

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
(повна назва)
Кафедра _____ Інформаційних управляючих систем _____
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____

Дослідження методів розподілу рекламного трафіку в системах управління
інтернет-рекламою _____
(тема)

Виконав:
студент 2 курсу, групи ІУСТм-18-1
_____ Савін С.А. _____
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки _____
(код і повна назва спеціальності)

Тип програми освітньо-професійна _____
(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні управляючі
системи та технології _____
(повна назва освітньої програми)

Керівник _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис) _____
(прізвище, ініціали)

2019р.

Завдання

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
(повна назва)

Кафедра _____ Інформаційних управляючих систем _____
(повна назва)

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Інформаційні управляючі системи та технології _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ

НА АТЕСТАЦІЙНУ РОБОТУ

студентові _____ Савіну Станіславу Андрійовичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Дослідження методів розподілу рекламного трафіку в системах управління інтернет-рекламою _____

затверджена наказом університету від _____ 20__ р. № _____

2. Термін подання студентом роботи до екзаменаційної комісії _____ 20__ р.

3. Вихідні дані до роботи _____ Науково-технічні публікації та інтернет джерела з тематики _____
атестаційної роботи _____

4. Перелік питань, що потрібно опрацювати в роботі Дослідження особливостей рекомендаційної підсистеми в системі управління інтернет-рекламою, аналіз гібридних методів фільтрації які можна застосувати для розподілу та персоналізації рекламного трафіку для рекламодавця, аналіз традиційних методів представлення знань, удосконалення методу управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань, Експериментальна перевірка удосконаленого методу для підтримки розподілу та персоналізації рекламного трафіку для рекламодавців в системі показу реклами компаній прокату автомобілів

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз літератури та інтернет-джерел		вик.
2	Постановка задачі		вик.
3	Обробка матеріалу		вик.
4	Дослідження особливостей рекомендаційних підсистем		вик.
5	Аналіз гібридних методів фільтрації		вик.
6	Аналіз традиційних методів представлення знань		вик.
7	Удосконалення методу управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань		вик.
8	Апробація результатів дослідження на прикладі		вик.
9	Написання пояснювальної записки		вик.
10	Підготовка презентації		вик.
11	Перевірка на плагіат		вик.
12	Нормконтроль		вик.
13	Захист		

Дата видачі завдання _____ 20__ р.

Студент _____
(підпис)

Керівник роботи _____ проф. каф. ІУС Чалий С.Ф.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до магістерської атестаційної роботи містить: 86 стор., 4 розділи, 38 рис., 3 таблиці, 31 джерела.

Об'єктом дослідження є процес розподілу та персоналізації рекламного трафіку для рекламодавця в системі управління інтернет-рекламою.

Метою роботи є дослідження гібридних методів фільтрації у процесах розподілу та персоналізації рекламного трафіку для рекламодавця системі управління інтернет-рекламою.

Предметом дослідження є методи персоналізації трафіку та формування рекомендацій для рекламодавця.

В результаті виконання магістерської атестаційної роботи було виконано дослідження процесу розподілу та персоналізації рекламного трафіку для рекламодавця, досліджені особливості рекомендаційної підсистеми в системі управління інтернет-рекламою, проаналізовані та порівняні гібридні методи фільтрації, які можна застосувати для розподілу та персоналізації рекламного трафіку для рекламодавця, були проаналізовані традиційні методи представлення знань та був удосконалений метод розподілу рекламного трафіку в системі управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань.

Результатом роботи є удосконалений метод, який підвищує ефективність розподілу рекламного трафіку поміж рекламодавців, який в практичному аспекті показав збільшення середнього прибутку.

РЕКЛАМА, РОЗПОДІЛ ТРАФІКУ, СИСТЕМА УПРАВЛІННЯ ІНТЕРНЕТ-РЕКЛАМОЮ, СИСТЕМА, РЕКЛАМНИЙ ТРАФІК, РЕКЛАМОДАВЕЦЬ.

ABSTRACT

Explanatory Note to master certification work contains 86 pages, 4 sections, 38 pictures, 3 tables, 31 sources.

The object of the study is the process of distribution and personalization of advertising traffic to the advertiser in the Internet advertising management systems.

The purpose of this work is to investigate hybrid filtering methods in the processes of distribution and personalization of advertising traffic for the advertiser online advertising management system.

The subject of the study is methods of personalizing traffic and formulating recommendations for the advertiser.

As a result of the performance of the master's appraisal work, a study of the process of distribution and personalization of advertising traffic for the advertiser was carried out, the features of the recommendation subsystem in the Internet advertising management system were investigated, hybrid filtering methods that could be applied to the distribution and personalization of advertising traffic for advertising were analyzed and compared traditional methods of presenting knowledge and the method of distribution of advertising traffic in the Internet management system has been improved internet-based advertising pre-filtering based on knowledge.

The result is an improved method that improves the distribution of advertising traffic among advertisers, which in practical terms has shown an increase in average revenue.

ADVERTISING, TRAFFIC DISTRIBUTION, INTERNET ADVERTISING MANAGEMENT SYSTEM, SYSTEM, ADVERTISING TRAFFIC, ADVERTISING.

ПЕРЕЛІК СКОРОЧЕНЬ, УСЛОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ ТА ТЕРМІНІВ

ІС – інформаційна система;

КФ – колаборативна фільтрація;

ЛМП – логічна модель представлення знань;

ЛПП – логічна модель представлення знань;

АСПР – автоматизованій системі підтримки прийняття рішень;

ЕС – експертна система;

ЗМІСТ

ВСТУП.....	7
1 ДОСЛІДЖЕННЯ МЕТОДІВ РОЗПОДІЛУ РЕКЛАМИ ЯК ОБ'ЄКТУ	
УПРАВЛІННЯ В СИСТЕМАХ УПРАВЛІННЯ ІНТЕРНЕТ-РЕКЛАМОЮ	10
1.1 Дослідження особливостей інтернет-реклами як об'єкту управління	10
1.2 Дослідження систем управління інтернет-рекламою.....	20
1.3 Дослідження методів розподілу рекламного трафіку як об'єкту персоналізації	24
1.4 Постановка задачі дослідження.....	30
2. ДОСЛІДЖЕННЯ ГІБРИДНИХ МЕТОДІВ ФІЛЬТРАЦІЇ	32
2.1 Особливості методу фільтрації на основі знань	32
2.2 Дослідження можливостей використання обчислювального інтелекту для формування правил в системах розподілення трафіку.	41
2.3 Удосконалення методу управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань	44
3 ДОСЛІДЖЕННЯ ОТРИМАНИХ НАУКОВИХ РЕЗУЛЬТАТІВ	48
3.1 Дослідження методу фільтрації на основі знань	48
3.2 Технологія використання удосконаленого методу розподілення трафіку в системах управління інтернет-рекламою	50
4. ПРАКТИЧНЕ ЗАСТОСУВАННЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ	52
4.1 Програмна реалізація методу розподілу рекламного трафіку в системах управління інтернет-рекламою.....	52
4.2 Експериментальна перевірка удосконаленого методу розподілу рекламного трафіку	62
ВИСНОВКИ	66
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ	67
ДОДАТОК А	71

ВСТУП

Інтернет-реклама є одним з найбільш вагомих елементів маркетингової діяльності, адже вона має суттєвий економічний вплив як в цілому на економіку підприємства, так і на продукти, споживачів, конкурентів і партнерів.

Реклама забезпечує такі можливості з просування товарів та послуг:

- реклама надає потенціал для отримання споживчого досвіду, адже її можна аналізувати не тільки як один з елементів бізнесу, але і як важливий елемент торгівлі;

- реклама проявляє великий вплив на сферу діяльності компанії, стимулює введення останніх розробок, допомагає збільшенню асортименту товарів і послуг, стимулює використання і впровадження методів менеджменту та новітніх методів контролю якості виробництва;

- прагнення до конкурентоспроможності підштовхує підприємців до вивчення і впровадження реклами в свою компанію, що призводить, при грамотному її використанні, до розширення і вдосконалення виробництва, а значить і значного збільшення прибутку підприємства;

Важливо відзначити, що основним в організації рекламної діяльності є створення та затвердження портрета користувача і пошук цільової аудиторії, яка буде бачити у себе рекламу вашого товару. Неправильно підібрані користувачі не будуть зацікавлені в товарі підприємства і гроші, витрачені на рекламну компанію будуть просто використані без вигоди для рекламодавця;

Отже ми можемо зробити висновок, що майже найголовнішим в показі реклами є пошук цільової аудиторії та розподілення рекламного трафіку між рекламодавцями.

Для вирішення наведених вище проблем та персоналізації реклами використовують рекомендаційні системи. Така система показує рекламу користувачам на основі їх профілів, забезпечуючи їм індивідуальні

рекомендації. Рекомендаційні систем застосовуються не тільки для персоналізації реклами для користувача, але й для розподілу реклами між рекламодавцями базуючись на параметрах продукту який він пропонує.

Рекомендаційні системи для надання актуальної реклами споживачу та надання відповідного трафіку рекламодавцям переважно використовують гібридні методи фільтрації. Гібридні методи можуть поєднувати в собі різноманітні методи фільтрації, але найчастіше поєднують методи колаборативної фільтрації, фільтрації на основі параметрів продуктів та фільтрації на основі знань.

Принцип колаборативної фільтрації передбачає, що є кластери користувачів, які ведуть себе подібним чином і мають порівнянні потреби та вподобання. Завдання колаборативного методу полягає в тому, щоб визначити подібних користувачів та отримати рекомендації щодо показу їм реклами. Тоді як парадигма, що базується на параметрах продуктів, передбачає рекомендацію продуктів, схожих і на ті, що сподобалися користувачеві в минулому. Рекомендація на основі знань потребує додаткового джерела інформації, а також явних знань про персональні вимоги користувачів. Ці знання можуть, наприклад, набувати форми логічних обмежень, які відображають потреби користувача на властивості предмета.

Багатокористувацькі схеми корисності та конкретні заходи подібності є альтернативними механізмами представлення знань. Ці моделі знань можуть бути отримані від третьої сторони, наприклад, доменних експертів; дізнаватися на основі даних про минулі транзакції; або використовувати комбінацію обох.

Сучасні гібридні методи поєднують у собі три базові методи фільтрації, але вони не вирішують ряд проблем які має кожен з цих методів та які впливають на якість рекомендацій, зокрема проблем холодного старту та спам-атак.

Проблема спам-атаки полягає в обмані рекомендаційної системи за рахунок великої кількості однотипних запитів.

Проблема холодного старту виникає при наданні новим користувачам рекомендацій. Нові користувачі ще не переглянули достатньо рекламних матеріалів, тому рекомендаційна система не може дати коректних прогнозів.

Для вирішення цієї проблеми в системах використовують початкові рейтинги, які засновуються на тому, що користувач повинен оцінити певну кількість товарів або зробити певну кількість дій для формування початкового рейтингу. Однак ці первинні рейтинги можуть ввести упередженість в систему, та не завжди можуть бути використані в системах. Проблема холодного запуску ще впливає на нові продукти, оскільки методи колаборативної фільтрації та аналізу продукту не є достатньо ефективними навіть за умови застосування їх обох у гібридних методах.

Для вирішення вказаної проблеми необхідно комбінувати інформацію про користувача, інформацію про ефективність показаної реклами та інформацію продуктів, що потребує дослідження та удосконалення рекомендаційних методів, що застосовуються для розподілу рекламного трафіку. Зазначене свідчить про актуальність розглянутої в роботі проблеми.

Отриманим результатом роботи є удосконалений метод, який підвищує ефективність рекомендацій показаної реклами користувачу та ефективність розподілення рекламного трафіку поміж рекламодавцями.

Отримані результати виконання магістерської атестаційної роботи оформлені у вигляді пояснювальної записки.[1]

1 ДОСЛІДЖЕННЯ МЕТОДІВ РОЗПОДІЛУ РЕКЛАМИ ЯК ОБ'ЄКТУ УПРАВЛІННЯ В СИСТЕМАХ УПРАВЛІННЯ ІНТЕРНЕТ- РЕКЛАМОЮ

1.1 Дослідження особливостей інтернет-реклами як об'єкту управління

Для того щоб зрозуміти, що таке інтернет-реклама та навіщо її використовувати заглянемо в історію інтернет-реклами, де стає ясно, що її роль в якості повноцінного інструменту маркетингу виникла зовсім недавно - якихось 15 років тому та лише на початку 90-х років минулого століття, з появою перших текстових повідомлень рекламного характеру, інтернет перестав бути тільки засобом передачі інформації[2].

Інтернет-реклама, також відома як цифрова реклама або веб-реклама - це форма маркетингу та реклами, яка використовує Інтернет для доставки рекламних маркетингових повідомлень споживачам. Інтернет-реклама є невід'ємною частиною електронної комерції.

Досить скоро виникло і поняття інтернет-маркетингу: спочатку як доповнення до традиційних форм бізнесу; потім як альтернативи і, нарешті, як повноцінного учасника економічного життя. Тоді ж відбулося і «перетворення» користувачів інтернету в потенційних і реальних покупців. Та й реклама в інтернеті трансформувалася з реклами товарів і послуг в електронний ринок: рекламу сайтів, що пропонують ці товари і послуги.

Для того щоб розібратися з типами реклами розглянемо що таке інтернет комерція та визначимо які типи вона має.

Електронна комерція, також відома як електронна комерція чи Інтернет-торгівля, стосується купівлі та продажу товарів чи послуг за допомогою Інтернету та передачі грошей та даних для здійснення цих операцій.

Електронна комерція часто використовується для позначення продажу фізичних товарів в Інтернеті, але вона також може описувати будь-який вид комерційної операції, який сприяє продажам через Інтернет.

Існує чотири основні типи моделей електронної комерції, які можуть описати майже кожну транзакцію, яка відбувається між споживачами та бізнесом.

Бізнес для споживачів (B2C). Це коли бізнес продає товар або послугу окремому споживачеві (наприклад, ви купуєте пару взуття у інтернет-магазину).

Бізнес для бізнесу (B2B). Це коли бізнес продає товар або послугу іншому бізнесу (наприклад, бізнес продає програмне забезпечення як послугу для інших підприємств, які можуть його використовувати).

Споживач до споживача (C2C). Це коли споживач продає товар або послугу іншому споживачеві (наприклад, ви продаєте старі меблі на eBay іншому споживачеві).

Споживач для бізнесу (C2B). Це коли споживач продає власні товари чи послуги бізнесу чи організації (наприклад, популярний блогер пропонує бізнесу рекламу своїй інтернет-аудиторії в обмін на плату або фотограф ліцензує свої фотографії для використання в бізнесі та продає їх).

Інтернет-реклама пропонує широкі можливості для зв'язку з споживачами, тому давайте розглянемо деякі найвпливовіші види інтернет-реклами на ринку:

- контекстна реклама;
- Email маркетинг;
- реклама в соціальних мережах;
- банерна реклама;

Маркетинг пошукових систем або контекстна реклама представляє собою текстові, або текстово-графічні модулі, які демонструються на майданчиках, що знаходяться в контексті аудиторії об'єкта реклами, або конкретного

користувача. Простіше кажучи, контекстна реклама може відповідати тематичним спрямуванням майданчики, на якій демонструється, або відображається виходячи з пошукових запитів / інтересів самого користувача. Контекстна реклама один з найбільш ефективних варіантів інтернет-реклами.

Великою перевагою контекстної реклами є її таргетованість - цільове орієнтування на інтереси аудиторії. Контекстні оголошення демонструються тільки в тому випадку, якщо користувач вже виявляв інтерес до даної тематики, або коли рекламований продукт відповідає тематиці площадки, на якій розміщується. При цьому оплата за демонстрацію реклами проводиться покліково, тобто за конкретну дію користувача.

Власні сервіси контекстної реклами є у більшості великих пошукових систем. У просторі Рунета найбільшою популярністю користуються мережі Google AdWords, Яндекс.Директ.

Використання контекстної реклами дає швидкий, але короткостроковий результат - реклама "працює", поки йдуть покази. Серед очевидних особливостей цього виду просування також варто відзначити високі розцінки. Майданчики відображають оплачені результати на основі аукціону з рейтингу оголошень. Компанії називають свою ціну за певне ключове слово, а майданчик аналізує якість та релевантність вмісту.

Формула рейтингу оголошення має вигляд: **Рейтинг оголошення = ставка ціни за клік × показник якості.**

Ціна "за клік" утворена за принципом аукціону призводить до досить серйозного рівня витрат на рекламу, особливо в популярних тематичних напрямках. Вона діє більш цілеспрямовано. У цьому випадку мова йде не стільки про відвідувачів, які почали пошук, скільки про зацікавлених потенційних клієнтів, які вивчають результати свого пошуку.

