

УДК 681.3+519.83

А. Г. Додонов¹, Д. В. Ландэ²

Інститут проблем реєстрації інформації НАН України
ул. Н. Шпака, 2, 03113 Київ, Україна
¹dodonov@ipri.kiev.ua, ²dwl@visti.net

Живучість інформаційних сюжетів як динамічних документальних систем

Розглянуті питання живучості інформаційних сюжетів в інформаційному просторі з точки зору теорії систем. Приведені приклади, алгоритми і моделі формування інформаційних сюжетів, розглянуті підходи до кількісної оцінки їх живучості.

Ключові слова: інфосюжет, система, живучість, інформаційне просторі, інформаційна безпека.

Визначимо інформаційний сюжет (далі — інфосюжет) як множину документів (інформаційних повідомлень), присвячених одній темі або одній події. «Середовищем обитання» інфосюжетів є інформаційне просторі, сьогодні повністю репрезентативно представлено мережею Інтернет, що, однак не обмежує авторів розглядом тільки цієї мережі. Інфосюжети можна трактувати як документальні або контентні системи (від англ. content — зміст), які, як буде показано нижче, повністю задовольняють загальному визначенню систем.

В цій статті розглядається таке властивість інфосюжетів як живучість. Очевидно, що живучість інфосюжета може, з однієї сторони, розглядатися як його об'єктивна властивість, яка залежить від тематики, аудиторії, часу, а з іншої сторони, як характеристика, яку хочуть надати йому в разі штучного формування, наприклад, в ході проведення інформаційних операцій.

Відомо, що системи в загальному випадку можуть розділятися на цілеспрямовані і нецелеспрямовані. Інфосюжети як системи можуть відноситися як до першого, так і до другого класу. Живучість штучно формованих інфосюжетів має вирішальне значення, наприклад, в ході рекламних кампаній і інших інформаційних операцій.

Візьмемо до уваги традиційне розуміння поняття живучості. Живучість — це фундаментальна властивість складних систем. Біологічні, соціальні і багато інших систем зазвичай мають властивість живучості, що дозволяє їм зберігати цілісність, виконувати свої функції і розвиватися незважаючи на деградацію і незалежно від наявності несприятливих впливів з боку зовнішнього середовища [1]. Методи і засоби забезпечення живучості успішно застосовуються при створенні складних технічних систем.

Інфосюжети звичайно не можна вважати ні біологічною, ні технічною системою, хоча окремі елементи цих систем необхідні для їх існування. Швидше

всего, инфосюжеты можно отнести к системам коммуникационным, на формирование которых существенное влияние оказывает так называемый «человеческий фактор», который, пожалуй, сложнее всего поддается формализации.

Живучие системы способны поддерживать непрерывное выполнение своих основных функций (в случае инфосюжетов — информировать о наиболее важных аспектах тематики или события), временно или постоянно отказываясь от выполнения менее важных функций информирования, изменять свою структуру и поведение, находить и выполнять новые функции, необходимые для успешного противостояния неблагоприятным воздействиям из внешней среды (информационного пространства), приспосабливаясь к условиям своего функционирования [2].

Для того, чтобы перейти к рассмотрению инфосюжета как системы и сопоставить его свойства со свойствами остальных 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 общем смысле рассматривается как динамически изменяемая совокупность сильно связанных объектов, обладающая свойствами организации, связности, целостности и членности.

Соответственно, информационный сюжет 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. Последнее относится к событиям, которые можно трактовать как нецеленаправленные системы. То есть действительно, инфосюжет — это совокупность сильно связанных объектов, обладающая свойствами организации, связности, целостности (определяемой тематикой или событием) и членности (на отдельные документы, их аспекты).

Рассмотрим в начале свойства систем, связанные с целями и функциями.

1. Синергетичность — однонаправленность действий компонентов усиливает эффективность функционирования системы. В случае инфосюжетов направленность отдельных документов усиливает информационную функцию всего инфосюжета.

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

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

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 инфосюжетов в случае их искусственного формирования (вместе с тем, существуют и нецеленаправленные системы, в том числе и инфосюжеты).

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 инфосюжетов следующие.

