НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ ПРОБЛЕМ РЕГИСТРАЦИИ ИНФОРМАЦИИ НАН УКРАИНЫ

НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ УКРАИНЫ
«КИЕВСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ»
ФАКУЛЬТЕТ СОЦИОЛОГИИ И ПРАВА
УЧЕБНО-НАУЧНЫЙ ЦЕНТР ИНФОРМАЦИОННОГО ПРАВА И
ПРАВОВЫХ ВОПРОСОВ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И БЕЗОПАСНОСТЬ

МАТЕРИАЛЫ XVI МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ

ВЫПУСК 16

Киев – 2017

Рекомендовано к печати ученым советом
Института проблем регистрации информации НАН Украины
(протокол № 7 от 10 января 2017 г.)

Информационные технологии и безопасность. Материалы XVI Международной научно-практической конференции ИТБ-2016. – К.: ИПРИ НАН Украины, 2017. – 194 с. ISBN: 978-966-02-7422-8

В сборник вошли материалы докладов, представленных на XVI Международной научно-практической конференции «Информационные технологии и безопасность» (ИТБ-2016, 1 декабря 2016 года, г. Киев, Украина).

В сборнике представлены статьи, посвященные вопросам создания и внедрения информационных технологий, актуальным проблемам технологического и правового обеспечения информационной и кибербезопасности, противодействия информационным операциям и кибертерроризму, проведения аналитических исследований на основе анализа контента сети Интернет.

Для специалистов в области информационных технологий, информационной безопасности, информационного права, а также для аспирантов и студентов старших курсов высшей школы соответствующих специальностей.

ISBN 978-966-02-7422-8

ИССЛЕДОВАНИЕ ИСТОЧНИКОВ ИНФОРМАЦИОННОГО ВЛИЯНИЯ ВЕБ-РЕСУРСОВ СЕТИ ИНТЕРНЕТ

Додонов А.Г., Ландэ Д.В.
Институт проблем регистрации информации НАН Украины, Киев

В статье описывается технология построения сети влияния источников информации на основе анализа контекстных ссылок. Технология включает методы и средства, базирующиеся на контент-мониторинге глобальных сетей, концепциях Complex Networks (комплексных сетей) и Text Mining (глубинный анализ текстов). В отличие от методов анализа гиперссылок в сетевых документах, применяемых для анализа популярности веб-страниц в Интернете, предлагаемая технология учитывает взаимное влияние источников информации, выраженное в виде ссылок в тексте или перепечаток существенных фрагментов текста. Представлены методы и средства анализа сетей взаимного влияния источников информации, отражающих различные тематические срезы, а также информационные операции.

Введение, постановка задачи

В настоящее время Интернет представляет собой значимый фрагмент информационного пространства, его влияние на людей постоянно возрастает. При этом следует учитывать различные механизмы распространения информации в сети, взаимного влияния источников информации, через которых и осуществляется воздействие на пользователей. Информационное пространство Интернета (веб-ресурсы, социальные сети) сегодня является мощнейшей площадкой для проведения информационных операций [1], признаки которых позволяют определять различные методики [2, 3].

В данной работе предлагается технология определения влиятельности сетевых источников информации на основе анализа контекстных ссылок. Представлена технология и методика ранжирования источников информации на основе оценки контекстных ссылок. В отличие от методов анализа гиперссылок в сетевых документах, применяемых для анализа популярности веб-страниц в Интернете, в предложенном подходе учитывается взаимное влияние источников информации, выраженное ссылками в

тексте и перепечатками существенных фрагментов текстов. При этом предполагается, что влияние источников информации друг на друга определяется наличием контекстных ссылок или перепечаток (см. рис. 1) Также предложен подход к оперативному выявлению информационных операций на основе анализа сетей взаимных ссылок источников информации.

Рисунок 1 – Гипотеза о соотношении наличия контекстных ссылок и влияния источников информации

Технологические этапы исследования взаимного влияния источников информации

Для эффективного исследования взаимного влияния источников информации из сети Интернет (веб-ресурсов, социальных медиа) предлагается последовательность шагов, этапов обработки информации, каждый из которых сам по себе обеспечивает получение аналитического продукта. Совокупность таких этапов, базирующихся на использовании необходимых и доступных инструментальных средств, специальных приемов, можно рассматривать как процедуру проведения действий, нацеленных на получение аналитических материалов, включающих построение и анализ сети их взаимного влияния.

При проведении данных информационно-аналитических исследований на базе контент-мониторинга к таким задачам можно отнести:

- Нахождение релевантных публикаций по заданной тематике.
- Выявление взаимных контекстных ссылок и перепечаток в документах, представленных разными информационными источниками.
- Построение сети влияния, анализ и визуализация взаимосвязей информационных источников, в том числе ранжирование узлов построенной сети по степени влиятельности.
- Выявление возможных информационных операций и построение сценария противодействия информационным операциям в сетевой среде.

Соответственно процедура исследования взаимного влияния источников информации, охватывает такие шаги:

1. Получение репрезентативного массива публикаций по выбранной тематике.
2. Выявление контекстных ссылок и перепечаток в тематическом информационном потоке.
3. Построение сети влияния источников информации на основе анализа контекстных ссылок и перепечаток.
4. Исследование сети влияния источников информации, ранжирование узлов по степени влиятельности.
5. Выявление возможных информационных операций и построение сценария противодействия информационным операциям в сетевой среде.

Рассмотрим эти шаги подробнее на конкретных примерах.

Получение репрезентативного массива публикаций

Для получения репрезентативного массива публикаций по выбранной тематике необходимо выбрать систему контент-мониторинга, предоставляющую поток информационных сообщений по определенной тематике. Тематика может выражаться запросом на языке информационно-поисковой системы.

В качестве системы контент-мониторинга авторами была выбрана система InfoStream, которая в настоящее время охватывает 10 тыс. источников информации на русском и украинском языках. В базы данных системы ежедневно поступает свыше 100 тыс. документов. Система InfoStream обеспечивает поиск, а также просмотр списка и полных текстов релевантных документов.

В приведенном на рис. 2 примере показан фрагмент интерфейса системы, через который отрабатывался запрос, относящейся к обсуждению в январе 2016 года вопроса от ставки премьер-министра Украины А. Яценюка. В результате был сформирован тематический информационный массив, охватывающий 3196 документов.

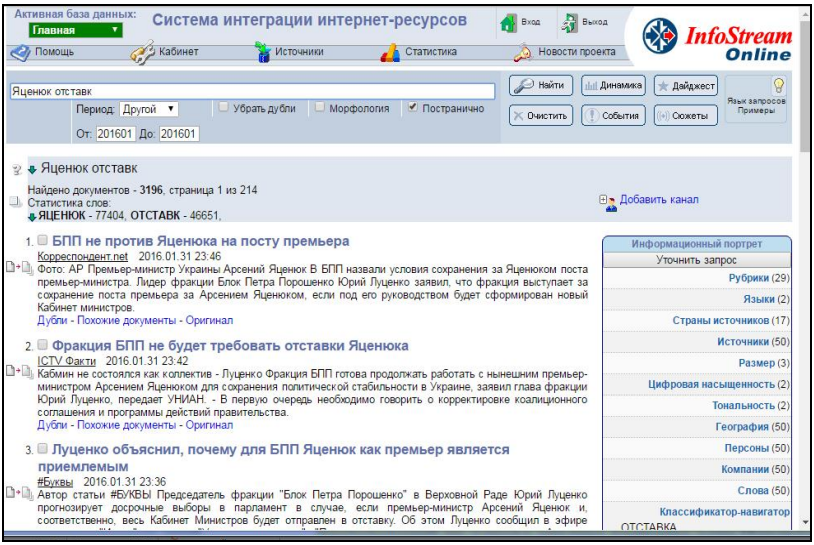


Рисунок 2 – Фрагмент интерфейса системы контент-мониторинга

Выявление контекстных ссылок

Основой построения сети влияния источников информации являются контекстные ссылки и перепечатки в тематическом информационном потоке. Контекстные ссылки выявляются путем идентификации шаблонов (Табл. 1) в документах выбранного информационного массива и признаков точных перепечаток, определяемых методами выявления плагиата [4, 5]. В свою очередь, сами шаблоны периодически определяются/дополняются экспертами в автоматизированном режиме путем анализа контекста потока документов системы контент-мониторинга методами Text Mining.

Построение сети влияния источников информации

Найденные в текстах контекстные ссылки и перепечатки позволяют сформировать матрицу цитирования, транспонируя которую в соответствии с приведенный выше гипотезой формируется матрица влияния. Данной матрице соответствует сеть влияния источников, пример визуализации которой для рассмотренного выше тематического информационного массива с помощью системы Gephi приведен рис. 3.

Таблица 1. Шаблоны названий информационных ресурсов (фрагмент)

№	*Код источника*	*Шаблон 1*	*Шаблон 2*
1	srd06193	Depo	"Депо"
2	srd03176	Українські національні новини	УНН
3	srd00045	Сегодня.ua	"Сегодня"
4	srd03076	ТСН.ua	"ТСН"
5	srd07509	112.ua	"112"
6	srd02348	Gazeta.ua	
7	srd00069	Корреспондент.net	"Корреспондент"
8	srd07487	Еспресо TV	"Еспресо ТВ"
9	srd02535	Телеканал новин "24"	"24"
10	srd06453	Телеграф.com.ua	"Телеграф"
11	srd01351	ЗІК	"ЗІК"
12	srd02514	РБК-Україна	РБК-Украина
13	srd04508	Українські Новини	
14	srd07686	"Антикор"	
15	srd00253	"Обозреватель"	
16	srd00057	Интерфакс	Інтерфакс
17	srd00404	ICTV Факти	ICTV
18	srd04125	РИА Новости Украина	
19	srd02732	УКРІНФОРМ	УКРИНФОРМ
20	srd00095	УНІАН	УНИАН
21	srd00094	Українська правда	
22	srd01408	Цензор.Нет	
23	srd00064	ЛІГАБізнесІнформ	
24	srd00039	Газета День	
25	srd07038	Вести.ua	

Построение сети влияния источников информации

Найденные в текстах контекстные ссылки и перепечатки позволяют сформировать матрицу цитирования, транспонируя которую в соответствии с приведенной выше гипотезой формируется матрица влияния. Данной матрице соответствует сеть влияния источников, пример визуализации которой для

рассмотренного выше тематического информационного массива с помощью системы Gephi приведен рис. 3.

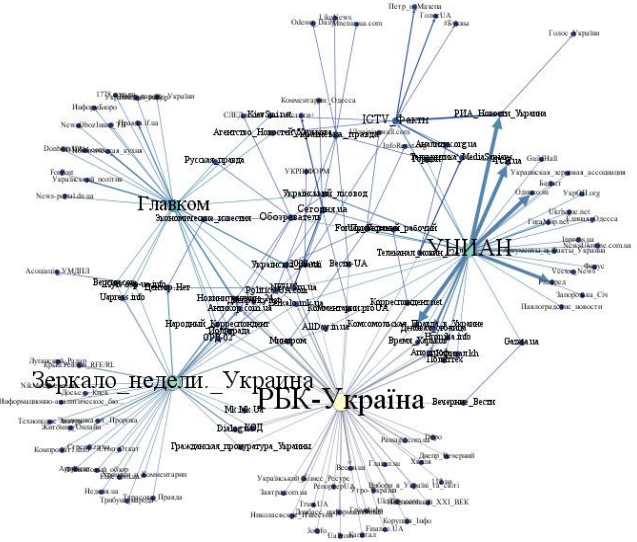


Рисунок 3 – Фрагмент сети связей источников по выбранной тематике

Исследование сети влияния источников информации

Построенную сеть влияния источников информации можно исследовать как с помощью общепринятых инструментальных средств (например, с помощью системы Gephi были получены такие параметры построенной сети, как количество узлов: 141, ребер: 196, плотность графа: 0,01, средний коэффициент кластеризации: 0,026, средняя длина пути: 1,26 и т.д.).

Для содержательного анализа большое значение имеет вес узлов сети, список самых весомых узлов по критерию исходящей мощности приведен в Таблице 2.

Перспективным подходом к ранжированию источников по уровню влияния является алгоритм HITS, предложенный Дж. Кляйнбергом [6].

Алгоритм HITS обеспечивает выбор из сети лучших «авторов» (узлов, на которые введут ссылки) и «посредников» (узлов, от которых идут ссылки включения). Узел является хорошим посредником, если от него идут связи на другие важные узлы, и

наоборот, хорошим автором, если на него ведут ссылки от важных узлов.

Таблица. 2. Наиболее влиятельные узлы по количеству цитирования

№	Веб-ресурс	Исходящая мощность
1	РБК-Україна	50
2	Зеркало недели	38
3	УНИАН	35
4	Главком	28
5	ICTV-Факты	10
6	Сегодня.ua	7
7	Українська правда	7
8	Обозреватель	6
9	Forbes-Украина	4
10	Цензор.Нет	3

В соответствии с алгоритмом HITS для каждого узла сети v_j рекурсивно вычисляется его значимость как автора $a(v_j)$ и посредника $h(v_j)$ по формулам:

$$a(v_j) = \sum_{i \to j} h(v_i); \; h(v_j) = \sum_{i \to j} a(v_i).$$

В данных формулах суммирование производится по всем узлам, которые ссылаются (или на которые ссылаются – во второй формуле) на данный узел.

Перефразируя обозначения, приведенные в [6], а именно заменяя «авторство» на «подверженность влиянию», а «посредничество» на «влиятельность» можно с небольшими вычислительными затратами вычислять соответствующие характеристики узлов сети влияния.

Также для выявления информационных влияний большое значение имеет определение «скрытых» связей, т.е. когда прямых связей между узлами нет, но прослеживаются связи через другие (вторые, третьи и т.п. узлы). Методика определения скрытых связей, скрытых влияний приведена в работе [7].

Выявление возможных информационных операций

Сеть информационного влияния источников информации позволяет оперативно идентифицировать возможные

информационные операции в соответствии с подходами, предложенными в работе [2]. Предполагается, что вероятность наличия информационной операции мала, если информация о событии вначале зарождается во влиятельном информационном источнике, а затем перепечатывается (со ссылками или без них) менее влиятельными источниками (рис. 4). Обратные явления, когда более влиятельные издания перепечатывают информацию у менее влиятельных, пусть и многочисленных, может являться признаком информационной операции, атаки (рис. 5). Именно такие картины наблюдались при сетевом анализе реальных тематических информационных потоков (рис.6).

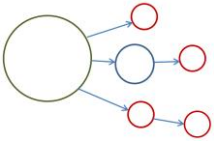


Рисунок 4 – Типовой сценарий распространения информации

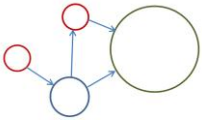


Рисунок 5 – Сценарий распространения информации, характерный для информационной операции

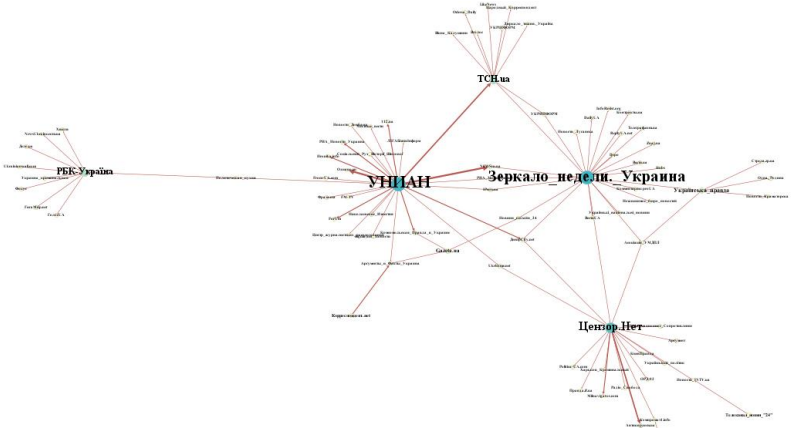


Рисунок 6 – Фрагмент сети связей источников по специальной тематике

В работе [1] приведены шаги противодействию выявленной (или возможной) информационной операции, некоторые из которых эффективно решаются в рамках предложенных технологических подходов, а именно:

1. Сбор информации с публикациями об объекте атаки;
2. Анализ контента публикаций в критических точках.
3. Определение источников, публикующих негативную информацию об объекте атаки.
4. Определение «первоисточников» – тех источников, которые первыми опубликовали негативную информацию.
5. С учетом реалий и публикаций оцениваются вероятные последствия.
6. Организуется информационное противодействие, диалог с наиболее влиятельными изданиями и т.д. Примеры публикаций в контексте информационного противодействия находятся в ретроспективной базе данных системы контент-мониторинга.

Выводы

Таким образом, представлена технология и методика ранжирования источников информации по влиятельности на основе оценки контекстных ссылок. Предложен подход к оперативному выявлению информационных операций на основе анализа сетей взаимных ссылок источников информации.

В работе также представлена технология выявления значимости информационного взаимного влияния различных источников информации – веб-ресурсов, а соответственно, и на конечных потребителей информации – пользователей сети Интернет. Данная технология базируется как на современных методах и инструментальных средствах контент-мониторинга глобальных сетей, так и на современных подходах Text Mining, распознавания образов, ранжирования узлов в информационных сетях, средствах анализа и визуализации информационных потоков.

Предложенную технологию можно использовать в качестве основы для выявления различных видов информационного влияния на основе исследования контента современных компьютерных сетей.

Публикация содержит результаты исследований, проводимых при грантовой поддержке Государственного фонда фундаментальных исследований по конкурсному проекту Ф73 № 23558 “Разработка методов и средств поддержки принятия решений при обнаружении информационных операций”.

Литература

1. Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. – К.: Інтертехнологія, 2009. – 164 с.
2. Потемкин А.В. Распознавание информационных операций средств массовой информации сети Интернет // Науковедение, 2015. –Том 7, №3. – URL: http://naukovedenie.ru
3. Додонов А.Г., Ландэ Д.В., Прищепа В.В., Путятин В.Г. Конкурентная разведка в компьютерных сетях.– К.: ИПРИ НАН Украины, 2013. – 248 с.
4. Ланде Д.В. Елементи комп'ютерної лінгвістики в правовій інформатиці. – К.: НДІІП НАПрН України, 2014. – 168 с.
5. Зеленков Ю.Г, Сегалович И.В. Сравнительный анализ методов определения нечетких дубликатов для Web-документов // Труды 9-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL, 2007. – Т. 1. – С. 166-174.
6. Kleinberg J.M. Authoritative Sources in a Hyperlinked Environment // Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 1998, and as IBM Research Report RJ 10076, May 1997.
7. Snarskii A.A., Zorinets D.I., Lande D.V. "Conjectural" links in complex networks // Physica A: Statistical Mechanics and its Applications, 2016. – Vol. 462. – pp. 266-273.

ПРОБЛЕМЫ БЕЗОПАСНОСТИ СОЦИОТЕХНИЧЕСКИХ СИСТЕМ

А.Г. Додонов, Е.С. Горбачик, М.Г. Кузнецова,

Институт проблем регистрации информации НАН Украины

Развитие и внедрение информационных технологий в повседневную жизнь привело к появлению большого количества так называемых социотехнических систем (СТС), которые представляют собой сложные системные образования, включающие технико-технологические подсистемы, системы ролей и функций обслуживающего и управленческого персонала (соответствующие системы деятельности, т.е. социальные подсистемы), внешнюю среду, активно взаимодействующую с СТС в условиях неопределенности факторов воздействия и перманентной изменчивости самой среды.

Наблюдается постоянное усложнение функций и структуры социотехнических систем, а также задач, решение которых обеспечивают СТС. Многие СТС являются звеньями в цепочках решения экономических задач, социальных и медицинских программ, поддержки государственной, оборонной, научной деятельности, системами управления высокотехнологическими объектами, критическими инфраструктурами и т.п. Качество и безопасность жизни все больше зависит от качества и эффективности функционирования СТС, поэтому уже сейчас остро стоит вопрос разработки методов и методик построения стратегий недопущения в СТС чрезвычайных кризисных явлений, приводящих к разрушению или неуправляемости как составляющих, так и СТС в целом. Необходимы методы и средства коррекции функционирования СТС в зависимости от действующих факторов внешней среды и внутрисистемного состояния.

Жесткость конкурентной среды с перманентно возникающими конфликтными ситуациями порождают серьезные риски и проблемы в сфере безопасности функционирования СТС, высокие требования к моделям и системам управления, к устойчивости развития СТС требуют наличия в них механизмов поддержки безопасности функционирования, коррекции и «программирования» поведения составляющих системы, своевременности изменения стратегии развития СТС [1].

Подходы, которые используются сегодня на практике для прогнозирования изменений в деловой среде и соответствующих

изменений поведения СТС, методы управления такими системами соединяют математические методы с элементами логико-эвристического программирования и предполагают непосредственное участие человека на всех этапах решения.

Критерии и оценки, на которые полагаются в процессе наработки решения, определяют исходя из анализа некоторых почти аналогичных ситуаций, возникавших ранее, но из-за неадекватности новой ситуации тем ситуациям, которые имели место в прошлом, надежность таких критериев и оценок, конечно же, не гарантируется.

Управленческие решения в СТС нарабатываются и реализуются в условиях неопределенности факторов воздействия и постоянной изменчивости среды функционирования, невозможности четкого предвидения реакции среды на действия системы, т.е. в условиях проявления таких фундаментальных системных свойств как устойчивость, адаптивность и живучесть. Используя возможности механизмов адаптации и живучести можно строить гибкие СТС, устойчивые к неизбежным непредвиденным изменениям внешней и внутренней среды функционирования системы.

Ставить задачу обеспечения живучести СТС в целом нецелесообразно, т.к. жизненный цикл многих СТС не продолжителен, требования по развитию и соответствию функционирования СТС требованиям деловой среды существенно важнее сохранности СТС как системы. Повышение живучести отдельных критических компонентов СТС, например, информационной инфраструктуры, являющейся базой для организации и реализации многих технологических процессов, взаимодействия различных подразделений СТС, бизнес-процессов и т.п., может позволить значительно повысить качество функционирования СТС, продлить жизненный цикл системы.

Достижение желаемого результата функционирования СТС зависит от своевременности принятия и реализации управленческого решения. Управление в СТС можно рассматривать как некоторое воздействие на систему, направленное на обеспечение необходимого поведения системы. Временные ограничения на процесс наработки решения зависят от допустимого времени управленческой реакции на динамику событий в деловой сфере (политике, экономике, бизнесе и т.п.). Управление должно сосредотачивать все ресурсы СТС на решении задач бизнеса, производства или других задач деловой среды, а также разрешении проблем дальнейшего развития собственно СТС. Затрата ресурсов на нейтрализацию или отражение

угроз безопасности усложняет, а иногда может даже сделать невозможным достижение цели управления, поэтому безопасность становиться условием эффективного управления. Кроме того, безопасность можно рассматривать и как цель управления, ведь управление СТС направлено на сохранение целостности системы, создание внутренних и внешних условий, гарантирующих стабильность и безопасность функционирования СТС, ее устойчивое развитие и совершенствование. Безопасность может также выступать и в качестве некоторого ресурса, необходимого для любого рода деятельности, и потеря которого может быть связана с серьезными материальными затратами.

Включение механизмов обеспечения живучести СТС позволит заблокировать широкий класс средств и способов неблагоприятного воздействия, минимизировать возможности целенаправленного изменения, уничтожения, копирования, тиражирования, блокирования, модификации информационных потоков и данных; значительно снизит риск дезорганизации работы СТС путем воздействия на системы защиты технико-технологической подсистемы, в частности, компьютерных систем и сетей. Учет возможностей механизмов обеспечения живучести позволяет иначе взглянуть на организацию безопасности в СТС и ее поддержание.

На практике меры по повышению безопасности СТС, как правило, направлены на усиление и усложнение систем защиты как в целом СТС, так и отдельных ее компонентов на организационном, нормативно-правовом, программно-техническом уровнях. Однако тотальный контроль и запретительные меры не устраняют «прорех» в технологиях реализации бизнес-процессов, не позволяют персоналу избежать ошибок или небрежного обращения с данными. Опыт свидетельствует, что даже если делаются существенные инвестиции в технологии защиты, это все равно не дает 100% гарантии безопасности. Так, несмотря на наличие развитых средств защиты, наблюдается постоянный рост ущерба от действий киберпреступников, который сегодня в глобальных масштабах оценивается в 400 млрд. долл. в год [2].

Главной причиной большинства инцидентов в СТС является человеческий фактор. Разграничение доступа, фильтрация, аутентификация и другие средства защиты технико-технологической подсистемы не позволяют обеспечить безопасность СТС на необходимом уровне. Как правило, доступ к конфиденциальным данным злоумышленники получают, используя знания психологии человека и социальные технологии. От знакомого человека

(заказчика), организации или компании приходит электронное письмо с вредоносным присоединением. Информацию об адресате часто получают из социальных сетей. Содержание письма провоцирует получателя открыть присоединение, после чего вредоносный код, попадая в компьютер, устанавливает связь с центром управления, находящимся на сервере, который управляется атакующим. Так устанавливается доступ в компьютерную сеть СТС. Кражей учетных записей атакующий может повысить свои права доступа и получить возможность изучить структуру сети, работающее ПО, средства защиты и т.п. Только на первых этапах проникновения используются какие-то методы взлома, в дальнейших действиях, как правило, пользуются стандартными инструментами ОС и сетевого ПО. В 60% случаев похищение данных занимает считанные часы, в то время, как 54% взломов остаются незамеченными месяцами [3]. По данным мировой статистики среднее время длительности выявления атак на компьютерные системы в 2015 г. составляла 146 дней, а для европейских компаний – около года [4].

Специалисты по безопасности считают, что следует не только защищаться от существующих угроз, но и готовиться к будущим, связанным с развитием коммуникационных технологий. Системы защиты и политика безопасности должны строиться в предположении, что атаки имеют место. Для противодействия новым возникающим, «мутирующим» угрозам необходимы адаптивные системы анализа, динамические модели распознавания атак, самонастраиваемые средства поддержания безопасности. В процессе управления безопасностью должна быть наработана система индикаторов «внутреннего предвидения» для того, чтобы, не оценивая все возможные альтернативы поведения СТС, определить подкласс альтернатив, которому принадлежит наиболее целесообразная (оптимальная) альтернатива поведения и число альтернатив в нем не превышало бы некоторого заданного числа. Кроме того, должны быть наработаны различные стратегии функционирования СТС при изменении факторов воздействия и характера взаимодействия с внешней средой.

Такие сложные, как СТС, системы, как правило, имеют внутренний механизм выбора рационального режима функционирования для достижения общесистемной цели (например, выполнения некоторого числа бизнес-процессов) в условиях неблагоприятных воздействий за счет внутренних ресурсов, перестройки структуры, изменения функций отдельных подсистем

или, возможно, алгоритмов их функционирования. Учитывая, что неблагоприятные воздействия могут вызвать существенные изменения СТС в целом или отдельных ее компонентов, выбор поведения должен осуществляться в соответствие с изменениями внешних условий и функциональным инвариантом системы, который можно определить как внутреннюю цель функционирования [5]. Выбор предполагает наличие некоторого множества возможных различных следствий, объединенных общим свойством соответствия одной внешней причине в данных условиях. Менять поведение, как известно, могут только системы, в принципе исключающие жесткую связь внешней причины выбора с фактическим поведением системы в результате выбора (внешние причины вызывают следствия, которые не могут быть предсказаны однозначно), что присуще СТС.

Базой создания технико-технологической подсистемы СТС достаточно часто являются компьютерные системы (КС), а функционирование многих СТС представимо как совокупность информационных взаимодействий. Использование информационных технологий в деловых процессах означает внедрение в СТС определенных стандартов безопасности [6], например, стандартов управления безопасностью ISO 15408, ISO 17799 (BS7799), BSI, стандартов аудита информационных систем и информационной безопасности COBIT, SAC, модель COSO, SAS 78/94. Согласно этим стандартам четко определяются цели обеспечения безопасности компьютерных систем; требования к системе управления информационной безопасностью; совокупность детализированных не только качественных, но и количественных показателей для оценки соответствия информационной безопасности задекларированным целям; определяется инструментарий обеспечения информационной безопасности и оценки ее текущего состояния; методики управления безопасностью с обоснованной системой метрик и мер обеспечения информационной безопасности, которые позволяют объективно оценить защищенность информационных активов и управлять информационной безопасностью.

Сегодня в архитектурных решениях закладываются возможности и средства повышения живучести (функциональной, структурной, информационной) КС, для чего задействуются механизмы распознавания, противодействия, восстановления, а также предлагаются специальные средства адаптации, реконструкции, реконфигурации и реорганизации [5,7-9]. Расходы на

обеспечение живучести КС можно рассматривать как расходы на повышение безопасности СТС. Благодаря повышению живучести КС можно обеспечить непрерывное функционирование инфраструктуры СТС, еще до анализа причин нарушения безопасности среагировать на неблагоприятное воздействие и перевести КС или ее отдельные ресурсы в безопасное состояние; значительно улучшить мониторинг КС; не прекращая функционирование, осуществлять реконфигурацию программного и аппаратного обеспечения, адекватную возникающим угрозам. В этом случае задача защищенности таких ресурсов СТС, как информации и технологий работы с нею, может формулироваться не как задача ограничения доступа к информационным ресурсам бизнес-процессов, а как задача прогнозирования критических ситуаций и ликвидации их последствий, т.е. в этом случае можно будет говорить о новом подходе к построению системы безопасности технико-технологической подсистемы СТС.

Литература

1. Додонов О.Г., Горбачик О.С., Кузнєцова М.Г. Системи організаційного управління: інформаційні технології та безпека // Сб. научн. трудов «Информационные технологии и безопасность: оценка состояния. Материалы Международной науч. конференции ИТБ-2013». – К.: ИПРИ НАН Украины, 2013. – Вып.13. – С.5-11.
2. Куликов Е. Эра безопасности [Электронный ресурс]. – Режим доступа: http://ko.com.ua/jera_bezopasnosti_117914.
3. Куликов Е. Защита после атаки [Электронный ресурс]. – Режим доступа: http://ko.com.ua/zashhita_posle_ataki_110371.
4. Бараш Л. Как защитить критические инфраструктуры [Электронный ресурс]. – Режим доступа: http://ko.com.ua/kak_zashhishhat_kriticheskie_infrastruktury_115808.
5. Додонов А.Г., Кузнецова М.Г., Горбачик Е.С. Живучесть и надежность сложных систем. Методическое пособие. – Международный научно-учебный центр ЮНЕСКО/МПИ информационных технологий и систем, 2001. – 163 с.
6. Петренко С. Методические основы защиты информационных активов компании [Электронный ресурс]. – Режим доступа: http://citforum.univ.kiev.ua/security/articles/zahita_aktivov/

7. Додонов А.Г., Кузнецова М.Г., Горбачик Е.С. Введение в теорию живучести вычислительных систем. – К.: Наук. думка, 1990. –184 с.
8. Robert J. Ellison, David A. Fisher, Richard C. Linger, Howard F. Lipson, Thomas A. Longstaff, Nancy R. Mead «Survivable Network Systems: An Emerging Discipline» [Электронный ресурс]. – Режим доступа: // http://www.cert.org\ research\97tr013.pdf
9. Robert J. Ellison, David A. Fisher, Richard C. Linger, Howard F. Lipson, Thomas A. Longstaff, Nancy R. Mead «Survivability: Protecting Your Critical Systems» [Электронный ресурс]. – Режим доступа: // http://www.cert.org\archive\html\protect-critical-systems.html

ДО ДЕЯКИХ АСПЕКТІВ ФОРМУЛЮВАННЯ ПОНЯТТЯ «ІНФОРМАЦІЙНІ ЗАГРОЗИ»

Головко О.М.,

Науково-дослідний інституту інформатики і права

НАПрН України

Формування інформаційного суспільства (далі – ІС) стає не просто фактом, а тому все більше впливає на формування державної політики інформаційної безпеки. Досягнення тих чи інших цілей виявилося можливим із застосуванням лише інформаційних технологій, які чинили б вплив на суспільну свідомість.

Безперечно, якщо ми говоримо про інформаційну війну, то вочевидь зрозумілим є той факт, що даний термін найбільш споріднений із воєнними діями. Тому, коли йдеться про інформаційну війну, то слід говорити про існування рішучої та небезпечної діяльності, пов'язаної із реальними бойовими діями. Отже, існування розвиненої системи інформаційної безпеки закладе фундамент для нормального існування людини в медіапросторі.

Цілі інформаційної війни є дещо іншими, аніж війни у звичному розумінні: не фізичне знищення противника та ліквідація його збройних сил, а широкомасштабне порушення роботи фінансових, транспортних і комунікаційних мереж і систем, руйнування економічної інфраструктури та підкорення населення країни, що зазнала атаки, волі країни-переможця [1, с. 340].

На початку 90-х років термін «інформаційна війна» з'явився у США та активно увійшов в загальносвітову практику. Сьогодні даний термін використовується в двох площинах:

у широкому розумінні – для визначення протиборства в інформаційній сфері, в засобах масової інформації для досягнення різних політичних цілей;

у вузькому розумінні – для визначення воєнного протиборства, у військовій інформаційній сфері для досягнення односторонніх переваг в отриманні зборі, обробці та використанні інформації на полі бою (в операції, битві).

У вітчизняній практиці в широкому розумінні найчастіше використовують термін «інформаційне протиборство»; у вузькому розумінні – «інформаційні воєнні дії».

Найбільш доречною видається позиція, відповідно до якої інформаційне протиборство – це форма боротьби сторін в інформаційному просторі з використанням політичних, економічних,

дипломатичних, військових та інших методів, способів та засобів впливу на інформаційне поле супротивника, а також захисту власного інформаційного поля в інтересах досягнення поставлених цілей.

При цьому, до «інформаційної зброї» варто відносити, по-перше, засоби інформаційно-технічного характеру, які знищують, перекручують або викрадають інформацію, не зважаючи на систему захисту, обмеження доступу до цієї інформації законних користувачів. По-друге, це безперечно, інформаційно-психологічні засоби, які дезорганізують інформаційні системи шляхом дезінформації, формування помилкових логічних інформаційних концепцій, інтерпретацій та ін., впливаючи таким чином на думку суспільства та на функціонування органів виконавчої влади.

Таким чином, інформаційна зброя – це сукупність способів, прийомів, засобів і технологій інформаційного впливу на інформаційну інфраструктуру, інфраструктуру ворожої держави та психіку, свідомість і підсвідомість її населення та особливого складу збройних сил.

Підсумовуючи, зазначимо, що інформаційні війна, інформаційне протиборство й інформаційна боротьба є проявами одного більш широкого поняття – загрози інформаційній безпеці. Саме цьому питанню – загрозам інформаційній безпеці людини в медіапросторі ми і приділимо увагу далі.

Слід зазначити, що аналізу змісту поняття «інформаційна безпека» зазвичай дослідники приділяють значну увагу, у той час як поняття «небезпека» і «загроза» розглядаються дещо спрощено і здебільшого у звуженому плані, відірваному від контексту поняття «інформаційна безпека».

На нашу думку, необхідність у розробці поняття «загроза» визначається:

1) відсутністю єдиного підходу до дослідження основних понять інформаційної безпеки;

2) недостатньою розробкою родового поняття «загроза» і питань його відмежування від інших споріднених понять, таких як «небезпека», «виклик», «ризик», відповідно і видового «інформаційна загроза» та його відмежування від таких понять, як «інформаційна війна», «інформаційне протиборство», «інформаційний тероризм»;

3) невирішеністю проблеми формування категорійно-понятійного апарату теорії інформаційної безпеки;

4) можливістю на підставі теоретичних розробок даного апарату формувати адекватну систему моніторингу та управління загрозами та небезпеками в інформаційній сфері.

В економічній літературі існує багато визначень поняттю «загрози», до найбільш розповсюджених належать:

– загрози як умови, які виникають або спричиняють вияв причин загрози для стратегічних можливостей підприємства [2, с. 52];

– загроза – це такий розвиток подій, внаслідок яких збільшується можливість або з'являється вірогідність порушення нормального функціонування підприємства та заподіяння збитків [3, с. 93];

– загроза як зміни в зовнішньому або внутрішньому середовищі, які приводять до небажаних змін предмета безпеки (підприємства)" [4, с. 45];

– загроза як найбільш конкретна і безпосередня форма небезпеки або сукупність умов і чинників, що створюють небезпеку інтересам різних суб'єктів" [5, с. 41; с. 52; с. 157];

– загрози як реальні чи потенційно важливі дії або умови навмисного чи випадкового (ненавмисного) порушення режиму функціонування підприємства шляхом заподіяння матеріальних (прямо або непрямо) збитків, що призводить до фінансових втрат, зокрема і до втрати вигоди [6, с. 10];

– загрози – це потенційні або реальні умови, чинники чи дії фізичних та юридичних осіб, що порушують нормальний фінансово-економічний стан суб'єктів підприємницької діяльності і здатні заподіяти великої шкоди аж до припинення його діяльності [5, с. 41; с. 52; с. 157];

– загроза як потенційна можливість завдання шкоди суб'єктам господарюючої діяльності з боку окремих чинників внутрішнього та зовнішнього середовища, тобто поява загрози визначає потенційну чинників економічних втрат для підприємства [7, с. 10];

– загроза – ще одна форма небезпеки – небезпека на стадії переходу з можливості у дійсність як наявна чи потенційна демонстрація готовності: відносно суб'єктів господарської діяльності – одних суб'єктів завдати шкоду іншим або по відношенню до процесів, явищ – негативно вплинути на господарську діяльність підприємства [8, с. 66];

– загроза – це зафіксований фірмою екзогенний чинник потенційно негативної дії [9, с. 37].

Майже всі вчені сходяться на думці, що загрози це: події, зміни або дії, тобто загрозам притаманна динаміка; вони спричиняють шкоду людині; загрози виникають під дією певних чинників (зовнішніх та внутрішніх), і тому потребують комплексу заходів з боку людини для їх нейтралізації та усунення.

Зазначимо, що в Україні поки відсутня розроблена концепція (система теоретико-методологічних засад, положень) забезпечення інформаційної безпеки. Більш того, аналіз сучасної геополітичної обстановки дає нам усі підстави зробити висновок, що проти України в режимі реального часу застосовується інформаційна зброя, спрямована на дискредитацію, дезорганізацію, підрив іміджу та дестабілізацію нашої держави. І передусім цей вплив чиниться на діяльність цілого ряду силових і фінансових відомств нашої країни, в тому числі шляхом широкомасштабного викривлення громадської думки психоемоційними інформаційними кампаніями.

Отже, такі поняття як «небезпека», «загроза» та «ризик» є дійсно взаємопов'язаними, а їх виникнення у більшості випадків спричинено невизначеністю зовнішнього середовища. В нашому дослідженні під цими поняттями ми будемо розуміти: небезпека – це об’єктивно існуюча реальність, яка може порушити стан рівноваги людини і призвести до негативних наслідків; загроза – це наслідок небезпеки у вигляді об’єктивного чинника потенційно негативної дії; ризик – це об’єктивно-суб'єктивна категорія, що пов'язана з певною мірою невизначеності результату внаслідок прийнятого рішення (дії і/або обставин).

Література

1. Почепцов Г.Г. Информационные войны / С.Л. Удовик (отв. ред.). – М.: Рефл-бук, 2000. — 576 с.
2. Клейнер Г. Б. Предприятия в нестабильной экономической среде: риски, стратеги, безопасность : [учебник] / Г. Б. Клейнер, В. Л. Тамбовцев, Р. М. Качалов. — М.: ОАО “Изд-во “Экономика”, 1997. — 288 с.
3. Гапоненко В. Ф. Экономическая безопасность предприятия: подходы и принципы / В. Ф. Гапоненко, А. А. Безпалько, А. С. Власков. — М.: Ось-89, 2006. — 208 с.
4. Тамбовцев В. Л. Объекты экономической безопасности России / В. Л. Тамбов- цев // Вопросы экономики. — 1994. — № 12. — С. 45–54.
5. Камлик М. І. Економічна безпека підприємницької діяльності. Економіко- правовий аспект / М. І. Камлик. — К. : Атіка, 2005.

— 432 с.; Мак-Мак В. П. Служба безопасности предприятия как субъект частной правоох- ранительной деятельности: дисс. ... канд. юрид. наук / В. П. Мак-Мак. — М., 2003. — 210 с. ; Дубецька С. П. Економічна безпека підприємств України / С. П. Дубецька // Не- державна система безпеки підприємництва як суб'єкт національної безпеки України : зб. матер. наук.-практ. конф., Київ, 16-17 травня 2001 р. — К.: Вид-во Європ. ун-ту, 2003. — С. 146-171.

6. Ярочкин В. И. Предприниматель и безопасность : Ч. 2. / В. И. Ярочкин. — М.: Изд- во «Экспертное бюро», 1994. — 112 с.
7. Основи економічної безпеки : [підручник] / [Бандурка О. М., Духов В. Є., Петрова К. Я., Червяков І. М.]. — Харків: Вид-во нац. ун-ту внутр. справ, 2003. — 236 с.
8. Горячова К. Фінансова безпека підприємства. Сутність та місце в системі еко- номічної безпеки / К. Горячова // Економіст. — 2003. — №8. — С. 65–67.
9. Королев М. И. Экономическая безопасность фирмы: теория, практика, выбор стратегии : [монография] / М. И. Королев. — М.: Экономика, 2011. — 284 с.

ЗАСТОСУВАННЯ ВЕЙВЛЕТ АНАЛІЗУ В ЗАДАЧАХ РОЗРІЗНЕННЯ СИГНАЛІВ, ДІАГНОСТУВАННЯ ТА ПРОГНОЗУВАННЯ ВІДМОВ

Г.Б. Жиров,

Військовий інститут КНУ ім. Т.Шевченка

На сьогоднішній час коло прикладних задач, які необхідно розв'язувати технічними засобами швидко зростає. До таких задач відносяться задачі розпізнавання, фільтрації, класифікації, діагностування, прогнозування і багато інших. Вирішення наведених задач базується на обробці сигналів, причому, проблема високоточної та високошвидкісної обробки сигналів та процесів у реальному масштабі часу на основі нових математичних алгоритмів стоїть досить гостро. Результати обробки повинні задовольняти багатьом умовам, включаючи швидкодію, достовірність та іншім умовам.

В кінці минулого століття виник і успішно розвивається новий важливий напрямок в теорії і практиці обробки сигналів, зображень, часових рядів та ін., що отримав назву вейвлет-перетворення (ВП), який добре пристосований для дослідження структури неоднорідних процесів. Термін вейвлет (wavelet) запропонував у своїй статті Гроссманн (Grossmann) і Морлі (Morlet) в середині 80-х років XX століття в зв'язку з аналізом властивостей сейсмічних і акустичних сигналів. Їх робота була початком інтенсивного дослідження вейвлетів в наступне десятиліття такими вченими, як Добеши (Dobechies), Мейер (Meyer), Малл (Mallat), Фарж (Farge), Чуи (Chui) та ін. [1].

При застосуванні вейвлет-аналізу (декомпозиції) процесу (або сигналу) при зміні масштабу, вейвлети здатні виявити відмінності в характеристиках процесу на різних шкалах, а за допомогою зсуву можна проаналізувати властивості процесу в різних точках на всьому досліджуваному інтервалі. Саме завдяки властивості повноти цієї системи, можна здійснити відновлення (реконструкцію) процесу за допомогою зворотного ВП. [2]

Підтвердженням значущості ВП є і той факт, що алгоритми ВП широко представлені у відомих математичних пакетах, таких як Mathcad, MATLAB, Mathematica. Міжнародні стандарти JPEC-200, MPEG-4 і графічні програмні засоби Corel Draw 9/10 широко використовують ВП для обробки зображень і, зокрема, для стиснення зображень для каналів з обмеженою пропускною

спроможністю, наприклад, для Інтернет. Крім того, фірмою Analog Devices розроблені і випускаються недорогі однокристальні мікропроцесори ADV6xx (ADV601, ADV601LC, ADV611, ADV612), засновані на ВП і призначені для стиснення і відновлення відеоінформації в реальному масштабі часу [1].

В доповіді розглядаються питання аналізу структури нестаціонарних сигналів за допомогою методів вейвлет перетворення. Узагальнено підходи до аналізу тонкої структури сигналу для вирішення практичних задач прогнозування відмов технічних об'єктів, а також розрізнення повітряних цілей засобами радіолокації.

В основі вейвлет-аналізу лежить теоретично обґрунтована можливість подання сигналу кінцевої енергії $S(t)$ у вигляді [3]:

$$S(t) = \sum_l a_l \varphi_l(t) + \sum_k d_k \psi_k(t) \qquad (1)$$

де $\varphi_l(t)$ - масштабуюча функція (phi-функція), визначає грубе наближення сигналу та породжує коефіцієнти апроксимації – a_l; $\psi_k(t)$ - вейвлет-функція (psi-функція), визначає деталі сигналу та породжує коефіцієнти деталізації – d_k.

На теперішній час, велика доля сучасних систем обробки використовують результати комп'ютерних розрахунків в яких використовується дискретне вейвлет-перетворення (ДВП або DWT – discrete wavelet transform). При застосуванні ДВП використовуються ортонормовані базисні вейвлет-функції. Для кожного рівня розкладання (декомпозиції) *m* існує система функцій:

$$\varphi_{m,k}(t) = \sqrt{2^m}\varphi(2^m t - k)\ ;\ \psi_{m,k}(t) = \sqrt{2^m}\psi(2^m t - k)\,. \qquad (2)$$

Число *m* характеризує рівень роздільної здатності. Чим більше *m*, тим більш дрібні носії має функція $\varphi_{m,k}(t)$ і коефіцієнти розкладання $\left(S(t), \varphi_{m,k}\right)$, та більш детально відображають властивості сигнальної функції $S(t)$.

