.

Терм (*от англ. — Term*) — слово или постоянное словосочетание. В математической логике понятия «терм» широко используется в качестве «символьного выражения».

Точка бифуркации — критическое состояние системы, при котором система становится неустойчивой относительно флуктуаций и возникает неопределенность: станет ли состояние системы хаотическим или она перейдет на новый, более дифференцированный и высокий уровень упорядоченности.

Устойчивость системы — способность системы возвращаться в начальное состояние после окончания влияния, которое вывело систему из этого состояния. Активное сохранение системой определенных характеристик безотносительно к тому, играют ли они какую-либо роль в общей системе.

Уязвимость — параметр, характеризующий возможность нанесения описываемой системе повреждений любой природы теми или иными внешними средствами или факторами.

Фазовая плоскость — координатная плоскость, в которой по осям координат откладываются какие-либо две переменные (фазовые координаты), однозначно определяющие состояние системы второго порядка. Фазовая плоскость является частным случаем фазового пространства, которое может иметь большую размерность. В физике колебаний на оси абсцисс фазовой плоскости откладывается значения параметра x , а на оси ординат — первая производная x по времени.

Фазовое пространство — множество всех состояний системы в фиксированный момент времени. Каждому возможному состоянию системы соответствует точка фазового пространства. Сущность понятия фазового пространства заключается в том, что состояние сколь угодно сложной системы представляется в нем одной единственной точкой, а эволюция этой системы — перемещением этой точки.

Фрактал (*от лат. Fractus — дробимый, состоящий из фрагментов*) — бесконечно самоподобный (точно или приблизительно) объект (множественное число), каждая часть которого повторяется при уменьшении масштаба. Другое определение — геометрический объект, размерность Хаусдорфа–Безиковича которого нецелая. Возможно и такое определение: фрактал — самоподобное множество нецелой размерности.

Фрактальный анализ — метод моделирования с помощью теории фракталов, который заключается в исследовании фрактальной размерности и других фрактальных свойств сигналов, наборов данных, объектов.

Целенаправленное поведение системы — стремление системы достичь цели функционирования.

Целостность — относительная независимость системы от среды и других аналогичных систем.

Целостность информации — термин в информатике и теории телекоммуникаций, который означает, что данные полны; условие то-

го, что данные не были изменены при выполнении любой операции над ними, будь то передача, хранение или представление.

Цель системы — предпочтительное для системы состояние, т.е. некоторое конечное состояние, к которому стремится система в силу своей структурной организации.

Эволюция системы — изменение структуры системы во времени.

Элемент (*от лат. — elementum — стихия*) — неразложимая далее (в данной системе при данном способе рассмотрения и анализа) компонента.

Эмерджентность (*от англ. Emergence — возникновение, появление нового*) в теории систем — наличие у какой-либо системы особых свойств, не присущих ее подсистемам и блокам, а также сумме элементов, не связанных особыми системообразующими связями; несводимость свойств системы к сумме свойств ее компонент; синоним — «системный эффект».