1. Целостность — первичность целого инфосюжета по отношению к отдельным документам.

2. Неаддитивность — принципиальная несводимость свойств инфосюжетов к сумме свойств составляющих их документов.

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

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

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

1) коммуникативность — существование сложной системы коммуникаций инфосюжетов с информационным пространством, в частности, отдельные документы из инфосюжета могут быть связаны не только с другими документами из того же инфосюжета, но и с другими частями информационного пространства;

2) взаимодействие и взаимозависимость инфосюжета и информационного пространства;

3) адаптивность — стремление к состоянию устойчивого равновесия, которое предполагает адаптацию параметров инфосюжета на определенных этапах его жизненного цикла к изменяющимся параметрам внешней среды;

4) надежность — существование инфосюжета при выходе из строя отдельных его компонент (документов), сохраняемость значений параметров системы в течение определенного периода;

5) интерактивность — взаимодействие с внешней средой и «ответная» изменчивость инфосюжетов.

Существует еще ряд системных свойств инфосюжетов, таких как:

1) интегративность — наличие системообразующих, системосохраняющих факторов;

2) эквивинальность — способность инфосюжетов достигать состояний независимых от исходных условий и определяющихся только параметрами системы;

3) наследственность;

4) развитие;

5) самоорганизация и т.д.

Для полноценной работы или сохранения минимального набора критически важных функций информирования инфосюжет должен 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 весь информационный сюжет, так и на отдельные документы, входящие в него, могут оказываться различные дестабилизирующие информационные воздействия, атаки, например, удаление отдельных материалов с веб-сайтов сети Интернет, уничтожение или отключение информационных серверов, публикация новых документов, которые в определенном направлении исказят исходный инфосюжет, или порождение нового инфосюжета, который может снизить актуальность или попросту уничтожить исходный сюжет.

Понятно, что для полноценной работы и сохранения минимального набора критически важных функций информирования инфосюжет должен обладать вполне определенным запасом устойчивости к внешним дестабилизирующим воздействиям. При этом нарушение целостности инфосюжета на фоне снижения актуальности его компонент влечет за собой дезорганизацию, одновременную потерю гибкости — понижение жи-

вучести и нарушение целостности, то есть потерю важнейших функций инфосюжетов (рис. 1) [3].

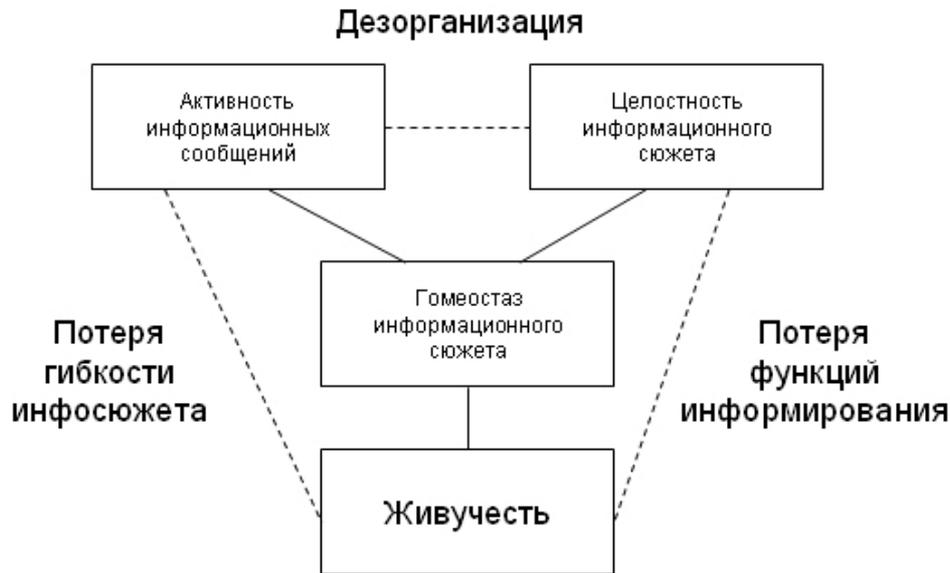


Рис. 1. Модель гомеостаза инфосюжета

Информационные сюжеты могут быть представлены как сетевые структуры, так называемые динамические сети [4]. Текущее состояние инфосюжета может быть представлено в виде графа $\langle M, L \rangle$, где M — это множество документов, входящих в сюжет, а L — множество ребер — 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 инфосюжетах важно ранжировать информационные сообщения, выделять основные из них. Тут на помощь могут прийти такие популярные методы ранжирования, как PageRank, HITS, Salsa, а также параметры, вычисляемые в рамках теории сложных сетей (complex networks).