Із зростанням рівня *m*, оператори проектування P_k дають більш точне наближення $P_k(S)$ до елементів функції $S(t) \in L^2(R)$.

$$P_m(S) = \sum\nolimits_{k \in Z} (S, \varphi_{m,k})\, \varphi_{m,k}(t)\,. \qquad (3)$$

З урахуванням ортогонального доповнення до підпростору V_m розкладання сигнальної функції визначається виразом

$$P_m(S) = \sum_{k\in Z} a_{m-1,k}\ \varphi_{m-1,k}(t) + \sum_{k\in Z} d_{m-1,k}\ \psi_{m-1,k}(t)\,, \qquad (4)$$

де $a_{m-1,k} = \left(S(t), \varphi_{m-1,k}\right) = \int S(t)\varphi_{m-1,k}(t)dt$ – коефіцієнти апроксимації (m-1)-го рівня декомпозиції; $d_{m-1,k} = \left(S(t), \varphi_{m-1,k}\right) = \int S(t)\psi_{m-1,k}(t)dt$ – коефіцієнти деталізації того ж рівня.

Якщо застосовувати швидке дискретне вейвлет-перетворення, то

$$a_{m-1,k} = \left(S, \varphi_{m-1,k}\right) = \sum_n h_n a_{m,n+2k},\ \ d_{m-1,k} = \left(S, \psi_{m-1,k}\right) = \sum_n g_n d_{m,n+2k}\,, \quad (5)$$

де $\{h_n\}_{n\in Z}$ та $\{g_n\}_{n\in Z}$ – коефіцієнти або фільтри вейвлетів $\varphi(t)$ та $\psi(t)$.

Отримані вектори коефіцієнтів прийнято позначати символами:

$$cA_1 = \{a_{m-1,k}\} \text{ и } cD_1 = \{d_{m-1,k}\}. \qquad (6)$$

При повторенні процедури розкладання за рівнем, до $m = N$, отримуємо кінцеве представлення сигналу у вигляді серії коефіцієнтів, тобто дискретне вейвлет перетворення (ДВП):

$$P_m(S) = \{cA_N, cD_N, ..., cD_1\}, \qquad (7)$$

де cA_N – коефіцієнти апроксимації розкладання (декомпозиції) глибини N;

де cD_m – коефіцієнти деталізації розкладання (декомпозиції) глибини $m = \overline{1, N}$;

$$\varphi_{m-1,k} = \sum_n h_n \varphi_{m,n+2k},\quad \psi_{m-1,k} = \sum_n g_n \psi_{m,n+2k}\,,$$

де $\{h_n\}_{n\in Z}$ та $\{g_n\}_{n\in Z}$ – коефіцієнти або фільтри вейвлетів $\varphi(t)$ та $\psi(t)$.

Таким чином, отримуємо формули швидкого ДВП [5]:

$$\begin{aligned} a_{m-1,k} &= \left(S, \varphi_{m-1,k}\right) = \sum_n h_n a_{m,n+2k}\,, \\ d_{m-1,k} &= \left(S, \psi_{m-1,k}\right) = \sum_n g_n d_{m,n+2k} \end{aligned}\ . \qquad (8)$$

Наведені алгоритми не тільки більш швидкі, з точки зору використання алгоритмічних процедур, але і при кожному перетворенні загальне число нових коефіцієнтів не збільшується у 2 рази, а залишається незмінним.

Процедуру можна повторити і таким чином визначити усі коефіцієнти $\{cA_N, cD_N, ..., cD_1\}$ вейвлет перетворення N-го рівня. На

практиці найменший можливий масштаб (найбільший рівень декомпозиції) визначається числом N дискретних значень сигналу.

Застосування вейвлет аналізу вирішує багато практичних завдань, одним з яких є прогнозування відмов пристроїв різного виду (електромеханічних, радіоелектронних, газотурбінних та ін.). Багатомасштабний метод вейвлет аналізу виявляється високоефективним методом аналізу складних фізичних сигналів на різних масштабах та у спеціально визначених точках (контрольних точках). Такий метод дозволяє виявити відмову значно раніше ніж при застосуванні інших методів [2].

Загальний підхід до прогнозування відмов з застосуванням вейвлет-аналізу та статистичної обробки результатів можна звести до наступної методики:

1.Визначається параметр (параметри), зміна якого (яких) призводить до відмови (руйнуванню) пристрою.

2. За допомогою датчиків значення обраного параметру перетворюється в цифрову форму.

3. Застосовується вейвлет-аналіз та аналізуються значення трансформованого параметру.

3.1. Для кожного виду об'єкту техніки та прогностичного параметру визначається вид материнського вейвлету.

3.2. Вейвлет аналіз проводиться при різних режимах роботи об'єкту техніки до настання моменту відмови. Час проведення аналізу визначається окремо для кожного виду техніки.

3.3. Проводиться статистична обробка результатів вейвлет аналізу. Оскільки, в більшості випадків досліджуємий сигнал флуктуює у часі, флуктуюють і його вейвлет коефіцієнти. Статистичною оцінкою флуктуацій можуть виступати дисперсії розподілів на різних масштабах: $\sigma(m,l)=\sqrt{\frac{1}{l-1}\sum_{k=0}^{l-1}[d_{m,k}-\overline{d_{m,k}}]^2}$, де l – кількість вейвлет коефіцієнтів на рівні m у визначеному для аналізу інтервалі часу.

4. Різка зміна (зменшення або збільшення) значення дисперсії говорить про знаходження об'єкту у передвідмовному стані.

Практичні досліди та об'єм опрацьованої літератури говорить, що: час до відмови залежить від об'єкту та прогностичного параметру; найбільш інформативні показники дисперсії, як правило, виявляються при аналізі вейвлет коефіцієнтів вищих масштабів.

Також, в якості статистичної оцінки можна використовувати середнє значення суми деталізуючих вейвлет коефіцієнтів:

$K_{m,k} = \frac{1}{D}\sum_{m=1}^{N}\sum_{k\in Z} d_{m,k}$, або проводити аналіз коефіцієнтів $K_{1,k}$, $K_{2,k}$, $K_{N,k}$ кожного рівня декомпозиції [4], де D – загальна кількість коефіцієнтів вейвлет аналізу.

Наступна прикладна задача, це задача розрізнення повітряних цілей засобами радіолокації.

Такий метод заснований на процедурі розпізнавання тонкої структури сигналу складної цілі, характерною ознакою якого, є наявність стрибків сигнальної обвідної в області перекриття сигналів від одиночних цілей за рахунок інтерференції, обумовленої випадковими змінами фази [5, 6]. В якості відбитого сигналу розглядається та аналізується сигнал у вигляді цифрової біквантованої пачки.

В узагальненому вигляді, задачу розрізнення можна звести до наступної методики:

1. Виділяється обвідна функція пачки відбитих імпульсів від цілі (цілей).

2. Перетворюється обвідна функція в цифрову біквантовану пачку імпульсів.

3. Отриманий сигнал розкладається на вектори апроксимуючих та деталізуючих коефіцієнтів $\{cA_N, cD_N, ..., cD_1\}$ до 4 рівня декомпозиції. Найбільш інформативними виявляються коефіцієнти декомпозиції 1-го нижнього рівню.

5. Вектор деталізуючих коефіцієнтів $\{cD_1\}$ 1-го рівня являє собою дискретну випадкову величину, яку можна охарактеризувати числовим статистичним параметром. Найбільш інформативним параметром було визначено [5] квадрат норми $norm^2(cD_1)$ вектору коефіцієнтів ВП нижнього рівня.

6. Порівнюючи статистичну оцінку з вибраним порогом, отримуємо рішення «ціль одинична», або «ціль групова».

Широкий спектр практичних задач базується на дослідженні та аналізі різноманітних сигналів, фізична природа яких може бути різною. Обробка сигналів із застосуванням вельвет перетворення виявляється більш ефективним у порівнянні з традиційним перетворенням Фур'є в задачах аналізу нестаціонарних сигналів, неоднорідних полів, зображень різної природи та ін.

Задача вибору того чи іншого вейвлету носить суб'єктивний характер, в основному базується на практичному досвіді науковця (інженера) та вимагає дослідження для кожного конкретного

випадку.

Найбільш інформативним, виявилися деталізуючи коефіцієнти вейвлет-аналізу нижніх рівнів декомпозиції.

Література

1. Яковлев А.Н. Введение в вейвлет-преобразования: Учеб. пособие / А.Н. Яковлев. – Новосибирск: НГТУ, 2003. – 104 с.
2. Дремин И.М. Вейвлеты и их использование / И.М. Дремин, О.В. Иванов, В.А. Нечитайло // Успехи физических наук. – 2001. – т. 171. – № 5. – С. 465–501.
3. Solonyna A., Arbuzov A. Cyfrovaja obrabotka sygnalov. Modelyrovanye v MATLAB (2008), S. – SPb.: BHV. Peterburg, 816.
4. Загірняк М.В. Діагностика пошкоджень стрижнів ротора асинхронних двигунів за аналізом електрорушійної сили в обмотках статора / М.В. Загірняк, Ж.І. Ромашихіна, А.П. Калінов // Електротехніка і Електромеханіка. – 2014. – №6. – С.34–42.
5. Долгушин В.П. Расспознавание класса целей методом оценки статистических параметров вектора вейвлет-декомпозиции сигнала / В.П.Долгушин, В.Н. Лоза, А.Н. Борзак, Б.Г. Жиров // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – 2014. – №45. – С.24–33.
6. Долгушин В.П. Статистический анализ спектрально-временных параметров эхо-пачки сосредоточенной парной цели / В.П.Долгушин, Е.С.Ленков, В.Н.Лоза, Р.Ю. Кольцов // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – 2014. – №43. – 35–44 с.

МОЖЛИВОСТІ ТА ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ЕКСПЕРТНИХ ТЕХНОЛОГІЙ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ У СФЕРІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Каденко С.В.,
Інституту проблем реєстрації інформації НАН України

Вступ

Протягом останніх десятиліть актуальність інформаційної безпеки як галузі національної безпеки стрімко зростала. Можна згадати, що інформаційні впливи були дієвим засобом політичної боротьби та пропаганди ще з давніх часів. Але в умовах сучасних реалій інформаційна безпека є актуальною як ніколи. Успіх в інформаційній війні є невід'ємною, навіть необхідною умовою успіху у війні звичайній. Прикладом можуть служити події в Україні останніх років. Через це інформаційна безпека, зокрема, попередження інформаційних операцій та впливів має бути обов'язковою складовою діяльності та важливим стратегічним пріоритетом будь-якої великої організації.

У даній доповіді пропонується дотримуватися визначення інформаційної операції, наведеного у [1]. За цим визначенням інформаційна операція, зазвичай, покликана вносити в свідомість суспільства та окремих людей певних ідей та поглядів, дезорієнтувати та дезінформувати реципієнтів інформації, послаблювати переконання громадян та суспільства взагалі, а іноді – залякувати маси.

Планування заходів з підсилення інформаційної безпеки, попередження негативних/ворожих інформаційних впливів та інформаційних операцій (а також планування успішних операцій у процесі інформаційної боротьби) вимагає чіткого розуміння та достеменного знання предметної області. Втім, сфера інформаційної безпеки являє собою слабко структуровану предметну область, яка важко піддається опису, зокрема, формальному, кількісному. Технології експертної підтримки прийняття рішень якраз являють собою засіб розв'язання задач в слабко структурованих предметних областях [2]. Тому, на думку автора, питання використання експертних технологій підтримки прийняття рішень в області інформаційної безпеки та конкретний контекст застосування окремих методів заслуговують на окремий розгляд. Такий розгляд пропонується у даній доповіді.

Інформаційна безпека як слабко структурована предметна область

Рисунок 1 – Ознаки слабко структурованих предметних областей

У праці [3] наводяться наступні ознаки слабко структурованих предметних областей (рис. 1): відсутність цілі функціонування, яка б піддавалася формалізації, відсутність оптимальності, унікальність, динамічність, неповнота опису, наявність суб'єктивного людського фактору, неможливість побудови аналітичної моделі, відсутність еталонів, велика розмірність.

У [1] зазначається, що на інформаційні операції впливає багато суто якісних (зокрема, соціально-психологічних) критеріїв, факторів та параметрів. Ці фактори важко піддаються формальному математичному (аналітичному) опису.

Там же йдеться про неможливість розробки та практичного застосування певної універсальної методики моделювання інформаційних операцій, насамперед, внаслідок слабкої формалізації понять і факторів. Автори наголошують, що у кожному випадку слід консультуватися з аналітиками, тобто експертами з аналізу інформаційних операцій, покладатися на їхню компетентність. При цьому іноді аналітики в змозі побудувати точні прогнози певних закономірностей, які потім підтверджуються практикою. До

аналітиків-експертів слід звертатися для опису суб'єктивних факторів. Коли ж йдеться про фактори об'єктивні, то їхній опис та аналіз пропонується здійснювати за допомогою відомих методів, які оперують детермінованими даними, зокрема, методів математичної статистики та аналізу часових рядів. Втім, ці методи спрямовані лише на опис формальних аспектів і не торкаються аспектів змістовних. З огляду на це, у [1] констатується необхідність розширення технологічного інструментарію, який може використовуватися для аналізу та моделювання інформаційних операцій.

Як бачимо, інформаційні операції (як, до речі, і будь-які операції, що передбачають суттєву людську участь) являють собою яскравий приклад слабко структурованої предметної області. Одним з технологічних засобів аналізу та моделювання інформаційних операцій, на думку автора даної доповіді, мають стати експертні технології підтримки прийняття рішень. На доцільність використання експертних знань в слабко структурованих предметних областях вказують також дослідження Delphi Group щодо складу знань організацій [4] (рис. 2).

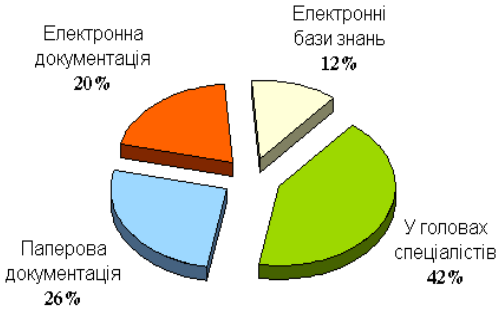


Рисунок 2 – Джерела знань про організацію, згідно з дослідженням Delphi Group

В результаті досліджень з'ясувалося, що значна частина знань знаходиться не в базах знань, або на паперових, чи електронних носіях, а саме в головах експертів (аналітиків, спеціалістів). Отже, в контексті опису та аналізу інформаційних операцій, безперечно, має сенс залучати експертні знання, особливо, коли йдеться про якісні фактори суб'єктивного характеру.

Ієрархічна декомпозиція та цільове динамічне оцінювання альтернатив

У [1] зазначено, що інформаційна операція являє собою міждисциплінарний набір методів та технологій, який охоплює багато сфер, від військової науки до соціології. При цьому не існує універсальної, стандартної технології проведення інформаційних операцій (яка могла б послужити не тільки військовим, а й керівникам великих урядових чи комерційних організацій). Отже, як зазначають автори праці, актуальним питанням лишається розробка наукової бази інформаційних операцій.

З урахуванням вищезгаданого міждисциплінарного характеру, на думку автора, для опису інформаційних операцій та впливів зручним інструментом має стати експертна ієрархічна декомпозиція. Зокрема, даний підхід складає основу методу цільового динамічного оцінювання альтернатив (МЦДОА), запропонованого В.Тоценком [2] та удосконаленого В.Циганком [5]. Метод якраз дозволяє об'єднати в єдину ієрархію (рис. 3) велику кількість критеріїв та заходів різної природи (з різних дисциплін), які впливають на певну головну ціль.

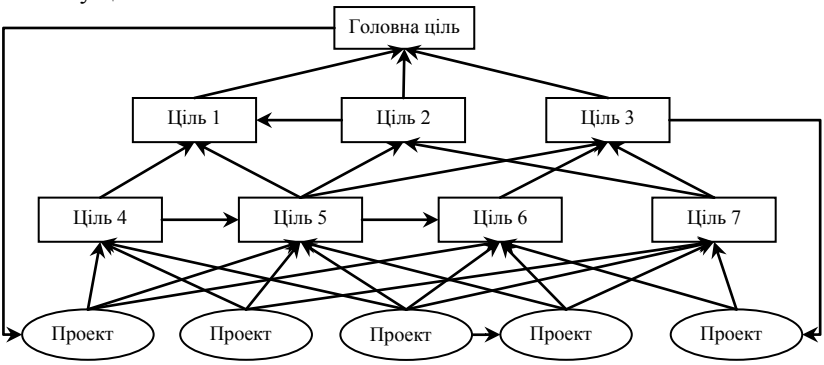


Рисунок 3 – Приклад графу ієрархії цілей

В залежності від типу конкретної інформаційної операції (наступального чи оборонного [1]), аналітик-експерт (або сама особа, що приймає рішення (ОПР)) може відповідним чином сформулювати головну ціль. Будь-яка інформаційна операція декомпозується на певні етапи або кроки. У [1] наводяться їхні детальні переліки. Зміст цих кроків може варіюватися в залежності, знов-таки, від типу операції та специфіки контексту. Наприклад, якщо йдеться про моделювання та декомпозицію (в рамках МЦДОА)

інформаційної атаки на академію наук (головна ціль), то ціль «Дискредитація наукової установи у ЗМІ» може декомпозуватися на цілі нижчого рівня «Дискредитація наукових праць та досягнень» та «Дискредитація наукових співробітників».

У МЦДОА декомпозиція відбувається до рівня атомарних підцілей (факторів, критеріїв), на які може впливати ОПР. Ці цілі називаються проектами (див. Рис. 3), і, зазвичай, можуть бути охарактеризовані певною величиною (абсолютною чи відносною чисельною, або логічного/булевського чи порогового типу).

Загальною принциповою задачею методів підтримки прийняття рішень, які передбачають ієрархічну декомпозицію задачі, зокрема МЦДОА та методів аналізу ієрархій та мереж (AHP/ANP) [6] є побудова рейтингу або ранжирування кількох об'єктів (альтернатив, проектів). На основі такого рейтингу ОПР може зробити обґрунтований вибір найкращої альтернативи (варіанту рішення) з заданої множини, або розставити пріоритети у своїй діяльності (тобто з'ясувати, які фактори чи заходи є найважливішими для досягнення заданої головної мети). Для побудови такого рейтингу необхідно визначити відносну важливість усіх цілей, що входять до графу ієрархії цілей, побудованого експертами (або інженерами зі знань на основі діалогів з експертами). Щоб визначити відносну важливість підцілей окремої цілі (її «нащадків» у графі ієрархії), експерти мають попарно порівняти їх між собою. Оцінки впливів (ваг) можуть здійснюватися експертами у різних шкалах парних порівнянь. Нещодавні дослідження В.Циганка [7,8] показали, що експерту слід надавати можливість вводити значення кожного окремого парного порівняння у найбільш зручній для нього шкалі. Коли експертами оцінені усі відносні ваги цілей у графі ієрархії, можна розрахувати відносний вплив кожної цілі та проекту на головну ціль (відносну ефективність), як показано в [2].

Якщо оцінки здійснюються кількома експертами, то слід враховувати кілька важливих аспектів. Першим з них є компетентність експертів. Якщо відомо, що експерти, які оцінюють критерії, проекти, чи альтернативи, мають різну компетентність, то її слід визначати на основі кількох складових: самооцінки, взаємної оцінки компетентності членів експертної групи, та об'єктивної компоненти (як показано у [2]). Різницею у відносній компетентності експертів можна знехтувати лише у випадку, коли розмір експертної групи є достатньо великим [9]. Другим важливим аспектом є узгодженість оцінок, отриманих від різних членів експертної групи.

Оцінки слід перевіряти на узгодженість, адже рекомендації, видані ОПР, на основі неузгоджених експертних даних закономірно викликатимуть недовіру. Для оцінки узгодженості експертних оцінок доцільно застосовувати так звані спектральні методи, описані, зокрема, у [10] та [11]. Перевага спектральних методів над іншими підходами до оцінки узгодженості (зокрема, тими, що запропоновані Сааті та колегами [12]) – наступна. За необхідності (якщо рівень узгодженості експертних оцінок у групі є низьким), спектральні методи дозволяють організувати конструктивний покроковий зворотній зв'язок з експертами. Тобто, експертам пропонується змінити відповідні оцінки («викиди») таким чином, щоб рівень узгодженості підвищився до необхідної величини. Коли оцінки, надані різними членами експертної групи досягають достатньо високого рівня узгодженості, їх можна агрегувати (отримувати на їхній основі узагальнені групові оцінки). Для агрегації експертних оцінок доцільно застосовувати комбінаторний метод агрегації [13]. З-поміж численних переваг цього методу слід назвати можливість його використання для агрегації неповних матриць парних порівнянь та максимальне використання надлишковості експертної інформації.

В контексті МЦДОА «зважена» ієрархія критеріїв (цілей) називається базою знань (БЗ) про предметну область. В даній доповіді увагу зосереджено на предметних областях, що стосуються інформаційної безпеки, зокрема, інформаційних операцій. За змістом, така БЗ являє собою один з типів моделей предметної області. БЗ будується експертами (або інженерами по знаннях у ході діалогів з експертами) за допомогою спеціальних програмних засобів – автоматизованих систем підтримки прийняття рішень (СППР).

Зазначимо, що МЦДОА не вимагає, щоб усі дані, які вводяться до БЗ СППР, були виключно експертними оцінками. Наприклад, якщо йдеться про порівняння кількох альтернатив за певним критерієм, то значення оцінок не обов'язково мають виражатися у шкалі парних порівнянь. Іноді це можуть бути абсолютні значення, доступні з відкритих джерел. Припустімо, для аналізу інформаційної політики чи кампанії часто використовуються абсолютні показники, такі як «кількість публікацій з негативним забарвленням протягом місяця». Такий показник цілком правомірно може бути включеним до ієрархії критеріїв, яка описує інформаційну політику організації.

Особливості роботи з експертом

Експерт (аналітик, спеціаліст), зазвичай, є представником вузько-орієнтованої предметної області, і, в загальному випадку, не орієнтується в технологіях і методах підтримки прийняття рішень. Тому процес отримання інформації від експерта має бути максимально комфортним для нього. Формальну сторону процесу слід делегувати інженеру по знаннях, а математичні обчислення – автоматизованій СППР. Для досягнення цієї мети в процесі експертизи доцільно враховувати кілька особливостей.

По-перше, якщо експерт не знайомий з технологіями підтримки прийняття рішень, його має сенс ознайомити хоча б з загальним ходом експертизи. В ідеальному випадку учасникам експертизи слід пояснити усю технологію, за допомогою якої оброблятимуться оцінки та формуватимуться рекомендації для ОПР щодо вибору варіанту рішення. Це підвищить рівень довіри експертів до процесу та дасть їм змогу вводити дані в СППР у прийнятному форматі. Отже, перед тим, як починати збір інформації у експертів, доцільно провести з ними ознайомчі тренінги (coaching sessions).

По-друге, ієрархія критеріїв (цілей) (рис. 3) має бути максимально збалансованою. Проекти бажано розташовувати на одному рівні (оскільки, як показано у [6], їхні ваги мають лежати у межах одного порядку). Слід уникати появи великої (більше 7±2) кількості «нащадків» у однієї цілі у графі ієрархії, з огляду на психофізіологічні обмеження людини [14]. Під час побудови ієрархії не бажано ставити експерту багато однотипних питань (зокрема, щодо позитивного чи негативного характеру впливу цілей на спільного «предка» у графі, або щодо попарної сумісності цілей-нащадків).

По-третє, під час оцінки впливів, слід віддавати перевагу не чисельним значенням, а їхнім вербальним еквівалентам (наприклад, «1» – рівнозначність, «2» – слабка перевага, ..., «5» – дуже сильна перевага).

Загалом, інтерфейс автоматизованої СППР має бути максимально дружнім до та зручним для користувача, інтуїтивно зрозумілим йому.

Більш детальний аналіз особливостей роботи з експертом наведений у [15].

Виявлення закономірностей розповсюдження інформаційних впливів на основі наявних даних та досвіду

В контексті дослідження інформаційних операцій автори праці [1] вказують на важливість задачі оцінки параметрів моделі на основі реальної поведінки певної залежності. На основі виявлених закономірностей можна буде прогнозувати перебіг інформаційної операції навіть якщо точне уявлення про нього відсутнє. Більш того, подібний прогноз може виявитися точнішим за дані традиційних експертиз (коли експерти роблять прогнози у вигляді оцінок).

У зв'язку з цим (вже у розрізі технологій підтримки прийняття рішень) хочеться згадати про можливість використання методів факторного аналізу та нейромережевих алгоритмів для визначення параметрів моделі на основі наявної вибірки «входів» та «виходів». На дану можливість вказував В.Тоценко у своїй монографії [2], називаючи конкретні методи факторного аналізу – метод групового урахування аргументів (МГУА), метод найменших квадратів (МНК), метод багатовимірної лінійної екстраполяції, та мінімальних нев'язок.

Слід наголосити, що у якості вхідних даних для роботи методів факторного аналізу можуть використовуватися як детерміновані дані реальної поведінки системи (зокрема, вхідні та вихідні параметри, що характеризують інформаційну операцію), так і дані, отримані експертним шляхом. При цьому, перелічені вище методи факторного аналізу (а також нейромережеві алгоритми) доцільно використовувати, якщо вхідні дані є кардинальними (чисельними). Якщо наявні лише дані ординального характеру (які несуть інформацію тільки про порядок слідування альтернатив у ранжируванні, а не про кількісне співвідношення між ними), то для факторного аналізу слід застосовувати підходи, запропоновані у [16,17]. Так, наприклад, після парламентських виборів, коли рейтинги та підсумкові ранжирування партій вже відомі, можна спробувати визначити відносну вагомість певних пунктів передвиборчих програм партій з огляду на ці рейтинги.

За допомогою методів кардинального та ординального факторного аналізу також можна визначати, які параметри визначають центральність елемента у певній мережевій структурі (мережевим методам та структурам у розрізі інформаційних операцій у [1] присвячено окремий розділ). Таку спробу було зроблено, зокрема, у праці [18].

До того ж, якщо говорити про елементи інформаційної мережної структури, то слід згадати, що вони часто відбивають певні соціальні відносини між членами мережі [1]. Так, учасників співтовариства у соціальній (або терористичній) мережі можуть єднати спільні ідеї, гасла, згадки, посилання, тощо. Визначення подібних зв'язків може здійснюватися на основі змістової подібності контенту, який стосується членів тої чи іншої мережної структури. Навіть якщо явних зв'язків немає, семантичний аналіз відповідного контенту може допомогти виявити латентні, приховані зв'язки. Методам визначення змістової подібності присвячено роботи О.Андрійчука, наприклад [19].

Стратегічне планування

Окрема інформаційна операція може мати потужний вплив. Втім, вона представляє, так би мовити, тактичний рівень. Якщо ж говорити про рівень стратегічний, то слід згадати, що стратегія інформаційної безпеки обов'язково входить до загальної стратегії національної безпеки [1].

Як показано, зокрема, у [20], стратегія може бути представлена у вигляді оптимального на заданий момент часу розподілу обмежених ресурсів між пріоритетними проектами у конкретній сфері. У [21] у якості предметної області, в якій будується стратегія обрано космічну діяльність та виробництво космічної техніки. У [20] прикладом предметної області є оборонна сфера. Аналогічну стратегію можна побудувати і у сфері інформаційної безпеки.

Процес стратегічного планування з використанням технологій багатокритеріальної підтримки прийняття рішень на основі експертних та об'єктивних даних об'єднує усі процедури, перелічені у попередніх розділах даної доповіді. Він включатиме наступні загальні етапи:

1) Формулювання ОПР головної цілі яка характеризує предметну область.

2) Вибір групи експертів (аналітиків, спеціалістів) для участі в експертизі.

3) Побудова в ході діалогів з експертами ієрархії критеріїв (факторів), які впливають на досягнення головної цілі.

4) Оцінка експертами відносних впливів критеріїв (проектів) ієрархії.

5) Розрахунок відносної ефективності проектів, тобто їхніх внесків у досягнення головної цілі.

6) Визначення оптимальної стратегії розвитку.

У даному випадку йдеться про розвиток сфери інформаційної безпеки. Стратегією, як зазначено вище, є такий розподіл обмежених ресурсів між окремими проектами, який забезпечує максимально ефективне досягнення головної цілі стратегії.

Висновок

Внаслідок міждисциплінарного характеру, слабкої формалізованості, наявності людського фактору, та інших причин, інформаційні операції слабко піддаються строгому аналітичному та математичному опису. Втім, такий опис є вкрай необхідним в контексті активної інформаційної боротьби. Поряд з мультиагентним моделюванням (за допомогою клітинних автоматів та інших засобів), більш чітке та формальне уявлення про інформаційні операції та їхній ефект дозволяють отримати експертні технології підтримки прийняття рішень. У той час, як мультиагентні підходи дозволяють досить успішно змоделювати процес розповсюдження інформаційного впливу або ефекту інформаційної операції, методи багатокритеріальної підтримки прийняття рішень на основі експертних та об'єктивних даних мають стати засобом формального опису та аналізу планування та здійснення інформаційних операцій.

Исследования выполнены в рамках проекта Ф73/23558 «Разработка методов и средств поддержки принятия решений при выявлении информационных операций» Государственного фонда фундаментальных исследований Украины.

Література

1. Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія – К., Інтертехнологія, 2009 – 164 с.
2. Тоценко В. Г. Методы и системы поддержки принятия решений. Алгоритмический аспект [Текст]/ В. Г. Тоценко; ИПРИ НАНУ. – К.: Наукова думка, 2002. – 382 с.
3. Таран Т.А. Искусственный интеллект. Теория и приложения / Т.А. Таран, Д.А. Зубов; Восточноукр. нац. ун-т им. Владимира Даля. — Луганск: ВНУ им. В.Даля, 2006. — 239 с.

4. Тузовский А.Ф. Системы управления знаниями (методы и технологии) / А.Ф. Тузовский, С.В. Чириков, В.З. Ямпольский – Томск: Изд-во НТЛ, 2005. – 260 с.
5. Циганок В.В. Удосконалення методу цільового динамічного оцінювання альтернатив та особливості його застосування / В. В. Циганок // Реєстрація, зберігання і обробка даних. - 2013. - Т. 15, № 1. - С. 90-99.
6. Saaty T.L. Relative measurement and its generalization in decision making. Why pairwise comparisons are central in mathematics for the measurement of intangible factors. The Analytic Hierarchy/Network Process / T.L.Saaty // Statistics and Operations Research, Vol. 102(2), 2008 – pp.251-318.
7. Циганок В.В. Агрегація групових експертних оцінок, що отримані у різних шкалах / В.В.Циганок // Реєстрація, зберігання і обробка даних, 2011. – т.13. – №4.– С.74-83.
8. Циганок В.В. Експериментальний аналіз технології експертного оцінювання / В.В. Циганок, П.Т. Качанов, С.В. Каденко, О.В. Андрійчук, Г.А. Гоменюк // Реєстрація, зберігання і обробка даних. – 2012. – Т. 14, № 1. – С. 91-100.
9. Tsyganok, V. Significance of expert competence consideration in group decision making using AHP. / V. Tsyganok, S. Kadenko, O.Andriichuk//International Journal of Production Research. Vol. 50, Issue 17, 2012 – pp. 4785-4792
10. Zgurovsky, M.Z. Group incomplete paired comparisons with account of expert competence / Zgurovsky, M.Z., Totsenko, V.G., & Tsyganok,V.V.// Mathematical and Computer Modelling, 39(4), 2004 – pp.349–361.
11. Olenko, A. Double Entropy Inter-Rater Agreement Indices / Olenko Andriy & Tsyganok Vitaliy // Applied Psychological Measurement 40(1), 2016 – pp. 37–55.
12. Forman, E.Aggregating individual judgments and priorities with the analytic hierarchy process/ Forman, E. & Peniwati, K // European Jouranl of Operational Research, Vol. 108, 1998 – pp.131-145.
13. Циганок В.В. Комбінаторний алгоритм парних порівнянь зі зворотним зв'язком з експертом / В.В.Циганок // Реєстрація, зберігання і обробка даних. – 2000. – Т.2, №2. – С.92-102.
14. Miller G. A. The Magical Number Seven, Plus or Minus Two / G. A. Miller // The Psychological Review, 1956, – vol. 63 – pp. 81-97.
15. Каденко С.В. Проблеми представлення експертних даних у системах підтримки прийняття рішень/ С.В. Каденко // Реєстрація, зберігання і обробка даних, Т. 18 № 3, 2016 – с. 67-74.

16. Каденко С.В. Про один підхід до прийняття кадрових рішень/ С.В.Каденко, В.В.Циганок// Реєстрація, зберігання і обробка даних, Т. 11 № 3, 2009 – с. 66-74
17. Каденко С.В. Определение относительных весов критериев оценки альтернатив на основе четких и нечетких ранжирований/ С. В. Каденко // Проблемы управления и информатики – 2013. - № 1. – с. 41-49
18. Горбов І.В. Визначення потенційних експертних груп науковців у мережі співавторства з використанням методів підтримки прийняття рішень / І. В. Горбов, С. В. Каденко, І. В. Балагура, Д. Ю. Манько, О. В. Андрійчук // Реєстрація, зберігання і обробка даних. – 2013. - 15, № 4. - С. 86-96.
19. Андрійчук О.В. Метод змістової ідентифікації об'єктів баз знань систем підтримки прийняття рішень / О.В. Андрійчук // Реєстрація, зберігання і обробка даних. — 2014. — Т. 16, № 1. — С. 65-78.
20. Циганок В.В. Інструментарій підтримки прийняття рішень як засіб стратегічного планування / В.В. Циганок, С.В. Каденко, О.В. Андрійчук, П.Т. Качанов, П.Д. Роїк // Озброєння та військова техніка – 2015, № 3(7). – С. 59-66.
21. Tsyganok, V.V. Using Different Pair-wise Comparison Scales for Developing Industrial Strategies / Tsyganok, V.V, Kadenko, S.V., & Andriichuk, O.V. // Proc. Int. J. Management and Decision Making, 14(3), 2015 – pp.224-250.

СКОРИНГОВІ ТЕХНОЛОГІЇ ОЦІНЮВАННЯ РИЗИКІВ ШАХРАЙСТВА В БАНКІВСЬКІЙ ДІЯЛЬНОСТІ

Н.В. Кузнєцова,
НТУУ «Київський політехнічний інститут»

Інформатизація сучасного світу за останні декілька років охопила майже всі куточки нашої планети; уявити світ без найсучасніших пристроїв зв'язку – мобільних телефонів, планшетів, ноутбуків, які за рахунок доступу до інтернету надають можливість залишатися на зв'язку і отримувати надзвичайно швидко і вчасно останню інформацію, просто неможливо. Останні тенденції розвитку суспільства підсилюють потребу отримувати лише найнеобхіднішу інформацію за декілька секунд, не перевантажуючись супутньою інформацією, головним стає принцип доступу до інформації лише «одним натисканням пальцю». Зрозуміло, що не уся інформація про роботу підприємства має бути поширеною і доступною для користувачів; частина її є таємною або конфіденційною, і потрапляння такої інформації до відкритих ресурсів одразу робить її легкодоступною для зловмисників, які можуть використати її в подальшому в злочинних цілях. Зокрема, наразі в Україні дедалі зростає процент шахрайських дій зловмисників в банківській сфері. Це і махінації з пластиковими картками, і спроби отримання кредиту за неправдивими документами. Частка таких кредитів вже зараз становить більше 10%. Українські банки давно почали розробляти стратегію боротьби з такими зловмисниками, впроваджуючи так звані «чорні списки» таких клієнтів та розповсюджуючи цю інформацію через відкриті засоби та банки кредитних історій. Однак, незважаючи на розробку таких баз даних, це не знижує частку кредитних договорів, що не планувалися і не були повернуті.

Тому зараз особливо актуальною для банківської сфери стало застосування спеціальних інформаційних технологій, які б дозволили виявити шахрайські плани зловмисників ще на етапі розгляду кредитних заявок. Для цього важливою є задача ризик-менеджменту з розрізнення кредитних ризиків на шахрайства (Fraud) та дефолти (неможливість подальшого виконання своїх кредитних зобов'язань). Разом з формуванням моделей для оцінки кредитоспроможності і виконанням попереднього скорингу (Application Score), моделей оцінки зміни кредитоспроможності існуючих позичальників для проведення скорингу поведінки (Behaviour Score), а також формування моделей оцінки ймовірності,

суми і оптимального способу впливу на поганих позичальників для виконання скорингу по дефолту (Collection Scoring), необхідно розглядати формування моделей виявлення шахрайства для проведення скорингу шахрайства (Application Fraud Score). Зокрема, в SAS Credit Scoring for Banking існує аналітичний модуль формування моделей оцінки кредитоспроможності позичальників та їх сегментації та можливості формування алгоритмів та моделей виявлення шахрайства.

Шахраїв розділяють на три основні групи: «побутові» (індивідуали) шахраї, професіональні шахраї та позичальники, які використовують послуги професіональних шахраїв. Побутові шахраї не повертають кредит через матеріальні труднощі. Ці боржники не будуть ховатися від банку і колекторів, а після суду змушені будуть повернути товар і банки не отримують прибуток від таких кредитів. Самі клієнти будуть визнані фінансово нестабільними і потраплять в чорний список банку, кредитного бюро і не зможуть отримувати кредити в майбутньому. Професійні шахраї весь час змінюють адреси, мобільні телефони, ніде офіційно не працюють. Страждають від таких клієнтів навіть консервативні банки, де діє жорстка політика перевірки клієнтів та їх документів. Третім типом є позичальники, які залучають шахраїв для отримання кредитів для відкриття бізнесу і через якийсь час не в змозі повернути кредит, і врешті-решт самі стають шахраями. Менша частина таких клієнтів – це люди, які співпрацюють з шахраями, що виготовляють неправдиві документи, і ділять з ними прибуток від шахрайства [1, 2].

Засобом зниження кредитних ризиків від шахрайства, особливо для продуктів без застави, є формування та використання моделей виявлення шахрайства Application Fraud Scoring. Застосування аналітичного модуля SAS для формування таких моделей допомагає банку створити послідовну і логічну базу для прийняття рішень, надати працівникам кредитного відділу більш чітку інтуїтивно зрозумілу міру кредитного ризику (рис.1).

Процес скорингового оцінювання може бути представлений у вигляді послідовності кроків:

1)за затвердженою банком скоринговою моделлю на основі анкетних даних заявника та даних по кредиту здійснюється розрахунок скорингового балу клієнта та перевірка його на схильність до шахрайства або банкрутства:

$$Score_i > ApprovalRateMin,$$

де $Score_i$ – скоринговий бал, розрахований для і-го позичальника розробленою скоринговою моделю; $ApprovalRateMin$ – поріг, що встановлюється банком як мінімальне допустиме значення скорингового балу, при перевищенні якого кредит видається позичальнику. Залежно від політики банки та ситуації на ринку поріг може змінюватися.

2) якщо скоринговий бал заявника менше від порогового балу у банку, то приймається рішення про банкрутство. Якщо з'являються свідчення щодо шахрайських або нетипових дій, то приймається рішення щодо відпрацювання дій по запобіганню шахрайству – передається інформація у відповідні органи та інші банки для виявлення та пошуку зловмисника.

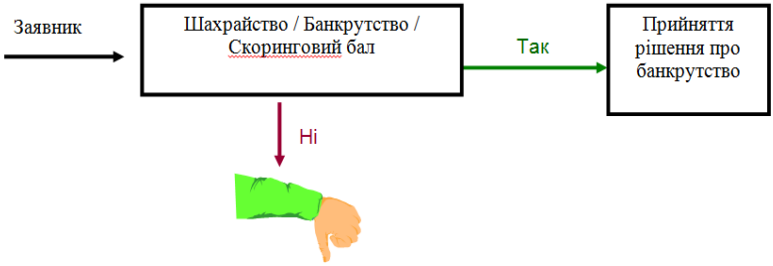


Рисунок 1 – Аналіз та виявлення шахрайських дій в SAS Credit Scoring for Banking

В результаті оцінки історичних даних (навчальної вибірки) формується кредитний портфель потенційного позичальника, що дозволяє розділяти потенційних позичальників на поганих (шахраїв) та «хороших», яким кредит може бути виданий. Цей результат закладається в історичний файл (навчальну вибірку) цільової змінної (target) і будується модель та профіль шахрая. Таким чином претенденти на отримання кредиту ранжуються за групами, кожній з яких присвоюється характеристика надійності позичальника від «високої» до ризикової (шахрай). Зазвичай, оцінка кредитного скорингу шахрая будується на основі 10-12 базових параметрів – сімейний стан, наявність персонального автомобіля, частота зміни роботи, тривалість проживання за останнім місцем тощо. Виходячи з результатів, отриманих за цими критеріями (частина інформації отримується з анкети клієнта, але потім уточнюється та перевіряється службою безпеки банка), система виставляє

потенційному клієнту певну кількість балів. Далі в автоматичному режимі співставляє отриману оцінку із заданим порогом відсікання. Клієнти, у яких оцінки виявились нижчими за порогове значення, не зможуть стати позичальниками банку. Аналіз кредитного скорингу (Score Fraud) дозволяє оцінити профіль шахрая і використовувати його на етапі прийняття рішення про видачу кредиту. Після проведення оцінки ймовірності шахрайства застосовуються традиційні моделі та скорингові карти для оцінки кредитоспроможності позичальника і ймовірності дефолту.

В разі виявлення скоринговою моделлю шахрайських факторів, тобто співпадіння характеристик, що характеризують зловмисників, інформація про такого позичальника перевіряється через доступні бюро кредитних історій на предмет співпадіння з відомими шахраями, які могли змінити персональні дані (підробити), місце проживання і вже розшукуються за скоєні зловмисні дії. В разі співпадіння інформація терміново передається службі безпеки та правоохоронним органам для проведення подальшого розслідування.

Ще однією можливістю для вчинення шахрайських дій залишається потрапляння до зловмисників інформації про особливість самої скорингової моделі, яка використовується на даний момент в банку для перевірки платоспроможності позичальників та відібраних характеристик, порогу відсікання тощо. Така інформація може передаватися шахраям безпосередньо працівниками банку. Засобом боротьби з такими видами ризику є підвищення безпеки скорингової моделі, періодична перевірка працівників банку, та періодичне оновлення скорингової моделі та відповідно скорингових карт для поведінкового скорингу та скорингу потенційних позичальників.

Аналізуючи статистичні дані декількох українських банків можна зробити висновок, що шахрайські дії зловмисників, які не були виявлені на етапі Fraud скорінгу, в подальшому будуть спостерігатися в якості постійної прострочки платежів, починаючи з першого місяця сплаті по кредиту. Всі дії, які спрямовуються банками на повернення такого кредиту (повідомлення, дзвінки тощо) є нерезультативними, бо скоріш за все в цей момент шахраї вже змінюють адреси проживання та мобільні телефони. Таким чином, єдиним ефективним способом боротьби з шахрайством при отриманні кредиту є розробка коректної скорингової моделі – скорингової карти – яка дозволить працівникам банку виявити фактори, що характеризують шахраїв, і відмовити у видачі кредиту.

На жаль, ще одним фактором підвищення ризиків шахрайства в банку залишається той факт, що прийняття рішення щодо видачі кредиту здійснюється самими працівниками фінансових установ після отримання автоматичного результату оцінювання за скоринговою картою. Тут залишається можливість маніпуляції та видачі кредиту тим особам, яким скорингова система відмовила в отриманні кредиту (наприклад, через наявність заборгованостей в інших банках чи негативної кредитної історії в бюро кредитних історій). Для уникнення людського фактору при прийнятті кредитних рішень рекомендується автоматизувати повністю процес прийняття рішень при аналізі кредитних заявок і забезпечити неможливість або складність зміни прийнятих рішень (наприклад, через штрафні санкції для працівників банку в разі виявлення кредитів, по яким приймалося рішення всупереч скоринговій карті, і по яким встановлені факти заборгованості або неповернення кредитів).

Література

1. Кузнєцова Н.В. Аналіз фінансових ризиків з використанням SAS-технологій обробки даних / Н. В. Кузнєцова, П.І. Бідюк // Електротехнічні і комп'ютерні системи. – 2016. – № 22(98). – С. 267 – 271.
2. Siddiqi N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring /N. Siddiqi// John Wiley & Sons, Hoboken. – 2005. – 208 p.

ПОСТРОЕНИЕ МОДЕЛИ ИНФОРМАЦИОННОГО СЕРВИСА НА БАЗЕ НАЦИОНАЛЬНОГО СЕГМЕНТА ИНТЕРНЕТ

Д. Ландэ[1], Б. Березин[1], О. Павленко[2],
[1]Институт проблем регистрации информации НАН Украины,
[2]Университет "Украина"

Актуальность и постановка проблемы

В настоящее время китайский сегмент Интернета является наибольшим в мире по количеству пользователей – более 688 млн. (что составляет более 50 % населения страны) и быстрорастущим сегментом Сети. Третий по количеству пользователей (после Китая и Индии) сегмент Интернета США насчитывает около 280 млн. пользователей, что составляет более 80 % населения страны. В ряде работ [1–7] отмечаются особенности китайского сегмента Интернета: большое число мобильных интернет-пользователей – в Китае они составляют около 90 % владельцев смартфонов, а в США около 40 %; большая активность и стабильность в публикации контента Интернета (для пользователей из группы стран, включающей Китай, среднее число публикаций на 20–50 % больше группы стран, включающей США); возраст основных групп пользователей – около 30 % составляет 20–29 лет; собственные социальные сети и поисковые системы.

