Taras Shevchenko National University of Kyiv
Software Systems and Technologies Department
Київський національний університет імені Тараса Шевченка
Кафедра програмних систем і технологій

France, LeMann , University LeMann,

Bulgaria, Sofia, University of Library Studies and Information Technologies

Czech Republic, Brno, University of Technology Institute of Mathematics and Descriptive Geometry

MSTIoE 2018-4
4-th East European Conference on Mathematical Foundations and Software Technology of Internet of Everything

4-та Східно-Європейська конференція Математичні та програмні технології Internet of Everything

(20-21.12.2018, Kyiv - Київ**)**

Proceedings - Збірник матеріалів

Збірник внесений до наукової електронної бібліотеки
Національної бібліотеки України імені Вернадського

УДК 002.5:004

MSTIoE 2018-4. 4-та Східно-Європейська конференція “Математичні та програмні технології Internet of Everything” (20-21.12.2018, Київ). Зб.матер., КНУ ім.Т.Шевченка, 56с.

У збірнику узагальнені матеріали конференції, яка проходила на базі кафедри програмних систем і технологій (ПСТ) Київського національного університету імені Тараса Шевченка 20-21.12.2018 року. В матеріалах висвітлюються актуальні питання розвитку теорії та практики програмування Internet of Everything (Всеохоплюючий інтернет, IoE) та розвітку суміжних технологій.

Напрямки конференції:

- Математичні технології IoE;
- Програмні технології IoE;
- Апаратні технології IoE;
- Інші технології та види забезпечення IoE.

Збірник розрахований на представників бізнесу та державних органів, викладачів та наукових працівників, аспірантів, студентів які займаються питаннями програмування Інтернет-додатків та розвитку інформаційних технологій в цілому.

MSTIoE 2018-4. 4-th East European Conference on Mathematical Foundations and Software Technology of Internet of Everything. (20-21.12.2018, Київ). Proceedings, Taras Shevchenko National University of Kyiv, 56 pp.

The collection summarizes the materials of the conference, which took place on the basis of the Department of Software Systems and Technologies (PST) of the Taras Shevchenko Kyiv National University on the 20-21.12.2018. The materials cover the topical issues of the development of the theory and practice of programming Internet of Everything (IoE) and the development of related technologies.

Directions of the conference:

- IoE Mathematical Technologies;
- IoE Software Technologies;
- Hardware IoE technology;
- Other technologies and types of IoE security.

The collection is intended for representatives of business and government bodies, teachers and researchers, postgraduates, students who are involved with the issues of programming of Internet applications and the development of information technologies in general.

MSTIoE 2018-4. 4-я Восточно-Европейская конференция "Математические и программные технологии Internet of Everything" (20-21.12.2018, Киев). Сб матер., КНУ им.Т.Шевченко, 56 с.

В сборнике обобщены материалы конференции, которая проходила на базе кафедры программных систем и технологий (ПСТ) Киевского национального университета имени Тараса Шевченко 20-21.12.2018 года. В материалах освещаются актуальные вопросы развития теории и практики программирования Internet of Everything (Всеобъемлющий интернет, IoE) и розвитку смежных технологий.

Направления конференции:

- Математические технологии IoE;
- Программные технологии IoE;
- Аппаратные технологии IoE;
- Другие технологии и виды обеспечения IoE.

Сборник рассчитан на представителей бизнеса и государственных органов, преподавателей и научных работников, аспирантов, студентов, которые занимаются вопросами программирования Интернет-приложений и развития информационных технологий в целом.

Матеріали подано в авторській редакції

Відповідальний за випуск: Шевченко В.Л.

Адреса редакції: 03116 Київ-116, вул. Ванди Василевської, 24
тел. (066) 430-54-32

http://pst.knu.ua/MSTIoE

ЗМІСТ

CONTENTS

СОДЕРЖАНИЕ

Bychkov O.S. *DSc, associated prof. bos.knu@gmail.com*
Taras Shevchenko National University of Kiyv

DEVELOPMENT OF INTERNET TECHNOLOGY FOR CONSULTATIONS IN CARDIOVASCULAR DISEASES DIAGNOSTICS

This article describes Internet software technology created and mathematical model of blood movement in large vessels of human circulatory system developed and software to conduct the experiments. Lattice-gas Boltzmann method with 3-dimentional model D3Q19 was taken as a basis. Based on the developed technology, it is proposed to create an Internet consultation center.

The components of blood are: plasma, erythrocytes and leucocytes, for which all the biological parameters were calculated and special conditions for interaction with vessel's walls were created (so called boundary conditions). By using this model and the effective software, authors have studied two types of flows, which are created by heart, namely laminar and spiral. The results of experiments show that as time goes by spiral flow becomes laminar, and this fact proves the adequacy of model to real processes in human vessels.

The role of the cardiovascular system in the human body is extremely important, because blood is a fluid that supplies all the substances necessary for the vital functions of cells of any tissue of any organ, is the basis of immunity and homeostasis in the body. Any violation of the blood flow causes a shortage of supply of vital substances, and therefore the disease of the tissue or even its death. That is why the study of hemodynamics, that is, the movement of blood, is given special attention in modern medicine.

With the advent of powerful computers, sensory and precision measuring devices, scientists have gained the opportunity to explore these phenomena not only at the macroscopic level, measuring pressure, temperature, etc., but also at the microscopic, molecular level. Only at this level is it possible to simulate blood as a multicomponent fluid and to study processes such as the exchange of substances in the capillaries between tissues and blood, the movement of red blood cells and the possible occurrence of thrombosis in capillaries [4].

Blood simulation as a multicomponent fluid in large vessels allows us to investigate the change in viscosity of blood, that is, the rate of hemodynamic resistance relative to the stationary surface, the inner layer of the vessels, and the resistance of the motion relative to adjoining moving fluid layers. As a result of this increase, the resistance to blood flow and, consequently, the heart load increases.

It is known that blood is a non-Newtonian fluid [3], because, unlike a simpler fluid, it has a complex structure, namely, elemental elements (erythrocytes, leukocytes and platelets). Because of the above reasons, blood modeling as a multicomponent fluid is a prerequisite for obtaining an adequate model of processes that occur in human vessels.

The mathematical model of blood flow as a three-component fluid is based on the Boltzmann lattice method for a multicomponent fluid [2]. For the model, special boundary conditions have been developed that take into account the specifics of the interaction of each component with the walls of the vessels. Aided by the special software product, two types of flow (laminar and spiral) created when a blood arrives with a certain periodicity in the vessel, was tested.

It is proposed to use the developed mathematical model in creating online consulting centers and training of personnel. This model allows you to demonstrate typical cardiovascular diseases.

1. Бандман О.Л. Клеточно-автоматные модели пространственной динамики. // сб. Системная информатика. – 2005. – Вып. 10. – С. 57-113.

2. Nicos S. Martys, Hudong Chen. Simulation of multicomponent fluids in complex three-dimensional geometries by the lattice Boltzmann method. // Phys. Rev. E 53. –1996. – pp. 743-750.

3. Ремизов А.Н., Максина А.Г., Потапенко А.Я.. Учебник по медицинской и биологической физике. – Москва: Дрофа, 2003.

4. Krzysztof Boryczko, Witold Dzwinel, David A. Yuen. Dynamical clustering of red blood cells in capillary vessels. – Springer-Verlag, 2003.

***Нестеренко О.В**., к.т.н., доц.*
Провідний науковий співробітник
***Нетесін І.Є**., к.ф.-м.н.*
Провідний науковий співробітник
Державне підприємство «Український науковий центр розвитку інформаційних технологій», м. Київ, Україна

ГРАФОВІ МОДЕЛІ КІБЕРБЕЗПЕКИ ІНТЕРНЕТУ РЕЧЕЙ

Незважаючи на те, що Інтернет речей (*Internet of Things*, IoT) знаходиться тільки на початку свого розвитку, але вже доводиться казати, що це нововведення додає серйозних проблем, пов'язаних з кібербезпекою. Керування пристроями за допомогою міжмашинної взаємодії несе великі загрози не лише для даних, з чим ми звикли мати справу в традиційних інформаційних системах, а й для функціонування підприємств, працездатності критичних інфраструктур і навіть для життя людини.

В даному аспекті нестача досліджень взаємозалежності складових безпекового середовища IoT, досліджень на відповідність умовам, що склалися та можливість адаптації до змін, що відбуваються, виступають тією проблемою, яка потребує розгляду. Досі залишаються невирішеними питання, що полягають у теоретичному вивченні та розкритті підходів до опису взаємодії загроз і різних властивостей інформаційних одиниць (об'єктів) для вибору засобів захисту, які необхідні для розв'язання задач забезпечення безпеки. Перспективним вважається представлення моделей у вигляді графів [1], але й цей підхід потребує подальшого розвитку.

У загальному випадку побудова моделі має за мету необхідність захисту від впливу загроз усім властивостям захищеності, насамперед від загроз, наслідком реалізації яких може бути неприпустимо високий чи високий рівень шкоди, оскільки такі загрози мають комплексний, тобто одночасний вплив на декілька властивостей захищеності. Такі загрози прийнято називати найбільш суттєвими (найбільш небезпечними, найбільш імовірними) загрозами. Виявлення найбільш суттєвих загроз та високого рівня шкоди, нанесеної ресурсу, є основою для визначення в подальшому потрібних засобів захисту, а отже визначення складу потрібних контрзаходів для забезпечення припустимої захищеності, необхідних для захисту засобів, підсистем, механізмів та функцій захисту, тобто дає змогу будувати відповідні моделі систем захисту.

В центрі будь-якої системи моделей знаходиться певна класифікація, яка описує поняття предметної області. Насамперед необхідно розглянути класифікацію загроз та нештатних ситуацій, що можуть виникнути відносно системи IoT. Для аналізу загроз необхідним є визначення можливих каналів та видів загроз, що можуть бути реалізовані відносно системи, а також аналіз основних джерел їх походження.

Одним з найнебезпечніших напрямів атаки, на які варто звернути увагу, є DDoS-атака. Також атака може бути реалізованою віддалено по Інтернету до обладнання IoT, яке має уразливості у безпеці своїх паролів, проблеми з шифруванням даних і з дозволом доступу. Машини, які передають інформацію в незашифрованому виді, можуть зберігати пароль мережі Wi - Fi, до якої вони приєднані. Також IoT-пристрої, які не мають належного захисту, піддаються атакам з метою зараження шкідливим кодом для створення ботнету.

Самі датчики як частини системи IoT є найуразливішими для зловмисників при спробах отримання доступу до основної мережі підприємства. В залежності від розташування датчиків, їх призначення (функціональності), виду з'єднань можна застосувати таку класифікацію уразливості датчиків, на які спрямовані загрози: стек мережевих протоколів (з тих, що використовуються в IoT), радіо-протоколи, засади роботи мікроконтролерів, інжиніринг прошивок та скомпільованих програм, web-уразливості. Враховуючи цю класифікацію, напрямки забезпечення кібербезпеки для IoT зводяться до трьох спрямувань: Hardware security, Software security та Radio security.

Тоді модель відношень множини загроз T і множини об'єктів O можна представити дводольним графом $G_{TO} = (V(T,O), E(T,O))$, у якому множини вершин його долей $T \cup O = V(T,O)$, $T \cup 0 = V(T,0)$, $\mathrm{T} = \mathrm{T1}, \mathrm{T2}, \ldots, T_{|T|}$, $O = 01, 02, \ldots, O_{|O|}$, де $|T|$, $|O|$ -

кількість елементів відповідно множин *T* та *O*, $T \cap O = \emptyset$ та множина ребер *E(T,O)*, в якому ребро $(T_p, O_q) \in E(T,O)$, якщо є загроза *Tp* об'єкту *Oq*. Спрямованість загроз до об'єктів визначається на основі їх описів (класифікації).

У відповідності до відомої моделі безпеки з повним перекриттям [2], яка будується виходячи з положення, що система безпеки повинна мати принаймні один засіб для забезпечення безпеки на кожному можливому шляху дії загрози на об'єкт, в нашій моделі з'являється третій набір, що описує механізми забезпечення безпеки

$M = M_1, M_2, \ldots, M_r, \ldots, r = 1, 2, \ldots.$

В ідеальному випадку набір механізмів *M* повинен усувати всі ребра T_p, O_q . На практиці ж M_r виконує функцію "бар'єру", забезпечуючи деяку міру опору спробам реалізації загрози. Включення в модель множини *M* перетворює граф G_{TO} в тридольний граф $G_{TMO} = V(T,M,O), E\ T,M\ , E(M,O)$.

Побудова графу G_{TMO} є нетривіальною задачею, зважаючи на складність зв'язків графу G_{TO}. Щоб полегшити розв'язання цієї задачі достатньо розшукати на графі G_{TO} компоненти зв'язності, тобто такі його підграфи $G_i = V_i, E_i$, що $G_{TO} = \cup Gi$, але $V_i \cap V_j = \emptyset$ та $E_i \cap Ej = \emptyset$, $i, j = 1,2,\ldots, i \neq j$, у той час як у будь-якому G_i будь-які вершини *u* та *v* з'єднані простим ланцюгом.

Для знаходження компонент зв'язності можливо застосувати відомі алгоритми. Більшість алгоритмів на графах використовують їх представлення за допомогою матриці суміжності або списків суміжних вершин. У разі представлення графу за допомогою списків суміжних вершин для пошуку компонент зв'язності зазвичай застосовують алгоритми, які базуються на алгоритмах пошуку в глибину та пошуку в ширину, які досліджують граф методом обходу усіх вершини і ребер, використовуючи механізми рекурсії, фарбування вершин або ребер, поняття предків і нащадків, міток часу тощо [3,4].

Тепер послідовно розглядаючи отримані підграфи G_i визначення необхідних механізмів захисту з набору *M* вочевидь значно спрощується.

Але залишається проблема оцінки пріоритетності та вартості реалізації вибраних механізмів захисту. Для цього традиційно відомим є визначення передусім оцінки ризиків, яка випливає з ймовірності здійснення загрози та рівня шкоди (збитків) від порушень по кожному з їх видів.

Таким чином створення графових моделей безпекового середовища IoT для визначення впливів різного типу загроз та захищеності об'єктів свідчить про зручність їх застосування, надаючи ефективний інструмент менеджерам і розробникам засобів захисту. Одночасне застосування елементів класифікації (онтологічних описів) підвищить рівень конкретності моделі та більш чіткого уявлення щодо стану середовища.

ЛІТЕРАТУРА

1. Качинський А.Б. Ієрархія факторів типових сценаріїв реалізації DDos-атак / А.Б. Качинський, В.М. Ткач, А.А. Поденко // Математичне моделювання в економіці, 2017. – №1-2. – С. 17-30, 2018. – №1. – С. 31-48.)

2. Хоффман Л.Дж. Современные методы защиты информации / Л.Дж. Хоффман. Пер. с англ. / М.: Советское радио, 1980. – 264 с.

3. Кормен Т. Алгоритмы. Построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест, К. Штайн / пер. с англ. – 2-е изд. – Москва–Санкт-Петербург–Киев: Издательский дом «Вильямс», 2013. – 1296 с.

4. Харари Ф. Теория графов / Ф. Харари / пер. с англ. – М: «Мир», 1973. – 302 с.

***Нестеренко О.В**., к.т.н., доц., провідний науковий співробітник*
***Зайцев С.М**., к.ф.-м.н.*, п*ровідний науковий співробітник*
***Артеменко О.Л**., провідний науковий співробітник*
Державне підприємство «Український науковий центр розвитку інформаційних технологій», м. Київ, Україна

УСЕОХОПЛЮЮЧИЙ ІНТЕРНЕТ, ВЕЛИКІ ДАНІ І КОМПЮТЕРНІ ОНТОЛОГІЇ

Глобалізація у сучасному світі - це використання для аналітики і прийняття ефективних рішень значних, вельми значних об'ємів даних. Ідеї, що з'являються в результаті проведення аналізу даних, використовуються для поліпшення клієнтського обслуговування, зниження витрат або навіть для запровадження нових бізнес-моделей, які «висаджують» галузі і витісняють конкуренцію. Але більш ніж будь-коли дані, на яких тримається це розуміння, стрімко множаться та поширюються від постійно зростаючої кількості джерел.

Навіть поверховий огляд різних видів і джерел даних в сучасній економіці малює картину виключно диференційованого ландшафту даних. Почнемо з традиційних джерел даних, що відносяться до систем управління підприємствами - системи ERP, CRM, керування ланцюжками постачань (*Supply Chain*), управління життєвим циклом продукту (PLM), колл-центри обслуговування клієнтів і т.ін. Додамо електронну пошту, статистику сайту компанії і матеріали електронної комерції. У підсумку маємо поєднання значної кількості структурованих і неструктурованих даних. Далі включимо сюди дані з нових каналів взаємодії з клієнтами - платформи продуктів, мобільні застосування, соціальні мережі.

Але все це дрібниці коли мова заходить про Інтернет Речей (*Internet of Things*, IoT) та Всеохоплюючий Інтернет, або Всеосяжний Інтернет (*Internet of Everything*, IoE). Якщо компанія почала використовувати датчики у виробничому устаткуванні, носимому/возимому приладді або у товарах на полицях - вважайте це ще одним важливим джерелом даних для управління. Але залишається одна проблема – як опанувати цей величезний об'єм даних?

Коли з'являється багато даних, по суті, старі способи управління ними для аналітики і зберігання вже не підходять. Тому ще наприкінці 90 років у обіг був уведений термін Великі дані (*Big Data*), під яким на сьогодні розуміються інтелектуальні технології аналізу значних об'ємів даних з різних джерел (бажано з усіх можливих у предметній області) з великою швидкістю і в режимі реального часу.

Стратегія «цифрового перетворення» (*Digital Transformation*) як нового ядра для тих, хто прагне конкурувати в цифровій економіці, надає можливостей значній частині великих світових компаній вже сьогодні отримувати дохід від послуг, що надаються на основі накопичених даних (*data-as-a-service*), таких як продаж необроблених даних, метрик, виявленого розуміння (знання) та рекомендацій, отриманих на основі технологій бізнес-аналітики (*Business Intelligence*, BI) та інтелектуального аналізу даних (*Data mining*).

