2. Doctrine OPvM [Електронний ресурс]. – Режим доступу: URL: http://doctrine-project.org (дата звернення: 04.08.2017).

3. Роберт Мартин Гибкая разработка программ на Java и C++. Принципы, паттерны и методики. /Роберт С. Мартин, Джеймс Ньюкирк, Роберт Косс - Изд-во: Диалектика-Вильямс, 2016. - 704с.

4. Джулій В.М. Методи та алгоритми розробки web-додатків / В.М. Джулій, Ю.О. Гунченко, Д.В. Чешун // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2017. – Вип. № 56. – С.107-115

Дослідження проблем ідентифікації об'єктів в базах даних

Мозолюк В.О., Джулій В.М.
Хмельницький національний університет

На даний момент СУБД широко використовуються в організації сучасних інструментальних, промислових, аналітичних та інформаційних систем. Однак такий бурхливий розвиток інформаційних технологій баз даних поставило також ряд нових проблем і визначило напрямки подальших досліджень у цій області. Не припиняюча робота дослідників та аналітиків відноситься до питань оптимізації виконання запитів і структур зберігання даних, новітніх способів виконання реляційних операцій, організації пошуку, і багато інших моментів, що визначають результативність роботи СУБД. Програмне забезпечення на даний момент розвивається в умовах швидкого зростання обчислювальних потужностей, апаратних можливостей, швидкості доступу до пам'яті, обсягу пам'яті, пропускної здатності та надійності каналів передачі даних. Все більшого значення набувають засоби, що забезпечують взаємодію в розподіленої системі функціонування інформаційних систем.

Розглянемо більш докладно основні напрямки розвитку сучасних баз даних і СУБД:

1. Стандартизація мови SQL. У сучасних СУБД на даний момент основною мовою написання запитів і доступу до баз даних є мова SQL (Structured Query Language). Міжнародний стандарт даної мови розроблений в 1989 році, і більшість виробників СУБД призвели свої системи у відповідність даному стандарту. Потрібна постійна актуалізація мови SQL до мінливих вимог сучасних програмних продуктів та апаратних засобів.

2. Використання мультипроцесорних організацій. Промислові комерційні СУБД реалізуються на основі архітектури "клієнт-сервер". При даній організації всі операції над базами даних виконуються на сервері, що володіє достатньою продуктивністю і набором обчислювальних ресурсів. Після появи мультипроцесорних симетричних апаратних архітектур в

багатьох СУБД була переглянута організація серверних платформ і реалізована можливість розпаралелювання обчислень.

3. Інтеграція та інтероперабельність. Залишається актуальним рішення проблеми використання баз даних попередніх поколінь і версій. Прагнення до спрощення технологічних процесів і необхідність інтеграції інформаційних ресурсів призвели до розробки СУБД здатних підтримувати поряд зі структурованими даними також і текстові документи і виконувати їх пошук по запитах користувачів. Розвинені засоби текстового пошуку присутні в даний час в DB2 (IBM), Oracle, Microsoft SQL Server та ін.

Ефективність управління сучасним бізнесом заснована на можливості отримання управлінським персоналом всебічної інформації з усіх напрямків діяльності. При цьому важливо встановлення контролю над зростаючими потоками інформації, прискорення процесу їх обробки, пошуку та аналізу даних. Існуючі в даний час і розроблювальні нові автоматизовані системи характеризуються великою різноманітністю підтримуваних інформаційних ресурсів, способів організації даних, функціональними можливостями користувацьких інтерфейсів та інших їх технологічних характеристик. В розробках інформаційних систем даної категорії затребуваний практично весь спектр ключових технологій управління інформацією. В даний час багато організацій накопичили значні обсяги інформації. Внаслідок цього стає актуальною проблема розробки корпоративної системи управління знаннями. Однак серйозною перешкодою на даному шляху є розрізненість інформації в корпоративній системі, нерідко дані по одному напрямку діяльності зберігаються в різних додатках і форматах. Крім того, обсяги оброблюваної електронної інформації наростають по експоненті – цьому сприяє активне впровадження мультимедіа, широке поширення корпоративних і глобальних мереж, відхід більшості компаній від паперового документообігу та перехід на автоматизовані системи управління. В подібній ситуації значно зросла необхідність у створенні та впровадженні ефективних систем пошуку та аналізу даних. Традиційними є системи пошуку, що розвиваються в тісному взаємозв'язку з СУБД і в основному орієнтовані на роботу зі структурованими текстовими даними. Однак інтегровані в СУБД системи пошуку слабо адаптовані для обробки мультимедійної і неструктурованої інформації. За статистикою, частка структурованих даних в сучасних базах даних становить не більше 35-50%, решта ж припадають на частку різних довідників, сканованих документів і іншої розрізненої інформації. У цьому випадку виникає проблема пошуку і вибірки необхідної інформації з великого неструктурованого масиву.