К потере живучести инфосюжета, например, может привести разрыв связей между его компонентами, например, устранением из информационного пространства наиболее весомых компонент, то есть таких, которые имеют 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.

Как и сеть террористов, восстановление которой после деструктивного воздействия описано в [5], информационный сюжет также является динамической системой, восстановление которой после уничтожения лучших «посредников» осуществляется за счет латентных связей с другими компонентами информационного пространства. После того как инфосюжет разделяется на изолированные фрагменты, он может использовать эти связи и быстро восстановить связность (рис. 2).

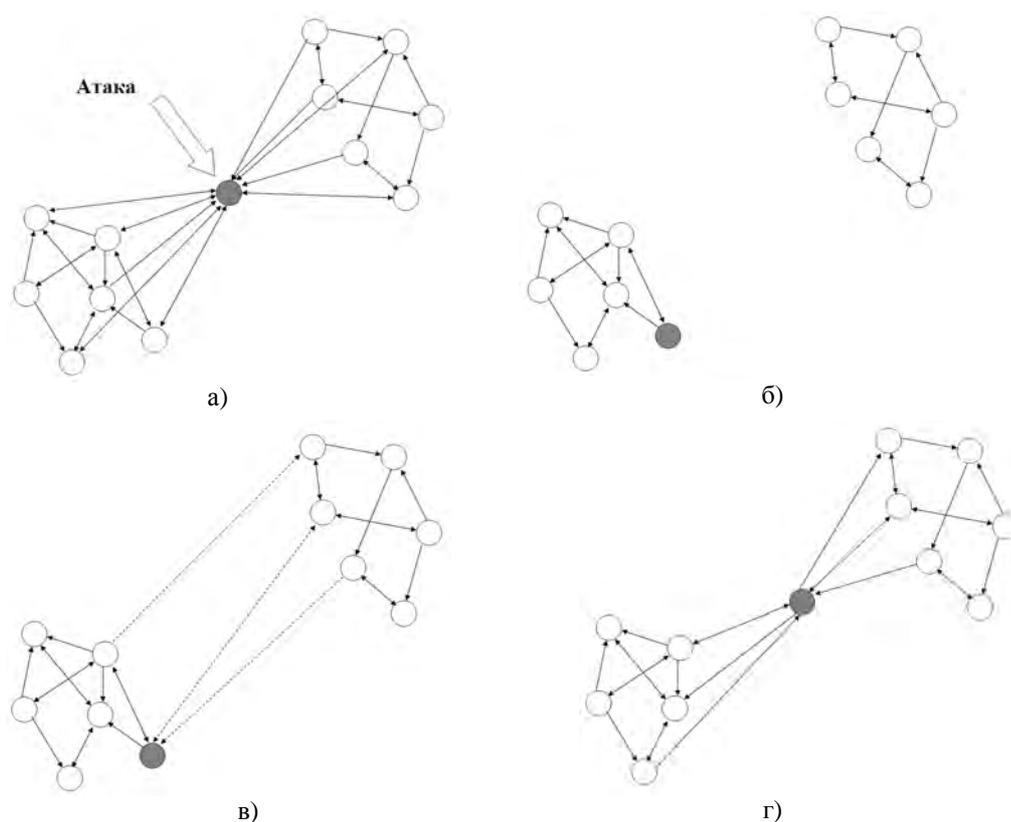


Рис. 2. Восстановление структуры сети путем выбора нового «посредника»: а) атака на сеть; б) несвязная сеть с изъятым посредником; в) у-возрождение скрытых (латентных) связей; г) связность сети восстановлена

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 инфсюжета зависит от того, смогут ли снова разъединенные части получить взаимные связи, недостаток которых наблюдается в этой части инфсюжета. Если часть инфсюжета 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.