Контент китайского сегмента Интернета представлен 4,23 млн. веб-сайтов и 212,3 млрд веб-страниц. Преимущественное использование китайского языка и незначительная доля английского в контенте китайского сегмента Интернета затрудняет непосредственное использование китайского контента в европейских и американских странах. Однако возможности Google и других онлайн переводчиков позволяют преодолеть языковый барьер и повышают актуальность сбора контента китайского сегмента Интернета.

Приведенные выше особенности китайского сегмента Интернета делают перспективным сбор и использование его контента в различных направлениях. Однако, в настоящее время, особенности контента и возможности его сбора исследованы и использованы недостаточно. Имеющиеся работы [1–5] и др., как правило, посвящены лишь отдельным характеристикам контента. В данной работе, на примере китайского сегмента сетевого информационного пространства показаны особенности контента национального сегмента Интернет. Предложено построение модели

информационного сервиса, обеспечивающего сбор контента китайского сегмента с помощью системы мониторинга веб-ресурсов на основе использования RSS-каналов, а также мониторинг ключевых слов социальной сети. Система мониторинга на основе каналов RSS является универсальной и может быть адаптирована к любому национальному сегменту Интернет.

Особенности контента китайского сегмента Интернет

Кроме высоких темпов роста количества веб-ресурсов и числа пользователей, китайский сегмент Интернета выделяется в мировой Сети наличием собственных социальных сетей, объемы которых соизмеримы с объемами аналогичных мировых; наличием собственной основной поисковой системы Baidu, ориентированной преимущественно на китайский язык и покрывающей значительную часть веб-ресурсов китайского сегмента Интернета; пока еще относительно небольшим, но развивающимся представлением ресурсов в формате RSS.

Социальные сети. Пользователи китайского сегмента Интернет являются активными участниками китайских социальных сетей. По данным [5], на конец 2015 г более 65 % пользователей Интернета использовали QQ.Zone.com, более 33 % – Weibo.com, около 15 % – использовали RenRen.com, Pengyou.com и Douban.com.

С помощью сервиса Alexa.com был проведен сравнительный анализ рейтинга и пользователей основных международных и китайских социальных сетей (таб. 1). Результаты подтверждают, что соизмеримые по размеру с Facebook и Twitter китайские сети по составу пользователей являются национальными сетями.

В работе [2] проведен сравнительный анализ трендов в китайских социальных сетях на основе списка 50 ключевых слов, которые появляются чаще всего в твитах пользователей weibo.com (ранжируются по частоте появлений за последний час) и анализа трендов тем в Твиттере. Показано, что среднее время нахождения каждого ключевого слова в списке трендинга составляет около 6-ти часов. Кроме того, распределение количества часов нахождения каждой темы в списке трендинга соответствует степенному закону (это показывает, что только нескольким темам свойственна долговременная популярность).

Поисковая система Baidu. Baidu.com была основана в 2000 г. и в 2004г. стала лидирующей поисковой системой в Китае. По количеству обрабатываемых запросов занимает 2 место в мире (с долей в глобальном поиске 18 %). В 2006 г. китайская поисковая

система Baidu предоставляла своим пользователям доступ на основе индекса более 740 млн веб-страниц, 80 млн изображений и 10 млн файлов мультимедиа.

Таблица 1.

	Facebook	Twitter	Weibo	Qzone.qq	Douban	RenRen
Рейтинг Alexa.com	3 gl; 2 US	8 gl; 8 US	20 gl; 5 Cn	10 gl; 2 Cn	868 gl; 72 Cn	1783 gl; 195 Cn
Год создания, краткие характеристики	2004 г. – более 1 млрд пользователей	2006 г., – более 200 млн пользователей	2009 г. – 250 млн аккаунтов, 90 млн постов/день	2005 г. – более 600 млн пользователей	2005 г. – около 200 млн. пользователей	2005 г. – более 160 млн пользователей
Страны пользователей	22 % – US 8 % – In 4 % – Br 3 % – GB 3 % – Gr	22 % – US 14 % – Jp 7 % – In 6 % – GB 4 % – Mx	97 % – Cn 0,7 % – US 0,6 % – Tw	98 % – Cn 0,5 % –US	92 % – Cn 3 % – US 0,8 % – Hk 0,8 % – Tw	92 % – Cn 4 % – US 0,8 % – Tw 0,6 % – Jp 0,5 % – Hk

Среди значительного числа сервисов, предоставляемых Baidu, возможности созданного в 2014 г. сервиса http://xueshu.baidu.com, который называют Baidu Scholar, во многом аналогичны сервису Google Scholar, который существует с 2004 г. Сравнительный анализ особенностей представления научных публикаций украинских авторов в результатах, формируемых поисковыми системами Baidu Scholar и Google Scholar, отражен в таблице. Таблица содержит результаты поиска англо-, украино- и русскоязычных научных публикаций по инициалам и фамилии авторов, перечисленных в таблице. (Приводятся результаты поиска для варианта инициалов без разделения точкой и с разделением. Другие варианты написания не рассматривались). В столбцах Baidu Scholar и Google Scholar каждому автору, по которым выполнялся поиск, соответствует количество полученных в результате поиска ссылок. Для некоторых из полученных в результате поиска количеств ссылок, с целью детализации, через дробь приводится также второе число – количество реально отобранных из них ссылок, соответствующих запросу. В случаях больших количеств полученных при поиске ссылок (более 150), отбор ссылок, соответствующих запросам, производился из первых 150 полученных результатов. При представлении результатов поиска в виде дроби, по соотношению числителя и знаменателя может оцениваться качество поиска. Для детализации результатов, кроме обычного поиска по всем доменам, в

Google Scholar также производился поиск отдельно по доменам .org, .ua, и .cn. Из результатов поиска и таблицы видно, что англоязычные публикации сосредоточены на серверах домена .org (например, arxiv.org), а украино- и русскоязычные на серверах домена .ua (nbuv.gov.ua и др.).

Сравнение результатов использования Baidu и Google показывает близость полученных ответов на запросы для англоязычных публикаций в обоих системах. (Число полученных в результате поиска ссылок в столбце 3 таблицы незначительно превышает число полученных ссылок в столбце 4, при этом число предлагаемых ссылок не намного больше количества реальных публикаций, которые могут быть по ссылкам отобраны. Для большинства показанных примеров, из нескольких десятков полученных ссылок могут быть отобраны десятки реальных публикаций). Для украино- и русскоязычных публикаций наоборот, получена значительная разница в ответах Baidu и Google по сравнению с реальным числом публикаций. (Для Baidu, в отличие от Google, число предлагаемых ссылок значительно превышает количество реальных публикаций, которые могут быть по ссылкам отобраны. Из нескольких десятков тысяч ссылок для украино- и русскоязычных публикаций, реально могут быть отобраны из первых 150 ссылок несколько десятков публикаций, соответствующих запросам).

RSS-каналы. RSS (Rich Site Summary – обогащенная сводка сайта) – это семейство XML-форматов, используемое для публикации и доставки часто изменяющейся информации (заголовков новостей, анонсов статей, новых записей в блогах и т.д.), это технология, которую применяют пользователи Интернета для получения обновлений с интересующих их веб-страниц. Природа RSS обусловила то, что одним из эффективных способов сбора контента информационных ресурсов Итернет является использование каналов RSS. Количество каналов RSS в 2004 г. составляло около 307 тыс., в 2016 г. директория Feedage.com (в которой RSS каналы со всего мира представлены 15 категориями с возможностью поиска) объединяет более 3,1 млрд каналов.

Анализ использования RSS-каналов на веб-ресурсах китайского сегмента Интернета и в мире показывает следующее. В работе [8] исследовано использование технологий Web 2.0 (социальных сетей, технологий wiki, блогов, RSS-каналов, обмен мгновенными сообщениями и функции каталогизации) в библиотеках 38 ведущих университетов Китая. Показано, что RSS

является второй по частоте использования технологией (представлена в 55% университетских библиотек). Отмечаются три основные цели использования RSS в библиотеках китайских университетов: уведомление об информации, представляющей интерес для читателей по инициативе библиотеки – новости и события библиотеки, доступность новых книги, т.е. информационная база данных; уведомление о личной информации про пользование библиотекой; синдикация тематической информации для легкого и своевременного доступа. Эти цели предполагают разные уровни технологической поддержки, поэтому большинство библиотек обеспечивают в основном базовые возможности RSS-каналов. Только RSS-каналы библиотек Шанхайского университета ориентированы на достижение всех трех целей. В работе [9] рассмотрено внедрение технологий Web 2.0, в том числе и RSS-каналов, в библиотеках 30-ти ведущих университетов Китая. Показано, что из всех технологий Web 2.0, каналы RSS получили наибольшее распространение в библиотечных проектах (далее следует передача сообщений, использование блогов и т.д.). Больше всего каналы RSS используются для распространения новостей и уведомлений – в 12-ти университетах из 30-ти, что составляет 43%.

В работе [10] исследуется использование приложений Web 2.0, в том числе и RSS-каналов (распространение информации библиотек для пользователей), на сайтах 120-ти крупнейших библиотек трех регионов – Северной Америки, Европы и Азии. В числе 40-ка крупнейших библиотек региона Азии рассматривались Гонконгская публичная библиотека и библиотеки Китайского университета Гонконга, а также Гонконгского университета науки и технологий, также библиотека Университета Цинхуа, Национальная центральная библиотека (Тайвань) и Национальная библиотека Китая. Из общего количества проанализированных 120-ти сайтов библиотек, распространение информации с помощью каналов RSS применяется на 28-ми сайтах университетов Северной Америки (что составляет около 70%), 17-ти и 15-ти сайтах университетов Европы и Азии (43% и 37% соответственно). В целом, исследование показало, что RSS-каналы используются примерно в 50-ти % крупнейших библиотек трех регионов и занимают второе место по популярности среди приложений Web 2.0 после блогов (примерно 57%). Также для сравнения, данные по использованию каналов RSS в 100-та ведущих академических библиотеках США приводятся в [11]. По данным этого исследования, из Web 2.0 технологий больше всего используются социальные сети – в 100% библиотек. Блоги

используются в 99%, а RSS-каналы в 97% исследованных академических библиотек США.

Результаты анализа использования RSS-каналов на веб-ресурсах китайского сегмента Интернета и в мире приведены на рис.16. Рисунок показывает, что около половины сайтов библиотек используют каналы RSS, это больше чем в среднем по странам Азии, но меньше чем в странах Европы и США.

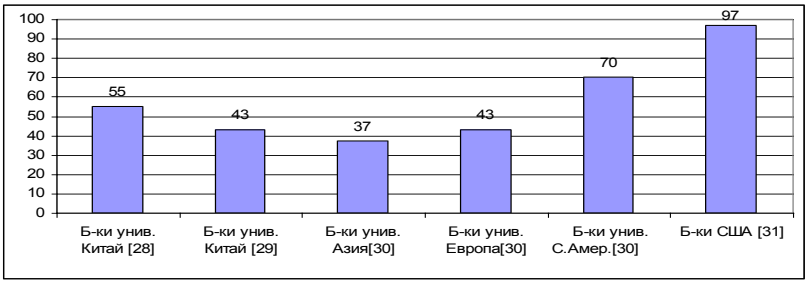


Рисунок 1 – Анализ использования RSS-каналов в составе Web 2.0 технологий на сайтах библиотек ведущих университетов Китая [8,9], библиотек ведущих университетов стран Азии, Европы, Сев. Америки [10], а также академических библиотек США [11]. Показан процент сайтов библиотек с использованием RSS-каналов от общего числа исследованных в указанной работе библиотек

Построение модели информационного сервиса

Предлагаемая модель информационного сервиса разрабатывается с учетом особенностей контента китайского сегмента Интернет [6,7]. Наличие RSS-каналов обеспечивает непрерывное получение обновлений веб-сайта как заинтересованными пользователями, так и автоматическими системами анализа веб-ресурсов. В частности, была разработана система мониторинга веб-ресурсов, базирующаяся на использовании RSS-каналов (рис. 2).

Система мониторинга веб-сайтов на основе использования каналов RSS обеспечивает сбор необходимого контента различной направленности, является универсальной и может быть с минимальными перестройками адаптирована к любому национальному сегменту Интернет.

Мониторинг ключевых слов. Кроме системы мониторинга веб-ресурсов на основе использования RSS-каналов, модель информационного сервиса предусматривает мониторинг социальной

сети Weibo. С целью мониторинга топ-50 ключевых слов сети Weibo было выбрано приложение для браузера Firefox – Alertbox, которое было настроено для мониторинга списка ключевых слов с периодом около 1 часа. В результате почти недельного мониторинга были получены примеры графиков изменения количества страниц weibo.com, содержащих определенные ключевые слова из списка топ-50.

Рисунок 2 – Интерфейс эксперта-аналитика подключения RSS-канала к системе мониторинга веб-сайтов

Например, ключевое слово pokemongo на страницах Вейбо 23.07.16 (в течение около 14 часов) занимало с 18 по 9 места рейтинга топ-50 с количествами страниц примерно от 10-ти тыс. до 160-ти тыс. (На рис. 3 нижний график, обозначение ключевое слово 1). Ключевое слово 快乐大本营 (Счастливый лагерь – популярное развлекательное шоу Китая) на страницах Вейбо с 23.07.16 до 24.07.16 (около 18-ти часов) занимало с 47 по 2 места рейтинга топ-50 с количествами страниц примерно от 20-ти тыс. до более 500 тыс. (На рис. 3 верхний график, обозначение ключевое слово 2).

На рис. 3 по оси абсцисс показаны номера сканирования страницы топ-50 Вейбо, а по оси ординат – количество страниц социальной сети Вейбо с соответствующими ключевыми словами.

Для более информативного мониторинга ключевых слов Weibo, целесообразно вместе с китайскими ключевыми словами отображать их перевод на английский, украинский, русский и др. языки. Примеры перевода ключевых слов с помощью Google, Yandex, Bing переводчиков приведены в табл. 3.

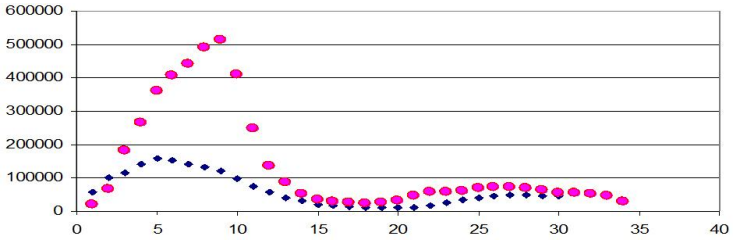


Рисунок 3. – Анализ количества появлений ключевых слов pokemongo (кл. слово 1) и 快乐大本营 (кл. слово 2 — Счастливый лагерь) в твитах социальной сети Вейбо на основе мониторинга страницы топ-50 Вейбо

С целью такого мониторинга используются три модуля на языке Perl: модуль выборки ключевых слов из веб-страницы, модуль обмена с API системы перевода (для получения соответствующих переводов ключевых слов с китайского на английский и русский языки), модуль представления ключевых слов вместе с переводами на веб-странице.

Таблица 3. Примеры онлайн перевода ключевых слов Weibo системами Google, Yandex, Bing

Ключ. слово	Рейтинг	Google	Yandex	Bing	Содержание
林依晨哭了	302947	Ariel крик	Ариэль плакала	Ариэль Крик	Актриса о награде
真的有蓝瘦的香菇	285822	Действительно тонкие голубые грибы	Правда синий тощие грибы	True blue тонкий гриб	Специалист о необычных грибах
你绑定的支付宝怎么办	285085	Вы связываете Alipay, как это сделать	Вы не привязаны к PayPal?	Какие привязки платежной карты вы	О переходе на новую систему оплаты
同班同学联名举报	124998	Совместный отчет Classmate	Одноклассница совместный доклад	Группы одноклассников сообщили	Одноклассники выступили против плагиата

Выводы

К особенностям китайского сегмента веб-пространства следует отнести:

– темпы роста веб-ресурсов и числа пользователей, которые по ряду характеристик превосходят всемирный сегмент Интернета;

– наличие собственных национальных социальных сетей, объемы которых соизмеримы с объемами аналогичных международных сетей в мире;

– наличие собственной основной поисковой системы Baidu (наряду с еще несколькими), ориентированной преимущественно на китайский язык (существенной проблемой является применение латиницы и кириллических кодов) и покрывающей значительную часть веб-ресурсов китайского сегмента Интернет;

– пока еще относительно небольшое представление ресурсов в формате RSS (связанное с некоторым запаздыванием внедрения интернет-технологий). Вместе с тем представление веб-ресурсов в RSS-формате в настоящее время возрастает и все шире используется в мобильных приложениях.

Предложенная с учетом особенностей контента китайского сегмента Интернет модель информационного сервиса обеспечивает мониторинг актуальных новостных тем и сбор необходимого контента различной направленности.

В то же время, система мониторинга на основе использования каналов RSS является универсальной и может быть с минимальными перестройками адаптирована к любому национальному сегменту Интернет.

Литература

1. Deans P.C. A framework to understanding social media trends in China / P.C. Deans, J.B. Miles // The 11-th Internation. DSI and APDSI Joint Meeting, Taipei, Taiwan. — 2011, July 12-16. — P. 12–16.
2. Yu L. Dynamics of trends and attention in chinese social media / L. Yu, S. Asur, B.A. Huberman // arXiv preprint arXiv:1312.0649, 2013. –P. 1–17.
3. Bolsover G. Social Foundations of the Internet in China and the New Internet World: A Cross-National Comparative Perspective / G. Bolsover, W.H. Dutton, G. Law. — Oxford Internet Institute, University of Oxford,. — 2013. — P. 1–22.
4. Internet Users by Country (2016) [Электронный ресурс]. — Режим доступа: http://www.internetlivestats.com/internet-users-by-country/. — Название с экрана.
5. CNNIC. (2016) // The 37-th Statistical Report on Internet Development in China. \
6. Ландэ Д.В., Березин Б.А., Додонов В.А. Обзор особенностей и возможности контент-мониторинга национального

сегмента сети Интернет // Реєстрація, зберігання і обробка даних, 2016. – Т. 18. – № 3. – С. 20-38.
7. Ландэ Д.В., Березин Б.А., Павленко О.Ю. Подход к мониторингу контента китайского сегмента Интернет // Міжнародна науково-практична конференція "Інтелектуальні технології лінгвістичного аналізу": Тези доповідей. – К.: НАУ, 2016. – С. 24.
8. Han Z. Web 2.0 applications in top Chinese university libraries / Zhiping Han, Yan Quan Liu // Library Hi Tech. 2010. — 28.1. — P. 41–62.
9. Chua A. A study of Web 2.0 applications in library websites / AYK Chua, DH Goh // Library & information science research. — 2010. — 32.3. — P. 203–211.
10. Si L. An investigation and analysis of the application of Web 2.0 in Chinese university libraries / L. Si, R. Shi, B. Chen // The electronic library 29.5. — 2011. — P. 651–668.
11. Boateng F. Web 2.0 applications' usage and trends in top US academic libraries / F. Boateng, Y. Quan Liu // Library Hi Tech 32.1. — 2014. — P. 120–138.

ОНТОЛОГІЧНИЙ ПІДХІД ДО РОЗРОБКИ НАЦІОНАЛЬНИХ СТАНДАРТІВ УКРАЇНИ З ОЦІНЮВАННЯ БЕЗПЕКИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Ю.В. Рогушина[1], А.Я. Гладун[1,2], Г.В. Снігирь[3],

[1]Інститут програмних систем НАН України,
[2]Міжнародний науково-навчальний центр інформаційних технологій та систем НАН України та МОН України
[3]ПрАТ "Міжнародні Авіалінії України", служби авіаційної безпеки, Київ

Розробка нормативно-правової бази України для підтримки інформаційної безпеки, яка повинна відповідати міжнародним стандартам та кращим зарубіжним практикам, потребує створення методів та засобів обробки і аналізу відповідної інформації на семантичному рівні. Аналіз публікацій вказує на велику кількість національних стандартів, безпосередньо пов'язаних з інформаційною безпекою, які вже розроблені або знаходяться на стадії узгодження. Між цими стандартами та їх елементами існує складна система ієрархічних відношень, взаємозв'язків, посилань та відповідностей. Через недостатньо глибоко стандартизовану та усталену україномовну термінологію виникають різні варіанти перекладу основних понять та назв. Тому виникає потреба в узгодженні термінологічної бази стандартів на семантичному рівні на основі сучасних засобів подання та обробки знань, які уможливлюють автоматизоване знаходження протиріч та розбіжностей. В роботі проаналізовано вимоги, що застосовуються до терміносистем та визначень термінів, що використовуються в галузевих терміносистемах, та пропонуються шляхи для задоволення цих вимог.

Для формалізації терміносистем стандартів пропонується використовувати онтологічний аналіз, який дозволить інтегрувати наявні знання із сфери інформаційної безпеки у відкритих ресурсах Web з досвідом експертів із стандартизації. Наявність онтологічних метаописів стандартів має забезпечити швидкий та ефективний доступ до релевантних

відомостей та коректну однозначну інтерпретацію контенту стандартів. Розроблено базові принципи побудови онтологічного опису стандарту.

Запропонований підхід детально розглянуто на прикладі стандарту "Інформаційні технології. Методи захисту. Оцінювання безпеки операційних систем", побудовано онтологію термінів цього стандарту, визначено їх властивості та відношення. Розглянуто перспективи та доцільність створення глобальної семантичної мережі стандартів, яка пов'язує окремі національні та міжнародні стандарти, сферу її застосування для семантичної обробки інформації. Така мережа має стати основою для еталонної семантичної розмітки різноманітних інформаційних ресурсів (як природномовних, так і мультимедійних), дозволить вдосконалювати онтології предметних областей та зробити більш ефективним доступ до інформації.

Сьогодні в Україні велике значення надається розробці нормативно-правової бази для підтримки інформаційної безпеки. Значна увага приділяється гармонізації нормативно-правового забезпечення України щодо оцінки відповідності систем управління інформаційною безпекою правилам та процедурам, які відповідають міжнародним стандартам та кращим міжнародним практикам. Це створює передумови для усунення технічних бар'єрів між країнами в сфері технічного регулювання, в тому числі з питань безпеки.

Практично усі сфери діяльності людини важко уявити без використання стандартів, які акумулюють передовий науково-технічний досвід багатьох країн, забезпечуючи єдність вимог до продукції, яка є предметом міжнародного товарообміну (взаємозамінність комплектуючих виробів, єдині методи випробувань і оцінювання якості виробів). Якість стандартів, що розробляються, залежить від багатьох чинників, включаючи компетентність розробників та узгодженість термінологічної бази, що безпосередньо впливає на ефективність їх використання.

Міжнародні організації доклали великих зусиль для забезпечення уніфікації засобів та систем інформаційної безпеки. Це відображено у великій кількості міжнародних стандартів ISO, зі значною частиною яких вже гармонізовано. національні стандарти України – ДСТУ. В Україні діяльність зі стандартизації ґрунтується на правових нормах Закону України «Про стандартизацію» [1],

Декрету КМУ «Про стандартизацію та сертифікацію», інших нормативних актах у цій сфері, з урахуванням принципів і положень міжнародних організацій зі стандартизації [2,3]. Стандартизація – це складна, міжгалузева, комплексна й багатоаспектна задача, рішення якої пов'язане із розробкою сучасних нових методів та технологій. Приміром, в [4] проаналізовано ієрархічну структуру більш ніж 100 стандартів, пов'язаних з методами і засобами забезпечення безпеки інформаційних технологій.

Постановка задачі

У зв'язку з важливістю розробки національного нормативно-правового забезпечення інформаційної безпеки, великим обсягом міжнародних стандартів в цій сфері та термінологічними проблемами гармонізації національних стандартів України виникає потреба у розробці методів та засобів узгодження термінологічної бази на семантичному рівні, з використанням сучасних засобів подання та обробки знань. Пропонується використовувати для цього онтологічний аналіз, щоб інтегрувати наявні знання в цій сфері, які містяться у відкритих ресурсах Web, із досвідом експертів зі стандартизації. Онтологічні метаописи стандартів, що характеризують їх семантику, забезпечать коректну однозначну інтерпретацію вмісту стандартів та їх автоматизовану обробку.

Вимоги до термінології, що застосовується в національних стандартах України

Термін – це слово або словосполучення, яке зіставляється з чітко окресленим поняттям предметної області (ПрО) і вступає в системні відношення з іншими одиницями мови, утворюючи разом з ними особливу систему – *термінологію*. Виділяють такі ознаки терміну: 1) чітку визначеність, зафіксовану в словнику; 2) однозначність в межах певної термінологічної системи; 3) точність, яка не залежить від контексту; 4) стилістичну нейтральність; 5) відсутність синонімів в межах обраної терміносистеми; 6) системність; 7) стислість. Принциповим у відборі термінів є цілісність системи понять та термінів, які змістовно підпорядковуються певній теорії або концепції.

Галузеві термінології (тобто сукупності термінів конкретних галузей) називають *терміносистемами*, або термінологічними системами. Характеризуючи сучасну україномовну термінологію в ІТ-сфері в цілому та окремо – пов'язану з інформаційною безпекою, необхідно відмітити як динаміку її розвитку та інтенсивність

збагачення новими лексичними одиницями, так і неоднозначні визначення багатьох ключових понять.

Для формалізації терміносистем певної ПрО широко використовують моделі знань щодо ПрО: тезауруси, таксономії, онтології. Визначення термінів створюються вручну експертом ПрО, будуються автоматично на основі обробки інформаційних джерел або здобуватися з інших баз знань (тезаурусів, онтологій тощо). В усіх випадках доцільно дотримуватися певних правил та удосконалювати отримані визначення відповідно до наведених нижче принципів [5].

Потрібно, щоб кожне визначення терміну:

– викладалося в однині (виключення складають поняття, які самі є множинними);

– визначало, чим саме є наведене поняття, а не тільки чим воно не є;

– мало вигляд описової фрази або речення;

– містило лише поширені скорочення;

– викладалося без використання визначень інших даних або базових понять;

– відображало суттєвий зміст поняття;

– було точним та однозначним; було стислим;

– припускало окреме використання;

– подавалося без пояснювальної інформації, функціонального використання або процедурної інформації;

– не містило циклічних посилань;

– використовувало однакову термінологію та логічну структуру для пов'язаних визначень. Для близьких або пов'язаних визначень повинні використовуватись одна й таж сама термінологія та синтаксис.

Терміносистема стандарту з оцінювання безпеки операційних систем

Розглянемо створення терміносистеми в ПрО інформаційної безпеки на прикладі стандарту "Інформаційні технології. Методи захисту. Оцінювання безпеки операційних систем" (ISO/IEC TR 19791:2010 (E), IDT), що базується на міжнародному стандарті "Information technology – Security techniques – Security assessment of operational systems" [6]. Цей стандарт надає настанови і розширені критерії оцінювання операційних систем – поєднання персоналу, процедур і процесів, інтегрованих за допомогою функцій і механізмів на основі технічних засобів, які використовуються для

обґрунтування застосовного рівня залишкового ризику у певній операційній системі. Для багатьох організацій інформація є основним ресурсом і вимагає захисту від загроз несанкціонованого розголошення, зміни або знищення.

Ці ресурси захищаються за допомогою поєднання технічних засобів керування та підтримки інфраструктур операційного контролю персоналу, політики, процедур і заходів фізичного захисту. При оцінюванні продукту послуги, пов’язані з безпекою, – це такі ІТ-функції, які впроваджені для досягнення цілей інформаційної безпеки технології. У контексті ОС можуть бути оцінені також процедурні та фізичні вклади у безпеку. Вони ідентичні функціям ІТ, оскільки вони являють собою такі засоби безпеки ОС, які разом сприяють досягненню цілей безпеки. Однак вони, як правило, не залежать від технології і більш пристосовані до оцінювання на стадії контролю життєвого циклу операційної системи і таким чином розглядаються окремо від функціональних вимог. Стандарт орієнтований на тих, хто бере участь в розробці, інтеграції, впровадженні і керуванні безпекою операційних систем, а також експертів, які бажають застосувати ISO/IEC 15408 в операційних системах.

В оцінюванні безпеки ОС можна виокремити наступні етапи:

- Проблема безпеки формулюється як множина ризиків, які слід усунути або пом'якшити, і група політик безпеки, які слід застосувати. Для цього слід провести попередній аналіз і визначити завдання ОС і провести оцінку ризиків для визначення таких ризиків, яким слід протидіяти за допомогою технічних та операційних засобів керування. Результати аналізу записуються в системний об’єкт безпеки

- Проблема безпеки ділиться на рішення по забезпеченню безпеки високого рівня і групу організаційних політик захисту, які слід застосувати. Ці цілі записуються в системний об’єкт безпеки.

- Цілі безпеки в подальшому уточнюються у вимогах щодо безпеки, які можуть оцінюватися незалежним експертом. Деякі цілі безпеки будуть віднесені до технічних, інші до операційних засобів керування. Деякі можуть вимагати як технічних, так і операційних засобів керування. Наприклад, контроль несанкціонованого доступу до інформаційних ресурсів часто доповнюється як через забезпечення фізичної безпеки об’єкта, в якому містяться ресурси (наприклад, замки, захисні пристрої), так і за допомогою функцій ІТ. Вимоги до безпеки записуються в системний об’єкт безпеки.

- Сукупність дій експерта, яких слід додержуватись при оцінюванні, визначається на підставі загальних цілей і загального забезпечення гарантії захисних заходів. Ці вимоги забезпечення гарантії записуються в системний об'єкт безпеки.

- Незалежний експерт визначає факт того, що ОС відповідає вимогам безпеки, на основі вимог, зазначених у системному об'єкті безпеки.

- Продовження оцінювання також є можливим з метою отримання гарантії того, що ОС відповідає власним вимогам при експлуатації. Вони фокусуються на операційних засобах керування ОС, оскільки вони залежать від поводжень людини, що менш керовані, ніж поведінка ІТ.

- Періодична повторне оцінювання ОС може визначити, що ОС продовжує відповідати своїм вимогам, незважаючи на зміни в ОС або її середовищі. Це визначається тим, які зміни мали місце, шляхом оцінювання впливу цих змін, оновлення системного об'єкту безпеки, і визначення того, що безпека все ще зберігається.

Першочерговою метою оцінювання ОС є гарантування того, що цілі безпеки ОС впроваджені правильно та ефективно. Однак, оцінювання як операційних, так і технічних засобів керування безпекою, не може абсолютно гарантувати, що ці засоби керування завжди будуть функціонувати належним чином. Процедура оцінювання виносить недиференційований висновок. Навіть коли оцінка не виявляє неприйнятних вразливостей, завжди залишається залишковий ризик неналежного функціонування засобів керування. Цей ризик можна зменшити за допомогою додаткових засобів керування, або використовуючи інші заходи забезпечення гарантії, які дають більшу впевненість. Залишковий ризик невірної або неефективної роботи можна визначити тільки шляхом безперервного моніторингу і оцінювання.

Концепція ISO/IEC 15408 з безпеки приділяє особливу увагу підтвердженню наявності існування функцій захисту і їх правильної і ефективної роботи. Високі рівні гарантії висувають більш докладні вимоги до змісту і стилю подання таких підтверджень. Крім того, більш високий ступінь забезпечення гарантії іноді вимагає більш строгого аналізу підтверджень як з боку розробника, так і експерта з оцінювання.

Оцінювання продукту ISO/IEC 15408 проводиться у порядку, що передбачає типове операційне середовище, в якому цей продукт може застосовуватись. При оцінюванні продукту головна увага зосереджується на контрольній перевірці таких функцій

захищеності, які цей продукт забезпечує незалежно від конкретних умов експлуатації. При оцінюванні продукту застосовуються різноманітні технічні характеристики, проектна документація та документація з тестування для обґрунтування висновку про коректність. При цьому вимоги щодо забезпечення гарантії, як правило, не випливають з проблеми безпеки. Натомість, вони вибираються аксіоматично або стратегічним рішенням. Основною метою оцінювання продукту є отримання гарантії того, що функції безпеки реалізовані правильно.

Базисна інформація для визначення коректності визначається вимогами безпеки, які містяться в цілях безпеки продукту, що включає в себе простежуваність проблеми безпеки. Проблема, зазначена в ST, ґрунтується на оцінці загрози для всіх видів оточення. Область застосування оцінювання продукту обмежується вимогами безпеки ІТ.

Термінологічна база цього стандарту включає терміни з ISO/IEC 15408-1, ISO/IEC 18045, а також наступні терміни з відповідними визначеннями:

- *Адміністративний контроль* (management controls): Контроль безпеки (тобто засоби захисту і заходи протидії) інформаційної системи, що зосереджується на керуванні ризиком і керуванні безпекою інформаційної системи.

- *Аналіз ризику* (risk analysis): Систематичне використання інформації для визначення джерела й оцінювання ризику.

- *Залишковий ризик* (residual risk): Ризик, що залишається після обробки ризику.

- *Захищена зона* (security domain): Частина операційної системи яка виконує функції політики забезпечення захисту.

- *Зовнішня операційна система* (external operational system): Окрема операційна система, яка взаємодіє з оцінюваною операційною системою.

- *Керування ризиком* (risk management): Скоординовані дії з керування й контролю структурою стосовно ризику.

- *Компонент* (component): Розпізнавана та окрема частина операційної системи, яка виконує частину функцій системи.

- *Контроль безпеки* (security controls): Керування, операційний і технічний контроль (тобто засоби захисту і заходи протидії), заздалегідь описані для інформаційної системи з метою захисту конфіденційності, цілісності й доступності системи та її інформації.

- *Контрольна перевірка* (verification): Процедура оцінювання, яка використовується для підтвердження того, що контроль безпеки операційної системи запроваджений і виконаний правильно і є ефективним при застосуванні.

- *Мета оцінювання системи* (system target of evaluation): Операційна система, якою керують у відповідності з її експлуатаційною настановою, включаючи як технічні, так і операційні засоби керування.

- *Обробка ризику* (risk treatment): Процедура вибору й застосування варіантів коригування ризику.

- *Операційна система* (operational system): Інформаційна система, яка виступає як інтерфейс між пристроями обчислювальної системи і прикладними програмами.

- *Операційний контроль* (operational controls): Контроль безпеки (тобто засоби захисту і заходи протидії, які переважно запроваджуються і виконуються людьми (на відміну від систем).

- *Оцінка ризику* (risk assessment): Уся процедура аналізу й оцінювання ризику.

- *Підсистема* (subsystem): Один чи більше ніж один компонент ОС, що можуть діяти окремо від решти системи.

- *Ризик* (risk): Поєднання ймовірності виникнення певної комбінації обставин та її наслідків.

- *Технічні засоби керування* (technical controls): Керування безпекою (тобто засоби захисту і заходи протидії), які переважно запроваджуються і виконуються інформаційною системою за допомогою механізмів, що містяться в апаратних, програмних або мікропрограмних компонентах забезпечення системи.

Однак ці конфігурації не приймають до уваги будь-які специфічні оточення. По завершенні оцінювання продукту все ще залишається необхідність інтегрувати оцінений продукт з іншими продуктами для створення ST, і в кінцевому рахунку, для звірки того, що ОС надає параметри безпечності та поведінки у своєму середовищі і операційних конфігурацій. При оцінюванні продукту, як правило, застосовуються ті ж самі заходи забезпечення гарантії по всіх визначених функціях безпеки. Хоча технічно можливо отримати різні домени безпеки в продуктах, але зазвичай такий метод для оцінювання типового продукту не застосовується.

Свідоцтво і звіти оцінювання можуть використовуватися для підтримки інтеграції ОС і контрольної верифікації. Хоча різниця між

властивостями продукту ІТ і ОС для цілей оцінювання безпеки невелика, але оцінювання ОС ускладнюють наступні причини:

- ОС може включати в себе багато комплектуючих та розроблених на замовлення ІТ-компонент, згрупованих в захищених доменах. Склад кожної системи захищеного домену може базуватися на декількох факторах, таких як технологія, функціональність і критичність захищених ресурсів;

- ОС може містити кілька конкретних зразків однакового продукту (приміром, кілька копій ОС від одного і того ж постачальника) або кілька різних зразків продуктів одного виду (приміром, кілька систем мережного захисту від різних постачальників);

- ОС може мати політики безпеки, які застосовуються лише на деяких захищених доменах, а на інших можуть не застосовуватися;

- Залишкові ризики можуть бути прийнятними в межах різних доменів ОС, тоді як сам продукт протидіє конкретним загрозам для конкретних типів активів без врахування ризику.

Всі ці фактори впливають на вимоги до безпеки ОС. Зокрема потрібні різні форми гарантії в різних зонах, залежно від інформації, або від видів обраних функціональних засобів керування. Це означає, що цілі забезпечення визначаються і пояснюються як частина вирішення проблеми.

Більш того, оцінювання ОС повинне включати усі засоби керування, які розглядаються як припущення при оцінюванні продуктів. Загалом види вимог до безпеки технічних засобів керування в ISO/IEC 15408-3 можуть бути розширені і на операційні засоби керування. Наприклад, концепція оцінювання проектної документації для технічних засобів керування перетворюється на оцінювання опису операційних засобів керування. Дії людей, які впроваджують операційні засоби керування, можна випробувати тим же шляхом, що і дії програми тестування технічних засобів керування.

Деякі вимоги ISO/IEC 15408-3, що стосуються розробки системи, можуть не застосовуватися в ОС, або їх оцінювання повинне відстрочуватися до настання фази інсталяції системи. Аналогічно, забезпечення безпеки в операційних засобах керування можна досягти у фактичному операційному середовищі, тоді як

технічні засоби керування аналізуються і випробовуються у своєму середовищі розробки.

Для оцінювання засобів керування в ОС необхідно узагальнити і змінити класи безпеки для технічної функціональності в ISO/IEC 15408-3. Області, в яких потрібні додаткові компоненти безпеки для роботи з ОС: – Вся архітектура безпеки та її компоненти; – Конфігурація складових компонентів ОС; – Політики керування, правила і процедури ОС; – Вимоги та правила для взаємодії з іншими довіреними і ненадійними ОС; – Моніторинг нетехнічних засобів керування на фазі експлуатації.

Існує п'ять основних способів досягнення гарантії в ОС: – Аналіз проекту ОС; – Тестування ОС; – Перевірка того, що ОС була встановлена і налаштована правильно; – Перевірка того, що ОС працює надійно; – Повторне використання результатів оцінювання.

Цей стандарт базується на триступеневому підході до встановлення необхідного рівня безпеки для операційної системи:

1. Оцінювання ризику, для визначення ризиків безпеки системи;

2. Зниження ризику, для вжиття заходів протидії або усунення ризиків безпеки шляхом вибору, застосування і оцінювання заходів безпеки;

3. Сертифікація, для підтвердження того, що залишкові ризики, які залишаються у рамках системи, після застосовування засобів керування, підходять для системи, яка буде використовуватися в живій роботі.

Модель оцінювання безпеки ISO/IEC 15408 виключає аналіз операційного середовища, яке оточує ІТ частину інформаційної системи. Як правило, безпека операційних систем залежить також від заходів безпеки ІТ, адміністративного або фізичного характеру. Тому потрібно визначити способи представлення та оцінювання таких вимог та засобів керування.

Використання онтологічного аналізу в стандартизації

Онтології ПрО, що характеризують певні стандарти, є потужним інструментом для обробки, аналізу та застосування знань, що містяться в цих стандартах, забезпечують їх автоматизовану обробку та інтеграцію [7].

Важливим питанням, пов’язаним з розробкою стандартів, є їх інтероперабельність та забезпечення їх автоматизованого пошуку та

порівняння. Тому в процесі розробки стандартів виникає необхідність в побудові їх семантичних метаописів. Такими метаописами можуть стати онтології ПрО стандарту, в яких формалізуються основні поняття ПрО, що відображаються в стандарті, та зв'язки між ними. Онтологічний підхід полегшує коректний переклад стандартів та дозволяє визначити, які саме близькі за значенням терміни потрібно використовувати у кожному окремому випадку [8].

Щоб виявити семантичну близькість між стандартами, потрібно зіставити їх онтології, знайти близькі за значенням поняття та на основі цього обчислити семантичну відстань між стандартами [9].

Саме онтології ПрО можуть стати джерелом знань для пошуку компетентних фахівців для розробки національних стандартів: за існуючим стандартом треба побудувати онтологію та порівняти її з онтологіями фахівців, які можна отримати з аналізу результатів їх науково-технічної та навчально-методичної діяльності, що відображається в їх публікаціях. Крім того, наявність онтології стандарту дозволяє надалі знаходити потрібні стандарти, здобувати з них необхідні користувачам знання та аналізувати їх вміст на семантичному рівні.

Ще одна сфера застосування онтологій стандартів – автоматизована побудова сфер компетенцій підкомітетів зі стандартизації, формалізованих через поєднання онтологій вже розроблених стандартів та наявність об'єктивних автоматизованих методів для класифікації нових стандартів до найбільш релевантного підкомітету. Аналіз відношень між онтологіям стандартів дозволить також визначити порядок розробки стандартів та визначення не тільки формальних, але й семантичних зв'язків між ними: приміром, стандарт, в якому визначаються певні терміни, потрібно розробляти раніше, ніж той, в якому ці терміни вже використовуються.

Розробка онтології стандарту

Розглянемо це на прикладі побудови онтології для стандарту «Інформаційні технології. Методи захисту. Оцінювання безпеки операційних систем». Усім визначеним у стандарті термінам мають відповідати класи онтології відповідної ПрО, яка описує цей стандарт та є підкласами класу «Термін стандарту». Спочатку в онтології створюється клас «Стандарт», який описує основні характеристики стандарту, такі як назва, код, тематика, рік прийняття, розробники, обсяг тощо.

Для відображення окремих термінів стандарту доцільно використовувати саме екземпляри онтології класу «термін стандарту», до яких можна застосовувати такі відношення, які реалізуються для екземплярів класів онтології (рис.1). Для інтероперабельного використання знань, які відображаються у стандартах, доцільно додати до опису класу «Стандарт» зв'язки даного стандарту з іншими. Крім того, можна вказати, які саме терміни відносяться до цього стандарту за допомогою властивості об'єктів «Стандарт використовує термін». Якщо онтологія створюється не для окремого стандарту, а для групи стандартів, то можна вказувати, що стандарт використовує термін, визначений в іншому стандарті, що описано даною онтологією. Це забезпечує однаковість термінів, що застосовуються, та спрощує розуміння вмісту стандартів. Щоб автоматизувати цей процес, доцільно використовувати спеціалізовані програмні засоби, орієнтовані на це [10].

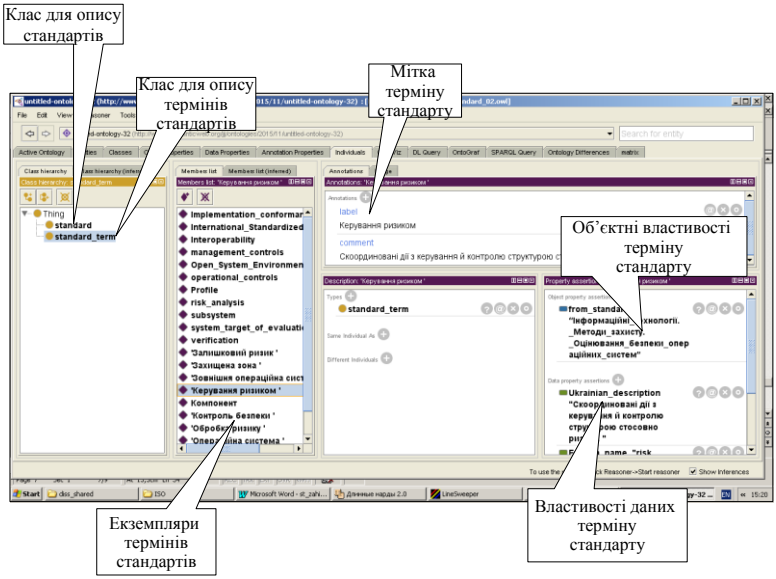


Рисунок 1 – Онтологія опису стандарту

Клас «Термін стандарту» має властивості даних (Data Properties) «Назва стандарту», «Назва терміну українською», «Назва терміну англійською», «Опис терміну українською», «Опис терміну

англійською» та «Примітка» типу «рядок символів» та властивості об'єктів (Object Properties) «Є підкласом» та «Є синонімом», які дозволяють встановлювати семантичні зв'язки між елементами термінами стандарту. Крім того, в таку онтологію можна додати семантичні зв'язки між термінами ПрО: приміром, якщо один термін є підкласом іншого або між ними існують специфічні для ПрО взаємини (бути компонентом, бути умовою тощо), які можна відобразити через об'єктні властивості відповідних класів та їх елементів. Щоб встановити однозначний зв'язок терміну з оригінальним англомовним терміном, доцільно використовувати англомовну назву в якості його імені, а для того, щоб забезпечити коректне розуміння україномовного контенту, рекомендується використовувати переклад терміну українською в якості його мітки. Це дозволить запобігти синонімічного дублювання термінів різними варіантами перекладу (в онтології неможливо створити два класи з одним ім'ям), але надати можливість використовувати кілька варіантів перекладу (в процесі вибору найбільш вдалого).

На рис.2 представлена онтологія ПрО, до якої відноситься терміносистема стандарту «Інформаційні технології. Методи захисту. Оцінювання безпеки операційних систем». Зв'язки між класами, екземплярами класів та їх властивостями в онтології стандарту відображаються візуально наступним чином.

