

УДК 004.67+519.83

ЛАНДЕ Д.В., доктор технічних наук, старший науковий співробітник

ЖИТТЄВИЙ ЦИКЛ ІНФОРМАЦІЙНИХ ОБ'ЄКТІВ

Анотація. Інформаційна компонента інформаційно-аналітичної системи – це змістовна складова її інформаційного забезпечення, сукупність інформаційних об'єктів. У статті наведено опис моделей життєвого циклу інформаційних об'єктів, основні методи підвищення живучості інформаційної складової інформаційно-аналітичних систем. Ці методи спрямовані на зменшення рівня вразливості інформаційних об'єктів у межах інформаційної інфраструктури.

Ключові слова: інформаційна складова, життєвий цикл, живучість, інформаційний об'єкт, інформаційні потоки, математичне моделювання.

Аннотация. Информационная компонента информационно-аналитической системы – это содержательная составляющая ее информационного обеспечения, совокупность информационных объектов. В статье приведено описание моделей жизненного цикла информационных объектов, основные методы повышения живучести информационной составляющей информационно-аналитических систем. Эти методы направлены на уменьшение уровня уязвимости информационных объектов в пределах информационной инфраструктуры.

Ключевые слова: информационная составляющая, жизненный цикл, живучесть, информационный объект, информационные потоки, математическое моделирование.

Summary. Information component of information-analytical system is a meaningful component of its information security, a set of information objects. The paper describes the model of the life cycle of information objects, the main methods to improve the survivability of the information component of information-analytical systems. These techniques are aimed at reducing the vulnerability of the information objects within the information infrastructure.

Keywords: information component, life cycle, survivability, information object, information flows, mathematical modeling.

Постановка проблеми. Сучасний інформаційний простір є динамічною документальною системою, що складається з пов'язаних за змістом елементів – інформаційних об'єктів, які утворюють в динаміці своєї еволюції інформаційні потоки [1].

На цей час у зв'язку з розвитком інформаційних технологій особливе місце серед завдань, що отримали актуальність, займають завдання, пов'язані з моделюванням їх життєвого циклу, а саме – окремих етапів:

- генерування інформації (формування і розвитку);
- передачі інформації;
- формування сховища інформаційних ресурсів (баз даних, інформаційних масивів, окремих документів тощо);
- безпосереднього використання інформаційної складової;
- реакції на деструктивні впливи, відновлення, руйнування;
- утилізації (архівування) інформації.

Питання життєвого циклу інформаційних об'єктів знаходить широке відображення у законодавчих актах України. Зокрема, у Законі України “Про інформацію” від 02.10.92 р. № 2657-ХІІ (ч.1 ст. 3) регламентуються інформаційні відносини на всіх етапах життєвого циклу інформаційних об'єктів, зокрема “створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації”.

Закон України “Про захист персональних даних” від 01.06.10 р. № 2297-VI “регулює правові відносини, пов’язані із захистом і обробкою персональних даних, і спрямований на захист основоположних прав і свобод людини і громадянина, зокрема права на невтручання в особисте життя, у зв’язку з обробкою персональних даних” (ст. 1).

Закон України “Про науково-технічну інформацію” від 25.06.93 р. № 3322-XII “порядок формування і реалізації” науково-технічної інформації в інтересах науково-технічного, економічного і соціального прогресу країни. Це далеко не повний перелік законодавчих актів у інформаційній сфері щодо означених питань.

Виклад основних положень. Основним об’єктом моделювання інформаційних потоків [2] сьогодні є їх тематичні зрізи, послідовності документів, що відповідають певній тематиці. Тематичним інформаційним потокам можна поставити у відповідність часові ряди, для вирішення завдань аналізу яких все частіше застосовуються статистичний, дисперсійний, фрактальний, Фур’є і вейвлет-аналіз [3].

Динаміка тематичних інформаційних потоків визначається комплексом як внутрішніх, так і зовнішніх нелінійних механізмів, що мають враховуватися при моделюванні (можливо, у неявному вигляді). Досить часто задовільним виявляється спрощене розуміння тематичного інформаційного сюжету як деякої залежної від часу величини $n(t)$, поведінка якої описується нелінійними рівняннями. Сьогодні при моделюванні інформаційних потоків використовуються переважно нелінійні моделі, застосовуються методи нелінійної динаміки, теорії клітинних автоматів, перколяції, самоорганізованої критичності [1, 4].