СПИСОК ЛІТЕРАТУРИ

1. *Shelton C.* A Framework for Scalable Analysis and Design of System-Wide Graceful Degradation in Distributed Embedded Systems / С. Shelton, P. Koopman, W. Nace // Eighth IEEE Int. Workshop on Object-Oriented Real-Time Dependable Systems (WORDS 2003). — Guadalajara (Mexico), Jan. 2003. — 8 p.
2. *Додонов А.Г.* Введение в теорию живучести вычислительных систем / А.Г. Додонов, М.Г. Кузнецова, Е.С. Горбачик. — К.: Наук. думка, 1990. — 184 с.
3. *Горбулін В.П.* Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія / В.П. Горбулін, О.Г. Додонов, Д.В. Ланде. — К.: Інтертехнологія, 2009. — 164 с.
4. *Broder A., Kumar R., Maghoul F. et al.* Graph structure in the Web / A. Broder, R. Kumar, F. Maghoul et al. // Proc. 9th Int. World Wide Web Conf. on Computer Networks: The International Journal of Computer and Telecommunications Networking. — Amsterdam, 2000. — P. 309–320.
5. *Додонов А.Г.* Живучість інформаційних сюжетів як динамічних документальних систем / А.Г. Додонов, Д.В. Ланде // Реєстрація, зберігання і оброб. даних. — 2010. — Т. 13, № 2. — С. 88–102.
6. *Додонов О.Г.* Живучість складних систем: аналіз та моделювання: навч. посіб. у 2-х ч. / О.Г. Додонов, М.Г. Кузнецова, О.С. Горбачик. — К.: НТУУ «КПІ», 2009. — 264 с.
7. Modeling the revolution in military affairs, autumn // JFQ. — Winter 1998–99.
8. *Иванов С.А.* Мировая система научной коммуникации как информационное пространство // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: 8-я Междунар. конф. «Крым 2002»: Материалы конф., Судак, 9–17 июня, 2001 г. — М., 2001. — Т. 1. — С. 1123–1126.
9. *Langton C.* Artificial Life: ser. SFI Studies in the Sciences of Complexity. — Red-Wood City, CA: Addison-Wesley, 1989. — P. 1–47.
10. *Epstein J.M.* Remarks on the foundations of agent-based generative social science // SFI (Santa Fe Institute) Working Paper. DOI: SFI-WP 05-06-024, SFI Working Papers, 2005.

11. *Карнов Ю.Г.* Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. — СПб.: БХВ-Петербург, 2006. — 400 с.
12. *Карнов Ю.Г.* Изучение современных парадигм имитационного моделирования в среде AnyLogic // Компьютерные инструменты в образовании. — 2005. — № 4. — С. 3–14.
13. *Борщев А.В.* Практическое агентное моделирование и его место в арсенале аналитика // Exponenta Pro. — 2004. — № 3–4.
14. *Штовба С.Д.* Муравьиные алгоритмы // Exponenta Pro. Математика в приложениях. — 2003. — № 4. — С. 70–75.
15. *Axelrod R.* Modeling Security Issues of Central Asia // R. Gerald Ford School of Public Policy University of Michigan (Project on «Security in Central Asia». — June 2004. — US Govt. Contract # 2003*H513400*000).
16. *Lustick I.* PS-I: A user-friendly agent-based modeling platform for testing theories of political identity and political stability // J. Artificial Societies and Social Simulations. — 2002. — Iss. 5, N 3.
17. *Forrester J.W.* System dynamics and the lessons of 35 years // A Systems-Based approach to Policymaking / Ed. by U.B. De Green U.B. — Boston: Kluwer, 1995. — P. 199–239.
18. *Grimm V.* Ten years of individual-based modeling in ecology: what have we learned and what could we learn in the future // Ecological Modelling. — 2002. — **115**, N 2–3. — P. 129–148.
19. *Epstein J. M., Axtell R.* Artificial societies and generative social science // Artificial Life and Robotics. — 1997. — N 1. — P. 33–34.
20. *Нейман Дж.* Теория самовоспроизводящихся автоматов. — М.: Мир, 1971. — 382 с.
21. *Theory and applications of cellular automats* / Ed. By S.A. Wolfram. — Singapore: World Scientific, 1986.
22. *Ландэ Д.В.* Моделирование электоральных процессов на основе концепции клеточных автоматов / Д.В. Ландэ, В.Н. Фурашев // Открытые информационные и компьютерные технологии. Харьков: НАКУ «ХАИ», 2007. — Вып. 36 — С. 17–34.
23. *Плотинский Ю.М.* Модели социальных процессов. — Изд. 2-е. — М.: Логос, 2001. — 296 с.
24. *Плотинский Ю.М.* Математическое моделирование динамики социальных процессов. — М.: Изд-во МГУ, 1992. — 133 с.
25. *Гарднер М.* Математические досуги. — М.: Мир, 2000. — 448 с.