Висновки та перспективи досліджень

В майбутньому результатом цієї роботи має стати створення глобальної семантичної мережі стандартів, яка пов'яже окремі національні та міжнародні стандарти; об'єкти, що використовують ці стандарти та посилаються на них (як матеріальні, так і інформаційні об'єкти); фахівців, що є експертами в сфері розробки стандартів, та організації різного рівня, що підтримують різні види діяльності, пов'язаної із розробкою та використанням стандартів. Інтероперабельне представлення знань та застосування відповідних технологій та форматів дозволить інтегрувати цю мережу знань із знаннями, представленими в Semantic Web.

На основі такої онтології можна виконувати семантичну розмітку природномовних текстів – як контенту окремих стандартів, пов'язаних з інформаційною безпекою, так і інших документів – приміром, описів конкретних систем, в яких реалізовані ці стандарти. Це має значно спростити пошук та аналіз таких документів та забезпечити можливість їх автоматичної обробки. Прикладом такої обробки може бути пошук операційних систем, що

задовольнять певним критеріям інформаційної безпеки відповідно до вказаного стандарту. При цьому користувачеві не потрібно буде самостійно відслідковувати зміни в останній редакції обраного стандарту або передивлятися опис кожної потенційно придатної системи – співставлення має виконуватися автоматизовано. Технологічною основою для такої семантичної розмітки та пошуку може стати середовище Wiki із семантичним розширенням [11].

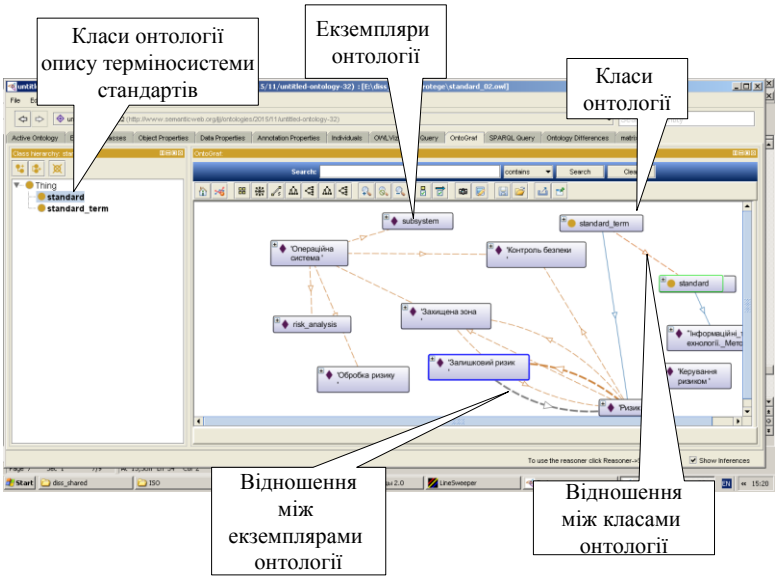


Рисунок 2 – Візуалізація зв'язків між класами, екземплярами класів та їх властивостями в онтології стандарту

Крім того, встановлення семантичної близькості між стандартами, що використовуються для розмітки, на основі онтологій дозволить визначити пріоритетні напрямки їх розвитку та забезпечити інформацію для подальшого їх вдосконалення.

Література

1. Закон України «Про стандартизацію». Верховна Рада України, Закон від 05.06.2014, № 13-15 –VII.
2. Державна система стандартизації. — К.: Держстандарт України, 1994.

3. ДСТУ 1.5:2003. Національна стандартизація правила побудови, викладання, оформлення та вимоги до змісту нормативних документів.
4. Цвілій О.О. Безпека інформаційних технологій: сучасний стан стандартів iso27k системи управління інформаційною безпекою // Телекомунікаційні та інформаційні технології. – No2, 2014. – С.73-79.
5. Грицик Н. Комп'ютерна термінологія та основні способи її перекладу/ Грицик Н. // XVIII-та Міжнародна науково-практична Інтернет-конференція «Проблеми та перспективи розвитку науки на початку третього тисячоліття», Переяслав-Хмельницький, 2013. – С.45-52.
6. ISO/IEC TR 19791:2010(en). – https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:19791:ed-2:v1:en
7. Gruber T., What is an Ontology? – http://www-ksl.stanford.edu/kst/what-is-an-ontology.html.
8. Гладун А.Я. Семантичні технології: принципи та практики (монографія)/ А.Я. Гладун, Ю.В. Рогушина.– К: Університет, 2016. – 314с.
9. Гладун А.Я., Рогушина Ю.В. Онтологічний підхід до проблем підвищення якості розроблення національних стандартів України // Стандартизація, сертифікація, якість, №2 (99), 2016. – С.19-28.
10. Лесько О.В. Рогушина Ю.В. Анализ семантики естественно-языковых законодательных документов с использованием онтологии предметной области // Проблеми програмування, № 4, 2015. – С.58-71.
11. Rogushina J. Semantic Wiki resources and their use for the construction of personalized ontologies // CEUR Workshop Proceedings 1631, 2016. – P.188-195.

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ ПРИ ПЛАНИРОВАНИИ МЕРОПРИЯТИЙ ПО ПРОТИВОДЕЙСТВИЮ ИНФОРМАЦИОННЫМ ОПЕРАЦИЯМ*

В.В. Цыганок,
Институт проблем регистрации информации НАН Украины

Современные методы поддержки принятия решений, основанные на наиболее полном использовании знаний в определенной предметной области (как формализованных, так и экспертных) предлагаются к применению при долгосрочном планировании. Предложен подход к построению стратегических планов, который включает иерархическую декомпозицию проблемы группой удаленно работающих экспертов под руководством инженера по знаниям, возможность использования разных шкал оценивания, позволяющая повысить достоверность результатов групповых экспертиз, целевое динамическое оценивание альтернатив, а также, метод оптимального распределения ресурсов. Применение подхода описывается на примере планирования противодействия информационным операциям.

Введение

Термин «информационные операции» (ИО) в современном противоборствующем мире приобрел значительное распространение в начале нынешнего века, когда информация стала важнейшим стратегическим ресурсом, недостаток которого приводит к значительным потерям во всех сферах жизни. Вероятно, что термин стал популярен после рассекречивания множества документов Департамента обороны США, где ИО упрощенно определялись как «действия, направленные на влияние на информацию и информационные системы противника, на защиту собственной информации и информационных систем». Затем, в «Дорожной карте информационных операций» [1], термин был уточнен как «Интегрированное применение основных средств радиоэлектронной борьбы, операций в компьютерных сетях, психологических

* Исследования выполнены в рамках проекта Ф73/23558 «Разработка методов и средств поддержки принятия решений при выявлении информационных операций» Государственного фонда фундаментальных исследований Украины

операций, военной маскировки и операций по обеспечению безопасности, в концепции со связанными с ними возможностями, с целью оказания влияния, разрушения, уничтожения или захвата у противника управления процессом принятия решений (как личностного, так и автоматизированного) при одновременной защите своих средств». Закладываемый в термин ИО смысл охватывает и раскрывает информационное влияние на массовое сознание (как на враждебное, так и на дружеское), влияние на информацию, доступную неприятелю и необходимую ему для принятия решений, а также на информационно-аналитические системы противника [2]. В современных условиях ИО, как неотъемлемая часть информационной войны, рассматривается в качестве нового вида боевых действий, активного противодействия в информационном пространстве, а информация при этом – в качестве потенциального оружия и цели для нанесения удара.

Принято различать два основных типа ИО – наступательные и оборонительные. Однако на практике большинство ИО являются смешанными, и большинство составляющих их процедур относятся одновременно к наступательным и оборонительным. Особенностью наступательных ИО (информационных атак) есть то, что объекты влияния таких операций определены и планирование основывается на довольно точной информации об этих объектах. Информационная атака чаще всего требует нахождения или создания информационного повода (для оборонительных ИО поводом может быть сама информационная атака неприятеля), раскручивание этого повода, т.е. пропаганда (в отличие от мероприятий контрпропаганды при оборонительных ИО), а также необходимость применения мероприятий по противостоянию информационному воздействию. Таким образом, ИО, вне зависимости от ее типа, можно разделить на следующие этапы: оценка, планирование, выполнение и завершающая фаза. Далее, в соответствии с целью данного исследования, более детально рассмотрим оборонительную ИО, соответствующую доктрине большинства прогрессивно развивающихся государств.

Типичная оборонительная ИО охватывает такие основные этапы:

- Оценка:
  - Анализ возможных уязвимостей (целей);
  - Сбор информации о возможных операциях;
  - Определение возможных «заказчиков» информационных влияний:

        - определение сфер общего интереса объекта и потенциальных «заказчиков»;
        - ранжирование потенциальных заказчиков за их интересами;
- Планирование:
    - Стратегическое планирование оборонительной операции (явное или неявное):
        - Определение критериев информационных влияний;
        - Моделирование информационных влияний с учетом:
            - связей объекта;
            - динамики влияния;
            - «особых» (критических) точек влияния;
        - Прогнозирование следующих шагов;
        - Расчет следствий;
    - Тактическое планирование контр-операций;
- Выполнение – реализация информационного влияния:
    - Выявление и «сглаживание» информационного повода;
    - Контрпропаганда;
    - Оперативная разведка;
    - Оценка информационной среды;
    - Корректирование информационного противодействия;
- Завершающая фаза:
    - Анализ эффективности;
    - Использование положительных результатов информационного влияния;
    - Противодействие отрицательным результатам.

Как можно видеть из предложенной детализации оборонительной ИО, основополагающим компонентом является стратегическое планирование. Очевидно, не существует единого «стандартного» плана проведения ИО. Можно лишь рассмотреть образцовую, полученную путем обобщения некоторых уже реализованных ИО, последовательность действий при их осуществлении. Причем выбор оптимального набора таких мероприятий в определенный момент времени зависит в первую очередь от наличия ресурсов для их проведения в этот текущий момент, а так же от результатов выполнения ранее выбранных мероприятий. Оптимальность здесь следует рассматривать с точки

зрения эффективности достижения целей проведения той, или иной оборонительной ИО.

Целью данного исследования есть усовершенствование имеющегося аппарата поддержки принятия решений (ППР) с учетом особенностей процесса стратегического планирования в слабо структурированных предметных областях. В контексте данной работы была поставлена более конкретная цель: разработать комплексную методику ППР, позволяющая повысить качество процесса стратегического планирования оборонительной ИО. Итак, в статье предлагаются методика формального построения стратегии ИО с привлечением группы специалистов, компетентных в этой области. На основе современных методов экспертной ППР предлагается возможность наиболее полного и без искажения получения знаний от специалистов, и использования их для построения адекватной модели предметной области. Работу методики предлагается показать на иллюстративном примере.

Сущность и общие этапы процесса построения стратегии

Как известно, в общем понимании стратегия представляет собой не детализированный план действий, рассчитанный на продолжительный период времени и направленный на достижение определенной главной цели. В то же время, план должен быть гибким, конструктивным, стойким к неопределенности условий среды и таким, что предусматривает конкретизацию путем декомпозиции этой главной цели.

В слабо структурированных предметных областях, к которым относятся управление, охрана окружающей среды, производство, социальная сфера и др., неотложной есть проблема построения долгосрочных не детализированных планов деятельности этих областей. Не возникает сомнений в том, что при создании таких стратегических планов нужно опираться на все имеющиеся знания в определенной предметной области. Поскольку, знание в каждой такой области не являются полностью формализованными и, поэтому, большей частью, находятся лишь в головах специалистов, то было бы безрассудно при планировании не использовать информацию, полученную от экспертов. Тем более, было бы неосмотрительным сводить оценки вариантов планирования только лишь к количественным (например, финансовым) показателям. Чтобы иметь реалистичные долгосрочные планы, их нужно адаптировать к неминуемым изменениям текущей ситуации и учитывать наличие ресурсов для их осуществления, необходимых в

каждый определенный момент. Поэтому стратегические планы могут быть рациональными лишь на определенном интервале времени.

Цель данного исследования – создать технологию, которая бы включала формальные механизмы построения стратегических планов в слабо структурированных предметных областях с привлечением групп экспертов и инженеров по знаниям.

Учитывая выше очерченные требования к стратегиям, а именно необходимость в реалистичных и динамических планах, предлагается при их построении использовать инструментарий распределения ограниченных ресурсов для определенных предложенных мероприятий. Ресурсы распределяются на заданный момент времени в зависимости от потенциального вклада определенного мероприятия в достижение стратегической цели. Фактически, результаты проведенной работы должны давать ответ на вопрос: «какие мероприятия должны быть выполнены при текущих условиях для наиболее эффективного достижения стратегической цели?».

Учитывая вышеуказанное, разработанная технология построения стратегии предусматривает несколько этапов.

1) Построение базы знаний (БЗ).

Этот этап реализован в виде веб-ориентированной программной системы, которая позволяет лицу, принимающему решение (ЛПР), инженерам по знаниям и экспертам работать удаленно для создания БЗ без необходимости собираться вместе.

Этот этап включает ряд под-этапов:

a. Подбор групп экспертов для проведения экспертизы.

Задача выбора экспертов в общем случае возлагается на ЛПР и на инженеров по знаниям. Причем, в рамках экспертизы при решении разных вопросов формируются разные группы специалистов, наиболее компетентных в каждой определенной области.

b. Построение (в ходе диалога с экспертами) иерархии целей, которая описывает предметную область.

На этом под-этапе ЛПР формулирует стратегическую цель, которая, в ходе проведения экспертиз инженерами по знаниям, подлежит декомпозиции на локальные цели (факторы), которые существенно влияют на ее достижение. В процессе декомпозиции, удаленно работающие в веб-ориентированной системе «Консенсус» [3] эксперты согласовывают свои мнения относительно состава множеств факторов влияния на ту или иную цель и приходят

к консенсусу в каждом вопросе. Инженеры по знаниям, имея функции организаторов экспертиз, для каждой декомпозиции локальной цели формируют отдельную группу экспертов. Созданная программная система позволяет разным экспертным группам работать одновременно, при этом каждый из экспертов может быть включен в состав разных групп. Преимуществом удаленного подхода является то, что в группе могут сотрудничать и предоставлять свои знания специалисты, которые могут быть несовместимыми при непосредственном личном контакте (работа предусматривает анонимность экспертов и это, в свою очередь, исключает диктаторское влияние суждений определенных "авторитетов"). Круг приобщенных к экспертизе специалистов может быть значительно расширен благодаря возможности в системе для каждого пользователя избирать наиболее удобный язык для общения, т.е. в экспертизе могут принимать участие специалисты, которые даже не смогли бы работать и общаться между собой без переводчика.

Решение о достаточном уровне детализации и прекращение дальнейшей декомпозиции стратегической цели принимают организаторы экспертизы в случае, когда нижний уровень иерархии целей будут составлять только лишь цели (факторы), представляющие собой готовые к реализации конкретные мероприятия (проекты).

Результатом данного этапа построения стратегии есть иерархическая структура, которая, в соответствии с мнением данной экспертной группы, в полной мере описывает предметную область.

Общий вид иерархии целей в СППР «Солон» [4] представлен на рис. 1. Демонстрационная версия СППР доступна на сайте лаборатории СППР (http://dss-lab.org.ua/).

с. Оценка экспертами относительных влияний целей в иерархии.

Относительное влияние каждой цели в графе иерархии определяется инженером по знаниям в случае наличия достоверных знаний об уровне этого влияния на достижение определенной цели, или же, в противном случае, – группой экспертов путем парных сравнений влияний целей (факторов).

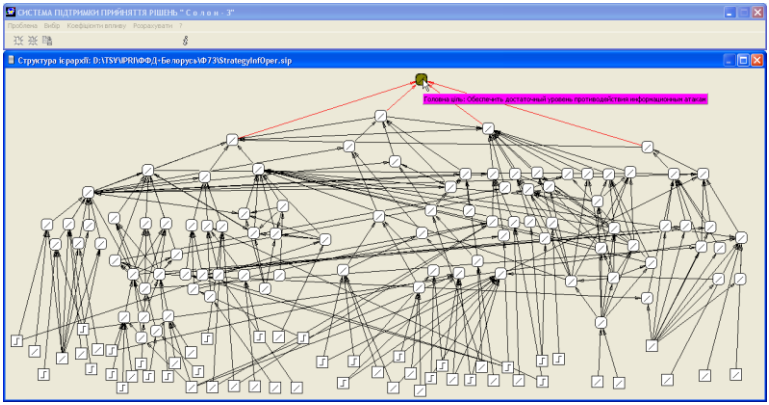


Рисунок 1 – Интерфейс СППР «Солон» и вид иерархии целей

Для повышения достоверности результатов экспертизы, разработано специальный программный инструментарий, который предоставляет эксперту возможность выполнять каждое отдельное парное сравнение в вербальной шкале оценивания, которая наиболее адекватно отображает представление/знание эксперта в обсуждаемом вопросе и уровень его/ее уверенности в собственных знаниях [5, 6]. Созданный для этого программный инструмент позволяет в процессе оценивания постепенно увеличивать уровень детализации шкалы, и уже окончательную оценку выполнить в наиболее приемлемой шкале (рис. 2).

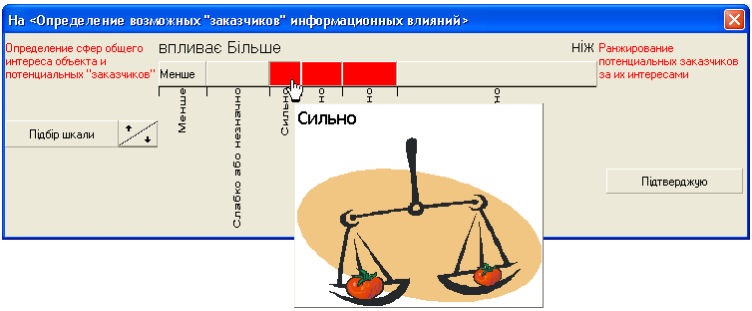


Рисунок 2 – Выбор экспертом уровня детальности шкалы парных сравнений

Результатом данного подэтапа есть относительные величины взаимных влияний целей, получаемые в результате агрегации индивидуальных экспертных оценок, выполненных в шкалах разной

подробности (детальности) и представленных в виде неполных матриц парных сравнений (МПС) влияний в рамках группы. Агрегацию предлагается производить разработанным В.В.Цыганком комбинаторным методом [7], преимущества которого в эффективности по сравнению с другими методами, подтверждено соответствующим экспериментальным исследованием [8].

Этот метод агрегации имеет несколько преимуществ над имеющимися подходами к обработке МПС:

- В методе максимально полно используется избыточность информации.
- Метод позволяет определять весомость альтернатив в случаях, когда часть элементов МПС отсутствуют (не заданы). Т.е., для определения весов альтернатив не является обязательным требованием наличие всех парных сравнений в матрицах. Необходимым условием есть лишь связность графа, соответствующего обобщенной МПС.
- Метод есть одноэтапным (в отличие от известных подходов, применяемых для вычисления весов в групповых методах оценивания [9]). Агрегация парных сравнений в таких групповых методах ППР представляет собой двухэтапную процедуру: или (1) сначала агрегируются индивидуальные МПС, а потом – на основе обобщенной матрицы вычисляется вектор весов альтернатив, или (2) сначала за каждой МПС вычисляется вектор весов, а потом все векторы агрегируются. В случае (1) согласованность всех индивидуальных МПС не гарантирует согласованности итоговой МПС. В случае (2) согласованность индивидуальных МПС не гарантирует согласованности векторов весов, вычисленных по каждой МПС. Если уровень согласованности недостаточный для корректного выполнения агрегирования и расчета весов, то в таком случае двухэтапность процедур делает невозможным организацию обратной связи с экспертами для повышения согласованности. При применении комбинаторного метода нет надобности в поэтапном достижении желательного уровня согласованности и, поэтому, не имеет места конфликт между двумя последовательными процессами согласования. Если нужно повысить согласованность парных сравнений, то определенные элементы индивидуальных МПС корректируются по согласию с экспертами, сформировавшими соответствующие МПС.

Заметим, что агрегацию допустимо выполнять лишь при достаточной согласованности экспертных суждений. Для оценки уровня согласованности парных сравнений предлагается

использовать Дважды-Энтропийный индекс согласованности [10], который на основе сформированного спектра экспертных оценок весов каждого из влияний определяет степень согласованности и отвечает всем поставленным требованиям, что положительно отличает его от других известных индексов. В случае недостаточной согласованности метод предусматривает возможность повышать уровень согласованности путем обратной связи с экспертами.

На этом построение БЗ заканчивается и предлагается этап построения оптимальной стратегии на основе знаний заложенных в БЗ.

2) Определение оптимальной стратегии.

Очевидно, что чем больше весомость определенного проекта, или мероприятия, тем существеннее он влияет на достижение стратегической цели. Поэтому, направление ресурсов на этот проект будет приносить более весомые и ощутимые результаты. В тот же время, не следует выделять на проект ресурсов меньше, чем необходимо для его старта и существования. Следовательно, в качестве оптимальной стратегии предлагается избирать оптимальный вариант распределения ресурсов между проектами (т.е. тот, который обеспечивает наиболее эффективное достижение стратегической цели).

Задача выбора оптимального распределения ресурсов между проектами является предметом отдельного исследования. Следует отметить, что поскольку проекты могут характеризоваться разными сроками реализации и, кроме того, цели могут иметь разные временные задержки влияния на главную цель, то оптимальное распределение ресурсов имеет место только для определенного заданного момента времени. Благодаря применению метода целевого динамического оценивания альтернатив [11, 12] в рамках заданного стратегического плана есть возможность оценить и сравнить совершенно разноплановые проекты/мероприятия: те, что дают моментальный (неотложный) эффект с теми, эффект от выполнения которых может появиться в далекой стратегической перспективе. Другой важный параметр, который характеризует проекты – диапазон необходимых объемов ресурсов. Например, если минимальный необходимый объем финансирования для осуществления проекта составляет 1 млн. грн., а запрашиваемый – 2 млн. грн., то нет смысла выделять на этот проект сумму, которая не принадлежит этому диапазону.

Учитывая указанные особенности, наиболее рациональным способом решения задачи распределению ресурсов между проектами

на заданный период усматривается целенаправленный перебор всех возможных распределений ресурсов с заданной точностью (предположим, до 10 тыс. грн.), например, с использованием Генетического алгоритма [13].

В зависимости от сложности предметной области и сформулированной цели, которая должна быть достигнута, процесс построения стратегического плана может быть более простым, или более сложным. Впрочем, предложенный математический аппарат и разработанные программные средства ППР дают возможность, опираясь на все доступные знания о предметной области, создавать довольно масштабные и содержательные, и главное, реалистичные перспективные планы.

Пример

Далее предлагается гипотетический пример, который показывает заключительные этапы процесса построения оптимального стратегического плана по обеспечению противодействия информационным операциям на 5-летнюю перспективу, при условии наличия финансовых ресурсов в объеме 200 млн. грн.

В рамках примера считаем, что иерархию с главной целью «Обеспечить достаточный уровень противодействия информационным атакам» уже построено и продолжается под-этап в) построения стратегического плана – оценка относительных влияний проектов на некоторую цель из графа иерархии целей.

Допустим, что оценивание на этом этапе примера выполняется группой из трех ровно-компетентных экспертов. Каждому эксперту формально предоставляется возможность определить наличие превосходства в каждой паре из 4-х проектов – выполнить ординальное сравнение („>”– больше; „<”– меньше), определиться с вербальной шкалой оценивания, выбрать количество делений для этой шкалы и, собственно, избрать номер конкретного деления.

В таблице 1 приведены данные экспертного оценивания важности мероприятий, входящих в состав цели «Реализация информационного влияния»: C_1 – Выявление и «сглаживание» информационного повода; C_2 – Контрпропаганда; C_3 – Оперативная разведка; и C_4 – Оценка информационной среды. Символом «*» в матрицах обозначены элементы, по которым эксперты, по той или иной причине, не предоставили информации.

Таблица 1. Пример экспертного оценивания относительных влияний проектов

Эксперт 1

Ординальное сравнение

	C_1	C_2	C_3	C_4
C_1	1	>	>	>
C_2		1	<	>
C_3			1	>
C_4				1

Количество делений шкалы

	C_1	C_2	C_3	C_4
C_1	1	5	9	9
C_2		1	5	9
C_3			1	7
C_4				1

Номер деления

	C_1	C_2	C_3	C_4
C_1	1	2	3	5
C_2		1	2	2
C_3			1	5
C_4				1

Номер шкалы

	C_1	C_2	C_3	C_4
C_1	1	1	3	1
C_2		1	2	1
C_3			1	2
C_4				1

Унифицированные значения парных сравнений

	C_1	C_2	C_3	C_4
C_1	1	2.500	1.732	5.000
C_2		1	0.739	2.000
C_3			1	3.138
C_4				1

Эксперт 2

	C_1	C_2	C_3	C_4
C_1	1	>	<	>
C_2		1	<	>
C_3			1	*
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	5	5	9
C_2		1	7	9
C_3			1	*
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	3	2	5
C_2		1	4	5
C_3			1	*
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	3	4	1
C_2		1	4	2
C_3			1	*
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	2.615	0.833	5.000
C_2		1	0.574	2.333
C_3			1	*
C_4				1

Эксперт 3

	C_1	C_2	C_3	C_4
C_1	1	*	>	>
C_2		1	<	<
C_3			1	>
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	*	8	9
C_2		1	9	3
C_3			1	9
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	*	4	8
C_2		1	2	2
C_3			1	3
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	*	2	3
C_2		1	3	5
C_3			1	1
C_4				1

	C_1	C_2	C_3	C_4
C_1	1	*	2.011	6.839
C_2		1	0.760	0.726
C_3			1	3.000
C_4				1

На основе унифицированных значений парных сравнений (нижний ряд матриц таблицы 1) вычисляются относительные веса влияний проектов (таблица 2).

Таблица 2. Рассчитанные относительные веса влияний проектов

Пометка проекта	Нормализованное значение веса
C_1	0.4455
C_2	0.1743
C_3	0.2919
C_4	0.0883

Для построения оптимальной стратегии на 5-летнюю перспективу воспользуемся инструментарием распределения ресурсов СППР «Солон» (см. рис. 3).

Из изображенной экранной формы можно увидеть, что для каждого проекта, который претендует на финансирование, вводятся экспертные оценки: минимально необходимое количество ресурсов для существования проекта (R min), процент выполнения проекта при минимальном финансировании (% min), количество ресурсов, которое запрашивается (R max) и запланированный процент выполнения при этом (% max – по обыкновению, равняется 100%). После выполнения расчетов (кнопка <Распределить>), количества

выделенных ресурсов отображаются в колонке «выделено».

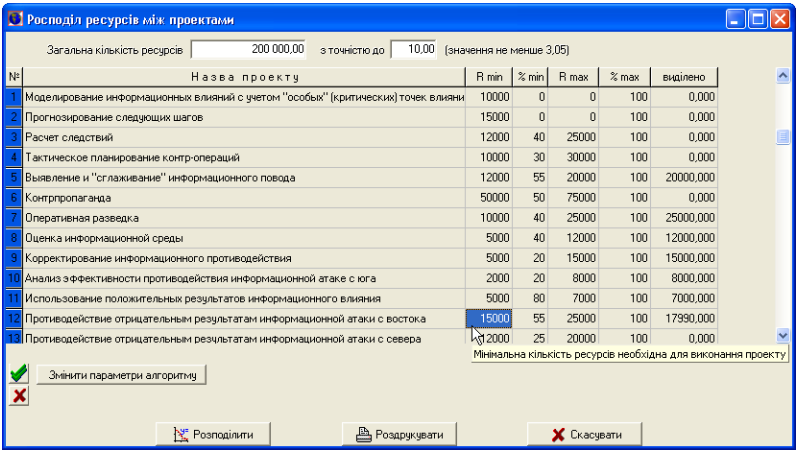


Рисунок 3 – Рассчитанное распределение ресурсов между проектами

Список рекомендованных действий для ЛПР в виде набора проектов с рассчитанными объемами финансирования будет базисом для оптимального стратегического плана по обеспечению достаточного уровня противодействия информационным атакам в 5-летней перспективе при условии обеспечения определенным количеством финансовых ресурсов.

Выводы

Предложена технология стратегического планирования в слабо структурированных предметных областях, которая основывается на использовании аппарата поддержки принятия решений. Преимуществами технологии есть возможность использования всех имеющихся знаний о предметной области (включая знания экспертов) и учет количественных и качественных факторов, влияющих на достижение стратегической цели, высокая достоверность групповых экспертиз за счет наличия механизма обеспечения достаточной согласованности экспертных данных, в том числе, неполных и полученных с использованием разных шкал оценивания, а также, возможность учета временных рамок выполнения проектов и наличия необходимых ресурсов. Данная технология позволяет на основе построенной базы знаний путем целенаправленного поиска, базирующегося на Генетическом алгоритме, среди множества вариантов определять рациональное на

некоторый момент времени распределение ресурсов для исходного множества возможных мероприятий с целью обеспечения наиболее эффективного достижения стратегической цели.

Перечисленные особенности делают технологию универсальным, удобным и гибким инструментом стратегического планирования. В качестве примера применения технологии рассмотрена текущая оценка конкретных мероприятий, направленных на противодействие возможным информационным операциям.

Дальнейшие исследования в данном направлении могут быть посвящены разработке новых алгоритмов определения оптимального распределения ресурсов в контексте заданной стратегического цели.

Литература

1. Information operations roadmap. – DoD USA, 30 october 2003. – 78 p. Retrieved 20/11/2016 from http://nsarchive.gwu.edu/NSAEBB/NSAEBB177/info_ops_roadmap.pdf
2. Горбулін В.П. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання / В.П. Горбулін, О.Г. Додонов, Д.В. Ланде. – К.: Інтертехнологія, 2009. – 164 с.
3. Свідоцтво про реєстрацію авторського права на твір №45894 Держ. служби інтелект. власності України. Комп'ютерна програма „Система розподіленого збору та обробки експертної інформації для систем підтримки прийняття рішень – «Консенсус»" / В.В.Циганок, П.Т.Качанов, О.В.Андрійчук, С.В.Каденко // від 03/10/2012.
4. Свідоцтво про держ. реєстрацію автор. права на твір №8669. МОН України Держ. деп. інтелект. власності. Комп'ютерна програма "Система підтримки прийняття рішень СОЛОН-3" (СППР СОЛОН-3) / В.Г.Тоценко, П.Т.Качанов, В.В.Циганок // зареєстровано 31.10.2003.
5. Циганок В.В. Агрегація групових експертних оцінок, що отримані у різних шкалах / В.В.Циганок // Реєстрація, зберігання і обробка даних. – 2011. – т.13. – №4.– С.74-83.
6. Tsyganok V.V. Using Different Pair-wise Comparison Scales for Developing Industrial Strategies / V.V.Tsyganok, S.V.Kadenko, O.V.Andriichuk // Int. J. Management and Decision Making. – 2015. – Vol. 14, No.3. – pp.224-250.
7. Циганок В.В. Комбінаторний алгоритм парних порівнянь зі зворотним зв'язком з експертом / В.В.Циганок // Реєстрація, зберігання і обробка даних. – 2000. – Т.2, №2. – С.92-102.

8. Tsyganok V.V. Investigation of the aggregation effectiveness of expert estimates obtained by the pair-wise comparison method // Mathematical and Computer Modelling. – 2010. – Vol.52(3-4). – pp. 538-544.
9. Forman E. Aggregating individual judgments and priorities with the analytic hierarchy process / E. Forman and K. Peniwati // European Jouranl of Operational Research. – 1998. – Vol.108. – pp.131-145.
10. Olenko A. Double Entropy Inter-Rater Agreement Indices / Andriy Olenko & Vitaliy Tsyganok // Applied Psychological Measurement. – 2016. – v.40(1). – pp.37-55.
11. Тоценко В.Г. Об одном подходе к поддержке принятия решений при планировании исследований и развития. Часть 2. Метод целевого динамического оценивания альтернатив / В.Г.Тоценко // Проблемы управления и информатики. – 2001. – №2. – С.127-139.
12. Циганок В.В. Удосконалення методу цільового динамічного оцінювання альтернатив та особливості його застосування / В. В. Циганок // Реєстрація, зберігання і обробка даних. – 2013. – Т. 15, № 1. – С. 90-99.
13. Holland J.H. Adaptation in natural and artificial systems. An introductory analysis with application to biology, control, and artificial intelligence. – London: Bradford book edition. – 1994. – 211 p.

МОДИФИЦИРОВАННЫЙ МЕТОД ФАКТОРИЗАЦИИ ФЕРМА И ИССЛЕДОВАНИЕ ЕГО ПРЕДЕЛЬНЫХ СВОЙСТВ

Е.В. Максименко,
Институт специальной связи и защиты информации НТУУ «Киевский политехнический институт»

В современном мире вопросы обеспечения информационной безопасности критически важных инфраструктур, оборонного и финансового секторов государства стоят на первом месте. В то же время, все большее внимание уделяется вопросам защиты информации в «не ключевых» областях экономики. Покупка товаров и оплата услуг через Интернет, электронная почта, IP-телефония, виртуальные сервера и облачные технологии, активно использующиеся во многих областях, также требуют обеспечения защищенности.

В большинстве зарубежных стран, включая США и страны Евросоюза, наиболее широкое применение в различных средствах и технологиях криптографической защиты современных ИТС получил алгоритм асимметричного шифрования RSA [1]. Оценка криптостойкости RSA алгоритма основывается на алгоритмической сложности решения задачи факторизации [2, 3, 4]. На сегодняшний день асимптотически самыми быстрыми из известных методов факторизации, которые на практике используются при криптоанализе RSA алгоритма, являются методы решета числового поля (GNFS) и квадратичного решета (QS), основанные на фундаментальных соотношениях алгоритма Ферма. Исходя из сказанного, можно предположить, что разработка новых или усовершенствование уже известных методов ускорения алгоритма Ферма позволят обеспечить снижение вычислительной сложности названных алгоритмов.

Пусть задано составное нечетное число N, которое следует разложить на множители:

$$N = p * q, \qquad (1)$$

где p и q некоторые нечетные числа, не обязательно являющиеся простыми.

В классическом методе Ферма для определения p и q решают уравнение

$$Y^2 = X^2 - N, \qquad (2)$$

где X и Y – целые положительные числа.

Если в уравнении (2) неизвестную X представить в виде $X = \left(\left[\sqrt{N}\right]+1\right)+ x = x_0 + x$, то решение уравнения (2) получают перебором значений x = 0, 1, 2, …, до тех пор, пока в (2) остаток $X^2 - N$ не окажется полным квадратом целого числа.

Ускорение работы такого алгоритма и обеспечение повышения эффективности метода Ферма возможно на основании следующих способов:

1. просеивания возможных значений x, т.е. отказ от проверки соотношения (2) (проверка остатка на полный квадрат) для случаев, когда иными, более простыми по вычислительной сложности, способами заранее установлено, что при таком значении x остаток $X^2 - N$ не может быть полным квадратом;

2. уменьшения коэффициента α за счет перехода от соотношения (1) к

$$kN = kp \cdot q, \qquad (3)$$

где коэффициент $\alpha_1 = \left|\sqrt{kp/q} - 1\right|$ более близок к нулю чем $\alpha = \left|\sqrt{q/p} - 1\right|$;

3. использования эффективных алгоритмов выполнения операций арифметических операций с многоразрядными числами. извлечения корня из $X^2 - N$ и возведения в квадрат больших чисел.

В докладе предложены способы ускорения метода Ферма за счет:

1. просеивания возможных значений x путем оптимизации выбора оснований модуля b при переходе от соотношения (2) к

$$\left((X \bmod b)^2 \bmod b - N \bmod b\right) \bmod b = (Y \bmod b)^2 \bmod b; \qquad (4)$$

2. обоснование выбора значений приближающих коэффициентов α при переходе от соотношения (1) к соотношению (3);

3. использования алгоритма прямого вычисления квадратного корня из $X^2 - N$, полученных на этапе предварительного просеивания, реализация которого не требует выполнения сложных операции умножения и деления больших чисел.

Литература

1. Горбенко І. Д. Прикладна криптологія: Терія. Практика. Застосування: монографія / І. Д. Горбенко, Ю. І. Горбенко. – Харків: Форт, 2012. – 880с.
2. Song Y. Yan. Cryptanalytic attacks on RSA / Song Y. Yan – Springer Science and Business Media, Inc. 2008. – P.255.
3. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. –М.: Триумф, 2002. – 816 с.
4. Кнут Д. Искусство программирования. Т. 2. Получисленные методы. – 3-е изд. – М.: Вильямс, 2007.
5. Винничук С.Д., Жилин А.В., Мисько В.Н. Алгоритм Ферма факторизации чисел вида N=pq методом прореживания / Электронное моделирование. – Т.36, №2. – 2014. – С. 3-14.
6. Терещенко А.Н. Быстрое вычисление квадратного и кубического корней без использования операций умножения и деления // Искусственный интеллект. – 2005. – Вип. 3. – С. 670-680.

ПІДХІД ДО ГЕНЕРАЦІЇ ОБ'ЄКТНИХ ПРОГРАМ РОЗВ'ЯЗКУ ЗАДАЧ ПРЕДМЕТНОЇ ОБЛАСТІ

Соколов В.В,
Інститут спеціального зв'язку та захисту інформації
НТУУ «КПІ ім. Ігоря Сікорського»

Одним із перспективних напрямків розвитку технологій програмування є автоматичне або автоматизоване генерування та виконання комп'ютерних програм по заданій постановці задачі на основі моделі предметної області (ПрО) та множини програмних компонентів [1-3]. Якщо в якості компонентів розглядати об'єкти класів, кожний з яких реалізує певну абстрактну функцію відображення значень вхідних властивостей у вихідні, і при цьому об'єкти здатні взаємодіяти між собою безпосередньо, утворюючи сполуки [4], то генерація потрібної сполуки об'єктів для розв'язку задачі стає цілком можливою.

Головною проблемою генерації програм є адекватне відображення семантики атрибутів і залежностей ПрО на властивості і функції відповідних класів.

Пропонується підхід до генерації об'єктних програм, при якому вважається заданими:

1. модель ПрО, яка задається у вигляді $P = (A, F)$, де A – домен атрибутів, F – множина функціональних залежностей (ФЗ) A;
2. множина класів, придатних для застосування в ПрО, яка задається у вигляді $C = \{CN\ (X) \rightarrow Y\}$, де CN – унікальне ім'я класу, X – множина імен вхідних властивостей класу, Y – множина імен вихідних властивостей, причому імена властивостей в X та Y унікальні тільки в межах класу;
3. постановка задачі в термінах ПрО $T = (In, Out)$, де In – підмножина атрибутів A, значення яких є вхідними даними, Out – підмножина атрибутів A, значення яких є шуканими результатами, для розв'язку якої потрібно згенерувати об'єктну програму.

Сутність запропонованого підходу полягає у наступному.

Для відображення семантики атрибутів та ФЗ ПрО на множину класів пропонується створення таблиці M, яка має $|F|$ рядків та $|A|+2$ стовпчиків, в якій кожний рядок містить назву ФЗ та ім'я відповідного класу, який реалізує цю ФЗ (останні два стовпчики). Кожний j-стовпчик, крім двох останніх, позначений іменем атрибута ПрО A_j та на перетині з i–рядком містить назву

відповідної властивості *i*-класу. Тобто, M_{ij} може приймати наступні значення:

- ім'я властивості з множини *X* класу C_i, якщо атрибут A_j є ключовим атрибутом ФЗ F_i;
- ім'я властивості з множини *Y* класу C_i, якщо атрибут A_j є залежним атрибутом ФЗ F_i.

Умовою мінімальної повноти відображення ПрО на множину класів є відповідність кожної ФЗ одному класу (багатьох до одного). Якщо для деяких ФЗ немає відповідного класу, то такі класи мають бути розроблені, причому в першу чергу розглядається можливість розробки потрібних класів шляхом інтеграції існуючих в межах одного класу. Верифікація повноти відображення тривіально виконується по таблиці *M*.

Пошук рішення задачі *T* полягає у визначенні такої підмножини класів в таблиці *M*, що залежність *In*→*Out* виводиться на основі аксіом Армстронга для ФЗ. Задача вважається розв'язаною, якщо в результаті редукції таблиці *M* залишились лише такі класи, що всі атрибути *In* є виключно входами, а всі атрибути *Out* є виходами хоча б одного класу.

Генерація програми полягає у створенні об'єктів відповідних класів (можливо, декількох об'єктів одного класу) та з'єднання їх у множину сполук, яка в сукупності являє собою програму розв'язку задачі. Програма може бути представлена у вигляді графічної схеми взаємодії об'єктів, безпосередньо у формі вихідного коду на об'єктно-орієнтованій мові програмування, наприклад, Сі++ або у вигляді декларативної мови, яку може автоматично виконати відповідне програмне середовище.

Таким чином, запропонований підхід дозволяє автоматизувати процес рішення множини обчислювальних задач в певній ПрО, підвищити коефіцієнт повторного використання програмного коду розроблених класів та зменшити кількість помилок. Очевидно, що для практичного застосування описаного підходу потрібно розробити метод проектування загальної структури системи рішення задач ПрО, який забезпечить формування домену унікальних атрибутів, оптимізацію множини ФЗ та проектування системи відповідних класів.

Література

1. Дорошенко А. Е., Иовчев В. А. Средства проектирования объектно-ориентированных программ на основе алгебры алгоритмики // Проблеми програмування. − 2012. − № 2-3. −

С. 241-250. − Режим доступу: http://nbuv.gov.ua/UJRN/ Progr_2012_2-3_31.
2. Лаврищева Е.М. Генерирующее и сборочное программирование. Аспекты разработки семейств программных систем // Кибернетика и системный анализ. — 2013. — Т. 49, № 1. — С. 129-144
3. Лаврищева Е.М., Грищенко В.Н. Сборочное программирование. Основы индустрии программных продуктов: Второе изд. – Киев: Наук. думка, 2009. -371с.
4. Соколов В.В. Технологія програмування активних динамічних сполук об'єктів // Тези доповідей на V науково-технічній конференції „Пріоритетні напрямки розвитку телекомунікаційних систем та мереж спеціального призначення". − К.: ВІТІ, 20-21.10.2010. – С. 232-233.

ОСНОВНІ ПРІОРИТЕТИ ВДОСКОНАЛЕННЯ СИСТЕМ ЗАПОБІГАННЯ ВТОРГНЕННЯМ В ІНФОРМАЦІЙНІ МЕРЕЖІ ОРГАНІВ ВІЙСЬКОВОГО УПРАВЛІННЯ

І.Ю. Субач[1,2], В.В. Фесьоха[1], В.Р. Прокопенко[2],
[1]Військовий інститут телекомунікацій та інформатизації,
[2]НТУУ «КПІ»

Фактична відсутність адміністративних обмежень в єдиному кібернетичному просторі в умовах ведення інформаційних війн (ІВ) створює нові передумови для несанкціонованих інформаційних руйнівних впливів (ІРВ) з метою порушення конфіденційності, доступності та цілісності інформаційних ресурсів, які підлягають обробці. Саме тому питання підвищення рівня захисту мережевих ресурсів від кібернетичних атак та спроб незаконного втручання в комп'ютерні системи (КС) та інформаційні мережі (ІМ) залишається відкритим.

Досвід розвинутих країн світу в області кібернетичного захисту ставить у пріоритет забезпечення безпеки саме ресурсів ІМ оборонного сектору держави, що тільки підкреслює рішення Ради національної безпеки і оборони України висвітлені в Доктрині інформаційної безпеки України.

Аналіз існуючих підходів та умов забезпечення захищеності ІМ органів військового управління (систем виявлення (запобігання) атак (вторгнень) (СВВ/СЗВ), аналізаторів мережевих протоколів, міжмережевого екранування, систем тестування навантаження, мережевого моніторингу, контролю цілісності тощо) показав доцільне застосування до запобігання та швидкого виявлення мережевих ІРВ саме СВВ/СВЗ.

Незважаючи на поширеність, використання, функціональні можливості покладених в їхній функціонал методів виявлення класифікованих атак (сигнатурний аналіз), параметричної реєстрації змін, евристичного і статистичного аналізу, нечіткої логіки, штучних нейронних мереж, штучних імунних систем, розпізнавання образів, аналізу поведінки об'єктів, методів заснованих на аномаліях для запобігання кібернетичним атакам, існують суттєві недоліки їх застосування:

досить високий рівень помилкових спрацьовувань та пропусків кібернетичних атак;

слабкий механізм виявлення нових кібернетичних атак;

більшість вторгнень неможливо визначити на початкових етапах;

практична відсутність змоги ідентифікації атакуючого та визначення цілі атаки;

слабкий механізм виявлення вже класифікованих атак, що використовують нові стратегії;

складність виявлення вторгнень у реальному часі з необхідною повнотою у високошвидкісних мережах;

значне завантаження систем при роботі в реальному часі;

неможливість інтерпретації адміністратором безпеки результатів класифікації поточної ситуації;

велика вартість побудови, налагодження, експлуатації та ремонту;

складність початкового "навчання" системи;

видача результату, точність ідентифікації якого не завжди відома та інших.

У зв'язку з цим, пріоритетами вдосконалення сучасних СЗВ є:

1. Відповідність концепції застосування нового покоління СЗВ (*NGIPS − Next generation intrusion prevention systems*) – застосування комплексного захисту від ІРВ та виявлення аномалій на основі кореляційного аналізу:

1.1 Поєднання переваг сигнатурного та інтелектуального аналізу даних виявлення вторгнень (гібридний підхід).