Технології Big Data дійсно дозволяють знаходити всілякі кореляції в будь-яких даних. Проте, в багатьох випадках відповідь може бути не настільки очевидною – чи спостерігаємо ми причинно-наслідковий зв'язок, чи випадковий збіг, що є суттєвою системною проблемою.

На цей час існують певні технологій, зокрема когнітивні семантичні технології і штучний інтелект (машинне навчання), що допомагають вирішувати проблеми аналізу великих даних. На підході й нові технологічні інструменти управління даними. Візьмемо, наприклад, обробку даних в пам'яті (*in-memory*) – швидкість обробки приголомшує. Гарною новиною є й те, що існують хмарні обчислення (*cloud computing*).

Однак наявність цього потужного інструментарію все ще не дозволяє вирішити вищезазначену проблему. У суспільних і соціальних сферах, зокрема в економіці, в державному управління – там, де широке поле для Великих даних – насправді немає скільки-небудь строгої теорії структуризації понять і даних, теорії в тому сенсі, як це розуміють, наприклад, природничі науки. Адже нова революція полягає не в комп'ютерах, які обробляють дані, а в самих даних і в тому, як ми їх використовуємо.

Цінність результату аналітичної діяльності вочевидь напряму залежить від повноти обробленої інформації [1]. Об'єм інформації, що опрацьовується, зазвичай складається з трьох частин: $I_A = I_П + I_Д + I_O$, де I_A - обсяг опрацьованої інформації; $I_П$ - обсяг пертинентної інформації; $I_Д$ - обсяг додаткової інформації, яка є релевантною, але не є пертинентною; I_O -

особиста інформація, яка згенерована власне аналітиком у процесі дослідження даних за рахунок індивідуальних знань, існуючого інструментарію аналізу даних і творчого підходу до аналітичних дій.

Найважливішою умовою успішної аналітичної роботи експерта є наявність інформаційного поля досліджуваної предметної області (ПдО), яке має уявляти множину структурованих і неструктурованих інформаційних масивів (інформаційних ресурсів), необхідних для витягу з них необхідних даних. Робота аналітика починається зі збирання інформації щодо проблеми, що розглядається, з різних джерел шляхом використання інформаційно-пошукових систем. Хоча можливості сучасних пошукових систем постійно зростають, однак поки що вони, як і раніше, є ще далекими від ідеалу, тому об'єм $I_Д$ іноді може значно переважати об'єм $I_П$, що ускладнює роботу аналітика і знижує ефективність систем аналізу. Водночас для формування I_O необхідно провести аналіз даних, що накопичені в базах даних автоматизованих систем, які функціонують в організації. Проблема зростає у зв'язку із тим, що об'єми цих БД постійно і стрімко зростають, і вже часто-густо стають для аналітика неосяжними.

Розвиток наукових досліджень і досвідів в сферах моделювання аналітичної діяльності свідчить, що одним з інструментів, за допомогою якого можливо досить ефективно спроектувати та реалізувати механізми управління ієрархією, яка відображає взаємодію усіх інформаційних компонентів, може бути онтологічна модель [2,3]. Вона у своєї інформаційній основі має механізм динамічного формування та використання ієрархій у вигляді певних таксономій.

Предметну область (ПрО) безпосередньо складають певні концепти та їх властивості. Концепти складають кінцеву множину *X={X1,X2,…,Xi,…,Xn}*, а множина властивостей *R* утворюється множиною декартових добутків множини ***X*** самої на себе: $\mathrm{R} = {}_{i}^{n} X_i$. Будемо вважати, що властивості *R* є інтерпретацією відношень, тобто існує перетворення, яке кожному відношенню встановлює відповідність певної властивості. Тоді множина *F* може бути утвореною декартовим добутком множин ***X*** і *R:* $F = X \times R$.

Таким чином онтологія деякого операційного середовища в загальному випадку формально представляється впорядкованою трійкою $O = X, R, F$.

На основі концептів ***X*** (понять, термінів) ПрО формується предметна складова операційного середовища аналізу даних. Функції інтерпретації *F* (визначень) ***X*** та/або *R* складають функціональну частину операційного середовища.

Розгляд граничних випадків множин: $R = \varnothing$; $R \neq \varnothing$; $F = \varnothing$; $F \neq \varnothing$ у всіх чотирьох комбінаціях значень R і F дає різні варіанти онтологічних конструкцій, починаючи від простого словника і таксономії до формальної структури концептуальної бази знань для високоінтелектуальних знання-орієнтованих систем.

ЛІТЕРАТУРА

1. Пархоменко Б.В. Роль творчості в процесі обґрунтування прийняття рішень / Б.В. Пархоменко // XII Международная научно-практическая конференция «Построение информационного общества: ресурсы и технологии», 6-7 июня (2007, Киев, Украина). - К.: УкрИНТЭИ, 2007. С. 106-108.

2. Емельянов С.В. Многокритериальные методы принятия решений / С.В. Емельянов, О.И. Ларичев. М.: Знание, 1985. 32 с. – (Новое в жизни, науке, технике. Сер. «Математика, кібернетика», №10).

3. Стрижак О. Є. Онтологічні інформаційно-аналітичні системи / О. Є. Стрижак // Радіоелектронні і комп'ютерні системи. – 2014. – № 3 (67). – С. 71–76.

Бражиненко М.Г. *студент 4 курсу*
Козачок П.А. *студент 4 курсу*
Федоренко Р.М. *к.е.н., асистент*
Шевченко В.Л. *д.т.н., професор*
кафедра Програмних систем і технологій
Київський національний університет імені Тараса Шевченка, м. Київ, Україна

M2M SIGNAL DATA COMPRESSION

Nowadays, it is unlikely to have a day without humming sensors, actuators, and smart meters. None of them is isolated anymore. The goal, is to connected them together and make our life and business easier. In the past, we expect that people shall take an action first. In difference, now when IoE becoming our reality we would like smart sensor and technologies to take an intelligent decisions and act on behalf of us. This expectation dictated by the global trend of aging society and our will to help people who could not use phones, smart devices due to physical inability. As an example, seniors suffering from heart attack may not be able to ask for an ambulance, child who stay alone at home or play in the courtyard may not be able to ask parents for help. List of such kind of scenarios could be expanded however, self-help ability in all of them were connected with three key constraints:

1. Person has a physiological ability to invoke information transfer.
2. Person transferring data understand how to use digital technologies.
3. Affordability and ability to carry everywhere (like watches).

Technologies and platforms implementing IoE phylosophy and values expected to solve above constraintes, enabling for most unprotected social groups new digital experience that any of us could have ever imagine before. However technology intelegince and ability to process and maintain devastating amount of sensor data require new breathtaking approach to well know paradigms and principles, like data compression and M2M data exchange. We believe that M2M communication and data exchange should enable an opportunity rather than create a constraints. Most of existing protocols and technologies applicable for IoE (similary to IoT) lack qulaities balance. In this work we propose technology that allow to solve most of existing chalanges with IoE M2M data exchange communication and in addition reduce network traffic for low bandwidth connections.

Conclusion: As a result of reserch reference qualities for embeded systems with connected sensors we propose new M2M communication protocol with compression for 1-D signal data transfer. The proposal includes message structures, specifications and data compression methodology using well known principles. Compression algorythm allow to get 4.7:1 compression ratio for worst cases. Compression algorythm is rather fast and easy to implement. Our new protocol minimize messages count and size of data sent within a system and enable to use low bandwidth networks like LoRa\LoRaWAN. Protocol traffic utilization is expected to be 80% on avarage for production usage.

Шевченко В.Л. *д.т.н., професор*
Климко В.В. *студент 4 курсу*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка
м. Київ, Україна

ВИЗНАЧЕННЯ ПОТОЧНОЇ АКТИВНОСТІ КОРИСТУВАЧА ЗА ДОПОМОГОЮ ВБУДОВАНИХ СЕНСОРІВ СМАРТФОНУ

На сьогоднішній день однією з основних потреб більшості мобільних додатків є визначення поточної активності користувача. Розпізнавання активності має широкий діапазон застосувань від відстеження фізичної форми і стану здоров'я користувача до контекстно-залежної реклами що виводиться, в залежності від попередніх дій користувача. Для цього смартфони мають не тільки потужні обчислювальні характеристики, але й різноманітні датчики, які дозволяють дізнатися про стан зовнішнього середовища та про стан самого пристрою в цьому середовищі.

Хоча були проведені дослідження в області контекстних додатків і систем, не було приділено достатньої уваги до розробки надійних та ефективних (з точки зору енергоефективності, детальності отриманих даних та обчислювальних витрат) алгоритмів, які автоматично виявляють поточну активність мобільних телефонів.

Серед основних задач, при отримані даних з сенсорів, є розширення діапазону можливих значень та очищення даних від шуму.

Для всіх сенсорів існують діапазони можливих значень:

- Акселерометр: для трьох координат від -10.0 до ~10.0 м/ с2;
- Сенсор магнітного поля: для трьох координат від -1000.0 до 1000.0 мкТл.;
- Сенсор температури: від -273.1 до 100 градусів за Цельсієм;
- Сенсор наближення: від 0 до 10 см.;
- Сенсор натиску: від 0 до 1100 гПа.;
- Сенсор освітлення: від 0 до 40 000 лк.;
- Сенсор відносної вологості: від 0 до 100%.

Час, за який мобільний пристрій може оптимально отримати один пакет даних від всіх сенсорів, складає 0.5 секунди. Маючи ці обмеження, може виникнути потреба розширити діапазон можливих значень та зменшити інтервал отримання даних, щоб отримати більш детальну інформацію.

Для збільшення діапазону можливих значень можна використати деякі аналітичні операції, а для того, щоб визначити значення, які були між відомими моментами часу, можна використати лінійну інтерполяцію. Таким чином, це може допомогти в обходженні обмежень з діапазоном максимально можливих значень з сенсорів та в збільшені їх інформативності.

Інша проблема, яка виникає під час збору даних з сенсорів – велика кількість шуму, яка заважає визначенню поточної активності мобільного пристрою. Для її вирішення, використовуються такі методи, як метод ковзаючого середнього та фільтр нижніх частот.

Метод середніх значень є одним з найпростіших методів фільтрації шуму. При використанні методу ковзаючого середнього були отримані помітні покращення в очищені даних від шуму, які дозволили побачити на графіку значень деталі, які було неможливо визначити через надмірну кількість шумів.

Фільтри нижніх частот - це група фільтрів, основною особливістю яких є здатність фільтрувати сигнали вище зазначеної частоти, тобто такі фільтри пропускають сигнали низької частоти, що дозволяє позбутися від шумових перешкод сигналу. Фільтр нижніх частот пригнічує сигнали вище деякої критичної частоти w і пропускає сигнали нижче цієї

частоти. Використання фільтру нижніх частот дає так само більш гладкий результат, як і при згладжуванні методом ковзаючого середнього.

Для виявлення та аналізу різких змін даних з сенсорів використовується наступний алгоритм:

1. Очищення від шуму. Шум виникає внаслідок неосновних рухів або невизначеного розташування мобільного пристрою. Можливе очищення шуму за допомогою додаткової обробки даних сенсорів за допомогою додаткового додатку у смартфоні або на сервері.
2. Виділення ознак. На цьому етапі потрібно виділити ознаки, значення яких відрізняються у об'єктів різних класів. Якщо розробляється алгоритм для класифікації в реальному часі, то бажано обчислювати прості ознаки, оскільки існують обмеження на обчислювальну швидкодію і час роботи алгоритму.
3. Відбір ознак. У реальних задачах часто виходить, що зайвих ознак виявляється істотно більше, ніж корисних. Спроба побудувати залежність з шуму може тільки погіршити якість алгоритму. Методи навчання повинні відрізняти шумові ознаки від інформативних і відкидати їх.

Таким чином, було розглянуто важливість та основні цілі визначення поточної активності користувача, методи для отримання більш детальних даних з сенсорів мобільного пристрою та їх очищення від шуму. В майбутньому, використовуючи попередні методи, планується провести дослідження на інших наборах даних та використати інші доступні сенсори мобільних пристроїв.

ЛІТЕРАТУРА

1. Automatically Detecting Phone Context through Discovery http://miluzzo.info/pubs/miluzzo-phonesense10.pdf
2. Phone Position/Placement Detection https://www.researchgate.net/publication/283104570_Phone_positionplacement_detection_using_accelerometer_Impact_on_activity_recognition

Шевченко В.Л. *д.т.н., професор*
Ільїн М. *студент 4 курсу*
Василенко В. *студент 4 курсу*
кафедра Програмних систем і технологій
Київський національний університет імені Тараса Шевченка, м. Київ, Україна

ПРОЕКТ PEERGAZE. КОМПОНЕНТ МОДЕЛЮВАННЯ ТА КЕРУВАННЯ МОБІЛЬНИМ ПРИСТРОЄМ

Основні принципи роботи компоненту

Після отримання даних щодо відстеження напрямку погляду користувача, вони передаються до компоненту моделювання та керування пристроєм.

Вся система цього компоненту спроектована на принципах реактивного програмування, тому є дуже швидкодіючою та оптимальною як на етапі супроводу системи, так і на етапі розробки.

Для моделювання поведінки напряму погляду користувача, відстеження точки фокусу погляду, роботи системи «жестів поглядом» та надання команд пристрою було обрано фреймворк Unity.

Хоча Unity вважається досить «важким» фреймворком для розробки та моделювання 3D контенту, його було обрано завдяки обсягу підтримуваних платформ, великому набору інструментів для моделювання та легкій інтеграції нативного коду.

Завдяки оптимізації нашого алгоритму відслідковування напряму погляду, ми можемо з легкістю використовувати цей фреймворк одночасно з системою відслідковування напряму погляду у реальному часі, що до цього часу не було реалізовано в жодному проекті.

Зв`язок системи з нативним компонентом

Для обміну даними необхідно будо реалізувати інтерфейс, який дозволить передавати компонентові актуальні дані про стан напрямку погляду та обличчя користувача у кадрі.

Для цього було розроблено інтерфейс на C, що дозволив нативному коду iOS викликати методи компонента моделювання та передавати актуальні дані.

Компонент приймає дані щодо напрямку погляду у вигляді двох нормалізованих векторів (x1, y1, z1) та (x2, y2, z2) для лівого та правого ока. Також система відстеження погляду передає координати положення обличчя в тривимірному просторі відносно камери пристрою у вигляді (x3, y3, z3).

Окрім цього, також між системами передається службова інформація про стан компонент, та команди запуску\відключення систем.

Принцип роботи відстеження точки фокусу погляду користувача.

У першу чергу, отримані дані від компоненту обробляються для побудови векторів напрямку у моделюванні фреймворку Unity завдяки вбудованій структурі **Vector3.**

Потім будується серединний вектор, для визначення вектору погляду з обох очей, бо дані про напрямок погляду можуть бути неточними. Графічно це показано на Рис. 1.

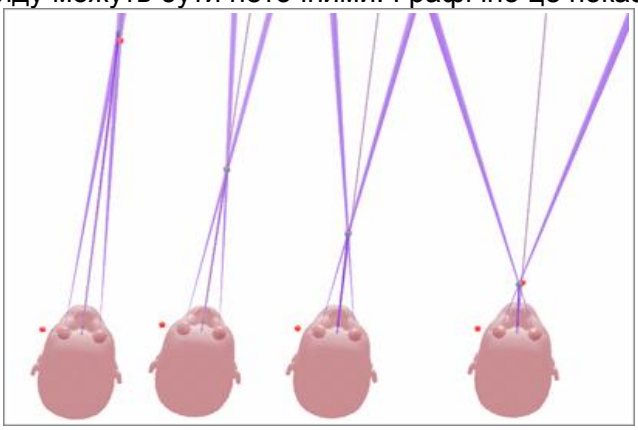


Рис. 1. Побудова серединного вектору напрямку погляду

Визначивши відстань обличчя користувача до екрану по координатам у обличчя у просторі, що передає компонент відслідковування напряму погляду, система буде готова до калібрування.

Система автоматично визначає модель мобільного пристрою запитом до системи, та визначає фізичний розмір його екрану. Якщо користувач захоче калібрувати систему у ручному режимі, то йому буде запропоновано вибрати модель свого пристрою. Так як сьогодні на ринку запропоновано лише три різні версії iPad (за розмірами екранів у них), таких як iPad mini / iPad Air / iPad Pro із розмірами екранів 7,9 / 9,7 та 12.9 дюймів відповідно, то вибір буде запропоновано з них. Також, є можливість задати розміри екрану пристрою вручну.

Після цього для калібрування системи потрібно закріпити планшетний пристрій статично відносно голови користувача, та дивитись у центр мобільного пристрою кілька секунд. За цей час система запам'ятає набір напрямків погляду до центру пристрою та визначить середній з них, який буде опорним.

Для визначення точки фокусу погляду користувача на екрані вимірюється відхилення поточних координат нормалізованого середнього вектору погляду відносно опорного, будується промінь у 3D середовищі, та за умови перетинання його з умовним екраном пристрою, малюється точка фокусу на інтерфейсі

Принцип керування мобільним пристроєм та надання команд

Наступним кроком у розробці проекту було визначення методу розпізнання команд, та найоптимальнішого методу вводу команд поглядом.

Серед усіх підходів до можливого надання команд користувачем було розроблено новий підхід, що впровадив найпродуктивніший метод розпізнання простих рукописних знаків при рукописному вводі в умовах вводу напрямком погляду користувача. Розробка та впровадження такого методу була зумовлена також обмеженими обчислювальними ресурсами системи, бо використовувати ще одну нейронну модель для розпізнання намальованих символів у реальному часі нажаль неможливо.

Алгоритм цього методу є досить простим та ефективним, бо символи які можна малювати очима є досить примітивними.

Графічно алгоритм зображено на Рис. 2.