Розміщення контекстної реклами специфічно: вона з'являється тільки у відповідь на певний пошуковий запит: або на сторінці з результатами в самій пошуковій системі, або на тематичному сайті. Короткий текстове повідомлення

(іноді з графікою) і обов'язкове посилання (шлях) до основного сайту - ось звичайна форма такої інтернет-реклами. Її мета - якщо не «переманити», то напевно перенаправити відвідувача в потрібному напрямку. Ефективність контекстної реклами висока і дозволяє точно визначати рентабельність рекламних вкладень: оплата йде за кількістю користувачів, «клікнувши» на оголошення.

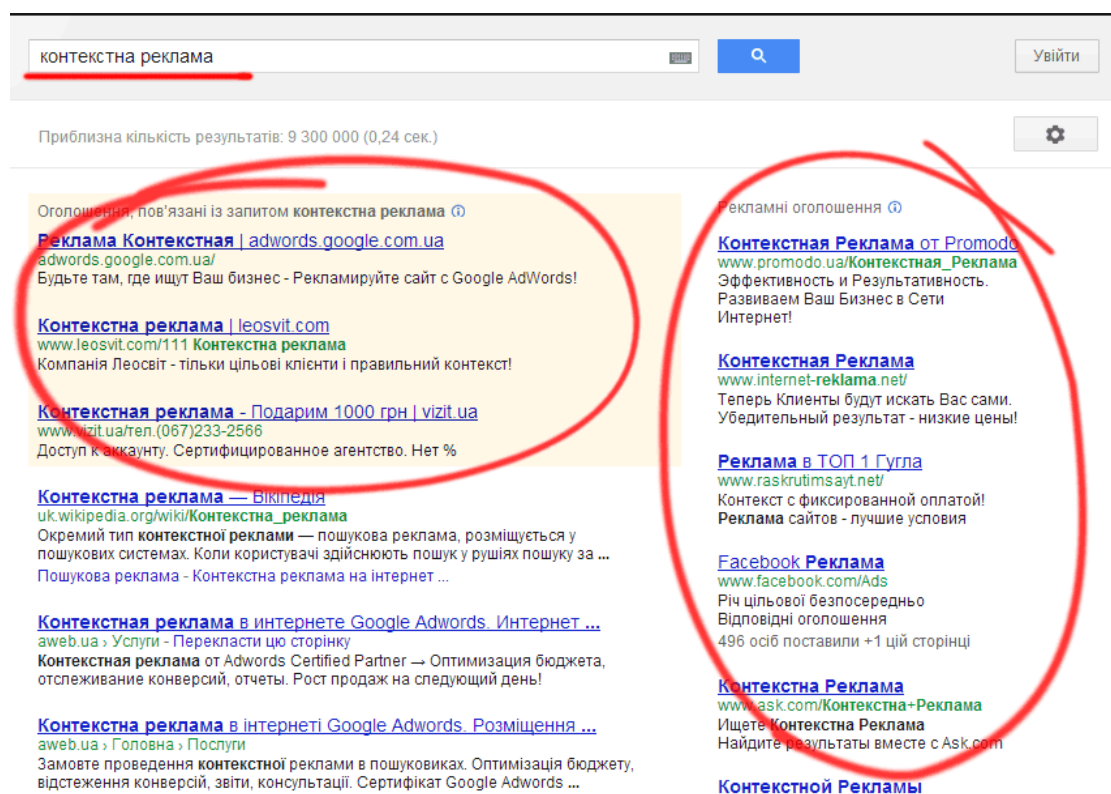


Рисунок 1.1 – приклад контекстної реклами

Email маркетинг це використання електронної пошти для просування товарів та послуг. Але кращим визначенням email маркетингу є використання електронної пошти для розвитку відносин з потенційними клієнтами. Маркетинг електронною поштою - це один із сегментів інтернет-маркетингу, який охоплює інтернет-маркетинг через веб-сайти, соціальні мережі, блоги тощо. Він по суті такий же, як пряма пошта, за винятком того, що замість того, щоб відправляти пошту через поштову службу, повідомлення надсилаються в електронному вигляді електронною поштою.

Більшу частину подібної електронної пошти ми всі відправляємо у папку «спам», і це хороший приклад поганого email-маркетингу, а у кращому випадку маркетинг по електронній пошті дозволяє підприємствам інформувати своїх клієнтів та адаптувати маркетингові повідомлення для своїх клієнтів.

Конкретні групи клієнтів так і єдиний клієнт можуть бути цільовими особами. Наприклад, пропонування індивідуальним клієнтам спеціальних пропозицій щодо товарів та послуг у день народження клієнта, наприклад, є одним із прикладів персоналізації маркетингу електронною поштою. (Ресторан може надіслати електронний лист клієнтам у день народження, пропонуючи 50% знижки,) Маркетинг по електронній пошті допомагає бізнесу розвиватися та підтримувати стосунки з клієнтом з часом, що, сподіваємось, призводить до збільшення продажів та підвищення лояльності клієнтів.

Найкращі практики маркетингу електронною поштою включають розробку власних стратегій та дизайну. Електронну пошту також слід оптимізувати для мобільного використання, оскільки згідно зі статистикою понад половину електронних листів відкриваються на мобільних пристроях.

Дві великі переваги email маркетинг - ціна та простота. Електронна пошта - це недорогий спосіб реклами вашої компанії та її товарів та послуг порівняно з багатьма іншими видами маркетингу. Також надзвичайно просто налаштувати та відстежувати маркетингову кампанію електронною поштою, що робить її дуже доступним типом маркетингу для малого бізнесу.

Також можна відправляти новини тим користувачам, які надали необхідну інформацію на веб-сайті, наприклад, надаючи цим потенційним клієнтам не тільки новини про вашу компанію, а й про майбутні події, спеціальні пропозиції та, звичайно, нагадувати користувачам про існування цього бізнесу та його продуктів.

M&S

EST. 1884

Women

Lingerie

Men

Food

Kids

Home

Offers

BUY TWO SHIRTS FOR £30



Stock up for summer with our
great deal on selected £17.50 shirts

SHOP THE OFFER



ORDER BY 10PM FOR NEXT-DAY DELIVERY*

Exclusions apply. [Find out more](#)

f t G+ p YouTube i Mobile app



Find your nearest M&S

Terms, conditions and exclusions apply. See [website](#) for full details.

If you no longer wish to receive marksandspencer.com emails you can [unsubscribe](#) at any time. If you want to update your Sparks email preferences, you can do so [here](#). For more information about how we use data, please see our [Data Protection Policy](#)

Marks and Spencer plc. Registered office: 35 North Wharf Road, London, W2 1NW, United Kingdom.
Registered number: 214436 (England and Wales)

Рисунок 1.2 – приклад email-реклами

У епоху розширення використання соціальних мереж для реклами, маркетинг електронної пошти все ще керує ціною, згідно з дослідженням, проведеним HostPapa:

- 94 відсотки сучасних користувачів мережі використовують електронну пошту, тоді як лише 61 відсоток користуються соціальними мережами;

- 75 відсотків користувачів мережі віком понад 30 років кажуть, що email-маркетинг є найбільш ефективним методом для них, а враховуючи що ця аудиторія найбільш платоспроможна не можна не брати цей факт до уваги;

- email-маркетинг дозволяє розподіляти рекламу за демографічними показниками (вік, дохід, та інше);

- email повідомлення мають набагато більший спектр можливостей форматування та дизайну, ніж реклама в соціальних мережах;

- email-маркетинг пропонує функціонал не менш широкий ніж в контекстній рекламі чи соціальних мережах. Звітування, аналітика, коефіцієнти кліків, відкриті показники, коефіцієнти відмов та конверсії;

Велика перевага електронної пошти над соціальними мережами в тому, що частіше потенційні клієнти швидше бачать email, ніж рекламу чи повідомлення в соціальній мережі. Якщо просто щось опублікувати, це не означає, що всі, хто хоче бачити ваше повідомлення, це побачать. Ваша публікація може навіть не відображатися в новинах користувача. Однак email зберігатиметься в папці "Вхідні", поки не буде прочитано (або видалено).

В ідеалі email маркетинг повинен іти одночасно з рекламою в соціальних мережах. Додавання кнопок "Подобається" або "Поділитися" у ваші маркетингові електронні листи дає додатковий спосіб клієнтам підключатися до вашого бренду. Фрагменти позитивних відгуків шанувальників соціальних мереж можуть бути включені до електронних листів, і, навпаки, публікації в соціальних мережах можуть використовуватися для заохочення шанувальників передплачувати ваші електронні розсилки.

Реклама в соціальних мережах це реклама що відображається соціальних мережах, таких як Facebook, Instagram, Twitter, Pinterest для B2C та LinkedIn для B2B. Компанії адаптують свої новини та акції до цільової аудиторії через соціальні медіа двома способами:

- органічним;
- передплаченим;

Органічний – це коли ви виробляєте кухонний посуд, ви можете розміщувати цінний вміст, як хаки, захоплюючі способи використання своїх продуктів або деякі рецепти для заохочення акцій та створення привабливого іміджу бренду.

- Передплачений - це коли ви можете використовувати функціональні можливості соціальних мереж для бізнесу, показуючи рекламні пости, орієнтовані на вашу аудиторію, залежно від віку, статі, улюблених заходів та інших спільних речей.

Ці невеликі рекламні пости дозволяють максимально використовувати всю інформацію, якою користувачі діляться у соціальних мережах, пропонуючи їм максимально персоналізовану рекламу.

Зараз все більше брендів роблять свою рекламу у соціальних мережах. Можливості та різновиди способів достукатися до цільової аудиторії здаються нескінченними, тому це чудовий спосіб реклами та її персоналізації. Social Ads - це неймовірно вигідний і універсальний рекламний канал, який дає нам можливість будувати конкретні кампанії в соціальних мережах для задоволення вибору різних цілей бізнесу за відносно низьких витрат.

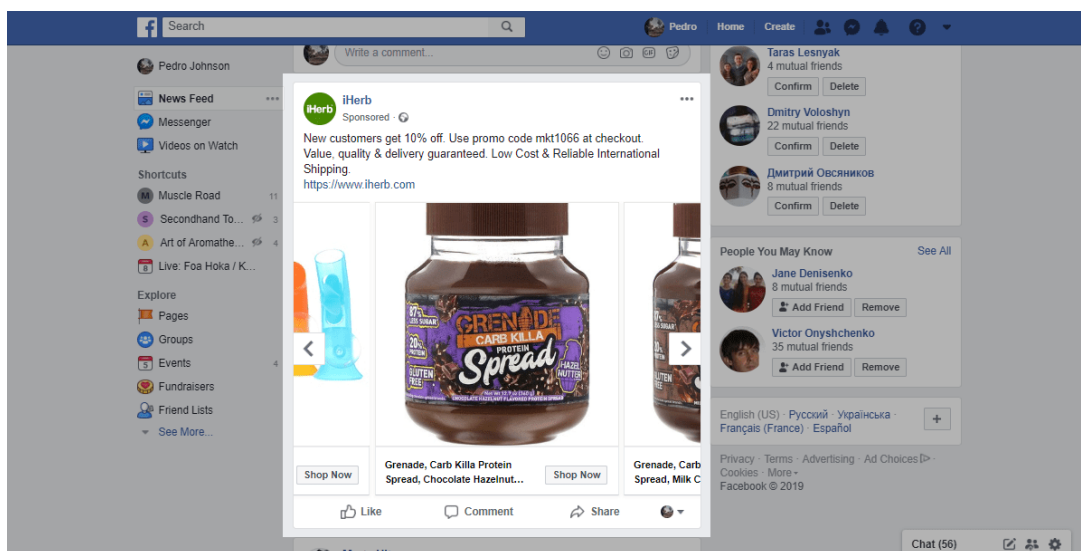


Рисунок 1.3 – приклад реклами в соціальних мережах

Банерна реклама, що розміщується на банерах (графічних носіях) - одна з перших, що з'явилися в рамках веб-простору. Статичні або анімовані рекламні об'єкти орієнтовані на залучення уваги аудиторії. Але назвати їх цільовими досить складно. Як правило, банери застосовують для підвищення впізнаваності товару, торгової марки, бренду, або для поширення інформації про діючі акції, знижки, спецпропозиції.

Серед інших видів інтернет-реклами банерна виділяється відносно невисокою клікабельністю - переходів тут не дуже багато, але зазвичай вони є цільовими. І при правильному виборі майданчика в результаті дають непогану віддачу.

Оплата розміщення банерної реклами зазвичай становить кількістю показів або часу трансляції - плата за клік тут неактуальна. Саме таку схему співпраці пропонують і централізовані майданчики, які організовують розміщення банерів, і приватні власники медіаресурсів.

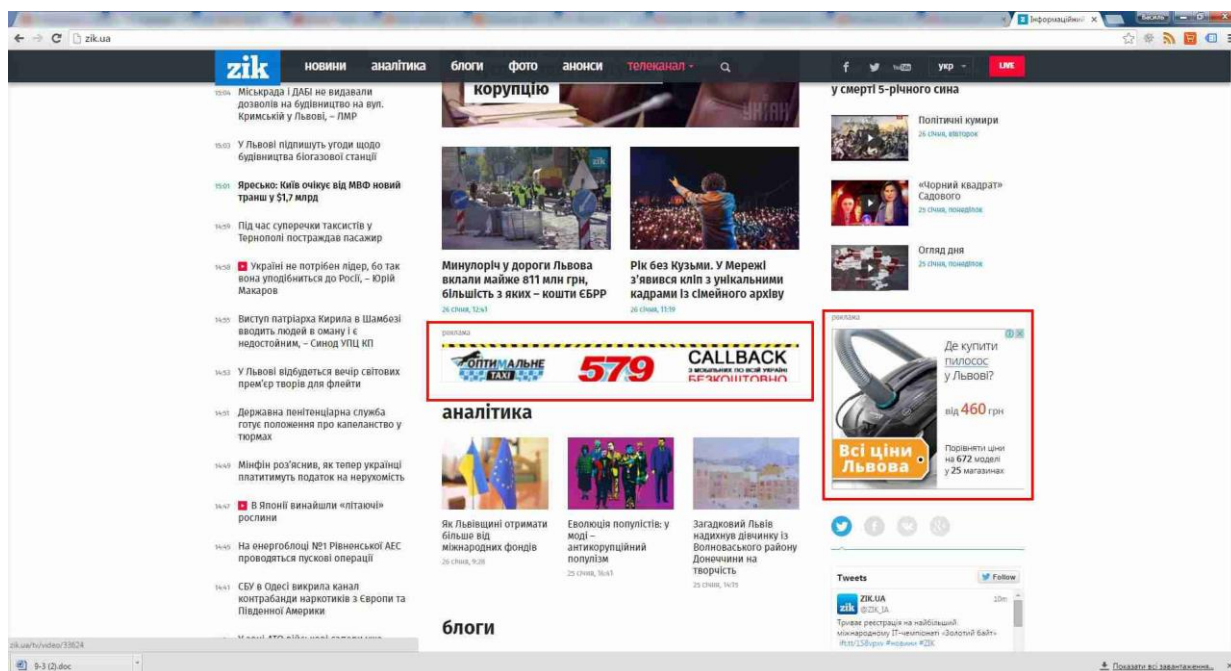


Рисунок 1.4 – приклад банерної реклами

Вона не настільки інформативна, як перераховані вище види реклами в інтернеті. Але і мета її в іншому: зробити бренд впізнаваним і надати йому позитивне емоційне забарвлення. Рекламний матеріал - яскравий графічний банер - містить зовсім небагато тексту, зате нагадує про назву компанії, має пряме посилання на сайт і заклик (купити, замовити, отримати знижку і т.д.). Ефективність медійної (банерної) реклами безпосередньо залежить від кількості показів. Але навіть при невеликій конверсії банери все одно успішно справляються зі своїм головним завданням - нагадують про бренд. Цікава іміджева реклама в кінці кінців обов'язково призведе відвідувача на сайт.

Як і інші рекламні засоби масової інформації, інтернет-реклама часто включає видавця, який інтегрує рекламу в свій Інтернет-контент, і рекламодавця, який забезпечує показ реклами у вмісті видавця. Інші потенційні учасники включають рекламні агентства, які допомагають створювати та розміщувати копію реклами, системи розподілення рекламного трафіку, які технологічно доставляють рекламу та відстежують статистику, та рекламні партнери, які виконують незалежну рекламну роботу для рекламодавця.

1.2 Дослідження систем управління інтернет-рекламою

Для кожного типу реклами існують свої системи управління. Системи управління забезпечують доставку або показ реклами кінцевому споживачеві, розподіляють рекламний трафік між рекламодавцями та забезпечують статистику як рекламодавця так і видавця реклами.

Для управління банерною рекламою можна використовувати DoubleClick Digital Marketing platform або DoubleClick for Advertisers (DFA) - це рішення для управління рекламою та розміщення реклами, яке допомагає агенціям та рекламодавцям керувати всією сферою цифрових рекламних програм.

