УДК 004.62

З.М. Пузняк, Д.А. Шеремет

МЕТОДИКА ВИЯВЛЕННЯ ВПЛИВУ НА ДОСТОВІРНІСТЬ ІНФОРМАЦІЇ В ІНФОРМАЦІЙНОМУ ПРОСТОРІ

Розглянуто класифікацію інформаційних загроз, визначено критерії оцінки достовірності інформації в інформаційному просторі. Також запропонована методика виявлення впливу на достовірність інформації в інформаційному просторі на основі інформаційно-орієнтованої моделі.

Ключові слова: вплив на достовірність інформації, достовірність інформації, інформаційна загроза, інформаційний простір.

Вступ

Виходячи з реалій сьогодення ефективність інформаційних ресурсів держави залежить не тільки від надійності функціонування інформаційно - телекомунікаційних систем, а й у значній мірі від захищеності її інформаційних ресурсів [1]. Однією з проблем, яка стримує впровадження ефективних систем захисту державних інформаційних ресурсів, є проблема створення достовірної класифікації атак. Зважаючи на це, підвищення ефективності виявлення атак на державні інформаційні ресурси, залишається актуальним завданням. Систематизація знань про атаки допомагає розробці заходів і систем захисту від них.

Основна частина

Під загрозою інформаційних ресурсів держави (ІРД) можна розуміти протиправні дії, які можуть призвести до спотворення, несанкціонованого використання або руйнування державних інформаційних ресурсів (безпосередньо їх властивостей: конфіденційність, цілісність, доступність), які є власністю держави та необхідність захисту яких визначено законодавством.

Загрози в інформаційній сфері можна розрізняти за такими ознаками: джерелами, об’єктами, засобами, методами та наслідками [8].

При аналізі класифікації інформаційних загроз, а також визначення інформаційної загрози, необхідно також мати визначення взаємно залежному з нею поняттю: атака.

Захист конфіденційності та цілісності інформації забезпечується законами України та нормативно-правовими документами: Закони України «Про інформацію», «Про захист персональних даних», НД ТЗІ 1.1-002-99 «Загальні положення щодо захисту інформації в комп’ютерних системах від несанкціонованого доступу» тощо. Про те відсутні правові норми та регламенти захисту достовірної інформації.

Дії щодо поширення в інформаційному просторі викривленої, недостовірної та упередженої інформації або негативні інформаційні впливи на суспільну свідомість в ІП відносяться до порушення цілісності інформації. Однак, слід сказати, що фактично цілісність інформації не змінюється, а змінюється відношення до інформації. Тобто відбувається вплив на достовірність інформації.

В результаті ІДР підлягають інформаційному ризику. Зазвичай при дослідженні ризиків не надається увага оцінці якості інформації, за допомогою якої оцінюється ризик. Складові оцінки якості інформації наступні:

- достовірність інформації – міра наближеності інформації до першоджерела;
- об’єктивність інформації – міра віддзеркалення інформацією реальності;
- однозначність;
- порядок інформації – кількість передавальних ланок між першоджерелом і кінцевим користувачем;
- повнота інформації – віддзеркалення вичерпного характеру відповідності одержаних відомостей цілям збору;
- корельованість – ступінь відповідності інформації поставленому завданню;
- актуальність інформації (значущість) – важливість інформації;
- вартість інформації.

До факторів інформаційного ризику слід віднести:

- неповну або недостовірну інформацію у ЗМІ;
- надмірний обсяг і хаотичність інформації, яка циркулює в суспільстві;
- штучне або природне перекручування інформації в процесі її відбору для аналізу й використання у ході прийняття відповідальних державних рішень;
- технологічні помилки щодо введення, збереження й обробки даних в інформаційних системах.

В результаті здійснення даних факторів, достовірність інформації підлягає сумніву.