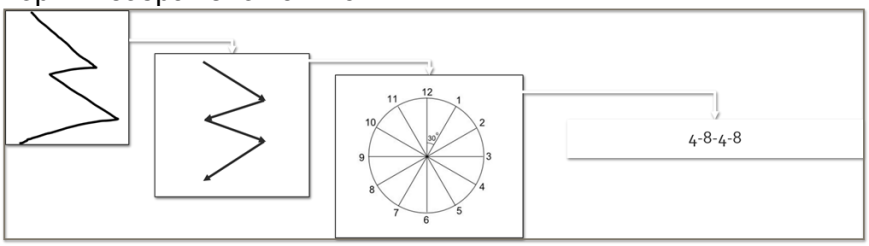


Рис. 2. Графічне зображення алгоритму розпізнання мальованих поглядом символів

Поетапно, алгоритм працює наступним чином:

- Спочатку крива, намальована поглядом розбивається на нормалізовані вектори;
- Координати початку цих векторів приводяться до початку координат;
- Уся координатна площина розбивається на 12 секторів з кутом відхилення у 30 градусів;
- Визначається якому з секторів належить кожен нормалізований вектор;
- Фігура записується у пам'ять як послідовність номерів секторів, та порівнюється з наборами заздалегідь запрограмованих команд;
- Якщо програма знаходить збіг, то виконується потрібна команда.

Для реалізації «Кліку» нами будо запропоновано використовувати різноманітні Bluetooth гарнітури, що дозволить користувачам з обмеженими можливостями без проблем користуватись нашою системою через різноманітність вибору таких приладів

Шевченко В.Л. д.т.н., професор
Ільїн М. студент 4 курсу
Василенко В. студент 4 курсу
кафедра Програмних систем і технологій
Київський національний університет імені Тараса Шевченка, м. Київ, Україна

ПРОЕКТ PEERGAZE. КОМПОНЕНТ РОЗПІЗНАВАННЯ НАПРЯМУ ПОГЛЯДУ КОРИСТУВАЧА

Етапи розпізнавання обличчя у кадрі у реальному часі

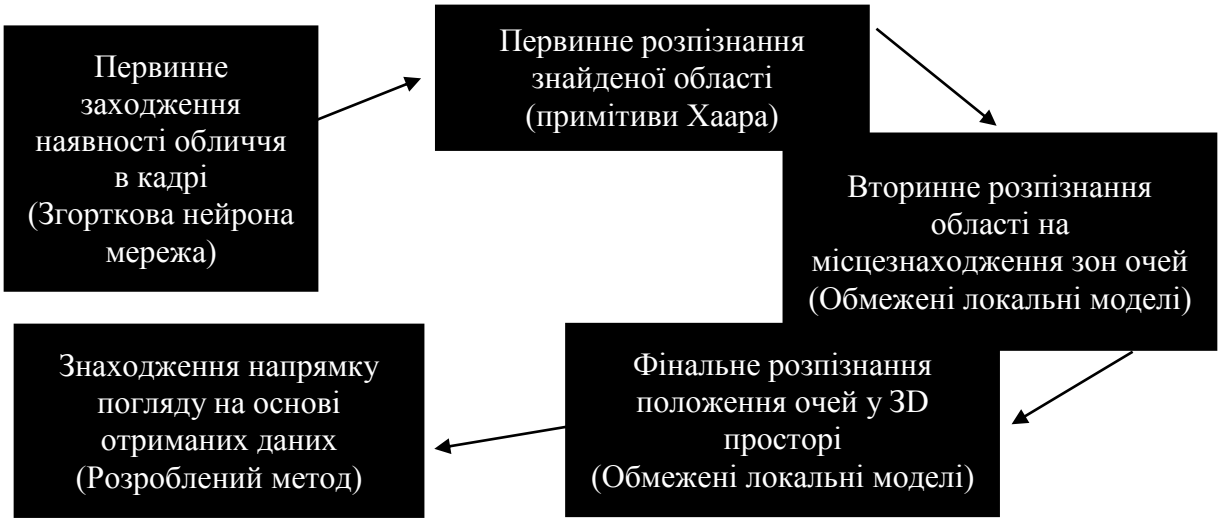


Рис 1. Схема етапів розпізнавання напряму погляду

Етап 1. Знаходження обличчя в кадрі

Цей метод називається методом Віоли-Джонса (так само відомий як каскади Хаара). Для того щоб реалізувати пошук обличчя різних розмірів використовується метод ковзаючого вікна. Саме всередині цього вікна і вираховуються примітиви. Вікно як би ковзає по всьому зображенню, та після кожного проходження зображення вікно збільшується, щоб знайти обличчя більшого масштабу

Етап 2. Знаходження області очей

Розпізнавання рис обличчя є наступним кроком роботи компоненту. Ця проблема розпізнавання обличчя має велику кількість рішень та є цікавою для вдосконалення на розробки нових рішень. До недавнього часу найбільш популярних методів розпізнавання рис обличчя є сукупність рішень під назвою Обмежені Локальні Моделі. Вони моделюють появу кожного орієнтира обличчя індивідуально за допомогою локальних детекторів і використовують модель форми для виконання оптимізації. Саме цей підхід має в собі багато переваг та розширень в порівнянні з іншими рішеннями а саме:

- Багатогранні детектори дають змогу ОЛМ справлятися с задачею розпізнавання рис обличчя в кадрі під різними кутами;
- Саме використання накладання 3Д форми та обличчя дозволяє стабільно відслідковувати положення та форму рис обличчя на відео.

Сучасні підходи розпізнавання рис обличчя можуть бути поділені на два основних типи: Працюючих на основі моделей або регрессорів.

Тип з використанням моделей визначається в пошуку кращих параметрів моделі обличчя в порівнянні з тим, що поступає на обробку цьому методу, тобто обличчя в кадрі.

Оригінальні моделі та детектори не мають змоги знаходити точні риси обличчя (тільки площину обличчя). Тому вдосконалення роботи детекторів та самої моделі було досягнуто за допомогою перетренуванні самих детекторів під нові набори даних, а саме набори: LFPW та Helen.

Етап 3. Відстеження напряму погляду.

Як тільки алгоритм визначає розташування очей і зіниці за допомогою моделі, ми використовуємо цю інформацію для обчислення вектора погляду індивідуально для кожного ока. Запускається умовний промінь від розташування камери у тривимірному просторі до центру зіниці в площині зображення і обчислюється його перетин зі оковою сферою. Це дає змогу визначити місце розташування зіниці в 3D координатах камери. Вектор від 3D центру окової сфери до місця розташування зіниці є очікуваним вектором напрямку погляду.

Література

1. Deep Neural Networks for Solution Problems of Recognition and Classification of Images. Victor Sineglazov

2. Rendering of Eyes for Eye-Shape Registration and Gaze Estimation. Erroll Wood, Tadas Baltrus˘aitis

3. P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Ku- mar. Localizing parts of faces using a consensus of exem- plars.

4. Feature Detection and Tracking with Constrained Local Models David Cristinacce, Tim Cootes

Шевченко В.Л. *Доктор технічних наук, професор,*
професор кафедри програмних систем і технологій
Петрівський В.Я., *Аспірант*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка
м. Київ, Україна

ПОРІВНЯЛЬНИЙ ОПИС ЗВУКОВИХ СЕНСОРІВ INTERNET OF EVERYTHING

Серед ключових характеристик для датчиків шуму виділяють чутливість та частотний діапазон. З наявних на ринку датчиків варто виділити наступні моделі Trema-module v2 та Waveshare Sound Senson. Порівняння характеристик даних датчиків [1, 2] наведено нанаступних гістограмах:

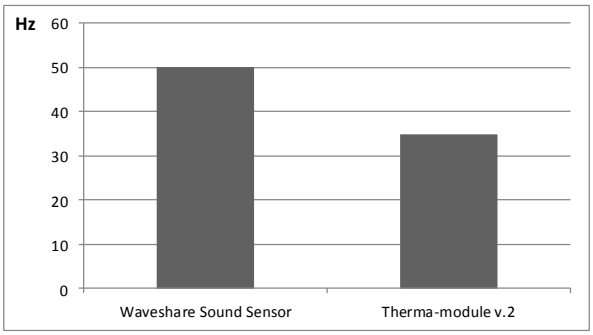


Рис. 1. Мінімальна частота.

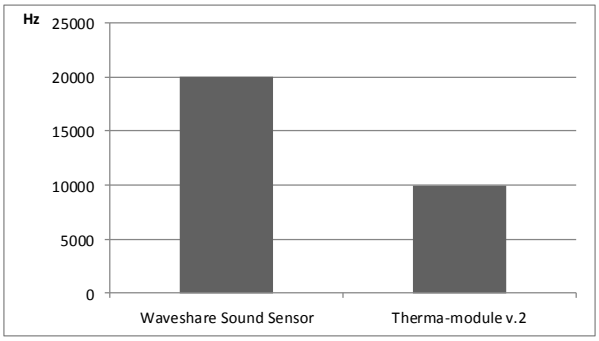


Рис. 2. Максимальна частота.

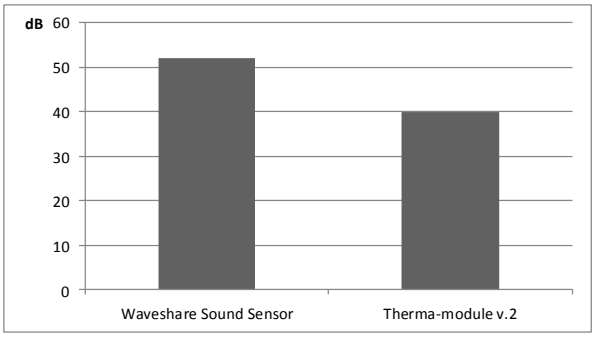


Рис. 3. Чутливість датчиків.

Сучасні мобільні пристрої(смартфони та планшети) на даному етапі розвитку технологій містять вбудовані датчики різного призначення[3], до яких відносяться сенсори руху(акселерометри, сенсори сили тяжіння, гіроскопи, сенсори обертання), сенсори навколишнього середовища(барометри, фотометри, термометри, мікрофон), сенсори положення(сенсори орієнтації, магнітометри, GPS-сенсор).

Аналогічно до випадку зі стаціонарними датчиками датчики звуку мобільних пристроїв мають ряд ключових параметрів. До них відносяться чутливість, сила звуку у акустичних децибелах та звуковий тиск. На світовому ринку серед великої кількості датчиків звуку для мобільних пристроїв великим попитом користуються моделі ICS-52000, ICS-51360, ICS-41350, INMP621, ICS-41352. Порівняльна характеристика параметрів[4] зображена на наступних рисунках.

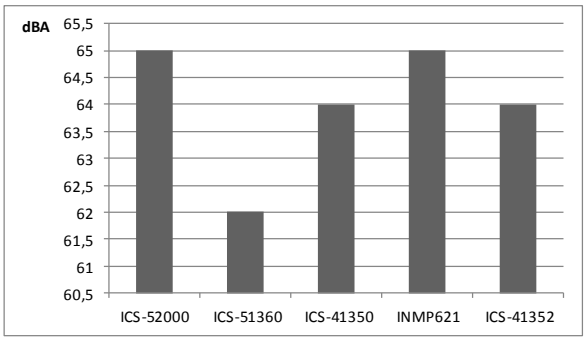


Рис. 4. Сила звуку.

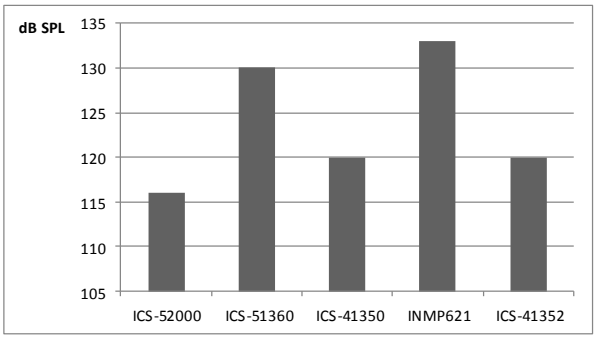


Рис. 5. Звуковий тиск.

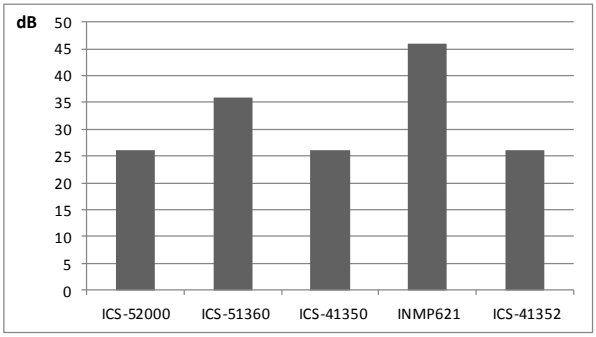


Рис. 6. Чутливість датчиків.

ЛІТЕРАТУРА

1. Waveshare. Sound Sensor [Електронний ресурс] / Waveshare – Режим доступу до ресурсу: https://www.waveshare.com/product/modules/sensors/fingerprint-sound/sound-sensor.htm.

2. Датчик звука (Trema-модуль v2.0) [Електронний ресурс] – Режим доступу до ресурсу: https://wiki.iarduino.ru/page/datchik-zvuka/.

3. From Machine-to-Machine to the Internet of Things. Introduction to a New Age of Intelligence., 2006. [Электронный ресурс] – Режим доступу до ресурсу: http://www.mforum.ru/news/article/110233.htm.

4. Products TDK [Електронний ресурс] – Режим доступу до ресурсу: https://www.invensense.com/products.

Шевченко В.Л. *Доктор технічних наук, професор,*
професор кафедри програмних систем і технологій
Фурса О.Е. *Студент 4-го курсу,*
Факультет Інформаційних Технологій
Київський національний університет ім.Тараса Шевченка, Київ
Кінзерський Д.С. *Студент 4-го курсу,*
Факультет Інформаційних Технологій
Київський національний університет ім.Тараса Шевченка, Київ

ТЕХНІКИ МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ЦІН АКЦІЙ: ФУНКЦІЇ ІНДИКАТОРІВ ТА АНАЛІЗ НОВИН.

Основний принцип полягає в моделюванні економічних показників всіх видів для можливості прогнозування базового активу. Для вирішення такого завдання потрібно сформулювати вимогу до емпіричний оцінки ефективності показників. Іншими словами-оцінка вплив показників на актив.

Мета наукової роботи - підвищення ефективності машинного навчання для прогнозування цін акцій на ринку.

Об'єкт досліджень – створення алгоритму навчання в процесі досліджень.

В роботі розроблено алгоритм машинного навчання на базі моделей економічних показників активів; для описуваного проекту в якості головного індикатора був обраний EMA, він дозволяє обробляти практично необмежений обсяг історичних даних, що дуже важливо для аналізу за допомогою тимчасових рядів. Були вирішені такі питання:

- Знаходження оптимального рішення щодо техніки машинного навчання
- Вибір між використанням алгоритмів машинного навчання.

Більш детально перелік здобутків при розробці системи такий:

- Отримали ***результати оцінок*** точності прогнозування ціни.
- Використали алгоритм AdaBoostM1 для суттєвого поліпшення точності

Для описуваного проекту в якості головного індикатора був обраний EMA - він дозволяє обробляти практично необмежений обсяг історичних даних, що дуже важливо для аналізу за допомогою часових рядів. Однак варто зауважити, що використання інших індикаторів може приносити і більшу точність прогнозів аналізованих акцій.

EMA (t) = EMA (t-1) + alpha * (Price (t) - EMA (t-1))

де, alpha = 2/ (N+1), в такому разі, for N=9, alpha = 0.20.$f\ \ x\ \ = q * y(x)$

В теорії, проблема передбачення цін акцій може буте розглянута, як оцінка функції F у часі T на основі попередніх значень F у часі t-1,t-2 … t-n, присвоюючи відповідну вагову фукнцію w на кожний момент F:

F (t) = w1*F (t-1) + w2*F (t-2) + … + w*F (t-n)

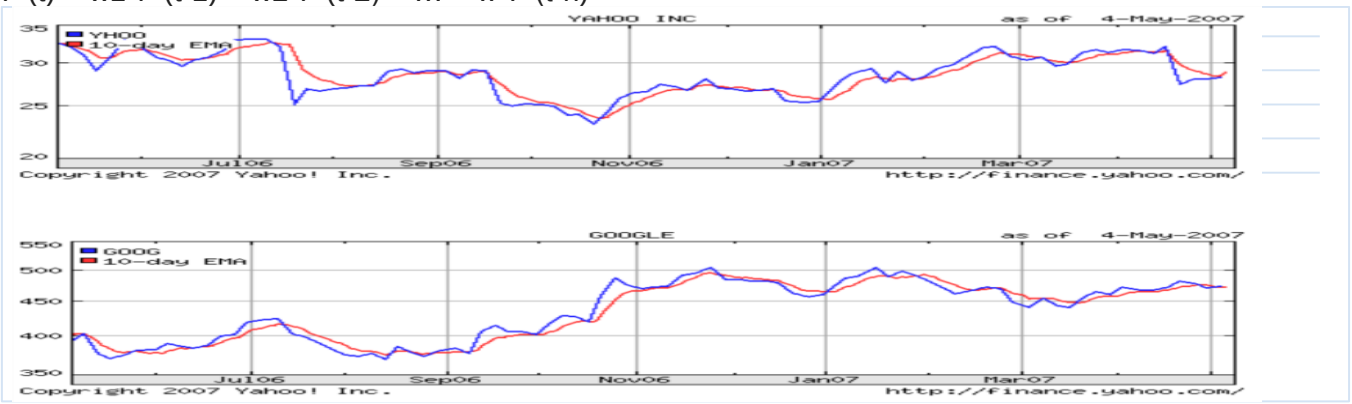


Fig.1. EMA model of current pricing

Алгоритм Decision Stump

Застосування простого алгоритму для передбачення EMA дозволило домогтися наступних результатів:

Коефіцієнт кореляції 0.8597
Середня абсолютна похибка 46.665
Корінь середньоквадратичної похибки 57.8192
Відносна абсолютна похибка 46.8704 %
Корінь середньоквадратичної відносної похибки 50.9763 %
Загальна кількість періодів 681

Линійна регресія

Застосування простої лінійної регресії (використовуються тільки чисельні атрибути) для передбачення EMA дали наступні результати: Коефіцієнт кореляції 0.9591

Середня абсолютна похибка 12.9115
Корінь середньоквадратичної похибки 32.0499
Відносна абсолютна похибка 12.9684 %
Корінь середньоквадратичної відносної похибки 28.2568 %
Загальна кількість періодів 681

Метод опорних векторів

Використання методу опорних векторів (C-Class) з прміененіе радіальної базисної функції ядра з параметром настройки C в діапазоні від 512 до 65536, дозволило отримати наступну точність прогнозування руху ціни:

Корінь середньоквадратичної похибки: 0.486 ± 0.012 Точність: 60.20 ± 0.49%

Бустинг

Після використання алгоритму C-SVC до набору даних був застосований алгоритм бустінга AdaBoostM1 - це дозволило добитися серйозного поліпшення точності.