Багато сучасних інформаційно-аналітичних систем містять у своєму складі засоби відображення статистики входження в бази понять, відповідних запитам користувачів. Зокрема, у рамках цих досліджень використовувалася система контент-мониторинга веб-простору InfoStream ([//www.infostream.ua](http://www.infostream.ua)), що має цю функціональність.

Моделювання інформаційних потоків.

Для вивчення інформаційних сюжетів як складних багатопараметричних систем, параметри яких ще мало вивчені, найбільш відповідною методикою є математичне моделювання. Життєвий цикл інформаційних об’єктів при цьому може описуватися, наприклад, моделлю дифузії інформації [5]. Процеси дифузії інформації, як і процеси дифузії, у фізиці досить точно моделюються за допомогою методів клітинних автоматів. Модель дифузії інформації, яку розглядатимемо надалі, є двовимірною.

У системі клітинних автоматів найближчими сусідами, що входять до околу елемента $y_{i,j}$, вважаються елементи, розташовані поряд з ним (так званий окіл Мура – кожна клітина має вісім сусідів). Значення клітини на кроці еволюції $t + 1$ в порівнянні з кроком t має вигляд:

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

У рамках цієї моделі поширення новин в інформаційному просторі, застосовуються окіл Мура й імовірнісні правила поширення новин за заданою тематикою. Передбачається, що клітина може перебувати в одному з трьох станів: 1 – “свіжа новина” (клітина забарвлюється у чорний колір); 2 – новина застаріла, але збережена у вигляді відомостей (сіра клітина); 3 – клітина не має інформації (клітина біла, інформація не дійшла або забута). Правила розвитку інформаційного сюжету наступні:

– спочатку усе поле складається з білих клітин за винятком однієї – чорної, яка першою “прийняла” новину;

–біла клітина може перефарбовуватися тільки в чорний колір або залишатися білою (вона може отримувати новину або залишатися такою, якою була);

–біла клітина перефарбовується, якщо виконується умова: $Cpm > 1$, де: p – псевдовипадкова величина ($0 < p < 1$), m – кількість чорних кліток в околі, C – константа ($C = 1,5$ при $m = 1$; $C = 1$ при $m \neq 1$);

–якщо клітина чорна, а навколо неї виключно чорні і сірі, то вона перефарбовується в сірий колір (новина застаріває, але зберігається як відомість);

–якщо клітина сіра, а навколо неї виключно сірі і чорні, то вона перефарбовується в білий колір (забування інформації при її загальновідомості).

Описана система клітинних автоматів цілком реалістично відбиває процес розвитку інформаційного сюжету (Рис. 1). Типові залежності кількості клітин (послідовності кількості однотипних клітин), що перебувають в різних станах, в залежності від кроку еволюції наведені на Рис. 2.

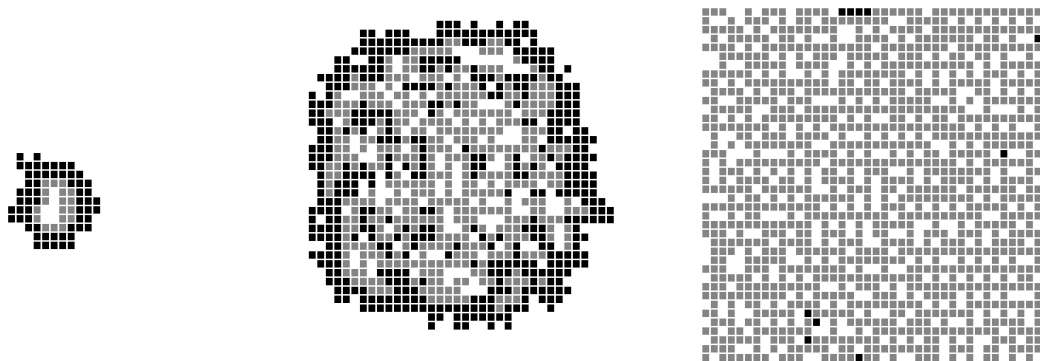


Рис. 1 – Стан еволюції системи клітинних автоматів.

