Кунцев, С.В. Технологія розробки програм аналізу даних за допомогою алгоритмів Data Mining бібліотеки Xelopes [Текст] / С.В. Кунцев // Системи обробки інформації : збірник наукових праць / Харківський національний економічний університет. – Харків, 2010. – Вип. 7 (88). – С. 224 – 225.

цього слід володіти певними відомостями про особу слухача.

Методи реєстрації і аутентифікації залежать від систем аутентифікації /ідентифікації, вживаних в конкретній СДО.

На даний момент широкий розвиток в світі отримали наступні системи реєстрации/ аутентифікації користувачів (Декларація ЄС 1999г., NIST 800-63, APEC Framework 2007 і так далі):

- програмні (паролі, компоненти ПО та ін.);
- технічні (смарт-карти, електронні ключі типа eToken і т.п.);
- біометричні (сканери сітківки ока, відбитку пальця, долоні, голосу, клавіатурного почерку і ін.).

Перераховані системи розрізняються своєю вартістю, складністю реалізації, часом реєстрації нового користувача і аутентифікації або ідентифікації, вірогідністю помилкового приймання законного користувача за порушника, і навпаки, а також наявністю або відсутністю деяких специфічних функцій, безпосередньо не пов'язаних з операцією розпізнавання користувача. Також їх застосовність залежить від необхідності особистої реєстрації користувача в інституті. Наприклад, пароль можна передати по електронній пошті, а смарт-карту передати по віртуальному простору неможливо, то ж стосується біометричних параметрів людини.

Застосування смарт-карт, електронних ключів, біометричних засобів вимагає наявності спеціального зчитуючого обладнання. Останнім часом пристрої зчитування відбитку пальця вбудовують в маніпулятор-мишу, а сітківка ока може розпізнаватися за допомогою web-камери. Паролі при нинішньому розвитку обчислювальної техніки стали ненадійними, вони часто розкриваються. Натомість використання таких біометричних параметрів людини, як голос або клавіатурний почерк, є надійним і недорогим способом ідентифікації користувачів.

Для ідентифікації по голосу необхідна наявність звукової плати, і, принаймні, мікрофону, тобто робоча станція має бути забезпечена мультимедіасистемою, вартість якої теж може бути невеликою.

При розпізнаванні по клавіатурному почерку жодного додаткового обладнання не потрібно, достатньо стандартної клавіатури.

У всіх випадках має бути реалізоване або куплене у сторонніх виробників ПО, що виконує функції реєстрації, аутентифікації користувачів, а також зберігання і обробки даних про користувачів системи, функціонує в глобальній мережі, подібній до Інтернету (найчастіше воно будується на клієнт-серверній архітектурі).

Оскільки більшість існуючих систем побудована на математичній моделі нейронних мереж, процес реєстрації займає значно більший період часу, чим розпізнавання користувачів. Оскільки активація нейронів у складі мережі здійснюється по деякій імовірнісній функції, існує певний відсоток можливої помилки як в ту, так і в іншу сторону.

Список літератури

1. Теренин А.А. Проблемы обеспечения безопасности систем дистанционного обучения / А.А. Теренин // Защита информации. Інсайд. – 2008. – № 4 (22). – С. 80-82.

2. Теренин А.А. Безопасность систем дистанционного обучения / А.А. Теренин // Защита информации. Інсайд. – 2008. – № 5 (23). – С. 86-89.

3. Теренин А.А. Создание защищенного канала передачи данных между распределенными ресурсами предприятия / А.А. Теренин // Защита информации. Инсайд. – 2005. – № 3. – С. 71-77.

УДК 004.853 (043.2)

С.В. Кунцев

Українська академія банківської справи Національного банку України, Суми

ТЕХНОЛОГІЯ РОЗРОБКИ ПРОГРАМ АНАЛІЗУ ДАНИХ ЗА ДОПОМОГОЮ АЛГОРИТМІВ DATA MINING БІБЛІОТЕКИ XELOPES

Сучасні інформаційні системи оснащуються аналітичними підсистемами – системами підтримки прийняття рішень (СППР). Згідно класифікації [1] розрізняють системи, засновані на телекомунікаціях і документах, на моделях, на даних і знаннях. Для аналізу даних і отримання нових знань в СППР використовують методи Data Mining. Найширше аналітичні технології Data Mining використовуються у галузі створення CRM-систем (11%), банківській справі (8%) і в маркетингу (5%) [2].

Основою для побудови СППР може служити програмний продукт компанії ZSoft [3] – бібліотека Xelopes, яка є відкритим, незалежним середовищем з гнучкою архітектурою, яка призначена для вбудовування в будь-яку інформаційну систему з метою забезпечення функціональності Data Mining. До складу бібліотеки Xelopes входять алгоритми "описові", які виконують пошук асоціативних правил і кластеризацію, і алгоритми "прогнозні", які виконують класифікацію і будують функцію регресії.

