УДК 004.9

А.Ю. Дяченко

КРИПТОГРАФІЧНИЙ ЗАХИСТ ІНФОРМАЦІЙНИХ СИСТЕМ НА ОСНОВІ ІНДУКТИВНИХ МОДЕЛЕЙ

Черкаський національний університет імені Богдана Хмельницького, dyachenko_a_u@mail.ru

Представлений новий метод криптографічного захисту інформації та основі індуктивних моделей. Розроблений метод може бути використаний для шифрування інформації у найрізноманітніших інформаційних системах. В якості ключа шифру використовуються тип, параметри та структура індуктивної моделі. Враховуючи той факт, що параметри та структура моделі можуть приймати значення з достатньо великого діапазону, то об'єм ключа забезпечує високий рівень криптостійкості

Ключові слова: *індуктивна модель, шифрування, криптографічна стійкість, захист інформації*

Вступ

В останні роки спостерігається бурхливе зростання обсягу потоків інформації в комп'ютерних системах. Відповідно збільшується кількість користувачів, що мають доступ до потоків інформації. Саме тому проблема захисту інформації є все більш актуальною. Для вирішення даної проблеми існую багато методів: програмні, технічні, організаційні, правові та інші методи [1]. Дуже важливу роль серед вищезазначених методів відіграють криптографічні методи, в тому числі методи шифрування інформації. Розвиток криптоаналізу та постійне зростання потужності комп'ютерів зумовлює висування все більшої кількості вимог до алгоритмів шифрування, зокрема: висока криптографічна стійкість, відносна простота програмної та апаратної реалізації, низька собівартість реалізації, висока швидкість операцій шифрування та розшифрування даних та ін. [2]. Тому актуальною задачею є розробка нових та удосконалення існуючих шифрів.

Аналіз досліджень. Забезпечення конфіденційності інформації є одним з основних завдань криптографії. Зазвичай конфіденційність реалізується за допомогою методів шифрування. Серед останніх особливий клас складають симетричні алгоритми шифрування, які призначені для обробки великих обсягів інформації.

На основі аналізу сучасного стану досліджень в області шифрування даних [1-5] можна зробити висновок, що в зв'язку з постійним зростанням потужності комп'ютерної техніки стали практично можливими такі методи криптоаналізу, зо раніш вважалися лише теоретичними. Тобто раніш їх реалізація вимагала практично недосяжних обчислювальних ресурсів. А зараз ці обчислювальні ресурси є все більш доступними, наприклад можливість використання розподілених обчислень через мережу Інтернет дозволяє поєднати потужності тисяч і мільйонів комп'ютерів в один суперкомп'ютер.

Крім того, постійно розвивається теорія криптоаналізу. Щороку винаходяться нові та оптимізуються існуючі методи криптоаналізу шифрів. Саме тому вдалося розкрити багато алгоритмів, які лише 20-30 років тому вважалися надійними. Тому проблема розробки достатньо надійного криптографічного методу захисту інформації є актуальною. Особливо актуальною є розробка принципово нових алгоритмів шифрування, наприклад з використанням досягнень теорії штучного інтелекту, зокрема нейронних мереж [6] та еволюційних алгоритмів. Також досить цікавими є алгоритми шифрування на основі рекурентних послідовностей [7].

Постановка завдання. Метою роботи є розробка нових методів криптографічного захисту на основі індуктивних моделей (ІМ).

Була сформульована гіпотеза про те, що для перетворення інформації при шифруванні можна використати ІМ.

Здатність ІМ до класифікації даних дозволяє розширити клас існуючих методів та засобів захисту інформації. Можливості спеціалізованих моделей сприяють ефективній практичній реалізації теоретично обгрунтованих розв'язків актуальних проблем захисту, зокрема шифруванню інформації.

Виклад основного матеріалу дослідження. При симетричному шифруванні безпосередньо перетворення блоку інформації здійснюється за допомогою шифруючого відображення:

$$E_k : A^* \to B^*, \ K \in Q \qquad (1)$$

де E_k – шифруюче відображення; ***A*** – алфавіт, що складається з множини слів**,** які називаються відкритими текстами ***M*** і утворюють простір відкритих текстів $A^*, M \in A^*$; ***B*** – алфавіт для запису зашифрованих текстів ***C,*** які становлять простір шифр текстів $B^*, C \in B^*$; ***Q*** – простір ключів, що складається з множини слів, які називаються шифруючими ключами ***K*** та обернених слів ***K***' деякого алфавіту [8].