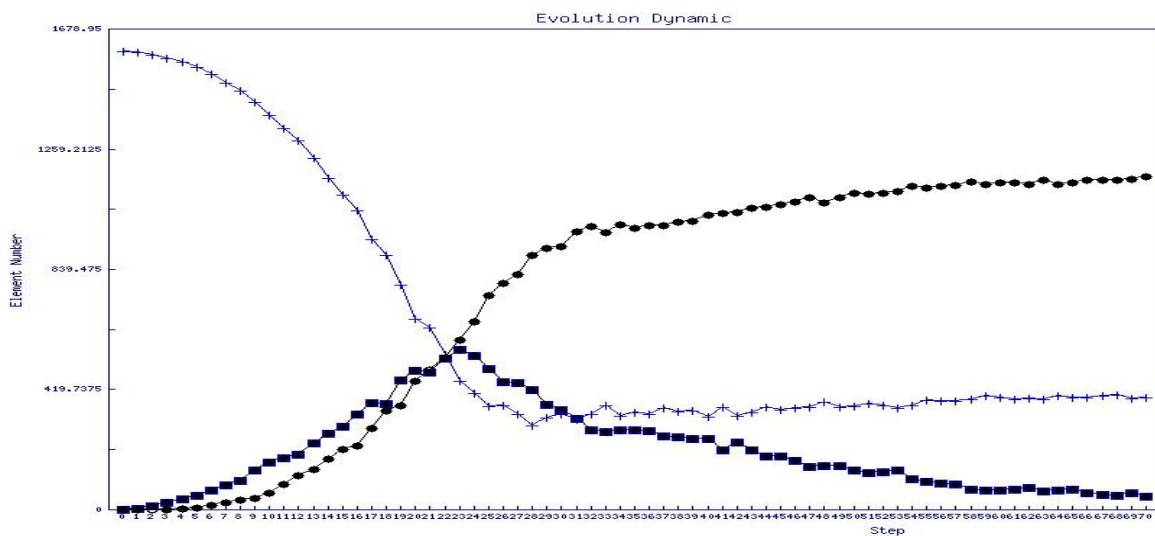


Рис. 2 – Розподіл клітин залежно від номера такту системи клітинних автоматів: білі клітини – (+); сірі клітини – (•); чорні клітини – (■).

При аналізі наведених графіків можна звернути увагу на такі особливості: 1 –сумарна кількість клітин, що перебувають в усіх трьох станах на кожному кроці ітерації постійна і дорівнює розміру поля; 2 – при стабілізації клітинних автоматів співвідношення кількості сірих, білих і чорних клітин приблизно становить: 0,75: 0,25: 0. Саме чорні клітини утворюють актуальний інформаційний сюжет, динаміка якого представлена на Рис. 1.

Цілком реалістичні профілі динаміки тематичних інформаційних потоків також були досягнуті за допомогою мультиагентної моделі [2]. У рамках цієї моделі окремі документи асоціюються з агентами, життєвий цикл агентів – з життєвим циклом документів в інформаційному просторі. Відповідно, простір агентів асоціювався з тематичним інформаційним потоком.

Передбачається, що протягом дискретних моментів часу відбувається еволюція популяції агентів. При цьому окремі агенти можуть:

- 1) самозароджуватися (народжуватися з причин, що виникають поза мультиагентного простору);
- 2) породжувати нових агентів;
- 3) “гинуть” – зникати з простору агентів;
- 4) отримувати посилення від інших агентів.

Кожен агент має “потенціал”, що залежить від його віку (часу життя на даний момент), авторитетності (посилань, проставлених на нього) і плодючості (кількості породжених безпосередньо ним агентів).

Управляючі параметри моделі наступні:

- 1) ймовірність “самозародження” P_1 ;
- 2) потенціал агента Pot , що залежить від кількості посилань на нього (ns), часу його життя (t) і кількості породжених ним агентів (k): $Pot = \frac{ns + k}{t}$;

3) ймовірність “народження” від існуючого: $P_2 \cdot Pot$;

4) ймовірність “загибелі” агента: P_3 / Pot ;

5) ймовірність посилення на агента: $P_4 \cdot Pot$.

