2010. – 1216 p.

10. Морозов А.А. Ситуационные центры – основа стратегического управления / А.А. Морозов, В.А. Ященко // Математичні машини і системи. – 2003. – № 1. – С. 3 – 14.

11. Экспертные системы. Принципы работы и примеры / А. Брукинг, П. Джонс, Ф. Кокс и др.; под ред. Р. Форсайта – М.: Радио и связь, 1987. – 224с.

12. Patent No.: US 6266579 Bl. System for reducing disaster damage / Mohammad Reza Baraty. – № 09/022.667; заявл. February 12, 1998; опубл. July 24, 2001.

Надійшла: 07.07.2012 р.

Рецензент: д.т.н., професор Хорошко В.О.

УДК 004.056.53

Баранов Г.Л., Захарова М.В., Горніцька Д.А.

МЕТОДОЛОГІЯ СИНТЕЗУ СИСТЕМ ОЦІНКИ РІВНЯ ЗАХИЩЕНОСТІ ДЕРЖАВНИХ ІНФОРМАЦІЙНИХ РЕСУРСІВ ВІД СОЦІОТЕХНІЧНИХ АТАК

У роботі представлено методологію синтезу систем аналізу та оцінки рівня захищеності державних інформаційних ресурсів від соціотехнічних атак, які у наш час становлять одну з найбільших загроз і суттєво впливають на загальний рівень інформаційної безпеки. Розроблена методологія є гнучким інструментом та дає можливість здійснювати оцінювання рівня підготовленості персоналу до соціотехнічних атак на різних за специфікою роботи, управління, технічною базою підприємствах. Перевагами запропонованої методології є застосування такого параметру, як якість експерта, з метою підвищення якості експертного оцінювання загороз, та використання логіко-лінгвістичного підходу і математичного апарату нечіткої логіки, що дає можливість формалізувати оцінку ризиків.

Ключові слова: методологія синтезу систем аналізу та оцінки ризику, рівень захищеності, інформаційна безпека, соціотехнічні атаки, ризик, оцінка ризику.

Розвиток інформаційного суспільства розширив можливості інформаційного обміну, що в свою чергу дало поштовх удосконаленню існуючих та розвитку нових методів атак на державні інформаційні ресурси (ДІР), під якими слід розуміти взаємопов'язану, впорядковану, систематизовану, втілену на матеріальних носіях інформацію, створену або зібрану на законних підставах органами державної влади або іншими суб'єктами за рахунок державного бюджету [1].

В останні роки набули розвитку методи соціального інжинірингу, які відносяться до соціотехнічних атак (СА) [2-5]. Всі атаки даного класу засновані на переконані персоналу в санкціонованості дій атакуючих, які видають себе за авторизованих співробітників, керівництво тощо [6]. Основною формою СА є запит, який не потребує складних алгоритмів підготовки, та отримання відповіді від атакованого персоналу, засновуючись на психологічних принципах [4]. Найчастіше виказується інформація, яку персонал вважає неважливою, проте за її допомогою в подальшому ДІР можуть бути скомпроментовані [6]. Таким чином можна зробити висновок, що саме персонал є найбільш уразливою ланкою, якою неможливо нехтувати при оцінці стану інформаційної безпеки (ІБ) ДІР. Отже, рівень підготовленості персоналу протистояти СА є визначним чинником, який впливає на ІБ. Важливою є можливість автоматизувати процес оцінки базуючись на методологічних засадах.

У зв'язку з цим розробка методології синтезу систем оцінки рівня підготовленості персоналу протидії атакам даного класу є актуальним питанням захисту ДІР, вирішення якого є метою даної роботи.

При вирішенні завдань визначення стану ІБ інформаційних систем та розробки методів прийняття рішень використовується логіко-лінгвістичний підхід на базі теорії нечітких множин, що дозволяє формалізувати розмиті поняття та вирішити проблему математичної обробки нечіткої інформації [7].

В основі цього підходу лежить поняття лінгвістичної змінної (ЛЗ), яка є зручним засобом опису складних систем що містять параметри, подані не тільки в кількісному, але і у

якісному вигляді. При цьому ЛЗ дозволяють поставити у відповідність якісним значенням певну кількісну інтерпретацію і таким чином формалізувати їх.