Корінь середньоквадратичної помилки: 0.467 ± 0.008Точность: 64.32% ± 3.99%

Шевченко В.Л. *Доктор технічних наук, професор,*
професор кафедри програмних систем і технологій
Фурса О.Е. *Студент 4-го курсу,*
Факультет Інформаційних Технологій
Київський національний університет ім.Тараса Шевченка, Київ
Кінзерський Д.С. *Студент 4-го курсу,*
Факультет Інформаційних Технологій
Київський національний університет ім.Тараса Шевченка, Київ

АЛГОРИТМ РОЗПІЗНАВАННЯ ІНСАЙДЕРСЬКИХ УГОД ЗА ДОПОМОГОЮ ТЕХНОЛОГІЙ DATA MINING

Основний принцип полягає в моделюванні економічних показників всіх видів для можливості прогнозування базового активу. Для вирішення такого завдання потрібно сформулювати вимогу до емпіричний оцінки ефективності показників. Іншими словами-оцінка вплив показників на актив.

Мета наукової роботи – розпізнавати інсайдерські угоди за допомогою технік машинного навчання.

Об'єкт досліджень – створення алгоритму розпізнавання в процесі досліджень.

В роботі розроблено алгоритм за допомогою застосування технологій Data Mining. В набір даних увійшла інформація про більш ніж 12 млн угод на американських біржах, які були здійснені більш ніж 370 тисячами інсайдерів з 1986 по 2012 рік. База зберігалася в SQLite, загальний обсяг даних склав 5,61 гігабайт.

Знаходження оптимального рішення щодо техніки машинного навчання

Розуміння ризиків – було проведено дослідження, якы довели, що ті керівники, які займаються купівлею і продажем на біржі акцій своєї компанії, в кінцевому підсумку починають приймати такі рішення по своїй основній роботі, які диктуються їх покупками або продажами на фондовому ринку.

Виявлення шахрайства та незаконних угод - застосування технік дата Майнінг дозволяє визначать можливі інсайдерські операції, метою яких є обман інвесторів компанії.

Нижче показано кумулятивний розподіл для числа компаній, до яких належать інсайдери і числа транзакцій ними скоєних. Як правило, інсайдер має відношення до невеликої кількості компаній і робить не особливо великі угоди, але є і невелика кількість людей, які пов'язані з багатьма організаціями і здійснюють об'ємні угоди.

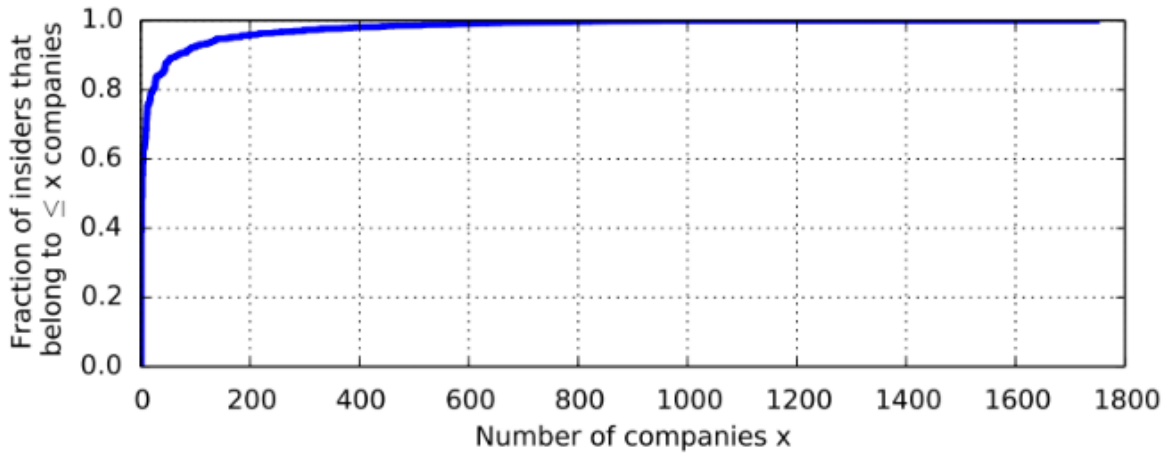


Рис. 1: Функція емпіричного кумулятивного розподілення числа компаній

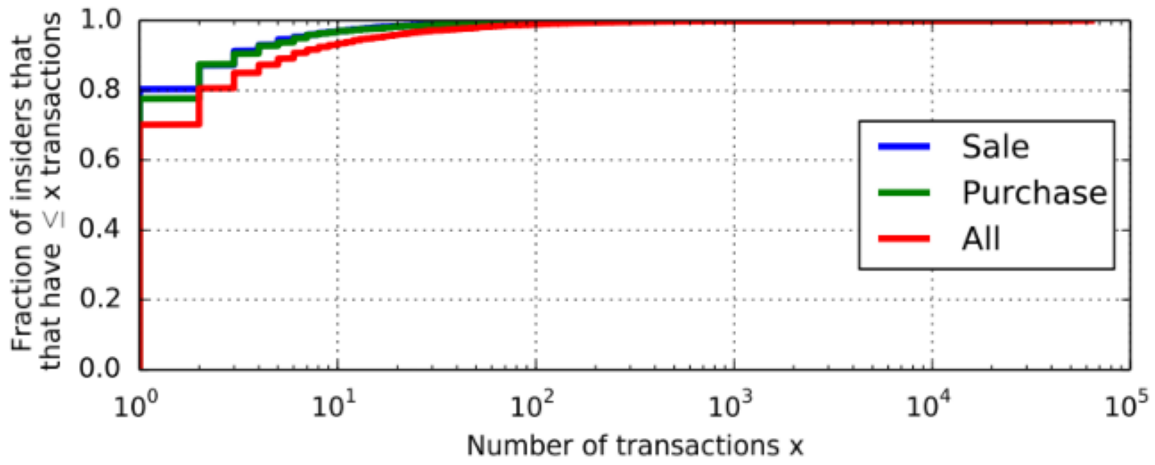


Рис. 2: Функція емпіричного кумулятивного розподілення числа здійсненних інсайдерами транзакцій (вісь X має логарифмічний масштаб)

Аналіз типів транзакцій дозволяє виявити деякі цікаві патерни. Наприклад, інсайдери частіше продають акції, а не купують - це пов'язано з тим, що багато керівників отримують акції за свою роботу за допомогою, наприклад, опціонів в компанії. Тому інсайдери частіше продають акції, щоб збалансувати.

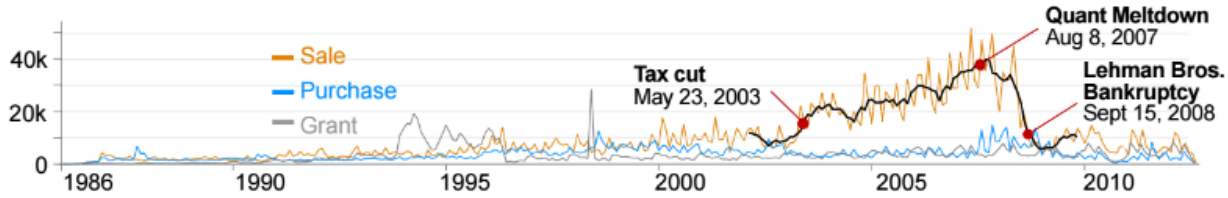


Рис. 4: Графік розподілу транзакцій різного типу (Sell - продаж, Purchase - покупка, Grant - отримання акцій інсайдерів за допомогою реализаци, наприклад, опціону).

Skovorodin Mykhailo
Kalyshka Dmytro
3-th course bachelor students,
Information technology faculty,
Taras Shevchenko National University of Kyiv

DETECTION AND ANALYSIS OF VULNERABILITIES IN WEB APPLICATIONS

IDENTIFICATION - is a statement about who you are. Depending on the situation, it can be a name, email address, account number, and so on.

AUTHENTICATION - providing evidence that you really are the one who identified (from the word “authentic” - true, authentic).

AUTHORIZATION - check that you are allowed to access the requested resource.

These terms are used in computer systems, where traditionally, identification means getting your account (identity) by username or email; under authentication — verification that you know the password for this account, and under authorization — verification of your role in the system and the decision to grant access to the requested page or resource.

Password authentication

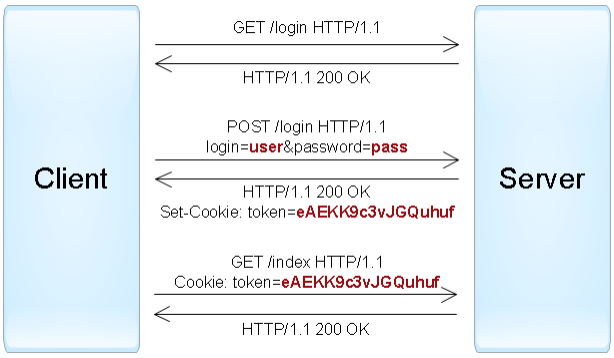
This method is based on the fact that the user must provide username and password for successful identification and authentication in the system. The username / password pair is set by the user upon his registration in the system, while the username may be the user's email address.

Basic is the simplest scheme in which the username and password of a user are transmitted in the Authorization header in an unencrypted form (base64-encoded). However, using the HTTPS (HTTP over SSL) protocol is relatively secure.

Forms authentication

There is no specific standard for this protocol, so all its implementations are specific for specific systems, and more specifically, for development framework authentication modules.

It works according to the following principle: an HTML form is included in the web application, in which the user must enter his username / password and send them to the server via HTTP POST for authentication. If successful, the web application creates a session token, which is usually placed in browser cookies. In subsequent web requests, the session token is automatically transmitted to the server and allows the application to obtain information about the current user to authorize the request.



Common vulnerabilities and implementation errors

Password authentication is considered not a very reliable way, since passwords can often be picked up, and users tend to use simple and identical passwords in different systems, or write them on pieces of paper. If the attacker was able to figure out the password, the user often does not know about it. In addition, application developers can make a number of conceptual errors that simplify hacking accounts.

Below is a list of the most common vulnerabilities when password authentication is used:

• Web application allows users to create simple passwords.

• The web application is not protected from brute-force attacks.

• The web application itself generates and distributes passwords to users, but does not require changing the password after the first login (i.e., the current password has been recorded somewhere).

• The web application allows passwords to be transmitted over an unprotected HTTP connection or in the URL string.

• The web application does not use secure hash functions for storing user passwords.

• The web application does not provide users with the ability to change the password or does not notify users about changing their passwords.

• The web application uses a vulnerable password recovery feature that can be used to gain unauthorized access to other accounts.

• The web application does not require re-authentication of the user for important actions: changing the password, changing the delivery address of goods, etc.

• The web application creates session tokens so that they can be matched or predicted for other users.

• The web application allows the transmission of session tokens over an unprotected HTTP connection, or in the URL string.

• The web application is vulnerable to session fixation attacks (that is, it does not replace the session token when the anonymous session of the user goes to the authenticated session).

• The web application does not set the HttpOnly and Secure flags for browser cookies containing session tokens.

• The web application does not destroy the user's session after a short period of inactivity or does not provide a function to exit the authenticated session.

Literature:

1. www.google.com
2. www.wikipedia.org
3. blogs.cisco.com

***Ткаченко М.В.**, к.т.н,*
асистент
***Федоренко Р.М.**, к.е.н.,*
асистент
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна.

РОЗПІЗНАВАННЯ ОБРАЗІВ ЗА ДОПОМОГОЮ ПАКЕТУ TENSORFLOW LITE НА МОБІЛЬНИХ ПЛАТФОРМАХ ТА У ВБУДОВАНИХ СИСТЕМАХ

Бібліотека машинного навчання TensorFlow вже працює на величезній кількості платформ: від стійок серверів до крихітних пристроїв Інтернету речей. Але так як впровадження моделей машинного навчання за останні кілька років має експоненціальний характер зростання, вже потрібно їх розгортання і на мобільних пристроях та у вбудованих системах. Саме тому Google представила нову версію бібліотеки – TensorFlow Lite, яка спеціально розроблена саме для невеликих пристроїв. Вихід нової версії бібліотеки був анонсований на конференції Google I/O 2017. Серед основних переваг Lite-версії варто виділити:

– малий об’єм: забезпечує швидку ініціалізацію і запуск моделей машинного навчання невеликого розміру на мобільних пристроях;

– можливість перенесення на будь-які платформи: навчання моделей можливо на великій кількості мобільних платформ (Android, iOS) та у вбудованих системах;

– швидкодія: бібліотека оптимізована для використання на мобільних пристроях та у вбудованих системах, підтримує апаратне прискорення.

На даний час все більше мобільних пристроїв використовують спеціально розроблені апаратні засоби для більш ефективної обробки робочих навантажень машинного навчання. TensorFlow Lite підтримує нейромережевий API (Android Neural Networks API) для прискорення машинного навчання на мобільних пристроях. Архітектуру TensorFlow Lite наведено на рисунку. Розглянемо кожен компонент окремо:

– TensorFlow Model: навчена модель TensorFlow, що збережена на диску;

– TensorFlow Lite Converter: програма, що конвертує модель у формат TensorFlow Lite;

– TensorFlow Lite Model File: формат файлу моделі на основі FlatBuffers, що оптимізовано для максимальної швидкості таі мінімального розміру.

Після проміжних кроків із переведення моделі у потрібний формат, файл із розширенням .tflite розгортається у мобільному додатку для однієї з ОС:

– Java API: оболонка навколо C ++ API для ОС Android;

– C++ API: завантажує файл моделі TensorFlow Lite та викликає інтерпретатор;

– інтерпретатор: запускає модель за допомогою набору операторів, підтримується вибіркове завантаження операторів;

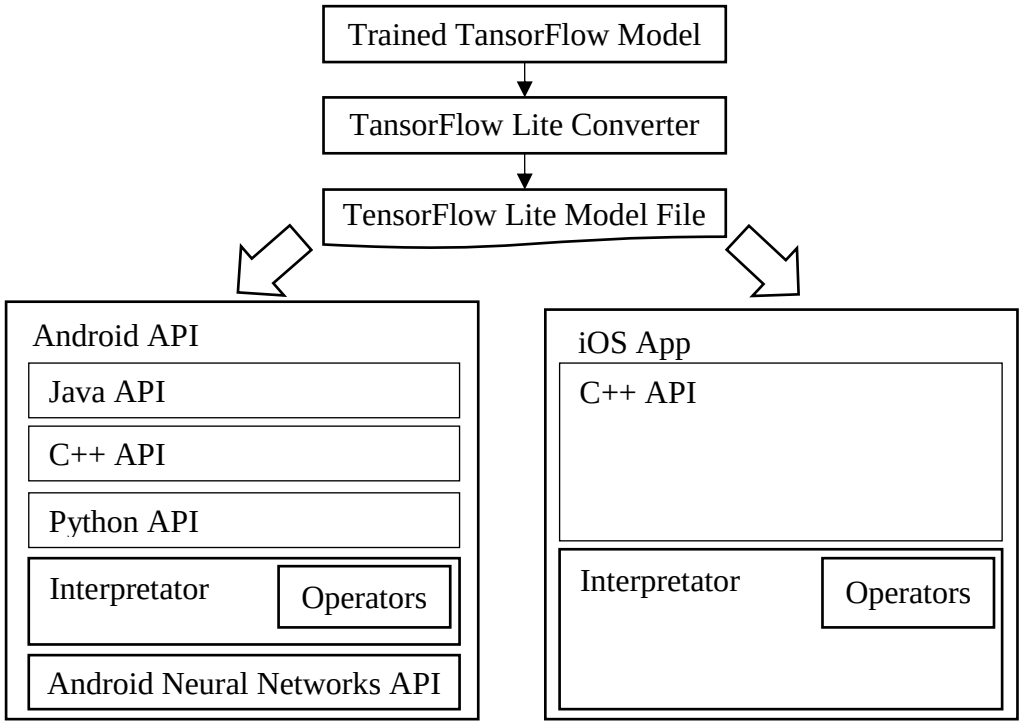
– на деяких пристроях інтерпретатор буде використовувати Android Neural Networks API для апаратного прискорення.

Основою для обчислень є граф, під яким розуміється структура даних, що у повному обсязі описує обчислення. При цьому досягаються наступні переваги:

– можливість перенесення, оскільки граф може бути виконаний негайно або збережений для використання пізніше, і він може працювати на різних платформах: процесорах, графічних процесорах, мобільних та вбудованих пристроях. Крім того, він може бути розгорнутий без залежності від будь-якого коду, який побудував граф;

– він трансформується і оптимізується, оскільки граф можна перетворити, щоб створити більш оптимальну версію для даної платформи. Крім того, можуть бути проведені оптимізація пам'яті або обчислень і знайдені компроміси між ними. Це корисно, наприклад, при підтримці більш швидкого мобільного виведення після навчання на більших машинах;

– підтримка розподіленого виконання.