Варіювання параметрами управління P_1 , P_2 , P_3 і P_4 дозволили змоделювати наведені на Рис. 2 профілі поведінки тематичними інформаційними потоками.

На Рис. 3 наведено приклад можливої динаміки мультиагентної системи: процеси народження нових агентів від існуючих позначені суцільними стрілками, процеси проставляння посилань на агентів представлені пунктирними стрілками, живі агенти – чорними кружками, “мертві” агенти до моменту $t = 5$ – незаповненими колами.

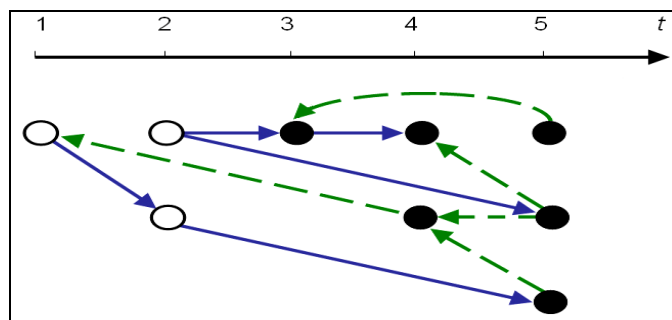


Рис. 3 – Фрагмент мультиагентного простору.

Слід зазначити, що дана модель не бере до уваги:

- 1) конкуренції агентів усередині агентного простору (передбачається тільки співпраця шляхом проставляння посилань і породження нових агентів);
- 2) конкуренції різних тематичних інформаційних потоків (враховується лише неявно, як причина, що обумовлює параметри функціонування даної мультиагентної системи).

Також слід зазначити, у запропонованій моделі враховується загальновідома практика проведення інформаційних кампаній в соціальних мережах, що полягає в реєстрації великого числа акаунтів-роботів (роїв), від імені яких проставляються посилання (лайки) на матеріали, що публікуються від імені акаунтів з того ж рою і на цільові інформаційні сторінки – документи.

В результаті проведених досліджень була реалізована програма еволюції простору агентів, досліджена еволюція мультиагентної системи при різних значеннях параметрів, знайдені аналогії з реальними тематичними інформаційними потоками, динаміка яких була визначена за допомогою системи InfoStream.

У разі інформаційних потоків, які асоціюються з конкретними тематичними інформаційними потоками, необхідно описувати динаміку кожного з таких потоків окремо, зважаючи на те, що зростання одного з них автоматично приводить до зменшення інших і навпаки. Тому обмеження на кількість повідомлень за усіма тематиками поширюється і на сукупність усіх тематичних сюжетів новин. У разі вивчення загального інформаційного потоку спостерігається явище “перетікання” публікацій з одних тематичних сюжетів, що втрачають актуальність, до інших.

Зазначимо, що запропонована модель дозволяє відрізнити інформаційні потоки, поведінка яких визначається природними закономірностями медійного простору, від потоків, висвітлення яких у медійних засобах має вплив зовнішніх чинників. Зокрема, таким індикатором може бути відхилення від характерних форм розподілу, поява періодичних зон нестабільності значень, відповідних динаміці тематичних інформаційних потоків, або, навпаки, суттєва локальна стабільність цих значень.

Як приклад, на Рис. 4 наведено динаміку публікацій в RUNet тематичних інформаційних потоків за запитами: “Банки Кипра”, “Офшор”, “Вирджинские острова” за березень-квітень 2013 року в період відомих кризових подій, яку отримано за допомогою системи InfoStream [6].