26. *Тоффоли Т., Марголюс Н.* Машины клеточных автоматов. — М.: Мир, 1991. — 280 с.
27. *Brown T.A.* Nonlinear politics // *Chaos Theory in the Social Sciences* / Eds. L.D. Kiel, E. Elliot. Ann Arbor. — The Univ. of Michigan Press, 1996. — P. 119–137.
28. *Casti J.L.* Searching for Certainty. New York: W.Morrow, 1990.
29. *Schelling T.S.* Models of segregation. *American Economic Review* // *Pap. and Proc.* — 1969. — **59**. — P. 488–493.
30. *Bhargava S.C., Kumar A., Mukherjee A.* A stochastic cellular automata model of innovation diffusion // *Technol. Forecast. and Soc. Change.* — 1993. — **44**, N 1. — P. 87–97.
31. *Tiebout Ch. M.* A Pure theory of local expenditures // *J. Political Economy.* — 1956. — **64**, N 5. — P. 416–424.
32. *Ландэ Д.В.* Модель диффузии информации // Информационные технологии и безопасность. Менеджмент информационной безопасности: Сб. науч. тр. Ин-та проблем регистрации информации НАН Украины. — 2007. — Вып. 10. — С. 51–67.
33. *Рябинин И.А., Черкесов Г.Н.* Логико-вероятностные методы исследования надежности структурно-сложных систем — М.: Радио и связь, 1984. — 238 с.
34. *Райнике К.* Модели надежности и чувствительности систем. — М.: Мир, 1979. — 454 с.
35. *Райнике К., Ушаков И.А.* Оценка надежности систем с использованием графов. — М.: Радио и связь, 1988. — 208 с.
36. *Додонов А.Г., Флейтман Д.В.* К вопросу безопасности информационных систем // *Інформаційні технології та безпека: Зб. наук. праць.* — К., 2004. — Вип. 6. — С. 26–29.
37. *Аверьянов В.Т.* Выбор и исследование свойств показателя уровня живучести системы пожаротушения // *Проблемы управления рисками в техносфере.* — 2009. — Т. 9–10. — № 1–2. — С. 55–62.
38. *Эшби У.Р.* Принципы самоорганизации. — М.: Мир, 1966. — 622 с.
39. *Петрушенко Л.А.* Единство системности, организованности и самодвижения. — М.: Мысль, 1975. — 286 с.
40. *Петров Ю.П.* Информация и энтропия в кибернетике. — Л.: ЛГУ, 1989. — 84 с.
41. *Шамбадаль П.* Развитие и приложения понятия энтропии. — М.: Наука, 1967. — 280 с.