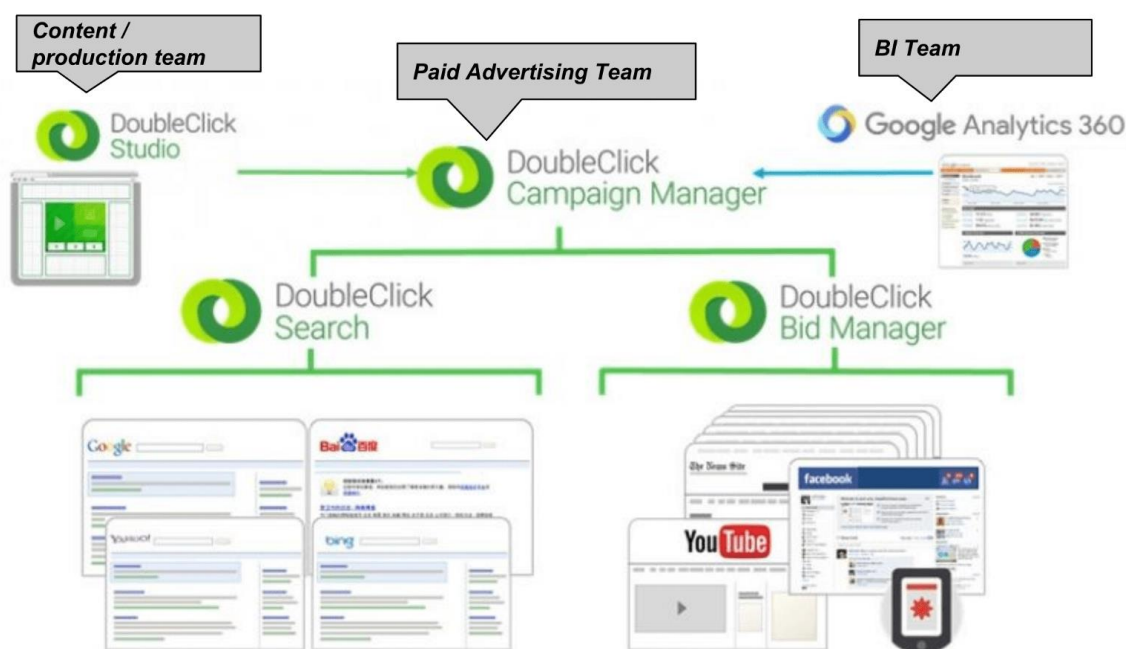


Рисунок 1.5 – структура DoubleClick Digital Marketing platform

DFA впорядковує робочий процес для планування, продажу реклами, таргету, обслуговування, оптимізації та звітності. DFA служить основою платформи управління цифровими оголошеннями DoubleClick для рекламодавців, яка включає DoubleClick Search, DoubleClick Rich Media та DoubleClick Ad Exchange. Разом ці рішення допомагають агенціям та маркетологам освоювати навіть найскладніші стратегії інтернет-реклами.

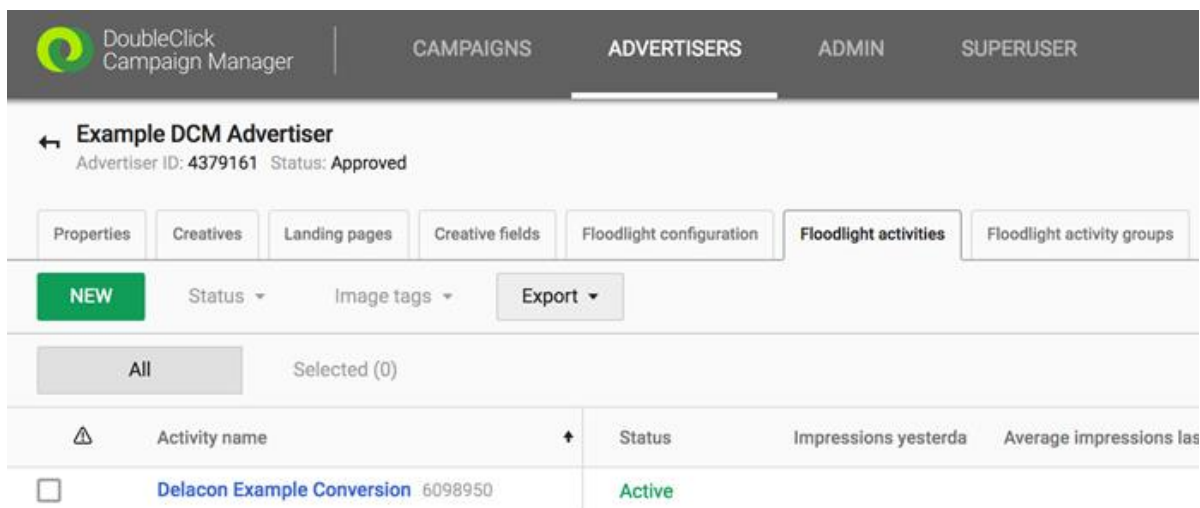


Рисунок 1.6 – Приклад інтерфейсу DoubleClick Digital Marketing platform

DoubleClick Search - це платформа для управління, яка допомагає маркетологам ефективно керувати великими пошуковими кампаніями на різних двигунах або каналах. Це дозволяє консолідувати робочий процес і звітувати про дані кампанії, одночасно підвищуючи ефективність ваших маркетингових кампаній.

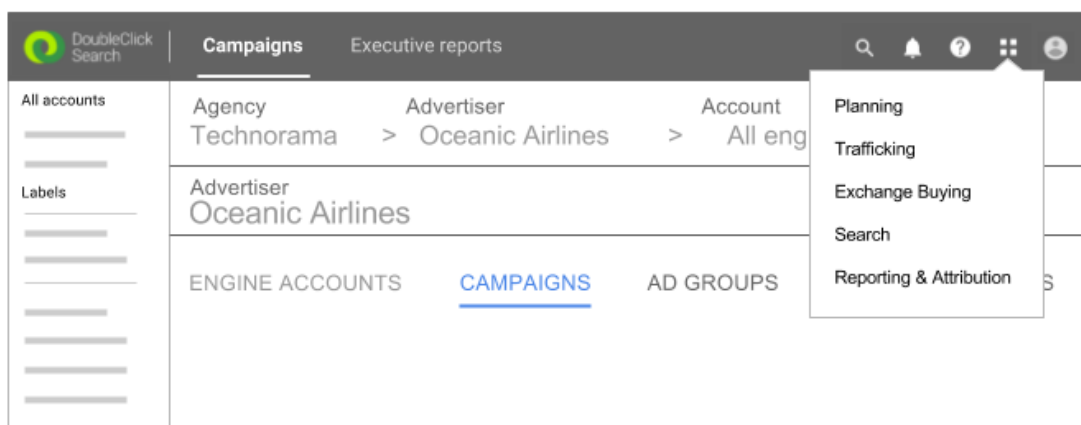


Рисунок 1.7 – Приклад інтерфейсу DoubleClick Search

Робочий процес без DoubleClick Search (DS) тягне за собою моніторинг ваших оголошень і ключових слів у кожній пошуковій системі окремо (насправді DS заявляє, що їх набір скорочує керування часом кампанії на

тринадцять відсотків.). Компіляція вашої ефективності на кожному каналі в аналітичних цілях була б кропітка за трудомісткою. Не кажучи вже про те, що ви б не мали автоматизованих можливостей призначення ставок.

DoubleClick Rich Media або DoubleClick Studio- це сервіс для створення різних типів рекламних матеріалів для відображення мультимедійних оголошень на цільовому веб-сайті. Творчі типи (також їх називають форматами) призначаються, коли дизайнер робить нове оголошення в студії. Мультимедійні оголошення включає в себе розширені функції, такі як відео, аудіо та інші елементи, які спонукають глядачів до взаємодії з вмістом.

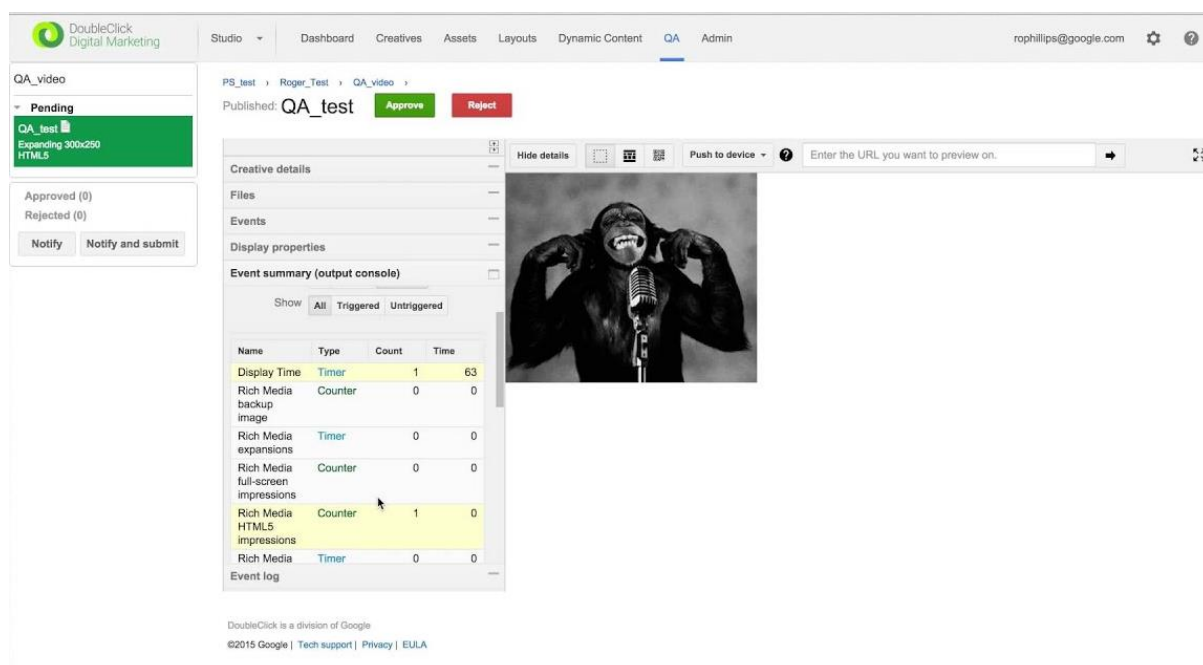


Рисунок 1.8 – Приклад інтерфейсу DoubleClick Studio

У той час як текстові оголошення продаються зі словами, а медійні - із зображеннями, рекламні мультимедіа пропонують більше способів залучення аудиторії до реклами. Оголошення може розширюватися, розповсюджуватися тощо. Ви можете отримати зведені показники щодо поведінки вашої аудиторії, включаючи кількість розширень, кількість взаємодій та доповнень відео, щоб отримати детальні дані про успіх вашої кампанії.

Мультимедіа дозволяє агенціям створювати складні реклами, які можуть викликати сильну реакцію користувачів. Використовуючи технологію HTML5, об'єкти можуть включати кілька рівнів вмісту в одне розташування: відео, ігри, твіти з оголошення тощо. Якщо у вас є проста мета - генерувати кліки або більш амбітну мету - створити поінформованість про бренд, мультимедіа - це дуже зручний формат.

DoubleClick Ad Exchange - одна з найбільших рекламних бірж в Інтернеті, якою керує Google. Це ринок рекламодавців та видавців. Видавці пропонують свій рекламний ресурс за мінімальну ціну за тисячу показів, а рекламодавці пропонують ставку на інвентар. Це чудова платформа, яка дозволяє попиту задовольнити пропозицію на ринку реклами.

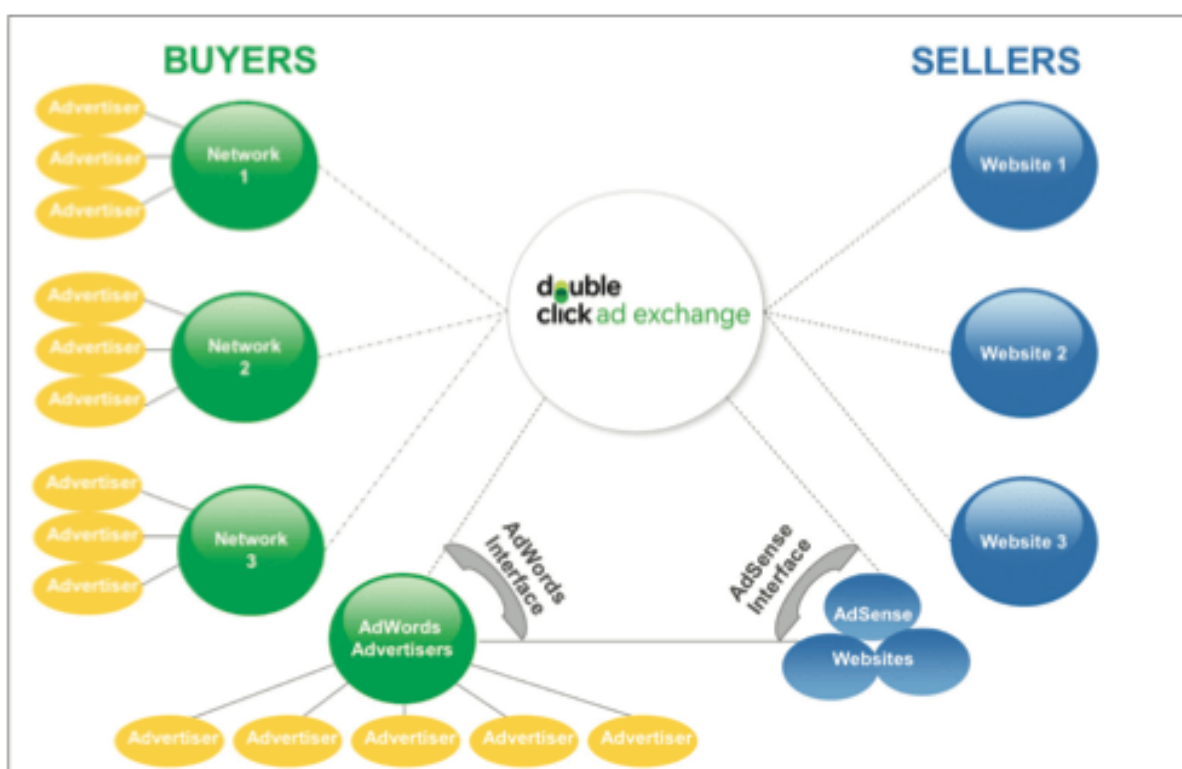


Рисунок 1.9 – структура DoubleClick Ad Exchange

DBM: DoubleClick Bid Manager - це платформа DoubleClick на базі попиту, яка надає рекламодавцям інструменти для придбання показів,

використовуючи торги в режимі реального часу та прямі програмні покупки (AdExchanges).

Таким чином, це як ваші медійні кампанії Google (GDN) під рекламними програмами, де ви можете запускати зображення та відео, але замість того, щоб використовувати лише рекламні ресурси Google, ви можете орієнтуватися на декілька рекламних ресурсів через AdExchanges через DoubleClick Bid Manager.

Рекламодавці можуть націлювати на конкретні аудиторії, конкретні типи сайтів та вміст для розміщення оголошень, таких як банери, відео, мобільні та динамічні оголошення.

Переглянувши усі ці системи управління інтернет-рекламою ми можемо зробити висновок, що найголовніше це правильний розподіл рекламного трафіку між рекламодавцями та максимальна персоналізація реклами для користувачів.

1.3 Дослідження методів розподілу рекламного трафіку як об'єкту персоналізації

Трафік веб-сайту стосується користувачів Інтернету, які відвідують веб-сайт. Веб-трафік вимірюється в відвідуваннях, іноді їх називають "сеансами" або "сесіями", і це звичайний спосіб вимірювання ефективності онлайн-бізнесу при залученні аудиторії [3,4].

Рекламний трафік є різновидом трафіку веб сайту, що використовується видавцями реклами для продажу її на своїх сайтах рекламодавцям або біржам реклами. Виходячи з цього виникає проблема грамотного розподілення, рекламного трафіку між рекламодавцями та біржами. Системи управління інтернет-рекламою повинні персоналізувати трафік який перенаправляється рекламодавцям за для максимальної ефективності відображення реклами. Вони

повинні аналізувати профілі користувачів та враховувати параметри продуктів які надають рекламодавці.

Для вирішення наведених вище проблем використовують рекомендаційні системи. Рекомендаційні системи використовуються для персоналізації реклами, але в випадку системи управління інтернет-рекламою може використовуватися й для розподілу трафіку між рекламодавцями. Система показує рекламу користувачам на основі їх профілів, забезпечуючи їм індивідуальні рекомендації та бере для відображення продукти найбільш відповідних рекламодавців на основі характеристик продукту та статистики. Тому рекомендаційні системи широко застосовуються не тільки для персоналізації реклами для користувача, але й для розподілу реклами між рекламодавцями.

