

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеська національна академія харчових технологій
Університет Інформатики і прикладних знань, м.Лодзь, Польща
Національний технічний університет України «Київський
політехнічний інститут»
Навчально-науковий інститут комп'ютерних систем і технологій
«Індустрія 4.0» ім. П.М. Платонова

XXI Всеукраїнська науково-технічна конференція
молодих вчених, аспірантів та студентів

«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ
ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»

Матеріали конференції



Одеса

22-23 квітня 2021 р.

Стан, досягнення та перспективи інформаційних систем і технологій / Матеріали XXI Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів. Одеса, 22-23 квітня 2021 р. - Одеса, Видавництво ОНАХТ, 2021 р. – 229 с.

Збірник включає матеріали доповідей учасників конференції, які об'єднані за тематичними напрямками конференції.

ОРГАНІЗАЦІЙНИЙ КОМІТЕТ

Голова - д.т.н., проф., Єгоров Б.В., ректор ОНАХТ.

Співголови:

Поварова Н.М. – к.т.н., доц., проректор з наукової роботи ОНАХТ,
Котлик С.В. – к.т.н., доц., директор ННІКСіТ "Індустрія 4.0" ОНАХТ,
Даріуш Долива, д.математичн.наук, уповноважений декана факультету Інформатики УІтаПЗ, м.Лодзь, Польща,
Ковалюк Т.В. - к.т.н., доц. кафедри АСОІтаУ НТУУ «Київський політехнічний інститут»

Члени оргкомітету:

Плотніков В. М. – д.т.н., проф., завідувач кафедри ІТтаКБ ОНАХТ,
Артеменко С.В. – д.т.н., проф., завідувач кафедри КІ ОНАХТ,
Хобін В.А. – д.т.н., проф., завідувач кафедри АТПтаРС ОНАХТ,
Тарасенко В.П. – д.т.н., проф., завідувач кафедри СКС НТУУ «Київський політехнічний інститут»,
Невлюдов І.Ш. – д.т.н., проф., завідувач кафедри КІТАМ ХНУРЕ,
Мельник А.О. – д.т.н., проф., завідувач кафедри ЕОМ НУ “Львівська політехніка”,
Жуков І.А. – д.т.н., проф., завідувач кафедри КСтаМ НАУ.

Матеріали подано українською, російською та англійською мовами.
Редактор збірника Котлик С.В.