Виявлення впливу на достовірність інформації в ІП ґрунтується на наступних критеріях [4, 6, 9]:

- правдоподібність викладених фактів, що визначається приховуванням джерел і авторів інформації, недостатньою аргументацією, посиланнями на думку широкого загалу, наявністю риторичних запитань;

1) Посилання на суб'єктивну точку зору F1 - відносний показник вживання у контенті оцінок фактів експертами, ученими, авторитетними джерелами тощо

$$F_1 = \frac{R_d}{W} \tag{1.1}$$

де Rd – кількість використаних цитат з основного джерела, W–кількість статей на дану смислову тему;

2) F2 – відносний показник використання лінгвістичних конструкцій, які відкидають необхідність підтвердження і доведення істинності контенту (наприклад, вочевидь, незаперечний факт тощо)

$$\mathrm{F_2} = \frac{\mathrm{G_d}}{\mathrm{W}} \tag{1.2}$$

де Gd – кількість лінгвістичних конструкцій із запереченням необхідності верифікації контенту;

3) Частка запитальних речень F3 – відношення кількості запитальних речень Sokl до загальної кількості речень Sz у текстовому контенті

$$\mathrm{F_3} = \frac{\mathrm{S_{okl}}}{\mathrm{S_z}} \tag{1.3}$$

4) Показник підготовки до негативного сприйняття F4–відношення кількості негативних висловлювань Bzдо загальної кількості статей на дану смислову тему W;

$$\mathrm{F_4} = \frac{\mathrm{B_z}}{\mathrm{W}} \tag{1.4}$$

5) Сумнівні висловлювання F5 – відносний показник використання лінгвістичних конструкцій, які припускають різні підходи до тлумачення (наприклад, можливо, ймовірно, завжди)

$$\mathrm{F_5} = \frac{\mathrm{F_z}}{\mathrm{W}} \tag{1.5}$$

де Fz – кількість неоднозначних висловлювань;

- емоційне забарвлення контенту, що використовується для відображення емоційного стану автора і проявляється у перенасиченні контенту образними засобами, прикметниками, порівняннями тощо

6) Окличні речення F6–відносний показник кількості окличних речень Sd в текстовому контенті

$$F_6 = \frac{S_d}{S_z} \quad (1.6)$$

7) Вигуки F7 – показник наявновності у текстовому контенті вигуків (наприклад, ну-ну, овва, ага тощо)

$$F_7 = \frac{E_d}{W} \quad (1.7)$$

де Ed– кількість виявлених вигуків у публікації;

8) Прислівники F8 – відносна кількість прислівників Ad у текстовому контекті, які використовуються для порівняння, пере фокусування читача публікації на його емоції (наприклад, наче, більше, назавжди тощо)

$$F_8 = \frac{A_d}{A_z} \quad (1.8)$$

де Az – загальна кількість прислівників у публікації;

9) Емоційний словник F9 – показник вживання у публікації лексем емоційного характеру (наприклад, безкарний, блокада, ганебний тощо)

$$F_9 = \frac{V_d}{W} \quad (1.9)$$

де Vd – кількість емоційних лексем;

- тональність контенту по відношенню до деякого об’єкту чи події, яка відображає оцінювальні судження автора і може проявлятися у використанні зображень тощо.

Метою є визначення позиції автора відносно досліджуваних об’єктів або подій. Задача оцінки тональності контенту розв’язується шляхом застосування методів машинного навчання та інформаційного пошуку. Даний крок [5, 7] зводиться до віднесення тональності публікації до попередньо визначеної категорії - негативна, позитивна, нейтральна тощо

Таблиця 1.1

Приклад нормованої шкали оцінки тональності

Клас тональності, який використовується	Інтервал нормованої шкали оцінок	Клас тональності, який використовується	Інтервал нормованої шкали оцінок
Негативна		Позитивна	
Виражено негативна	-(1,00-0,70)	Виражено позитивна	1,00-0,70
Помірно негативна	-(0,71-0,50)	Помірно позитивна	0,71-0,50
Нейтральна	-(0,51-0,40)	Нейтральна	0,51-0,40
Помірно позитивна	-(0,41-0,20)	Помірно негативна	0,41-0,20
Виражено позитивна	-(0,21-0,00)	Виражено негативна	0,21-0,00

- сенсаційність контенту, яка має на меті привернути увагу завдяки підвищенню тривожності та ін.