1.2 Впровадження ешелонної системи захисту з поєднанням переваг застосування міжмережевих екранів, СВА, СЗВ (ешелонний або багатошаровий підхід).

2. Зміна їхньої структури:

2.1 Додавання підсистеми підтримки прийняття рішень (ППР) – розробка спеціального математичного забезпечення для ППР адміністратором безпеки щодо виявлення та запобігання кібернетичним атакам на ІМ органів військового управління, з метою підвищення обґрунтованості та оперативності рішень, які він приймає в процесі виконання своїх функціональних обов'язків.

2.2 Додавання модуля візуального аналізу, в основу роботи якого покладено багатовимірне представлення даних – представлення інформації про ситуацію, що складається в ІМ у багатовимірному вигляді за багатьма параметрами та порівняння отриманого зображення з відомими графічними шаблонами кібернетичних атак, що знаходяться у базі даних (БД). Це дозволить адміністратору безпеки оперативно виявляти кібернетичні вторгнення в систему, навіть при відсутності значного практичного

досвіду та середньої кваліфікації, а також створювати нові графічні шаблони кібернетичних атак та завантажувати їх до БД системи.

3. Застосування нових моделей та методик виявлення та запобігання кібернетичним вторгненням, побудованих на основі теорії прийняття рішень, теорії нечітких множин, теорії баз даних, моделей та методів інтелектуального аналізу даних, а також методів інженерії знань.

Таким чином запропоновані шляхи удосконалення СВВ дозволяють не тільки підвищити оперативність та обґрунтованість прийняття рішень адміністратором безпеки в режимі реального часу по виявленню кібернетичних атак, а й є підґрунтям для реалізації нових механізмів ідентифікації кібернетичних атак та застосування їх під час реалізації систем виявлення вторгнень наступного покоління з метою реагування них на раніше невідомі типи кібернетичних атак.

ДОСТУПНІСТЬ – ГОЛОВНИЙ ЧИННИК ІНФОРМАЦІЙНОГО НАПОВНЮВАННЯ ВЕБ-САЙТІВ

О.М. Юдін,
Полтавський університет економіки і торгівлі

Веб-сайт є особливим комунікаційним каналом державної установи. Останнім часом все частіше розробники веб-сайтів згадують про таке поняття, як доступність [1]. Багато хто розуміє під цим терміном зручність використання сайту людьми з проблемами зору. Тобто то, як сайт читається програмами читання з екрану. Але за фактом доступність сайту - це набагато більш велика область. Під доступністю варто розуміти не тільки читабельність сайту і зручність його використання групами людей з певними обмеженнями, а взагалі зручність використання сайту звичайними користувачами: будь-який відвідувач сайту повинен зручно отримати необхідний контент. Доступність знаходиться в одному ряду з таким загальноприйнятими практиками як юзабіліті, адаптивний дизайн, SEO-оптимізація. Дослідження довели, що доступні сайти збільшили охоплення аудиторії, піднявшись на більш високі позиції в органічному пошуку. Крім того, доступність допомагає збільшити соціальну інтеграцію літніх людей, людей з країн, що розвиваються і сільських районів. У багатьох розвинених країнах можливість використовувати Інтернет відноситься до основних прав людини. Це означає, що всім соціальним групам людей повинні бути надані рівні умови доступу до інформації, зручності її отримання. Тому питання забезпечення належної якості комунікації, що створюється веб-сайтом, має для державної установи ключове значення.

В даний час оцінка якості комунікації сайту державного органу виконавчої влади здійснюється за визначеною методикою, що встановлює перелік параметрів, кожний з яких відображає окремий вид інформації [2]. Згідно з даною методикою веб-сайти міністерств та інших центральних органів виконавчої влади оцінюються за 27 параметрами, веб-сайти обласних адміністрацій – за 31 параметром. Кожний параметр, в свою чергу, оцінюється за такими коефіцієнтами:

1) K_p – коефіцієнт розміщення, визначає обов'язковість розміщення інформації;

2) $K_н$ – коефіцієнт наявності, визначає наявність на веб-сайті інформації, яка визначається параметром;

3) $К_п$ – коефіцієнт повноти, визначає рівень висвітлення своєї діяльності органом виконавчої влади за визначеним параметром та означає, що інформація, розміщена на веб-сайті з цього питання, є вичерпна та достатня для розуміння;

4) $К_а$ – коефіцієнт актуальності, визначає рівень відповідності інформації дійсності;

5) $К_д$ – коефіцієнт доступу, визначає рівень простоти та зручності пошуку інформації на веб-сайті;

6) $К_я$ – коефіцієнт якості, визначає рівень якості розміщеної інформації, розраховується як середнє значення критеріїв $К_п$, $К_а$, $К_д$.

Після встановлення значень коефіцієнтів для всіх параметрів, визначається показник наявності інформації на веб-сайті – $П_н$, а також показник якості інформаційного наповнення веб-сайту – $П_я$, за якими і оцінюється веб-сайт. Аналіз результатів застосування методики показав, що майже для всіх досліджених сайтів були отримані значення показника $П_н$ близькими к максимальному. Можна зробити висновок, що найбільшу значимість цей показник мав на початковому етапі існування сайтів державних органів виконавчої влади. На даний час показник втратив свою актуальність, став чисто формальним. Крім того, методикою припускається, що коефіцієнти $К_д$, $К_а$, $К_п$ мають однакову важливість. Разом з тим, проведений аналіз показує, що між ними існує певна впорядкованість і залежність: повнота має значення у випадку, коли інформація є актуальною, в свою чергу, актуальність інформації має значення, коли її можна без зайвих перешкод знайти на сайті. Отже, коефіцієнти потрібно впорядкувати з урахуванням їх важливості: $К_д > К_а > К_п$. Таким чином, методика потребує вдосконалення.

Дослідження умов рішення задачі оцінки інформаційного наповнення сайтів, дозволило зробити висновок, що дана задача є багатокритеріальною та залежною від експертної інформації. Аналіз методів рішення багатокритеріальних задач показав, що в даному випадку доцільно застосувати лексикографічний метод. Сутність методу полягає у виділенні спочатку з множини альтернатив найкращої альтернативи за найважливішим показником, що визначається коефіцієнтом доступу ($К_д$). Якщо такий сайт один, то вин вважається найкращим, якщо сайтів декілька, то з їх підмножини виділяються ті, які мають кращу оцінку за другим показником, що визначається актуальністю інформації ($К_а$). Якщо знову залишаються кілька варіантів, то перевагу отримує той, який має кращу оцінку за останнім показником – повнотою інформації ($К_п$).

Для врахування важливості інформації сайту було припущено, що за частотою звернень до неї користувачів її можна розбити на три групи: інформація, що затребувана частіше, середнє і рідко.

Відповідно, для кожної групи було введено ваговий коефіцієнт, що визначає її важливість. Для спрощення розрахунків було припущено, що перші десять параметрів характеризують інформацію, що затребувана частіше, наступні десять – середнє, останні – рідко. Формули для розрахунку показників мають вигляд:

$$П_д = (\frac{\sum_{i=1}^{10} K_{дi}}{10}) * \omega_h + (\frac{\sum_{i=11}^{20} K_{дi}}{10}) * \omega_m + (\frac{\sum_{i=21}^{30} K_{дi}}{10}) * \omega_l .$$

$$П_a = (\frac{\sum_{i=1}^{10} K_{ai}}{10}) * \omega_h + (\frac{\sum_{i=11}^{20} K_{ai}}{10}) * \omega_m + (\frac{\sum_{i=21}^{30} K_{ai}}{10}) * \omega_l .$$

$$П_n = (\frac{\sum_{i=1}^{10} K_{ni}}{10}) * \omega_h + (\frac{\sum_{i=11}^{20} K_{ni}}{10}) * \omega_m + (\frac{\sum_{i=21}^{30} K_{ni}}{10}) * \omega_l ,$$

де $П_д$, $П_a$, $П_n$ – відповідно показники доступності, актуальності та повноти інформації сайту; ω_h, ω_m, ω_l – вагові коефіцієнти для інформації сайту, що затребувана частіше, середнє і рідко; $K_д$, K_a, K_n – коефіцієнти оцінки певної інформації на сайті (певного параметру).

У формулу розрахунку показника якості інформаційного наповнення сайту $П_я$ також було введено вагові коефіцієнти, що визначають важливість показників $П_д$, $П_a$, $П_n$:

$$П_я = П_д * \omega_д + П_a * \omega_a + П_n * \omega_n ,$$

де $\omega_д$, ω_a, ω_n – вагові коефіцієнти показників доступності, актуальності й повноти. Розрахунок вагових коефіцієнтів було виконано за методом попарних порівнянь [3].

Таким чином, шляхи вдосконалення методики оцінки якості інформаційного наповнення сайтів державних органів виконавчої влади полягають у такому: відказатися від розрахунку показника наявності інформації на веб-сайті ($П_н$); впорядкувати коефіцієнти оцінки інформаційних параметрів за важливістю; найважливішим коефіцієнтом вважати коефіцієнт, що визначає доступ до інформації – $К_д$; оцінку параметру за певним коефіцієнтом здійснювати за десятибальною шкалою; для кожного сайту розраховувати показники $П_д$ – доступності, $П_а$ – актуальності, $П_п$ – повноти, $П_я$ – показник якості; при розрахунку показників враховувати важливість певного виду інформації з точки зору частоти звернень до неї за допомогою вагових коефіцієнтів; визначати найкращий сайт за допомогою лексикографічного методу [4].

За підходом, що пропонується перевагу отримують сайти, які в цілому не мають проблем із доступністю інформації, особливо з доступністю до важливої інформації. Такий підхід на перше місце ставіть потреби користувача в інформації і дозволяє веб-майстру зосередити зусилля, в першу чергу, на вирішенні проблем з доступом користувача саме до важливої інформації на сайті державного органу виконавчої влади. В зв’язку із запровадженням в Україні електронного урядування, прогнозується зростання кількості звернень громадян до сайтів органів влади, що, в свою чергу, обумовлює зростання актуальності питання доступності інформації на сайтах даної категорії. Нажаль, існуюча методика, застосовує спрощений підхід до визначення такого важливого показника інформації сайту, як доступність. Вимоги до доступності сайту визначаються відповідним стандартом, який докладно описує всі вимоги, містить посилання на роз'яснення та технології, що використовуються, а також основні помилки [1]. Це дозволяє повністю спиратися на нього при аналізі доступності веб-сторінок і їх адаптації. Крім того, в загальному доступі є великий інструментарій для напівавтоматичного виявлення помилок: валідатори і доповнення для браузерів. Враховуючі вище зазначене, наступним кроком вдосконалення методики оцінювання інформаційного наповнення веб-сайтів державних органів виконавчої влади буде впровадження в процес оцінювання вимог до доступності з боку визначеного стандарту та існуючого інструментарію.

Література

1. Руководство по обеспечению доступности веб-контента (WCAG) 2.0 [Електронний ресурс] – Електрон. дані. – Режим доступу: https://www.w3.org/Translations/WCAG20-ru/ – Назва з екрана.
2. Порядок проведення моніторингу інформаційного наповнення веб-сайтів органів виконавчої влади [Електронний ресурс] – Електрон. дані. – Режим доступу: http://www.publicpolinfo.gov.ua/ informational_policy/informational_society – Назва з екрана.
3. Саати, Т. Метод анализа иерархий [Текст] / Т. Саати. – М. : Радио и связь, 1993. — 278 с.
4. Юдін О.М., Яначек С.П. Вдосконалення методики оцінювання інформаційного наповнення веб-сайтів [Електронний ресурс] – Електрон. дані. – Режим доступу: http://dspace.puet.edu.ua/ bitstream/ 123456789/4042/1/Vdos1713.pdf

ІСТОТНІ ОСОБЛИВОСТІ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ В КІБЕРПРОСТОРІ

Ю.І. Хлапонін,
Київський національний університет будівництва і архітектури

Розвиток інформаційних послуг вимагає рішення завдань ефективного управління інформаційними ресурсами з одночасним розширенням функціональності інформаційно-телекомунікаційних систем (ІТС).

Одночасно з розвитком технологій постає питання безпеки в інформаційно-телекомунікаційних мережах. З точки зору забезпечення безпеки найбільш важливими властивостями мереж є: конфіденційність (використання інфраструктури або її частини); цілісність (інфраструктури); доступність (служб та сервісів); спостереженість (за використанням інфраструктури або її частини); прихованість (використання та управління інфраструктурою) [1].

Захист інформації це діяльність, яка спрямована на забезпечення безпеки оброблюваної в ІТС інформації та ІТС в цілому, і дозволяє запобігти або ускладнити можливість реалізації загроз, а також знизити величину потенційних збитків внаслідок реалізації загроз. Основним завданням захисту інформації є протидія порушенню таких властивостей як: конфіденційність **інформації**; цілісність **інформації**; доступність використання ІТС та оброблюваної **інформації**; спостереженість **за діями користувачів** та керованість ІТС.

На даний час існує кілька визначень поняття кіберпростір. Наприклад, кіберпростір – це інформаційне середовище (простір), яке виникає (існує) за допомогою інформаційно-телекомунікаційних систем під час взаємодії людей між собою, взаємодії інформаційно-телекомунікаційних систем та управління людьми цими системами [2] або кібернетичний простір (кіберпростір) – середовище, утворене організованою сукупністю інформаційних процесів на основі взаємопоєднаних за єдиними принципами та правилами інформаційних, телекомунікаційних та інформаційно-телекомунікаційних систем [3] та інші. На сьогоднішній день проект Закону [3] “Про кібернетичну безпеку України”, поданий ще 04.06.2013, досі не прийнятий в якості Закону.

Кіберзахист – це діяльність, яка спрямована на забезпечення безпеки кіберінфраструктури. Кіберінфраструктура може характеризуватися рядом властивостей. З точки зору забезпечення

безпеки найбільш важливими властивостями кіберінфраструктури є: конфіденційність **(використання інфраструктури або її частини);** цілісність **(інфраструктури);** доступність **(служб та сервісів);** спостереженість **(за використанням кіберінфраструктури або її частини);** прихованість **(використання та управління кіберінфраструктурою).**

Якщо для захисту інформації найбільш важливими заходами є запобігання загрозам конфіденційності та цілісності, то в кіберпросторі основні зусилля повинні бути направлені на запобігання загрозам доступності служб (в кіберпросторі атаки з метою порушення доступності реалізуються простіше) та спостереженість за використанням інфраструктури (або її частини).

При побудові системи захисту інформації властивість "доступність" розглядається насамперед як доступність самої **інформації**, а доступність використання визначеної АС – в контексті захисту конкретної **інформації**. Наприклад, в АС класу 1 ненавмисне або навмисне форматування жорсткого диску призводить до того, що всі дані, які зберігаються на носієві інформації стають недоступними, хоча фізично з носія не видаляються.

В кіберпросторі, насамперед в великих розподілених системах, рідко застосовується безпосередній доступ до жорсткого диску віддаленого комп'ютера. Доступ до інформації відбувається шляхом формування та обробки запитів до відповідних служб, які функціонують на різних серверах в цьому кіберпросторі. В питанні запобігання загрозам спостереженості в кіберпросторі, серед визначених задач, найбільш відповідальною та складною є задача взаємної аутентифікації і авторизації користувачів або окремих елементів кіберінфраструктури, до яких визначений користувач намагається отримати доступ.

Таким чином, безпека в кіберпросторі має істотні відмінності від забезпечення безпеки конкретної інформації в будь-якій визначеній системі. На сьогоднішній день, враховуючи необхідність взаємодії та функціональної сумісності окремих структур, задіяних в зоні проведення антитерористичної операції з зовнішніми користувачами (до яких в першу чергу будуть відноситися силові структури), можуть порушуватися окремі лінії зв'язку, можлива втрата боєздатності окремими елементами управління військами, порушення керованості озброєнням та військовою технікою та ін.

Література

1. Хлапонін Ю.І. Загальні характеристики загроз в кіберпросторі / Ю.І. Хлапонін, В.В. Овсянніков, Н.А. Паламарчук – Пріоритетні напрямки розвитку телекомунікаційних систем та мереж спеціального призначення: VI наук.-практ. сем. Військового інституту телекомунікацій та інформатизації НТУУ "КПІ", 20 жовтня 2011 р.: тези доп. – К., 2011. – С. 157.
2. http://cybercop.in.ua/index.php/naukovi-statti/80-naukovi-statti/176-ponyattya-kiberprostoru-ta-kiberzlochiniv
3. Проект Закону "Про кібернетичну безпеку України", № 2207а від 04.06.2013. http://w1.c1.rada.gov.ua/pls/zweb2/webproc4_2?id=&pf3516=2207%E

АНАЛІЗ ТА МОНІТОРИНГ ТЕЛЕКОМУНІКАЦІЙНОЇ МЕРЕЖІ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ

Хлапонін Ю.І.[1], Жиров Г.Б.[2]
[1] Київський національний університет будівництва та архітектури
[2] Військовий інститут Київського національного університету ім. Т. Шевченка

Світові інфокомунікаційні мережі надають відповідні сервіси великій кількості кінцевих абонентів, які можуть бути розташованими на великих територіях, у межах усієї земної кулі. Інфокомунікаційні сервіси повинні задовольняти критеріям якості, а для цього необхідно постійно забезпечувати необхідну смугу пропускання каналу мережі та постійно підтримувати в працездатному стані програмно-апаратні вузли мережі, які можуть бути розташовані на великій території. Однією з головних вимог щодо якості мережі є забезпечення користувачам можливості доступу до ресурсів всіх комп'ютерів, об'єднаних в мережу. Для виконання даної вимоги необхідний постійний аналіз стану та моніторинг телекомунікаційних мереж.

Під аналізом телекомунікаційної мережі розуміється процес порівняння поточного і нормального станів телекомунікаційної мережі в заданий часовий інтервал. Моніторинг телекомунікаційної мережі включає в себе: спостереження, відбір за визначеними ознаками, оброблення та реєстрація сеансу зв'язку в мережі. Результатом моніторингу є збір первинних даних про роботу мережі: статистика щодо кількості циркулюючих в мережі пакетів різних протоколів, про стан портів концентраторів, комутаторів і маршрутизаторів і т. п.

В даній статті на основі проведеного аналізу практичного використання телекомунікаційних систем встановлена необхідність більш широкого і науково обґрунтованого впровадження статистичних методів їх аналізу і моніторингу на основі відкритої потокової інформації.

Аналіз телекомунікаційних мереж, що використовує технології передачі даних ATM 1/0, Fast Ethemet l/0, Fast

Ethernet 4/0 показав, що ефективність роботи мережі залежить від наступних характеристик: завантаження каналу на вході і виході (байт); число пакетів на вході і виході; число помилок в їх реєстрації; завантаження процесора (%); обсяг вільної пам'яті процесора і системи введення-виведення для маршрутизатора (байт). Найбільш інформативним параметром є завантаження каналу.

Запропоновано перспективний підхід до організації обробки неявних форм подання знань, який базується на застосуванні технології нейромережевих структур. Архітектура нейронних мереж дозволяє реалізувати їх із застосуванням технологій надвисокого ступеня інтеграції. Доведена можливість успішного використання нейронних мереж та їх аналогових моделей для вирішення задачі апроксимації неперервних функцій багатьох змінних та прогнозу процесів, які відбуваються у телекомунікаційних мережах протягом часу. Також розроблені процедури первинної обробки значень параметрів телекомунікаційних мереж для подальшого використання в якості вхідних даних для нейронної мережі. Розроблені процедури дозволяють більш детально розглядати і аналізувати динаміку зміни інформаційних потоків, циркулюючих в мережах, та визначати характерні особливості випадкових послідовностей, а впровадження нейронних мереж дозволяє прогнозувати поведінку мережі в залежності від сезонності та тренду.

1. Вступ

На сьогоднішній час глобальні мережі (Wide Area Networks, WAN) використовуються для того, щоб надавати інфокомунікаційні сервіси великій кількості кінцевих абонентів, розкиданих по великій території – в межах регіону, держави, континенту або всієї земної кулі. Зважаючи на велику протяжність каналів зв'язку потрібно забезпечувати необхідну смугу пропускання каналу, постійну підтримку в працездатному стані програмно-апаратні вузли мережі, які розкидані по великій території.

Головною вимогою до мереж, є забезпечення користувачам можливості доступу до ресурсів всіх комп'ютерів, об'єднаних в

мережу [1,2]. Дослідження і аналіз стану та постійний моніторинг телекомунікаційних мереж є актуальною задачею.

Моніторинг телекомунікацій являє собою: спостереження, відбір за визначеними ознаками, оброблення та реєстрацію сеансу зв'язку в мережах телекомунікацій із застосуванням системи моніторингу мережі телекомунікацій [3]. На етапі моніторингу виконується процедура збору первинних даних про роботу мережі: статистика про кількість циркулюючих в мережі пакетів різних протоколів, про стан портів концентраторів, комутаторів і маршрутизаторів і т. п. За станом первинних параметрів телекомунікаційної мережі визначають вищевказані параметри функціонування мережі. Назвемо поточним станом телекомунікаційної мережі сукупність поточних параметрів телекомунікаційної мережі, виміряних в заданий часовий інтервал. Таким чином, під моніторингом телекомунікаційної мережі будемо розуміти збір і фіксацію поточного стану телекомунікаційної мережі в заданий часовий інтервал.

Назвемо нормальним станом телекомунікаційної мережі сукупність параметрів телекомунікаційної мережі, які встановлені регламентом її функціонування в заданий часовий інтервал. Таким чином, під аналізом телекомунікаційної мережі будемо розуміти процес порівняння поточного і нормального станів телекомунікаційної мережі в заданий часовий інтервал. Під результатом аналізу будемо розуміти сукупність зафіксованих значень розбіжностей між поточним і нормальним станом телекомунікаційної мережі по кожному параметру.

Завдання аналізу вимагає більш активної участі людини і використання таких складних засобів, як експертні системи, що акумулюють практичний досвід багатьох мережевих фахівців. Отриману інформацію про роботу телекомунікаційної мережі можна аналізувати з різним ступенем глибини або деталізації.

Стандартами ISO/ITU-T, визначені п'ять основних функціональних груп завдань системи управління [4–6].

Перша група: управління конфігурацією мережі і ім'ям – ці завдання полягають в конфігурації параметрів, як окремих елементів мережі, так і телекомунікаційної мережі в цілому. Для елементів мережі за допомогою цієї групи завдань визначаються мережеві адреси, ідентифікатори (імена), географічне положення. Для мережі в цілому управління конфігурацією зазвичай починається з побудови карти мережі, тобто відображення реальних зв'язків між елементами мережі і зміни зв'язків між цими елементами мережі.

Статистичні результати роботи телекомунікаційної мережі можуть служити основою при прийнятті рішення в рамках даної групи завдань управління.

Друга група: обробка помилок – дана група завдань включає виявлення, визначення та усунення наслідків збоїв і відмов мережі, в тому числі і на основі статистичних результатів роботи мережі.

Третя група: аналіз продуктивності та надійності – завдання цієї групи пов'язані з оцінкою на основі накопичувальної статистичної інформації таких параметрів, як час реакції системи, пропускна здатність реального або віртуального каналу зв'язку, інтенсивність трафіку в окремих сегментах і каналах мережі, імовірність спотворення даних при їх передачі через мережу, а також коефіцієнт готовності мережі. Функції аналізу продуктивності і надійності мережі потрібні як для оперативного управління мережею, так і для планування розвитку мережі.

Четверта група: управління безпекою – завдання цієї групи включають в себе контроль доступу до даних при їх зберіганні і передачі через мережу. Базовими елементами управління безпекою є процедури аутентифікації користувачів, призначення і перевірка прав доступу до ресурсів мережі, управління повноваженнями і т.д. При вирішенні завдань управління даної групи слід враховувати результати статистичної обробки інформації про атаки на мережу і спробах несанкціонованого доступу до її ресурсів.

П’ята група: облік роботи мережі – завдання цієї групи займаються реєстрацією часу використання різних ресурсів мережі – пристроїв, каналів і транспортних служб, ***в тому числі з урахуванням статистичних параметрів роботи телекомунікаційної мережі.***

2. Визначення статистичних характеристик та головних компонент в системі статистичного аналізу телекомунікаційних мереж

2.1. Визначення статистичних характеристик для аналізу телекомунікаційної мережі

Нехай $x_1,\ldots,x_n$ – вибірка з n значень змінної x. Якщо дану вибірку впорядкувати по зростанню величин, то вийде так званий ряд порядкових статистик $x_{(0)} <.... < x_{(n)}$ (номер в дужках (0), (1), ..., (n) називається рангом відповідного значення).

У статистичній системі аналізу телекомунікаційної мережі будемо оцінювати такі одновимірні статистичні характеристики.

Характеристики положення:

• середнє значення

$$m(x) = \bar{m}_x = \sum_{i=1}^{n} x_i / n ;$$

• медіана

$$med(x) = \begin{cases} x_{(\frac{n-1}{2+1})}, & \text{якщо } n \text{ непарне} \\ \dfrac{x_{\frac{n}{2}-1} + x_{\frac{n}{2}+1}}{2}, & \text{якщо } n \text{ парне} \end{cases}$$

тобто оцінка медіани є точкою, ліворуч і праворуч від якої знаходиться однакове число точок вибірки;

• серединна точка (центр найкоротшої половини). Нехай *xL* и *xR* відповідно координати лівої і правої кінцевих точок "найкоротшої половини" (визначення дано далі в характеристиках розкиду); тоді серединна точка визначається як

$$Mid(x) = xL + (xR - xL)/2 .$$

Медіана і серединна точка є характеристиками положення центру розподілу, більш стійкими до наявності викидів (аномальних спостережень), ніж середнє значення. Серединна точка є також оцінкою моди розподілу змінної *x*.

Характеристики розкиду

• дисперсія

$$s^2(x) = s_x^2 = \frac{1}{n}(x_i - m(x))^2;$$

• стандартне відхилення *s*, дорівнює квадратному кореню з дисперсії.

2.2. Визначення головних компонент в системі статистичного аналізу телекомунікаційних мереж

Перші *q* головних компонент $z_1,..., z_q$ багатовимірної ознаки *X* визначаються як лінійні ортогональні нормовані комбінації вихідних показників, тобто

$$z_j(X) = u_{1j}(x_1 - m_1) + ... + u_{pj}(x_p - m_p);$$

$$\left.\begin{aligned}\sum_{i=1}^{p} u_{ij}^2 = 1,\ (j=\overline{1,q}) \\ \sum_{i=1}^{p} u_{ij}u_{ik} = 0,\ (j,k=\overline{1,q}, j\neq k)\end{aligned}\right\} \quad (1)$$

де m_i – середнє значення ознаки x_i, а в якості міри інформативності q – мірної системи показників $(z_1(X), z_2(X),..., z_q(X))$ взята величина:

$$I_q(Z(X)) = \sum_{i=1}^{q} Dz_i \Big/ \sum_{j=1}^{p} Dx_j \quad (2)$$

де Dz_i та Dx_j – дисперсії відповідних показників [7].

Таким чином, головні компоненти – це система лінійних ортогональних комбінацій вихідних змінних, яка характеризується тим, що дисперсії цих комбінацій мають екстремальні значення. Так, першим головним компонентом є нормована лінійна комбінація вихідних змінних (сума квадратів коефіцієнтів нормованої лінійної комбінації дорівнює одиниці) з найбільшою дисперсією, тобто

$$Dz_i = \max_{|U|^2=1} D(U^{'}, X - M)\ .$$

Це означає, що перша головна компонента орієнтована вздовж напрямку найбільшого розкиду точок даної сукупності.

Друга головна компонента має найбільшу дисперсію серед всіх лінійних комбінацій виду (1), некорельованих з першою головною компонентою. Вона являє собою проекцію на напрям найбільшого розкиду спостережень серед напрямків, перпендикулярних першої головної компоненти, і т.д.

Обчислення векторів коефіцієнтів $U_1,...,U_q$ лінійних комбінацій, що відповідають основним компонентам засноване на тому, що вектори $U_1,...,U_q$ є власними векторами коваріаційних матриць досліджуваної сукупності. Коваріаційна матриця зазвичай нам невідома, і ми можемо використовувати тільки її оцінку *S*. Далі розглядаються оцінки головних компонент та головні компоненти вибірки (матриці даних).

Як власні вектори матриці *S* вектори *U* задовольняють рівняння:

$$SU_j = I_j U_j, \quad j = \overline{1,q} \tag{3}$$

Відповідні власні числа дорівнюють дисперсії (строго кажучи, оцінкам дисперсії) головних компонент. Якщо тепер власні вектори впорядкувати в порядку убування власних чисел $l_1 \geq l_2 \geq,...,l_q$ то i-й головній компоненті відповідає власний вектор U_i з власним числом I_i. Компоненти вектора U_i, тобто величини U_{ij}, $j = \overline{1,p}$ називаються навантаженнями i-ї головної компоненти (фактора) на змінні X_j, $j = \overline{1,p}$. Сам вектор U_i, є вектором навантажень або, на ряду з z_i, i-ю головною компонентою. Вектори навантажень можна розглядати як колонки матриці навантажень U.

Методи обчислення власних чисел і власних векторів симетричної матриці детально викладені, наприклад, в роботі [8]. В системі статистичного аналізу телекомунікаційної мережі використовується ***QR***-метод з попереднім приведенням матриці до трехдіагонального виду.

Всього існує p власних векторів матриці S *i*, отже, p головних компонент. Мають місце наступні співвідношення.

$$\det S = \prod_{i=1}^{p} l_i, \qquad SpS = \sum_{i=1}^{p} l_i \tag{4}$$

де *det S i Sp S* позначають відповідно визначник і слід матриці ***S.***

Формула (2) може бути записана у вигляді

$$I_q(Z) = \sum_{i=1}^{q} l_i \Big/ SpS, \tag{5}$$

тобто величина $I_q(Z)$ дорівнює частці сумарної дисперсії змінних X_j, $j = \overline{1,p}$ (сліду матриці S), що "пояснюється" першими q головними компонентами. Чим ближче значення цього критерію до 1, тим менше буде спотворена картина взаємного розташування спостережень при переході в простір головних компонент.

Використання головних компонент найприродніше і плідне, коли всі змінні X_j, $j = \overline{1,p}$ мають загальну фізичну природу і виміряні в одних і тих самих одиницях. Саме такий випадок і досліджується для цілей статистичного аналізу телекомунікаційної мережі. Результат може істотно залежати від вибору масштабу виміру. Тому зазвичай переходять до безрозмірних величин,

нормуючи значення змінних, що також характерно для статистичного аналізу телекомунікаційної мережі [8]. Наприклад, вдалими нормуваннями є нормування розкидом значень змінної і нормування стандартним відхиленням. У разі, коли основна мета застосування головних компонент полягає в описі структури залежності між змінними, перехід до нормованих даних стає практично необхідною умовою.

2.3. Використання головних компонент

Процедура головних компонент може бути використана, наступним чином:

а) для скорочення розмірності даних з мінімальною втратою інформації в сенсі критерію (2); дані скороченої розмірності можуть бути використані, наприклад, в процедурі кластер-аналізу; зазвичай з цією метою відкидають головні компоненти з мінімальними власними числами; найбільш важливими для відбору є величини частки сліду (дисперсії), що відповідає головній компоненті і накопиченій частці сліду;

б) для візуального аналізу особливостей розташування точок вихідної вибірки; з цією метою використовується інтерактивна графічна система;

в) як різновид факторного аналізу; в цьому випадку виділені головні компоненти розглядаються як оцінки деяких прихованих чинників (тут виникає проблема змістовної інтерпретації виділених чинників).

3. Застосування нейромережевих технологій для обробки статистичної інформації телекомунікаційних мереж

Аналіз експлуатації телекомунікаційних мереж показує, що на даному етапі розвитку забезпечити їх ефективну роботу досить складно. Практика використання гетерогенних телекомунікаційних систем та комп'ютерних мереж пов'язана з недостатньою їх прозорістю, складністю, організаційними обмеженнями і специфікою, що визначає необхідність більш широкого і науково обґрунтованого впровадження статистичних методів їх аналізу і моніторингу на основі відкритої потокової інформації [4–7], особливо при вирішення складних задач та виникнення надзвичайних ситуацій [8].

Проведений аналіз робіт [9, 10] показує, що для вирішення поставлених завдань доцільно та необхідно застосовувати інтелектуальні технології.

На сьогоднішній час, швидкими темпами розвиваються технології створення нейромережевих структур. Архітектура нейронних мереж дозволяє реалізувати їх із застосуванням технологій надвисокого ступеня інтеграції. Різниця елементів мережі невелика, а їх повторюваність величезна. Це відкриває перспективу створення універсального процесора з однорідною структурою, здатного переробляти різноманітну інформацію і не вимагає обов'язкової наявності програми обробки, достатня тільки постановка задачі.

Таким чином, аналіз опрацьованої літератури дозволяє зробити висновок, що є області застосування нейронних мереж в телекомунікаційних системах, які розкриті не в повному обсязі.

Нейронні мережі являють собою один з найбільш універсальних підходів для побудови правил класифікації і прогнозу [2, 10]. Однак їх основним недоліком є досить складна процедура налаштування архітектури мережі і оцінки її параметрів, які забезпечують прийнятну якість прогнозу (класифікації).

3.1. Структура нейронної мережі

У статистичній системі аналізу телекомунікаційної мережі використовуються мережі з декількома впорядкованими шарами нейронів. При цьому взаємодія між нейронами, що належать до одного і того ж шару, відсутня [10]. Нейрони кожного шару отримують дані (сигнали) від нейронів попереднього шару, обробляють їх і передають результат обробки до наступного шару. Винятком є нейрони вхідного шару. Число нейронів у вхідному шарі дорівнює числу змінних відібраних для вирішення завдання прогнозу або класифікації, так що кожному нейрону відповідає одна з змінних. Таким чином, сигнали, що надходять на вхідний шар, являють собою значення цих змінних.

Сигнали, на виході останнього (вихідного) шару нейронів є результат роботи нейронної мережі. Тому, якщо нейронну мережу передбачається використовувати для класифікації об'єктів в одну з M груп, то число нейронів у вихідному шарі має дорівнювати M.

3.2. Обробка сигналів нейронами проміжних шарів

На вхід кожного нейрона будь-якого проміжного шару надходять сигнали від усіх нейронів попереднього шару. Обробка сигналів полягає в тому, що спочатку проводиться зважене підсумовування сигналів, що надійшли. Якщо ця зважена сума

перевищує певний поріг, то вихідний сигнал нейрона дорівнює 1, в іншому випадку – 0.

Таким чином, якщо $z_{jl},...,z_{jn_{k-1}}$ сигнали, що надійшли на вхід *j*-го нейрона *k* шару від n_{k-1} нейронів попереднього шару, а $W_{jl}^{(k)},...,W_{jn_{k-1}}^{(k)}$ – ваги, даних нейронів, то для формування суми використовують вираз:

$$S_k^{(k)} = W_{jl}^{(k)} z_{jl} + ... + W_{jn_{k-1}}^{(k)} z_{jn_{k-1}} \quad (6)$$

Нехай $t_j^{(l)}$ – граничне значення. Вихідний сигнал даного нейрона визначається як величина $\theta(S_j^{(k)} - t_j^{(k)})$, де функція стрибка $\theta(x) = 1$ якщо $x > 0$; та 0, якщо $x \le 0$, тобто якщо $S_j^{(k)} > t_j^{(k)}$.

На практиці функція стрибка $\theta(x)$ замінюється певною функцією. Найбільш часто використовується логістична функція:

$$L(x) = \frac{e^x}{1+e^x} \quad (7)$$

Оскільки на вхід кожного нейрона в *k*-му шарі надходять сигнали від усіх нейронів попереднього (*к*-1)-го шару, кількість вагових коефіцієнтів і граничних значень для обробки вхідних сигналів усіма нейронами дорівнює $(n_k + 1)n_{k-1}$, де n_k – число нейронів в *k*-му шарі. Сукупність вагових коефіцієнтів всіх нейронів *k*-го шару утворює матрицю зв'язку $\mathbf{W}^{(k)}$ між *k*-м та *(k*-1*)*-м шарами.

3.3 Створення нейронної мережі

Для створення нейронної мережі, яку можна було б використовувати для класифікації багатовимірних об'єктів або для передбачення значень незалежної змінної (в разі завдання регресійного аналізу або прогнозу часових рядів), що особливо важливо в разі статистичного аналізу телекомунікаційної мережі, необхідно:

- задати архітектуру мережі, тобто задати кількість шарів і кількість нейронів в кожному з них;
- оцінити вагові коефіцієнти для всіх нейронів мережі (ваги в матрицях зв'язку $\mathbf{W}^{(k)}$).

Архітектура мережі. Нейронна мережа повинна містити як мінімум два шари: вхідний і вихідний. Кількість нейронів у вхідному шарі визначається кількістю використовуваних змінних. Якщо всі

змінні – безперервні кількісні, то число нейронів просто дорівнює числу змінних. Якщо ж серед змінних є номінальні, то для кожної такої змінної, наприклад, змінної *y*, відводиться *(l*-1*)* вхідних нейронів, де *l* – число градацій (категорій) змінної *y* в *i*-му нейрону (з цих *(l*-1*)* нейронів) та присвоюється значення 1, якщо змінна приймає *i*-е значення, і 0 в іншому випадку.

Отже, кількість нейронів у вхідному шарі однозначно визначено, як тільки обрані активні змінні для вирішення задачі класифікації, регресії або прогнозу.

Кількість нейронів у вихідному шарі визначається типом розв'язуваної задачі, при вирішенні задач класифікації об'єктів в одну з *M* груп, вихідний шар містить *M* нейронів. При вирішенні задачі прогнозу (регресії) кількість нейронів дорівнює числу залежних змінних. Число проміжних шарів і кількості нейронів в кожному з них задається дослідником перед етапом оцінки вагових коефіцієнтів.

Оцінка вагових коефіцієнтів (навчання). Для оцінки вагових коефіцієнтів в статистичній системі аналізу телекомунікаційної мережі застосовні процедури безумовної оптимізації за методом сполучених градієнтів. Для вирішення проблеми локальних мінімумів використовується генерація деякої кількості стартових точок.

3.4. Процес збору інформації про роботу телекомунікаційної мережі

Розглянемо процедуру первинної обробки значень параметрів телекомунікаційної мережі підприємства. Для аналізу мережевого трафіку на сервері системи управлінської інформації підприємства системним адміністратором мережі здійснювався збір даних з допомогою протоколу SNMP. На сервері використовувалися такі технології передачі даних, як: ATM 1/0, Fast Ethemet l/0, Fast Ethernet 4/0. Дані про функціонування телекомунікаційної мережі реєструвалися за допомогою чотирьохбайтового лічильника з інтервалом 5 хв. Для аналізу були визначені наступні характеристики:

- завантаження каналу на вході і виході (байт);
- число пакетів на вході і виході;
- число помилок в їх реєстрації;
- завантаження процесора (%);
- обсяг вільної пам'яті процесора і системи введення-виведення для маршрутизатора (байт).

Збір та реєстрація параметрів телекомунікаційної мережі здійснювалися протягом тривалого періоду часу за допомогою чотирьохбайтових лічильників, при переповненні лічильників відбувалося їх обнулення (або скидання), це призводило до пилкоподібності в поданні значень параметрів телекомунікаційної мережі і не дозволяло безпосередньо використовувати відомі методи статистичної обробки інформації.

3.5. Процедура перетворення первинної інформації в випадкову послідовність

Дана процедура передбачає аналіз безпосередньо первинної інформації «накопичувального» типу. Випадковою величиною, в даному випадку, є момент «обнулення». Для нормальних періодів роботи мережі можна розглядати також число «обнулень» n_k, для k-го періоду часу T_k або ж частоту «обнулень». В цьому випадку характеристиками випадкової послідовності є функції розподілу або ж їх числові характеристики. Першим кроком моніторингу є візуалізація даних – графічне відображення реєстрації інформації в процесі надходження, де P – показання лічильника, в байтах (або кількість пакетів); t – час реєстрації інформації.

Серед сильних сторін даного дослідження необхідно відзначити те, що показано можливість застосування нейронної мережі у статистичній системі аналізу параметрів телекомунікаційної мережі.

Додаткові можливості, що забезпечують досягнення мети дослідження, криються в тому, що нейронні мережі та їхні аналогові моделі можуть бути успішно використані для вирішення задачі апроксимації неперервних функцій багатьох змінних та прогнозу процесів у часі.

Складнощі у впровадженні отриманих результатів дослідження пов'язані з тим, що на сьогодні майже завжди моделювання нейронних мереж проводиться на цифрових обчислювальних машинах архітектурою Неймана. Це має велику кількість переваг: надзвичайну універсальність, велику точність (а отже передбачуваність алгоритму), стабільність та багато інших. Але за всі ці переваги доводиться платити дуже малою швидкістю та продуктивністю.

З іншого боку сучасні операційні підсилювачі можуть працювати на частоті у кілька гігагерц. Максимальна частота обрахунку функції операційним підсилювачем у декілька разів менша за його граничну частоту. Якщо кількість зв’язків модельного

нейрона збільшити у два рази, продуктивність цифрової моделі зменшиться приблизно в таку ж кількість разів, проте продуктивність аналогової майже не зміниться.

Висновки

1. При вирішенні задач аналізу та моніторингу мереж в першу чергу розглядається первинна потокова інформацію та вирішуються такі завдання, як апроксимації функцій, прогнозування, оптимізація та ін. Для вирішення таких завдань можна та необхідно використовувати нейронні мережі.

2. Після аналізу інформації про роботу телекомунікаційної мережі, що використовує технології передачі даних ATM 1/0, Fast Ethemet l/0, Fast Ethernet 4/0 встановлено, що ефективність роботи мережі залежить від наступних характеристик: завантаження каналу на вході і виході (байт); число пакетів на вході і виході; число помилок в їх реєстрації; завантаження процесора (%); обсяг вільної пам'яті процесора і системи введення-виведення для маршрутизатора (байт). Найбільш інформативним параметром є завантаження каналу.

3. Розроблено процедуру перетворення первинної інформації телекомунікаційної мережі, сутність якої полягає в перетворенні вихідної інформації з кількості байтів (пакетів) в частоти скидання або «обнулення» за певний період.

4. Доведено можливість використання нейронних мереж для аналізу процесів, що протікають у телекомунікаційних мережах в часовій області.

Література

1. Кокс Д. Статистический анализ последовательности событий / Д. Кокс, П. Льюис. – М.: Мир, 1969. – 312 с.
2. Комашинский В.И. Нейронные сети и их применение в системах управления и святи / В.И. Комашинский, Д.А. Смирнов. М.: Горячая линия – Телеком, 2002. – 94 с.
3. Винницкий В.П Методы системного анализа и автоматизации проектирования телекоммуникационных сетей: (Монографія) / В.П Винницкий, В.В Хиленко. – К.: Интерлинк, 2002. – 192 с.
4. Домрачев В.Г. Нечеткие методы в задачах мониторинга сетевого трафика / В.Г.Домрачев, Д.С. Безрукавный, Э.В. Калинина, И.В. Ретинская Ж, Информационные технологии, JV23 2006 , С. 2–10.
5. ETSI ETR 003 «Сетевые аспекты. Общие аспекты качества обслуживания и эффективности сети».

6. ITU-T Recommendation E.800 (09.08) Definitions of terms related to quality of service (Визначення термінів, що стосуються якості послуг).
7. Бугай А.И. Некоторые особенности моделирования сетевого трафика. Теоретические проблемы информатики и ее приложений: / А.И. Бугай, Э.В.Калинина, И.В.Ретинская // Сб. науч. тр. Под ред. проф. А.А.Сытника – Саратов: Изд-во Сарат. ун-та, 2003. – Вып 5. – С. 30–41.
8. Корнейчук Н.П. Аппроксимация с ограничениями / Н.П. Корнейчук, А.А. Лигун, В.Г. Доронин. – Киев: Наукова думка, 1982. – 252 с.
9. Zakasovskaya E.V. Restoration of point influences by the fiber-optical network in view of a priori information / E.V.Zakasovskaya, V.V.Fadeev // SPIE Proc. APCOM. – 2007. – Vol. 6675.
10. Хлапонін Ю.І. Побудова апроксимаційної функції на основі алгоритму зворотного розповсюдження помилки як методу навчання штучних нейронних мереж / І.І. Бех, С.О. Новак, Ю.І. Хлапонін // Вісник інженерної академії. – 2016. – № 1. – С. 198–201.

РОЗРОБКА СЦЕНАРІЇВ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ ОНТОЛОГІЧНОЇ МОДЕЛІ

А.В. Бойченко,
Інститут проблем реєстрації інформації НАН Украины

Розглянуто використання онтологічної моделі в задачі сценарного моделювання інформаційної безпеки. Запропонований підхід дозволяє в автоматизованому режимі на базі аналізу вхідного пакету документів вирішувати задачу розробки та дослідження сценаріїв інформаційної безпеки на об'єкти, які відповідають вибраним ключовим поняттям.

Сценарний аналіз найбільш повно відповідає завданням дослідження і прогнозування поведінки складних процесів, до яких належить і інформаційна безпека.

Пропонується технологія розробки сценаріїв інформаційної безпеки, заснована на онтологіях, отриманих із масиву документів, які описують певну тематичну область. Технологія включає наступні етапи:

1. Аналіз масиву документів, які описують предметну область.
2. Виділення ключових понять: окремих слів, біграм (пар слів), триграм (трійок слів).
3. Формування графу зв'язків між ключовими поняттям за допомогою вагових критеріїв[1]. Як такий граф може розглядатись граф входження понять у близькі фрагменти тексту.
4. Експертну оцінку взаємовпливу понять та присвоєння числових значень графу предметної області.
5. Аналіз отриманої мережі понять за допомогою комплексу пакетів моделювання (Gephi, Protégé) та розробленого на мові Perl програмного комплексу, що дозволяє інтегрувати засоби моделювання.