На цей час вже розроблені певні моделі – наприклад, MobileNet, яка призначена для класифікації зображень та являє собою глибоку нейронну мережу, яка навчена розрізняти 1000 різних класів зображень. Моделі складаються із файлів двох типів: файл конфігурації моделі, що зберігається в форматі JSON, та ваги моделей, збережені в двійковому форматі. Для класифікації використовується трикольоровий канал (RGB), альфа-канал ігнорується. Цей код перетворює двоканальний масив у вірну трьохканальну версію. Модель класифікує зображення висотою та шириною по N-пікселів. Вхідні тензори повинні містити значення з плаваючою точкою в діапазоні від -1 до 1 для кожного з трьох канальних значень кожного пікселя. На виході моделі даються імовірнісні значення, що характеризують шанси знайти дані об'єкти на зображенні, що класифікується.

Отже, бібліотека TensorFlow Lite відкриває перед розробниками можливості глибокого навчання та надає змогу легко надбудовувати в мобільних додатках та вбудованих системах нові можливості для вирішення складних завдань машинного навчання з мінімальними зусиллями та лаконічним кодом.

ЛІТЕРАТУРА

1. Ron Bekkerman, Mikhail Bilenko, and John Langford.Scaling Up Machine Learning:Parallel and Distributed Approaches. New York, NY, USA: Cambridge University Press, 2011.

2. Jeffrey Dean et al. “Large Scale Distributed Deep Networks”. In:Advances in NeuralInformation Processing Systems 25. Ed. by F. Pereira et al. Curran Associates, Inc., 2012, pp. 1223-1231.

3. Xavier Glorot and Yoshua Bengio. “Understanding the difficulty of training deep feed-forward neural networks”. In: vol. 9. 2010, pp. 249-256.

4. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet Classification withDeep Convolutional Neural Networks”. In:Advances in Neural Information ProcessingSystems 25. Ed. by F. Pereira et al. Curran Associates, Inc., 2012, pp. 1097-1105.

***Ткаченко М.В.**, к.т.н,*
асистент
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка
м. Київ, Україна,
***Федоренко Р.М.**, к.е.н.,*
асистент
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка
м. Київ, Україна,
***Федоренко Т.О.**,*
аспірант,
кафедра економічної теорії
Київський національний економічний університет імені Вадима Гетьмана
м. Київ, Україна,

SEMANTIC WEB: МЕТОДИ ЗВ'ЯЗУВАННЯ ДАНИХ

Ключові слова: Semantic Web, опис властивостей, пошуковий запит, пошукова система, правила обробки даних, семантичні зв’язки, семантична розмітка, формування бази даних.

Виконано огляд методів та основних принципів зв’язування даних, що використовуються для структурування та систематизації веб-даних. Semantic Web або Семантичний Web – це глобальна мережа взаємопов’язаних даних у форматі для машинного читання. Усі неструктуровані дані мають бути об’єднані у єдину базу даних. Семантична мережа побудована на синтаксисі RDF, що регулює правила обробки даних, використовує протокол SPARQL для здійснення запитів і звертається до словників OWL [1].

Основою для взаємозв’язку даних є загальна схема опису ресурсів RDF, яка має стандартний формат на мові XML і базується на ідентифікаторах URI. Метою RDF є побудова взаємозв’язку веб-даних, на основі трьох компонентів – ресурсів (веб-адреса), атрибутів (назва, властивість, ознаки об’єкта) і їх значення. При цьому, для відображення семантичних зв’язків потрібні наступні компоненти: словники (збірники термінів), таксономії (ієрархічно побудовані словники термінів), онтології (концепції та зв'язки).

Для більш детального опису взаємозв’язків між ресурсами використовується словник RDF Schema (RDFS), який вводить поняття класу, виділяє типи даних (у тому числі строки із зазначенням мови XML або HTML-розміткою), визначає класи, підкласи ресурсів, властивості, необхідні для запису трійок (Triplets).

Додатково RDF звертається до інших словників, зокрема словник контактів людей, соціальних мереж FOAF (Friend of a friend), Dublin Core (відео, малюнки, книги, CD), Schema.org, SKOS. В результаті система отримує більш гнучкі запити для отримання інформації. Зазначена технологія використовується для таких цілей [2]:

- формування баз даних в електронних бібліотеках, університетах;
- для більш детального опису результатів пошуку, Google використовує семантичний опис сторінок в RDF-форматі у вигляді JSON-LD (зі словником Schema.org) [3];
- для опису вмісту, рейтингу веб-зображень, графіків часу для веб-подій;
- для опису властивостей сторінок соціальні мережі (Facebook (opengraph);
- для короткого представлення нових записів на сайтах новин, і-журналів, в блогах (у формі RSS 1.0);
- дані для аналізу та/або інтеграції в інформаційних системах підприємств (SPARQL).

Мова опису онтології – ще один компонент побудови зв’язків між даними на основі описової логіки. Метою застосування онтології є класифікація елементів з точки зору

семантики, сенсу. При цьому, порівняно з RDF, OWL має більшу інтерпретаційну здатність і більший словник.

OWL має три модифікації –Lite (надає просту класифікацію), DL (з повним доступом, але деякі функції обмежені), Full (повний). Кожна наступна модифікація є логічним розширенням попередньої. Таким чином, OWL формалізує веб-інформацію, надає можливості створювати нові твердження на базі вже існуючих, а також виводити нові залежності і зв'язки між фактами на основі відомих. З OWL комп'ютери матимуть можливість інтегрувати інформацію з Інтернету [4].

Отже, метою реалізації семантичної розмітки є зручність і простота використання інформації, всі подорожі та ділові зустрічі в єдиному форматі та місці. Ці технології корисні для покупців і продавців інтернет магазинів, для онлайн торгівлі, малий бізнес матиме більший ступінь захисту та автоматизації транзакцій.

SPARQL – мова запитів до даних, представлених у RDF, а також протокол для передачі цих запитів та відповідей на них. SPARQL підтримує чотири загальнодоступних формати обміну результатів запитів в машиночитаній формі, а саме: XML (Extensible Markup Language), JSON (JavaScript Object Notation), CSV (comma-separated values), TSV (tab-separated values) [5].

Запити SPARQL надають один або декілька шаблонів щодо зв'язку між ресурсами. Використання SPARQL дозволяє отримувати більш складну інформацію (посилання на зв'язки між ресурсами, табличні дані). Як наслідок, отримуємо потужний інструмент для пошукових систем, які включають дані, що походять з Semantic Web.

Формат взаємодії правил RIF є ще одним компонентом інфраструктури семантичного веб за специфікацією W3C. Формат RIF є набором діалектів, який дозволяє обмінюватися правилами між різними системами правил. RIF визначає механізми формального подання синтаксису логічних діалектів, використовується для визначення семантики діалектів та демонстрації зв'язку. Головною метою є полегшення обміну правилами за рахунок мережевого ефекту, більш широке застосовування специфікації підвищує ефективність обміну правилами.

Отже, логіка Semantic Web є комбінацією математичних та інженерних рішень для опису складних властивостей об'єктів. Основними вимогами до системи є принцип простоти, щоб система не зависала від неоднозначних запитів. Основним принцип отримання результату є максимізація у вигляді доступності інформації, зручності / корисності, масштабності та ефективності.

ЛІТЕРАТУРА

1. Semantic Web [Електронний ресурс] // W3C – Режим доступу до ресурсу: https://www.w3.org/standards/semanticweb/
2. RDF – Examples of Use [Електронний ресурс] // W3C – Режим доступу до ресурсу: http://w3schools.sinsixx.com/rdf/rdf_intro.asp.htm
3. Data Model [Електронний ресурс] – Режим доступу до ресурсу: https://schema.org/docs/datamodel.html
4. Introduction to OWL [Електронний ресурс] // W3C – Режим доступу до ресурсу: http://w3schools.sinsixx.com/rdf/rdf_owl.asp.htm
5. Query [Електронний ресурс] // W3C – Режим доступу до ресурсу: https://www.w3.org/standards/semanticweb/query

Ковтун О.І., *к ф.-н., доц.*
кафедра програмних систем і технологій
Лещенко О.О.., *к т.н., доц.*
кафедра мережевих та інтернет технологій
Духновська К.К., *асис.*
кафедра прикладних інформаційних систем
Птушкін С.Д. *студент*
кафедра програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка, м. Київ, Україна

Мобільний застосунок для для розв'язання типових задач з комбінаторики

Використання програмного забезпечення з метою навчання набуло розвитку в багатьох передових навчальних закладах сучасного світу. Впровадження подібних заходів є важливим інноваційним кроком для закладів вищої освіти в Україні. Необхідність в подібних програмних застосунках у навчальному процесі сучасних освітніх закладів формує нову нішу для багатьох розробників програмного забезпечення.

Під час виконання цієї роботи спроектовано та розроблено програмну реалізацію кросплатформного мобільного застосунку для розв'язання типових задач комбінаторики [1]-[2], яка працює на платформах Android та iOS, може бути розповсюджена за допомогою QR-коду. Для досягнення поставленої мети були розроблені оптимізовані алгоритми вибору [3] та розв'язання задач комбінаторики.

Створене програмне забезпечення легко масштабується та модифікується. Ця особливість застосунку відкриває велику кількість архітектурних рішень для його зміни відповідно до потреб будь-яких навчальних напрямків. Таким чином створений програмний продукт може бути використаним у якості програмного каркасу засобів вивчення комбінаторики.

Розуміння задач комбінаторики вимагає певних початкових знань від користувача, оскільки переважна більшість кінцевих користувачів застосунку може не мати відповідної кваліфікації, виникає необхідність у спрощенні процесу вибору необхідної задачі комбінаторики.

Для цього було створено алгоритм (рис. 1.1.), що базується на групі запитань до користувача [4]. На основі відповідей на задані запитання система згідно до алгоритму автоматично обирає необхідну задачу комбінаторики.

Рис. 1. Алгоритм вибору задачі комбінаторики

Алгоритм використовує прості відповіді «Так» чи «Ні» для переходу до однієї з результуючих задач комбінаторики: комбінації, комбінації з повтореннями, розміщення,

розміщення з повтореннями, перестановки та перестановки з повтореннями. Таке рішення дозволяє студентам будь-яких напрямів легше розуміти суть та логіку типових задач комбінаторики, не упираючись у математичні формули та складні формулювання.

Після обрання задачі комбінаторики, програмний застосунок [5] автоматично виконує розрахунки на основі попередньо введених даних, виводить на екран математичну формулу, що відповідає розв'язку задачі, кількість розв'язків, випадаючий список з модальним вікном, відповідне посилання на веб-сторінку з докладною інформацією щодо обраної задачі комбінаторики. Таке рішення дозволяє ініціювати роботу системи без необхідності у повторному натисканні кнопок для розрахунку, пришвидшуючи розрахунки, без збільшення навантаження на систему.

Для розробки кросплатформного мобільного програмного забезпечення було обрано технологію React Native [6], що базується на фреймворку React JavaScript. Для тестування, відладки, сборки проекту та його розповсюдження було використано платформу Expo SDK.

Застосунок, отриманий в результаті проведеної роботи, забезпечить студентів та викладачів закладів вищої освіти зручним інструментарієм для проведення розрахунків в області комбінаторики.

ЛІТЕРАТУРА

1. В.А. Вишенський, М.О. Перестюк. Комбінаторика: перші кроки. — Кам'янець-Подільський // Аксіома, 2010. — 324 с.

2. Ландо С. К., Лекции по комбинаторике. — МЦНМО // 1993. – 58с.

3. Pang, Jong-Shi; Olvi Mangasarian, Computational Optimisation - Kluwer Academic Publishers. // 1999. - 273 с.

4. Розробка застосунків для мобільниих пристроїв – Wikipedia, the free encyclopedia - Режим доступу: http://en.wikipedia.org/wiki/ Розробка_застосунків_для_мобільних_пристроїв

5. Building standalone apps – Expo, Develop RN Apps – Режим доступу: https://docs.expo.io/versions/v27.0.0/distribution/building-standalone-apps

6. React Native – GitHub – Режим доступу: https://github.com/facebook/react-native

Адамчук Д. *Студент 3 курсу*
Кафедра інженерії програмного забезпечення
Національний авіаційний університет м. Київ, Україна
Поперешняк С.В. *Доцент, кандидат фізико-математичних наук.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

ОСОБЛИВОСТІ ПОБУДОВИ ТА КЕРУВАННЯ СПЕЦІАЛІЗОВАНОЇ СИСТЕМИ ДЛЯ МАЙНІНГУ

Майнінг, також видобування (від англ. mining — видобуток корисних копалин) — діяльність з підтримки розподіленої платформи і створення нових блоків з можливістю отримати винагороду в формі емітованої валюти і комісійних зборів у різних криптовалютах, зокрема в Біткоїнах.

Вироблені обчислення потрібні для забезпечення захисту від повторного використання одних і тих же одиниць валюти, а зв'язок майнінгу з емісією стимулює людей витрачати свої обчислювальні потужності і підтримувати роботу мереж.

Сам термін Майнінг походить від слова Mine (видобуток). Якщо говорити складними словами, то цей процес являє собою діяльність по підтримці роботи мережі шляхом закриття і створення блоків в Blockchain з використанням обчислювальних потужностей (Рис. 1). Майнер використовує потужності заліза для виконання спеціальних обчислень з пошуку цифрового підпису (хешу), яка закриє блок.

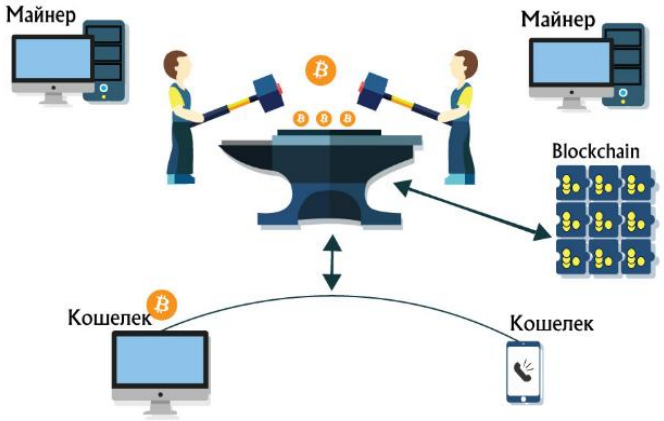


Рис. 1 Схема роботи майнингу

Майнер, який «знайде» цифровий підпис, отримує винагороду у вигляді 1 одиниці криптовалюта. Майнінг підтримує роботу мережі, гарантує її захист від дублюючих транзакцій.

Розглянемо налаштування ферми. Для кожної майнінг ферми потрібно вибрати комплектуючі. Розглянемо з чого складається стандартна майнінг ферма

- SSD-диск (його переваги полягають в стабільності і швидкості роботи);
- процесор;
- Відкрита GPU (оптимально 4-6 шт.);
- Материнська плата;
- Блок живлення (з запасом на 20-30% від максимального споживання всієї системи);
- Райзер (по одному на кожну відеокарту); О
- оперативна пам'ять (достатньо 4 Гб);
- USB WatchDog;
- Емулятор монітора.

На рис. 2 наведено приклад власної ферми для якої були обрані відеокарти GTX1080 MSI.

Рис. 2. Майнінг ферма на відеокарті

Розглянемо недоліки з якими прийшлося зіткнутися.

- Підхід Win 10 до збереження системних налаштувань. Наприклад, відключивши Windows Defender та службу оновлення, але на одній з ригів система все одно постійно завантажувала оновлення, пару раз зависнув при цьому, а Захисник із завидною регулярністю намагався «відрізати» інтернет Клеймору.
- Пару раз TieamViwer не мог з'єднатися з віддаленою машиною, яка просто відмовлялася в авторизації. При цьому сам риг, по показаниям пула, правильно «копал». Відсутність можливості удаленого управління мені сильно напружено!
- Оновлення Драйверів проходить досить довго для оновлення драйверів на відеокарти було витрачено більше години.
- Не можливо відключити оновлення Windows. Оновлення все одно будуть встановлюватись.
- Живлення операційної системи більше 800w.

Майнінг HiveOS

Система представляє собою Linux збірку, засновану на дистрибутиві Ubuntu, яка відразу після розгортання пропонує своїм користувачам ряд приємних моментів:

1. Можливість віддаленого моніторингу та управління ригом через сайт hifeos.farm.
2. Встроенний хешрейт-вотчдог, який перегружає майнер або весь риг при падінні хешрейта. До речі, підтримуються і технічні вотчдоги, в тому числі і популярна рішення від open-dev.ru.
3. Підтримка великого числа майнерів для відеокарт AMD, NVIDIA і навіть для центрального процесора.
4. Можливість запускати та працювати з будь-якими флешками об'ємом не менше 8 Гб краще 16 Гб.

Спочатку було встановлено образ на SSD диск.

Установки як такої не потрібно. Можно просто стартувати клунь з флешки, а Hive OS сама сконфігурує себе в автоматичному режимі. Причому, на відміну від Windows, ми відразу запускалив і налаштовували клунь з встановленими 7-ю картами! Ніяких послідовних додавань в ріг по одній карті, як в Windows, не треба було робити.

В роботі було розглянуто особливості побудови, моніторингу та керування спеціалізованої системи для майнінгу на операційній системі Windows 10 та HiveOS.

Встановити Hive OS на ріг просто і управляти дуже зручно. На випадок будь-яких складнощів у Hive OS є чат підтримки в Telegram в якому можна отримати оперативну допомогу по запуску, налаштування або конфігурації системи.

Залишилося розповісти про ціну. Так, система не безкоштовна. Вірніше, до 3 рігов Ви можете використовувати Hive OS абсолютно безкоштовно , але, починаючи з четвертого рігу, розробник просить оплату за 3 $ за кожен ріг, включаючи перші три. Тобто за 3 ріги Ви не платите нічого, а за 4 ригі вже необхідно платити 12 $ в місяць. Багато це чи мало? Це справедлива ціна за систему, яка динамічно розвивається. Наприклад, буквально днями додали графічну статистику по ригі. Та можливість перегляду затрат на електроенергію

Завдяки можливості працювати з флешки, ми економимо на SSD або жорсткому диску, що, погодьтеся, приємно.