42. *Бестужев-Лада И.В.* Рабочая книга по прогнозированию. — М.: Мысль, 1982. — 430 с.
43. *Survivable network systems: an emerging discipline* CMU/SEI-97-TR-013 ESC-TR-97-013. — Software Engineering Institute Carnegie Mellon University Pittsburgh. — May 1999.
44. *Newman M.E.J.* The structure and function of complex networks // *SIAM Rev.* — 2003. — **45**. — P. 167–256.
45. *Dorogovtsev S.N., Mendes J.F.F.* Evolution of networks: from biological networks to the Internet and WWW. — Oxford University Press, 2003.
46. *Erdős, P., Renyi A.* On Random Graphs. I // *Publ. Math.* — 1959. — **6**. — P. 290–297.
47. *Erdős P., Renyi A.* On the evolution of random graphs // *Publ. Math. Inst. Hungar. Acad.* — 1960. — *Sci. 5*. — P. 17–61.
48. *Watts D.J., Strogatz S.H.* Collective dynamics of «small-world» networks // *Nature*. — 1998. — **393**. — P. 440–442.
49. *Albert R., Jeong H., Barabasi A.* Attack and error tolerance of complex networks // *Nature*. — 2000. — **406**. — P. 378–382.
50. *Gol'dshtein V., Koganov G.A., Surdutovich G.I.* Vulnerability and hierarchy of complex networks // *Phys. Rev. Lett.* — 2004.
51. *Zhou S., Mondragon R.J.* The rich-club phenomenon in the internet topology // *Commun. Lett. IEEE*. — 2004. — **8**. — P. 180–182.
52. *Boguna M., Pastor-Satorras R., Vespignani A.* Statistical mechanics of complex networks / *Lecture and notes in physics*. — Springer Berlin. — 2003. — P. 127–147.
53. *Newman M.E.J.* Assortative mixing in networks // *Phys. Rev. Lett.* — 2002. — **89** (208701).
54. *Barabasi A., Albert R.* Emergence of scaling in random networks // *Science*. — 1997. — **286**. — P. 509–512.
55. *Bjorneborn L., Ingwersen P.* Toward a basic framework for webometrics // *J. Amer. Soc. Inform. Sci. and Technol.* — 2004. — **55**(14). — P. 1216–1227.
56. *Milgram S.* The small world problem // *Psychology Today*. — 1967. — **2**. — P. 60–67.
57. *Watts D.J., Strogatz S.H.* Collective dynamics of «small-world» networks // *Nature*. — 1998. — **393**. — P. 440–442.
58. *Newman M.E.J., Watts D.J.* Scaling and percolation in the small-world network model // *Phys. Rev. E*. — 1999. — **7332**.

59. *Снарский А.А.* Процессы переноса в макроскопических неупорядоченных средах: От теории среднего поля до перколяции. / А.А. Снарский, И.В. Безсуднов, В.А. Севрюков. — М.: УРСС, Изд-во ЛКИ, 2007. — 304 с.
60. *Тарасевич Ю.Ю.* Перколяция: теория, приложения, алгоритмы. — М., УРСС, 2002. — 112 с.
61. *Broadbent S.R., Hammersley J.M.* Percolation processes: I. Crystals and Mazes // Proc. Cambridge Phil. Soc. — 1957. — P. 629–641.
62. *Cohen R., Erez K., ben-Avraham D., Havlin S.* Resilience of the Internet to Random Breakdown // Phys. Rev. Lett. — 2000. — **85**. — P. 4626.
63. *Snarskii A.* FreeBSD Stack Integrity Patch. 1997. URL: <ftp://ftp.lucky.net/pub/unix/local/libc-letter>
64. *The Deep Web: Surfacing Hidden Value*, 2000 BrightPlanet.com LLC, 35 p. URL: <http://www.dad.be/library/pdf/BrightPlanet.pdf>
65. *Брайчевский С.М., Ландэ Д.В.* Современные информационные потоки: актуальная проблематика // Научно-техническая информация. Сер. 1. — 2005. — Вып. 11. — С. 21–33.
66. *Додонов О.Г.* Інформаційні потоки в глобальних комп'ютерних мережах / О.Г. Додонов, Д.В. Ланде, В.Г. Путятін. — К.: Наук. думка, 2009. — 295 с.
67. *Ланде Д.В., Фурашев В.М., Григор'єв О.М.* Програмно-апаратний комплекс інформаційної підтримки прийняття рішень: Науково-методичний посібник. — К.: Інжиніринг, 2006. — 48 с.
68. *Ландэ Д.В.* Структура новостного Web-пространства // Научно-техническая информация. Сер. 2. — 2006. — № 8. — С. 17–20.
69. *Ландэ Д.В.* Основы интеграции информационных потоков: Монография. — К.: Инжиниринг, 2006. — 240 с.
70. *Григорьев А.Н., Ландэ Д.В.* Адаптивный интерфейс уточнения запросов к системе контент-мониторинга InfoStream // Компьютерная лингвистика и интеллектуальные технологии: Труды междунар. конф. «Диалог-2005» (Звенигород, 1–6 июня, 2005 г.). — М.: Наука, 2005. — С. 109–111.
71. *Luisi P.L.* Emergence in Chemistry: Chemistry as the Embodiment of Emergence // Found. Chem. — 2002. — **1**. — P. 183–200.
72. *Feigenbaum M.J.* Quantitative universality for a class of nonlinear transformations // J. Stat. Phys. — 1978. — **19**. — P. 25–52.