За даним критерієм оцінюється здатність текстового контенту своїм змістом зацікавити, вразити і привернути увагу. Критерій зводиться до виявлення таких ознак:

10) Приверенення уваги F10–відносний показник вживання слів, що привертають

увагу індивіда, підвищують тривожність та створюють атмосферу швидкоплинності (наприклад, вбивство, шок, сепаратизм; миттєво, швидко, несподівано)

$$F_{10} = \frac{U_d \times O_d}{W} \quad (1.10)$$

де U_d – кількість виявлених слів-ідентифікаторів уваги,

O_d – кількість виявлених слів для позначення оперативності;

- прихований (імпліцитний) зміст контенту пов’язаний з його глибинним змістом.

Метою даного критерію є виявлення прихованої теми повідомлення в результаті тематичного моделювання. Для визначення такого змісту використаємо приховане розміщення Діріхле, що належить до породжуючих моделей, які дозволяють побудувати речення відповідно до правил заданої мови. Публікація розглядається як набір різних тем, апріорно розподілених за діріхле. Метод ефективний для опису кластер них структур, спрощує вивід апостеріорних ймовірностей публікацій і їх тем. Недоліком є відсутність лінгвістичних обґрунтувань методу та можливість перенавчання моделі.

$$p(d, w) = \sum_{t \in T} p(d) \times p(w|t) \times p(t|d) \quad (1.11)$$

В узагальненому вигляді зв'язок між частинними ознаками виявлення впливу на достовірність інформації зобразимо у вигляді ієрархії (рис. 1.1).

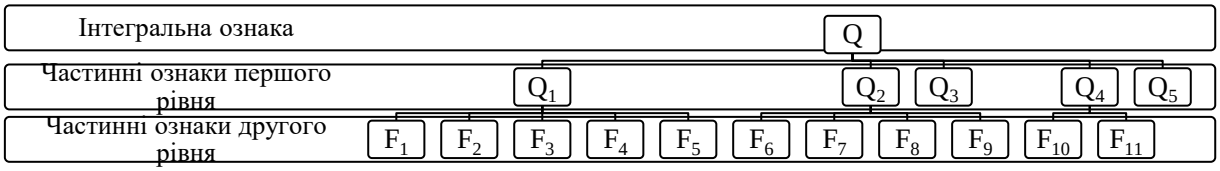

Рис. 1.1. Дерево рішень

Тоді критерій виявлення компрометації у текстовому контенті запишемо у вигляді нерівності

$$H = -\sum_{v=1}^{k} \sum_{l=1}^{g} Q_l^v \, log_2 Q_l^v \quad (1.16)$$

де H – граничне значення інформаційної ентропії (невизначеності);

Q_l^v - числове значення прояву ознаки впливу на достовірність інформації;

$l = \overline{1, g}$ - індекси частинних ознак впливу на достовірність інформації другого рівня;

$v = \overline{1, k}$ - індекси частинних ознак впливу на достовірність інформації першого рівня.

Для зручності інтерпретації розрахованих значень введемо нормоване значення ентропії H_n

$$H_n = \frac{H_{max} - H}{H_{max}} \quad (1.17)$$

де $H_{max} = \log N$- максимальне значення ентропії.

Таким чином, зміст критерію виявлення недостовірності інформації зводиться до оцінки інформаційної ентропії текстового контенту, тобто встановлення рівня невизначеності щодо наявності у контенті прихованого впливу на достовірність інформації

та порівняння його числового значення із допустимим граничним. Інформаційна ентропія [8] зменшується при зростанні частот появи ознак впливу на достовірність інформації в ІП інформаційна невизначеність зростає. Якісна шкала оцінки загроз на достовірність інформації сформована в результаті обчислювального експерименту та узагальнення і адаптації підходів до оцінки загроз у галузі інформаційної безпеки (табл. 1.2) [2, 3].