На практике при поиске новостной информации всегда возникает задача выявления инфсюжетов, состоящих из отдельных документов, и их ранжирования по некоторым признакам, что должно обеспечить, не только выявления самой важной темы, но и «веерное» многоаспектное освещение всех наиболее значимых аспектов инфсюжетов. Эта задача, решаемая 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, то есть инфсюжетов, «взвешивание» (оценка важности, актуальности) и наглядная визуализация самых весомых из них [6].

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. [6] Выделенные классы документов и представляют собой инфсюжеты.

Для предъявления пользователям инфсюжеты 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 инфсюжете, а размер отражает общий интерес к конкретной теме. Во всех этих подходах центральная задача состоит в отождествлении документов, относящихся к одному сюжету и выделение «непересекающихся» сюжетов. На рис. 3 представлен типовой алгоритм выявления инфсюжетов. Последнее по времени генерации информационное сообщение (документ) сравнивается с предыдущими, оценивается уровень их подобия. Если уровень подобия с некоторым ранее сгенерированным документом превышает некоторый порог, анализируемый документ считается принадлежащим инфсюжету, к которому относится ранее сгенерированный документ. Если подобных документов не находится, фиксируется новый сюжет, состоящий на текущий момент из одного документа.

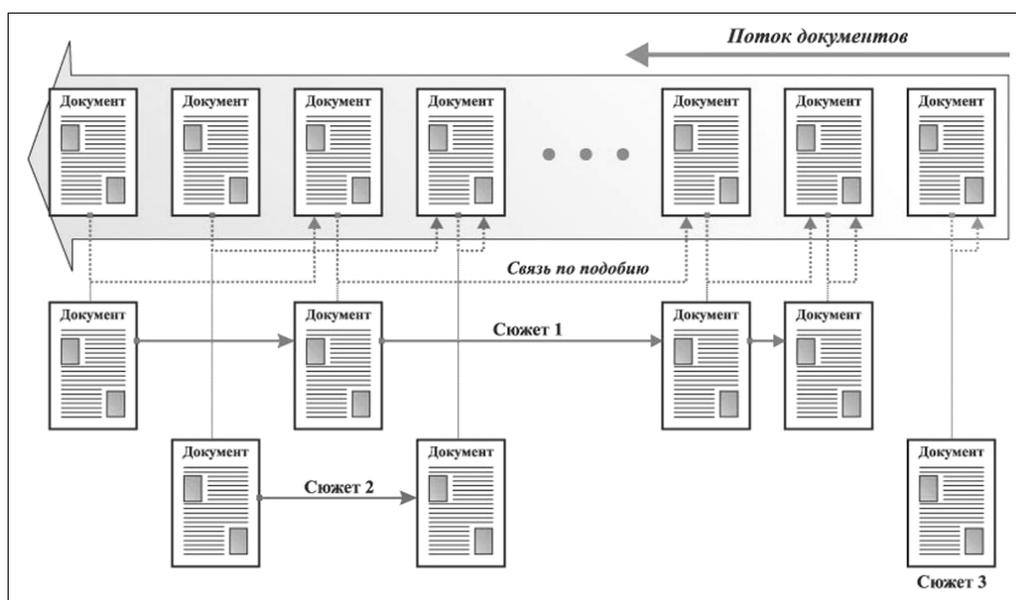


Рис. 3. Типовой алгоритм выявления информационных сюжетов

Для результирующего отображения каждого отдельно взятого инфсюжета используются отобранные по содержательной близости документов 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. То есть 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 инфсюжета строится матрица попарной близости документов, которая обрабатывается алгоритмом кластеризации с эмпирически подобранными параметрами (в частности, ра-

диусом метрики близости). Для того, чтобы увеличить связность крупных сюжетов, в Яндекс.Новости дополнительно используется кластеризация второго уровня, обеспечивающая сбор атомарных кластеров в более крупные. Все сообщения в результатах поиска на сайте Яндекс.Новости сгруппированы, при этом ранжирование инфосюжетов построено на стандартных для Яндекса принципах ранжирования выдачи. Оно основано на числе и ранге новостей внутри новостных сюжетов, при этом 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 [7]. 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 взвешиваются по длине и оперативности, после чего пользователю предъявляется определенное количество самих важных тематических инфосюжетов. Для репрезентации сюжетной цепочки, заголовки документов также взвешиваются относительно ключевых слов, соответствующих сюжету, а затем из всех заголовков выбираются наиболее «весомые» для отображения (рис. 4).