Коротін Д.С. *Студент 3 курсу*
Поперешняк С.В. *Доцент, кандидат фізико-математичних наук.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

ПІДХОДИ ЩОДО ДОЦІЛЬНОСТІ ВИКОРИСТАННЯ СИСТЕМИ DEMAND-SIDE PLATFORM

В статті проведено аналіз доцільності використання системи автоматизованої купівлі / продажу реклами з боку рекламодавців та покупців медіа продукту з метою економії часу та матеріальних ресурсів.

Аналіз процесів купівлі або продажу цифрової реклами протягом останніх 20 років був дорогим і дуже ненадійним. Встановлено, що сучасні Demand-Side Platform (DSP) допомагають зробити цей процес не таким витратним і максимально ефективним, видаляючи на певних етапах фактор людського втручання, а також необхідність безпосереднього узгодження розцінок і пов'язаних з ним “мануальних” операцій [1].

В якості об’єкта дослідження буде розглянуто процес автоматизації взаємодії з Supply Side Platform (SSP), Ad Network, Ad Exchange і, безпосередньо, з паблішерами - DSP система. Головна мета DSP - купити за мінімальною ціною покази користувачам, максимально точно відповідними запитами рекламодавців. Якщо коротко, то DSP дозволяє рекламодавцям купувати рекламні покази на різних сайтах, але при цьому «цілитися» в певних користувачів на основі даних, таких як їх місце розташування і історія поведінки на сайті. Паблішери роблять рекламні покази доступними через рекламні біржі (ad exchanges) і DSP самостійно вирішує, який з цих показів має сенс купити. Часто ціна показу визначається в рамках аукціону в режимі реального часу, відомого як Real Time Bidding (RTB). Це означає, що немає ніякої необхідності в тому, щоб продавці домовлялися про ціни безпосередньо з покупцями, тому як рекламні покази йдуть автоматично. Цей процес займає мілісекунди, поки комп'ютер користувача завантажує веб-сторінку. На рис. 1 представлена блок-схема об’єктної моделі взаємодії DSP-SSP систем компанії EPOM.

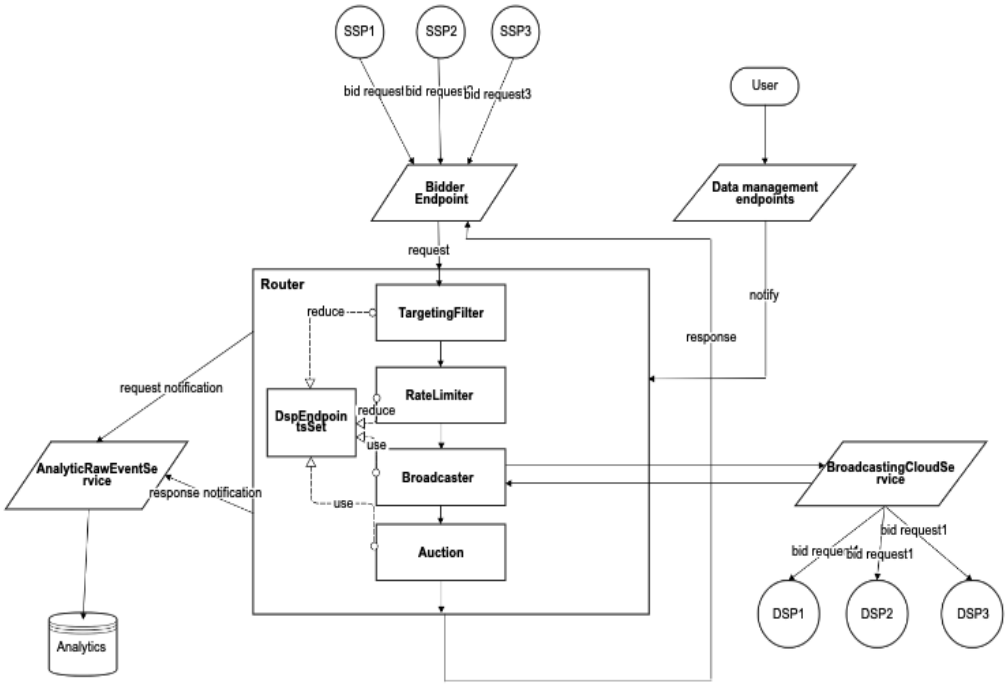


Рис. 1. Блок-схема об’єктної моделі взаємодії DSP-SSP систем

Ця модель показує, що DSP влаштовує аукціон, де за рекламну платформу SSPs виставляють ціну і DSP вибирає найпривабливіший (оптимальний) для нього лот. Після цього аукціону кожний SSP та DSP можуть подивитися аналітику в своєму особистому кабінеті.

В деякій мірі DSP роблять багато з того, що раніше пропонували «ad networks», включаючи доступ до широкого спектру інвентарю і таргетингу. Але їх перевагою, на відміну від рекламних мереж, є можливість купити, показати і відстежити оголошення, використовуючи єдиний інструмент [2]. В результаті це дозволяє досить легко оптимізувати рекламні кампанії. Тут все крутиться навколо даних. Як правило, рекламні мережі ставлять націнку на медійну рекламу, яку вони продають. DSP ж стягують невелику плату за суттєве спрощення транзакції. DSP тепер часто працюють безпосередньо з клієнтами рекламодавця, ефективно замінюючи агентства, коли справа доходить до закупівель медійної реклами. Клієнти кажуть про те, що вони як і раніше співпрацюють з агентствами для розробки стратегій або отримання консультацій, але починають активно звертатися і до третіх сторін, включаючи DSP, які допомагають закуповувати рекламу [3-4].

Supply Side Platform – це платформа, яка торгує рекламним інвентарем або рекламними позиціями інтернет-майданчиків. SSP агрегує пропозиції майданчиків, «збирає» залишковий трафік, а також встановлює таку мінімальну вартість, по якій майданчик готова реалізувати показ. SSP проводить аукціонний торг з DSP, максимально вигідно продаючи інвентар паблішер [3].

DSP не володіють інвентарем, не купують і не продають його, а просто з’єднують паблішера з Ad Exchange або SSP, щоб він сам міг продати свій інвентар. SSP - це та ж DSP, але з точки зору паблішера або продавця контенту. Це технологічна платформа, яка автоматизує продажу показів для паблішерів [3].

Отже, DSP є платформою, яка автоматизує медіабаїнг та дозволяє економити час та матеріальні ресурси (кошти).

ЛІТЕРАТУРА:

1. Demand-Side Platform. [Електронний ресурс] – Ресурс доступу: https://ru.wikipedia.org/ wiki/ Demand-Side_Platform.
2. DSP, SSP и Ad exchange – что это такое и как они все дружат? [Електронний ресурс] – Ресурс доступу: https://maddata.agency/mad-blog/dsp-ssp-i-ad-exchange-chto-eto-takoe.
3. Что Такое Supply-Side Platform или SSP в Маркетинге? [Електронний ресурс] – Ресурс доступу: https://www.mobidea.com/academy/ru/shto-takoe-ssp.
4. S. Popereshnyak The method of data exchanging between smartphone and smart watch/ S. Popereshnyak, O. Suprun // Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 2017 14th International Conference. – 2017. – С. 392-395

Поперешняк С.В. *Доцент, кандидат фізико-математичних наук.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

СТАТИСТИЧНИЙ АНАЛІЗ БІТОВИХ ПОСЛІДОВНОСТЕЙ СЕРЕДНЬОЇ ДОВЖИНИ

Сучасні інформаційні системи вимагають особливого підходу до передачі електронних документів по відкритих каналах зв'язку, що забезпечується з використанням систем безпеки. Ефективна система безпеки інформації повинна забезпечувати:

- таємність інформації або важливою її частини;
- автентичність суб'єктів і об'єктів інформаційної взаємодії;
- захист від несанкціонованого доступу;
- захист прав власників інформації;
- оперативний контроль процесів управління, обробки і передачі інформації.

Всі перераховані пункти, яким повинна задовольняти ефективна система безпеки інформації, вирішуються з використанням генератора псевдовипадкових послідовностей. Такі генератори є детермінованими алгоритмами і застосовують в якості вхідних даних початкове значення, а на виході породжують послідовність значень, яка дуже схожі на випадкову. [1]

В роботі [2] запропонована наступна класифікацію вибірок за чисельністю, виходячи з вимог представлених в програмі критеріїв:

- дуже малі вибірки - від 5 до 12,
- малі вибірки - від 13 до 40,
- вибірки середньої чисельності - від 41 до 100,
- великі вибірки - від 101 і вище.

Для виявлення закономірностей аналізованих ПВП (або до їх відрізків різної довжини) застосовують широкий спектр різних статистичних тестів, розроблених в останні десятиліття. Попередній аналіз розглянутих статистичних тестів дозволив визначити, які з них найбільш придатні для використання в різних завданнях розробки засобів криптографічного захисту інформації. Розвиток методів статистичного тестування для деяких класів ПВП. Аналіз існуючих пакетів для статистичного тестування призводить до висновку, що багато тести можуть успішно використовуватися для дослідження ПВП на випадковість. Разом з тим, особливості прикладних задач показують, що класична математична модель статистичного тестування не цілком адекватно відображає потреби в дослідженні деяких об'єктів на випадковість. Така ситуація виникає, коли генеруються послідовності невеликої довжини (до 100 біт). В такому випадку досліджувану ПВП неможливо протестувати за допомогою одновимірної статистики. У цій ситуації пропонуємо протестувати ланцюжок на випадковість використовуючи двох і / або тривимірні статистики. Наведемо відповідну математичну модель.

Нехай є випадковий або псевдовипадковий процес і пов'язані з ним випадкові величини

$$\gamma_1, \gamma_2, \ldots, \gamma_n, \qquad (1)$$

де $\gamma_i = \ 0,1\ $, i = 1, 2, . . . , n, n > 0.

Підпослідовність $\gamma_j, \gamma_{j+1}, \ldots, \gamma_{j+s-1}$, послідовності (1) називається s-ланцюжком , $j = 1, 2, \ldots, n-s+1, s = 1, 2, \ldots, n$.

Позначимо $\eta\ t_1, t_2, \ldots, t_s$ число s-ланцюжків в послідовності (1), які співпадають з $t_1, t_2, \ldots, t_s$, де $t_i = \ 0,1\ , i = 1, 2, \ldots, s$.

Теорема. Нехай послідовність (1) складається з n, n > 0, незалежних однаково розподілених випадкових величин; $\mathrm{P}\ \gamma_i = 1\ = p$, $\mathrm{P}\ \gamma_i = 0\ = q$, p + q = 1, i = 1, 2, . . . , n, та

виконуються умови (У), k_1, k_2, k_3, t, t_1 – цілі числа такі, що $k_1 \geq 0, k_2 \geq 0, n \geq k_1, k_3 \geq 0, t, t_1 \epsilon\ 0,1$. Тоді

$$\mathrm{P}\ \eta\ t_1, t_1^*\ = k_1, \eta\ t_1\ t\ t_1^*\ = k_2\ = \sum_{m_1=k}^{n-k_1} p^{m_1} q^{m_0}\ C_{k_1}^{k_2} C_{m_1-k_1}^{k_2} C_{m_{t^*}}^{k_1}, \qquad (2)$$

де $m_0 = n - m_1$, символ $\sum$ позначає додавання за всіма цілими невід'ємними числами δ_0 и δ_1 такими, що $\delta_0 + \delta_1 = 2k_1 - k_2$, $t^* \stackrel{\text{def}}{=} 1 - t$.

Приклади застосування двомірної статистики для аналізу бітової послідовності

Розглянемо послідовності довжиною n = 13. Побудувавши набори всіх можливих ланцюжків довжиною 13 елементів отримаємо кількість ланцюжків рівним 2^{13}

Проаналізуємо твердження (2). Зафіксуємо t = 0 і t_1 = 0. Для цього порахуємо кількість входжень ланцюжків вигляду k_1 та k_2, для кожної з можливих послідовностей довжини 13, де k_1 це кількість входжень підпослідовності вигляду (01) і k_2 кількість входжень підпослідовності типу (001). Таким чином, для кожного з наборів k_1 і k_2 була знайдена ймовірність, що наведено на рисунку 1, а також

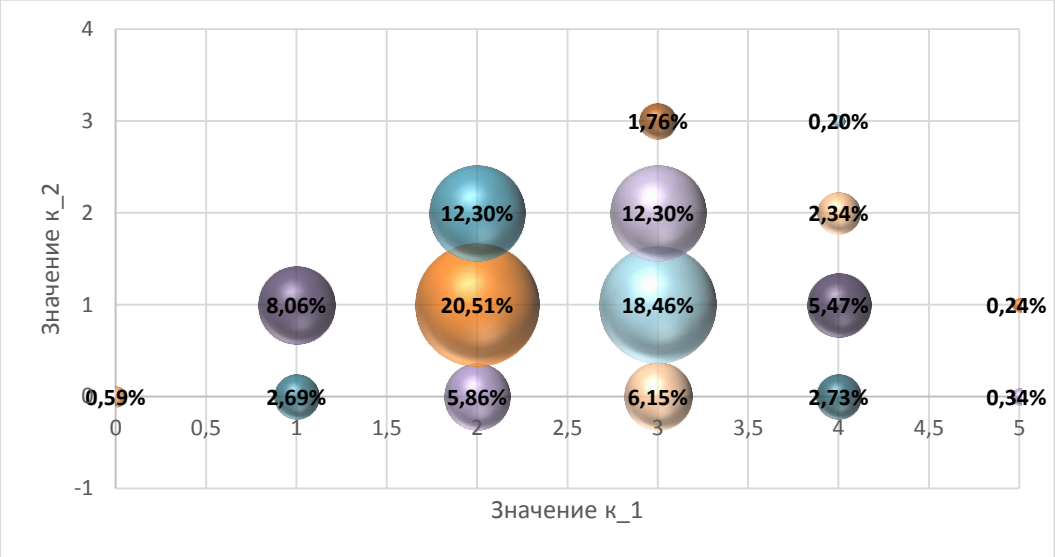


Рис. 1. Діаграма розподілу ймовірностей випадкової величини $\eta\ t_1, t_1^*\ = k_1, \eta\ t_1\ t\ t_1^*\ = k_2$

Проаналізувавши рисунок 1 можна зробити висновок що для аналізу послідовності ланцюжків малої та середньої довжини (від 13 до 100 елементів) одномірна статистика не завжди дає коректний результат. Наприклад, якщо розглянемо послідовність (0 0 0 0 1 0 0 1 0 0 0 1 0), де k_1 = 3, то можна зробити висновок з високим ступенем ймовірності про випадковості послідовності з даними характеристиками, однак якщо звернути увагу при k_1 = 3 і k_2 = 3, можна стверджувати, що дана послідовність є невипадковою, тому що з рисю моємо $\mathrm{P}\ \eta\ t_1, t_1^*\ = k_1, \eta\ t_1\ t\ t_1^*\ = k_2\ = 1{,}76\%$. Що і показує недостатність використання одновимірних статистик для аналізу коротких і середніх бітових послідовностей.

Підхід до тестування з використанням n-мірних статистик дозволяє розраховувати на більш глибоке обґрунтування випадковості послідовностей, що генеруються. Перевагою тестування n-мірними статистиками є також принципова можливість високого ступеня розпаралелювання обчислення сімейства статистик. Ця область є перспективною для наукових досліджень.

ЛІТЕРАТУРА:

1. Устименко Т. А. О случайности псевдослучайных последовательностей // Молодой ученый. — 2010. — №10. — С. 8-11. — URL https://moluch.ru/archive/21/2189/ (дата обращения: 10.01.2019).

2. Гайдышев И.П. Программное обеспечение анализа данных AtteStat. Руководство пользователя. Версия 13 – 2012. – 505 с.
3. V Masol, S Popereshnyak A theorem on the distribution of the rank of a sparse Boolean random matrix and some applications // Theory of Probability and Mathematical Statistics, V.76 – 2008. – p. 103-116.

Вєчерковська А.С. *Асистент*
Поперешняк С.В. *Доцент, кандидат фізико-математичних наук.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

МОДЕЛЮВАННЯ ОБ'ЄКТУ ПРОМИСЛОВОГО ВИРОБНИЦТВА

Технології управління сучасними вітчизняними промисловими підприємствами має суттєві відмінності від виробництв у промислово-розвинених країнах світу і на даний час не задовольняє вимог сучасних промислових виробництв.

Для управління сучасних підприємств з формування фільтруючих поліпропіленових волокнистих елементів (ФПВЕ) необхідно використовувати дані з різноманітних автоматизованих систем CAD/CAM/CAE, PDM, CAPP та інших, які при інтеграції між собою утворюють інтегроване інформаційне середовище (ІІС) підприємства, яке поєднує дані про повний життєвий цикл виробу (ЖЦВ).

Життєвий цикл промислових продуктів, як і життєвий цикл програмного забезпечення включає в себе ряд етапів, починаючи від зародження ідеї нового продукту, до його утилізації на закінчення терміну використання.

Розробка та впровадження інтегрованих інформаційних систем виробничого призначення – CAD/CAM/CAE, PDM та ERP-систем надало певні інструментальні засоби та інструменти для автоматизації завдань управління даними у ході виробництва. Разом із тим, до даного часу відсутні інформаційні технології та системи, які б дозволили комплексно реалізувати управління даними і процесами в умовах сучасних підприємств з виробництва ФПВЕ [1-4].

В якості об'єкта моделювання обрано кільцевий нагрівач, розглядаємо канал керування температурою нагрівача, за допомогою зміни його потужності. Об'єктом керування є кільцевий нагрівач, який знаходиться на початку тіла (циліндра) екструдера, кільцеві нагрівачі забезпечують нагрівання та плавлення гранул поліпропілену до температури яка визначається технологічним регламентом .