Для багатьох організацій інформація є основним активом. Спотворення або пошкодження важливої інформації може призвести до суттєвих фінансових втрат і репутаційним ризикам. Аналізуючи дані, отримані з відкритих джерел і наукових публікацій, можна виділити основні

види втрат, що виникають внаслідок помилок і спотворень інформації в базах даних: втрати внаслідок невірного, поганого надання послуг («брак» в інформації). Даний вид втрат присутній майже в будь-якій організації. В середньому організація втрачає 25-40% часу співробітників, від втрат даного виду; втрати оплачуваного часу співробітників на непродуктивну діяльність. В тому чи іншому виді даний вид втрат зустрічається в будь якій організації, може досягати, наприклад, у менеджерів середньої ланки більше 50% робочого часу, у менеджерів низової категорії до 80%; втрати внаслідок використання «не оптимальних технологічних ланцюжків. Даний вид втрат присутній майже в будь-якій організації. За цими причинами в середньому організація втрачає близько 35% робочого часу задіяних співробітників і це може призвести до подорожчання однієї операції до 100%; втрати часу, грошових коштів, клієнтів по причині відсутності або дублюванні інформації. Даний вид втрат присутній майже в будь-якій організації. Втрати становлять близько 15% часу співробітників, що спричиняє збільшення вартості виконуваної операції.

Основним чинником, що стимулює розвиток технологій пошуку, є поява великої кількості електронних бібліотек і архівів, що містять значні обсяги актуальних знань. Продуктивність і ефективність будь-якої системи зберігання інформації безпосередньо залежить від ефективності та продуктивності пошукових систем. Саме пошукова система визначає, чи перетворяться в знання численні розрізнені дані, що надходять по різних каналах зв'язку і накопичуються в різноманітних базах даних та електронних архівах. Найбільш поширеним видом інформаційних ресурсів для організацій, що працюють з персональними даними (бюро кредитних історій, банки, страхові організації, будь-які організації з досить крупним штатом співробітників) є тексти на природних мовах. Цим обумовлено широке застосування в таких системах технологій текстового пошуку. Дані технології використовуються при цьому не тільки в системах, побудованих за принципом традиційних текстових систем, але і для пошуку в колекціях, організованих у вигляді веб-сайтів, а також для пошуку в глобальної мережі Інтернет.

При організації пошуку в базах персональних даних клієнтів виникають характерні проблеми, пов'язані з наявністю в запитах орфографічних і фонетичних помилок, помилок введення інформації, а також відсутністю єдиних стандартів транскрипції з іноземних мов. Внаслідок цього задача пошуку в базах персональних даних не може бути повною мірою вирішена тільки методами перевірки на точну відповідність. Стає актуальною задача розробки спеціальних методів і технологій текстового пошуку з використанням нетривіальних рішень, в тому числі на основі операцій несуворої відповідності. Однак універсальної методики пошуку в умовах зашумленості даних не існує, оскільки кожна проблема має власну

оригінальну специфіку. Для рішення виниклих проблем потрібно використовувати алгоритми здатні відшукати всі лексикографічно близькі до шаблону пошуку слова, що відрізняються замінами, пропусками і вставками символів. Таким чином, автоматично стає допустимою помилка, як у вхідних даних, так і в термінах запиту. В даний час можливості виконання пошуку за подібністю не використовуються в СУБД. Таким чином, виникає задача розробки алгоритмів виконання спеціальних реляційних операцій, що виникають в задачі ототожнення записів. Проведений аналіз напрямків розвитку сучасних баз даних показує, що склалися і формуються за останні роки тенденції розвитку інформаційних технологій істотно впливають, у тому числі і на функціональні можливості автоматизованих систем. Задача встановлення відповідності між окремими об'єктами - побудова процедур ототожнення в даний час не має задовільного рішення. Існуючі роботи, присвячені інтеграції БД, дозволяють здійснити тільки інтеграцію схем БД, але не пропонують способів побудови процедур ототожнення. Побудова процедур ототожнення ускладнюється відсутністю серед загальних атрибутів відповідних один одному таблиць різних БД первинних ключів і наявністю помилок операторського введення. Існуючі СУБД не пропонують можливості для використання пошуку за подібністю, що усуває викликані помилки операторського введення.