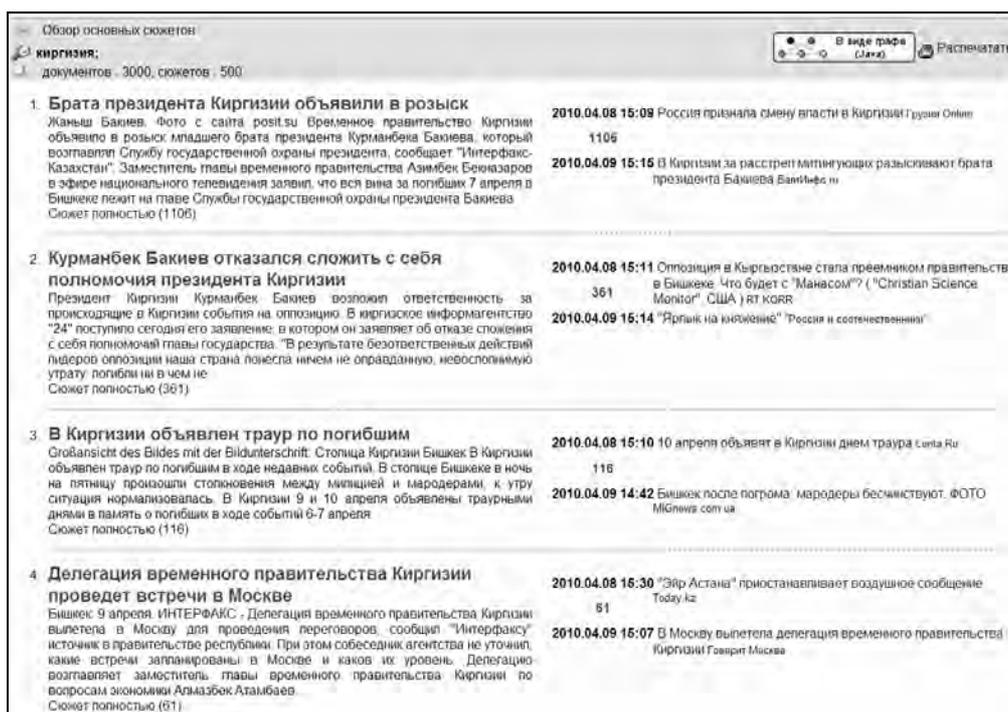


Рис. 4. Пример отображения инфосюжетов в системе InfoStream

Следует обратить внимание, что проблема автоматического построения качественных инфосюжетов на основе тематических потоков сетевой новостной информации сегодня практически решена.

Для изучения проблем живучести инфосюжетов как сложных многопараметрических систем, параметры которых еще малоизученны, наиболее подходящей методикой является математическое моделирование. Жизненный цикл информационных сюжетов может описываться, например, моделью диффузии информации [8]. 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 случае двумерной решетки, элементами которой являются квадраты, ближайшими соседями, входящими в окрестность элемента $y_{i,j}$, можно считать либо только элементы, расположенные вверх-вниз и влево-вправо от него, либо добавленные к ним еще и диагональные элементы (окрестность Мура). В модели Мура каждая клетка имеет восемь соседей. Это позволяет определять общее соотношение значения клетки на шаге $t + 1$ по сравнению с шагом t :

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

В рамках данной модели, которая относится к распространению новостей в информационном пространстве, применяются окрестность Мура [9] 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, то есть 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 кле-