Рекомендаційні системи для надання актуальної реклами споживачу та надання відповідного трафіку рекламодавцям переважно використовують різні методи фільтрації, але найчастіше використовуються колаборативна фільтрація або гібридні методи фільтрації на її базі [5]. Гібридні методи можуть поєднувати в собі різноманітні методи фільтрації, але найчастіше поєднують методи колаборативної фільтрації, фільтрації на основі параметрів продуктів та фільтрації на основі знань[6].

Як вже було сказано, ми зосередимось на трьох базових методах фільтрації: колаборативної, на основі параметрів продуктів та на основі знань.

Принцип колаборативної фільтрації передбачає, що є кластери користувачів, які ведуть себе подібним чином і мають порівнянні потреби та вподобання [7]. Завдання колаборативного методу полягає в тому, щоб визначити подібних користувачів та отримати рекомендації щодо показу їм реклами. Тоді як парадигма, що базується на параметрів продуктів, дотримується "більш того ж" підходу, рекомендуючи продукти, схожі на ті, що сподобалися користувачеві в минулому, рекомендація на основі знань передбачає додаткове джерело інформації та явні знання про персоналізацію.

Ці знання можуть, наприклад, набувати форми логічних обмежень, які відображають потреби користувача на властивості предмета.

Існує ще велика кількість методів фільтрації, які при об'єднанні з колаборативною фільтрацією, фільтрацією на основі пам'яті або на основі моделі можуть значно вдосконалити результати.

Наведемо кілька методів які можна використовувати з колаборативною фільтрацією та фільтрацією на основі пам'яті або на основі моделі [8, 9].

Таблиця 1.1 – порівняння гібридних методів фільтрації

Гібрид КФ з:	КФ об'єднана з	Сильні сторони	Слабкі сторони
Фільтрація на основі контенту	На основі пам'яті	– Зменшення проблеми нового елемента – Зменшення проблеми розрідження – Зменшення проблеми сірих овець – Більш висока точність, ніж чиста КФ на основі пам'яті	– Нова проблема з користувачем – Погана масштабованість для великих наборів даних – Різноманітність рекомендацій невисока
	На основі моделі	– Зниження проблеми сірих овець – Підвищення точності прогнозування щодо чисто модельованої КФ – масштабований	– Нижня точність порівняно з гібридом КФ на основі пам'яті з фільтрацією на основі контенту – Різноманітність рекомендацій невисока
Демографічна фільтрація	На основі пам'яті	– Полегшення проблеми нового користувача – Більш висока точність, ніж чиста КФ на основі пам'яті	– Демографічні дані важко отримати – Погана масштабованість для великих наборів даних – Проблема сірих овець
	На основі моделі	– Підвищення точності прогнозування щодо чистого КФ на основі моделей – масштабованість	– Демографічні дані важко отримати – Проблема сірих овець – Низька точність порівняно з гібридом КФ на основі пам'яті з демографічною фільтрацією

Продовження таблиці 1.1

Гібрид КФ з:	КФ об'єднана з	Сильні сторони	Слабкі сторони
Фільтрація на основі знань	На основі пам'яті	– Зниження проблеми холодного пуску – Проблема зменшення розрідженості – Зниження проблеми сірих овець – Більш висока точність, ніж чистий КФ на основі пам'яті – Швидко реагуючи на зміни налаштувань користувача	– Погана масштабованість для великих наборів даних – Інженерія знань
	На основі моделі	– Зниження проблеми сірих овець – Підвищення точності прогнозування щодо чистої КФ на основі моделей – Масштабованість – Швидко реагує на зміни налаштувань користувача	– Низька точність порівняно з гібридом КФ на основі пам'яті з фільтрацією на основі знань – Інженерія знань
Семантична фільтрація	На основі пам'яті	– Зниження проблеми холодного пуску – Зменшення проблеми розрідженості – Зниження проблеми сірих овець – Більш висока точність, ніж чистий КФ на основі пам'яті – Швидко реагує на зміни налаштувань користувача – Більша різноманітність рекомендацій	– Погана масштабованість для великих наборів даних – Інженерія знань
	На основі моделі	– Зниження проблеми сірих овець – Підвищення точності прогнозування щодо чистої масштабованої КФ на основі моделей – Швидко реагує на зміни налаштувань користувача – Більша різноманітність рекомендацій	– Низька точність порівняно з гібридом КФ на основі пам'яті з фільтрацією на основі знань – Інженерія знань

Продовження таблиці 1.1

Гібрид КФ з:	КФ об'єднана з	Сильні сторони	Слабкі сторони
Контекстно-залежна фільтрація	На основі пам'яті	– Створення цінних рекомендацій у динамічних умовах	– Проблема з холодним стартом – Проблема зрідженості – Проблема сірих овець – Погана масштабованість для великих наборів даних
	На основі моделі	– Створення цінних рекомендацій у динамічних умовах – масштабованість	– Проблема сірих овець – Низька точність порівняно з гібридом КФ на основі пам'яті з контекстно-залежною фільтрацією
Колаборативна фільтрація	На основі пам'яті та моделі	– Проблема зменшення розрідженості та холодного запуску КФ на основі пам'яті – Поліпшення масштабованості КФ на основі пам'яті – Підвищення точності моделювання КФ на основі моделей	– Проблема сірих овець
	На основі даних користувача та продукту	– Проблема зменшення розрідженості – Полегшення проблеми холодного запуску – Більш висока точність	– Проблема сірих овець – Погана масштабованість для великих наборів даних

В таблиці 1.1 порівняні фільтрація на основі контенту, демографічна фільтрація, фільтрація на основі знань, семантична фільтрація та контентно-залежна фільтрація

Більшість комерційних систем рекомендацій на практиці базуються на техніці КФ. Системи КФ покладаються виключно на рейтинги користувачів (а іноді і на демографічну інформацію) як єдині джерела знань для створення

пропозицій для своїх користувачів [10, 11]. Таким чином, ніяких додаткових знань - таких як інформація про доступні фільми та їх характеристики - не потрібно вводити та підтримувати в системі. Методи рекомендацій на основі змісту використовують різні джерела знань, щоб прогнозувати, чи сподобається користувачеві якийсь предмет [12, 13]. Основні джерела знань, що використовуються на основі контенту систем, включають інформацію про категорії та жанри, а також ключові слова, які часто можуть бути автоматично вилучені з описів текстових елементів.

Подібно до КФ, головна перевага рекомендацій, що базуються на змісті — порівняно низькі витрати на здобуття та підтримку знань. Як КФ, так і алгоритми рекомендацій на основі контенту мають свої переваги. Однак існує багато ситуацій, для яких ці підходи - не найкращий вибір. Нові рекламодавці приходять не дуже часто, тому система буде працювати не ефективно через низьку кількість доступних рейтингів. Також рекламодавці можуть висувати свої вимоги та змінювати їх, що буде призводити до повного перерозподілення трафіку. Формулювання таких вимог не є типовим для чистої рекомендаційної структури на основі КФ та на основі змісту. Тому є сенс звернути увагу на фільтрацію на основі знань.

Рекомендаційні системи на основі знань допомагають нам вирішувати вищезазначені проблеми [15]. Перевага цих систем полягає в тому, що проблем з нарощуванням не існує, оскільки для розрахунку рекомендацій не потрібні дані про рейтинг. Рекомендації розраховуються незалежно від індивідуальних рейтингів користувачів: або у формі подібності між вимогами рекламодавця та параметрами користувача, або на основі явних правил рекомендацій. Традиційні інтерпретації того, що система рекомендацій орієнтована на аспект фільтрації інформації, в якому користувачі з певними параметрами, які можуть зацікавити певного рекламодавця, фільтруються. На противагу цьому, процес рекомендацій на основі знань є високоінтерактивним. Цей аспект інтерактивності викликав незначний зсув від інтерпретації як системи

фільтрації до ширшої інтерпретації, де рекомендації визначаються як системи, які «персонально орієнтують користувача на цікаві чи корисні об'єкти у великому просторі можливих варіантів або створюють такі об'єкти як вихід» [16]. Рекомендації, які покладаються на джерела знань, які не використовуються підходами, що базуються на основі КФ та базі вмісту, за замовчуванням визначені як рекомендації на основі знань. Два основних типи рекомендаційних систем на основі знань це системи засновані на основі обмежень та системи на основі конкретних випадків. Обидва підходи схожі з точки зору процесу рекомендацій: користувач повинен вказати вимоги, а система намагається визначити рішення. Якщо рішення не вдається знайти, користувач повинен змінити вимоги [17]. Система також може надавати пояснення щодо рекомендованих пунктів. Однак ці рекомендації відрізняються способом використання наданих знань: рекомендації, що базуються на конкретних випадках, зосереджуються на пошуку подібних предметів на основі різних типів заходів подібності, тоді як рекомендації, що базуються на обмеженнях, покладаються на чітко визначений набір правил рекомендацій. У системах на основі обмежень набір рекомендованих елементів визначається, наприклад, пошуком набору елементів, які відповідають рекомендаційним правилам. Системи, засновані на кейсах, з іншого боку, використовують показники подібності для отримання елементів, подібних (у межах попередньо визначеного порогу) за даним вимогам замовника.

1.4 Постановка задачі дослідження

Проблема дослідження полягає у невідповідності між можливостями існуючих методів статичного розподілу рекламного трафіку в мережі Інтернет та потребами рекламодавців у динамічному перерозподілу реклами на основі

комбінування інформації про користувача, про ефективність показаної реклами та про продукти.

Об'єктом дослідження в рамках магістерської атестаційної роботи є процес розподілу та персоналізації рекламного трафіку для рекламодавця.

Предметом дослідження являються методи персоналізації трафіку та формування рекомендацій для рекламодавця.

Метою даної роботи є дослідження гібридних методів фільтрації у процесах розподілу та персоналізації рекламного трафіку для рекламодавця.

Для досягнення мети, необхідно досліджувати наступні питання:

- дослідження особливостей рекомендаційної підсистеми в системі управління інтернет-рекламою;
- аналіз гібридних методів фільтрації які можна застосувати для розподілу та персоналізації рекламного трафіку для рекламодавця;
- аналіз традиційних методів представлення знань.
- удосконалення методу управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань;
- експериментальна перевірка удосконаленого методу для підтримки розподілу та персоналізації рекламного трафіку для рекламодавців в системі показу реклами компаній прокату автомобілів.[1]

2. ДОСЛІДЖЕННЯ ГІБРИДНИХ МЕТОДІВ ФІЛЬТРАЦІЇ

2.1 Особливості методу фільтрації на основі знань

Як вже було показано у попередньому розділі, не всі методи фільтрації можна застосувати до систем розподілення трафіку. Враховуючи той факт що нові рекламодавці приходять не дуже часто, тому система буде працювати не ефективно через низьку кількість доступних рейтингів за умови використання КФ. Також рекламодавці можуть висувати свої вимоги та змінювати їх, що буде призводити до повного перерозподілення трафіку. Формулювання таких вимог не є типовим для чистої рекомендаційної структури на основі КФ та на основі змісту [18]. Тому розглянемо фільтрацію на основі знань.

Системи рекомендацій на основі знань - це специфічний тип системи рекомендацій, який базується на явних знаннях про продукти, уподобання користувачів та критерії рекомендацій (тобто, який товар слід рекомендувати, у якому контекст). Ці системи застосовуються в сценаріях, коли альтернативні підходи, такі як КФ та фільтрування на основі вмісту, неможливо застосувати [19, 20].

Щоб застосовувати фільтрацію на основі знань, спочатку треба розглянути моделі представлення знань. Традиційні моделі представлення знань: семантична мережа або як її ще називають семантична павутина, логічна модель представлення знань, продукційна модель або модель заснована на правилах та модель фреймів.

Семантична мережа - це техніка представлення знань, яка використовує позначення графа для опису структури знань [27]. Згідно з цим поняттям, екземпляри та властивості відповідають вузлам графіків, тоді як відносини між ними (тобто таксономія) задаються через спрямовані дуги. Формально семантична мережа визначається як кортеж $\langle I, R_1, R_2, \dots, R_n, M \rangle$, де I виступає як інформаційні сутності, $R_1, R_2, \dots, R_n$ 1 позначає типи відносин між

інформаційними об'єктами, і М відповідає карті, яка встановлює відносини між суб'єктами інформації. Різноманітність характеру елементів у кортежі та операціях підтримка дозволяє виділити наступні найпоширеніші види семантичних мереж: визначальні, твердження, імплікаційні, виконувані, навчальні та гібридні. Незважаючи на значні відмінності в інформаційних сутностях та застосуванні мереж, усі вони використовують декларативно-графічний підхід до подання знань. Існують різні види смислових мереж, тому давайте докладніше розглянемо їх [23].

Визначальні мережі використовуються для опису понять, примірників та зв'язків між ними. Вони сильно використовують узагальнення типів "is-a", "has-a", "instance-of"(екземпляр), щоб сформувати таксономію знань, яка є важливою для зовнішнього механізму міркування.

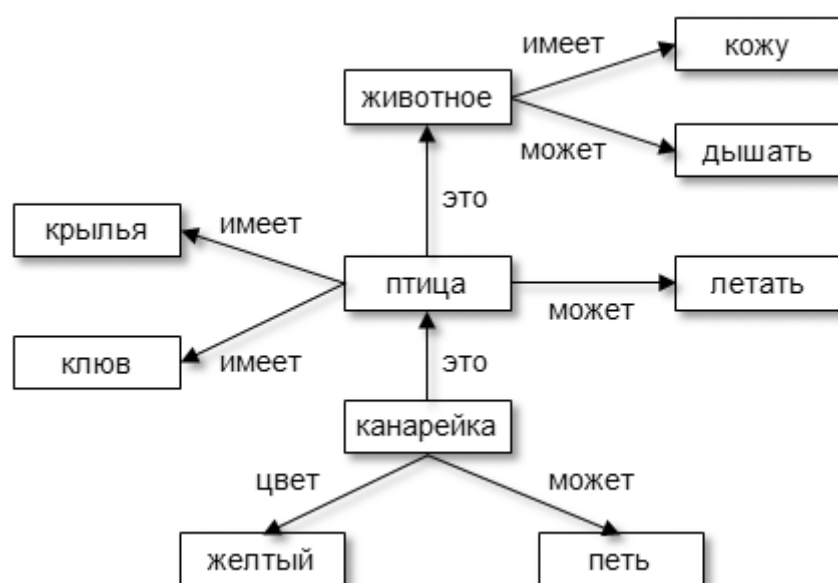


Рисунок 2.1 – пример визначальні семантичної мережі

Ассертаційна мережа представляє графічно логічні пропозиції, тому її елементи виконують додаткові смислові ролі, що відповідають логічним операторам, таким безстатевим кількісним показником, кон'юнкцією, диз'юнкцією та ін.

Імплікаційні мережі - це особливий випадок апропозиційної семантичної мережі, яка використовує імплікацію як основний зв'язок для з'єднання вузлів; інші відносини можуть бути вкладені всередині пропозиційних вузлів, але вони ігноруються процедурами висновку.

Виконавчі мережі містять механізми, які можуть викликати певну зміну самої мережі. Найпоширенішими виконавчими механізмами є передача повідомлень, додана процедура та перетворення графіків.

Навчальні мережі налаштовують своє внутрішнє представлення у відповідь на нову інформацію, щоб ефективно працювати в їхньому середовищі. Навчальна мережа може бути скоригована трьома способами: рухом, зміною ваги та перебудовою.

Нарешті, гібридні мережі - це ті, що одночасно об'єднують кілька типів семантичних мереж, щоб повторно використовувати переваги та усунути недоліки вищезгаданих мереж.

Серед головних переваг смислових мереж — легкість уявлення про людину, що забезпечується графічним зображенням. Спочатку графічне позначення показує прямі зв'язки між спорідненими поняттями, тоді як лінійні позначення повинні базуватись на часто зустрічаються змінних чи імен, щоб відображати однакові зв'язки. По-друге, графічні позначення часто мають евристичну цінність, допомагаючи людським читачам виявити закономірності, які було б важко або неможливо побачити в лінійній формі; це могло пояснитись кластеризованою появою пов'язаних знань. Говорячи про недоліки семантичних мереж, важливо виділити відсутність стандартизації значень вузлів і дуг, що вводить складність обслуговування. Крім цього, деякі типи семантичних мереж страждають від конкретних питань, таких як множинні помилки успадкування, відносини один до багатьох, відсутність формальної семантики, змішування різних рівнів абстракції. Тим не менш, "природний" підхід до представлення знань робить семантичні мережі широко використовуваною технікою в інженерії систем знань.