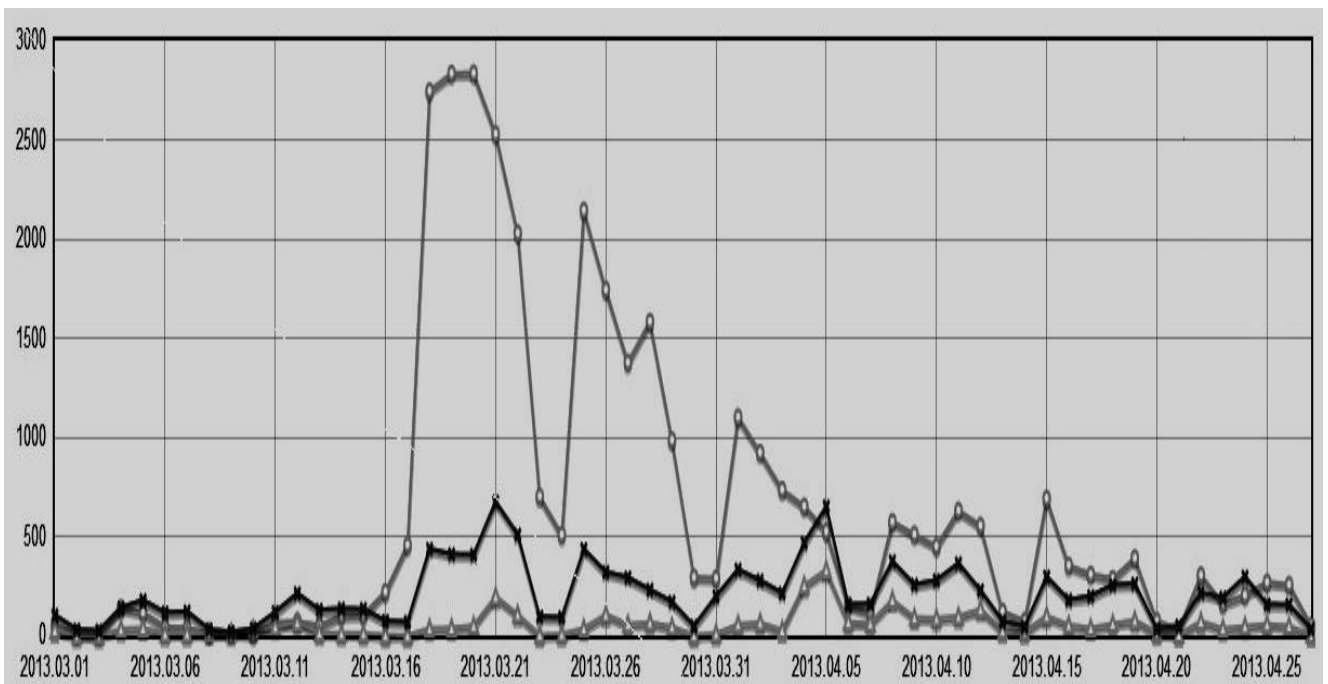


Рис. 4 – Діаграма динаміки тематичних інформаційних потоків за запитами:
о – “Банки Кипра”; Δ – “Вирджинские острова”; х – “Офшор”.

Як видно з Рис. 4, пік публікацій, пов’язаних з банківською кризою на Кіпрі припадає на 17 – 18 березня 2013 року, в той час, як більшість публікацій по островах Вірджинії з’явилися 4 – 5 квітня, коли там у значно менших масштабах, стали відбуватися події, подібні до кіпрських. При цьому слід зазначити слабку взаємну кореляцію динаміки інформаційних потоків, пов’язаних з Кіпром і островами Вірджинії. В цьому випадку коефіцієнт взаємної кореляції відповідних числових рядів становив всього 0,3. При цьому відзначається високий рівень взаємної кореляції рядів, відповідних тематикам “Офшор” і “Банки Кипра” (0,73), а також “Офшор” і “Вирджинские острова” (0,77).

Мабуть, прояви інформаційних операцій в області офшорних банків у даному випадку краще за все побачити при аналізі загальнішої тематики – “Офшори”. На графіці відповідного числового ряду чітко видно дві області локальних екстремумів, відповідних кризовим ситуаціям на Кіпрі і на островах Вірджинії. Можна зробити припущення, що якщо динаміка часткового інформаційного потоку в якийсь момент починає істотно відрізнятися від динаміки потоку, відповідного загальнішій тематиці (як в даному випадку – “Банки Кіпру” і “Офшор”), то можливий прояв ознак початку інформаційної операції, що відноситься до вузької тематики.