точных 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

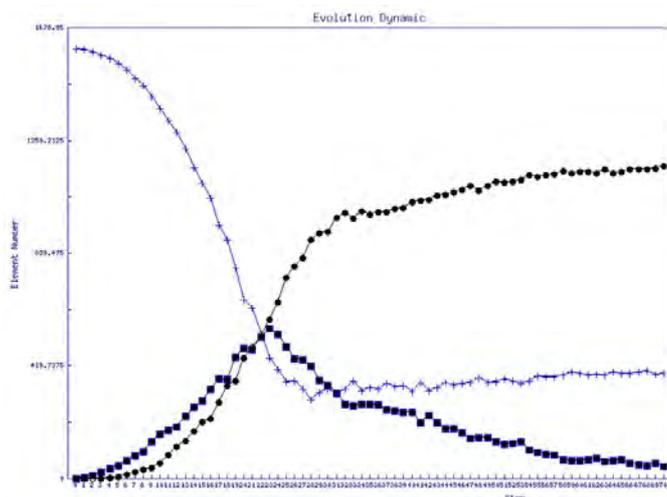


Рис. 6. Распределение клеток в зависимости от такта системы клеточных автоматов: белые клетки — (+); серые клетки — (•); черные клетки — (■)

Именно черные клетки образуют актуальный инфосюжет, динамика которого представлена на рис. 5.

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 живучести инфосюжетов, также как и 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.

В булевой модели можно предположить, что инфосюжет состоит из n элементов (документов), при этом i -му элементу соответствует булева переменная x_i , которая может принимать значения $\{0, 1\}$, то есть:

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

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

Состояние инфосюжета определяется структурной булевой функцией его работоспособности (действенности), зависящей от переменных x_1, x_2, \dots, x_n :

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

Если активность элемента инфосюжета (документа) 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 инфосюжета, как для системы, определяется вероятность его работоспособности по известным правилам [10, 11].

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.

Инфосюжет 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 фиксированного 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 инфосюжета в заданный момент времени t (работоспособность системы) [10].

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 компонентов 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 [12].

Пусть $E = \{e_i\}$ — множество деструктивных воздействий на инфосюжет; $\sigma_j(e_i)$ — показатель эффективности (качества) функционирования j -го варианта инфосюжета при воздействии e_i ; $H_j(E) = \min_{e_i \in E} \sigma_j(e_i)$ — показатель живучести инфосюжета для множества возможных воздействий на него E . Тогда при целенаправленном формировании инфосюжета задача проектирования состоит в том, чтобы из множества вариантов инфосюжетов 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 инфосюжетов при неблагоприятных для инфосюжета воздействиях, часто ставится задача сохранения определенного уровня его 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 рассмотрим инфосюжет, состоящий из четырех документов ($n = 4$). Деструктивное воздействие на инфосюжет — 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 инфосюжета. Остальные три документа считаются равноправными. Устранение любых двух из них ($k = 4$) также ведет к потере функциональности инфосюжета.

Если обозначить состояние инфосюжета 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, то есть равновероятны.

Если состояния инфосюжета не являются равновероятными, то живучесть G инфосюжета 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.

Более общую оценку живучести инфосюжета можно построить, исходя из цели его 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 информирования инфосюжета. 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 инфосюжета 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 инфосюжета будем характеризовать тем, на выполнение каких функций информирования 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 инфосюжета 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 инфосюжета введем понятие характеристического вектора состояния — n -мерного вектора (n — мощность множества элементарных функций системы). Начальной конфигурации инфосюжета при условии, что выполняется все множество функций F , будет соответствовать характеристический вектор состояния $(0, 0, \dots, 0)$. Некоторой текущей конфигурации инфосюжета 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 , то есть d_i — количество реконфигураций инфосюжета из-за «отказа» функции информирования $f_i \in F$.

Решим задачу нахождения множества характеристических векторов состояний инфосюжета, 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 инфосюжета S_f , определяющих состояния, в которых возможен выбор конфигурации, обеспечивающей выполнение множества элементарных функций F . Пусть некоторый начальный инфосюжет 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, то есть:

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

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

В качестве критерия оптимизации естественно выбрать

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

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, то есть изменяется состояние инфосюжета. 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:

$$\begin{aligned} \Phi &= \sum_{i=1}^n \sum_{j=1}^p b_{ij} \varphi_{\text{эф}}(i, j, B) = \max(\min), \\ \sum_{i=1}^n b_{ij} &= 1, \forall j = \overline{1, p}, \\ \sum_{j=1}^p b_{ij} &= 1, \forall i = \overline{1, n}, \\ \Phi &\geq \Phi^* (\Phi \leq \Phi^*). \end{aligned}$$

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 инфосюжета S_f выделим подмножество S_q , определяющее состояния системы, в которых возможен выбор конфигураций, обеспечивающих выполнение цели функционирования.