Логічна модель представлення знань (ЛМП) - це сукупність формальних систем, які складаються з опису формальних мов опису тверджень та дедуктивної системи для міркування. Існує безліч варіацій ЛМП, включаючи інтуїціоністичні, багатосортивні, модальні тощо; тут ми обговоримо класичну логіку першого порядку (ЛПП) [25]. Логіка першого порядку є стандартом для формалізації математики в аксіоми, а також відіграє вирішальну роль у представленні знань практично будь-якого виду. Це постало як наступник логіки пропозицій, яка не володіє достатньою виразністю для представлення реального знання. Основна відмінність логіки першого порядку від логіки пропозицій полягає в онтологічному зобов'язанні: ЛПП, як і природна мова, передбачає, що світ містить об'єкти, відносини та функції, тоді як пропозиційна логіка діє лише з фактами. ЛПП корисний для показу логічних зв'язків та їх міркувань; більше того, це дозволяє логічне твердження поділити на слова. Як приклад, розглянемо два логічні твердження "Церква — математик" та "Тьюрінг — математик". У пропозиційній логіці ці твердження розглядаються як не пов'язані між собою, тоді як в логіці передбачуваності нам дозволяється ввести предикат "Математик (x)", який робить взаємозв'язок між твердженнями явним.

Формальна мова ЛПП складається з синтаксису та семантики. Синтаксис визначає, які колекції символів із алфавіту є валідними виразами, тоді як семантика визначає значення цих виразів. Алфавіт - це набір символів, з яких можуть утворюватися рядки формальної мови. У логіці першого порядку символи алфавіту поділяються на дві групи, логічні та нелогічні. Різниця полягає в тому, що логічні символи завжди мають однакове значення, тоді як значення нелогічних символів залежить від інтерпретації. Логічні символи у ЛПП зазвичай представлені кванторами ($\forall, \exists$), логічними сполучниками ($\neg, \leftrightarrow, \rightarrow, \vee, \wedge$), пунктуацією (дужки), змінними (нескінченна множина), рівністю ($=$). складаються з предикатів і функцій з валентністю (числом аргументів) більшою або рівною до нуля.

Предикат валентності 0 відповідає пропозиційній змінній, а функція валентності 0 називається константою. Правила утворення визначають два типи валідних синтаксичних виразів у ЛПП, терміни та формули. Терміни представлені змінними та функціями, тоді як формули складаються з предикатів (приймають аргументи), символу рівності (діє на терміни), логічних сполучників та кількісних показників. Формула може бути зв'язаною або вільною (або обома) залежно від того, чи лежить вона в межах кількісного показника чи ні. Формула без вільних змінних називається реченням першого порядку. Речення першого порядку - це формули, які матимуть чітко визначені значення істини в інтерпретації, тому вони використовуються для представлення знань; решта - це лише підтримка синтаксичної техніки. Семантика першого порядку забезпечується інтерпретацією, яка визначає область дискурсу та позначення нелогічних символів.

Міркування у ЛПП виконується в рамках дедуктивної системи, яка використовується, щоб показати, що одна формула є логічним наслідком іншої формули. Більшість дедуктивних систем складаються з логічних аксіом (можливо, жодних) та правил умови вводу, які можуть бути використані для отримання теорем системи. Ці системи мають важливу особливість - вони суто синтаксичні, а значить, виводи можна перевірити, не враховуючи жодної інтерпретації. У результаті кожне словотворення є синтаксичним наслідком усіх виразів, які передують йому.

Існує багато дедуктивних систем для логіки першого порядку, включаючи систему в стилі Гільберта, природну дедукцію, послідовне обчислення, табличний метод, роздільну здатність. Дедуктивна система називається звуком, якщо вона виробляє лише логічно обгрунтовані формули, ті, що гарантовано є логічним наслідком. Дедуктивна система називається повноцінною, якщо гарантовано створювати всі логічно дійсні формули. Звучність і незавершеність відносно семантики, придатної для мови системи дедукції. На жаль, неавтоматизований процес міркування для ЛПП може бути як надійним, так і

створювати проблеми, оскільки проблема визначення того, яке одне речення є логічним наслідком інших у ЛПП, загалом є нерозв'язним. Логіка першого порядку є хорошою відправною точкою для представлення знань та міркувань, оскільки вона має багато металогічних властивостей, яких сильніші логіки не мають. Відповідно до теореми Ліндстрьома, ЛПП - це максимальна логіка, яка задовольняє спадну властивість Левенгайма-Сколема та лічильну властивість компактності. Проте він також зазнає значних обмежень. Спочатку на нього впливає паралеокс Сколема, з чого випливає, що нескінченні структури (тобто реальна лінія, натуральні числа) не можуть бути категорично аксіоматизовані в логіці першого порядку. Це питання вирішується в більш сильних логіках, таких як логіка другого порядку. Натомість, логіки першого порядку недостатньо для вираження деяких типових особливостей природної мови, подібної кількісної оцінки властивостей, що вимагає наявності кількісного показника над предикатами. Нарешті, обчислювальна складність ЛПП є одним з основних факторів, які призводять до розгляду різних інших варіантів представлення знань та міркувань.

Фрейм — це об'єктно-орієнтована форма подання знань, яка робить його онтологічно прихильним до об'єктів властивостей, цінностей і відносинам між ними. Як результат, його часто називають представленням "об'єкт-властивість-цінність". Об'єкти - це природний спосіб організувати знання про фізичні об'єкти та ситуації. Важливе значення полягає в тому, що об'єкт дає можливість групувати процедури для визначення власних властивостей, а також властивостей інших об'єктів. У ЛПП твердження про світ розкидані навколо, тому важко контролювати послідовність та всебічність прийнятих знань. Об'єктно-орієнтоване представлення можна досягти шляхом об'єднання багатьох властивостей одного і того ж типу в одну структуру.

У системах на основі фрейму, об'єкт, що містить властивості, заповнені значеннями, відповідає "фрейму", який містить "слоти", заповнені "наповнювачами" тому фрейм можна визначити як структурне та реляційне

представлення об'єкта. Кожен фрейм має назву і складається з набору слотів із пов'язаними наповнювачами. Наповнювач може приймати значення примітивних типів (число, рядок, дата тощо), складних типів (набір, список) або зберігати як посилання на інший каркас. Виходячи з характеру об'єкта опису, розрізняють два типи фреймів - індивідуальні та загальні. Індивідуальний фрейм являє собою єдиний об'єкт, який є екземпляром деякої категорії, тоді як загальний фрейм використовується для опису абстрактної категорії. Отже, окремі фрейми за замовчуванням мають спеціальний слот під назвою "INSTANCE-OF", де наповнювач - назва (або вказівник на) аналогічного загального кадру. У той же час, загальні фрейми можуть мати спеціальний слот під назвою "IS-A", який має вказівник на більш загальний фрейм. Наявність цих спеціальних прорізів має вирішальне значення, оскільки вони включають механізм збереження.

Продукційна модель, або модель заснована на правилах - також відомі як "IF-THEN" правила, є однією з найбільш широко використовуваних форм представлення знань, особливо в експертних системах (ЕС) та автоматизованій системі підтримки прийняття рішень (АСПР). Вони характеризуються високою модульністю, що приносить гнучкість до процесу інсталяції та обслуговування. Кожне правило визначає відносно невеликі і незалежні знання, крім того, його можна легко додавати та видаляти незалежно від інших правил. Простота синтаксису та природний спосіб вираження знань призводять до простої інтерпретації та простоти розуміння. Типове виробниче правило складається з попереднього та наступного. Антецедент відповідає приміщенню (умова, IF-стаття) і, як наслідок, означає дію (висновок, ТО-пункт). Щоб включити механізм розподілу, правила зазвичай групуються в так звану виробничу систему, яка має свої структурні та функціональні особливості.

Продукційна система складається з бази знань і механізму висновку, внаслідок чого утворюються два типи правил - декларативний та процедурний [26]. Декларативні правила містять усі факти та стосунки щодо проблеми, а

процедурні правила - як вирішити проблему, враховуючи, що відомі певні факти. Вони також відрізняються місцем зберігання: перші зберігаються у базі знань, а другі - включений у механізм висновку. База знань виробничої системи має свою специфіку і поділяється на робочу пам'ять і базу правил. Робоча пам'ять складається з сукупності фактів, які в основному є твердженнями про світ, але можуть також представляти різні структури даних. Синтаксис і представлення фактів не є строго формалізованими, тому можуть відрізнятися в різних системах. База правил формується набором правил, які належать до особливого типу правил " IF-THEN", де попередній податок представлений поєднанням умов, а наступний являє собою послідовність дій. Структура випадкового запису в базі правил визначається за допомогою формули $p_1 \wedge p_2 \wedge \dots \wedge p_n \Rightarrow a_1, a_2, \dots a_k$. Старовинні члени можуть приймати форму заперечених або негативних предикатних висловлювань та умов на окремий об'єкт або кілька об'єктів. Наступні, як правило, полягають у наступних діях: "ADD" факт робочої пам'яті, "REMOVE" факт з робочої пам'яті, "MODIFY" поле атрибутів, "QUERY" користувача для вводу;

$$A(x) \wedge B(x) \wedge C(y) \Rightarrow \text{REMOVE} : D(x), \text{ADD} : B(y).$$

Механізм виводу у продукційних системах, як правило, реалізується за допомогою методу прямого ланцюга. Це починається з наявних у робочій пам'яті фактів і використовує набір активних правил бази знань для отримання більшої кількості фактів, поки мета не буде досягнута. Правило, як кажуть, є активним, якщо його попередник задоволений; це є необхідною умовою для виконання наступних. Це звичайна ситуація, коли багато правил активні одночасно час, таким чином вводячи необхідність вказувати порядок їх виконання. Порядок виконання важливий, оскільки дії ADD та REMOVE можуть змінити набір активних правил, так що наступне активне правило в черзі може бути деактивоване попереднім. Стратегія вибору правила для виконання серед можливих кандидатів називається вирішенням конфлікту. Нижче перераховані деякі найпопулярніші стратегії вирішення конфліктів:

- не дублювати — не виконайте одне і те ж правило двічі;
- недавність — правила, що стосуються фактів, щойно додані до робочої пам'яті, мають перевагу;
- специфічність — правила, більш конкретні, віддають перевагу перед більш загальними;
- пріоритетні рівні — призначають рівні пріоритету для правила; правило з більш високим пріоритетом має перевагу.

Ще одне питання, що з'являється в аргументаціях продукційних систем, пов'язане з процесом уніфікації[28, 29]. Деякі правила можуть вимагати великої кількості уніфікацій, особливо якщо вони містять змінні тестування всіх випадків, коли правила активні. Крім того, умови багатьох правил можуть перетинатися, що призводить до багаторазового повторення одних і тих же об'єднань. Для вирішення питань неефективності та вирішення конфліктів було введено алгоритм Рете. Основні принципи його функціонування полягають у наступному: правила складаються в мережу, яка об'єднує умови декількох правил взагалі, уникаючи, таким чином, дублювання; розповсюджуються лише дійсні об'єднання; У висновку, що правила виробництва забезпечують велику модульність і гнучкість до завдання представлення знань. Вони виразні, прості для розуміння і підтримуються потужним висновком, що доступний у продукційних системах. Крім того, їх корисність вже була доведена успіхом широкої комерційної експертної системи. Говорячи про недоліки, важливо зазначити, що не всі знання можуть бути виражені у формі правил. Продукційні правила є поганими у представленні структурованих описових знань, і важко дотримуватися ієрархії. Великі системи, як правило, складаються з тисяч правил, що ускладнює консолідацію знань і значно знижує продуктивність роботи системи.

Саме продукційна модель представлення знань була використана в системі розподілення трафіку. Для кожного типу реклами, кожного користувача та кожного рекламодавця існують правила в базі знань, але через велику

кількість даних такий спосіб представлення знань стає все менш ефективним для використання в цій системі.

Стає потреба використовувати гібридні методи фільтрації, де будуть об'єднані фільтрація на основі знань та ще один метод який би автоматизував зміни в деяких правилах базуючись на здобутих даних за певний період часу.

2.2 Дослідження можливостей використання обчислювального інтелекту для формування правил в системах розподілення трафіку.

В системах розподілення трафіку, де використовується фільтрація на основі знань та використовується продукційна модель збереження знань, з масштабуванням починають виникати проблеми з консолідацією знань та редагуванням правил тому для автоматизації зміни правил та аналізу даних використовують обчислювальний інтелект.

Одним з відгалужень обчислювальний інтелекту є задача прогнозування. В системі розподілення трафіку з представленням знань у вигляді правил прогнозування для редагування деяких правил може суттєво повисити ефективність роботи системи. Найпростішим, в нашому випадку, рішенням для прогнозування буде лінійна регресія [30].

$$y = \beta_0 + \beta_1 X + \epsilon \quad (2.1)$$

формула 2.1 – формула простої лінійної регресії

Проста модель лінійної регресії - це модель з одиничним регресором x , яка має зв'язок з відповіддю y , яка є прямою лінією. Завдання полягає у тому щоб розв'язати рівняння цієї прямої лінії де β_0 та β_1 прості коефіцієнти лінійної регресії, де перехоплення(перетин з віссю) це β_0 і нахил β_1 - невідомі постійні, а

ε - складова випадкової помилки. Припускається, що помилки мають середню нульову та невідому дисперсію σ_2 . Крім того, ми зазвичай припускаємо, що помилки є некорельованими. Це означає, що значення однієї помилки не залежить від значення будь-якої іншої помилки [31].

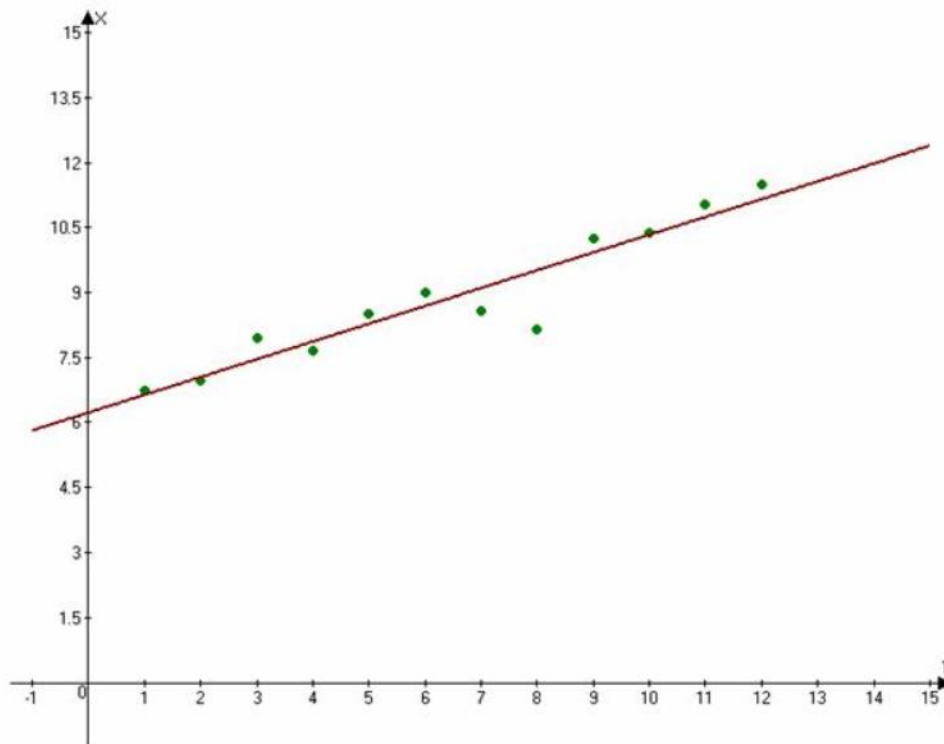


Рисунок 2.2 – приклад графіку лінії простої лінійної регресії

Зручно розглядати регресор x як контрольований даними і вимірювати з незначною помилкою, тоді як відповідь y - випадкова величина. Тобто існує розподіл ймовірностей для y при кожному можливому значенні для x . Середнє значення цього розподілу:

$$E(y|x) = \beta_0 + \beta_1 x \quad (2.2a)$$

Формула 2.2a - середнє значення розподілу ймовірностей

Та формула дисперсії:

$$\text{Var}(y|x) = \text{Var}(\beta_0 + \beta_1 x + \varepsilon) = \sigma_2 \quad (2.26)$$

Формула 2.2б – формула дисперсії

Таким чином, середнє значення y є лінійною функцією x , хоча дисперсія y не залежить від значення x . Крім того, оскільки помилки некорельовані, відповіді також некорельовані.

Параметри β_0 і β_1 зазвичай називають коефіцієнтами регресії.