Таблиця 1.2

Адаптована інтервальна шкала

Клас загрози	Інтервальні значення нормованої ентропії H_n
Дуже низький	0,00 – 0,20
Низький	0,21 – 0,49
Значний	0,50 – 0,74
Високий	0,75 – 0,90
Дуже високий	0,91 – 1,00

Для того, щоб описати методику виявлення впливу на достовірність інформації в інформаційному просторі, розглянемо інформаційно-орієнтовану модель. Для початку розглянемо – стандартну її модель. Простором у ній є двовимірна сітка з рівними клітками – квадратами. У кожен момент часу t існує постійне кінцеве число інформації, розташованих у просторі. У момент часу t кожна клітка (x, y) може містити інформацію $a(a_t (x, y) = a)$, тобто інформація a перебуває в клітці (x, y), або не містити інформації $a(a_t (x, y) = ø)$. Кількість інформації у клітці (x, y) в момент часу (x, y) становить $r_t (x, y)$.

Інформація з'являється на поле з двома параметрами: достовірність (кількість кліток у клітках, що він може бачити) і значення ентропії (кількість інформації в певний момент часу). Інформація може бути актуальною і неактуальною. Якщо інформація немає попиту в інформаційному просторі, то вона зникає. Правила поведінки інформації наступні:

- вивчається окіл бачення інформації (4 або 8 напрямів кліток) і визначається вільна клітка, що має найбільший попит;
- після цього інформація переміщається у цю клітку та стає актуальною.
- На основі інформаційно-орієнтованої моделі отримуємо дані, що цілком відповідають звичайній поведінці інформації у інформаційному просторі. З цією метою в модель вводиться рівень загрози як результат отримання та сприйняття інформації. Тобто в цьому випадку кожна клітинка містить попит на інформацію і деякий клас загрози. За новими правилами інформація пересувається у вільну клітку, де співвідношення (попит/клас загрози) максимальне.

Висновки

Існує необхідність у створенні методики виявлення впливу на достовірність інформації в інформаційному просторі.Завдяки такій методиціз'являється можливість досліджувати інформаційні ресурси на наявність впливу на них, своєчасно виявляти проблеми ІБ в використовуваних програмних і апаратних засобах, виробляти рекомендації щодо їх усунення, інформувати відповідні підприємства, рекомендувати перевірені рішення до використання на критично важливих об'єктах, організовувати і проводити спеціальні навчання на державному рівні і т.п.

Список використаної літератури

1. Бурячок В.Л. Політика інформаційної безпеки: підручник / В.Л. Бурячок, Р.В. Грищук, В.О. Хорошко; під заг. ред. В.О. Хорошка. – К.: ПВП «Задруга», 2014. – 222 с.

2. Грищук Р.В. Метод оптимізації рівномірності потоку вхідних даних для систем захисту інформації / Р.В. Грищук, В.М. Мамарев // Інформаційна безпека. – 2012. - № 2 (8). – С. 27-34

3. Грищук Р.В. Основи кібернетичної безпеки: моногр. / Р.В. Грищук, Ю.Г. Даник; під заг. ред. проф. Ю.Г. Даника. – Житомир: ЖНАЕУ, 2016. – 636 с.

4. Жарков Я.М. Інформаційно-психологічне протиборство (еволюція і сучасність): моногр. / Я.М. Жарков, В.М. Петрик, М.М. Присяжнюк, Є.Д. Скулиш, Л.Ф. Компанцева: Київський національний університет ім. Т. Шевченка. – К.: Віпол, 2013. – 247 с.

5. Жураковський Ю.П. Теорія інформації та кодування: підр. / Ю.П. Жураковський, В.П. Полторак. – К.: Вища школа, 2001. – 255 с.

6. Компанцева Л.Ф. Интернет-лингвистика: коммуникативно-прагматический и лингво культурологический подходы: моногр. / Л.Ф. Компанцева. – Луганск: Знание, 2008. – 528 с.

7. Ланде Д.В. Интернетика: навигация в сложных сетях: методы и алгоритмы / Д.В. Ланде, А.А. Снарский, И.В. Безсудов. – М.: Книжный дом «ЛИБРОКОМ», 2009. – 264 с.

8. Молодецька-Гринчук К.В. Методика виявлення маніпуляцій суспільною думкою у соціальних інтернет-сервісах / К.В. Молодецька-Гринчук // Інформаційна безпека. – 2016. - №3(23). – С.80-92

9. Морозов А.М. Психологическая война / А.М. Морозов. – К.: «Логос», 1996. – 140 с.

Надійшла 05.07.2017 р. Рецензент: д.т.н., проф. Дудикевич В.Б.