В качестве оценки живучести инфосюжета можно *взять мощность множества* S_q . В случае инфосюжетов 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 каждого конкретного инфосюжета, исходя из его специфики. *Количественный показатель живучести конкретного инфосюжета 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 ин-

формирования; Δ — множество функций инфосюжета, выполняемых в заданных условиях ($\Delta \subseteq \Theta$); α_n — критичность n -й функции. Таким образом, количественная оценка живучести инфосюжета будет изменяться в интервале $[0, 1]$. Живучесть тем выше, чем больше ее количественная оценка.

Понятие *живучести* системы (инфосюжета) подразумевает ее способность своевременно выполнять свои функции (в рассматриваемом случае информирования) в условиях действия дестабилизирующих факторов. В случае информационных сюжетов такими факторами могут выступать устранение отдельных документов из информационного пространства, потеря их актуальности, доступности. Привлечение внимания аудитории к другой теме, порождение другого информационного сюжета также может снизить актуальность текущего инфосюжета. Вместе с тем, с точки зрения теории живучести, происхождение неблагоприятного воздействия играет значительно меньшую роль, чем его последствия.

Инфосюжеты как системы могут быть как целенаправленными, так и нецеленаправленными. При этом целенаправленные сюжеты могут порождаться в ходе активных рекламных кампаний, в качестве отвлекающих факторов, информирования и других элементов информационных операций [5]. При этом живучесть, проявляющаяся как способность целенаправленных инфосюжетов выполнять свои функции информирования на заданном отрезке времени без отказов, определяет их минимальный порог устойчивости, за которым без восстановления компонент и функций инфосюжет может потерять свою актуальность и возможность влияния. Вследствие этого и многих других факторов живучесть информационных сюжетов имеет важнейшее значение для информационной безопасности.

1. *Shelton C. A Framework for Scalable Analysis and Design of System-Wide Graceful Degradation in Distributed Embedded Systems* / C. Shelton, P. Koopman, W. Nace // Eighth IEEE International 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. *Modeling the Revolution in Military Affairs, Autumn // Winter 1998–99 / JFQ.*
4. *Survivable Network Systems: An Emerging Discipline CMU/SEI-97-TR-013 ESC-TR-97-013 Software Engineering Institute Carnegie Mellon University Pittsburgh.* — May 1999.
5. *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.
6. *Ландэ Д.В.* Поиск знаний в Internet. Профессиональная работа / Д.В. Ландэ. — М.: Диалектика, 2005. — 272 с.
7. *Григорьев А.Н.* Адаптивный интерфейс уточнения запросов к системе контент-мониторинга InfoStream / А.Н. Григорьев, Д.В. Ландэ // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2005» (Звенигород, 1–6 июня, 2005 г.). — М.: Наука, 2005. — С. 109–111.
8. *Ландэ Д.В.* Модель диффузии информации / Д.В. Ландэ // Информационные технологии и безопасность. Менеджмент информационной безопасности: сб. науч. тр. Ин-та проблем регистрации информации НАН Украины. — 2007. — Вып. 10. — С. 51–67.
9. *Нейман Дж.* Теория самовоспроизводящихся автоматов / Дж. Нейман. — М.: Мир, 1971. — 382 с.
10. *Райншке К.* Модели надежности и чувствительности систем / К. Райншке. — М.: Мир, 1979. — 454 с.
11. *Райншке К.* Оценка надежности систем с использованием графов / К. Райншке, И.А. Ушаков. — М.: Радио и связь, 1988. — 208 с.
12. *Додонов А.Г.* К вопросу безопасности информационных систем / А.Г. Додонов, Д.В. Флейтман: зб. наук. праць «Інформаційні технології та безпека». — К., 2004. — Вип. 6. — С. 26–29.

Поступила в редакцию 15.05.2010