Ці коефіцієнти мають просту і часто корисну інтерпретацію. Нахил β_1 — це зміна середнього значення розподілу y , виробленого зміною одиниці в x . Якщо діапазон даних на x включає $x = 0$, то перехоплення β_0 - це середнє значення розподілу відповіді y , коли $x = 0$. Якщо діапазон x не включає нуль, то β_0 не має практичної інтерпретації.

Параметри β_0 і β_1 невідомі і повинні бути оцінені, використовуючи вибіркові дані. Припустимо, у нас є n пар даних, скажімо $(y_0, x_0), (y_1, x_1), \dots, (y_n, x_n)$. Ці дані можуть бути результатом контрольованого експерименту, призначеного спеціально для збору даних, спостережного дослідження або існуючих історичних записів (ретроспективне дослідження).

Оцінка параметрів β_0 і β_1 може бути проведена методом найменших квадратів. Де коефіцієнти повинні бути підібрані так, що сума квадратів різниць між спостереженнями y_i та прямою є мінімальною. Базуючись на формулі (2.1) ми можемо написати:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2.3)$$

Формула 2.3 – визначення y_i для простої моделі лінійної регресії

Якщо формула (2.1) може розглядатися як загальна модель регресії, тоді як формула (2.3) є зразковою регресійною моделлю, записаною через n пар даних (y_i, x_i) де $(i = 1, 2, \dots, n)$. Таким чином, функцією найменших квадратів є:

$$S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 + \beta_1 x_i)^2 \quad (2.4)$$

Формула 2.4 – рівняння що використовується в методі найменших квадратів.

Мінімум функції визначають через обчислення часткової похідної від $S(\beta_0, \beta_1)$ що до коефіцієнтів β_0 та β_1 і прирівнювання їх до нуля.

2.3 Удосконалення методу управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань

Результатом роботи є удосконалений гібридний метод фільтрації на основі знань, який доповнений методом автоматизованої зміни правил ваги в рекламних кампаніях на підставі прогнозувань, що були здобуті за допомогою використання простої моделі лінійної регресії в методі розподілу рекламного трафіку системи управління інтернет-рекламою. (рисунок 2.3).

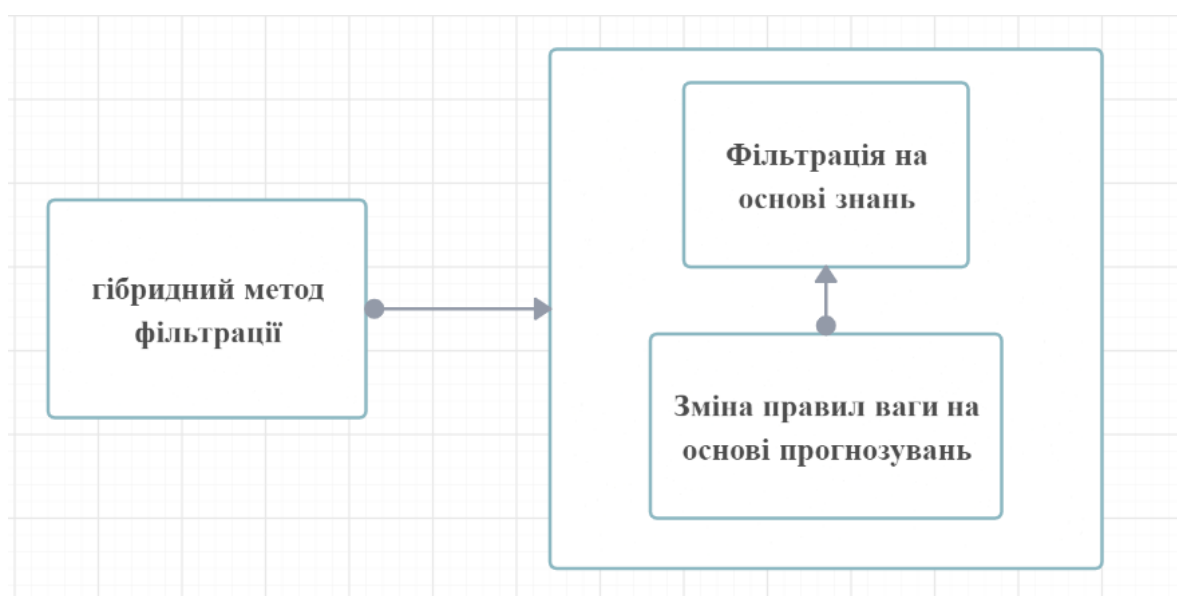


Рисунок 2.3 – Структура гібридного методу

Запропонований метод використовує наступні вхідні дані.

Використовуються дані запиту $P = \{p_i\}$ в такому вигляді:

$$p_k = \{l, u_j, pd, dd\}, \quad (2.5)$$

де l - вибрана споживачем локація; u_j - дані споживача; pd , dd - дата початку та закінчення оренди відповідно.

В свою чергу, кожен споживач $u_j \in U$ в системі має такі характеристики:

$$u_j = \{id, d, p, g, a\}, \quad (2.6)$$

де id - унікальний ідентифікатор споживача, d - тип пристрою користувача, p - платформа пристрою, g - геолокація користувача, a - ідентифікатор рекламодавця.

Шаблон відображення характеризується ідентифікаційним кодом, даними споживача та вагою.

$$t_i = \{it, u_j, tw\}, \quad (2.7)$$

де it - ідентифікаційний код, u_j - дані користувача, tw - вага шаблону встановлена експертом.

Рекламна кампанія характеризується кодом, даними споживача, шаблоном відображення та вагою та використовуються дані кампаній $C = \{c_i\}$ у такому вигляді:

$$c_n = \{ic, u_j, t_i, cw\}, \quad (2.8)$$

де ic - ідентифікаційний код, u_j - дані користувача, t_i - шаблон відображення реклами, cw - вага кампанії встановлена експертом чи системою.

Запропонований метод містить у собі такі етапи.

Етап 1. Відбір шаблону для відображення t за умовами користувача та вагою.

Етап 1.1 попередній відбір підмножини шаблонів T^* .

$$T^* = \{t_k : \forall u_j, d_j \in D_k, p_j \in P_k, a_j \in A_k\}, \quad (2.9)$$

де d_j – тип пристрою користувача, D_k – підмножина типів пристроїв що відносяться до шаблону, p_j – платформа пристрою користувача, P_k – підмножина платформ що відносяться до шаблону, a_j – ідентифікатор рекламодавця від якого перейшов користувач, A_k – підмножина дозволених ідентифікаторів рекламодавців для цього шаблону.

Етап 1.2 Вибір шаблону з підмножини T^* з найбільшою вагою tw_k .

Етап 2. Вибір підмножини рекламних кампаній $C^* \subseteq C$ за ознаками користувача та шаблону:

$$C^* = \{c_l : \forall u_j t_j = t_l^*, d_j \in D_l, p_j \in P_l, a_j \in A_l\}, \quad (2.10)$$

де d_j – тип пристрою користувача, D_l – підмножина типів пристроїв що відносяться до кампанії, p_j – платформа пристрою користувача, P_l – підмножина платформ що відносяться до кампанії, a_j – ідентифікатор рекламодавця від якого перейшов користувач, A_l – підмножина дозволених ідентифікаторів рекламодавців для цієї кампанії, t_l^* – ідентифікатор вибраного на етапі 1 шаблону.

Етап 3. Корегування ваг кампаній з підмножин C^* .

Це новий етап, який корегує ваги кампаній на підставі прогнозів які були виявлені при побудові лінійної регресії по даним статистики.

Етап 3.1. Для кожної рекламної кампанії підмножини C^* виберемо підмножину $S^* \subseteq S$:

$$S^* = \{s_m : \forall c_j ic_m = C^*, d_m \geq d_{now} - \varepsilon\}, \quad (2.11)$$

де ic_m – це ідентифікатор кампанії для якої призводиться вибірка даних, а ε – це кількість днів за яку буде братися статистика.

Етап 3.2. Для використання прогнозів на основі отриманих даних раз в день будується нова модель лінійної регресії для кожної рекламної кампанії на підставі підмножин S^* . Знаходимо β_0 та β_1 та враховуючи обмеження ваги для кожної кампанії виставляємо ваги використовуючи функцію:

$$\begin{aligned} r &= \beta_0 + \beta_1 cw \\ \min_{cw} &< cw < \max_{cw}, \end{aligned} \quad (2.12)$$

де β_0 та β_1 - знайдені коефіцієнти, cw – це вага у вказаному рекламодавцем діапазоні, r – середній прибуток на одну сесію.

Через те що функція лінійна, ми можемо визначити на якому з кінців відрізка буде максимальний результат та виставити ваги у те положення.

Етап 4. Вибір рекламної кампанії за вагою.

Тепер вибираємо рекламну кампанію серед підмножини S^* з найбільшою вагою.

Отримана рекомендована рекламна кампанія буде мати оптимальні для прибутку ваги та надалі використана в генерації продуктів та відображені їх на вибраному шаблоні.

Відмінність нового методу полягає в тому, що частина правил на основі знань буде автоматично редагуватися за допомогою метода зміни правил ваги на основі прогнозувань, що збільшить ефективність розподілу рекламного трафіку за рахунок корегування ваги рекламних кампаній.

Перевага удосконаленого методу в тому, що експерт не завжди може правильно перерозподілити ваги компаній, тому що сума вагів компаній є відносною величиною до загальної кількості відвідувачів та вага кожної кампанії не є персоналізована для кожного шаблону відображення даних, тому метод що зможе автоматично спрогнозувати вагу, при якій прибуток на кожну сесію буде максимальним для кожної рекламної кампанії значно полегшить управління розподілення рекламного трафіку.

Недолік удосконаленого методу полягає у високій обчислювальній складності при великій кількості рекламних кампаній.

3 ДОСЛІДЖЕННЯ ОТРИМАНИХ НАУКОВИХ РЕЗУЛЬТАТІВ

3.1 Дослідження методу фільтрації на основі знань

Не всі методи фільтрації можна застосувати в системі розподілу трафіку через те що нові рекламодавці приходять не дуже часто, тому система буде працювати не ефективно через низьку кількість доступних рейтингів за умови використання КФ. Також рекламодавці можуть висувати свої вимоги та змінювати їх, що буде призводити до повного перерозподілення трафіку. Тому було прийняте рішення розглянути фільтрацію на основі знань.

Системи рекомендацій на основі знань, як вже було сказано раніше, це специфічний тип системи рекомендацій, який базується на явних знаннях про продукти, уподобання користувачів та критерії рекомендацій (тобто, який рекламний продукт слід рекомендувати, у якому контекст). Ці системи застосовуються в сценаріях, коли альтернативні підходи, такі як КФ та фільтрування на основі вмісту, неможливо застосувати.

В системі застосовується продукційна модель представлення знань. Тому для кожного типу реклами, кожного користувача та кожного рекламодавця існують правила в базі знань. Розглянемо самі вхідні дані більш детально.

Дані запиту або пошук – це структуровані дані про користувача, що зайшов на сайт. Пошук містить дані про геолокацію де шукає користувач, та дати взяття на прокат та повернення авто.

Дані користувача – це дані, що аналізуються, коли користувач заходить на сайт, та зберігаються у системі. Ці дані надалі є основними при фільтрації трафіку, тому що головну роль в рекомендації реклами відіграє персоналізація, але використання цих даних для розподілення показаної реклами користувачу є не менш важливим, тому ці дані будуть використані і для розподілення трафіку серед рекламних кампаній.

Шаблон - це тип відображення рекламних продуктів. Шаблон може бути вибраний в залежності від, правил, які ми можемо встановити для кожного з шаблонів на основі даних користувача. До кожної шаблону може бути прив'язана необмежена кількість правил. Система буде слідкувати щоб правила не повторювались.

Рекламна кампанія - це дуже гнучкий інструмент що дає можливість розподілити трафік між шаблонами, типами рекламного трафіку, партнерами та врахувати максимальну кількість факторів використовуючи налаштування кампанії та правила вибору рекламної кампанії серед інших кампаній.

Рекламна кампанія має такі основні параметри:

- Унікальний ідентифікатор кампанії;
- Назва кампанії;
- Ідентифікатор партнера;
- Шаблон;
- Вага;

Рекламна кампанія має ще й інші параметри, але вони не відіграють важливої ролі в нашому методі.

До кожної рекламної кампанії прив'язані правила. Їх може бути необмежена кількість, але система слідкує щоб правила не повторювались. Все залежить від розмірів довідкові таблиці, які містять типи девайсів та платформи. Ці таблиці постійно розширюються.

Як ми бачимо правила активно застосовуються в системі та їх кількість невпинно росте, саме через велику кількість даних такий спосіб представлення знань стає все менш ефективним для використання в цій системі.

Стає потреба використовувати гібридні методи фільтрації, де будуть об'єднані фільтрація на основі знань та ще один метод який би автоматизував зміни в деяких правилах базуючись на здобутих даних за певний період часу.

Удосконалений гібридний метод фільтрації базується на фільтрації на основі знань, яка була доповнена методом автоматизованої зміни правил ваги в рекламних кампаніях на підставі прогнозувань.

Перевага удосконаленого методу полягає в автоматизації аналізу прибутків рекламних кампаній та встановлення вагів відповідно до результатів цього аналізу. Це дозволить впоратись з проблемою оцінки ефективності великої кількості кампаній та перерозподілення трафіку між ними.

3.2 Технологія використання удосконаленого методу розподілення трафіку в системах управління інтернет-рекламою

Технологія використання удосконаленого методу розподілення трафіку починається з аналізу даних запиту та користувача потім йде створення сесії та прив'язки її до користувача, якщо він зареєстрований в системі. Далі на основі даних користувача та партнера йде аналіз та перерозподіл трафіку. Тепер, ми можемо виділити основні функції системи:

- аналіз даних запиту та користувача. Обираємо та аналізуємо дані користувача. Після цього створюємо сесію та зберігаємо туди весь набір даних;
- вибір шаблону базується на даних користувача та правилах, що можна задати в адмін панелі;
- вибір підмножини рекламної кампанії та корегування вагів на підставі
- вибір рекламної кампанії базується на правилах, шаблону, даних пошуку користувача та вагових коефіцієнтах;

Для більш чіткого відображення структури використаємо діаграму IDEF0.

На контекстній діаграмі IDEF0 (рисунок 2.1) можемо побачити загальний блок, який визначає головну мету розробки – «Система розподілу рекламного трафіку». На вихід задачі надходять відомості про показану рекламу та про прибутки з ефективністю.

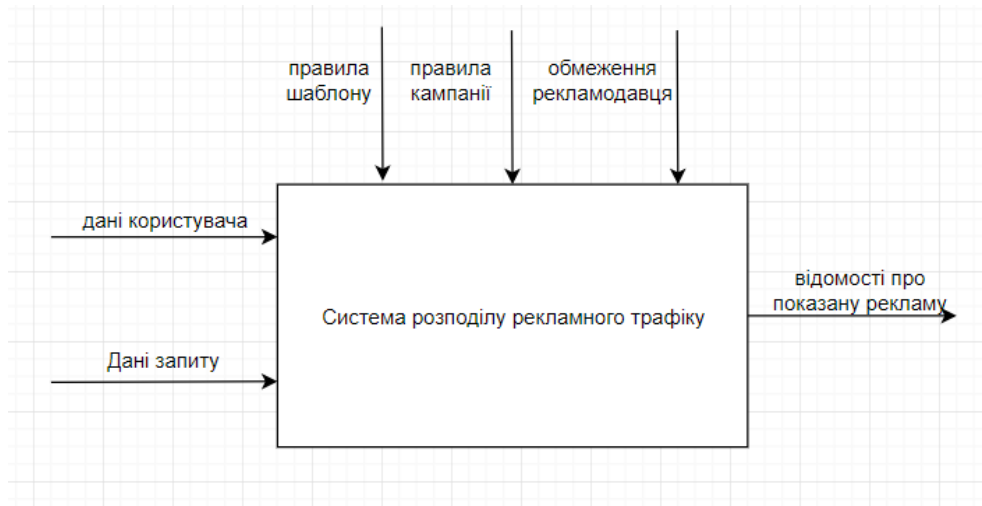


Рисунок 2.1 – Контекстна діаграма IDEF0 задачі «Система розподілу рекламного трафіку»

Діаграма декомпозиції першого рівня IDEF0 (рисунок 2.2) буде містити 4 блоки, які включають в себе усі описані функції системи.