Проте, процес виявлення знань в даних є складним навіть за наявності відповідного програмного забезпечення. Очевидно, що для вирішення проблем потрібна ретельна підготовка фахівця-аналітика.

Для придбання практичних навичок аналізу даних за допомогою бібліотеки Xelopes рекомендується лабораторний практикум [4].

Особливе місце в практикумі займає розробка програми для інтелектуального аналізу даних. Важливо помітити, що технологія розробки програми є універсальною і не залежить від початкових даних або методу, який використовується.

Процес створення програми складається з наступних основних етапів:

1.Сбор даних і подання їх у форматі ARFF (Attribute Relation File Format).

2.Створення екземпляра класу MiningInputStream для читання початкових даних.

3.Створення екземпляра алгоритму класу MiningAlgorithm.

4.Створення екземпляра класу MiningSettings і MiningAlgorithmSpecification – їх параметри залежать від задачи.

5.Перевірка параметрів за допомогою методу verifySettings().

6.Передача алгоритму початкових даних і параметрів.

7.Побудова моделі MiningModel за допомогою методу buildModel().

8.Збереження моделі у форматі PMML (predictive modeling mark up language).

9.Застосування моделі до нових даних (виконується у тому випадку, якщо модель є керованою – відноситься до класу SupervisedMiningModel).

В процесі побудови модель настроюється, її властивості коригуються. Для кожної моделі задаються свої параметри: 1) для моделі, що представляє асоціативні правила: мінімальна підтримка, мінімальна довіра, атрибут ідентифікації транзакції, атрибут ідентифікації елементів; 2) для моделі, що представляє задачу класифікації: атрибут класифікації, максимальна глибина дерева, максимальне число замін, мінімальний розмір вузла дерева; 3) для моделі, що представляє задачу кластеризації: максимальна кількість кластерів, параметри для обчислення відстані між об'єктами; 4) для моделі, що представляє математичну залежність: атрибут класифікації даних, тип моделі, вид функції, параметри ядра, загальні параметри.

Таким чином, розробка програм з використанням бібліотеки Xelopes дозволяє аналітику на практиці вивчити основні принципи створення систем інтелектуального аналізу даних.

Список літератури

1. Академия АйТи [Електронний ресурс]. – Режим доступу до ресурсу: http://www.academy.it.ru/ru/.

2. KDnuggets. Лучшие ресурсы Data Mining [Електронний ресурс]. – Режим доступу до ресурсу: http://www.kdnuggets.com/.

3. Компания ZSoft. Библиотека Xelopes [Електронний ресурс]. – Режим доступу до ресурсу: http://www.zsoft.ru/page.php?14.

4. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP: учебное пособие / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – 2-е изд. – СПб.: БХВ-Петербург, 2007. – 384 с.

УДК 681.3.06

И.В. Московченко

Национальный технический университет "ХПИ", Харьков

ОЦЕНКА ЭФФЕКТИВНОСТИ РАЗРАБОТАННОЙ ВЕРОЯТНОСТНОЙ МОДЕЛИ ФОРМИРОВАНИЯ НЕЛИНЕЙНЫХ УЗЛОВ ЗАМЕН, ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ СОВЕРШЕНСТВОВАНИЯ СИММЕТРИЧНЫХ КРИПТОГРАФИЧЕСКИХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ

Использование предложенной вероятностной модели формирования нелинейных узлов замен (далее НУЗ) позволяет формировать блоки нелинейной подстановки для симметричных криптографических средств защиты информации (далее КСЗИ). Установлено, что формируемые НУЗ обладают улучшенными свойствами, их применение в симметричных КСЗИ позволяет улучшить показатели статистической безопасности.

Предложенные алгоритмы вероятностного формирования блоков нелинейной подстановки и программная реализация позволяют строить сбалансированные НУЗ с заданными показателями нелинейности, высокими корреляционными, спектральными и другими свойствами для симметричных КСЗИ. Показано, что практическое использование разработанных алгоритмов и их программной реализации позволяет формировать блоки нелинейной подстановки различного размера для различных криптографических приложений.

Проведенные исследования криптографических свойств сформированных блоков нелинейной подстановки показали, что по основным и дополнительным показателям стойкости они не уступают лучшим известным мировым аналогам, а по некоторым показателям (нелинейность и автокорреляция) превосходят их. Таким образом, разработанный вычислительный метод вероятностного формирования криптографических булевых функций и вероятностную модель формирования НУЗ целесообразно ис-