Було запропоновано узагальнену послідовність вимірювання рівня безпеки, яку доцільно використовувати у якості базової для синтезу систем оцінки рівня захищеності, в тому числі і захищеності ДІР від СА [7].

Методологія містить такі етапи: визначення характеристик ІБ, аналіз загроз, визначення базового експертного запиту, ранжирування вхідних даних, формування лінгвістичних термів, вибір методу обробки нечітких чисел, вибір нечіткої моделі (НМ), обчислення та інтерпретація рівня ІБ [7].

Визначення рівня захищеності згідно з НМ визначається за наслідками відповідей користувачів системи на складений експертами запит. Складові зазначеного запиту заздалегідь ранжируються експертами шляхом визначення коефіцієнтів важливості (КВ) P_j ($j=\overline{1,n}$, де n – кількість складових). Експертна оцінка та ранжирування складових за ступенем небезпеки є одним з найголовніших моментів проведення експертизи захищеності ДІР від СА. При цьому ключову роль відіграють вибір шкали оцінювання [4], методу визначення КВ та урахування такого параметру як якість експерта (ЯЕ), тому означені чинники мають бути відображені у розроблюваній методології синтезу систем оцінки рівня захищеності ДІР (рис. 1).

Розроблена методологія включає одинадцять етапів: 1) оцінка ЯЕ; 2) визначення ДІР; 3) аналіз СА; 4) визначення подій ІБ; 5) визначення базового експертного запиту; 6) вибір шкали обчислень; 7) формування еталонних рівнів захищенності; 8) вибір методу та визначення КВ [9,10]; 9) інтерпретація результатів запитів; 10) визначення рівня захищеності ДІР. Питання урахування ЯЕ під час проведення експертиз у сфері ІБ є одним з базових. Методи оцінки ЯЕ ґрунтовно аналізувались в роботах [9-11] і з проведеного дослідження випливає той факт, що апріорна оцінка ЯЕ з практичної точки зору є більш зручною, оскільки не потребує розробки додаткових тестів або повторного опитування експертів з метою виявити з часом відхилення їх суджень. Ці методи дозволяють сформувати експертну групу (ЕГ) з фахівців, якість яких заздалегідь є достатньою, тобто відповідає певному рівню.

1. Оцінка ЯЕ. Згідно із розробленою методологією (рис. 1) безпосередній оцінці рівня захищеності ДІР від СА передує формування ЕГ з фахівців достатнього рівня якості [2]. Модель оцінки ЯЕ та формування ЕГ детально описана в [11]. Дії, що відбуваються на даному етапі, полягають у визначенні пріоритетних критеріїв вибору методу апріорної оцінки ЯЕ з-поміж наступних: тривалість підготовки ($t_{пк}$), тривалість проведення ($t_{пр}$), об'єктивність (b) та кількість експертів в ЕГ (G). Після цього з множини методів $M_1 ... M_S$ обирається метод M_S', за яким визначається ЯЕ та формується ЕГ ($E=E_1 ... E_G$).

2. Визначення ДІР. На даному етапі аналізу та оцінки ризиків (АОР) відібрана ЕГ визначає комплекс ДІР, які повинні бути захищенні. Для цього з переліку можливих інформаційних ресурсів, які можуть міститись в базі даних інформаційних ресурсів (БДІР) експерти відбирають ті R_t $t=\overline{1,x}$, де t – індекс ДІР, а x – кількість ДІР, які характерні в конкретному випадку. Наприклад, можна отримати наступний перелік ДІР: R_1="Веб-сайт", R_2="Поштовий сервер", R_3="Сервер баз даних" і т.д.

3. Аналіз СА. Далі ЕГ проводить аналіз всіх можливих СА відносно ДІР, визначених на попередньому етапі. Із множини всіх можливих СА, які доцільно зберігати у вигляді бази даних соціотехнічних атак (БДСА) експерти виділяють підмножину SA_a $a=\overline{1,i}$ a – індекс; i – кількість СА, які можуть бути реалізованими відносно вже визначених ДІР. Ці СА будуть слугувати вхідними даними (ВД) для побудови експертного запиту. Наприклад, для визначеного ДІР R_1="Веб-сайт" можливо ідентифікувати наступні SA_a $a=\overline{1,3}$: SA_1="Повідомлення пароля адміністратора сайту несанкціонованій особі", SA_2="Переконання адміністратора у внесенні несанкціонованих змін у сайт", SA_3="Відновлення паролю адміністратора на електронну пошту загального користування" і т.д.