Розглянемо налаштування регулятора за допомогою програмного засобу SISOTool

Інтерактивне середовище SISOTool пограмного пакету MatLab можна застосовувати для налаштування регуляторів і аналізу системи. В даному інтерактивному середовищі можна задати параметри самому, або вибрати їх зі списку за допомогою автоматичного налаштування.

Задаємо передавальну функцію об'єкту:

```
num=[200 1];
den=[50000 14000 85 1];
Wp=tf(num, den)
Transfer function:
200 s + 1
---------------------------------
50000 s^3 + 14000 s^2 + 85 s + 1
sisotool(Wp)
```

Команда "sisotool(Wp)" викликає інтерактивне середовище «SISOTool» та завантажує в нього передавальну функцію об'єкта Wp. В реультаті відкривається вікно редактора, за допомогою кнопки "Control Architecture…" відкриваємося вікно вибору системи (Рис. 1.), в якому можна задати потрібну архітектуру, де G – plant (математична модель об'єкта), H – sensordynamics (давач, що вимірює вихідну величину), F — prefilter (фільтр), C — compensator (регулятор).

На рис. 2-4 представлено перехідні характеристики контуру керування кільцевого нагрівача, отриману в середовищі MatLab з його передавальної функції з налаштованим різними типами регуляторів.

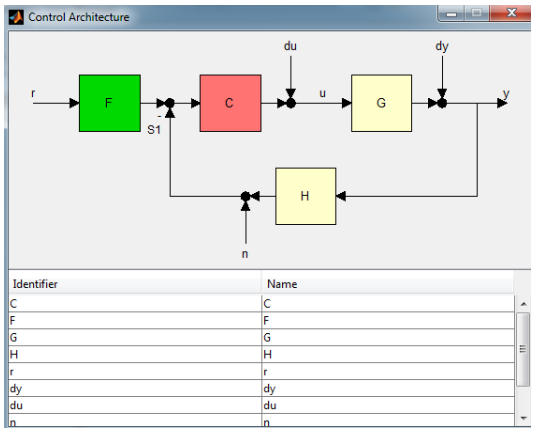


Рис. 1. Архітектура системи.

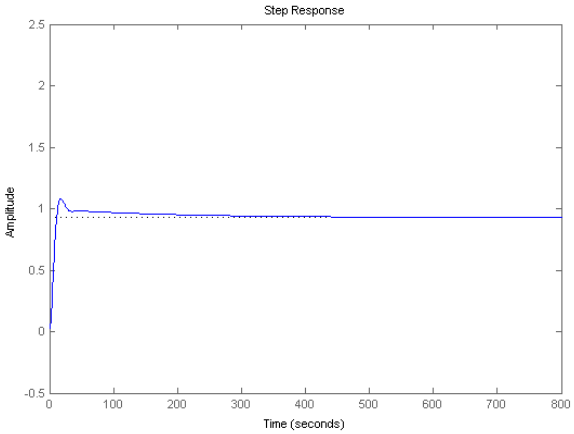


Рис. 2. Перехідна характеристика системи з налаштованим П-регулятором

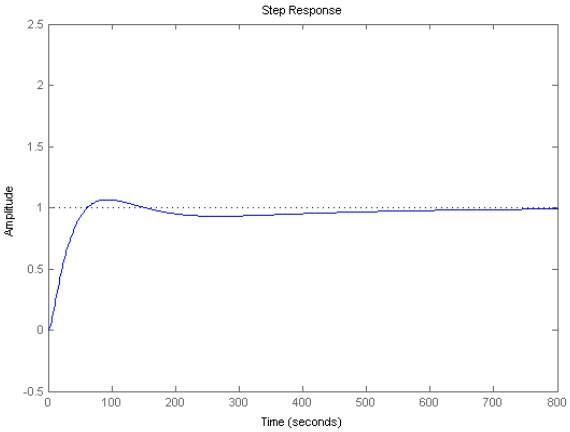


Рис. 3. Перехідна характеристика системи з налаштованим ПІ-регулятором.

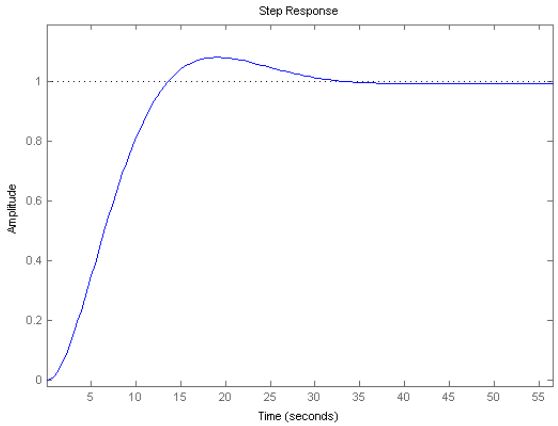


Рис. 4. Перехідна характеристика системи з налаштованим ПІД-регулятором.

З даних графіків, використовуючи налаштування за методом SISOTool можна побачити що:

- в результаті використання П-регулятора система не виходить на задане значення усталеного режиму;
- в результаті використання ПІ-регулятора в системі виникне перерегулювання;

- найдоцільніше використовувати ПІД-регулятор, тому що цей регулятор найефективніше виводить систему керування на задане значення усталеного режиму.

ЛІТЕРАТУРА:

1. Vecherkovskaya A. Software for calculation of productivity of polypropylene filtering element in dependence from its application / A. Vecherkovskaya, S. Popereshnyak // Технологический аудит и резервы производства, 2018 – №1 (39) – с. 14-23.
2. Vecherkovskaya A. Development of the management system of orders of the companies and organization of the staff work/ A. Vecherkovskaya, S. Popereshnyak // «EUREKA: Physics and Engineering», 2018. – p.12-20.
3. A Vecherkovskaya, S Popereshnyak Comparative analysis of mathematical models forming filter elements //- 2017 XIIIth International Conference on Perspective …, 2017. – p. 113-115
4. Vecherkovska, A., Popereshniak, S. (2018). Avtomatyzatsiia vyrobnytstva elementiv z porystoho polipropilenu metodom pnevmoekstruzii. Visnyk natsionalnoho tekhnichnoho universytetu «KhPI». Seriia: Mekhaniko-tekhnolohichni systemy ta kompleksy, 44 (1266), 116–121. Available at: http://mtsc.khpi.edu.ua/article/view/124899

Ємець Є.Я. *Студентка 3 курсу*
Поперешняк С.В. *Доцент, кандидат фізико-математичних наук.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

АВТОМАТИЗОВАНА СИСТЕМА УПРАВЛІННЯ ФАКУЛЬТЕТОМ

Наразі, загальні закономірності процесу автоматизації в різних сферах суспільного виробництва є дуже відомими. Але, так як галузь освіти відноситься до соціально-культурній сфері, то на відміну від виробничої сфери, результати цієї діяльності не є очевидними: виробничий продукт практично не піддається кількісному вимірюванню, а показники ефективності освітньої діяльності не мають чіткого і однозначного розуміння. Це багато в чому визначено той факт, що існуюча теорія автоматизованих систем управління була різноспрямованою [1-7].

Основним завданням даної наукової роботи є автоматизація основних інформаційних процесів освітнього закладу, а саме:

- автоматизація роботи з педагогічними кадрами;
- ведення бази даних по студентам (статистичні дані та досягнення студента);
- адміністрування освітнього процесу (складання розкладу і т. д.);
- автоматизація адміністративно-господарської і фінансової діяльності установи.

Також, важливим кроком для ефективного ведення статистики досягнень є створення системи, яка зможе постійно відстежувати досягнення студентів та визначати їх майбутні показники, наприклад, їх остаточні оцінки за заліки або екзамени, враховуючи поточний прогрес.

Однак, прогнозування успішності студентів у рамках однієї спеціальності істотно відрізняється і стикається з новими проблемами, наприклад:

- студенти можуть значно відрізнятися за рівнем освіти, а також обраними спеціалізаціями, в результаті чого вибираються різні курси. З іншого боку, той самий курс може бути прийнятий студентами в різних областях. Оскільки прогнозування успішності студентів у конкретному курсі залежить від успішності студентів в інших курсах, ключовим завданням для підготовки ефективного предиктора є те, як обробляти різнорідні дані студентів через різні сфери та інтереси. Аналогічно, прогнози успішності студентів на курсах часто ґрунтуються на оцінках, які проводяться на курсах, які мають бути однаковими для всіх студентів;
- студенти можуть проходити багато курсів, але не всі курси є однаково інформативними для прогнозування майбутньої діяльності студентів. Використання минулої роботи студента у всіх курсах, які виконав студент, не тільки збільшує складність, але й вводить шум у прогнозуванні, тим самим погіршуючи ефективність прогнозування.

Багато дослідників намагалися передбачити результати учнів на основі різних даних. Прогнози були зроблені з використанням різних статистичних методів, таких як багатофакторна регресія, аналіз шляхів або дискримінантний аналіз. Жоден з цих методів не має можливості виявляти потенційні моделі даних як нейронні мережі. Провідні нейронні мережі застосовуються у багатьох галузях, таких як фінансове прогнозування, медична діагностика, прогнозування банкрутства, розпізнавання в цілях регресії або класифікації, оскільки вони є одним з кращих функціональних показників. Гарні результати застосування нейронних мереж у задачах класифікації змушують нас використовувати їх для прогнозування результатів учнів у вищій освіті.

Актуальність роботи полягає в тому, що на українському ринку наразі немає автоматизованих систем управління вищим навчальним закладом або окремим факультетом. Це обумовлено тим, що провідним напрямком діяльності освітнього закладу є навчальний процес, а більшість представлених на ринку систем орієнтовані на виробництво і торгівлю, що і стало причиною того, що навчальні заклади залишились найменш автоматизованими. Також, розроблювана система необхідна для того, щоб набагато більше студентів закінчили навчання своєчасно через вплив на студентів, чия робота не відповідатиме критеріям закінчення освітньої програми.

ЛІТЕРАТУРА:

1. Neural Networks in Foreign-Language Based Higher Education, The Research Bulletin of Jordan ACM – Ресурс доступу: https://dl.acm.org/citation.cfm?id=521706

2. Naik, B., Ragothaman, S., 2004. Using Neural Network to Predict MBA Student Success, College Student – Ресурс доступу: https://eric.ed.gov/?id=EJ701993

3. Kaur K, Kaur K (2015) Analyzing the effect of difficulty level of a course on students performance prediction using data mining. In: 1st international conference on next generation computing technologies 2015 – Ресурс доступу: https://doi.org/10.1109/NGCT.2015.7375222

4. С.В. Поперешняк Проблеми підготовки IT-спеціалістів // Системи обробки інформації. № 7. – 2010. – С. 127-131

5. С.В. Поперешняк Актуальна проблема електронного документообігу– нестача дискового простору / С.В. Поперешняк, О.І. Недбайло // Вісник соціально-економічних досліджень. – 2013. – С. 147-152

6. Поперешняк С.В Сучасні проблеми електронного документообігу на прикладі системи «ДІЛО» /Недбайло О.І. Поперешняк С.В // Економіка підприємства: сучасні проблеми теорії та практики. – 2012. – С. 97-98

7. С.В. Поперешняк Інформаційні системи холдингових організацій та система управління взаємовідносинами з клієнтами // ДонНТУ. – 2010. – URI : http://ea. donntu. edu.ua /handle/123456789/18533

Касіяненко Д.В.
Студент 3 курсу
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка
м. Київ, Україна

СКЛАДАННЯ РОЗКЛАДУ З ВИКОРИСТАННЯМ ГЕНЕТИЧНОГО АЛГОРИТМУ

Інформатизація сфери освіти - один із пріоритетних напрямів процесу розвитку суспільства. Підвищення ефективності навчального процесу неможливо досягти без використання автоматизованих робочих місць для вирішення завдань, що виникають у вищих навчальних закладах. До таких завдань відносяться: облік персоналу, матеріально-технічних ресурсів, управління документацією, розподіл викладацького навантаження і т.д. Особливо слід відзначити важливість проблеми створення інформаційного забезпечення для організаційних процесів навчальних закладів, від якого певною мірою залежить ефективність використання науково-педагогічного потенціалу. У цей напрям включені завдання призначення і розподілу навчального навантаження студентів і викладачів - складання розкладу занять та іспитів. Ця складова навчального процесу регламентує трудовий ритм і впливає на творчу віддачу викладачів і студентів, тому її можна розглядати як фактор оптимізації використання обмежених трудових ресурсів. Технологію ж розробки розкладу слід сприймати не лише як трудомісткий технічний процес або об'єкт автоматизації з використанням комп’ютера, але і як процес оптимального управління. Таким чином, завдання отримання оптимальних розкладів занять та іспитів мають очевидний соціальний і економічний ефект.

Завдання планування завантаження викладацького складу і розподілу аудиторного фонду виникає у всіх навчальних закладах з певною кількістю викладачів і деяким числом аудиторій. Непросто організувати навчальний процес в рамках завдання розподілу ресурсів матеріальних, людських, трудових та ще й так, щоб побудована система була оптимальною за багатьма критеріями і не виходила б за рамки оголошених обмежень.

Діаграма на рис. 1 ілюструє в загальному, що управляє програмою генерування розкладу занять (зверху), що подається на вхід (зліва), що використовується для його генерування (знизу) та який результат отримаємо на виході (справа).

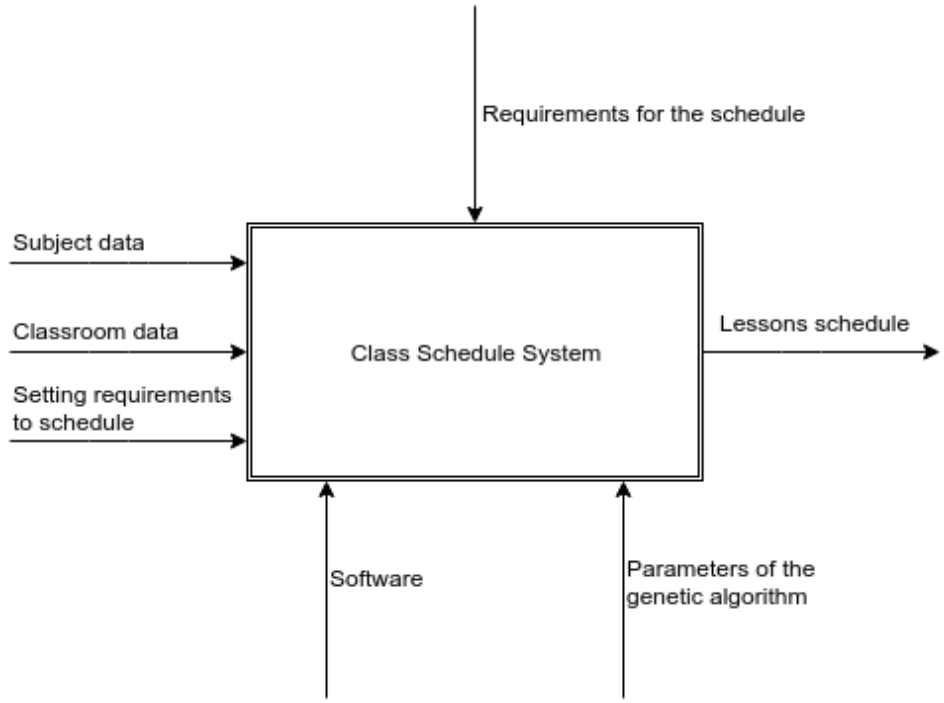


Рис. 1 - Загальна DFD діаграма

Розглянуто застосування генетичного алгоритму в задачі автоматизації складання розкладу занять, сформульовано задачу оптимізації розкладу в рамках генетичних алгоритмів. На основі цієї задачі реалізовано програмний продукт на мові програмування Python. За допомогою програмного продукту і ручного рішення складено розклади занять для груп кафедри програмних систем і технологій. На їх основі проведено порівняльну характеристику, визначено переваги та недоліки самостійно розробленого алгоритму.

ЛІТЕРАТУРА

1. Проблемы составления расписания в ВУЗе. – Режим доступу: http://www.pulsar.ru/wiki/detail.php?title=Проблемы+составления+расписания+в+вузе. – Дата доступу: 20.02.2019

2. Бевз С. В. Розробка автоматизованої системи формування розкладу магістратури / Бевз С. В., Войтко В. В., Бурбело С. М., Шоботенко А. М. // Наукові праці ВНТУ. – 2009. – No1. – С. 1-10.

3. Астахова И. Ф. Разработка информационной системы построения расписания / И.Ф. Астахова, Т.В. Курченкова // Математика. Образование.: Материалы межд. конф.– Т.2.– М.: Прогресс-Традиция, 2001.– С. 287-290.

Suprun O.M. *Candidate of physical and mathematical sciences,*
Associate Professor of software systems and technologies department
Liuzniak A.V. *4-th course bachelor student,*
Information technology faculty
Taras Shevchenko National University of Kyiv, Kyiv

IMPLEMENTATION OF THE POST-QUANTUM CRYPTOGRAPHIC ALGORITHMS

Existing cryptography is mostly based on the difficulty of factoring and the difficulty of calculating elliptic curve discrete logarithms. The invention of a large-scale quantum computer will however provide a possibility to efficiently solve these two kinds of problems, therefore the task of inventing cryptography approaches that appear to be resistant to an attacker who has access to a quantum computer becomes more and more urgent [1].

Post-quantum cryptography is the study of cryptosystems which can be run on a classical computer, but are secure even if an adversary possesses a quantum computer. Recently, NIST initiated a process for standardizing post-quantum cryptography and is currently reviewing first-round submissions. The most promising of these submissions included cryptosystems based on lattices, isogenies, hash functions, and codes [2].

The security of most cryptographic systems resides in the computational difficulty to the cryptanalyst of discovering the plaintext without knowledge of the key. This problem falls within the domains of computational complexity and analysis of algorithms, two recent disciplines which study the difficulty of solving computational problems. Using the results of these theories, it may be possible to extend proofs of security to more useful classes of systems in the foreseeable future [3].