73. *Потев М.И.* Электронный учебник по дисциплине: «Концепции современного естествознания». — СПбГУ ИТМО, URL: http://de.ifmo.ru/bk_netra/start.php?bn=12
74. *Пригожин И., Стенгерс И.* Порядок из хаоса. — М., 1986. — 432 с.
75. *Арнольд В.А.* Теория катастроф. — М.: Наука, 1990. — 128 с.
76. *Коковин С.Г.* Лекции по теории игр и политологии. Часть 1. Введение в теорию игр. — 66 с. (<http://www.math.nsc.ru/~mathecon/Kokovin/mltigran.pdf>)
77. *Расторгуев С.П.* Философия информационной войны. — М.: Вузовская книга, 2001. — 468 с.
78. *Интриллигатор М.* Математические методы оптимизации и экономическая теория. — М.: Прогресс. — 1975. — 607 с.
79. *Stackelberg H. von.* Marktform und Gleichgewicht. Faksimile der 1934 in Wien und Berlin erschienenen Erstausgabe. Reihe: Klassiker der Nationalökonomie VI, 142 S., Gebunden.
80. *Фурсова П.В., Левич А.П., Алексеев В.Л.* Экстремальные принципы в математической биологии // Успехи современной биологии. — 2003. — Т. 123, № 2. — С. 115–117.
81. *Webb J.N.* Hamilton's variational principle and ecological models // Ecol. Modelling. — 1995. — 80. — P. 35.
82. *Приц А.К.* Принцип стационарных состояний открытых систем и динамика популяций. — Калининград, 1974. — 123 с.
83. *Lurie D., Valls J., Wagensberg J.* Thermodynamic approach to biomass distribution in ecological systems // Bull. Math. Biol. 1983. — 45. — P. 869.
84. *Левич А.П.* Структура экологических сообществ. — М.: Изд-во Моск. ун-та, 1980. — 181 с.
85. *Свирижев Ю.М.* Феноменологическая термодинамика взаимодействующих популяций // Журн. общ. биологии. — 1991. — Т. 52, № 6. — С. 840.
86. *Ханин М.А.* Энергетика и критерии оптимальности онтогенетических процессов. Математическая биология развития. — М.: Наука, 1982. — 177 с.
87. *Ханин М.А., Дорфман Н.Л., Кухаров И.Б., Левадный В.Г.* Экстремальные принципы в биологии и физиологии. — М.: Наука, 1978. — 256 с.

88. *Евдокимов Е.В.* Проблемы регулярного поведения и детерминированного хаоса в основных моделях популяционной динамики (Теория и эксперимент) // Автореф. дис. на соискание уч. степени доктора биологических наук. — Красноярск, 1999. — 40 с.
89. *Малков С.Ю.* Математическое моделирование исторической динамики: подходы и модели. — М.: РГСУ, 2004. — 188 с.
90. *Жаботинский А.М.* Концентрационные колебания. — М.: Наука, 1974. — 179 с.
91. *Вольтера В.* Математическая теория борьбы за существование. — М.: Наука, 1976. — 288 с.
92. *Гаузе Г.Ф.* Борьба за существование. — М.: УРСС, 2002. — 160 с.
93. *Гаузе Г.Ф.* Экология и некоторые проблемы происхождения видов // Экология и эволюционная теория. — Л., Наука, 1984. — С. 5–108.
94. *Чернавский Д.С., Чернавская Н.М., Малков С.Ю., Малков А.С.* Математическое моделирование геополитических процессов // Стратегическая стабильность. — 2002. — № 1. — С. 60–66.
95. *Ландэ Д.В.* Поиск знаний в Internet. Профессиональная работа. — М.: Диалектика, 2005. — 272 с.
96. *Фурашев В.М., Ланде Д.В.* Інформаційні операції крізь призму системи моніторингу та інтеграції Інтернет-ресурсів // Права інформатика. — 2009. — № 2(22). — С. 49–57.
97. *Хартли Р.* Передача информации. Теория информации и ее приложения. — М.: Физматгиз, 1959.
98. *Чуи К.* Введение в вэйвлеты. — М.: Мир, 2001. — 416 с.
99. *Астафьева Н.М.* Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. — 1996. — Т. 166, № 11. — С. 1145–1170.
100. *Давыдов А.А.* Вейвлет-анализ социальных процессов // Социологические исследования. — 2003. — № 11. — С. 97–103.
101. *Давыдов А.А.* Системная социология. — М.: КомКнига, 2006. — 192 с.
102. *Мандельброт Б.* Фрактальная геометрия природы. — М.: Институт компьютерных исследований, 2002. — 656 с.
103. *Федер Е.* Фракталы. — М.: Мир, 1991. — 254 с.