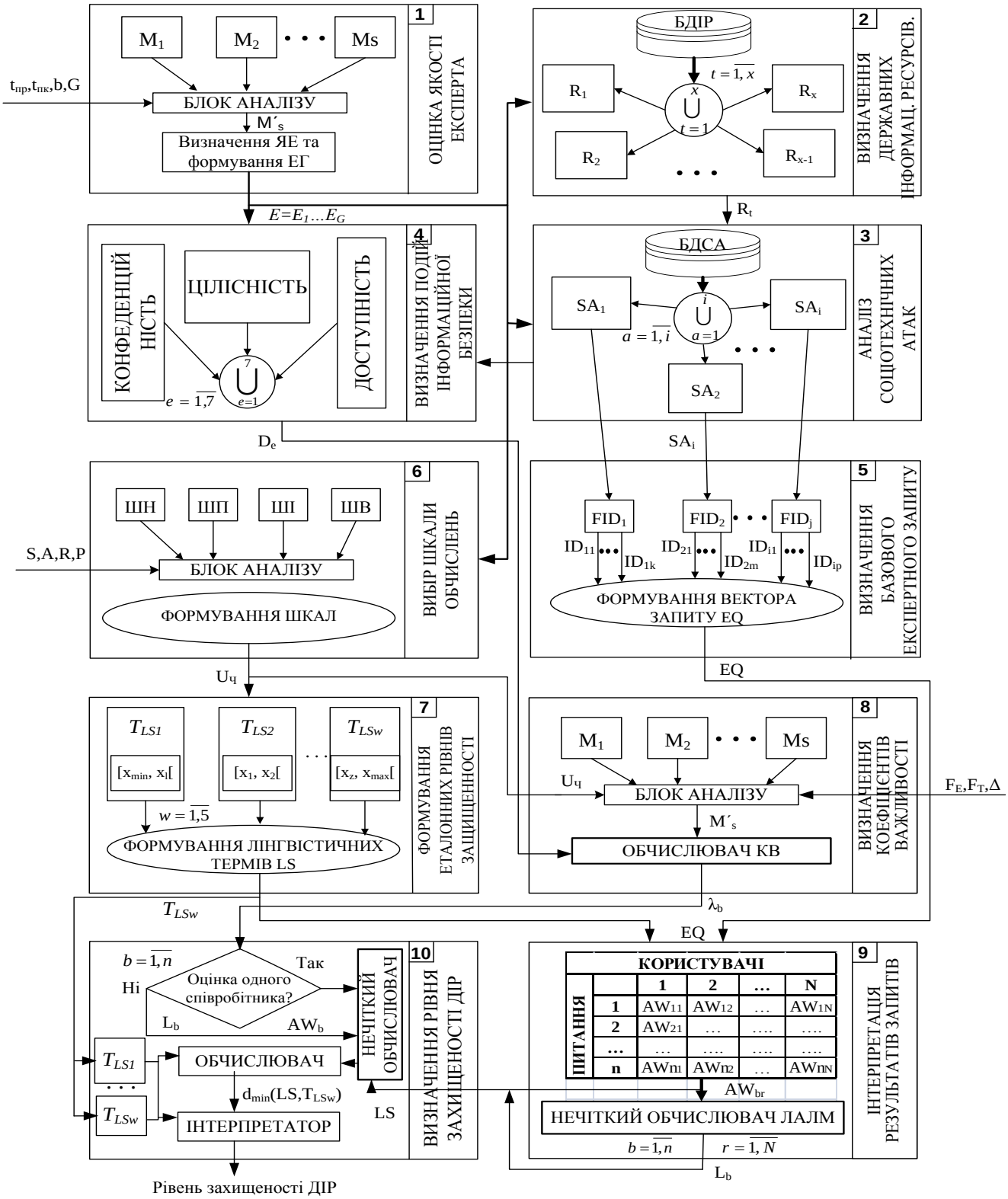

Рис. 1 Методологія синтезу систем аналізу та оцінки рівня захищеності ДІР від СА

4. Визначення подій ІБ. Тут ЕГ має визначити, на які саме характеристики безпеки – конфіденційність (К), цілісність (Ц), доступність (Д), визначені в [12], – впливає реалізація СА, визначених на попередньому етапі. При цьому визначаються базові події ІБ D_e, де $e = \overline{1,7}$ – поточний індекс події: D_1="Порушення конфіденційності (ПК)", D_2="ПЦ", D_3="ПД", D_4="ПКЦ", D_5="ПЦД", D_6="ПКД" і D_7="ПКЦД". Результатом даного етапу є набори даних R_t, SA_i, D_e. Наприклад, для R_1="Веб-сайт" визначені SA_1, SA_2, SA_3, які відповідно можуть призвести до D_7="ПКЦД", D_2="ПЦ" та D_7="ПКЦД" відповідно. Ці дані будуть застосовані ЕГ на етапі 8 при ранжируванні компонентів експертного запиту.

5. Визначення базового експертного запиту. При визначені базового експертного запиту з формувачів вхідних даних (FID_j) ($j=\overline{1,i}$) надходять масиви даних $ID_{11}...ID_{1k}$, $ID_{21}...ID_{2m}$, $ID_{i1}...ID_{ip}$, де i – кількість загроз (СА), визначених раніше, а k, m, p – кількість складових запиту першої, другої та i-ї СА відповідно, які слугують компонентами вектору запиту. Під цим слід розуміти окреме питання тесту, наприклад: ,,Як часто змінюються паролі?”. Таким чином, на виході блоку маємо базовий експертний запит

$$EQ=\begin{bmatrix} ID_{11} & ... & ID_{1k} \\ ... & ... & ... \\ ID_{i1} & ... & ID_{1p} \end{bmatrix},$$ що складається з $n=k+m+p$ окремих питань. Для розставлення акцентів під час оцінювання необхідно провести ранжирування загроз за ступенем небезпеки, адже на практиці ніколи не зустрічається ситуація, коли б всі загрози (СА) були рівноцінні, що буде реалізовано на етапах 8 та 10.

Вибір шкали обчислень. Експерти мають провести оцінку небезпеки кожного компоненту експертного запиту. Для цього слід обрати оптимальну шкалу оцінювання та оптимальний для даної експертизи метод визначення КВ [11,14].

Існує декілька типів шкал, які прийнято класифікувати за типом відношень, що відображає шкала, та за типами даних, що вимірюються, які визначають множину допустимих математичних операцій над цими даними [7].