One of the classes of post-quantum algorithms is code-based cryptographic systems. This term describes the cryptosystems in which the algorithmic primitive (the underlying one-way function) uses an error correcting code C. This primitive may consist in adding an error to a word of C or in computing a syndrome relatively to a parity check matrix of C. The first of those systems is a public key encryption scheme and it was proposed by Robert J. McEliece in 1978. The private key is a random binary irreducible Goppa code and the public key is a random generator matrix of a randomly permuted version of that code. The ciphertext is a codeword to which some errors have been added, and only the owner of the private key (the Goppa code) can remove those errors.

Similar ideas have been used to design other cryptosystems, like the Niederreiter encryption scheme, the CFS signature scheme, identification schemes, random number generators or a cryptographic hash function.

As for any class of cryptosystems, the practice of code-based cryptography is a trade-off between security and efficiency. For example, in McEliece's scheme there is an issue concerning the large size of the public key (100 kilobytes to several megabytes), but apart from the key size the McEliece encryption scheme has many strong features. First, the security reductions are tight and the system is very fast, as both encryption and decryption procedures have a low complexity [4].

The trapdoor for the McEliece cryptosystem is the knowledge of an efficient error correcting algorithm for the chosen code class (which is available for each Goppa code) together with a permutation. The McEliece PKC is summarized in the following algorithm:

1.System Parameters: $n, t \in \mathrm{N}$, where $t << n$.

2.Key Generation: Given the parameters n, t generate the following matrices:

– $G : k \times n$ generator matrix of a code $\mathrm{G}\bar{\mathrm{G}}$ over F of dimension k and minimum distance $d \geq 2t+1$ (A binary irreducible Goppa code in the original proposal).

– $S : k \times k$ random binary non-singular matrix
– $P : n \times n$ random permutation matrix

Then compute the $k \times n$ matrix $G^{pub} = SGP$.

3.Public Key: G^{pub}, t .

4.Private Key: S, D_{G}, P , where D_{G} is an efficient decoding algorithm for G.

5.Encryption $\left(E_{G^{pub},t} \right)$: To encrypt a plaintext $m \in \mathrm{F}^k$ choose a vector $z \in \mathrm{F}^n$ of weight t randomly and compute the ciphertext c as follows: $c = mG^{pub} \oplus z$.

6. Decryption $D_{S,D_{\mathrm{G}},P}$: To decrypt a ciphertext c calculate $cP^{-1} = mS\ G \oplus zP^{-1}$.

first, and apply the decoding algorithm $D_{G^{pub}}$ for G to it. Since cP^{-1} has a hamming distance of t to G we obtain the codeword $mSG = D_{\mathrm{G}}\ cP^{-1}$.

Let $J \subseteq 1,\mathrm{K}\ ,n$ be a set, such that G_J^{pub} is invertible, then we can compute the plaintext $m = mSG_{\ J}\ G_J^{\ -1} S^{-1}$ [4].

The invention of the quantum computer is fast approaching, so it is important to develop cryptosystems whose security relies on different, hard mathematical problems that are resistant to being solved by a large-scale quantum computer in the shortest possible time.

REFERENCES

1. Post-quantum cryptography [Electronic resource]. – 2019 – Available: https://www.microsoft.com/en-us/research/project/post-quantum-cryptography/
2. A Guide to Post-Quantum Cryptography [Electronic resource]. – 2018. – Available: https://blog.trailofbits.com/2018/10/22/a-guide-to-post-quantum-cryptography/
3. Whitfield Diffie, Martin E. Hellman. New directions in cryptography. IEEE Transactions on Information Theory, 644–654, 1976.
4. Daniel J. Bernstein, Johannes Buchmann Erik Dahmen. Post-Quantum Cryptography. 95-97, 2009.
5. Post-quantum cryptography [Electronic resource]. – 2018. – Available: https://pqcrypto.org/

Олійник А.М. *Студентка 3 курсу*
Супрун О.М. *к.ф.-м.н., доц.*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

ДОЦІЛЬНІСТЬ ВИКОРИСТАННЯ АЛГОРИТМІВ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО ЗАБАРВЛЕННЯ ТЕКСТУ

Сентимент-аналіз інформаційних потоків має великий потенціал застосування для моніторингових, аналітичних і сигнальних систем, систем документообігу і рекламних платформ, націлених на тематику веб-сторінок.

На рис. 1 представлена блок-схема роботи одного із алгоритмів для визначення тональності тексту, а саме - словарний метод

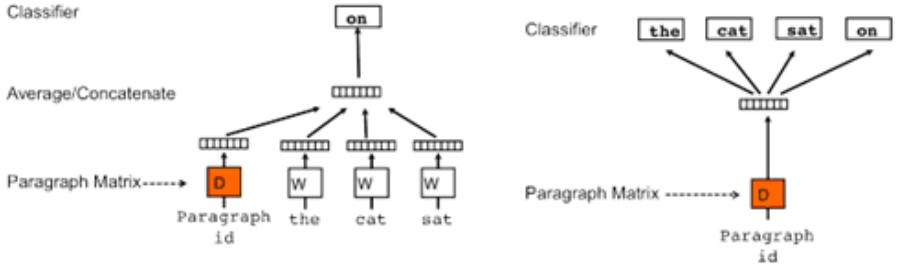


Рис. 1. Блок-схема роботи одного із алгоритмів для визначення тональності тексту

Виражена в тексті емоційна оцінка називається тональністю або сентиментом (від англ. Sentiment) тексту. Людина оцінює світ відразу за багатьма шкалами (хороший-поганий, сильний-слабкий, великий-маленький, щасливий-нещасливий, веселий-сумний, швидкий-повільний і т.п.), які емоційно по-різному навантажені. Для простоти можна вважати, що емоційна оцінка зводиться до шкалою хороший-поганий або позитивний-негативний.

Але це не єдиний і не визначальний тип завдання, яке має вирішувати сентимент-аналіз тексту. В даний час користувачів цікавить не загальна емоційна оцінка тексту, а відношення сентименту до конкретного об'єкта, або відношення суб'єкта висловлювання до обговорюваного об'єкту.

Об'єкт, щодо якого виражається емоційна оцінка, прийнято називати об'єктом тональності. Такий вид сентимент аналізу називається об'єктна тональність (object-based).

Носієм вираженої в тексті емоційної оцінки також зазвичай є цілком певний особа, в загальному випадку це автор тексту. Однак, якщо автор тексту посилається на чию-небудь думку, або цитує висловлювання іншої людини, то носієм емоційної оцінки або суб'єктом тональності буде той, на чиє думка посилаються.

Таким чином, тональність висловлювання визначається трьома компонентами: суб'єктом тональності (хто висловив оцінку), об'єктом тональності (про кого або про що висловлена оцінка) і власне тональною оцінкою (як оцінили).

Крім самої тональності, текст можна оцінювати по суб'єктивності / об'єктивності судження (Opinion Mining). Якщо ця думка автора висловлювання містить суб'єктивну оцінку описуваного, то текст вважається суб'єктивним. І навпаки, якщо це повідомлення або думка, за замовчуванням розділяється учасниками діалогу, то воно вважається об'єктивним.

Отже, за допомогою визначення емоційної тональності тексту, можна відслідковувати реакцію споживачів на будь-який продукт, враховуючи усі їх відгуки, що є наразі однією із головних задач маркетингового дослідження.

ЛІТЕРАТУРА

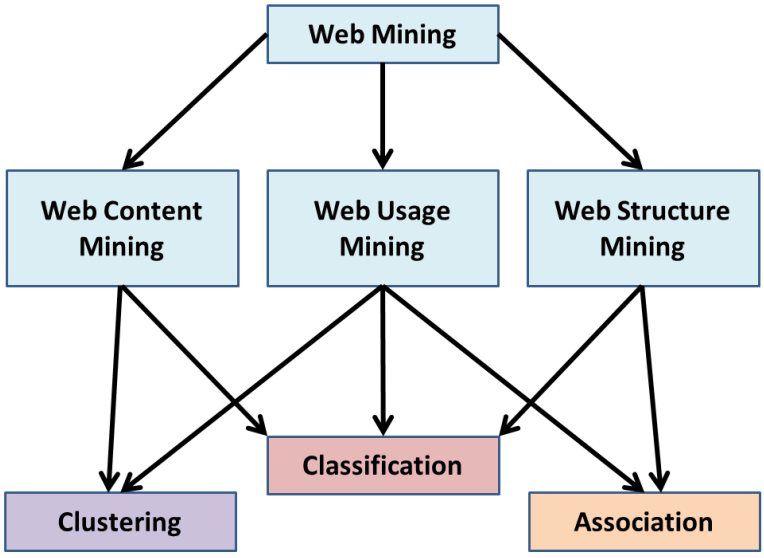
1. Варінти реалізацій для аналізу тональності тексту [Електронний ресурс] – Ресурс доступу: http://datareview.info/article/sovremennyie-metodyi-analiza-tonalnosti-teksta/
2. Що таке аналіз тональності тексту? [Електронний ресурс] – Ресурс доступу: https://uk.wikipedia.org/wiki/Аналіз_тональності_тексту
3. Word2Vec от Google [Електронний ресурс] – Ресурс доступу: https://code.google.com/archive/p/word2vec/

Podanenko Denys *3-rd course bachelor student,*
Information technology faculty
Taras Shevchenko National University of Kyiv, Kyiv

AUTOMATIC CLASSIFICATION OF INTERNET USERS BY THEIR NETWORK ACTIVITY

1) *Data mining* is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. Data mining is an interdisciplinary subfield of computer science and statistics with an overall goal to extract information (with intelligent methods) from a data set and transform the information into a comprehensible structure for further use. Data mining is the analysis step of the "knowledge discovery in databases" process, or KDD. Aside from the raw analysis step, it also involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

2) *Web mining* is the application of data mining techniques to discover patterns from the World Wide Web. As the name proposes, this is information gathered by mining the web. It makes utilization of automated apparatuses to reveal and extricate data from servers and web2 reports, and it permits organizations to get to both organized and unstructured information from browser activities, server logs, website and link structure, page content and different sources.



3) *Web Usage Mining* is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site.

Web usage mining itself can be classified further depending on the kind of usage data considered:

-Web Server Data: The user logs are collected by the Web server. Typical data includes IP address, page reference and access time.

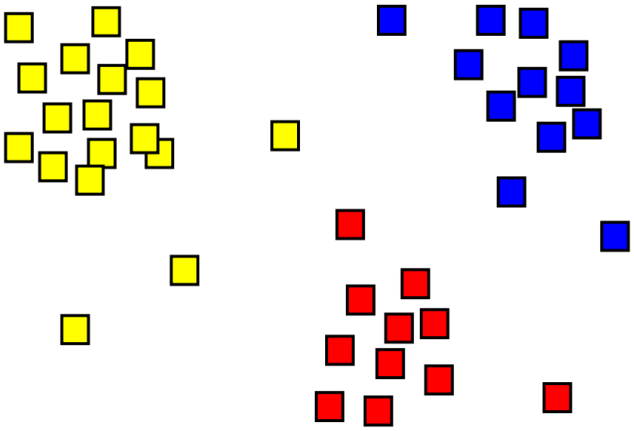
-Application Server Data: Commercial application servers have significant features to enable e-commerce applications to be built on top of them with little effort. A key feature is the ability to track various kinds of business events and log them in application server logs.

-Application Level Data: New kinds of events can be defined in an application, and logging can be turned on for them thus generating histories of these specially defined events. It must be

noted, however, that many end applications require a combination of one or more of the techniques applied in the categories above.

Studies related to work are concerned with two areas: constraint-based data mining algorithms applied in Web Usage Mining and developed software tools (systems). Costa and Seco demonstrated that web log mining can be used to extract semantic information (hyponymy relationships in particular) about the user and a given community.

4) *Cluster analysis or clustering* is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, data compression, and computer graphics.



REFERENCE

1. https://www.wikipedia.org/
2. https://www.intuit.ru/studies/courses/6/6/info
3. Wang Y. Web Mining and Knowledge Discovery of Usage Patterns

Решетняк М.А. *Студент 4 курсу*
кафедра Програмних систем і технологій
Факультет інформаційних технологій
Київський національний університет імені Тараса Шевченка м. Київ, Україна

ЗАСІБ ЗАБЕЗПЕЧЕННЯ ДОСТУПУ ТА ОБРОБКИ МУЛЬТИМЕДІЙНОГО КОНТЕНТУ У ХМАРНОМУ СЕРЕДОВИЩІ

Alexa - це хмарна голосова служба Amazon і мозок за десятками мільйонів пристроїв, включаючи сімейство пристроїв Echo, FireTV, Fire Tablet і сторонні пристрої з вбудованим Alexa. Ми створили можливість або навичку, які роблять Alexa більш розумним і роблять щоденні завдання більш швидкими, простішими і більш чудовими для клієнтів. Десятки тисяч розробників створили навички за допомогою комплекту Alexa Skills (ASK), колекції API для самообслуговування, інструментів, документації та зразків коду. З ASK, будь-хто може використовувати знання Amazon в голосовому дизайні для побудови швидко і легко. Початок будівництва сьогодні, щоб переосмислити ваш досвід клієнтів для голосу і досягти клієнтів, де вони є.

Навички Alexa включають інтерфейс голосового користувача або VUI, щоб зрозуміти наміри клієнта, а також хмарну службу заднього виду для обробки інтенцій і повідомити Алекса, як відповісти. Ми використали можливості в комплекті Alexa Skills, щоб доставити обидва.

Alexa перетворює вимовлені слова в текст за допомогою автоматичного розпізнавання мови (ASR), виводить значення мови, використовуючи природне розуміння мови (NLU), і надає основним клієнтам намір до вашої майстерності.

Ми скористалися перевагами власних голосових користувацьких інтерфейсів для розумного вдома, музики, відео та флеш-брифінгу або визначили власний користувальницький інтерфейс користувача.

Навички можуть підтримувати єдину мову або будь-яку комбінацію доступних мов. В нашій розробці використовується англійська.

Для відповідей на запити клієнтів можна використовувати текстові та аудіопотоки. Ми хочемо в ході розвитку проекту додати візуальні елементи та сенсорні входи на сумісних з Алексами пристроях з дисплеями, а також можна додати підтримку Echo Buttons.

Ми створили захоплюючі голосові враження, і спробуємо задовльнити клієнтів за допомогою пристроїв з підтримкою Alexa. Використали пакети SDK від Амазон, API та проаналізували для власного проекту зразки коду для створення ігор, навичок, корисних для дитини, вмісту вмісту, музичних навичок.

Навички Alexa включають інтерфейс голосового користувача або VUI, щоб зрозуміти наміри клієнта, а також хмарну службу заднього виду для обробки інтенцій і повідомити Алекса, як відповісти.

Alexa перетворює вимовлені слова в текст за допомогою автоматичного розпізнавання мови (ASR), виводить значення мови, використовуючи природне розуміння мови (NLU), і надає основним клієнтам намір до вашої майстерності. Для відповідей на запити клієнтів можна використовувати текстові та аудіопотоки.

На даний момент прописаний функціонал, який дозволяє прослуховувати за допомогою Smart Drive аудіокниги, музику і флеш-брифінги для вашої роботи. Так само він може бути використаний для певних понад як бізнесу, наприклад ваша компанія збирає дані

для якості обслуговування клієнтів. Наша розробка спростить і заощадить час пошуку розмови, компаніям, які в разі проблем з клієнтом аналізують його.

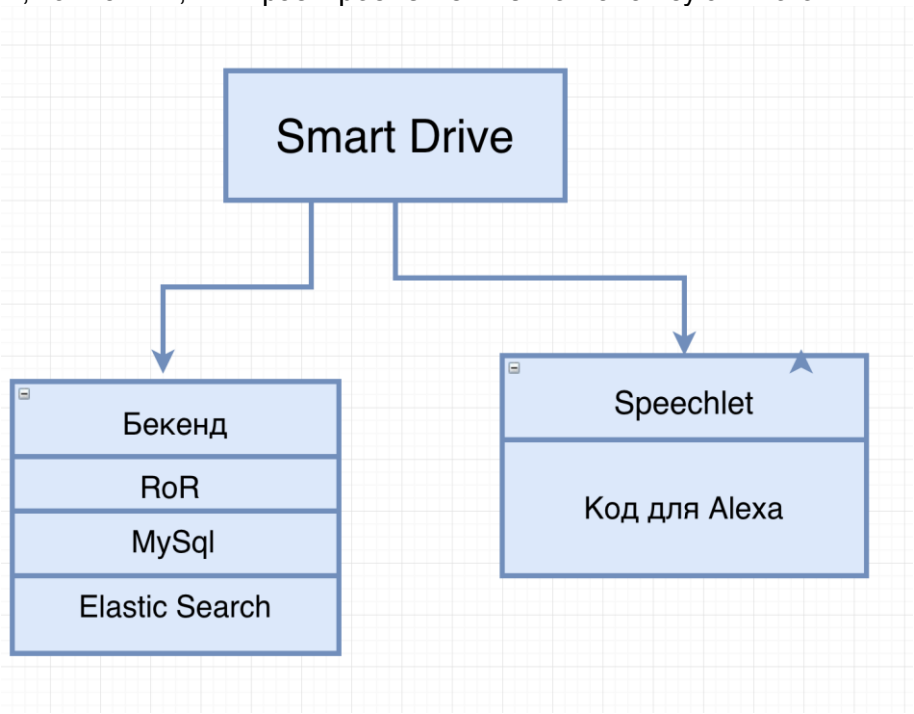


Рис. 1

Нові вклади в науку про голосових технологіях (текст в мову, розуміння природної мови і автоматичне розпізнавання мови).

Вже існує зростаючий список невеликих компаній, які використовують вигоду з фонду Alexa, наприклад, компанія під назвою Or-ange Chef, яка розширює Інтернет речей в професійну кухню.

По завершенню проекту як один зі способів реалізації готового програмного продукту ми можемо подати заявку до фонду Amazon для потенційного продажу або подальшого розвитку проекту та реалізацію його на ринку.

https://developer.amazon.com/