При проведенні вейвлет-аналізу (Рис. 5) було прийнято рішення щодо використання вейвлету “Мексиканський капелюх” як найбільш близького за формою до діаграми, відповідної тренду інформаційних операцій [6].

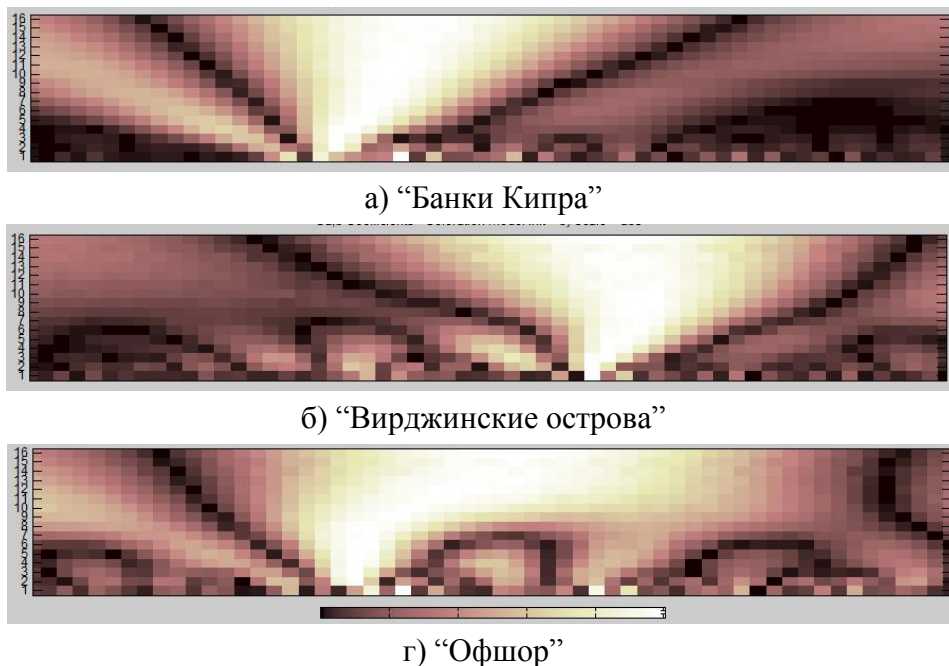


Рис. 5 – Вейвлет-спектограми, що відповідають динаміці тематичних інформаційних потоків за запитами.

Дані процеси є чітко видимими як на вейвлет-спектограмах, так і на відповідних їм скелетонах (графіках ліній екстремумів).

Живучість інформаційних об’єктів.

Поняття живучості інформаційної складової мережі Інтернет може розглядатися як здатність інформаційних об’єктів (новинних повідомлень, статей, документів, відеороликів і т. д.) своєчасно виконувати свої функції інформування в умовах дії

дестабілізуючих чинників. Такими чинниками можуть бути усунення окремих об’єктів з інформаційного простору, втрата інформаційними об’єктами властивостей актуальності, доступності [7, 8].

Оцінка живучості інформаційних об’єктів може проводитися на усіх етапах їх життєвого циклу. Існує декілька підходів до проведення оцінки живучості, що мають найбільш загальний характер. Живучість можна оцінити стосовно деякої стандартної зовнішньої дії або щодо множини зовнішніх дій. У цьому випадку вирішується завдання знаходження множини характеристичних векторів станів інформаційного об’єкта (у простому випадку – розподіл даних по серверах), в яких реалізується конфігурація, що забезпечує виконання мети функціонування. Потужність цієї великої кількості може служити мірою живучості усього інформаційного об’єкта.

При аналізі живучості інформаційних об’єктів виникає проблема інформування за різними аспектами незалежно від наявності або відсутності несприятливих чинників. У зв’язку з цим як кількісний критерій оцінки живучості доцільно використовувати відношення кількості функцій, що виконуються об’єктом за наявності певних несприятливих дій або множини таких дій, до загальної кількості функцій інформаційного об’єкта, з урахуванням критичності виконуваних і невиконуваних функцій. Критичність кожної конкретної функції визначається індивідуально для кожного конкретного інформаційного об’єкта виходячи з його специфіки. Кількісний показник живучості конкретного інформаційного об’єкта в заданих умовах можна обчислювати за формулою:

$$S = \sum_{i \in \Delta} \alpha_i / \sum_{j \in \Theta} \alpha_j,$$

де: Θ – множина всіх функцій інформування, Δ – множина функцій інформаційного об’єкта, що виконуються при заданих умовах ($\Delta \subseteq \Theta$), α_n – критичність n -ої функції.