З точки зору технології експертного оцінювання найбільше значення мають шкали найменувань (ШН), порядку (ШП), інтервалів (ШІ) та відношень (ШВ). ШН слугують для того, щоб відрізнити об’єкти один від одного. За своєю сутністю ШН є просто класифікаційною шкалою, єдина операція, яка може бути проведена над отриманими даними у ШН, це перевірка на співпадіння або неспівпадіння. ШП дозволяє впорядковувати об’єкти у порядку зростання чи спадання кількісної характеристики будь-якої з їх властивостей (ознак). Слід зазначити, що ШП зберігає свої властивості при ізотонічних (монотонних) перетвореннях.

Характерним для ШВ є чітке визначення ,,нуля”, початку відліку. Результати виміру за ШВ володіють усіма властивостями чисел, і з ними можливо проводити будь-які стандартні операції. Тобто для ШВ числові значення числової системи $U_ч$ визначаються з точністю до перетворень подібності $\varphi(x)=\alpha x, \alpha>0$. Для результатів експертної оцінки, отриманих у ШВ, можливо застосувати будь-які статистичні методи оцінки [7,13].

6. Шкала може характеризуватися валідністю (можливість застосування шкали для виміру конкретного параметра), повнотою (здатністю виявити відношення експерта до критерію з заданим ступенем диференціації), чутливістю (S) (кількістю градацій), точністю (A) (в якій мірі результат оцінки за шкалою відповідає дійсності), надійністю (R) (сталість результатів у часі), можливістю перетворень, можливістю застосування статистичних методів обробки результатів (P) та простотою застосування експертом.

З-поміж цих параметрів найважливішими є можливість статистичної обробки, можливість перетворень, чутливість, надійність та простота застосування. ШН є скоріше якісною, ніж кількісною, адже оцінки мають нечисловий характер і результати оцінювання практично не піддаються математичним методам дослідження, отже даний тип шкали не має перспектив використання у експертному оцінюванні. Практично те саме можна сказати і про ШП, яка дає лише грубу оцінку параметра – рангові оцінки не піддаються арифметичним діям, хоча ШП можна перетворити в ШІ (нормалізація розподілу, заміна інтервалів на середні значення але це суттєво знижує точність результатів експертизи.