Для оберненого перетворення блоку інформації (розшифрування) використовують дешифруюче відображення: $D_k : B^* \to A^*, \ K \in Q$ (2)

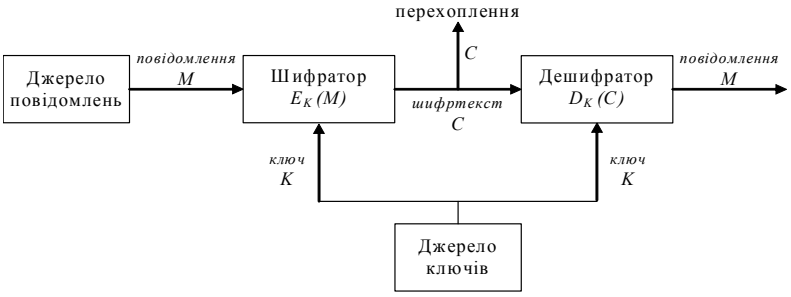

Рис. 1. Загальна структура симетричної системи шифрування

Разом E_k, D_k, A, B, Q утворюють симетричну систему шифрування (рис. 1). Як правило шифруюче відображення здійснюється шляхом послідовного застосування базових операцій над відкритим текстом. В якості базових операцій можуть застосовуватися: циклічний зсув, заміна, перестановка, додавання за модулем, логічна інверсія, підстановка, множення за модулем та інші.

Шифрування та розшифрування разом з правилами побудови ключів утворюють криптосистему. Для вибору певного криптографічного перетворення із сукупності можливих в криптосистемі застосовується ключ. Тобто ключ визначає тип криптосистеми і її характеристики.

Також E_k та D_k є взаємно оберненими:

$$D_k(E_k(M)) = M, \forall M \in A^*, K \in Q \quad (3)$$

$$D_k(E_k(M)) = E_k(D_k(M)) = M \quad (4)$$

При індуктивному моделюванні поведінка системи описується виразом виду:

$$F^* = F(x_1^*, x_2^*, \ldots x_n^*), \ a_i \le x_i \le b_i, \ F_{\min} \le F^* \le F_{\max} \quad (5)$$

де ***F*** - цільова функція, яка характеризує поведінку системи; x_i – параметри (атрибути) моделі; a_i та b_i – відповідно мінімальне та максимальне значення ***i***-го параметру з первинного опису об'єкта моделювання; F_{min} та F_{max}. – мінімальне та максимальне значення функції.

Слід зазначити, що ІМ можуть використовувати в якості генератора псевдовипадкових послідовностей (ПП). Причому отримана ПП є придатною до криптографічних застосувань [9].

Перед шифруванням інформації з використанням ІМ в якості ключа шифру необхідно підготувати групу моделей-ключів для шифрування, а також групу моделей-ключів для розшифрування.

Спочатку будуємо моделі, що реалізують шифруюче перетворення. Тобто кожна модель-ключ приймає на вхід незашифровані дані (відкритий текст), а на виході отримуємо певний символ (частинку) зашифрованого. Разом моделі-ключі дозволяють отримати на виходах певну множину символів, що разом утворюють зашифрований текст. Зазвичай розмір частинки становить 8 біт, тобто 1 байт, але для підвищення криптостійкості можна використовувати частинки більшого розміру: 16 біт, 32 біт або іншої кількості бітів. Для забезпечення достатнього рівня криптостійкості необхідно спеціальним чином генерувати первинний опис. Алгоритм генерації первинного опису залежить від типу ІМ, але загальне правило наступне: набори значень параметрів первинного опису вибираються з множини можливих значень незашифрованого тексту та підлягають обов'язковому перетворенню за допомогою певних математичних функцій. А в якості значень цільової функції виступає значення частинки незашифрованого тексту, яке теж підлягає перетворенню за допомогою певної математичної функції. Математичні функції для перетворення слід обрати такі, щоб ІМ могла здійснити апроксимацію функції.

За аналогічним принципом будуємо групу моделей-ключів для розшифрування, кожна з яких отримує на вході зашифровані дані, а на виході певний символ (частинку) розшифрованих

даних. В якості значень параметрів первинного опису використовуємо блоки даних зашифрованого тексту, що отримані як результат прогнозу групи моделей, що шифрують. А в якості значень цільової функції підставляємо значення частинки незашифрованого тексту.