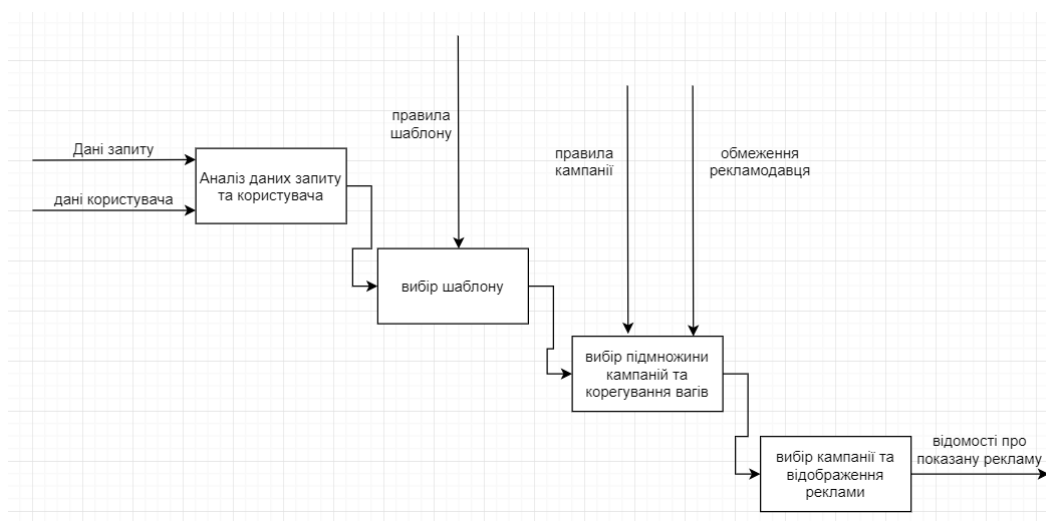


Рисунок 2.2 – Діаграма декомпозиції першого рівня IDEF0 задачі «Система розподілу рекламного трафіку»

4. ПРАКТИЧНЕ ЗАСТОСУВАННЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ

4.1 Програмна реалізація методу розподілу рекламного трафіку в системах управління інтернет-рекламою

Розглянемо приклад реалізації гібридного методу розподілу трафіку який використовується в системі управління інтернет рекламою. Реклама в конкретному випадку буде надаватися компанія по прокату авто. За для наочності використання методу розглянемо наступні вхідні дані, які представлені в таблиці 4.1

Таблиця 4.1 – Вихідні дані пошуку

Поле	Значення
Локація	Харків
Початок оренди	22.12.2019
Кінець оренди	25.12.2019

Для комфортного користування системою був створений графічний інтерфейс для вводу даних.

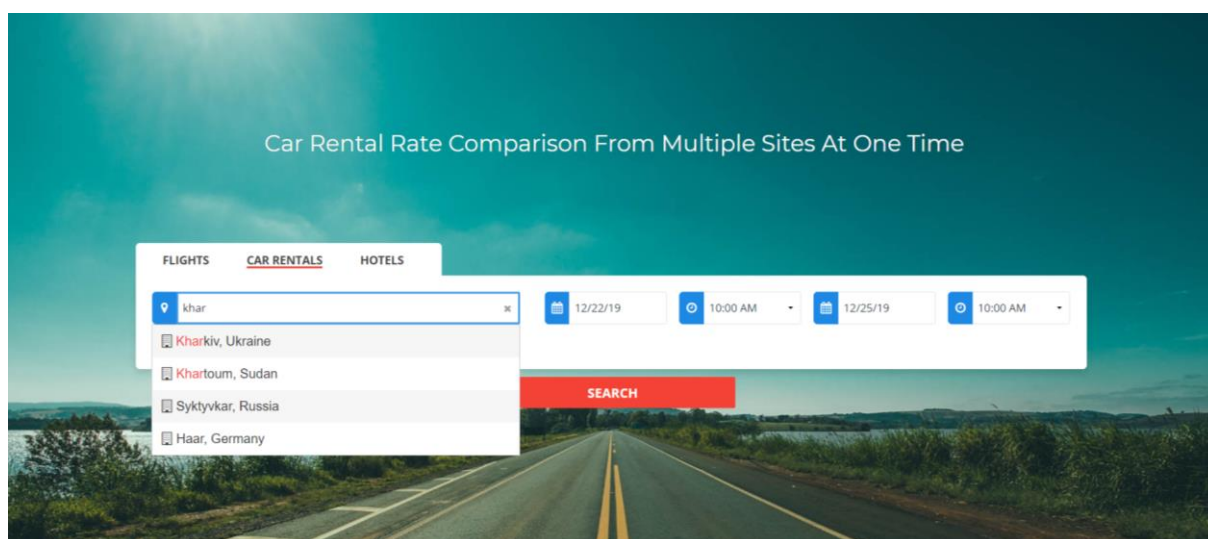


Рисунок 4.1 – створення пошуку

Після створення пошуку система починає аналізувати вхідні дані користувача, такі як тип девайсу, браузер та звідки прийшов трафік.

```
$helper := new CommonHelper();

$session['country'] := $helper->getCountryByIp($session['ip_address']);
$session['flag_https'] := APP_SECURE == true ? 1 : 0;
$browser_type := $request->getHeader('User-Agent');
$session['browser_type'] := $browser_type[0];
$mobile_detect := new Mobile_Detect();
if ($mobile_detect->isMobile() || $mobile_detect->version('Android') || $mobile_detect->is('iphone')) {
    if ($mobile_detect->isTablet()) {
        $session['device_type'] := 2;
    } else {
        $session['device_type'] := 1;
    }
} else {
    $session['device_type'] := 0;
}
```

Рисунок 4.1 – визначення типу девайсу

```
$platform := (empty($user_agent['platform']) ? 'undefined' : $user_agent['platform']);
if (strlen($platform) > 50) {
    $platform := substr($platform, 0, 50);
}
$session['platform'] := $platform;

$platform_id := CacheManager::cacheRead("platform_id_{$session['platform']}");
if ($platform_id === false) {
    $platform_id := $this->getModel('Refplatform')->getIdByName($session['platform']);
    if (empty($platform_id)) {
        $platform_id := $this->getModel('Refplatform')->save(['platform' => $session['platform']]);
    }
    CacheManager::cacheWrite("platform_id_{$session['platform']}", $platform_id);
}
$session['platform_id'] := (empty($platform_id) ? 0 : $platform_id);
```

Рисунок 4.2 – визначення платформи

```

$advertiser_code := '';

if ($this->request->getQueryParam('cid', .null)) {
    » $advertiser_code := $this->request->getQueryParam('cid', .null);
}

if (!empty($advertiser_code)) {
    » $advertiser := CacheManager::cacheRead("advertiser_{$advertiser_code}");
    » if ($advertiser === false) {
    »     » $advertiser := $this->getModel('Avrecords')->getAdvertiserByCID($advertiser_code);
    »     » CacheManager::cacheWrite("advertiser_{$advertiser_code}", $advertiser);
    » }
}

if (!empty($advertiser)) {
    » $session['advertiser_id'] := $advertiser['advertiser_id'];
    » $session['advertiser_code'] := $advertiser_code;
    » $session['advertiser_keyword'] := $this->request->getQueryParam('kw', .null);
} else {
    » //set organic
    » $session['advertiser_id'] := 2;
    » $session['advertiser_code'] := '';
}

```

Рисунок 4.3 – визначення походження трафіку

Після цього, на підставі зібраних даних та правил шаблону буде вибраний шаблон відображення. Для управління правилами шаблону був створений графічний інтерфейс.

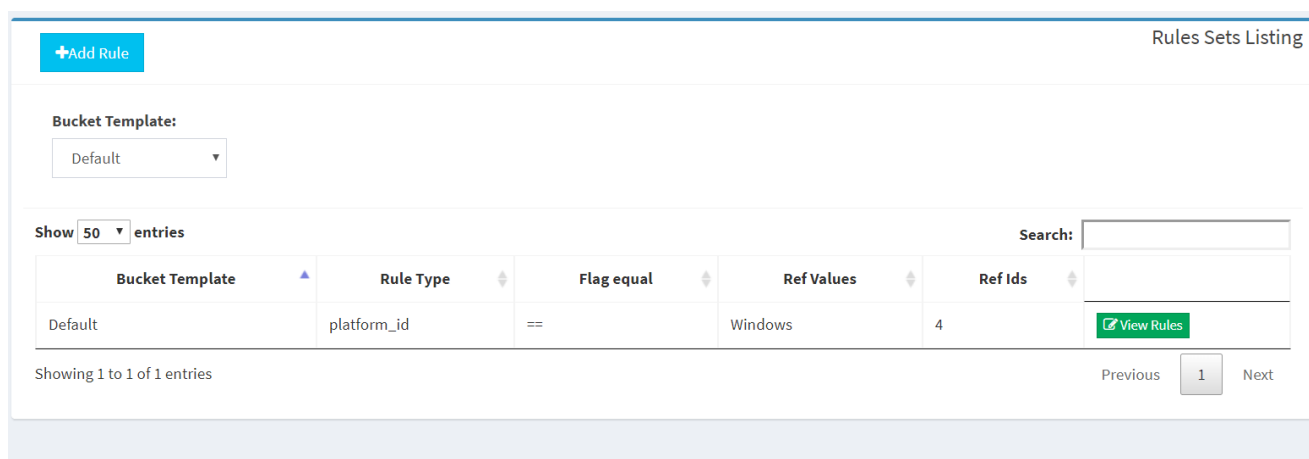


Рисунок 4.4 – графічне відображення правил шаблону

Також був створений інтерфейс для редагування та створення нових правил.

Add Rule

Bucket template:
Default

Rule type:
device_type_name

Ref id:
desktop

Flag Equal:
==

Status:
active

Add **Cancel**

Рисунок 4.5 – приклад створення правила для шаблону

Додамо до шаблону нове правило, що буде перенаправляти на цей шаблон користувачів тільки стаціонарних комп'ютерів та ноутбуків. Після додавання ми зможемо побачити нове правило у списку правил для цього шаблону.

Якщо адміністратор захоче додати ще одне таке ж правило система не дозволить це зробити та видасть попередження. Цей захід спрощує роботу з правилами шаблонів коли кількість правил стає значною.

Rules Sets Listing

Bucket Template:
Default

Show 50 entries

Search:

Bucket Template	Rule Type	Flag equal	Ref Values	Ref Ids	
Default	device_type	==	desktop	0	View Rules
Default	platform_id	==	Windows	4	View Rules

Showing 1 to 2 of 2 entries

Previous 1 Next

Рисунок 4.6 – приклад відображення нового правила серед правил шаблону

Edit Rule

ID: 15083

Bucket template: Default

Rule type: device_type_name

Ref id: desktop

Flag Equal: ==

Status: inactive

Edit **Cancel**

Рисунок 4.7 – приклад редагування правила шаблону

```

$group_arr := $this->getModel('Buckettemplate')->getTemplate($session['ad_bucket_id'], $data_bucket_templat);
$session['group_id'] := $group_arr['group_id'];
$session['bucket_template_id'] := $group_arr['bucket_template_id'];
$this->getModel('Sessionbreakdown')->updateSB($session);

```

Р

Рисунок 4.8 – виклик функції вибору шаблону та збереження його

Edit Bucket Template

ID: 1

Probe Flag: No

Bucket: Base Config

Template: Default|All other browsers|iframed

Status: active

Weight: 100

Name: Default

Edit **Cancel**

Р

Рисунок 4.9 – приклад редагування параметрів шаблону

```

$rules_model := new RulesetsModel();
$res_array := array();
$rules := array();
$group_id := 1; //default template
$bucket_template_id_res := 0;

$sql_query := "SELECT * FROM `{$this->db_name}`.`{$this->table_name}` as b WHERE
»   »   »   b.`ad_bucket_id` = ".intval($bucket_id)." AND
»   »   »   b.`status` = ".intval(20);

$result := $this->db->query($sql_query);
if ($result && $result->num_rows > 0) {
»   $res_array := $result->fetch_all(MYSQLI_ASSOC);

»   foreach ($res_array as $key => $bucket_template) {
»       //get rules by bucket template id
»       $rules := $rules_model->getRulesByBucketTemplate($bucket_template['bucket_template_id']);
»       if (!empty($rules)) {
»           $status := $this->checkBucketTemplate($rules, $data);
»           if (!$status) {
»               unset($res_array[$key]);
»           }
»       }
»   }
}

if (empty($res_array)) {
»   $sql_query := "SELECT * FROM `{$this->db_name}`.`{$this->table_name}` as b WHERE
»   »   »   b.`ad_bucket_id` = 0 AND
»   »   »   b.`status` = ".intval(20);

»   $result := $this->db->query($sql_query);
»   if ($result && $result->num_rows > 0) {
»       $res_array := $result->fetch_all(MYSQLI_ASSOC);

```

Рисунок 4.10 – Фрагмент коду вибору списку шаблонів по заданим параметрам та правилам.

На рисунку 4.10 був наведений приклад вибірки підмножини шаблонів на основі правил та параметрів, але щоб вибрати один конкретний шаблон бу реалізований алгоритм «random weighted strategy» що бере вагу кожного шаблону помічає діапазон в якому знаходиться даний шаблон та сумує всі ваги підмножини шаблонів. Після цього формується випадкове число поміж 0 та сумою всіх шаблонів. Після формування числа робиться перевірка в який з діапазонів попадає згенероване число та вибирається той шаблон.

Цей метод дозволяє більш рівномірно розподілити трафік поміж усіх шаблонів. Приклад реалізації наведений на рисунку 4.11.

```

//set.group.id.using.random.weighted.strategy
if (!empty($res_array)) {
    » $sum := 0;
    » $random_val := 0;
    » $res_id := 1;
    » $weight_array := array();
    » foreach ($res_array as $arr_id => $bucket_template_val) {
    »     » if ($bucket_template_val['weight'] > 0) {
    »         » $weight_array[$bucket_template_val['bucket_template_id']] := array(
    »             » 'bucket_template_id' => $bucket_template_val['bucket_template_id'],
    »             » 'start' => $sum,
    »             » 'end' => $sum + $bucket_template_val['weight'],
    »             » 'group_id' => $bucket_template_val['group_id']
    »         );
    »         » $sum += $bucket_template_val['weight'];
    »     }
    » }

    » if ($sum > 0) {
    »     » $random_val := random_int(0, ($sum - 1));
    »     » foreach ($weight_array as $bucket_template_id => $weight) {
    »         » if ($weight['start'] <= $random_val && $weight['end'] > $random_val) {
    »             » $group_id := $weight['group_id'];
    »             » $bucket_template_id_res := $bucket_template_id;
    »             » break;
    »         }
    »     }
    » }
}

```

Рисунок 4.11 - приклад реалізації методу «random weighted strategy» для вибору шаблону

Наступним кроком в реалізації методу буде пошук підмножини рекламних кампаній які відповідають вхідним даним користувача. Рекламні кампанії як і шаблони мають правила. Додавання, редагування та відстежування цих правил також було додано до графічного інтерфейсу адмін панелі.

Ad Campaign

+ Add "Ad Campaigns"

Pages: All

Buckets: All

Status: active

Show exit pop: All

Show 50 entries

Search:

Ad Campaign id	Ad Campaign Name	Ad Bucket Name	Ad Page Name	Limit	Megapop Limit	Megapop Delay	Show exit pop	TabScroll Limit	Weight	Switch	Template Name
1	DefaultCar ALL - MA	Default	Car rates page	12	999	0	False	10	1000		default
2	DefaultFlight - MA	Default	Flight rates page	12	999	0	False	10	100		default
3	DefaultHotel	Default	Hotel rates page	12	999	0	False	10	100		default
7	ListingCar	Default	Car listing page	10	999	0	False	10	100		default
8	ListingFlight	Default	Flight listing page	10	999	0	False	10	100		default

Рисунок 4.12 – відображення списку рекламних кампаній

Для більш зручного користування системою з кожної рекламної кампанії можна перейти на перегляд правил для неї.