Єдиною перевагою ШП є значне спрощення процедури упорядкування експертних оцінок при її використанні. З точки зору трудомісткості ШІ і ШВ можуть розглядатися як правомірні. Проте ШВ має суттєву перевагу, оскільки точність отриманих з її допомогою оцінок вища, ніж для ШІ. Оцінки, отримані за ШВ, підлягають абсолютно всім статистичним методам дослідження, у той час к для ШІ існують деякі обмеження.

Із проведеного дослідження типів шкал обчислень можна зробити висновок, що основною шкалою для методу експертного оцінювання повинна бути ШВ, з метою спрощення процедури експертного оцінювання допустиме застосування ШП. ШН та ШІ не є ефективними та зручними, тому їх потрібно уникати.

7. Формування еталонних рівнів захищеності. Для оцінки рівня захищеності ДІР була обрана НМ з лінгвістичною шкалою [4], яка передбачає, що *N* співробітників дадуть відповідь на *n* компонент експертного запиту за шкалою, складеною експертом.

Експерти повинні побудувати нечіткі еталони, які відображають ЛЗ *(LS)* „рівень безпеки", що задається кортежем <*LS, T*$_{LS}$*, X*$_{LS}$>, і слугують зразком при порівнянні нечітких чисел. Базова терм-множина задається п'ятьма нечіткими термами $T_{LS\,w}$ = {„Низький" (Н), „Нижче середнього" (НС), „Середній" (С), „Вище середнього" (ВС), „Високий" (В)}, де $w = \overline{1,5}$. Це лінгвістична характеристика компонентів експертного запиту, яка може бути відображена на універсальну множину $X_{LS} \in \{0, x_{\max}\}$. Для кожного терму $T_{LS\,w}$ визначається інтервал [x_{min}; x_1[, …, [x_z; x_{max}] з використанням шкали обчислень, обраної на етапі 6. Визначені на даному етапі інтервали, терми та нечіткі числа будуть використані на етапах 9 та 11.

8. Визначення КВ. Для побудови системи визначення захищеності ДІР від СА були досліджені існуючі методи визначення КВ та виявлені параметри, які впливають на доцільність використання того або іншого методу в конкретних умовах проведення експертизи.

Найголовнішим чинником, що впливає на вибір методу обрахунку КВ, є фізична суть параметрів та відношення між ними. Параметри визначаються, виходячи з мети експертизи.

Далі слід визначитись із ступенем взаємозв'язків між параметрами – залежність або незалежність – та характером взаємозв'язків (незалежність по корисності, по перевазі, по байдужості і т.д.). Важливу роль відіграє ступінь узгодженості у оцінках експертів (Δ). Суттєво впливає на вибір методу складність проведення експертизи (F_E) і трудомісткість отримання експертної інформації (F_T), які визначаються реальними умовами та можливостями проведення експертизи [14].

Отже, на даному етапі з множини методів визначення КВ обирається метод визначення КВ M_S', який є оптимальним в умовах конкретної експертизи, та визначаються КВ λ_b для ранжирування експертного запиту на етапі 10.

9. Інтерпретація результатів запитів. На даному етапі *N* опитаних користувачів дають відповіді на *n* тестових питань експертного запиту за шкалою, визначеною експертами, наприклад „Ні", „Частіше ні", „Посередньо", „Частіше так", „Так". Кожній відповіді користувачів протиставляється у відповідність одне еталонне нечітке число. Значення нечітких чисел, що оцінює відповіді всієї групи користувачів на *b*-е питання, визначається за формулою:

$$L_b = \left(\sum_{r=1}^{N} AW_{br} \right) \Big/ N \quad ,$$ де $\sum_{r=1}^{N}$ – нечітке складання методом лінійної апроксимації за локальними максимумами (ЛАЛМ); AW_{br} – відповідь *r*-го користувача на *b*-е питання тесту; $b = \overline{1,n}$; $r = \overline{1,N}$.

10. Визначення рівня захищеності ДІР. Сумарна оцінка безпеки ДІР визначається за рахунок вирахуваних на етапі 8 КВ $LS = \sum_{b=1}^{n} (\lambda_b \times L_b)$.