Таким чином, кількісна оцінка живучості інформаційного об’єкта буде вимірюватися в інтервалі [0, 1], живучість тим вище, чим більша її кількісна оцінка.

Методи підвищення живучості інформаційної складової спрямовані на зменшення рівня її вразливості в мережах і системах інформаційної інфраструктури. Поняття живучості включає поняття надійності, безпеки, відмовостійкості тощо. Тому, зокрема, методи забезпечення живучості інформаційно-аналітичних систем (далі – ІАС) включають до свого складу також методи забезпечення цих характеристик, але не обмежуються ними. Нижче наведено основні методи, які застосовуються для підвищення живучості інформаційної складової [9]:

1. Регулярна перебудова метаданих та індексів інформаційних ресурсів ІАС. У системі організації інформаційних ресурсів одним з ключових чинників є метадані (класифікації, переліки об’єктів, термінологічні словники, тезауруси, уніфіковані форми представлення даних, стандарти, патенти й інші форми нормативних і правових документів).

2. Багатократне дублювання даних (реалізація надмірності даних). Надмірність вводиться штучно при проектуванні баз даних у цілях підвищення надійності системи в умовах роботи із збоями. При цьому передбачається регулярне здійснення реплікації дублюючих блоків інформації з перевіркою ідентичності.

3. Реалізація резервного та архівного копіювання інформаційних ресурсів ІАС.

4. Застосування децентралізованих систем зберігання інформації разом із дублюванням критичних блоків інформації.

5. Резервне дублювання не тільки інформації, а й цілих апаратно-програмних комплексів.
6. Застосування сертифікованого антивірусного захисту.
7. Застосування засобів електронного цифрового підпису, хешування даних (як засіб проти спотворення даних).
8. Постійна перевірка валідності джерел інформації.
9. Видалення (архівування) зайвої, непотрібної інформації, інформаційного шуму.
10. При здійсненні процедур обробки даних з можливою втратою інформації (автоматичне реферування, переклад, JPEG-перетворення зображень) необхідно зберігати першоджерела документів в інформаційному сховищі (депозитарії).
11. Забезпечення доступу аналітиків (аналітичних програмних модулів) до будь-якого фрагмента інформаційного репозитарію ІАС з метою оперативного виявлення впливу на інформаційну складову, здійснення регулярного цілеспрямованого аналізу, контролю і коригування стану програмного та інформаційного забезпечення.
12. Ранжирування інформації. Інформація, що зберігається в ІАС, має бути ранжирувана за ступенем важливості, рівнем конфіденційності та належністю відповідним підрозділам.
13. Розподіл прав доступу до інформації. Реалізація визнаних політик безпеки.
14. Фільтрація вихідної інформації, яка поступає з ІАС безпосередньо для прийняття рішень, шляхом цензурування з метою відсіювання несприятливої для зовнішнього світу інформації.
15. Розподіл частин ІАС (“відкритої”, корпоративної тощо) можливо реалізувати на різних рівнях (паролі, файрволи, фізичне роз’єднання).
16. Обмеження доступу до інформації. Цей метод припускає наявність розвиненого периметра безпеки, що має як централізоване, так і децентралізоване керування, який був би в змозі оцінювати вхідну інформацію (“службову” і “корисну”), і виключати інформацію небезпечного змісту.
17. Протоколювання подій в системі, ведення системних журналів з метою виявлення можливих фрагментів даних, які зазнали втручання.
18. Забезпечення працездатності апаратних засобів, коректне відключення носіїв даних при зупинці системи, окремих блоків і модулів.
19. Зберігання в системі різних версій програм обробки/візуалізації, що допоможе уникнути помилок при реалізації принципу “наслідування”.
20. Наявність конверторів форматів даних і підсистем інтеграції ПЗ (підсистеми імпорту, експорту і синхронізації даних між різними додатками) для забезпечення використання вихідних даних у різних форматах.
21. Застосування засобів термінового зберігання критично важливих документів, доступних у режимі “тільки читання”. Резервне зберігання інформації на знімних носіях, розміщення в захищених приміщеннях. Задовольнити нові вимоги можна, використовуючи декілька підходів.
22. Зберігання історії (версій) інформаційних документів. Майже всі сучасні документальні сховища підтримують версійність інформаційних документів.
23. Використання надійних каналів передачі вхідних даних. Безпека має забезпечуватися шифруванням трафіку між серверами додатків і клієнтами. Використання декількох каналів інформації з подальшим порівнянням при занесенні в інформаційне сховище (реплікації).
24. Застосування систем виявлення зовнішнього впливу (вторгнення), за допомогою яких можна зафіксувати факт атак на інформаційну інфраструктуру, оцінити