104. *Peng C.K., Buldyrev S.V., Havlin S., Simons M., Stanley H.E., Goldberger A.L.* Mosaic organization of DNA nucleotides. // *Phys Rev E*. — 1994. — **49** (2). — P. 1685–1689.
105. *Ландэ Д.В., Снарский А.А.* Динамика отклонения элементов ряда измерений от локальных линейных аппроксимаций // Реестрация, зберігання і оброб. даних. — 2009. — Т. 11, № 1. — С. 27–32.
106. *Mandelbrot B.B.* Fractals and multifractals // *Noise, Turbulence and Galaxies, Selecta*. — New York: Springer-Verlag, 1989. — **1**.
107. *Павлов А.Н., Сосновцева О.В., Зиганин А.Р.* Мультифрактальный анализ хаотической динамики взаимодействующих систем // *Изв. вузов, Прикладная нелинейная динамика*. — 2003. — Т. 11, № 2. — С. 39–54.
108. *Снарский А.А., Ландэ Д.В., Брайчевский С.М., Дармохвал А.Т.* Свойства распределения релевантности в документальных массивах // *Труды 9-й Всероссийской научной конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL'2007, Переславль-Залесский, Россия, 2007*. — С. 151–155.
109. *Information operations roadmap*. — DoD US, 30 oct., 2003. — 78 p.
110. *Hill J.M.D., Surdu J.R., Ragsdale D.J., Schafer, J.H.* Anticipatory planning in information operations // *Systems, Man, and Cybernetics: 2000 IEEE Int. Conf.* — 2000. — Vol. 4. — P. 2350–2355.
111. *Burke M.M.* Knowledge Operations: above and beyond Information Operations // *6th Int. Command and Control Research and Technology*. — June 19–21, 2001.
112. *Lasswell H.D.* The structure and function of communication in society // *The Communication of Ideas / Ed. by L. Bryson*. — New York: Harper and Brothers, 1948.
113. *Schramm W., Roberts D.F.* The Process and Effects of Mass Communication. — Univ. of Illinois Press, 1974.
114. *Osgood Ch.E.* Psycholinguistics. A survey of theory and research problems // *Supplement to the Int. J. Amer. Linguistics*. — 1954. — **20**, N 4. — Baltimore: Waverly Press, 1954.
115. *Latane B.* The psychology of social impact // *American Psychologist*. — 1981. — **33**. — P. 343–356.
116. *Latane B., Nowak A.* Causes of polarization and clustering in social groups // *Progr. commun. sci.* — 1997. — **13**. — P. 43–75.