З урахуванням специфіки роботи з персональними даними пропонується вирішення наступних прикладних задач: повна ідентифікація клієнта при наявності спотворень інформації в базі даних або в пошукових запитах; усунення дублікатів записів при надходженні до БД з множинних джерел зі слабоструктурованою інформацією; пошук і коректування помилок в персональних даних клієнтів (фізичних і юридичних осіб).

В області технологій можна виділити появу принципово нових програмних засобів аналізу фрагментарної, слабоструктурованої, пошкодженої, нечіткої, неповної інформації. До таких засобів можна віднести технологію інформаційного моніторингу комплексних процесів і засоби Business Intelligence (бізнес аналітики). Використання принципово нового інструментарію на основі алгоритмів нечіткого пошуку в міжнародних компаніях і державних організаціях, великих корпораціях і фінансових установах показало їх ефективність і величезний потенціал для вирішення прикладних задач.

Перелік посилань

1. Васильєв В.І. Інтелектуальні системи захисту інформації: навч. посібник / В. І. Васильєв. - 2-е изд., Испр. - М .: Машинобудування, 2012. - 171 с.

2. Гордейчик С.В. Безпека бездротових мереж. / С.В. Гордейчик, В.В. Дубровін - М .: Гаряча лінія - Телеком, 2008. - 288 с.

3. Гузаіров М.Б. Управління захистом інформації на основі інтелектуальних технологій: навчальний посібник./ М.Б. Гузаіров, І.В. Машкіна - М .: Машинобудування, 2013. - 241 с.

Проблеми та перспективи побудови систем управління ресурсами інформаційних комунікаційних мереж

Просянюк В.В
Науковий керівник – д.т.н.,проф. Андрощук О.С
Хмельницький національний університет

Дослідження параметрів інформаційного потоку сучасних комунікаційних мереж показує, що припущення про справедливість властивостей процесів надходження та обслуговування потоків можуть реалізовуватися на практиці при малих навантаженнях, в умовах низького навантаження і невисоких швидкостях передавання потоків даних. Актуальними завданнями для дослідження та моделювання роботи інформаційних комунікаційних мереж є:

- моніторинг роботи інформаційної мережі на інфраструктурному, проміжному і базовому рівні моделі;
- аналіз потоку навантаження на кожному рівні, що враховує характер навантаження;
- синтез методів розрахунку ймовірнісно-часових параметрів мережевих вузлів при обробці потоку навантаження у мережевих пристроях із метою побудови адекватного прогнозу значень імовірнісно-часових параметрів для вироблення управляючих впливів на режим функціонування інформаційної мережі в реальному масштабі часу.

Для оцінки якості функціонування інформаційної комунікаційної мережі необхідно оперуватися набором критеріїв висунутих для забезпечення ефективного використання мережевих ресурсів. Це дасть змогу забезпечити низькі затримки обслуговування, високі пропускні здатності, захищеності даних в процесі передавання потоків інформації. Традиційний підхід побудови інформаційних комунікаційних систем та мереж, пов'язаний із чіткою регламентацією на всіх рівнях мережевої взаємодії, для гарантування високої якості обслуговування необхідно планувати інформаційну мережу із значним запасом ресурсів [1]. По скільки такий підхід викликаний властивістю само подібності потоку навантаження, що характеризується суттєвими локальними флуктуаціями пропускної здатності та “тяжкими хвостами”. Тому в процесі організації та плануванні міжмережевої взаємодії крім середнього значення пропускної здатності інформаційної мережі необхідно враховувати її пікові значення. У результаті в такій мережі потрібно передбачити значні запаси за критерієм пропускної здатності, як