можливі збитки і виконати адекватні дії у відповідь. При цьому зовнішні впливи на інформаційну складову нині прийнято вважати інформаційними операціями [6], відповідно для забезпечення її живучості необхідно застосовувати методи моніторингу і протидії інформаційним операціям.

Висновки.

Забезпечення живучості аналітичної складової охоплює усі ланки життєвого циклу інформаційних об'єктів, а саме:

- отримання аналітиками вихідної інформації;
- аналіз інформації за визначеною проблемою, що зібрана;
- обробка інформації;
- підготовка документів (інформаційних об'єктів);
- верифікація інформаційних об'єктів;
- використання інформаційних об'єктів.

Наведені вище моделі і методи придатні для опису загальних тенденцій динаміки розвитку інформаційних процесів, однак проблема прогнозування залишається відкритою. Мабуть, більш реалістичні моделі можуть бути отримані з урахуванням додаткового набору чинників, більшість яких не відтворюються в часі.

Використана література

1. Ланде Д.В. Інформаційні потоки в глобальних комп'ютерних мережах / О.Г. Додонов, Д.В. Ланде, В.Г. Путятін. – К : Наукова думка, 2009, – 295 с.
2. Додонов А.Г., Ландэ Д.В. Мультиагентная модель поведения тематических информационных потоков : материалы VI Всероссийской мультиконференции по проблемам управления (30 сентября – 5 октября 2013 г.). – Том. 4. – Ростов-на-Дону : Издательство Южного федерального университета, 2013. – С. 102-107.
3. Ланде Д.В. Основи інформаційного і соціально-правового моделювання : монографія / Д.В. Ланде, В.М. Фурашев. – К. : ТОВ “ПанТот”, 2012. – 144 с.
4. Ландэ Д.В. Интернетика : Навигация в сложных сетях : модели и алгоритмы / Д.В. Ландэ, А.А. Снарский, И.В. Безсуднов. – М. : Либроком (Editorial URSS), 2009. – 264 с.
5. Ландэ Д.В. Модель диффузии информации / Д.В. Ландэ : сб. науч. тр. Ин-та проблем регистрации информации НАН Украины [Информационные технологии и безопасность. Менеджмент информационной безопасности]. – 2007. – Вып. 10. – С. 51-67.
6. Додонов А.Г., Ландэ Д.В. Динамика информационных потоков при выявлении информационных операций : материалы XIII Международной научно-практической конференции [“Информационная безопасность”]. – Ч. 1. – Таганрог : Изд-во ТТИ ЮФИ, 2013. – С. 42-49.
7. Ландэ Д.В. Живучесть информационных систем / А.Г. Додонов, Д.В. Ландэ. – К. : Наук. думка, 2011. – 256 с.
8. Knight J.C., Strunk E.A., Sullivan K.J. Towards a Rigorous Definition of Information System Survivability // Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX'03), 2003.
9. Додонов О.Г., Ланде Д.В. Методи підвищення живучості інформаційної складової корпоративних інформаційно-аналітичних систем підтримки прийняття рішень // Реєстрація, зберігання і обробка даних. – 2012. – № 2-14. – С. 48-58.

~~~~~ \* \* \* ~~~~~