THROUGH THE INTRODUCTION OF SMART PARKING. <i>O.N.DOLININA, M.E. MANSUROVA, Z.E. BAIGARAYEVA, S.A. BAYAZITOVA</i> (Al-Farabi Kazakh National University, Almaty, Kazakhstan)	
АНАЛІЗ ОСОБЛИВОСТЕЙ ТА ЕТАПІВ ВИКОНАННЯ ЗАВДАННЯ ПРОГНОЗУВАННЯ ПРИ АНАЛІЗІ ДАНИХ. <i>ГЕЖА М.І., ТИЩЕНКО С.Є., РУДНІЧЕНКО М.Д.</i> (Державний Університет «Одеська Політехніка»)	183
ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖЕВОЇ АРХІТЕКТУРИ ДКЧП ДЛЯ ЗАДАЧІ ВІДСТЕЖЕННЯ ТОЧКИ МАКСИМАЛЬНОЇ ПОТУЖНОСТІ. <i>НЕЧАХІН В.В.</i> (Чорноморський національний університет ім. Петра Могили)	185
НАНОРОБОТОТЕХНІКА: УТОПІЯ ЧИ РЕАЛЬНІСТЬ? <i>ЛЯШУК Т.Г.</i> (Рівненський державний гуманітарний університет)	186
РОЗРОБКА ПРОТОТИПУ СИСТЕМИ ДОСЛІДЖЕННЯ СТАНУ МІКРОКЛІМАТУ ПРИМІЩЕНЬ НА ОСНОВІ ПЛАТИ РОЗРОБНИКА TI-RSLK. <i>КРАВЧУК О.О., ЧЕКУБАШЕВА В.А., ГЛУХОВ О.В., ЛЕВЧЕНКО Є.В., РОГОВЕЦЬ В.Є.</i> (Харківський національний університет радіоелектроніки)	188
УТИЛІТА КАЛІБРУВАННЯ 3D ПРИНТЕРІВ, ЗІБРАНИХ НА БАЗІ ARDUINO MEGA. <i>КОТЛИК Д.В., СОКОЛОВА О.П., КОТЛИК С.В.</i> (Одеська національна академія харчових технологій)	190
АНАЛІЗ МЕТОДІВ ТА ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ОБЛИЧЬ В СИСТЕМАХ ШТУЧНОГО ІНТЕЛЕКТУ. <i>УЛЬЯНОВСЬКА Ю.В., ТХОРЖЕВСЬКИЙ Д.О., КОЗЛОВ Є.С.</i> (Університет митної справи та фінансів.)	193
АВТОМАТИЗАЦІЯ МАРКЕТИНГУ ПІДПРИЄМСТВА ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ РЕФЕРАЛЬНИХ МАТРИЦЬ. <i>ЖМАЙ О.В.</i> (Громадська організація «Молодіжна організація “Енектус” при Одеському національному університеті імені І.І.Мечникова), <i>КОРКІНА А.О.</i> (Одеський національний університет імені І.І.Мечникова)	196
ПРОЕКТИРОВАНИЕ ИГРОВОГО БОТА. <i>РУДЬ А.В.</i> (Белорусский Государственный Университет, Республика Беларусь)	198
СУЧАСНИЙ СТАН ТА ПЕРСПЕКТИВИ РОЗВИТКУ НЕЙРОННИХ МЕРЕЖ У WEB-РОЗРОБЦІ. <i>ЗИБІНА К.В., РУСАКОВА Н.Є.</i> (Харківський Національний Університет Радіоелектроніки)	200
АЛГОРИТМ КЛАСТЕРИЗАЦІЇ СЛІВ В ДОКУМЕНТАХ З БЛОЧНОЮ СТРУКТУРОЮ. <i>МАСАЛЬСЬКИЙ Р.О., МАЗУРОК І.Є.</i> (Одеський національний університет імені І.І.Мечникова)	202
MODERN CLOUD STORAGE TECHNOLOGIES. <i>TASHU A.A., TARNAVSKYI Y.A.</i> (National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”)	203
METHODS FOR DETERMINING SPATIAL ORIENTATION IN AUGMENTED REALITY USING MARKERS. <i>RADOUTSKA A.K.</i> (Kharkiv National University of Radio Electronics)	205
TACOTRON 2 I WAVEGLOW ДЛЯ ПЕРЕТВОРЕННЯ ТЕКСТУ ДО РЕЧІ ДЛЯ ПЕРСОНАЖІВ КОМП'ЮТЕРНИХ ІГОР. <i>ГРИГОРЯН К.А., МАЗУРОК І.Є., ВОЛКОВ К.С., МАСАЛЬСЬКИЙ Р.О.</i> (Одеський національний університет імені І.І.Мечникова)	207
АНАЛІЗ СУЧАСНИХ МОЖЛИВОСТЕЙ ТЕХНОЛОГІЙ ДОПОВНЕНОЇ РЕАЛЬНОСТІ В МЕДИЦИНІ. <i>ВАЛЬТЕР Х.Є.</i> (Харківський національний університет імені Василя Назаровича Каразіна)	208
Розділ 7.	
Комп'ютерні ігри і WEB-дизайн	
АНАЛІЗ ВІДМІННОСТЕЙ <i>PBR</i> І <i>RAY TRACE</i> МЕТОДІВ РЕНДЕРИНГУ. <i>ЖУКОВЕЦЬКА С.Л., БОГДАНОВ С.Ю.</i> (Одеська національна академія харчових	210

УДК 004.021

АЛГОРИТМ КЛАСТЕРИЗАЦІЇ СЛІВ В ДОКУМЕНТАХ З БЛОЧНОЮ СТРУКТУРОЮ

МАСАЛЬСЬКИЙ Р.О. (*masalskyi@stud.onu.edu.ua*),

МАЗУРОК І.Є. (*igor@mazurok.com*)

Одеський національний університет імені І. І. Мечникова

Розпізнавання документів одна з найважливіших автоматизацій в компаніях. Але інколи на розробку такої системи виділяють малий датасет, через небажання розголошувати конфіденційну інформацію. В роботі пропонується алгоритм вирішення проблеми кластеризації слів в документах, які мають блочну структуру, маючи малу кількість тестових даних.

Ключові слова: кластеризація, розпізнавання документів, машинне навчання, комп'ютерний зір.

Вступ. Розглянемо проблему розпізнавання документів із дуже невеликою кількістю вхідних прикладів, до декількох одиниць документів. Ми розглянемо документи, що мають блокову структуру. Під структурою блоків ми розуміємо документ, що складається з блоків (кластерів) слів. У блокових документах можна припустити (залежно від документів), що один блок повністю належить певному ключовому слову. Потім, якщо ви знайдете кластери [1] слів, ви можете вирішити проблему зіставлення блоку з якимось ключовим словом. Тут виникають такі проблеми: ввести відповідну функцію відстані між словами. На основі функції відстані розробити алгоритм швидкої кластеризації слів.

Проблема відстані. Відстань між двома словами слід описувати не одним числом, а двома - вертикальною та горизонтальною (W, H). Порогова відстань визначається щодо розміру документа. Різні типи документів можуть вимагати різної порогової відстані. Для морських коносаментів оптимальним буде порогове значення $W_t \in (0.0125; 0.02)$, $H_t \in (0.006; 0.01)$.