Отримані в процесі моделювання онтології розглядаються як когнітивні карти, що дозволяють застосовувати розвинені технології побудови сценаріїв здійснення впливів.

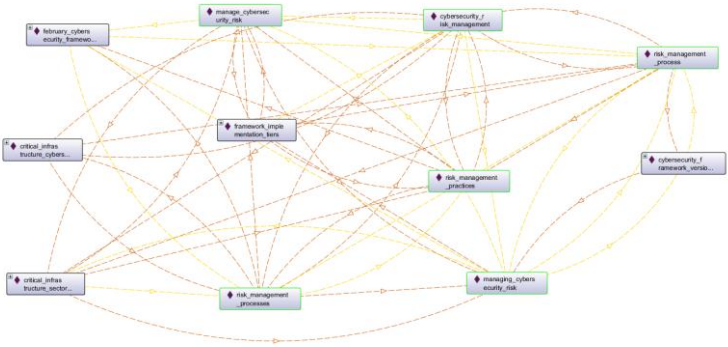


Рисунок 1 – Приклад отриманої онтології

Когнітивна карта – це знаковий орієнтований граф: G =< V,E >, де: V – множина вершин V_i є V, i = 1,2,...k, які є елементами досліджуваної системи; E – множина дуг e_{ij} є E, i, j = 1,2,...N, які відображують взаємозв'язок між вершинами V_i і V_j ; вплив V_i на V_j може бути позитивним, коли збільшення (зменшення) одного фактора приводить до збільшення (зменшення) іншого; негативним, коли збільшення (зменшення) одного фактора веде до зменшення іншого, чи бути відсутнім (0).

Ребра графа мають ваги +1 або -1, скорочено позначаються знаками "+" чи "-". Знак + позначає позитивний зв'язок, знак - позначає негативний вплив. Вага шляху дорівнює добутку ваг його ребер, тобто позитивний, якщо число негативних ребер в ньому парне, і негативний, якщо це число непарне. При позитивній зв'язку зростання чинника-причини призводить до зростання фактора-слідства, а при негативному зв'язку зростання чинника-причини призводить до зменшення фактора-слідства. Якщо ж від вершини a_i до вершини b_j ведуть як позитивні, так і негативні шляхи, то питання про характер впливу фактора a_i на фактор b_j залишається невизначеним.

Аналіз когнітивної карти включає дослідження змісту складових її блоків, цільових і керуючих факторів, аналіз шляхів і циклів, взаємозв'язків між. При формуванні сценаріїв експерт вибирає, які вершини в отриманій моделі слід виділити, а та які зв’язки між вершинами враховувати, а які – ні. Із отриманої таким

чином множини можливих сценаріїв поведінки моделі обираються ті, що забезпечують прийнятний рівень захищеності.

Література

1. Ландэ Д.В.,Снарский А.А. Подход к созданию терминологических онтологий // Онтология проектирования, 2014. – № 2(12). – С. 83-91.
2. Ланде Д.В., Бойченко А.В. Використання моделей предметних областей у задачах сценарного аналізу // Міжнародна науково-практична конференція "Інтелектуальні технології лінгвістичного аналізу": Тези доповідей. – Київ: НАУ, 2016. – С. 9.

ВЫЯВЛЕНИЕ СОБЫТИЯ, ЕГО СУБЪЕКТА И ОБЪЕКТА В ТЕКСТОВЫХ ДОКУМЕНТАХ

С.В. Прищепа,
Институт проблем регистрации информации НАН Украины

В данной статье рассматриваются задачи и проблемы выявления событий и их фигурантов в неструктурированных или слабоструктурированных текстах. Проанализировали некоторые из современных подходов в выявлении событий и предложили свой метод выявления новых событий. В статье также представлен алгоритм данного метода. Для выявления событий и связанных с ними сущностей, используется классификация текстов SVM методом. Для выявления событий в нашем алгоритме происходит разбиение текстов на униграммы, биграммы и триграммы. Каждому из них присваиваем вес для конкретной предметной области по TF-IDF (или его разновидностями) методу. В случае работы с текстами очень большого и малого размера одновременно – предложено проводить двойную нормализацию частоты встречаемости слова.

Для выявления событий, их субъектов и объектов создаются специальные шаблоны правил разбора предложений по определенной тематике и словари индикаторов событий по определенной теме, которые заполняются экспертом с учетом различных лингвистических признаков.

Индикаторы тематического события представлены словарями (униграммами, биграммами и триграммами) и могут одновременно входить сразу в несколько тематик. При этом используется не один большой словарь индикаторов события по определенной теме, а два. Это позволяет строить более сложные шаблоны условий и вывести показатели точности экстрагирования событий на более высокий уровень.

Рассмотрено проблему и предложено вариант её решения в задаче выявления фигурантов события – физических или юридических лиц.

При наличии хорошего обучающего корпуса и грамотного выбора признаков, мы рассчитываем на качество выявления

события и его фигурантов на уровне 0.7 по F1-мере для текстов на русском / украинском языке.

В качестве дальнейшего улучшения модели выявления новых событий и их фигурантов, можно использовать кросс-текстовое выявление событий, их субъектов и объектов. Это делается с помощью выделения сразу нескольких документов за определенный промежуток времени с одинаковыми или связанными событиями и принятия решения по событию уже по кластеру предложений содержащим информацию о конкретном событии.

Введение

Рост количества информации в сети Интернет, информационный шум, большое количество "фейковой" информации – все это приводит к росту временных и других затрат на поиск и выявление ценной информации в сети Интернет, как для различных государственных органов, так и для конкурентной бизнес-разведки. Разработка и внедрение системы автоматического выявления событий, их субъектов и объектов – решение данной проблемы. Выявление событий, их субъектов и объектов, а также других сущностей, которые связаны с событием – одна из самых важных и ценных задач в разборе не структурированных или слабо структурированных текстов.

Актуальность разработки методов экстрагирования событий с информационных потоков растет во всем мире, подтверждением этому является количество научных публикаций и патентов по данной теме за прошедший год. В данной статье, мы презентуем наше виденье и наработки по экстрагированию новых событий и их фигурантов из не структурированных или слабо структурированных текстовых данных. В своем подходе мы используем классификацию текстовых документов, выделение и взвешивание слов как терминологической основы, создаем сеть понятий и выявляем событие, субъект события и его объект с помощью двойных словарей и специальных шаблонов условий. Наша предварительная экспериментальная оценка показывает, что такой подход осуществим, и позволяет обнаруживать события, выявлять их субъектов и объектов с высокой точностью.

Событие – значительное происшествие, явление или иная деятельность как факт общественной или личной жизни. Нас интересуют события, где есть хотя бы один фигурант, который его

совершил – субъект или фигурант над которым оно совершалось – объект.

Мы рассматриваем задачу извлечения событий, как выявление индикаторов (триггеров) событий заданных типов и выявление их аргументов и связи с фигурантами в предложении, а затем отобранная информацию об этих событиях должна быть распознана и объединена в единое представление для каждого обнаруженного события [1].

Это важная и сложная задача обработки естественного языка, так как одинаковое событие может присутствовать в различных выражениях, и при этом может выражать различные события в зависимости от контекста. Также события часто имеют вложенность – одно событие может привести ко второму, к примеру, событие “преступление”, приводит к “расследование”, а оно, в свою очередь, может привести к событию “задержание” или “арест”[2].

Проблемы обнаружения событий и их фигурантов:

- Огромное количество возможных имен (названий) фигурантов;
- Большое количество фигурантов в тексте, которые не относятся напрямую к событию;
- Один индикатор события может выражать различные события в разном контексте;
- Сложные предложения с большим количеством индикаторов событий;
- Большое количество псевдо событий (Шум);
- Различные лексические и семантические сложности разбора.

Некоторые подходы идентифицируют и классифицируют триггеры событий используя большие наборы функций [3], без использования шаблонов. И хотя данные методы могут быть весьма полезными, использование шаблонов в выявлении событий – до сих пор является незаменимым во многих случаях и это показывают оценки точности выявления по F1 у многих научных работ(4).

Два основных подхода в выявлении событий:

Совместный подход – одновременно прогнозирует индикаторы событий и их аргументы в предложениях, с точки зрения структуры. Глобальные особенности выделения зависимостей между индикаторами событий и аргументами – доступны в совместном подходе.

Конвейерный подход – сначала выполняет прогнозирование индикатора (триггера) события, а затем идентифицирует его аргументы в отдельных этапах. Локальные особенности инкапсулирования характеристик для отдельных заданий (индикатор события и маркировка ролей аргументов) – конвейерный подход.

Преимущества совместной системы (совместного подхода) носят двоякий характер: смягчение распространения ошибки от вышестоящего компонента (идентификации триггера) к нисходящему классификатору (аргументу), и получения выгод от взаимозависимости между индикатором событий и ролями аргументов в глобальных функциях. [4]

Для разбора модели экстрагирования событий мы выделили массив документов по предметной области “Безопасность”, а именно по трем конкретным категориям – убийства и покушения, кражи, ДТП – по 100 материалов для каждой категории.

Для классификации документов проводится их первичный стемминг, убираются стоп-слова из словаря стоп-слов – слова «и», «но», «по», «из», «под» и т.д.

Классификация документов проводится методом опорных векторов (Support Vector Machine). Метод работает не только для слов, но и для n-грамм, например, биграмм.

Это существенно повышает точность. При этом, естественно, при переходе к биграммам объем словаря растет. Выбор был остановлен на SVM благодаря точным показателям классификации и наличию бесплатных программ SVM^{light} и SVM^{perf} от University of Dortmund [5].

Конкретно в данном материале, в качестве примера рассмотрим массив текстовых данных, определенный как “убийства и покушения”. В текстах с событиями такого рода – часто упоминаются различного рода лица, которые совсем не причастны к событию, а лишь заявили о нем общественности или выступают свидетелями.

После выделения массива документов по предметной области классификатором, происходит разбиение текста на предложения, очистка от знаков препинания (html и других тэгов) и выделение слов как терминологической основы. Затем происходит разбиение текстов на униграммы, биграммы и триграммы. Каждому из них присваиваем вес для данной предметной области по TF-IDF методу или его вариантами. В случае работы с текстами очень большого и малого размера одновременно проводим двойную нормализацию частоты встречаемости слова таким образом, что:

TF термина A = 0.5 + 0.5 * (Количество раз, когда термин A встретился в тексте / Количество раз, когда встретился самый частотный термин этого текста).

Данный вид измерения TF нивелирует возможную ошибку, вызванную влиянием размера текста.

При разборе текста, уже с определенной категорией (темой), мы ищем предложения (или абзацы) в которых есть индикаторы тематического события из словаря по определенной тематике. Если таких индикаторов не найдено – текст пропускается.

Для выявления событий, их субъектов и объектов мы создаем специальные шаблоны правил разбора предложений по определенной тематике и словари индикаторов событий по определенной теме, которые заполняются с экспертом с учетом различных лингвистических признаков.

Индикаторы тематического события представлены словарями (униграммами, биграммами и триграммами) и могут одновременно входить сразу в несколько тематик. При этом используется не один большой словарь индикаторов события по определенной теме, а два. Это позволяет строить более сложные шаблоны условий и вывести показатели точности экстрагирования событий на более высокий уровень.

Выделение индикаторов событий происходит с помощью поиска униграмм и n-грамм в предложениях документа из одного из двух словарей присвоенной документу категории.

Использование двойных словарей позволяет присвоить индикатору события специальный аргумент принадлежности к одному из словарей, это позволяет строить более сложные шаблоны условий и вывести показатели точности экстрагирования событий на более высокий уровень.

В дальнейшем – можно автоматизировать наполнение словарей индикаторов событий, например, биграммами и триграммами такого рода как: “был жестоко убит“, “был зверски убит“, “была похищена и убита“ и т.д., автоматически разбирая и сравнивая предложения с предположительными синонимами словаря индикатора описания события, при условии, что TF веса данных слов по всему массиву текстов определенной тематики достаточно велики.

Для выявления возможных фигурантов физ. лиц – создается общий словарь имен русских / украинских и отчеств с различными окончаниями, затем берутся биграммы (триграммы) из выборки текстов, где присутствует одно из имен словаря (или отчеств) и

смотрится слово после него (или до), если это слово начинается с большой буквы, то с большой долей вероятности можно утверждать, что это слово фамилия (или имя) – это и выступает как фигурант1 или фигурант2, при условии, что длина слов не менее 3 символов.

Для выявления возможных фигурантов юр. лиц, берутся все биграммы и триграммы из выборки текстов в которых хотя бы одно из слов начиналось с большой буквы и вторым или третьим словом была указана или форма деятельности (ООО/ ПП / ТОВ/ ЗАО …или слово типа Компания, Компаний, холдинг и т.д.).

Выделение фигурантов событий на данном этапе – это просто выявление их наличия.

Если в предложении найден индикатор события и хотя бы один фигурант – идет дальнейший разбор предложения – нумерация порядка всех слов в предложении и присвоение каждому слову своего порядкового номера.

Окончания фигурантов разбираются и записываются как отдельные аргументы. Они нужны для присвоения статуса фигурантам события. Если в предложении фигурант имеет окончание “ем/ой/ым/ми/им”, то делаем вывод, что фигурант с таким окончанием – субъект. Если нет, то – объект (неизвестно).

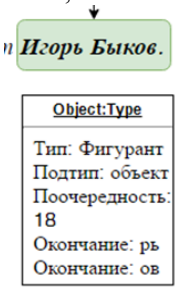


Рисунок 1 – Пример выделения аргументов для одного из фигурантов события

Затем, когда всем сущностям присвоены типы, подтипы, аргументы и порядковые номера в предложении – идет подбор одного из списка шаблонов.

Шаблон экстрактора событий состоит из шаблонов такого вида:

Варианты шаблонов – индикатор тематического события из словаря №1 или словаря №2 (может быть как униграмм, так и n-грамм) и варианты расположения других переменных с ним для

определенной категории. Шаблоны для словаря №1 и №2 могут отличаться.

- Вариант 1: Фигурант1 (Субъект) – Индикатор тематического события – Фигурант2 (Объект) – Место и/или Время
- Вариант2: Место и/или Время – Фигурант1 (Субъект) – Индикатор тематического события – Фигурант2 (Объект)
- Вариант3: Фигурант1 (Объект) – Место и/или Время – Индикатор тематического события – Фигурант2 (Субъект)
- Вариант4: Фигурант1 (Объект) – Индикатор тематического события – Место и/или Время – Фигурант2 (Субъект)

Причем, Фигурант1 или Фигурант2, Место и/или Время могут отсутствовать вовсе. Обязательным является разбор только текстов с присвоенной им категорией (темой) и наличие хотя бы одного Фигуранта и Индикатора тематического события в предложении.

Важно отметить, что модель принимает решение о том, какой именно шаблон экстракции использовать учитывая такие условия (правила) как:

- К какой категории отнесен весь текст
- В каком из двух словарей индикаторов события определенной тематики находится слово индикатор.
- Окончания имен и/или фамилий в предложении. (Если в предложении фигурант имеет окончание “ем/ой/ым/ми/им”, то делаем вывод, что фигурант с таким окончанием – субъект.
- Порядок употребления индикатора события, субъекта и объекта в предложении.

if a ^ b then c
except if d then e
else if f ^ g then h

Данное правило можно интерпретировать как: если a и b истина, то мы принимаем решение c, за исключением случая, когда d не истина. Если d истина (исключение), то принимаем решение e. Если a и b не истина, то мы переходим к другому правилу и принимаем решение h, если f и g истина.

При наличии хорошего обучающего корпуса и грамотного выбора признаков, мы рассчитываем на качество выявления события и его фигурантов на уровне 0.65-0.75 по F1-мере для текстов на русском / украинском языке.

F1-мера:

$$F = 2\frac{Precision \times Recall}{Precision + Recall}$$

Результатом выявления события является:

- Категория (тема) события;
- Субъект (если есть);
- Индикатор события;
- Объект (если есть);
- Остальные аргументы – время, место, дата (если есть).

В качестве дальнейшего улучшения модели выявления новых событий и их фигурантов, можно использовать кросс-текстовое выявление событий, их субъектов и объектов. С помощью выделения документов за определенный промежуток времени с одинаковыми или связанными событиями и принятия решения уже по кластеру предложений с определенным событием. Также, с помощью использования WordNet [6] можно разработать модель взаимосвязи событий с предметами и другими сущностями в информационных потоках. Для увеличения количества выявлений события, можно проводить разбор не предложений, а целых абзацев текста, если они присутствуют в текстовых документах. А для обеспечения большей полноты – использовать фоновые знания [7], добытые ранее с различных баз данных и сайтов типа Wikipedia. Для этого необходимо строить сети слов. Для построения сетей слов можно использовать дисперсионную оценку важности слов [8].

Литература

1. The ACE 2005. Evaluation Plan Evaluation of the Detection and Recognition of ACE Entities, Values, Temporal Expressions, Relations, and Events. 2005.
2. Nate Chambers and Dan Jurafsky. Unsupervised Learning of Narrative Schemas and their Participants. Proceedings of ACL. 2009.
3. Qi Li, Heng Ji, and Liang Huang. Joint event extraction via structured prediction with global features. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 73–82, Sofia, Bulgaria, August. Association for Computational Linguistics. 2013.

4. Lei Sha1, Jing Liu, Chin-Yew Lin , Sujian Li , Baobao Chang, Zhifang Sui. RBPB: Regularization-Based Pattern Balancing Method for Event Extraction. 2016.
5. Thorsten Joachims, *Learning to Classify Text Using Support Vector Machines*. Dissertation, Kluwer, 2002.
6. Josu Goikoetxea, Eneko Agirre, and Aitor Soroa. Single or Multiple? Combining Word Representations Independently Learned from Text and WordNet. 2016.
7. Heng Ji. Relation extraction event extraction. 2014
8. Ortuño M., Carpena P., Bernaola P., Muñoz E., Somoza A.M. Keyword detection in natural languages and DNA // Europhys. Lett, – 57(5). – P. 759-764. 2002.

РАНЖИРОВАНИЕ ПОНЯТИЙ, ИЗВЛЕКАЕМЫХ ИЗ ПОТОКОВ СЕТЕВЫХ НОВОСТЕЙ

А.А. Снарский[1,2], Д.В. Ланде[2,1], Д.И. Зоринец[2]

[1]Национальный технический университет Украины им. И. Сикорского,
[2]Институт проблем регистрации информации НАН Украины

Ранжирование – один из методов упорядочения объектов, как физических, так и информационных. В том случае, когда каждому объекту из совокупности можно приписать некоторое численное значение, задача ранжирования становится формально тривиальной, так как объекты можно ранжировать по величине этого значения. Сложность, однако, заключается в том, что, во-первых, не всегда понятно как определить такое численное значение, а, во-вторых, таких численных значений может быть много и не всегда ясен критерий, по которому нужно выбирать одно из них. Другими словами, наиболее сложной, плохо формализуемой частью задачи ранжирования является выбор критерия, по которому объекту приписывается численные значения (формализация объектов).

В работе, на примере группы ведущих мировых политиков (далее будем называть их персонажи), предложен новый метод ранжирования, позволяющий оценить наиболее «влиятельных» и наиболее «подверженных влиянию» политиков в данный период времени.

Метод состоит из двух этапов. На первом этапе для каждого персонажа определяется число его цитирований в некотором пуле печатных изданий на каждый день, т.е. формируется временной ряд. Для этого нами была использована система InfoStream. Далее каждый член этих временных рядов (соответствует дню) нормируется на полное число цитирований всех персонажей за этот день. Для полученных нормированных числовых рядов вычисляются взаимные временные корреляторы, задающие среднее значение по множеству полученных значений для совместного распределения двух процессов (персонажей). В зависимости от сдвига временного интервала при вычислении коррелятора, полученные корреляторы задают вероятность появления цитирования одного персонажа от того, был ли процитирован ранее другой. Полученные для всех персонажей корреляторы можно представить как набор сложных сетей. Каждая сложная сеть представляет собой набор узлов (персонажей), соединенных направленными связями, вес которых

соответствует численному значению коррелятора. Каждой сети соответствует свой временной сдвиг (запаздывание) при вычислении корреляторов – день, два и т.п. Таким образом получается полный направленный граф. Связь, направленная от одного персонажа к другому определяет прямое влияние (положительное или отрицательное, в зависимости от знака численного значения коррелятора).

На втором этапе для определения полного влияния одного персонажа на другого, учитывается не только непосредственное влияние (связь между персонажами), но и опосредованное, через цепочку. Для этого полученная сложная сеть представляется как когнитивная карта. Эта когнитивная карта может быть исследована различными методами, например, импульсным или фаззи-методом. Нами был использован К-метод [1].

В результате такого анализа вычисляются попарные влияния узлов, что позволяет рассчитать суммарное влияние каждого узла на все остальные. Именно это численное значение и было определено как критерий ранжирования.

Адекватность предложенного подхода была экспериментально проверена на группе ведущих мировых политиков, в частности, получен нормированный ряд цитирований для двух персонажей – Трампа и Клинтон для периода, соответствующего октябрю 2016 г., при этом ранговый коэффициент первого оказался выше. Также предложенным методом был получен ранжированный ряд для всех персонажей.

Литература

1. Snarskii A.A., Zorinets D.I., Lande D.V., Levchenko A.V. K-method of cognitive mapping analysis // E-Preprint ArXiv:1605.08243, 2016.

ВЫЯВЛЕНИЯ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ И АГЕНТОВ ВЛИЯНИЯ НА ДИСКРЕДИТАЦИЮ РУКОВОДСТВА УКРАИНСКОЙ АРМИИ

Шнурко-Табакова Э.В.,

Издательский дом «СофтПресс»

Интернет как отражение и проводник процессов реальной жизни не только порождает новые инструменты коммуникации с гражданами, но и служит средой рождения экспертов, героев, предателей и огромного числа сообществ, объединенных однородными ценностями, идеями, героями и врагами. При этом эти самые эксперты или враги в информационной войне могут выполнять на самом деле совершенно другую функцию – например, быть личными киллерами темы или общественного деятеля. Известны и другие яркие случаи, когда реальные должностные лица, отсутствуя в интернете, фактически не известны гражданам и выпадают из политического процесса. Или, что еще страшнее, за них говорят другие, формируя образы плохих или хороших героев.

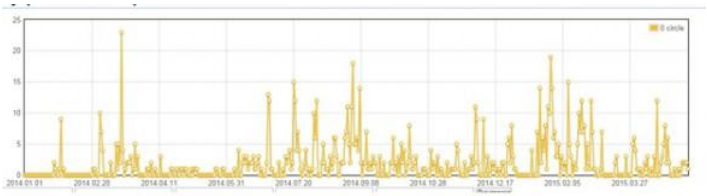
Если набрать в гугле «украинские генералы», то ничего хорошего мировой лидер поиска вам не предложит: только негативные ассоциации запросов и соответствующие им ссылки. Главные предложенные ассоциации будут «Слили, воруют, некомпетентность или предательство». Известно, что результаты поиска базируются на запросах народа. Откуда берется народный мазохизм в отношении руководства украинской армии – мы и попытаемся разобраться.

Лично для автора странности начались в один прекрасный четверг конца мая 2014 года – именно в тот обычный день в «Фейсбуке» появились похожие сообщения про трех действующих генералов украинской армии. Сделаны сообщения были по одинаковому шаблону: официальный портрет, официальная биография и искусно вставленные в текст свидетельства патриотического окружения об ужасах воровства и предательства засвеченных особ. Надо ли говорить, что все патриоты «Фейсбука» тут же подхватили волны негатива и отправились в плавание. Как оказалось – в длительное.

Конечно же, методы анализа интернет-среды показывали искусственность ситуации и полгода назад. Но именно сейчас, когда ситуация застыла и происходят выборы сценария дальнейшего хода войны, есть смысл успокоиться, посмотреть на ситуацию сверху и

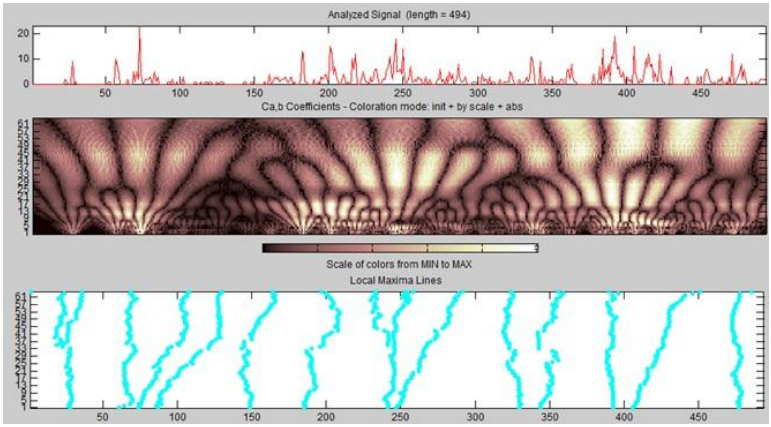
выработать план. Личный план каждому гражданину, политическому/общественному деятелю или чиновнику своего участия в противодействии информационным атакам на Украину.

Действительно ли в мае исполняется год информационной операции по дискредитации руководства украинской армии? Воспользуемся методами анализа контента и выявления информационных операций, ранее описанными в марте 2014 года – «Дело в шляпе». Строим график упоминаний украинских генералов в системе «Инфострим» с 1 января 2014 года по простому запросу «украинские генералы», который уже стал устоявшимся мемом.



Напомним матчасть. Для того чтобы обнаружить в информационном шуме спланированные искусственные информационные операции, применяется обработка данных с целью обнаружения определенных закономерностей. Информоперации, как правило, имеют форму кривой с амплитудными колебаниями определенной формы [1]. Такую форму ближе всего передает так называемый вейвлет Морле $\left(\psi(r) = \exp\left(ik_0 r - \frac{r^2}{2}\right)\right)$, который используется как шаблон в статистических данных аналогичных кривых:

Процесс обработки данных состоит из нескольких стадий: сглаживание данных (нормирование по рабочим и выходным дням, вычеркивание повторных тестов и т. п. – первая диаграмма), применение вейвлет-анализа для идентификации возможных информационных операций (средний рисунок) и определение локальных максимумов (скелетонов) на нижней диаграмме.



Максимумы показывают, на какой день, начиная с 1 января 2014 года, попадают пики информопераций, при этом чем светлее область средней диаграммы – тем более динамика информационного процесса соответствует форме выбранного нами вейвлета, т. е. вероятна информатака на просторах Интернета. Зная даты, мы можем проверить, какие события в реальной жизни происходили в эти дни.

75-й день (26 февраля): первая информоперация после Майдана ознаменовалась материалом «Война впотьмах». Основная тема – отставные украинские генералы координируют действия неофашистов. Т. е. шла подготовка крымских событий: подготовка населения к понятиям производных от фашизма с одной стороны, и запуск мема «украинские генералы» – с другой.

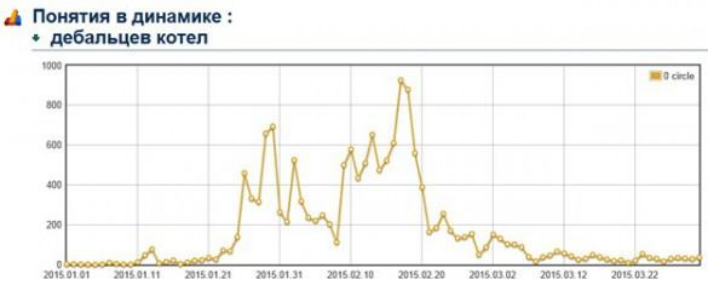
150-й день (1 июля): возобновление наступления на Востоке Украины, прекращение перемирия. Первый материал – «Украинские генералы проверяют солдат на живучесть: военные боятся нападений».

180-й день (1 августа): масса публикаций о прибыльном бизнесе украинских генералов. О чем же говорить в затишье, если не о деньгах? Тема зависти и раздражения вприкуску с воровством – один из лучших способов заронить подозрения даже у самых адекватных личностей. Практически одновременно с гибелью малазийского Боинга широко распространяется версия коварного плана украинских генералов по уничтожению гражданского самолета.

220–250-й день (август – сентябрь 2014): Иловайск. Здесь самое интересное то, с какой завидной синхронностью мочат

украинских генералов украинские эксперты и ресурсы – и все ресурсные солдаты Московии. Причем для гиперболизации описаний Иловайского котла масса материалов запускается с описанием трех котлов, сколько стоят генеральские погоны и т. п. Фактически именно в этот период к мему «украинских генералов» привязывается начальник Генерального Штаба Виктор Муженко.

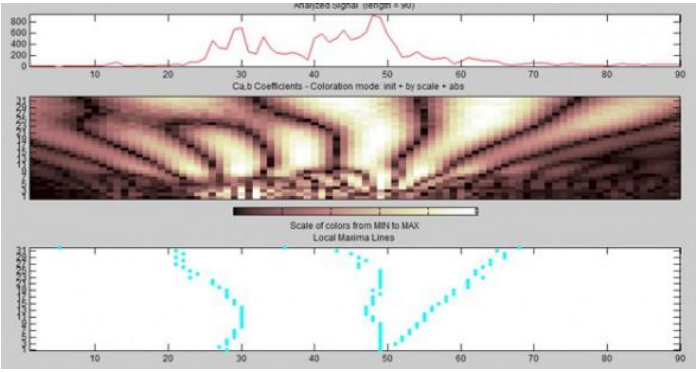
330–350-й день (декабрь 2014): Донецкий аэропорт. Все, что может быть плохого, – от «украинских генералов», которые не могут воевать, договариваться и занимаются контрабандой.



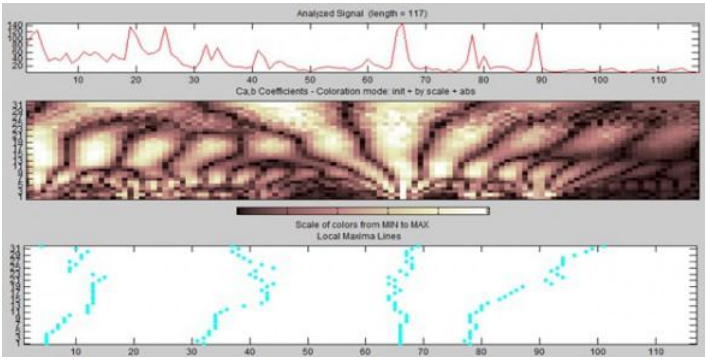
380–420-й день (январь – февраль 2015): Дебальцево. Сопровождающий фон – «Звезды украинского крепостного театра». Говорят, украинские генералы нынче становятся долларовыми миллионерами за пару недель. Прожорливы, что кадавры. Заинтересованы слить фронт, солдат, страну. И, конечно же, опять трагический котел с гибелью тысяч людей. В эти даты особенно интригующе выглядит то, что формулировка «котел Дебальцево» возникла задолго до каких-либо реальных военных операций в этом районе. Ничто не мешает исследовать вопрос отдельно – делаем запрос «Дебальцево Котел».

Даже без особой обработки видно, что проверочный вброс был еще 12 января, затем два пика ярко выраженной информационной операции: первый – подготовка, второй – разворачивание паники, подкрепленной действиями по «окружению».

Вейвлет Морле выявляет два максимума операции. Собственно информационно суть операции выглядит вполне по-шариковски: мы их окружали-окружали, окружали-окружали… Зачем так долго и однообразно? Потому что готовили сильные переговорные позиции в Минске, применяя как информационную обработку, так и локальные военные действия и операции.



Параллельно с информационными операциями по обезличенным «украинским генералам» происходят и спецоперации по конкретным должностным лицам. Так, начальник генерального штаба Виктор Муженко живет в интернете своей виртуальной жизнью – образ формируется со слов экспертов, очевидцев, очень «военных» журналистов. Только в этом году регистрируется четыре максимума неестественных информационного шума.



Кроме обезличенных сайтов и неизвестных авторов, волн перепечаток негативных материалов, можно также проанализировать известных спикеров, говорящих и пишущих о войне. Для сравнения можно сопоставить количество упоминаний в материалах об украинской армии и двух ее генералах.

Спикер	Украинская армия	Виктор Муженко		Руслан Хомчак	
		генерал-полковник		генерал-лейтенант	
	Количество публикаций		Процент в объеме		Процент в объеме
Геращенко	4352	494	11,4	226	5,2
Бирюков	3892	799	20,5	42	1,1
Бутусов	3275	2007	61,3	551	16,8
Касьянов	961	168	17,5	9	0,9
Мачанов	256	119	46,5	5	2,0
Семенченко	5095	1591	31,2	601	11,8
Тимчук	6755	488	7,2	25	0,4

Это как раз тот случай, когда размер имеет значение и в процентах, и в отношении к цифрам коллег. Как по мне, когда есть что-то большее, чем 30 %, то это уже смычка, а не сводки, журналистика или анализ. А если и тональность однобока – то слово «заказ» получает свою красивую визуализацию.

Какие цели реализуются с помощью расшифрованных информационных операций? Кроме очевидных – тотальное неверие граждан Украины в боеспособность нашей армии, единство руководства и рядовых, – формируется равнодушие «ничегонемогусделать» и единственный выход – «третий майдан». Но есть и более глубокие результаты – например, появились эксперты, которые, ссылаясь на недоверие народа и солдат к руководству армии, объясняют, что именно поэтому Украина не получает оружие от США: дескать продадут, потеряют, поломают. Такие информволны также используются политическими партиями на выборах: на уже расшатанном психологически электорате намного легче «продавать» необходимость смены власти и руководства. Можно также устраивать кампании по смене руководства армии в пользу различных влиятельных группировок, в том числе бывалых распильщиков.

Роль каждого сознательного гражданина очень высока и ценна – именно за наши умы и настроение сражается пропагандистская машина агрессора. И это только кажется, что бетонный забор в интернете хорошо виден и осязаем. Технологии разрушающего патриотизма отлично работают с «полезными идиотами» (полезные идиоты – это такие невинные дебилы, которые поддаются манипуляциям на основе эмоциональных информационных воздействий [2]). Подхватывая негатив, сомнительный компромат или отвлекаясь на искусственно раздутое сострадание, мы перестаем

заниматься конструктивными делами, теряем вдохновение, четкие и понятные цели растворяются в тумане войны. И уж по крайне мере – не надо перепечатывать и ссылать на источники, попавшие в рейтинги участников информационных операций.

Топ-50 ресурсов по странам, проявивших себя в операциях по «украинским генералам». Рейтинг – по количеству публикаций

Украина	Россия
Trust.UA	Newsland
Вести.ua	Антимайдан
Комитет.net.ua	Politikus.ru
Народный Корреспондент	Конт
Elise.com.ua	WorldPristav
ОРД-02	Око Планеты
Главком	Военное обозрение
Киевская Правда	Newsli.ru
Кременчуцький Телеграф	Novorus.info
Навигатор	Свободная Пресса
ICTV Факти	Вести.Ru
Цензор.Нет	НТВ
Антифашист	Русская весна
Информационно-аналитическое бюро города Одессы и области	NewsFront
UAinfo	Вопросик.net
Обозреватель	ПравдоРУБ
News-portal.dn.ua	Биржевой Лидер
Винница Реал	Forum.msk
ИнформБюро	Накануне.RU
Politica-UA.com	Красная звезда
	Rusfact.ru
	Российская газета

Топ-50 ресурсов, задействованных в информоперации «Дебальцево.Котел»

Антимайдан
Южный Федеральный
РИА Новости Украина
Novayagazeta-ug.ru
Newsland
Монависта
NewsFront
Pravda News
Империя.by
DNR24.com
Информационно-аналитическое бюро города Одессы и области
ПравдоРУБ
PolitRussia
Русская весна
News-portal.dn.ua
ИнформБюро
Голос Севастополя
Новости@MAIL.RU
Novorus.info
NahNews.com.ua
Око Планеты
РИА Новости
Politikus.ru
NTer.Net
Накануне.RU
Газета.Kg
Военное обозрение
Комитет.net.ua
Информационный портал Донецкой народной республики
WorldPristav
Крымские новости
Контрольный выстрел
Конт
Свежий ветер
ГигаМир.net
Вести.Ru
Антифашист
Rusfact.ru
Коммунистическая партия РФ
ФАН No1
LikeNews

Таким образом, мы в очередной раз убедились, что внутренние и внешние враги вполне сознательно уже полтора года занимаются целенаправленной дискредитацией руководства украинской армии.

Методы статистического анализа указывают на искусственность происходящего и эффективное сочетание наземных

и информационных операций. Эти операции не всегда идут синхронно и про одно и тоже – психологически иногда очень эффективно вбросить негатив на фоне героизма, заранее настроить наблюдателей на поражение и, конечно же, внедрить рефрен «всепропало». Украинцы уже победили, пропаганда русского мира не прошла. Осталось только не подставляться под псевдопатриотические акции, в основе которых лежат генетическое недоверие к власти, комплексы неполноценности и отсутствие механизмов эффективного участия в жизни страны гражданского общества.

Литература

1. Горбулін В.П. , Додонов О.Г. , Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. – К.: Інтертехнологія, 2009. – 164 с.
2. Шнурко-Табакова Э.В. Тыловое, лирическое... или несколько поводов спустить курок // http://blogs.pravda.com.ua/authors/ shnurko-tabakova/ 5441193920bcb/

МЕТОДИКА ПРИМЕНЕНИЯ ИНСТРУМЕНТАРИЯ ЭКСПЕРТНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ИДЕНТИФИКАЦИИ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ

Андрейчук О.В., Качанов П.Т.
Институт проблем регистрации информации НАН Украины

В докладе приведен анализ целесообразности применения инструментария экспертной поддержки принятия решений при идентификации информационных операций. Предложена методика применения инструментария экспертной поддержки принятия решений при идентификации информационных операций. Данная методика базируется на использовании экспертной информации, полученной путём проведения групповых экспертиз с помощью соответствующего инструментария для работы специалистов-экспертов через глобальную сеть. Для получения экспертной информации в полной мере и без искажений применяются средства адаптивного экспертного оценивания. На основании полученной экспертной информации, инженер по знаниям строит базу знаний предметной области, путём применения соответствующего инструментария поддержки принятия решений. Согласно построенной базы знаний, уточняются запросы для анализа динамики соответствующих информационных сюжетов путём применения средств текстовой аналитики. Используя результаты анализа и построенную базу знаний, средствами поддержки принятия решений вычисляется степень достижения цели информационной операции, как сложной системы, компонентами которой являются конкретные информационные мероприятия. Далее, базируясь на проведённых расчётах, лица принимающие решения могут разрабатывать стратегические и тактические меры по противодействию информационной операции, оценивать её эффективность, а также и эффективности отдельных её компонентов. Применение предложенной методики детально продемонстрировано на примере для информационной операции против Национальной академии наук Украины с использованием системы распределённого сбора и обработки экспертной информации "Консенсус-2", комплекса программных средств для экспертного оценивания путём парных сравнений „Уровень", системы поддержки принятия решений "Солон-3" и системы контент-мониторинга InfoStream.

Введение

При сегодняшнем уровне развития информационных технологий сложно переоценить их влияние на жизнь людей. Информационная среда, в которую погружен каждый отдельный человек, социальная группа, население, формирует соответствующее мировоззрение, влияет на поведение и принятие решений. Поэтому вопросы, касающиеся формирования и изменения этой информационной среды, приобретают в наше время чрезвычайную актуальность.

Под информационной операцией (ИО) [1-3] подразумевается комплекс информационных мероприятий (новостные статьи в интернете и газетах, новости по телевизору, комментарии в социальных сетях, форумах и т.п.), нацеленный на изменение общественного мнения об определённом объекте (личность, организация, институт, страна и т.п.). Например, распространив слухи о проблемах в банке, можно спровоцировать его вкладчиков на возврат вкладов, что в свою очередь может привести к его банкротству. В основном – это мероприятия дезинформированного характера. Информационная операция относится к так называемым слабо структурированным предметным областям [4, 5], поскольку ей присущи некоторые характерные для таких областей свойства: уникальность, невозможность формализации цели функционирования и, как следствие, невозможность построить аналитическую модель, динамичность, неполнота описания, наличие человеческого фактора, отсутствие эталонов. Для работы с такими предметными областями применяют экспертные системы поддержки принятия решений (СППР) [6].

В [1-3] показаны методы идентификации ИО, которые базируются на анализе временных рядов, построенных на основе мониторинга тематического информационного потока. Отметим ряд проблемных ситуаций, которые могут возникнуть при идентификации ИО из-за недостатков соответствующих методов и технологий:

1. На фоне достаточно большого общего количества публикаций об объекте ИО количество публикаций (информационных вбросов) об одном отдельном его компоненте может быть весьма незначительной и, как следствие, не будут выявлены соответствующие системные нарушения типичной динамики информационных сюжетов (такие как, например,

выявленные вейвлеты “мексиканская шляпа” и Морле на соответствующей вейвлет-скейлограмме). Некоторые ИО могут быть комплексными и соответствующие информационные вбросы могут быть поэтапными, касаться различных компонентов объекта ИО на разных периодах времени. Если их количество будет размыто на фоне общего количества публикаций про объект ИО ("информационного шума") и соответствующие информационные атаки не будут идентифицированы, то может быть пропущено начало информационной кампании по дискредитации объекта ИО и определенная информационный ущерб его имиджу не буде учтен.

2. Средства контент-мониторинга отрабатывают запросы, состоящие из ключевых слов, в результате чего будут найдены соответствующие публикации. Ключевые слова формулируются, исходя из названия объекта ИО. Но сложный объект ИО может иметь значительное количество компонент с соответствующими названиями, которые не учтены в запросах и, как следствие, не все публикации по тематике будут найдены.

3. Запросы, касающиеся объекта ИО, имеют разную степень важности в соответствии с компонентами ИО, которых они касаются. Отсутствие информации о значении этих степеней важности (то есть их равнозначность) приводит к снижению адекватности модели ИО.

Для предотвращения вышеописанных недостатков предлагается использовать следующую методику применения инструментария экспертной поддержки принятия решений при выявлении ИО.

1 Сущность методики применения инструментария экспертной поддержки принятия решений при идентификации информационных операций

Сущность предлагаемой методики применения инструментария экспертной поддержки принятия решений при идентификации ИО заключаются в следующем:

1. Проводится предварительное исследование объекта ИО, выбираются его целевые параметры (показатели). Далее предполагается, что ранее в ретроспективе уже имели место ИО против объекта и его состояние (соответствующие целевые показатели) от них ухудшалось.

2. Проводится групповая экспертиза по определению и декомпозиции целей информационной операции, а также оценке

степени влияния. Таким образом, объект ИО декомпозируется как сложная слабо структурированная система. Для этого используются средства системы распределенного сбора и обработки экспертной информации (СРСОЭИ). Для получения экспертной информации в полной мере и без искажений используется система экспертного оценивания.

3. Строится соответствующая база знаний (БЗ) средствами СПП на основании результатов проведенной средствами СРСОЭИ групповой экспертизы, а также имеющейся объективной информации.

4. Проводится анализ динамики тематического информационного потока средствами системы контент-мониторинга (СКМ). Дополняется БЗ СППР.

5. Рассчитываются рекомендации средствами СПП на основании построенной БЗ. Для этого вычисляются степени достижения целей ИО в ретроспективе и сопоставляются с соответствующими изменениями состояния объекта ИО. Вычисляются среднее значение степеней достижения целей ИО, при которых происходило ухудшение значений целевых показателей объекта ИО. Таким образом, путем мониторинга состояния объекта ИО за текущий период времени, можно предположить ухудшение значений целевых показателей объекта ИО на основании сравнения вычисленного за текущий период времени значения степеней достижения целей ИО с вышеуказанным средним значением. В случаях наличия достаточной для статистики объема выборки, а также достаточной корреляции между значениями степеней достижения целей ИО и ухудшением значений целевых показателей объекта ИО, можно даже прогнозировать количественное значение ухудшения целевых показателей объекта ИО на текущий период времени.

Преимуществами предложенной методики являются:

1. Большая детализация модели - на фоне большого количества публикаций про объект ИО вообще, изменения динамики количества публикаций, вызванных вбросом про один из компонентов ИО, будут незначительными и, как следствие, не будут выявлены.

2. Увеличивается объем найденных тематических публикаций, так как запросов и ключевых слов будет больше.

3. Взвешенность компонентов ИО позволяет избежать ситуации, когда все компоненты имеют одинаковую важность. Построенная таким образом модель ИО будет более адекватной.

4. Построенная один раз БЗ может использоваться в дальнейшем на протяжении значительного периода времени без необходимости заново проводить экспертизы.

5. Использование инструментария СРСОЭИ позволяет экспертам работать через глобальную сеть, что обеспечивает экономию времени и средств.

Недостатки предложенной методики заключаются в следующем:

1. Применение экспертных технологий требует временных и финансовых усилий на проведение групповой экспертизы. Кроме того, необходимо проводить своевременную актуализацию БЗ для ее повторного использования в будущем.

2. Сложность и, порой, неоднозначность представления некоторых достаточно сложных формулировок компонентов ИО в виде запросов в системе контент-мониторинга.

2 Пример использования методики применения инструментария экспертной поддержки принятия решений при идентификации информационных операций

Продемонстрируем в деталях предложенную в предыдущем разделе методику на примере информационной операции против Национальной академии наук (НАН) Украины. Как известно, НАН Украины сейчас переживает не лучшие времена. В последние годы дела с финансированием все хуже и хуже: уменьшается бюджет НАН Украины и уменьшается доля бюджета НАН Украины в бюджете страны. Это видно из данных о распределении расходов Государственного бюджета Украины, например для 2014-2016 годов [7-9]. Предположим, что это уменьшение финансирования является результатом информационной операции против НАН Украины.