У випадку, коли аналізуються відповіді лише одного співробітника, тоді сумарна оцінка захищеності одразу розраховується за наступною формулою $LS = \sum_{b=1}^{n} \lambda_b \times AW_b$. Визначена *LS* порівнюється з еталонним нечітким числом, для чого використовуємо відстань

Хеммінга: $d\left(LS, T_{LS_w}\right)=\sum_{w=1}^{Q}\left|\mu_{\mathrm{LS}}(x_w)-\mu_{T_{LS_w}}(x_w)\right|$, де $\mu_{\mathrm{LS}}(x_w)$ і $\mu_{T_{LS_w}}(x_w)$ – відповідно значення функцій приналежності поточного *LS* і еталонного T_{LS_w} чисел; $w=\overline{1,Q}$ [7].

Для розрахунку значень функції приналежності еталонного нечіткого числа для конкретного носія використовується формула прямої, яка проходить через дві точки, з якої випливає: $\mu(T)=\left(\left(LS-T_{LS_w}\right)/\left(T_{LS_{w+1}}-T_{LS_w}\right)\right)\cdot\left(\mu\left(T_{LS_{w+1}}\right)-\mu\left(T_{LS_w}\right)\right)+\mu\left(T_{LS_w}\right)$, де $T=LS$, $T_{LS_w}<T<T_{LS_{w+1}}$.

Після цього розраховується власне відстань Хемінга між поточним нечітким числом і еталонним, при цьому мінімальне зі значень *d* буде свідчити про більшу близькість поточного числа до еталона [7]. Найближчий еталон T_{LS}' і відображатиме визначений рівень безпеки ДІР. На основі запропонованих методології та алгоритмів реалізації операцій нечіткої арифметики і НМ можливо створити програмний продукт системи оцінки рівня захищеності ДІР від СА, призначений для проведення оцінки підготовленості персоналу державних органів до протидії атакам даного класу, за допомогою якого можливо провести опитування керівництва, адміністраторів безпеки та інших співробітників, які працюють з ДІР. Запитання є компонентами заздалегідь підготовленого масиву експертних еталонних запитів з визначеними одним з обраних методів КВ. За допомогою даного програмного продукту можливо визначити персонал з низьким або недосить високим рівнем підготовленості та вжити заходів з підвищення рівня безпеки ДІР від СА. Результати даної роботи мають практичну цінність та можуть застосовуватись під час проведення реальних експертиз з метою підвищення захисту ДІР, в тому числі і інформації, що становить державну таємницю.

ЛІТЕРАТУРА

1. Корченко А. Г. Методы анализа и оценки рисков потер государственных информационных ресурсов / А. Г. Корченко, В. П. Щербіна, С. В. Казмірчук // Захист інформації. – К.: 2012. – №1. – С.126 – 140.
2. Горніцька Д. А. Система социотехнических атак в информационной среде / Д. А. Горніцька, О. Г. Корченко, В. П. Харченко // Проблемы экономики и управления на железнодорожном транспорте. Материалы второй международной научно-практической конференции. – К.: ЭКУЖТ, 2007. – С. 137-138.
3. Корченко О. Г. Класифікація методів соціального інжинірингу / О. Г. Корченко, Є. В. Паціра, Д. А. Пуха // Захист інформації. – 2007. – Вип.4(36). – С. 37-45.
4. Горницкая Д. А. Атаки на ресурсы информационных систем в современном информационном обществею / Д. А. Горницкая, А. Г. Корченко, Е. В. Пацыра // Информационные технологии в гуманитарном образовании. Материалы I Между народной научно-практической конференции. – Пятигорск: 2008. – Ч.II. – С.224-233.
5. Корченко О. Г. Методи протидії соціотехнічним атакам. / О. Г.Корченко, Є. В. Паціра, Д. А. Пуха // Защита информации: Сб. науч. тр. – К.: НАУ, 2007. – С. 176-180.
6. Пуха Д. А. Методи соціотехнічних атак / Д. А. Пуха, Є. В. Паціра // Світ інформації та телекомунікацій–2007: Мат. IV Міжн. конференції студентства та молоді. – К.: ДУІКТ, 2007. – С. 133.
7. Корченко А. Г. Построение систем защиты информации на нечетких множествах. Теория и практические решения. / А. Г. Корченко – К.: МК-Пресс, 2006. – 320 с.
8. Хоффман Л. Дж. Современные методы защиты информации: Пер. с англ. / Под ред. В. А. Герасименко. – М.: Сов. Радио, 1980. – 264 с.
9. Горніцька Д. А. Дослідження методів апріорної оцінки якості експерта для реалізації експертиз у сфері інформаційної безпеки / О. Г. Корченко, Д. А. Горніцька, Т. Р. Захарчук // Захист інформації. – Київ, 2010. – №4. – С.53–60.
10. Дослідження апостеріорних методів оцінки якості експерта для сфери інформаційної безпеки / Корченко О. Г., Горніцька Д. А., Чепілко М. М. [та ін.] // Захист інформації. – Київ, 2011. – №1(50). – С.69 – 74.
11. Горніцька Д. А. Модель оцінки якості експерта для підвищення об'єктивності експертиз у сфері інформаційної безпеки / Д. А. Горніцька, О. Г. Корченко // Захист інформації. – Київ, 2011. – №2(51). – С.115 – 121.