Проблема швидкої Кластеризації. Кластеризацію за допомогою створеної матриці відстаней можна виконати алгоритмами dfs або bfs. Основна проблема полягає в тому, щоб знайти матрицю відстані швидше ніж $O(n^2)$, де n це кількість слів в документі. Нехай поріг вертикальної відстані W_t , та горизонтальної H_t . Витягнемо прямокутники горизонтально до $\frac{W_t}{2}$ з кожної сторони і вертикально до $\frac{H_t}{2}$ з кожної сторони. Тепер, якщо два прямокутники перетинаються, це означає, що відстань між ними нижче порогового значення і їх можна об'єднати в один кластер. Проблема зводиться до знаходження всіх перетинів прямокутників. Використовується техніка sweeping line [2]. Побудована координатна площина. Вертикаль відповідає висоті пікселів документа. Горизонталь відповідає подіям. Події бувають двох типів: початок прямокутника, кінець прямокутника. Коли трапляються події першого типу, додаються до сегмента вертикальної осі, що відповідає даному прямокутнику, позначки, що вказують на відкриття прямокутника на цьому сегменті в даний момент часу. Тепер, при іншій події першого типу, перевіряється наявність на вертикальній лінії, яка відповідає цьому прямокутнику, позначок, що позначають поточне відкриття інших прямокутників у цьому сегменті. Якщо мітки були знайдені, то було знайдено перетин. Для події другого типу видалить мітки з сегмента. Споживання пам'яті збільшується. Споживання пам'яті стає $O(n \log n)$ через те, що кожна вершина дерева містить усі мітки, що лежать на відрізку, за який ця вершина відповідає. Як результат, асимптотична поведінка дорівнює $O(n^2 \log n)$ через те, що другий тип запиту повинен бути зроблений n разів, що гірше, ніж квадратний пошук всіх відстаней. Але через специфіку завдання це не зовсім так. При запиті

другого типу кожного разу запитується відносно невеликий сегмент. Тоді під час підйому в дереві об'єднуються не всі елементи, а лише ті, що лежать на запитуваному сегменті, яких на практиці не більше 3-4. У результаті асимптотична поведінка $O(n \log n)$.

Результат кластеризації на морському коносаменті (Рис. 1,2)

Shipper COOPERATIVA AGRARIA CAFETALERA PANGO A LTDA AV. LOS CAFETALEROS N 600 SAN MARTIN DE PANGO A SATIPO-JUNIN-PERU PHONE: +51(064)543017 FAX +51-(064)543144		Booking No. 855291456	Svc Contract 182020
Consignee (negotiable only if consigned "to order", "to order of" a named Person or "to order of bearer") COOPERATIVE COFFEES, INC. 302W. LAMAR ST. SUITE E 31709 AMERICUS, GEORGIA, USA TEL.: +1-229-924-3035 FAX: +1-229-924-6250 Atencion: BILL HARRIS		Notify Party (see clause 22) EXCELCO TRADING L.P. 17 BATTERY PLACE, SUITE 1010 NEW YORK, NY 10004 TEL. 212 344 6771 FAX.: 212 344 8723 George.TslatsiosGtslatsios@exceltrade.com	Onward inland routing (Not part of Carriage as defined in clause 1. For account and risk of Merchant)
Vessel (see clause 1 + 19) MAERSK RIO GRANDE	Voyage No. 0766	Place of Receipt. Applicable only when document used as Multimodal Transport B/L. (see clause 1)	
Port of Loading Callao	Port of Discharge Newark	Place of Delivery. Applicable only when document used as Multimodal Transport B/L. (see clause 1)	

Рис. 1. Оригінальний морський коносамент.

The image shows a scanned version of the bill of lading from Figure 1. Several fields are highlighted with green rectangular boxes: the shipper information, the booking number (855291456), the consignee information, the notify party information, the vessel name (MAERSK RIO GRANDE), the voyage number (0766), the port of loading (Callao), and the port of discharge (Newark).

Рис. 2. Результат кластеризації на морському коносаменті.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Duran B., Odell P. Cluster analysis. – Resource access mode: <https://www.twirpx.com/file/2062638/>
2. Diane Souvaine. Line Segment Intersection Using a Sweep Line Algorithm. — 2008.

УДК 004.75

MODERN CLOUD STORAGE TECHNOLOGIES

TASHU A.A., TARNAVSKYI Y.A.
(tashualbert@gmail.com)

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Cloud technologies are developing rapidly. With usage of cloud computing, even small organizations can afford a cross-regional, distributed architecture. One of the most important parts in the cloud infrastructure is data storage. This work is aimed to consider problems associated with storing data in the cloud and provide an overview of algorithms and technologies utilized to solve these problems.

With the growing applicability of information technology in all spheres of human life, the amount of data that must be stored and processed inevitably increases. Storing such painful amounts of data requires the development of scalable and fault tolerant storage systems that can meet the

**XXI Всеукраїнська науково-технічна конференція
молодих вчених, аспірантів та студентів**

**«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ
ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»**

Одеса

22-23 квітня 2021 р.

Збірник включає доповіді учасників конференції. Тези доповідей публікуються у вигляді, в якому вони були подані авторами.

Відповідальність за зміст і форму подачі матеріалу несуть автори статей.

Редакційна колегія: Котлик С.В., Корнієнко Ю.К.

Комп'ютерний набір і верстка: Соколова О.П.

Відповідальний за випуск: Котлик С.В.