2.1 Групповая экспертная декомпозиция в системе «Консенсус-2»

В качестве СРСОЭИ для групповой экспертной декомпозиции используем систему «Консенсус-2», которая предназначена для проведении оценивания территориально распределенными экспертными группами. Эта система является усовершенствованной версией системы «Консенсус» [10]. В СРСОЭИ «Консенсус-2» реализована технология построения БЗ для слабо структурированных неформализованных предметных областей в

виде иерархии целей. Система «Консенсус-2» состоит из двух автоматизированных рабочих мест: организатора экспертизы (инженера по знаниям) и эксперта.

Групповая экспертиза состоит из ряда стадий, каждая из которых контролируется инженером по знаниям. Он же инициирует переход между стадиями экспертизы. Сначала эксперту предлагается ответить на следующий вопрос: сформируйте перечень существенных факторов, влияющих на достижение цели "Информационная операция против Национальной академии наук Украины". Для этого эксперт вводит новые формулировки факторов или выбирает факторы из списка уже имеющихся объектов БЗ. Далее, когда экспертная группа ввела достаточно формулировок, инженер по знаниям выделяет группы одинаковых по смыслу формулировок среди всех введенных на текущей декомпозиции. На следующей стадии в каждой смысловой группе экспертами выбираются лучшие (по их мнению) формулировки и проводится соответствующее голосование. Затем инженер по знаниям определяет типы воздействий факторов. Если фактор способствует достижению цели, то влияние считается положительным, если препятствует – отрицательным. На финальной стадии декомпозиции отображается соответствующий граф иерархии целей (рис. 1).

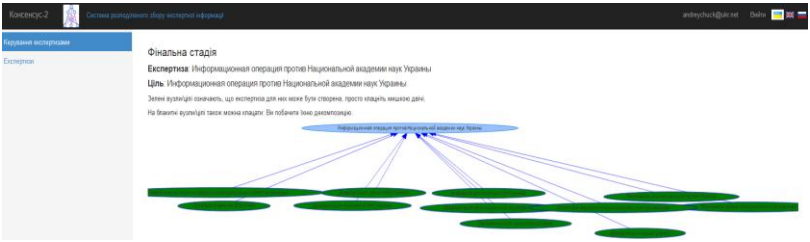


Рисунок 1 – Декомпозиция главной цели в системе «Консенсус-2»

Далее, инженер по знаниям выбирает на графе цель, которую будет раскрывать экспертная группа на следующей декомпозиции. Причем параллельно может раскрываться сразу несколько целей.

Аналогичным образом проходит декомпозиция всех остальных целей иерархии. В результате работы системы формируются соответствующие таблицы базы данных. В дальнейшем СППР формирует БЗ на основе интерпретации этих промежуточных данных.

2.2 Построение базы знаний в СППР «Солон-3»

СППР «Солон-3» [11] предназначена для поддержки решений при планировании крупных комплексных долгосрочных целевых программ, в том числе для построения стратегических планов в различных сферах деятельности. Система дает возможность оценивать и выбирать различные политические, социальные, экономические и другие меры (варианты решений) в зависимости от их влияния на достижение главной и промежуточных целей программы. Система «Солон-3» также позволяет оптимально распределять имеющиеся ресурсы и планировать проведение мероприятий. В ходе оценки учитываются многочисленные сложные взаимосвязи факторов, влияющих на достижение цели программы. СППР "Солон-3" является системой коллективного пользования, база знаний которой формируются многими экспертами-специалистами высшей квалификации в различных областях знаний.

В СППР «Солон-3» используется оригинальный метод, основанный на декомпозиции главной цели программы, построении базы знаний (иерархии целей) и динамическом целевом оценивании альтернатив [6].

На основе интерпретации базы данных, сформированной в результате работы СРСОЭИ «Консенсус-2» в предыдущем пункте, СППР «Солон-3» формирует соответствующую БЗ. Инженер по знаниям имеет возможность редактировать БЗ средствами системы, а именно: вводить новые цели/проекты/связи, редактировать/удалять имеющиеся цели/проекты/связи, вводить частичные коэффициенты влияния в соответствии с результатами групповой экспертизы, а также вводить «объективную (не экспертную)» информацию и другие функции.

2.3 Экспертное оценивание в системе «Уровень»

Комплекс программных средств для экспертного оценивания Комплекс программных средств для экспертного оценивания путем парных сравнений «Уровень» («Рівень») [12] предназначен для проведения экспертизы в системах поддержки принятия решений и позволяет получать знания от экспертов путем предоставления им возможности попарно сравнивать объекты между собой. Он позволяет эксперту указывать наличие предпочтения между объектами с возможностью дальнейшего постепенного уточнения степени этого предпочтения до достижения уровня,

соответствующего реальным знанием эксперта об объекте. В каждом отдельном парном сравнении предусматривается возможность для эксперта выбирать свою, наиболее удобную шкалу с соответствующим количеством делений. То есть, информация от эксперта получается в полной мере и без давления, которое могло бы ее исказить относительно собственных представлений эксперта.

На рис. 2 показано, как происходит экспертное парное сравнение в системе «Уровень», а именно: уточнение степени превосходства влияния в целочисленной вербальной шкале с 9-ю делениями.

Рисунок 2 – Экспертное парное сравнение в системе «Уровень»

2.4 Анализ тематического информационного потока средствами системы контент-мониторинга InfoStream

В результате описанной в предыдущих пунктах групповой экспертизы было получено 15 экспертных формулировок составляющих или компонентов ИО против НАН Украины, а именно:

1) Бюрократия в НАН Украины;
2) Неэффективная кадровая политика НАНУ;
3) Коррупция в НАН Украины;
4) Принижение уровня научных результатов НАН Украины;
5) Отсутствие внедрений научных разработок в производство;
6) Принижение уровня международного сотрудничества;
7) Нецелевое и неэффективное использование недвижимости НАНУ;
8) Нецелевое и неэффективное использование земельных ресурсов НАНУ;

9) Дискредитация президента НАН Украины;

10) Дискредитация управляющего делами НАН Украины;

11) Дискредитация других известных личностей НАН Украины;

12) Противопоставление научных результатов МОН к НАН;

13) Противопоставление научных результатов других академических организаций к НАН Украины;

14) Противопоставление достижений украинских фирм к НАН Украины;

15) Противопоставление научных результатов зарубежных организаций к НАН Украины.

Средствами СКМ InfoStream [13] проводится анализ динамики тематического информационного потока. Для этого, согласно каждого из вышеперечисленных компонентов ИО, на специализированном языке формируются запросы, по которым в дальнейшем и будет происходить вышеупомянутый процесс – анализ динамики публикаций по целевой тематике.

Ниже приведены результаты экспресс-анализа [2] тематического информационного потока, который соответствует объекту ИО – НАН Украины. В результате анализа средствами СКМ InfoStream был получен соответствующий тематический информационный поток из украинского сегмента веб-пространства. Для выявления информационных вбросов с помощью имеющихся аналитических средств анализировалась динамика публикаций по целевой тематике. На рис. 3 представлен некоторый характерный ее фрагмент (за период с 01.07.2015 по 31.12.2015).

Рисунок 3 – Динамика публикаций по целевой тематике

Для выявления степени подобия фрагментов соответствующего временного ряда к диаграмме ИО в разных масштабах используют "вейвлет-анализ". Вейвлет коэффициенты показывают, насколько поведение процесса в определенной точке подобно вейвлету в определенном масштабе. На соответствующей

вейвлет спектрограмме (рис. 4) видно все характерные особенности исходного ряда: масштаб и интенсивность периодических изменений, направление и значение трендов, наличие, расположение и продолжительность локальных особенностей

Рисунок 4 – Вейвлет спектрограмма (вейвлет Морле) информационного потока

Динамику ИО наиболее точно отражают вейвлеты "мексиканская шляпа" и Морле [14]. Поэтому анализируются временные ряды, в соответствии с каждым из 15 компонентов ИО на протяжении 4-х периодов (01.01.2013-31.12.2013, 01.01.2014-31.12.2014, 01.01.2015-31.12.2015 и 01.01.2016-15.12.2016) и идентифицируется наличие вышеупомянутых вейвлетов.

2.5 Расчет рекомендаций средствами СППР «Солон-3»

На основании выявленных в предыдущем пункте информационных вбросов и их параметров (расположение и продолжительность) инженер по знаниям дополняет БЗ СППР «Солон-3». В частности, был идентифицирован вброс по компоненту объекта ИО – "Принижение научных результатов НАН Украины", расположенный 30.11.2015, продолжительностью 14 дней. Соответственно вводится в качестве характеристики проекта "Принижение научных результатов НАН Украины" параметр продолжительность выполнения проекта сроком в 14 дней, а также вводится в качестве характеристики влияния проекта "Принижение уровня научных результатов НАН Украины" на цель "дискредитации научных результатов НАН Украины" параметр задержки в распространении влияния на срок 10 месяцев. Для остальных выявленных информационных вбросов характеристики проектов и воздействий вводятся аналогичным образом.

Таким образом, в частности для периода 01.01.2015–31.12.2015, дополнена БЗ имеет следующую структуру, как показано на рис. 5.

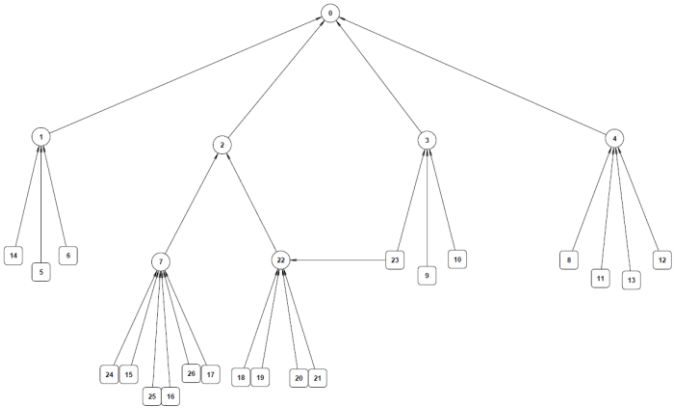


Рисунок 5 – Структура БЗ

Соответственно, Таблица 1 содержит список формулировок всех целей и проектов БЗ.

Следует отметить, что для некоторых компонентов ИО, а именно: "Коррупция в НАН Украины", "Бюрократия в НАН Украины", "Неэффективная кадровая политика НАНУ", "Нецелевое и неэффективное использование земельных ресурсов НАНУ" и "Нецелевое и неэффективное использование недвижимости НАНУ" были обнаружены на протяжении 2015 по 2 информационных вброса, поэтому в БЗ вводились соответствующие проекты по 2 раза. Например: для компонента ИО "Бюрократия в НАН Украины" – проекты "Бюрократия в НАН Украины 1" и "Бюрократия в НАН Украины 2", но каждый из них имеет различные характеристики продолжительности выполнения (9 и 15 дней) и соответствующие влияния имеют разные характеристики задержки в распространении (9 и 11 месяцев).

Далее в СППР «Солон-3» вводятся степени выполнения проектов. Если для некоторых компонентов ИО не было обнаружено никаких информационных вбросов, как в частности для "Противопоставление достижений украинских фирм к НАН Украины" и "Дискредитации деятельности Управления делами НАНУ", то для соответствующих проектов устанавливаются степени выполнения 0%. Для всех остальных проектов – 100%..

Далее получаются результаты расчета рекомендаций, а именно: степень достижения главной цели ИО (рис. 6) и эффективности проектов (относительный вклад в достижение главной цели).

Таблица 1 Список формулировок целей

№	Формулировка цели
0	Информационная операция против НАН Украины
1	Дискредитация научных результатов НАН Украины
2	Дискредитация структуры НАН Украины
3	Дискредитация известных личностей НАН Украины
4	Превозношение научных результатов конкурирующих с НАНУ организаций
5	Отсутствие внедрений научных разработок в производство
6	Принижение уровня международного сотрудничества
7	Дискредитация организационной структуры НАНУ
8	Противопоставление научных результатов МОН к НАН
9	Дискредитация президента НАН Украины
10	Дискредитация других известных личностей НАН Украины
11	Противопоставление научных результатов других академических организаций к НАН Украины
12	Противопоставление научных результатов зарубежных организаций к НАН Украины
13	Противопоставление достижений украинских фирм к НАН Украины
14	Принижение уровня научных результатов НАН Украины
15	Коррупция в НАН Украины 2
16	Бюрократия в НАН Украины 2
17	Неэффективная кадровая политика НАНУ 2
18	Нецелевое и неэффективное использование недвижимости НАНУ 1
19	Нецелевое и неэффективное использование недвижимости НАНУ 2
20	Нецелевое и неэффективное использование земельных ресурсов НАНУ 1
21	Нецелевое и неэффективное использование земельных ресурсов НАНУ 2
22	Дискредитация деятельности Управления делами НАНУ
23	Дискредитация управляющего делами НАН Украины
24	Коррупция в НАН Украины 1
25	Бюрократия в НАН Украины 1
26	Неэффективная кадровая политика НАНУ 1

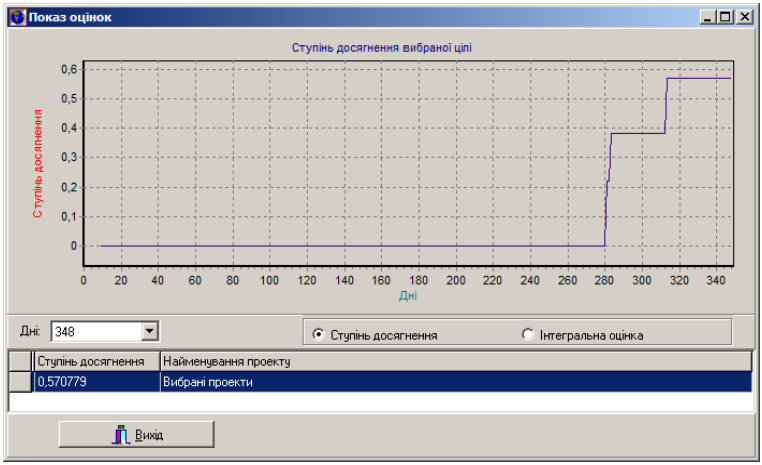


Рисунок 6 – Степень достижения выбранной цели
в СППР «Солон-3»

За периоды 01.01.2013-31.12.2013, 01.01.2014-31.12.2014, 01.01.2015-31.12.2015 и 01.01.2016-15.12.2016 степени достижения главной цели имеют значения: 0.380492, 0.404188, 0.570779 и 0.438703 соответственно.

Для ретроспективы среднее значение степени достижения главной цели равно: (0.380492 + 0.404188 + 0.570779) / 3.0 ≈ 0.45182.

Итак, поскольку среднее для ретроспективы и текущее значение степеней достижения главной цели ИО достаточно близки (отличаются не более чем на 3%), то можно сделать вывод, что ИО за текущий период весьма вероятно может вызвать ухудшение значений целевых показателей объекта.

Выводы

1. Показана целесообразность применения инструментария экспертной поддержки принятия решений в процессе идентификации информационных операций.
2. Предложена методика применения инструментария экспертной поддержки принятия решений при выявлении информационных операций, позволяющая на основании анализа ретроспективы прогнозировать изменение значений целевых показателей объекта на текущий период.

3. Предложенная методика продемонстрирована на примере информационной операции против Национальной академии наук Украины.

Исследование проведено в рамках проекта Ф73/23558 "Разработка методов и средств поддержки принятия решений при выявлении информационных операций". Проект является победителем конкурса Ф73 на грантовую поддержку научно-исследовательских проектов Государственного фонда фундаментальных исследований Украины и Белорусского республиканского фонда фундаментальных исследований.

Литература

1. Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія – К., Інтертехнологія, 2009 – 164 с.
2. Ландэ Д.В., Додонов В.А., Коваленко Т.В. Информационные операции в компьютерных сетях: моделирование, выявление, анализ // МОДЕЛИРОВАНИЕ-2016: материалы пятой Международной конференции МОДЕЛИРОВАНИЕ-2016, Киев, 25-27 мая 2016 г. / ИПМЭ НАН Украины, 2016. – С. 198-201.
3. Додонов А.Г., Ландэ Д.В., Коваленко Т.В. Модели предметных областей в системах поддержки принятия решений на основе мониторинга информационного пространства // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2016): материалы VI междунар. науч.-техн. конф. (Минск 18-20 февраля 2016 года) / – Минск: БГУИР, 2016. – С. 171-176.
4. Таран Т.А. , Зубов Д.А. Искусственный интеллект. Теория и приложения // Восточноукр. нац. ун-т им. Владимира Даля. – Луганск: ВНУ им. В.Даля, 2006. — 239 с.
5. Глибовець М.М., Олецікий О.В. Штучний інтелект. – К.: Видавничий дім «КМ Академія», 2002 – 366 с.
6. Тоценко В. Г. Методы и системы поддержки принятия решений. Алгоритмический аспект. – К.: Наукова думка, 2002. – 382 с.

7. Закон України Про Державний бюджет України на 2014 рік [Електрон. ресурс]. Режим доступу: http://zakon4.rada.gov.ua/laws/show/719-18.
8. Закон України Про Державний бюджет України на 2015 рік [Електрон. ресурс]. Режим доступу: http://zakon4.rada.gov.ua/laws/80-19.
9. Закон України Про Державний бюджет України на 2016 рік [Електрон. ресурс]. Режим доступу: http://zakon4.rada.gov.ua/laws/show/928-19.
10. Цыганок В.В., Качанов П.Т., Андрійчук О.В., Каденко С.В. Свідоцтво про реєстрацію авторського права на твір № 45894 Державної служби інтелектуальної власності України. Комп'ютерна програма “Система розподіленого збору та обробки експертної інформації для систем підтримки прийняття рішень - «Консенсус»” від 03.10.2012.
11. Тоценко В.Г., Качанов П.Т., Циганок В.В. Свідоцтво про державну реєстрацію авторського права на твір №8669. Міністерство освіти і науки України державний департамент інтелектуальної власності. Комп’ютерна програма "Система підтримки прийняття рішень СОЛОН-3" (СППР СОЛОН-3) від 31.10.2003.
12. Циганок В.В., Андрійчук О.В., Качанов П.Т., Каденко С.В. Свідоцтво про реєстрацію авторського права на твір № 44521 Державної служби інтелектуальної власності України. Комп'ютерна програма “Комплекс програмних засобів для експертного оцінювання шляхом парних порівнянь «Рівень»” від 03.07.2012.
13. Григорьев А.Н., Ландэ Д.В., Бороденков С.А. и др. InfoStream. Мониторинг новостей из Интернет: технология, система, сервис: Научно-методическое пособие – К.: Старт-98, 2007. – 40 с.
14. Додонов А.Г., Ландэ Д.В., Бойченко А.В. Сценарный подход при исследовании динамики информационных потоков в сети Интернет // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2015): материалы V междунар. науч.-техн. конф. (Минск 19-21 февраля 2015 года) / – Минск: БГУИР, 2015. – С. 225-230.

ВСЕСВІТНІ ІНТЕРНЕТ-ПРОВАЙДЕРИ В УКРАЇНСЬКІЙ МЕРЕЖІ ОБМІНУ ТРАФІКОМ: ВИКЛИКИ ТА МОЖЛИВОСТІ

В.Ю. Зубок,
Інститут спеціального зв'язку та захисту інформації КПІ ім. І.Сікорського,
ТОВ "Інформаційний центр "Електронні вісті"

Перші мережі обміну трафіком — «network access points» — з'явилися в ARPANET ще до того, як термін Інтернет набув популярності. Ідея їхнього заснування — додержуватися чітко визначених процедур підключення та правил взаємодії між мережами учасників. Сьогодні в Європі функціонує багато мереж обміну трафіком. Вони мають різну кількість учасників та обсяги трафіка, різні про-цедури підключення, а головне — різну політику маршрутизації та взаємодії між учасниками. Але кожна з них має свій вплив на топологію зв'язків між автономними системами в Інтернеті.

У топології сучасного Інтернету мережі обміну трафіком чи біржі трафіка (Internet Exchange Points, IXPs) відіграють дуже важливу роль. Однією з «класичних» мереж обміну трафіком є Українська мережа обміну трафіком (UA-IX), заснована в 2001 році. Наприкінці 2000 років вона перестала бути монопольною мережею обміну трафіком, коли з'явилась мережа DTEL-IX, а згодом – мережа GigaNet, огляд яких було опубліковано в [1].

Станом на квітень 2013 р. в UA-IX приймали участь 132 прямі учасники, що безпосередньо підключені до цієї мережі. Вони обмінювались через UA-IX анонсували маршрути до майже 7000 мереж (чи «мережевих префіксів»), які походять від 1984 вузлів — автономних систем, що з'єднані 2240 безпосередніми зв'язками.

Станом на жовтень 2016 року параметри UA-IX суттєво змінились. Так, кількість учасників зросла до 197, з'явились деякі досить великі учасники, такі як Google, Akamai, Яндекс та Mail.ru. В табл.1 наведені характеристики моделі UA-IX, побудованої за даними таблиць маршрутизації. Показане порівняння характеристик 2013 та 2016 року (до включення Hurricane Electric). Тепер мережа менш схожа на «тісний світ» [2].

В UA-IX базовою і обов'язковою політикою маршрутизації є так званий «відкритий пірінг», тобто обов'язкова передача всіх маршрутів на центральний вузол і отримання всіх маршрутів інших учасників.

Таблиця 1. Характеристики моделі UA-IX в 2013 та 2016 роках

№	Назва параметра	2013	2016
1	Кількість прямих учасників	132	172
2	Кількість префіксів	6867	17926
3	Кількість AS, що є джерелами анонсів	1984	3469
4	Загальна кількість AS, що зустрічаються в шляхах	n/a	3535
5	Діаметр мережі	9	9
6	Кількість тупикових вузлів	1677	2689
7	Кількість транзитних вузлів	307	845
8	Середній коротший шлях	4,05	5,25
9	Глобальна ефективність	0,264	0,207
10	Транзитивність	0,039	0,018

В UA-IX базовою і обов'язковою політикою маршрутизації є так званий «відкритий пірінг», тобто обов'язкова передача всіх маршрутів на центральний вузол і отримання всіх маршрутів інших учасників.

У листопаді 2016 року до мережі UA-IX приєднався міжнародний оператор Hurricane Electric, ідентифікатор AS 6939. Цей оператор за багатьма даними (довжина власних мереж, кількість підключених мереж, обсяги трафіку) належить до провайдерів першого рівня (Tier 1 providers). За опублікованими відкритими даними проекту Center for Applied Internet Data Analysis (CAIDA), він безпосередньо взаємодіє з приблизно 5000 автономними системами [3]. Інші учасники UA-IX з напруженням очікували початку взаємодії з Hurricane Electric. Повна кількість префіксів, якими оперує цей провайдер, достамено невідома, бо в різних мережах обміну трафіком він анонсує різні множини префіксів.

І ось підключення відбулось. Hurricane Electric анонсував в мережу UA-IX 39540 префіксів, збільшивши їхню загальну кількість до понад 57000, тобто більш ніж втричі. За даними таблиці маршрутизації побудовано модель сегменту мережі, що приєдналась до UA-IX. Її характеристики наведені в табл.2.

Отже, цей сегмент подібний з характеристиками до UA-IX, але дещо щільніший, що відображується більшою транзитивністю (середній показник коефіцієнта кластеризації по мережі) та меншим середнім шляхом. Співвідношення транзитних та тупікових AS є приблизно однаковим. На рис. 1а та 1б показана різниця, як

виглядало ядро (група вузлів з найбільшою кількістю зв'язків) мережі UA-IX до приєднання Hurricane Electric, та ядро мережі Hurricane Electric, що було побудовано за даними таблиці маршрутизації.

Таблиця 2. Характеристики сегменту Інтернет, що приєднався до UA-IX після підключення Hurricane Electric.

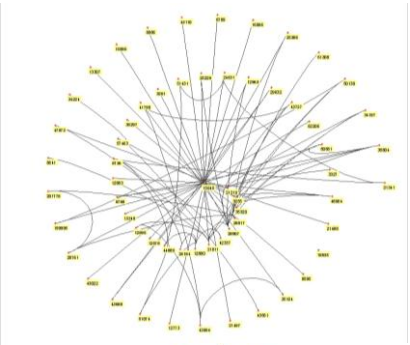
№	Назва параметра	Значення
1	Кількість прямих учасників, підключених до H.E.	1028
2	Кількість префіксів	39540
3	Кількість AS, що є джерелами анонсів	5172
4	Загальна кількість AS, що зустрічаються в шляхах	5255
5	Діаметр мережі	9
6	Кількість тупикових вузлів	4213
7	Кількість транзитних вузлів	1043
8	Середній коротший шлях	4,17
9	Глобальна ефективність	0,259
10	Транзитивність	0,021

Отже, мережа UA-IX відтепер наче охоплює втричі більшу кількість вузлів. Протягом тижня велись спостереження за навантаженням трафіком підключення Hurricane Electric в UA-IX. Було з'ясовано, що середньодобовий вхідний та вихідний трафік не перевищують 3 Гбіт/с. Цей обсяг є дуже незначним порівняно з відомими українськими Інтернет-провайдерами. Наприклад, Датагруп (AS 21219) надсилає всього 928 префіксів, що походять від 344 автономних систем, продукує в декілька разів більше трафіку (рис.2).

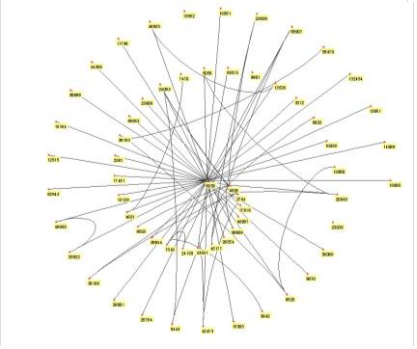
Це наштовхує на необхідність аналізу структури сегменту мережі, що підключений за Hurricane Electric, а також прогнозам використання цього підключення для певних категорій Інтернет-споживачів.

Проаналізуємо склад ядра мережі Hurriacane Electric. Він наведений в табл.3. Перша п'ятірка за кількістю транзитних префіксів є також першою п'ятіркою за кількістю зв'язків з іншими автономними системами. Отже, це досить великі регіональні хаби. Але вони обслуговують специфічні регіони: Північну Америку, Південний Схід, Австралію та Нову Зеландію.

а)



б)



в)

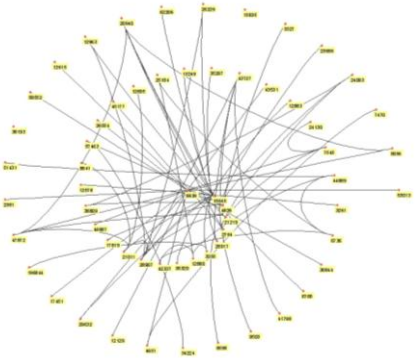
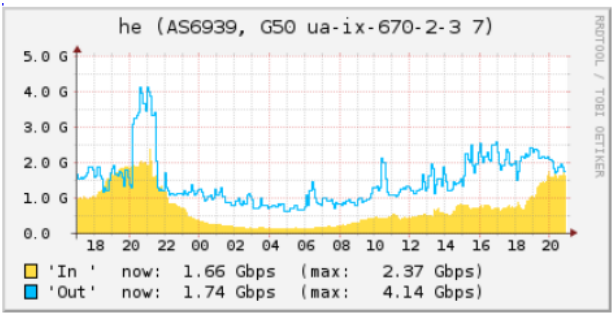
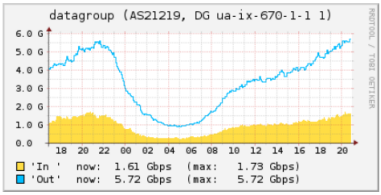


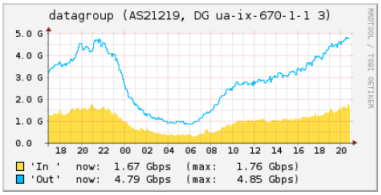
Рисунок 1 – Ядро ядро мережі Hurricane Electric (а), мережі UA-IX до приєднання Hurricane Electric (б), мережі UA-IX після включення Hurricane Electric (в)

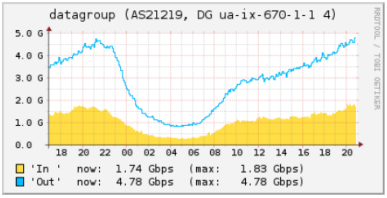
а)



б)







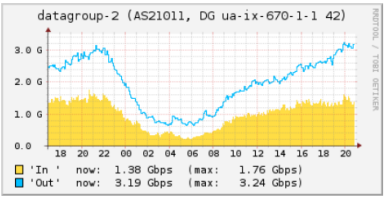


Рисунок 2 – Порівняння денних графіків трафіку від Hurricane Electric (а) та Датагруп (б).

Вони, крім взаємодії з Hurricane Electric, є учасниками інших мереж обміну трафіком, зокрема європейських. Тому, попри анонси їхніх префіксів в UA-IX, коротші маршрути досить часто пролягають не через UA-IX, а через DE-CIX та AMS-IX.

Нарешті, можна продемонструвати як змінилися характеристики мережі UA-IX після підключення настільки великого, але дуже специфічного учасника (табл. 4).

Таблиця 3. Опис основних мереж, що анонсуються в Hurricane Electric.

№	Номер AS	Назва, країна та основний регіон обслуговування	Кількість зв'язків	Кількість префіксів
1	6939	Hurricane Electric (Північна Америка, Південно-східна Азія, Австралія)	1029	39540 транзит, 147 власні
2	46887	Lightower (Північна Америка)	407	1126 транзит, 211 власні
3	4826	Vocus Connect Intl (Австралія)	273	4305 транзит, 102 власні
4	2764	AAPT Limited (Австралія, Нова Зеландія)	174	1939 транзит, 351 власні
5	26554	US Signal Company (Північна Америка)	84	414 транзит, 38 власні
6	7470	TrueInternet (Таіланд, Півд-Сх. Азія)	75	667 транзит, 299 власні

Таблиця 4. Характеристики мережі UA-IX після підключення Hurricane Electric.

№	Назва параметра	До	Після
1	Кількість прямих учасників	172	173
2	Кількість префіксів	17926	57466
3	Кількість AS, що є джерелами анонсів	3469	8641
4	Загальна кількість AS, що зустрічаються в шляхах	3535	8790
5	Діаметр мережі	9	9
6	Кількість тупикових вузлів	2689	6837
7	Кількість транзитних вузлів	845	1953
8	Середній коротший шлях	5,25	5,05
9	Глобальна ефективність	0,207	0,21
10	Транзитивність	0,018	0,02

Висновки

Попри значні топологічні зміни, спричинені в UA-IX приєднанням великого міжнародного Інтернет-провайдера, суттєвого перерозподілу трафіку на нові напрямки не спостерігається через географічну специфіку мереж, які приймають участь в обміні трафіком.

Дослідження моделей сегментів Інтернету, побудованих на основі глобальних таблиць маршрутизації, дозволяє спостерігати зміни топології, пояснювати та прогнозувати їхній вплив на обсяги і напрямки трафіку, і на цій базі розробляти рекомендації з побудови глобальної міжмережевої взаємодії.

Література

1. Зубок В. Ю. Порівняння топології звязків в нових мережах обміну Інтернет-трафіком України / В. Ю. Зубок // Информационные технологии и безопасность: оценка состояния : материалы междунар. науч. конф. ИТБ-2013. – К.: ИПРИ НАН Украины, 2013. – Вып. 13. – С. 70-77.
2. Ландэ Д.В. Параметры украинского сегмента Интернет как сложной сети / Д. В. Ландэ, В. Ю. Зубок, В. Н. Фурашев // Открытые информационные и компьютерные технологии : Сб. науч. трудов. – 2008. – № 40. – С.235-242.
3. AS Rank: AS Relationship Table [Електронний ресурс]. – Режим доступу : URL : http://as-rank.caida.org/?mode0=as-info&mode1=as-table&as=6939 . – Назва з екрану.

СТАТИСТИЧНІ ВЛАСТИВОСТІ МАТЕМАТИЧНОЇ МОДЕЛІ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ

А.М. Грайворонська,
Інститут проблем реєстрації інформації НАН України

Моделювання розповсюдження інформації в соціальних мережах дозволяє дослідити відповідні інформаційні процеси, проаналізувати механізми передачі інформації, а також встановити статистичні закономірності процесу розповсюдження інформації, які можуть бути використані при детектуванні аномалій в динаміці всього інформаційного потоку або життєвого циклу окремого повідомлення при інформаційних операціях [1].

Мультиагентна модель розповсюдження інформації

Документ, опублікований в інформаційному просторі, може викликати різні види реакції: позитивні та негативні коментарі, копіювання, посилання. Отже, документ у інформаційному просторі можна розглядати як агента з певними властивостями та правилами поведінки [3].

Будемо розглядати мультиагентну модель [2], в якій агенту відповідає повідомлення, і еволюція агента буде пов'язана з подіями, які з ним відбуваються. В якості основної характеристики агента введемо енергію (E), яка відображає актуальність повідомлення.

Формалізуємо правила еволюції агента в моделі. Агент з'являється з початковою енергією E_0 і з кожним дискретним відліком часу його енергія зменшується на 1. Будемо розглядати події, які є типовими для соціальних мереж: like (вираження позитивної реакції), dislike (вираження негативної реакції), repost (копіювання та поширення). Ці події впливають на енергію агента наступним чином: like підвищує енергію на 1, dislike зменшує на 1, repost підвищує на 2.

Ймовірності того, що з повідомленням з енергією E відбулася певна подія, в такий спосіб:

$$p_{like}^{(E)} = p_{l_0}\varphi(E),\ p_{dislike}^{(E)} = p_{d_0}\varphi(E),\ p_{repost}^{(E)} = p_{r_0}\varphi(E),$$

де p_{l_0}, p_{d_0}, p_{r_0} – параметри моделі, а φ – це деяка монотонно неспадна функція від поточної енергії агента зі значеннями в [0, 1]. При падінні енергії агента до 0, повідомлення «вмирає» і більше не розглядається.

Моделювання динаміки всього тематичного інформаційного потоку починається з одного агента. Поява нового агента можлива двома способами: створення копії існуючого агента (repost) або поява нового агента (випадкова подія з ймовірністю p_s).

Статистичні властивості життєвого циклу агента

Агент з'являється з початковим значенням енергії E_0 і далі його енергія змінюється в залежності від подій, які з ним відбуваються. Позначимо ε_t значення енергії агента в момент часу t. Тоді значення енергії в наступний момент часу можна записати в такий спосіб

$$\varepsilon_{t+1} = \varepsilon_t + \delta_t,$$

де δ_t є випадковою величиною зі значеннями в {-1, 0, 1, 2}. Отже стохастична послідовність значень енергії агента є марківським ланцюгом, з графом, який зображений на рис. 1.

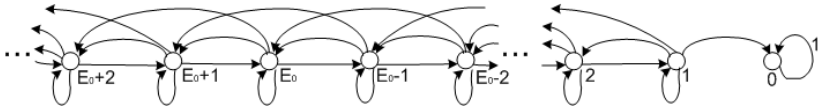


Рисунок 1 – Граф марківського ланцюга, який відповідає зміні енергії агента

Дослідження властивостей випадкового блукання енергії дозволяє зробити певні висновки про вплив параметрів моделі на результати моделювання. Наприклад, нерівність $p_{r_0} > \frac{1-p_{l_0}}{2+p_{l_0}}$ є достатньою умовою для того, щоб час випадкового блукання енергії був обмеженим. На рис. 2а показана область параметрів, яка задовольняє цій нерівності.

Важливою характеристикою опублікованого інформаційного повідомлення є час, протягом якого воно залишається актуальним. В моделі це час життя агента. На рис. 2б показано розподіл часу життя агентів, отриманий в результаті моделювання.

При аналізі життєвого циклу повідомлення важливе значення мають кількості лайків та репостів [4, 6]. Були розглянуті розподіли кількості лайків та репостів, отримані в результаті моделювання. На рис. 3 показана щільність ймовірнісного розподілу кількості репостів.

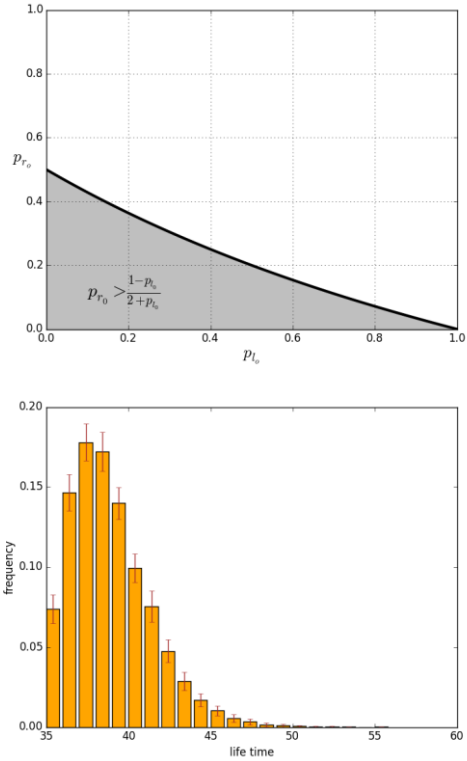


Рисунок 2. а – Область параметрів моделі, при яких час життя агента буде обмеженим з ймовірністю 1; б – Розподіл часу життя агентів

Розподіл, отриманий в результаті моделювання, має чіткий передній фронт, і з високою точністю апроксимуються розподілом Вейбула. Отримані результати моделювання порівнювались з результатами дослідження життєвого циклу повідомлень новин в мережі мікроблогів Twitter, які приведені в [5]. Коефіцієнт форми розподілу Вейбулла в обох випадках близький до 1,9.

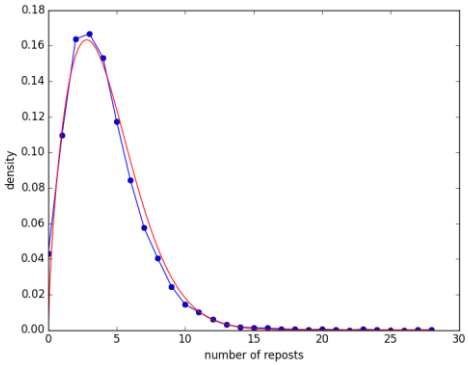


Рисунок 3 – Щільність розподілу кількості репостів для агента

Література

1. Додонов А. Г., Ландэ Д. В., Прищепа В. В., Путятин В. Г. Конкурентная разведка в компьютерных сетях. — К.: ИПРИ НАН Украины, 2013. — 248 с. — с. 34–38.
2. Epstein J. M. Remarks on the foundations of agent-based generative social science / Joshua M. Epstein // Handbook on Computational Economics, Volume II, K. Judd and L. Tesfatsion, eds. North Holland Press, 2005.
3. Додонов А. Г., Ландэ Д. В., Додонов В. А. Распознавание информационных операций: мультиагентный подход . — Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2016): материалы VI междунар. науч.-техн. конф. (Минск 18-20 февраля 2016 го- да). – Минск: БГУИР, 2016. С. 253-256.
4. Dashun Wang, Zhen Wen, Hanghang Tong Information spreading in context. — Proceedings of the 20th international conference on World wide web . — 2011. — pp. 735-744.
5. Ландэ Д.В Мультиагентная модель распространения информации в социальной сети / Ландэ Д.В., Грайворонская А.Н., Березин Б.А. // Реєстрація, зберігання і обробка даних, 2016. - Т. 18. - N 1. - С. 70-77
6. C.TEDAS: A Twitter-based Event Detection and Analysis System / Li R., Lei K.H., Khadiwala R., Chang K.C. // Data Engineering (ICDE), 2012 IEEE 28th International Conference, 2012. — P. 1273–1276.

МЕТОД МАРКОВСЬКИХ ЛАНЦЮГІВ У МЕРЕЖЕВИХ МОДЕЛЯХ ПРОЕКТІВ ДЛЯ ПРОГНОЗУВАННЯ ТА РИЗИК-АНАЛІЗУ

А. І. Кузьмичов,
Інститут проблем реєстрації інформації НАН України

Методологія проектного менеджменту (ПМ) базується на засобах комп'ютерного моделювання, ставши основою проектно-орієнтованого та проектно-керованого організаційного проектування, планування, управління та супроводження діяльності у будь-якій сфері. За визначеними міжнародними стандартами (PMBOK) будь-яка продукція (товари, послуги) розробляється й розповсюджується із врахуванням часових, ресурсних, вартісних обмежень для прийняття рішень в умовах визначеності, невизначеності і ризику. ПМ увійшов у склад фундаментальних аналітичних засобів організаційного управління, що застосовуються для прийняття рішень.

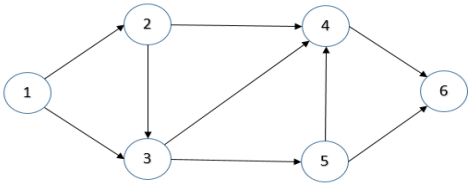
Мережева модель проекту – направлений граф без циклів і контурів, це: дугова (робота-дуга, подія-вузол) або вузлова (робота-вузол, зв'язок-дуга) мережа, де принципово не передбачений реверс до попередніх робіт/подій. Однак практика свідчить, що в умовах невизначеності й ризиків завжди є ситуації, за якими на певній стадії вимушено необхідно повернутися до попередніх етапів, щоби зробити певні зміни. Саме ця проблема змушує шукати відповідні засоби, не порушуючи принципів побудови мережевих моделей проектів певних типів і властивостей.

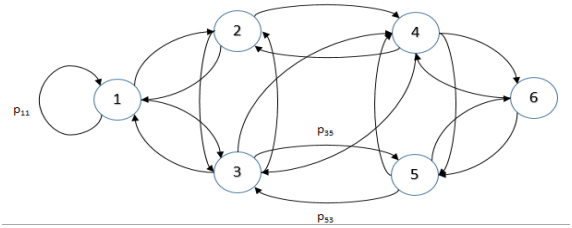
Приклад. Система – проектна направлена мережа та її модель – марковський ланцюг із реверсом між 6 станами (із заданими ймовірностями). За n кроків перебування системи у різних станах визначаються фінальні ймовірності існування кожного стану, в прикладі – стану 6, завершення проекту (на рис. показано не усі дуги реверсу «сам до себе»).

Проектна мережа як ланцюг Маркова:
є система станів, $S = (s_1, s_2, ..., s_6)$ зі зв'язками між ними.

Динамічний процес починається в стані s_1 і розвивається послідовним переходом із одного стану в інший згідно мережі, покроково й із поверненням. Якщо ланцюг в момент t в стані s_i, тоді

він на наступному кроці переходить в стан s_j з ймовірністю p_{ij} і ця ймовірність не залежить від того, у якому стані ланцюг перебував до цього.





Ймовірності p_{ij} – це ймовірності переходу. Процес на певному кроці може залишитися в тому ж стані, де він зараз знаходиться, із ймовірністю p_{ii}. Початковий розподіл ймовірностей на S визначається квадратною матрицею переходів. Через n кроків визначаться фінальні ймовірності усіх станів.

Марковський аналіз: прогноз рівня ризику. Експертний метод

	Очікувані стани							
Поч. стани	*Допустимий*	*Мінімальний*	*Підвищений*	*Критичний*	*Катастрофічний*	*Втрати, до*	*Сума*	*P*
Допустимий	3	5	2	0	0	0	10	0,20
Мінімальний	4	5	2	1	0	25%	12	0,24
Підвищений	1	2	5	2	1	50%	11	0,22
Критичний	0	1	4	3	2	75%	10	0,20
Катастрофічний	0	1	3	2	1	100% і більше	7	0,14
						Експертів =	**50**	
Прогноз ризиків. Тенденція								
Початкова йм.	16,0%	28,0%	32,0%	16,0%	8,0%			
	↑	↑	↓	↓				
Фінальна йм.	18,1%	29,4%	31,4%	14,4%	6,7%			

Оцінювання ризику: формування матриці експертних оцінок; обчислення матриці ймовірностей переходу; формування покрокового вектору рівня ризиків; побудова графіка (тренд).

Результат: прогноз ризиків на основі порівняльного аналізу початкових й фінальних ймовірностей станів.

МЕТОД ФОРМИРОВАНИЯ ПРОЕКТНЫХ ТРЕБОВАНИЙ К СИСТЕМЕ УПРАВЛЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ

В. Мохор[1], А. Богданов[2], А. Бакалинский[2], В. Цуркан[2],
[1]ИПМЭ им. Г.Е. Пухова НАН Украины,
[2]ИСЗЗИ КПИ им. И. Сикорского

Для современных организаций требованием времени является построение и использование систем управления информационной безопасностью. Это обусловлено такими аспектами как исключение неприемлемых рисков, эффективное использование имеющихся средств, повышение осознанности и управляемости процессов обеспечения информационной безопасности. Построение и использование систем управления информационной безопасностью рассматривается на основе риск-ориентированного подхода. Как следствие, за основу берется двухкомпонентная модель риска, которая представляется на плоскости. Благодаря этому определяется вероятностный критерий и его значение, задаваемое в качестве проектного требования при построении систем управления информационной безопасностью в виде «карты риска». Она позволяет «владельцам риска» задавать приемлемые уровни рисков и разделять их на приемлемые и неприемлемые. Однако, «карты рисков» оперируют единичными проявлениями событий и не учитывают их возможного повторного (многократного) проявления. Из этого, делается вывод о неконструктивности проектного требования к системе управления информационной безопасностью, основанного на концепте «обеспечить уровень риска не выше». Поэтому корректное проектное требование формулируется в контексте обеспечения системой управления информационной безопасностью обработки потока рисковых событий с уровнями риска и заданной вероятностью появления таких событий. То есть показывается возможность оценивания вероятности появления события с рисками для заданного уровня приемлемого риска. Или по заданному уровню приемлемого риска оценивается вероятность появления событий с рисками. Решение данной задачи осуществляется путем использования понятия и

методов геометрической вероятности. Применение геометрического подхода к оцениванию вероятности попадания произвольных значений нормированного риска в зону приемлемого риска, дало возможность получить точную количественную оценку этой вероятности. Благодаря такому подходу субъективный показатель риск-аппетита «владельца риска», отображаемый в виде приемлемого уровня риска, трансформируется в формализованный вероятностный критерий, на основе которого можно сформулировать проверяемые проектные требования к созданию систем управления информационной безопасностью.