117. *Nowak A., Szamrej J., Latane B.* From private attitude to public opinion: A dynamic theory of social impact // *Psychological Review*. — 1990. — **97**. — P. 367–376.
118. *Lewenstein M., Nowak A., Latane B.* Statistical mechanics of social impact // *Phys. Rev. A*. — 1993. — **45**. — P. 763–776.
119. *Kacperski K., Holyst J.A.* Phase transitions as a persistent feature of groups with leaders in models of opinion formation // *Physica A. Statistical Mechanics and its Applications*. — 2000. — **287**, Iss. 3–4. — P. 631–643.
120. *Sobkowicz P.* Effect of leader's strategy on opinion formation in networked societies. — 2003a. — URL <http://arxiv.org/pdf/condmat/0311566>.
121. *Sobkowicz P.* Opinion formation in networked societies with strong leaders. — 2003b. — URL <http://arxiv.org/pdf/condmat/0311521>.
122. *Кульба В.В., Малюгин В.Д., Шубин А.Н., Вус М.А.* Введение в информационное управление. Учебно-методическое издание. — СПб.: Изд-во С.-Петербургского ун-та. 1999. — 116 с.
123. *Кононов Д.А., Кульба В.В., Шубин А.Н.* Базисные понятия моделирования информационного управления в социальных системах // Тр. междунар. научно-практической конф. «Теория активных систем». — М.: Ин-т проблем управления им. В.А. Трапезникова РАН, 2003. — Т. 2. — С. 125–129.
124. *Кульба В.В., Кононов Д.А., Косяченко С.А., Шубин А.Н.* Методы формирования сценариев развития социально-экономических систем. — М.: СИНТЕГ, 2004. — 296 с.
125. *Чхартишвили А.Г.* Теоретико-игровые модели информационного управления. — М.: ЗАО «ПМСОФТ», 2004. — 227 с.
126. *Bak P.* How nature works: The Science of Self-Organized Criticality. — New York: Springer-Verlag Inc., 1996.
127. *Катица С.П., Курдюмов С.П., Малинецкий Г.Г.* Синергетика и прогнозы будущего. — М.: Наука, 1997. — 288 с.
128. *Van Raan A.F.J.* Fractal geometry of information space as represented by cocitation clustering // *Scientometrics*. — 1991. — **20**, N 3. — P. 439–449.
129. *Иванов С.А.* Стохастические фракталы в Информатике // Научно-техническая информация. Сер. 2. — 2002. — № 8. — С. 7–18.

130. Ландэ Д.В. Фрактальные свойства тематических информационных потоков из Интернет // Реєстрація, зберігання і оброб. даних. — 2006. — Т. 8, № 2. — С. 93–99.

131. Чертов О.Р. Поліноми Кунченка для розпізнавання образів // ВІСНИК НТУУ КПІ. Сер. Інформатика, управління та обчислювальна техніка. — 2009. — № 51. — С. 103–108.

132. Кунченко Ю.П. Приближения в пространстве с порождающим элементом. — К.: Наук. думка, 2003. — 243 с.

133. Почепцов Г.Г. Информационные войны. М.: Рефл-бук, К.: Ваклер, 2000. — 576 с.

Наукове видання

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ РЕЄСТРАЦІЇ ІНФОРМАЦІЇ

ДОДОНОВ Олександр Георгійович
ЛАНДЕ Дмитро Володимирович

ЖИВУЧІСТЬ ІНФОРМАЦІЙНИХ СИСТЕМ

(Російською мовою)

Київ, Науково-виробниче підприємство
«Видавництво «Наукова думка» НАН України», 2011

Комп'ютерна верстка *М.Д. Рассоленко*

Підп. до друку 11.01.2011. Формат 60×84/16.
Папір офс. № 1. Гарн. Таймс. Друк офс.
Обл.-вид. арк. 16,4. Ум. друк. арк. 14,88
Наклад 300 прим. Замовлення №

НВП «Видавництво «Наукова думка» НАН України»
Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру — серія ДК № 2400 від 15.03.2006 р.
01601 Київ 1, вул. Терещенківська, 3

Видруковано у друкарні ТОВ «Інфодрук»
03113, м. Київ, вул. М. Шпака, 2