Для ускладнення криптоаналізу рекомендується додавати в первинний опис розшифровуючих моделей певну секретну складову, тобто додаткові параметри, що не можливо відтворити з перехопленого зловмисником зашифрованого тексту. Параметри алгоритму генерації секретної складової тримається в секреті і є частиною множини розшифровуючих ключів.

Алгоритм симетричного шифрування на основі ІМ, полягає в послідовному виконанні наступних кроків:

1. Розбиття вхідного тексту на блоки певного розміру. В загальному випадку розмір блоку може бути довільним, але для забезпечення достатнього рівня криптостійкості рекомендується обирати розмір не менше 96 біт. Кожен блок розбивається на символи (підблоки) розміром 8,16, 32 (або інша кількість біт).

2. Набори значень символів послідовно підставляємо в обрану математичну функцію. Результат обчислень функції подаємо на вхід шифруючої моделі. Далі шляхом нормування вхідного значення в діапазони $a_i...b_i$ (для кожного з параметрів) отримуємо набір значень атрибутів моделі. В результаті прогнозу за цим набором параметрів отримуємо деяке дійсне число F_j.

3. Дії з кроку №2 послідовно виконуємо для кожної з шифруючих моделей. Отримуємо деяку множину значень ***F,*** що і є зашифрованим текстом.

4. Гамування, тобто накладання гама-послідовності на зашифрований текст. В якості гама-послідовності використовуємо псевдовипадкову послідовність, отриману з ІМ, а в якості функції гамування – виключне АБО (XOR).

Дешифрування здійснюється аналогічно, але в якості ключів використовуються розшифровуючі моделі, на входи яких послідовно подаються зашифровані дані, а на виходах отримуємо розшифровані дані.

Аналізуючи описану криптосистему слід зазначити, що тип, параметри та структура ІМ разом утворюють ключ шифру. І тип, і параметри, і структура ІМ можуть приймати значення з досить великого діапазону, отже розмір ключа є достатнім. Також оскільки для дешифрування необхідно знати певну секретну складову і обсяг ключових даних є великим, то довільний обсяг перехопленої зашифрованої інформації не дозволяє здійснити криптоаналіз. Тобто криптосистема відповідає оптимальному критерію кількості секретності [10]. Враховуючи те що ІМ функціонують на базі простих математичних операцій, то шифрування та розшифрування не потребує суттєвих обчислювальних ресурсів.

В запропонованій системі захисту на основі індуктивних моделей, рекомендується використовувати наступні алгоритми: МГУА [11]; Степаненка; 3) нейронні мережі; 4) генетичні (в яких генетичний алгоритм використовується для формування аналітичного опису системи) ; 5) багаторівневі.

Для дослідження запропонованих в статті методів криптографічного захисту на основі індуктивних моделей було розроблено спеціальний програмний комплекс з назвою «Crypto ІМ». Програмний комплекс складається з наступних основних модулів: навчання шифруючих ІМ, навчання розшифровуючих моделей, шифрування та дешифрування. Кожен модуль розроблений засобами C++ Builder 2007 та реалізований у вигляді окремої програми в форматі EXE для ОС Windows. Модуль навчання шифруючих ІМ призначений для підготовки набору моделей-ключів, що надалі передаються в модуль шифрування. Аналогічно модуль навчання розшифровуючих моделей готує моделі-ключі для модуля дешифрування. Модуль шифрування отримує на вхід текст і з використанням шифруючих моделей-ключів шифрує його, таким чином на виході модуля отримуємо файл двійкового формату з зашифрованим текстом. А модуль дешифрування отримує на вхід двійковий файл з зашифрованим текстом і з використанням розшифровуючих моделей моделей-ключів дешифрує його, відповідно на виході цього модуля отримуємо звичайний текст. Серед характеристик програмного комплексу слід виділити: дружній україномовний віконний інтерфейс користувача, представлення інформації в графічному та табличному вигляді, експорт та імпорт інформації в різноманітні формати (Word, Excel, TXT, CSV, XML, HTML ті інші) для обміну даними з іншими програмами, багатий вибір налаштувань ІМ, відносно низькі затрати часу для побудови ІМ в межах заданих параметрів.