12. Критерії оцінки захищеності інформації в комп'ютерних системах від несанкціонованого доступу: НД ТЗІ 2.5-004-99. — [Чинний від 1999.04.28]. — К. : ДСТСЗІ СБУ, 1999. — № 22. — (Нормативний документ системи технічного захисту інформації).

13. Литвак Б. Г. Экспертная информация. Методы получения и анализа. – М.: Радио и связь, 1982. – С. 23–28.

14. Горніцька Д. А. Визначення коефіцієнтів важливості для експертного оцінювання у галузі інформаційної безпеки / О. Г. Корченко, Д. А. Горніцька, В. В. Волянська // Захист інформації. – Київ, 2012. – №1. – С.108-121.

Надійшла: 12.07.2012 р.

Рецензент: д.т.н., професор Конахович Г.Ф.

УДК 004.056.53 **Бабенко Т.В., Сушко С.О.**

ПРО ЕНТРОПІЮ УКРАЇНСЬКОЇ МОВИ

У статті виконано аналіз підходів до формалізації природних мов та викладено результати експериментальних досліджень теоретико - інформаційних характеристик різних стилів української мови, визначена її надлишковість та приведено результати порівняння ентропії української мови з іншими природними мовами.

Ключові слова: ентропія, мова, дослідження, параметри.

У даний час існують різні підходи до формалізації природних мов сформульовані у вигляді математичних конструкцій, але не існує універсальної моделі мови, що дозволяла б з достатньою точністю апроксимувати реальну мову. Також залишається відкритим питання створення моделі, спроможної оцінити природність мови. Як відомо, ця задача є актуальною при вирішенні проблем оптимізації роботи пошукових систем у мережі Інтернет, задач криптографічного аналізу та інших.

Однією із задач, що потребує вирішення при синтезі моделі відповідної мови є визначення її теоретико-інформаційних характеристик, зокрема ентропії. Як відомо, знання ентропії відповідної мови є важливим при дослідженні асимптотики кількості осмислених відкритих текстів фіксованої довжини, при обчисленні відстані єдиності шифрів, для виявлення атипових зразків даних, що можуть нагадувати шкідливий код, при проведенні частотного аналізу, зокрема, ентропією оцінюють складність пароля в комп'ютерній індустрії.

Доцільно відзначити, що спроби обчислити ентропію із задовільною точністю виконувались для багатьох мов. Так, ентропію англійської мови досліджував класик теорії інформації К.Шеннон [1]. Радянський академік Піотровський Р.Г. із співробітниками одержали багато цікавих інформаційно-статистичних параметрів російської та інших мов колишнього СРСР [2–4]. Деякі дослідження статистичних властивостей української мови проводились в Інституті мовознавства ім. О.О. Потебні НАН України.

У роботі [5] зроблена спроба оцінити для української мови значення умовних ентропій розподілу ймовірностей біграм, триграм і чотириграм. Нажаль, у своїх розрахунках автори обмежились аналізом виключно українських та зарубіжних (у перекладі) літературних творів при цьому загальний обсяг текстового матеріалу незначно перевищив 12 млн. символів. Відповідно в зазначених роботах були отримані виключно інформаційні характеристики української абетки.

У даному досліджені авторами поставлено за мету оцінити значення ентропії української мови базуючись на вивченні текстів, що представляють всі стилі сучасної української мови.