Постановка проблемы

Требованием времени для современных организаций является построение и использование систем управления информационной безопасностью, а особенно для тех, функционирование которых зависит от стабильной работы информационных технологий или иной критической инфраструктуры [1]. Это связано с тем, что построение и использование обозначенных систем обусловлено такими аспектами как исключение неприемлемых рисков, оптимизация затрат на обеспечение информационной безопасности за счет более эффективного использования имеющихся средств, повышение осознанности и управляемости процессов обеспечения информационной безопасности [1, 2].

При построении систем управления информационной безопасностью руководствуются требованиями международного стандарта ISO/IEC 27001:2013 «Информационные технологии. – Методы обеспечения безопасности. – Системы управления информационной безопасностью. – Требования» [3]. Этот стандарт предопределяет целесообразность использования риск-ориентированного подхода к управлению информационной безопасности [4-6]. С целью конкретизации требований по управлению рисками в рамках группы стандартов серии ISO/IEC 27k принят международный стандарт ISO/IEC 27005:2011 «Информационные технологии. – Методы и средства обеспечения безопасности. – Управление риском информационной безопасности» [6]. В нем, в частности, предопределено, что «риски должны быть идентифицированы, количественно определены или качественно описаны и расставлены в соответствии с приоритетами согласно критериям оценивания риска и уместным для организации целям». Поэтому для формирования корректных и конструктивных

требований к построению систем управления информационной безопасность важным является приведенное в этом стандарте определение риска: «Риск представляет собой комбинацию последствий, вытекающих из нежелательного события, и вероятности возникновения события». В частности, если такая комбинация принимает мультипликативную форму, то соотношение для вычисления уровня риска может быть записано в следующем виде:

$$R = H \cdot p, \qquad (1)$$

где R – уровень (величина) риска, H – оценка величины последствий (ущерба), являющихся следствием нежелательного события, которые (речь идет о последствиях) в случае событий информационной безопасности принимают форму ущерба, p – вероятность возникновения события информационной безопасности. Трехмерный график этой зависимости представлен на рис.1.

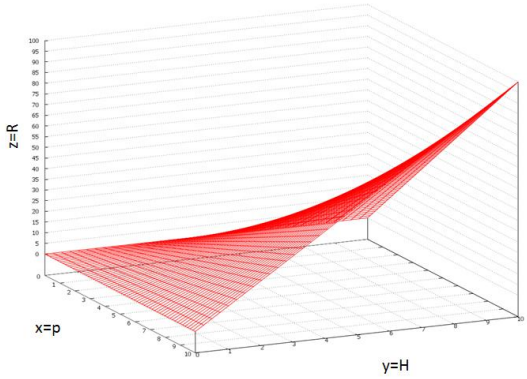


Рисунок 1 – Зависимость уровня риска R от вероятности его реализации p и оценки величины ущерба H

На рис.1 видно, что зависимости уровня риска R от вероятности его реализации p и стоимости ущерба H имеют нелинейный характер. Анализ систем с нелинейностями представляет большие сложности, а если представить, что на практике величина риска зависит от многих факторов (множеств H и p), то анализ подобных систем является исключительно сложным. С другой стороны, задача формирования проектных требований к системе управления информационной безопасностью может быть поставлена и в линейном виде. Рассмотрим это на простейшем

примере двухкомпонентной модели риска представленного на плоскости (1).

В частности, на основе соотношения (1) можно сформировать тривиальный критерий ранжирования рисков. Но, кроме того, можно предположить, что опираясь на соотношение (1) и понятие приемлемого риска $R = R_0$ можно определить вероятностный критерий и его значение, задаваемое в качестве проектного требования при построении систем управления информационной безопасностью. Однако, такой вероятностный критерий не может быть установлен очевидным соотношением $p = R_0 / H$, поскольку величина H является неизвестной. Для этого применяется идея подхода, использующего так называемые «карты риска», которые позволяют «владельцам риска» задавать приемлемые уровни риска $R = R_0$ и разделять все риски на приемлемые и неприемлемые, проведя на «картах риска» линии, соответствующие $R = R_0$ [3, 6].

Тем не менее, следует отметить, что «карты рисков» оперируют единичными проявлениями событий и не учитывают их возможного повторного (многократного) проявления. Накопление последствий совокупности событий, каждое из которых попадает в зону приемлемых, может привести к ущербу более высокому, чем тот, который ассоциирован с каждым из составляющих рисков заданного уровня, даже без учёта такого явления, как провокация одним риском появления другого. Все это приводит к осознанию того, что уровень приемлемого риска единичного события не может быть использован в качестве корректного проектного требования к построению системы управления информационной безопасностью. Иными словами, существующие в настоящее время методики ее построения не имеют возможности трансформировать уровень приемлемого риска, задаваемый собственником, в корректные формальные требования к построению системы управления информационной безопасностью. Из этого, следует вывод о неконструктивности проектного требования к системе управления информационной безопасностью, основанного на концепте «обеспечить уровень риска не выше R_0».

Поэтому корректное проектное требование следует сформулировать иначе, а именно так: система управления информационной безопасностью должна обеспечивать обработку потока рисковых событий с уровнями риска $R \geq R_0$ и заданной вероятностью P_0 появления таких событий. Для обоснования

корректности такого требования необходимо показать возможность определить по заданной величине приемлемого риска $R = R_0$ величину вероятности P_0, с которой проявляются события, ассоциированные с рисками $R \geq R_0$. Иными словами, нужно показать возможность оценивания вероятности P_0 появления события с рисками $R \geq R_0$ для заданного уровня приемлемого риска $R = R_0$. Или по заданному уровню приемлемого риска $R = R_0$ оценить вероятность P_1, с которой могут появляться события с рисками $R < R_0$.

Формирование проектного требования к системе управления информационной безопасностью

Для оценки вероятности P_1 используем двумерную декартову систему координат, по горизонтальной оси которой будем откладывать значения вероятностей p, а по вертикальной оси – значения ущерба H [7]. Очевидно, что значения вероятностей изменяются в диапазоне от $p = 0$ до $p = 1$, а значения ущерба в диапазоне от $H = 0$ до некоторого $H = H_{\max}$. Для единообразия диапазона изменения величины ущерба с диапазоном изменения вероятностей введем в рассмотрение нормированную величину ущерба

$$h = \frac{H}{H_{\max}}.$$

Тогда нормированная величина ущерба будет изменяться в диапазоне от $h = 0$ (при $H = 0$) до $h = 1$ при

$$H = H_{\max}.$$

В декартовых координатах ($h0p$) определим «единичный квадрат» $OACE$ (см. рис. 2), как геометрическое место точек, соответствующих любым возможным значениям нормированного риска r:

$$r = h \cdot p, \qquad (2)$$

где r подчиняется условию $0 \leq r \leq 1$ вследствие выполнения условий $0 \leq h \leq 1$ и $0 \leq p \leq 1$. Поскольку длина каждой из сторон квадрата $OACE$ равна единице, то и площадь $S_{общ}$ квадрата $OACE$ равна 1.

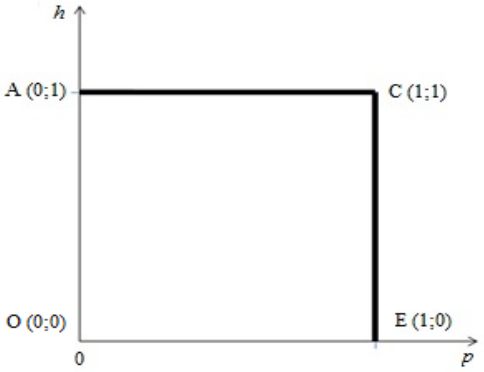


Рисунок 2 – Геометрическое место точек множества любых возможных значений нормированных рисков r

Зададим уровень приемлемого нормированного риска $r = r_0$. Тогда из соотношения (2) очевидно следует функциональная зависимость

$$h = r_0 \cdot \frac{1}{p}, \tag{3}$$

графиком которой является гипербола $h = (1/p)$, сдвигаемая коэффициентом r_0 от начала координат $(0,0)$ по направлению к точке с координатами $(1,1)$. Если наложить гиперболу $h = (1/p)$ на единичный квадрат $OACE$, геометрическое место точек множества всех рисков разделяется на два подмножества (см. рис. 3), а именно: фигура $OABDE$ определяет геометрическое место точек множества значений рисков, для которых выполняется соотношение $r < r_0$, а фигура BCD определяет геометрическое место точек множества значений рисков, для которых выполняется соотношение $r \geq r_0$.

В таком случае вероятность P_1 того, что значение произвольного нормированного риска r не будет превышать значения заданного уровня нормированного риска $r = r_0$, определяется отношением площади фигуры $OABDE$ к площади «единичного квадрата» $OACE$

$$P_1 = \frac{S_{ф}}{S_{общ}}, \tag{4}$$

где $S_ф$ – площадь фигуры $OABDE$, а $S_{общ}$ – площадь «единичного квадрата».

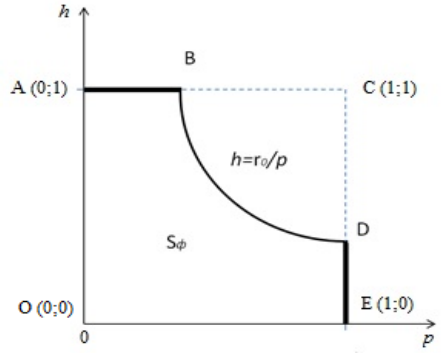


Рисунок 3 – Геометрическое место точек множества значений рисков, разделенное гиперболой $h = (1/p)$

Так как ранее было показано, что $S_{общ} = 1$, то соотношение (4) принимает вид:

$$P_1 = S_ф . \qquad (5)$$

Таким образом, вероятность P_1 того, что для произвольного риска будет выполняться условие $R > R_0$ равна площади фигуры $OABDE$. Остаётся рассчитать площадь этой фигуры.

Для этого разобьём фигуру $OABDE$ на две части (см. рис.4): часть первая – фигура $OABG$ с площадью S_1 и часть вторая – фигура $GBDE$ с площадью S_2. Очевидно, что

$$S_ф = S_1 + S_1 .$$

Площадь S_1 рассчитывается как площадь прямоугольника со сторонами OA и AB. Длина стороны OA, как было ранее обусловлено, равна 1. А длина стороны AB определяется численным значением вероятностной координаты точки B. Точка B есть точка пересечения прямой $b = 1$ с гиперболой, определяемой соотношением (3). Тогда численное значение вероятностной координаты точки B можно определить, подставляя значение $h = 1$ в левую часть соотношения (3):

$$1 = r_0 \cdot \frac{1}{p} . \qquad (6)$$

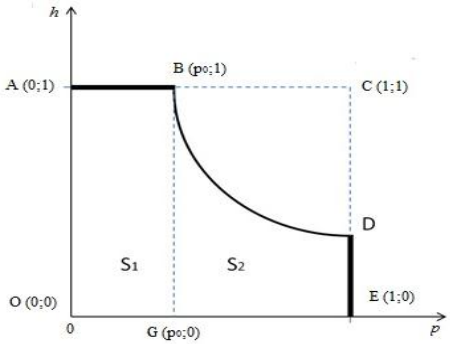


Рисунок 4 – Разбиение фигуры $OABDE$ на две фигуры: прямоугольник $OABG$ и фигуру $GBDE$

Из этого соотношения следует, что численное значение вероятностной координаты $p = p_0$ точки B есть:

$$p_0 = r_0 .$$

Тогда площадь S_1 может быть выражена следующим соотношением:

$$S_1 = 1 \cdot r_0 = r_0 . \qquad (7)$$

Площадь S_2 второй фигуры $GBDE$, которая образована гиперболой, заданной соотношением (3) и тремя прямыми: $h = 0$, $p = p_0 = r_0$ и $p = 1$, вычисляется как определенный интеграл по следующей формуле:

$$S_2 = \int_{r_0}^{1} \frac{r_0}{p} dp = r_0 \int_{r_0}^{1} \frac{1}{p} dp = r_0 \ln p \Big|_{r_0}^{1} = r_0 \left(\ln 1 - \ln r_0\right).$$

Поскольку $\ln 1 = 0$, то формула для вычисления площади S_2 принимает следующий вид:

$$S_2 = r_0 \left(\ln 1 - \ln r_0\right) = -r_0 \ln r_0 . \qquad (8)$$

Тогда для вычисления площади фигуры $OABDE$ подставим в (6) значения (7) и (8) и получим:

$$S_\phi = S_1 + S_2 = r_0 - r_0 \ln r_0 = r_0 \left(1 - \ln r_0\right). \qquad (9)$$

Итак, с учетом (5) получается формула для оценки вероятности P_1 того, что нормированные значения величины возможных рисков не будут превышать заданной величины приемлемого риска r_0:

$$P_1 = r_0\left(1 - \ln r_0\right). \qquad (10)$$

Проанализируем полученное соотношение.

Во-первых, поскольку для значений r_0 выполняется условие $0 \le r_0 \le 1$, постольку функция $\ln r_0$ в формуле (10) принимает отрицательные значения $\ln r_0 < 0$. За счет этого вычитаемая величина $(-r_0 \ln r_0)$ в формуле (10) превращается в положительное слагаемое. Для того, чтобы этот факт отразить явным образом, формулу (10) представим в следующем виде:

$$P_1 = r_0\left(1 + \ln\left(r_0^{-1}\right)\right). \qquad (11)$$

Пример положения графика этой функции относительно графика линии $P = r_0$ показано на рис.5.

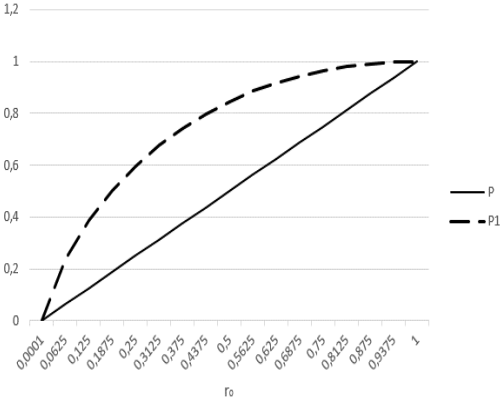


Рисунок 5 – Положение графика функции $P_1 = r_0\left(1 + \ln\left(r_0^{-1}\right)\right)$ по отношению к графику функции $P = r_0$

Из соотношения (11) следует, что вероятность P_1, с которой могут возникать нормированные риски $r < r_0$, почти всегда превышает значение заданной величины этого приемлемого нормированного риска r_0, за исключением единственного случая $r_0 = 1$. В этом крайнем случае $\ln r_0 = 0$ и соотношение (11) принимает вид

$$P_1 = r_0\left(1 + \ln\left(r_0^{-1}\right)\right) = 1 \cdot (1 + \ln 1) = 1 \cdot (1 + 0) = 1,$$

и это является формальным отражением того тривиального факта, что если максимальную величину ущерба $H = H_{\max}$ задавать в качестве приемлемой, то тогда любые значения рисков являются допустимыми.

Во-вторых, можно определить максимальную погрешность замены вероятности P_1 риском r_0 (т.е. вероятностью $P = r_0$), как отклонение функции, заданной соотношением (11), от линии $P = r_0$, взяв следующую разность:

$$P_1 - P = r_0\left(1 + \ln\left(r_0^{-1}\right)\right) - r_0 = r_0 \ln\left(r_0^{-1}\right).$$

График функции, соответствующей такой разности, приведен на рис. 6 и из него можно непосредственно получить, что:

1) максимальное значение погрешности оценивания вероятности ненамного превышает значение 0.36 (а если точно, то оно равно 0.3678) от единицы нормированного уровня риска;

2) максимальное значение погрешности достигается в окрестности значений нормированного риска $r_0 = 0.36$;

3) превышение уровня 10% погрешности оценивания вероятности может наблюдаться на 80% возможных значений r_0;

4) уровень погрешности, превышающий 36%, возможен более чем на 10% всех значений r_0.

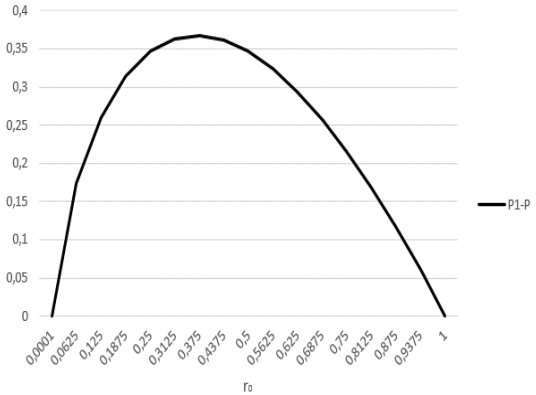


Рисунок 6 – График разности функции $P_1 = r_0\left(1 + \ln\left(r_0^{-1}\right)\right)$ и $P = r_0$

Выводы

Применение геометрического подхода к оцениванию вероятности P_1 того, что произвольные значения нормированного риска r угроз безопасности информации будут попадать в зону $r < r_0$, дало возможность получить точную количественную оценку этой вероятности в виде формулы (11). Как следствие, установлено,

что такая вероятность P_1 практически всегда превышает уровень r_0. При этом в большинстве случаев это отличие достигает 30%, а более чем на 10% всех случаев различие даже слегка превышает 36%.

Благодаря этому стало возможным трансформировать субъективный показатель риск-аппетита владельца риска, отображаемый в виде приемлемого уровня риска, в формализованный вероятностный критерий, на основе которого можно сформулировать проверяемые проектные требования к построению систем управления информационной безопасностью.

Литература

1. ISO 27001 – Information Management Security System [Electronic resource]. – Access mode : http://www.enhancequality.com/iso-standards/iso-27001-information-security-management-system/. – Access data : June 2016. – The title of the screen.
2. Дмитриев А. Менеджмент информационной безопасности [Электронный ресурс]. – Режим доступа: http://www.comizdat.com/index_.php?in=ksks_articles _id&id=568. – Дата доступа : сентябрь 2016. – Название с экрана.
3. Information technology. Security techniques. Information security management systems. Requirements : ISO/IEC 27001:2013. – Second edition 2013-10-01. – Geneva, 2013. – P. 23.
4. Information technology. Security techniques. Code of practice for information security controls : ISO/IEC 27002:2013. – Second edition 2013-10-01. – Geneva, 2013. – P. 80.
5. Risk management. Principles and guidelines : ISO 31000:2009. – First edition 2009-11-01. – Geneva, 2009. – P. 24.
6. Information technology. Security techniques. Information security risk management : ISO/IEC 27005:2011. – Second edition 2011-06-10. – Geneva, 2011. – P. 68.
7. Кендалл М. Геометрические вероятности / М. Кендалл, П. Моран. – М. : Наука, 1972. – 192 с.

ІНФОРМАЦІНІЙНІ РЕСУРСИ ДОСТУПУ ТА ОБМІНУ НАУКОВОЮ ІНФОРМАЦІЄЮ, СИСТЕМИ ІДЕНТИФІКАЦІЇ НАУКОВЦІВ - МОЖЛИВОСТІ, НЕДОЛІКИ, ПЕРЕВАГИ

В.Б. Андрущенко, І.В.Балагура, Д.В. Ланде,

Інститут проблем реєстрації інформації НАН України

Вступ

Наука є одним із важливіших чинників, що забезпечує розвиток світу. На сьогодні інформаційні технології є основним інструментом в кожному аспекті існування людини, науки і її проегресу зокрема. Інформаційні технології – є вагомим інструментом наукової взаємодії: це не тільки прискорення обміну інформації, а й вдосконалення процедур її пошуку та отримання.

В той же час одним із наріжних каменів інтернаціоналізаційного процесу в науковій сфері є обмеження наукової діяльності рамками окремих осередків, спільнот і держав.

Необхідно розглянути причини та шляхи вирішення проблем видимості накових досягнень у світовому науковому просторі.

Кожна країна впроваджує власні схеми фінансування і відповідно висуває відповідні вимоги до проведення, реалізації, популяризації та впровадження наукових досліджень.

Одна із найбільш розповсюджених схем – часткова підтримка лабораторій, колективів вчених та окремих проектів із залишковою грантовою підримокю проектів.

На сьогоднішній день існує низка як національних так і глобальних грантових программ – одноосібні гранти, гранти для колективів та міжнародних колаборацій. Серед вимог, що висуваються до вченого, колективу та організації, що бере участь у проекті або у формуванні колаборації – певна кількість опублікованих наукових праць належної якості – публікація у видання що індексується провідними наукометричними базами даних із відповідними показниками цитованості.

З огляду на це і на стрімкий розвиток глобальних мереж і інформаційних систем в межах них достатньо розповсюдженими є наукометрині та реферативні бази даних, репозиторії наукових праць та препринтів, системи індетифікації науковців та соціальні мережі для науковців. Для реалізації задач інтернаціоналізації науки тієї чи іншої країни важливим є видимість досягнень світовій науковій спільноті, що можливо реалізувати за рахунок популяризації власних

досягнень за рахунок участі у інформаційних наукових ресурсах глобальних мереж.

Мета роботи – опрацювати масив інснуючих інформаційних наукових ресурсів для пошуку, поширення та обміну науковою інформацією, пошуку партнерів для формування наукових колаборацій, публікації результатів та визначення наукометричних показників. Аналіз систем дозволить визначити переваги та недоліки з огляду на ергономічність ресурсів, а також розширення можливостей для формування додаткових даних та створення передумов для вдосконалення проаналізованих інформаційних ресурсів. Також окремий акцент в роботі буде зроблено на участь українських вчених у вищезазначених системах.

Порівняльний аналіз наукових ресурсів

На сьогоднішній день не існує стандартів та вимог до реферативних ресурсів. Оцінки подібних баз даних можна реалізувати тільки через призму стандартів розроблених для програмного забезпечення. Можна виділити лише інтуїтивні та загальні критерії для порівняння сучасних реферативних баз даних. Серед таких критеріїв варто виділити такі:

- предметна область,
- повнота та достовірність інформації,
- структурованість даних,
- наявнсть посилань на повні тексти документів,
- швидкість відображення першоджерел,
- зручність інтерфейсу (функцій пошуку зокрема),
- наявність наукометричних показників,
- можливість реєстрації у системі з отриманням додаткових переваг,
- можливості створення власного профілю з додатковими функціями,
- управління приватністю наданої інформації,
- інше [1].

На сьогоднішній день було достатньо докладно висвітлено у публікаціях досілдників порівняльний аналіз провідних світових наукометричних ресурсів Web of science, Scopus та Google Scholar [2,3]. Але подібні роботи присвячені переважно інформації, доступ до якої забезпечує той чи інший ресурс, та способам обчислення наукометричних показників і набору даних, що опрацьовується системами. В той же час не було приділено достатньо уваги

структурам цих систем та інших факторів, за якими реалізоване дане дослідження.

В даній роботі на прикладі відкритих ресурсів буде проілюстровано порядок порівняння систем та висновки в рамках проведеного аналізу.

Для провадження грунтовного аналізу і формування поступових результатів із подальшим їх опрацюванням було виокремлено низку ресурсів, що користуються попитом і можуть виступати допоміжними інфтрументами в роботі науковців при формуванні стратегій не тільки власних досілджень, а й стратегій побудови власної кар'єри дослідника.

В Таблиці 1 наведено перелік та короткий опис ресурсів – наукометричних баз даних, реферативних баз даних, баз даних накових видань окремих галузей, бази даних повних текстів наукових статей, систем ідентифікації науковців, архів препринтів.

Таблиця 1

Коротка характеристика інформаційних ресурсів

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
Scopus	Реферативна база даних і наукометрична платформа, що була створена в 2004 р. Присутня можливість створення профілю науковця і організації. Scopus не містить, але надає посилання на повні тексти документів, або на умови доступу до текстів. Доступ до бази даних є передплачуваним. Science Direct – ресурс відкритого доступу на базі Scopus, що дозволяє проводити пошук публікацій за кількома параметрами.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
 Web of Science (WoS)	Реферативна наукометрична база компанії ThomsonReuters. Наукометричний аппарат платформи забезпечує відстеження показників цитованості публікацій з ретроспективою до 1900 р. Одним з ключових концептів наукометричного апарату платформи є імпакт-фактор (індекс впливовості) наукового видання. Профіль актора може бути створений в безкоштовному ресурсі Thomson Reuters – Researchers ID, що корелюється з ORCH ID.
 Index Copernicus	IndexCopernicus (IC) наукометрична база даних, створена в 1999 році в Польщі. База даних має кілька інструментів оцінки продуктивності, які дозволяють відслідковувати вплив наукових робіт і публікацій, окремих вчених або науково-дослідних установ. Система дозволяє архівувати і багатовимірно аналізувати досягнення вчених від імені установи, забезпечує доступ до зовнішніх баз даних і додаткових інструментів для наукової співпраці.На додаток до продуктивності, індекс Коперникус також пропонує традиційне реферування та індексування наукових публікацій.
 Math Sci Net	Одна з найбільш авторитетних реферативних баз даних по математиці, підтримувана Американським математичним суспільством (AMS). Індексується більше 1800 математичних журналів, крім того є записи на 85000 монографій і 300000 доповідей з наукових конференцій. Всього більше 3 млн. записів, 2,2 млн. з них забезпечені рефератом/рецензією. Охоплення - з початку 1900 рр. по теперішній час. (Доступ обмежений, за передплатою).

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
ebsco.com	EBSCO host - служба, що надає доступ до баз даних англомовних періодичних видань. Частина статей в базах представлена у вигляді повних текстів, частина - тільки у вигляді анотацій. У EBSCO включені як найсвіжіші номери журналів, так і архів - для деяких видань аж до 1950-х років. EBSCO host підключає користувача до кількох баз даних різної тематики. Найбільш корисними є бази даних Academic Search Premier, Buisness Source Premier і Master FILEP remier - ті, в яких представлені журнали з економіки, менеджменту, соціології, політології, права та інші.
Academic Search Premier	База даних наукових журналів. Тематика універсальна. Більш 3600 найменувань, включаючи повні тексти з більш ніж 2700 журналів, відрецензованих науковою громадськістю. Хронологічний обхват з 1975 року по теперішній час. База даних щодня оновлюється.
Business Source Premier	База даних по бізнесу та економіці, включаючи, фінанси, менеджмент, бухгалтерський облік, міжнародний бізнес та ін. повні тексти. Більш 2800 наукових журналів, включаючи більше 900 видань, відрецензованих науковою громадськістю, та реферати з 3350 журналів. Містить більш ніж 5000 описів найбільших світових компаній, а також економічні звіти країн світу. Щодня оновлюється.
Master FILE Premier	База даних універсального змісту, що забезпечує доступ до бібліографічних посиланнях, рефератів і повних текстів на публікації з наукових та науково-популярних журналів, починаючи з 1975году по теперішній час. Включає також повні тексти книг (164 найменування), переважно довідників, близько 100 000 біографій, офіційні документи, колекцію фотографій, карт, прапорів. Щодня оновлюється.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
 Wilson Business Abstracts	Пропонує множину ділових і наукових журналів, є ідеальним ресурсом для тих, хто хоче провести дослідження або знайти інформацію в будь-якій області бізнесу. Облік; Придбання і злиття; Реклама; Банківська справа; Будівництво і Конструювання; Хімічна та фармацевтична промисловість; комунікації; Комп'ютери; косметична промисловість; Економіка; електроніка; індустрія розваг; фінанси; фінансові послуги; постанови Уряду; Охорона здоров'я; Гостинність і туризм; людські ресурси; трудові відносини; Страхування; Міжнародний бізнес; інвестиції; управління; маркетинг; ЗМІ; Охорона праці та безпека; Нафта і Газ; Паперова та целюлозно-паперова промисловість; комунальні підприємства; Видавництво; Купівля; Нерухомість; Роздрібна торгівля; малий бізнес; оподаткування; технологія; транспорт.
 Econlit	Ця база даних містить більше одного мільйона записів, з цитатами і тезами з 1886 року. Ринки капіталу; країнознавство; економетрія; економічне прогнозування; економіка природокористування; постанови Уряду; економіка праці; теорія грошей; економіка міста.
 Regional Business News	Повнотекстова база даних регіональних новин в області бізнесу, забезпечує доступ до публікацій з 75 журналів з бізнесу, газетам, телеграфним повідомленнями з усіх регіонів США. Щодня оновлюється.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
 Wilson Applied Science & Technology	База даних надає вичерпну індексацію, що охоплює широкий спектр міждисциплінарних галузей на основі широкого масиву в науково-технічних журналах. Акустика; Повітроплавання; Прикладна математика; Атмосферні науки; Хімічне машинобудування; Цивільне будівництво; Зв'язок та інформаційні технології; Інженерні та біомедичні матеріали; Енергетичні ресурси та наукові-дослідження; Моделювання експлуатаційних умов; Геологія; Промислове проектування; Морські технології; Машинобудування; Металургія; Гірниче машинобудування; Нейронні мережі; Ядерна техніка; Океанографія; Оптичні і нейронні обчислення; Фізика; Робототехніка; Космічна Наука; Транспорт; Поводження з Відходами.
 Wilson Social Sciences Abstracts	База даних освітлює найостанніші концепції, теорії та методи щодо прикладного і теоретичного аспектів соціальних наук.
 Wilson Humanities Abstracts	Ця база даних містить реферати та бібліографічні індексації найвідоміших наукових джерел у галузей гуманітарних наук.
 Scimago Journal & Country Rank (SJR)	Сайт показника рівня цитованості наукових журналів більше 230 країн світу на базі інформаційної системи Scopus (Elsevier BV). Можливий пошук за багатьма параметрами, отримання різних варіантів візуалізації результатів.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
Google Scholar	Вільно доступна пошукова система, яка індексує повний текст наукових публікацій всіх форматів і дисциплін. Google Scholar включає статті, що опубліковані в журналах, зберігаються в репозиторіях або знаходяться на сайтах наукових колективів чи окремих вчених. Передбачає можливість створення та управління власним профілем наковця чи організації.
Російський індекс наукового цитування (РІНЦ)	Національна інформаційно-аналітична система, безкоштовний загальнодоступний інструмент вимірювання та аналізу публікаційної активності вчених і організацій. У базу також включені доповіді на конференціях, монографії, навчальні посібники, дисертації. База містить відомості про вихідні дані, авторів публікацій, місцях їх роботи, ключові слова і предметні рубрики, а також анотації та пристатейні списки літератури. Хронологічне охоплення - з 2005 р. Загальний обсяг публікацій, що надходять у РІНЦ щорічно, складає більш 280 000 статей. Крім того, понад 2500 журналів представлені повними текстами, у тому числі 1400 журналів - у відкритому доступі.
«ZBMATH – Thedatabase Zentralblatt MATH»	Метою є збір, систематизація, публікація та розповсюдження бібліографічних даних та рефератів книг і статей, що присвяченні всім розділам математики та її прикладне застосування в інформатиці, механіці і фізиці. Реферуються більше 2300 журналів і періодичних видань різних країн, щорічно публікується близько 80000 анотацій і рецензій, написаних більш ніж 5000 вченими. Більшість рефератів публікується англійською мовою, деякі - французькою чи німецькою.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
 INSPEC	Ведуча англомовна реферативна науково-технічна база даних. Ресурс створюється Лондонським Інститутом інженерів з електротехніки (The Institution of Electrical Engineers, IEE) і містить в даний час більше 8 млн. записів: реферати публікацій з більш 3500 наукових журналів з фізики, електроніці, інформатиці, комп'ютерних технологій і технічних наук, майже 2 тис. матеріалів наукових конференцій. У базу включаються також описи книг , технічних звітів і дисертацій. Поповнення становить близько 400 000 записів щорічно. Хронологічний обхват: з 1969р. по теперішній час. База даних INSPEC відповідає друкованим виданням IEE: ScienceAbstractsseries, PhysicsAbstracts, Electrical&ElectronicsAbstracts,Computer&ControlAbstracts.
 ERIC (Educational Resource Information Center)	База даних з проблем освіти. Надає доступ до повних текстів більш ніж 2200 збірників статей з проблем освіти, а також містить реферати та описи статей із понад 1000 наукових журналів з освітньої тематики. Створена міністерством освіти США та відділом досліджень і розвитку в галузі освіти.
Research Gate	Соціальна мережа для науковців. Передбачає реєстрацію, створення власного профілю. Надає можливості переглядати профілі інших вчених відповідно до налаштувань приватності. Відстежувати публікаційну активність та цитування.
PubMed	Реферативна база даних медико-біологічного спрямування. Запроваджена Національною медичною бібліотекою США. Містить посилання на повні тексти публікацій відповідного напрямку.

Найменування наукометричної бази даних	Коротка відомість про наукометричну базу даних
OrchID	Система ідентифікації дослідника. Передбачає створення власного профілю із зазначенням короткого резюме: освіта, місце роботи, перелік публікацій та інших наукових досягнень. Дозволяє за умов налаштувань приватності переглядати профілі інших учасників.
ReserchерID	Система ідентфікації досілдника. Передбачає створення власного профілю із можливістю налаштувань автоматичного додавання власних публікацій, що індексуються наукометричною базую даних WebOfScience. Інтегрується із профілями у системі OrchID
ArXiv	Електронний архів препринтів наукових публікацій з фізики, комп'ютерних наук, біології, статистики, фінансів. Містить повні тексти документів.
PlosOne	Перше в світі міждисциплінарне видання відкритого доступу. Містить повні тексти документів.

На прикладі порівняльного аналізу двох відкритих ресурсів було продемонстровано окремі етапи реалізації глобальної роботи з оцінки систем та в подальшому формуванню пропозицій щодо вимог, які можуть бути висунуті до подібних систем.

До розгляду було взято системи, що є доступними в глобальній мережі і не передбачають передплати: Google Scholar Citations (Google) та Science Direct (Elsevier). Для систем було застосовано однакові теги для пошуку. З огляду на застосування тегів, було визначено спільне у представленні інформації наукометричними системами і в той же час зауважити обмеження у доступі інформації.

Порівняння проводилося за 5-ма наочними критеріями, обробка яких надасть можливість для визначення недоліків і шляхів побудови алгоритмів для оптимізації роботи систем: доступність для користувача (потреба в реєстрації), спосіб організації пошуку вихідної інформації, представлення інформації, сортування

результатів пошуку та можливості щодо зміни порядку сортування даних, можливості уточнення пошук

Результати порівняльного аналізу, що є необхідними для формування унікального алгоритму, що на даному етапі дисертаційного дослідження може бути реалізований та застосований для кількох систем, зокрема Google Scholar та Science Direct, відображено у таблиці 2.

Таблиця 2.

Порівняльний аналіз структурних елементів наукометричних систем

Науко–метрична система	Доступ–ність для користу–вача	Пошук інформа–ції	Представлення інформації	Сорту–вання	Можли–вості уточнення пошуку
Google Scholar	Не обов'яз–кова реєстрації в системі	В один рядок	- Назва публікації, монографії, що містить тег пошуку - Назва видання та реквізити номеру - Автори - Випадаюче меню для абстракту та основних здобутків	Довільне	- Рік - Часовий проміжок - Дата - Релевант–ність
Science Direct	Не обов'яз–кова реєстрації в системі	Можли–вість зазначен–ня кількох параметрі в пошуку	- Назва публікації, монографії, що містить тег пошуку - Автори та реквізити видання - Випадковий текст абстракту - Кількість цитувань, пов'язані публікації, текст для посилання, способи відображення/зазна–чення публікації в системі - Посилання на доступ до публікації	За датою публікації, від нових	- Рік - Назва видання - Тема - Видання (журнал, моногра–фія, довідкові матеріали)

Висновки

1. Проведено порівняльний аналіз наукових інформаційних ресурсів, які є доступними для користування в Україні, серед безлічі ресурсів для більш детального аналізу обрано безоплатні Google Scholar та ScienceDirect.

2. Відповідно до отриманих даних можна зауважити, що системи є подібні у доступі до інформації, жодна не передбачає створення власного профілю користувача для доступу до інформації, доступ є необмежений, порівняно із зареєстрованими у системах користувачами. І в той же час кожен зареєстрований користувач отримує додаткові пріоритети у вигляді можливості збереження отриманої і скорегованої в результаті пошукових операцій інформації.

3. Пошук інформації в системі Google Scholar є спрощеним, і в той же час сервіси Science Direct пропонує користувачеві одразу кілька параметрів для здійснення пошуку.

4. Представлення інформації згідно пошуку є подібним і в той же час доступ до інформації про абстракт та основні результати в системі Science Direct є доступними зі сторінки результатів пошуку і не вимагають додаткових кроків.

5. Сортування публікацій у системі Science Direct є досконалішим, що дозволить уникнути певних етапів алгоритму при реалізації методології розширення можливостей систем.

6. Можливості уточнення пошуку за тегом є реалізованими більш орієнтованими на користувача і дозволяють за рахунок чітких уточнень швидше отримати інформацію для реалізації запланованих надбудов.

7.

Література

1. Балагура І.В. Перспективи розвитку реферативної бази даних «Україніка наукова» та реферативного журналу «Джерело» / Наукові праці Державної науково-педагогічної бібліотеки Ук раїни ім. В.О. Сухомлинського. Науково-методичні та організаційні засади інформаційно-аналітичного забезпечення педагогічної науки, освіти і практики України: стан та перспективи. – 2012. – Вип.3. – с.115-125.

2. Moed Henk F., Judit Bar-IIan, Gali Halevi A new methodology for comparing Google Scholar and Scopus /Journal of Informetrics– 2016. – V.10,I.2 – P.533-551.

3. Franceschini Fiorenzo, Maisano Domenico, Mastrogiacomo Luca. Empirical analysis and classification of database errors in Scopus and Web of Science/Journal of Informetrics– 2016. – V.10,I.4 – P.933-953.

СОДЕРЖАНИЕ

Додонов А.Г., Ландэ Д.В.
ИССЛЕДОВАНИЕ ИСТОЧНИКОВ ИНФОРМАЦИОННОГО ВЛИЯНИЯ ВЕБ-РЕСУРСОВ СЕТИ ИНТЕРНЕТ 3

Додонов А.Г., Горбачик Е.С., Кузнецова М.Г.
ПРОБЛЕМЫ БЕЗОПАСНОСТИ СОЦИОТЕХНИЧЕСКИХ СИСТЕМ 13

Головко О.М.
ДО ДЕЯКИХ АСПЕКТІВ ФОРМУЛЮВАННЯ ПОНЯТТЯ «ІНФОРМАЦІЙНІ ЗАГРОЗИ» 20

Жиров Г.Б.
ЗАСТОСУВАННЯ ВЕЙВЛЕТ АНАЛІЗУ В ЗАДАЧАХ РОЗРІЗНЕННЯ СИГНАЛІВ, ДІАГНОСТУВАННЯ ТА ПРОГНОЗУВАННЯ ВІДМОВ 25

Каденко С.В.
МОЖЛИВОСТІ ТА ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ЕКСПЕРТНИХ ТЕХНОЛОГІЙ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ У СФЕРІ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ 31

Кузнєцова Н.В.
СКОРИНГОВІ ТЕХНОЛОГІЇ ОЦІНЮВАННЯ РИЗИКІВ ШАХРАЙСТВА В БАНКІВСЬКІЙ ДІЯЛЬНОСТІ 43

Ландэ Д., Березин Б., Павленко О.
ПОСТРОЕНИЕ МОДЕЛИ ИНФОРМАЦИОННОГО СЕРВИСА НА БАЗЕ НАЦИОНАЛЬНОГО СЕГМЕНТА ИНТЕРНЕТ 48

Рогушина Ю.В., Гладун А.Я., Снігирь Г.В.
ОНТОЛОГІЧНИЙ ПІДХІД ДО РОЗРОБКИ НАЦІОНАЛЬНИХ СТАНДАРТІВ УКРАЇНИ З ОЦІНЮВАННЯ БЕЗПЕКИ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ 58

Цыганок В.В.
ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ ПРИ ПЛАНИРОВАНИИ МЕРОПРИЯТИЙ ПО ПРОТИВОДЕЙСТВИЮ ИНФОРМАЦИОННЫМ ОПЕРАЦИЯМ 72

Максименко Е.В.
МОДИФИЦИРОВАННЫЙ МЕТОД ФАКТОРИЗАЦИИ ФЕРМА И ИССЛЕДОВАНИЕ ЕГО ПРЕДЕЛЬНЫХ СВОЙСТВ 87

Соколов В.В.
ПІДХІД ДО ГЕНЕРАЦІЇ ОБ'ЄКТНИХ ПРОГРАМ РОЗВ'ЯЗКУ ЗАДАЧ ПРЕДМЕТНОЇ ОБЛАСТІ 90

Субач І.Ю., Фесьоха В.В., Прокопенко В.Р.
ОСНОВНІ ПРІОРИТЕТИ ВДОСКОНАЛЕННЯ СИСТЕМ ЗАПОБІГАННЯ ВТОРГНЕННЯМ В ІНФОРМАЦІЙНІ МЕРЕЖІ ОРГАНІВ ВІЙСЬКОВОГО УПРАВЛІННЯ 93

Юдін О.М.
ДОСТУПНІСТЬ – ГОЛОВНИЙ ЧИННИК ІНФОРМАЦІЙНОГО НАПОВНЮВАННЯ ВЕБ-САЙТІВ 96
Хлапонін Ю.І.
ІСТОТНІ ОСОБЛИВОСТІ ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ В КІБЕРПРОСТОРІ 101
Хлапонін Ю.І., Жиров Г.Б
АНАЛІЗ ТА МОНІТОРИНГ ТЕЛЕКОМУНІКАЦІЙНОЇ МЕРЕЖІ НА ОСНОВІ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ 104
Бойченко А.В.
РОЗРОБКА СЦЕНАРІЇВ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ ОНТОЛОГІЧНОЇ МОДЕЛІ 118
Прищепа С.В.
ВЫЯВЛЕНИЕ СОБЫТИЯ, ЕГО СУБЪЕКТА И ОБЪЕКТА В ТЕКСТОВЫХ ДОКУМЕНТАХ 121
Снарский А., Ланде Д., Зоринец Д.
РАНЖИРОВАНИЕ ПОНЯТИЙ, ИЗВЛЕКАЕМЫХ ИЗ ПОТОКОВ СЕТЕВЫХ НОВОСТЕЙ 130
Шнурко-Табакова Е.В.
ВЫЯВЛЕНИЯ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ И АГЕНТОВ ВЛИЯНИЯ НА ДИСКРЕДИТАЦИЮ РУКОВОДСТВА УКРАИНСКОЙ АРМИИ 132
Андрійчук О.В., Качанов П.Т.
МЕТОДИКА ПРИМЕНЕНИЯ ИНСТРУМЕНТАРИЯ ЭКСПЕРТНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ИДЕНТИФИКАЦИИ ИНФОРМАЦИОННЫХ ОПЕРАЦИЙ 141
Зубок В.Ю.
ВСЕСВІТНІ ІНТЕРНЕТ-ПРОВАЙДЕРИ В УКРАЇНСЬКІЙ МЕРЕЖІ ОБМІНУ ТРАФІКОМ: ВИКЛИКИ ТА МОЖЛИВОСТІ 156
Грайворонська А.М.
СТАТИСТИЧНІ ВЛАСТИВОСТІ МАТЕМАТИЧНОЇ МОДЕЛІ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ 163
Кузьмичов А. І.
МЕТОД МАРКОВСЬКИХ ЛАНЦЮГІВ У МЕРЕЖЕВИХ МОДЕЛЯХ ПРОЕКТІВ ДЛЯ ПРОГНОЗУВАННЯ ТА РИЗИК-АНАЛІЗУ 167
Мохор В., Богданов А., Бакалинский А., Цуркан В.
МЕТОД ФОРМИРОВАНИЯ ПРОЕКТНЫХ ТРЕБОВАНИЙ К СИСТЕМЕ УПРАВЛЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ 169
Андрущенко В.Б, Балагура І.В., Ланде Д.В.
ІНФОРМАЦІНІЙНІ РЕСУРСИ ДОСТУПУ ТА ОБМІНУ НАУКОВОЮ ІНФОРМАЦІЄЮ, СИСТЕМИ ІДЕНТИФІКАЦІЇ НАУКОВЦІВ - МОЖЛИВОСТІ, НЕДОЛІКИ, ПЕРЕВАГИ 180

Национальная академия наук Украины
Институт проблем регистрации информации

Национальный технический университет Украины «КПИ»
Факультет социологии и права
Учебно-научный центр информационного права и правовых вопросов информационных технологий

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И БЕЗОПАСНОСТЬ

Материалы XVI Международной научно-практической конференции

Выпуск 17

Підп. до друку 10.01.2017. Формат 60x84 1/16. Папір офс. Гарнітура Times.
Спосіб друку – ризографія. Ум. друк. арк. 12,65. Обл.-вид. арк. 23,66. Наклад 100 пр.
Зам. № 15-200.