Оскільки переважна більшість методів ІМ відносяться до наближених методів то необхідно дослідити, наскільки розшифрований текст відповідає початковому. Для з'ясування відповідності текстів було використано посимвольне порівняння. В таблиці 1 наведені результати дослідження.

Тексти отримані с сайтів www.dinternal.com.ua/topics/ та infoenglish.info/publ/topics/31 мережі Internet.

Таблиця 1

Відносна похибка методу криптографічного захисту на основі індуктивних моделей.

Назва тексту	Мова тексту	Довжина тексту, (символів)	Кількість невірно розпізнаних символів	Відносна похибка, (%)
1. My Friend	англійська	1601	2	0,1249
2. Мой друг	російська	1648	2	0,1214
3. Мій друг	українська	1626	2	0,1230
4. Easter in Ukraine	англійська	3222	4	0,1241
5. Пасха в Украине	Російська	3152	4	0,1269
6. Великдень в Україні	українська	3254	4	0,1229
7. London	англійська	4408	5	0,1134
8. Лондон	Російська	4862	6	0,1234
9. Лондон	українська	4847	6	0,1238
10. Olympic Games	англійська	3236	4	0,1236
11. Олимпийские игры	Російська	3168	4	0,1263
12. Олімпійські ігри	українська	3203	4	0,1249

За результатами досліджень видно, що усереднене значення відносної похибки складає 0,1232%, тобто є цілком достатнім для значної кількості прикладних задач, оскільки тексти характеризуються суттєвою надлишковістю інформації. Проте існує великий клас задач, де вказане значення похибки є неприпустимо великим, тому з метою розширення меж застосувань даного методу слід проводити дослідження по зменшенню відносної похибки, для чого пропонуються наступні методи: калібрування моделі, багаторядне та багаторівневе моделювання.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямку

Захищеність і криптографічна стійкість запропонованого алгоритму шифрування зумовлюється тим, що використовується секретний ключ великого об'єму, тобто атака типу повний перебір ключів вимагає недосяжно великих обчислювальних ресурсів. Розроблений алгоритм є придатним до програмної реалізації на практично всіх типах сучасних комп'ютерів, а також до апаратної реалізації на спеціалізованих мікросхемах. Подальші дослідження доцільно здійснювати в напрямку виявлення типів та параметрів ІМ, що підвищують крипостійкість шифру та не збільшують суттєво витрати обчислювальних ресурсів.

Перспективи застосування індуктивних моделей для задач захисту інформації: розробка нових та вдосконалення існуючих систем цифрового підпису на основі ІМ, розробка алгоритмів хешування даних, перевірка криптостійкості існуючих шифрів.

Список літературних джерел

1. Венбо Мао. Современная криптография: теория и практика / Венбо Мао. – М.: Вильямс, 2005. – 768 с.
2. Смець В., Мельник А., Попович Р. Сучасна криптографія. Основні поняття. - Львів: БаК, 2003. - 144 с.
3. Анин Б.Ю. Защита компьютерной информации. - СПб.: БХВ - С.-Петербург, 2000. - 384 с.
4. Шнайер Б. Практическая криптография / Н. Фергюсон, Б. Шнайер. – М.: Вильямс, 2005. – 425 с.
5. Грушо А.А., Тимонина Е.Е. Теоретические основы защиты информации. - М.: Изд-во агентства "Яхтсмен", 1996. - 130 с.
6. Томашевський О.М. Криптографічний захист інформаційних систем на базі штучної нейронної мережі // Труды Одесского политехнического университета.– 2001.– Вып. 4 (16).– С. 74–77.
7. Яремчук Ю.Є. Криптографічні методи та засоби шифрування інформації на основі рекурентних послідовностей. Монографія - Вінниця: «Книга-Вега», 2002. - 136 с.
8. Вербіцький О.В. Вступ до криптології. - Львів: Вид-во наук.-техн. л-ри, 1998. - 247 с.
9. Дяченко А.Ю., Голуб С.В., Квасніков В.В. Генератор псевдовипадкових послідовностей на основі індуктивних моделей. Збірник «Вісник Інженерної академії України». — Київ, 2009, том 2. — С. 87- 89.
10. Шеннон К. Работы по теории информации и кибернетике. s М.: Изд-во иностр. лит., 1963. s С. 333 - 369.
11. Івахненко А.Г., Зайченко Ю.П., Димитров В.Д. Принятие решений на основе самоорганизации. –М.: Сов. радио, 1976. – 280 с.