Template Name	Status	Groups		
default	Active	1 - - 4 -	Edit	Products Rules

Рисунок 4.13 – перехід до правил кампанії

Campaigns Rules Listing

+ Add Rule

Ad Campaign: DefaultFlight - FBT

Rule Type: All

Status: active

Show 50 entries

Search:

ID	Ad Campaign	Rule Type	Ref id	Ref Val	Flag equal	Status	
15	DefaultFlight - FBT	device_type	0	desktop	==	Active	Edit
16	DefaultFlight - FBT	device_type	2	tablet	==	Active	Edit

Showing 1 to 2 of 2 entries (filtered from 39 total entries)

Previous1Next

Рисунок 4.14 – приклад відображення всіх активних правил для вибраної рекламної кампанії

Рисунок 4.15 приклад створення нового правила для рекламної кампанії

Розглянемо частину програмного коду яка відображає загальний метод вибору рекламної кампанії на рисунку 4.16.

```

//.main.campaign.data
if (!empty($ad_campaign_id)) {
    » $ad_campaign := CacheManager::cacheRead("campaigns_ad_campaign_id_{ad_campaign_id}");
    » if ($ad_campaign === false) {
    »     » $ad_campaign := $this->getModel('Campaigns')->get($ad_campaign_id);
    »     » CacheManager::cacheWrite("campaigns_ad_campaign_id_{ad_campaign_id}", $ad_campaign);
    » }
} else if (empty($switch_id) && !empty($ad_campaign := $this->getModel('Campaigns')->getCampaignsByRulesPriority($page_id, $session))) {
    » $ad_campaign_id := (isset($ad_campaign['ad_campaign_id']) ? $ad_campaign['ad_campaign_id'] : 0);
} else {
    » $ad_campaign := (!empty($ad_campaigns := $this->getModel('Campaigns')->getCampaign($page_id, $session, $switch_id))) ? $ad_campaign : (isset($ad_campaign['ad_campaign_id']) ? $ad_campaign['ad_campaign_id'] : 0);
}

$ad_units := $this->getModel('Prunits')->prepareAdUnits($ad_campaign, 0, array(), $ad_unit_type_id, $session);

//.default.bucket.campaigns
$fallback_campaigns := $this->getModel('Campaigns')->getFallbackCampaigns($page_id, $session, $ad_campaign_id);

```

Рисунок 4.16 – вибір рекламної кампанії

```

/**
 * Get campaigns after filtering by rules
 */
@param-mixed $ad_campaigns_array
@param-mixed $session
*/
private function getCampaignsAfterRulesFiltering($ad_campaigns_array, $session)
{
    » $campaigns_rules_model := new CampaignsRulesModel();
    » $campaigns_with_rules := $campaigns_rules_model->checkCampaignsRules($ad_campaigns_array, $session);
    » return array_values($campaigns_with_rules);
}

```

Рисунок 4.17 – метод вибору кампаній враховуючи правила

Після виконання методу «getCampaignsAfterRulesFiltering» система одержує список кампаній, що були відфільтровані з урахуванням правил.

Тепер розглянемо сам вдосконалений метод, що аналізує дані та виставляє потрібні ваги перед вибором кампанії методом «random weighted strategy».

```
public function getCampaignsByRulesPriority($ad_page_id, $session) {
    {
        $campaigns_by_suitable_buckets := $this->getCampaignsByBuckets($ad_page_id, $session, 0, 0, 1);
        if (!empty($campaigns_with_rules := $this->getCampaignsAfterRulesFiltering($campaigns_by_suitable_buckets, $session))) {
            //check and update weight of each campaign
            $campaigns_with_rules := $this->setCampaignsWeightUsingStat($campaigns_with_rules);
            //get campaign by weighted random
            $campaigns_with_rules := $this->sortWeightedRandom($campaigns_with_rules, 1);
            return reset($campaigns_with_rules);
        } else {
            return [];
        }
    }
}
```

Рисунок 4.18 - приклад реалізації вдосконаленого методу

Після вибору потрібної рекламної кампанії робиться запит до рекламодавців що прив'язані до неї та формується відображення рекламних продуктів у шаблоні.

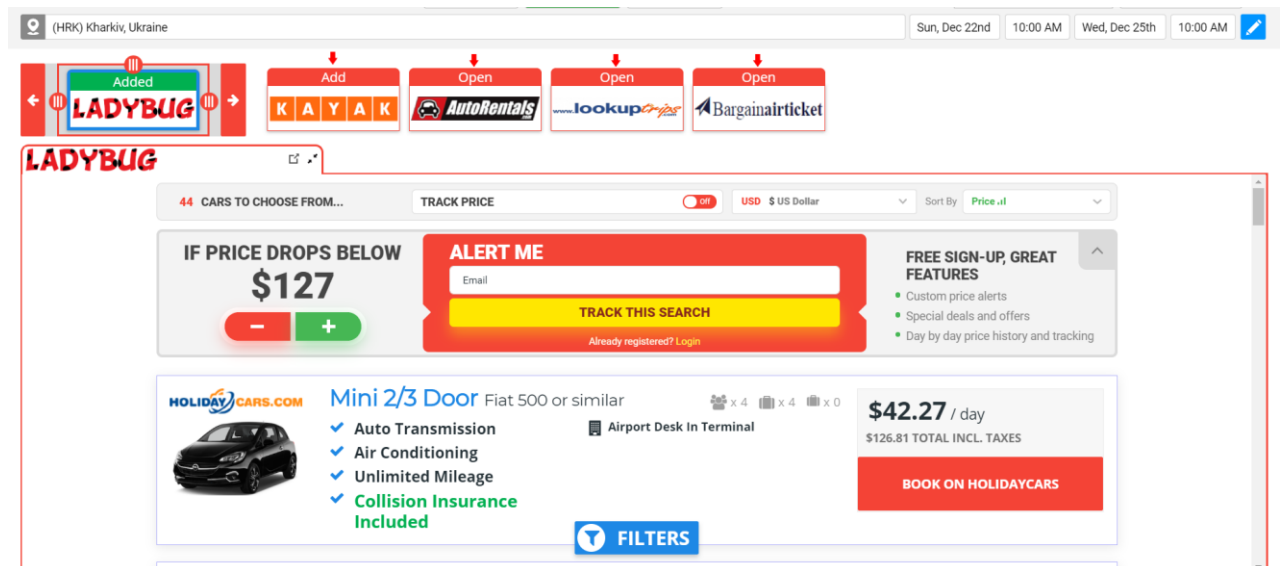


Рисунок 4.18 - приклад відображення рекламного продукту в шаблоні.

Для роботи методу зміни вагів використовується статистика, що щоденно аналізується та прораховуються коефіцієнти для правильного вибору ваги. Ця статистика зберігається в базі даних.

ad_dailystats_id	date	ad_page_id	device_type	advertiser_id	ad_campaign_id	total_ads	total_shown	total_searches	daily_b_1	daily_b_2
39022215	2019-01-12	2	0	2	18	8	8	8	0.0026000000	0.0000580000
39022216	2019-01-12	2	0	2	27	24	24	8	0.0032000000	0.0000310000

Рисунок 4.19 – приклад статистичних даних

4.2 Експериментальна перевірка удосконаленого методу розподілу рекламного трафіку

В цьому експерименті будемо використовувати реалізовану систему управління інтернет рекламою, яка займається показом рекламної продукції компаній по прокату автомобілів. В цю систему нещодавно був впроваджений удосконалений гібридний метод фільтрації на основі знань об'єднаний з методом який реалізує модель лінійної регресії.

Спочатку в структурі системи був реалізований тільки метод фільтрації на основі знань, що давало приблизно такі результати(рисуюнок 4.20).

Show entries Search:

User accounts	Sessions	Total Search Count	Sessions With Search	Sessions With Search %	Sessions With Revenue	Sessions With Revenue %	Rev/Session	Template
459	55562	64189 Avg/Search 1.1553	42161	75.881%	13807	24.850%	0.1296	Probe iframed - mega

Рисунок 4.20 – приклад денної статистики по вибраному шаблону до впровадження методу

Головним показником тут є прибуток на одну сесію (Rev/Session). Прийmemo його за середній результат. Потім було прийняте рішення впровадити удосконалений метод розподілення трафіку.

На наступний день ми отримали такі результати(рисуюнок 4.21)

Applied Filters: [x Group By = Template](#)

Show entries Search:

User accounts	Sessions	Total Search Count	Sessions With Search	Sessions With Search %	Sessions With Revenue	Sessions With Revenue %	Rev/Session	Template
391	54907	58150 Avg/Search 1.0591	38423	69.978%	12710	23.148%	0.1151	Probe iframed - mega

исунок 4.21 – результат першого дня тестування

Як ми бачимо, результат трохи погіршився, але це може бути пов'язане як з нормальною похибкою в роботі системи, так і з недостатньою кількістю даних. Тому почекавши ще 2 дні ми здобули таких результатів(рисуюнок 4.22)

Applied Filters: [x Group By = Template](#)

Show entries Search:

User accounts	Sessions	Total Search Count	Sessions With Search	Sessions With Search %	Sessions With Revenue	Sessions With Revenue %	Rev/Session	Template
380	52123	54389 Avg/Search 1.0435	36482	69.992%	12305	23.608%	0.1550	Probe iframed - mega

Рисуюнок 4.22 – 3 дні після початку тестування

Показник прибутку став більше більш ніж на 20% у порівнянні з першим пошуком, але був виставлений аналіз статистики за останні 7 днів, а з початку тестування пройшло 3. Це означає що, можливо результати можуть ще покращитися.

Після семи днів тестування ми переглянули статистику системи й отримали такі результати(рисунок 4.23).

Applied Filters: Group By = Template

Show 50 entries

Search:

User accounts	Sessions	Total Search Count	Sessions With Search	Sessions With Search %	Sessions With Revenue	Sessions With Revenue %	Rev/Session	Template
462	54010	53847	35989	66.634%	12489	23.123%	0.1697	Probe iframed - mega
		Avg/Search 0.9970						

Рисунок 4.24 – результат після 7 днів тестів.

На цей раз показник прибутку зріс навіть у порівнянні з попереднім тестом, що дуже позитивно впливає на загальний прибуток кампанії.

Для більшої наочності зробимо таблицю з результатами тесту.

Таблиця 4.2 – результати тесту

Днів з початку тесту	Прибуток за сесію	Різниця з першим тестом в %
0 (перший тест)	0.1296	0%
1	0.1151	-10.7%
3	0.1550	+20.1%
7	0.1697	+31.5%

З отриманих результатів можна побачити, що застосування удосконаленого методу розподілу рекламного трафіку більш ніж на 30 відсотків покращило показники прибутку що, безперечно, є дуже хорошим результатом, та значно покращило роботу системи.

ВИСНОВКИ

В результаті виконання магістерської атестаційної роботи було досліджено процес та методи розподілу та персоналізації рекламного трафіку для рекламодавця.

Під час виконання магістерської роботи були вирішені наступні задачі.

- досліджені особливості рекомендаційної підсистеми в системі управління інтернет-рекламою;
- проаналізовані та порівняні гібридні методи фільтрації, які можна застосувати для розподілу та персоналізації рекламного трафіку для рекламодавця;
- були проаналізовані традиційні методи представлення знань;
- був удосконалений метод розподілу рекламного трафіку в системі управління інтернет-рекламою з урахуванням попередньої фільтрації на основі знань;
- була проведена експериментальна перевірка удосконаленого методу для підтримки розподілу та персоналізації рекламного трафіку для рекламодавців в системі показу реклами компаній прокату автомобілів.

Отриманим результатом роботи є удосконалений метод, який підвищує ефективність розподілу рекламного трафіку між рекламодавцями, який в практичному аспекті показав збільшення середнього прибутку більш ніж на 30%. Вдосконалений метод і далі успішно використовується в системі.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Методичні вказівки з дослідницької практики за спеціальністю 122 – Комп’ютерні науки та інформаційні технології, спеціалізація - Інформаційні управляючі системи і технології. Освітньо-кваліфікаційний рівень – магістр / Упоряд.: В.М.Левикін, Д.К. Міхнов, – Харків: ХНУРЕ, 2016. – 7 с.
2. СУЧАСНІ ОСОБЛИВОСТІ РОЗВИТКУ ІНТЕРНЕТ-РЕКЛАМИ:
<http://elib.hduht.edu.ua/bitstream/123456789/3665/1/10.pdf>.
3. Web Page Recommendation Models: Theory and Algorithms:
<https://www.pdfdrive.com/web-page-recommendation-models-theory-and-algorithms-synthesis-lectures-on-data-management-e161157704.html>.
4. Adaptive Advertisement Recommender Systems for Digital Signage:
<https://www.pdfdrive.com/adaptive-advertisement-recommender-systems-for-digital-signage-e37280680.html>.
5. Recommender Systems: The Textbook:
<https://www.pdfdrive.com/recommender-systems-the-textbook-d187744586.html>
6. : Recommender Systems: An Introduction:
<https://www.pdfdrive.com/recommender-systems-an-introduction-e156651559.html>
7. Multiword keyword recommendation system for online advertising:
<https://www.pdfdrive.com/recommender-systems-the-textbook-d187744586.html>
8. Web page recommendation models theory and algorithms synthesis lectures on data management: <https://www.pdfdrive.com/web-page-recommendation-models-theory-and-algorithms-synthesis-lectures-on-data-management-e161157704.html>
9. Online Advertising Research in Advertising Journals: A Review:
https://www.researchgate.net/publication/241729514_Online_Advertising_Research_in_Advertising_Journals_A_Review
10. Литвин В. В. Задачі оптимізації структури та змісту онтології та методи їх розв’язування / В. В. Литвин // Вісник НУ “Львівська політехніка”. Серія: Інформаційні системи та мережі. –2011. – № 715 – С. 203–214.

11. C. Schank R., Abelson R.. Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures. – Lawrence Erlbaum Associates, 1977 – 256pp.
12. Чалий С.Ф., Лещинський В.О., Лещинська І.О. Моделювання контексту в рекомендаційних системах. Науковий журнал «Проблеми інформаційних технологій», 2018, №. 1(023). С. 21-26.
13. Чалий С.Ф., Лещинський В.О., Лещинська І.О. Персоналізація переліку товарів та послуг в системах електронної комерції з використанням контекстних фільтрів. Матеріали дев'ятої міжнародної науково-технічної конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління». С.35. Харків
14. С.Ф. Чалий, І.Б. Прібильнова Ситуаційна модель користувацького вибору в рекомендаційній системі.// Системи управління, навігації та зв'язку. – 2019. – Вип. 2(54). – С.159-163. doi:<https://doi.org/10.26906/SUNZ.2019.2.159>.
15. Chalyi S. Доповнення вхідних даних рекомендаційної системи в ситуації циклічного холодного старту з використанням темпоральних обмежень типу «next» / S. Chalyi, V. Leshchynskyi, I. Leshchynska // Системи управління, навігації та зв'язку. Збірник наукових праць. – Полтава: ПНТУ, 2019. – Т. 4 (56). – С. 105-109. – doi:<https://doi.org/10.26906/SUNZ.2019.4.105>.
16. Chalyi S. Концепція формування пояснень в рекомендаційних системах за принципом білого ящика / S. Chalyi, V. Leshchynskyi, I. Leshchynska // Системи управління, навігації та зв'язку. Збірник наукових праць. – Полтава: ПНТУ, 2019. – Т. 3 (55). – С. 156-160. – doi:<https://doi.org/10.26906/SUNZ.2019.3.156>.
17. Чалий С.Ф., Лещинський В.О., Лещинська І.О. Інтеграція локальних контекстів споживачів в рекомендаційних системах на основі відношень еквівалентності, схожості та сумісності. Process mining Materials of the VII International Scientific Conference «Information-Control System and Technologies» 17th-18th September, 2018, Odessa. С.142-144.

18. Чалий С.Ф., Прібильнова І.Б. Ситуаційне представлення споживачів рекомендаційної системи. Матеріали дев'ятої міжнародної науково-технічної конференції. С.35.
19. Чалий С.Ф., Прібильнова І.Б. Побудова багат шарового ситуаційного представлення вибору споживача рекомендаційної системи. Третя міжнародна науково-технічна конференція «Комп'ютерні та інформаційні системи і технології». Харків: ХНУРЕ. 2019. С. 114.
20. Попередня Обробка Вхідних Даних Рекомендаційної Системи з Урахуванням Змін у Вимогах Користувачів
21. Чалий С. Ф., Лещинський В. О. Узагальнене представлення знань для ситуації циклічного холодного старту в рекомендаційній системі. V Міжнародної науково-технічної конференції «Комп'ютерне моделювання та оптимізація складних систем» С. 170-173. м. Дніпро
22. Левикін В. М, Чала О. В. Розробка представлення причинно-наслідкових залежностей для бази знань системи процесного управління/ В. М. Левикін, // Вісник Національного технічного університету «Харківський політехнічний інститут». Серія: Системний аналіз, управління та інформаційні технології, 2018, № 21 (1297). С. 48-53.
23. Чала О. В. Розробка представлення знань на основі марківських логічних мереж в системі процесного управління. Вісник Національного технічного університету «Харківський політехнічний інститут». Серія: Системний аналіз, управління та інформаційні технології, 2018, № 22 (1298). С. 22-26.
http://repository.kpi.kharkov.ua/bitstream/KhPI-Press/37357/1/vestnik_KhPI_2018_22_Chala_Rozrobka_predstavlennia.pdf
24. Чала О. В. Побудова темпоральних правил для представлення знань в інформаційно-управляючих системах. Науково-технічний журнал, Сучасні інформаційні системи, 2018, Том 2, № 3. С. 54-59.
<http://ais.khpi.edu.ua/article/view/2522-9052.2018.3.09/147739>

25. The Univalent Foundations Program. Homotopy Type Theory: Univalent Foundations of Mathematics. – <https://homotopytypetheory.org/book>, 2013. – 475 p.
26. Semantic Network.: https://upload.wikimedia.org/wikipedia/commons/6/67/Semantic_Net.svg.
27. Noy N., McGuinness D. Ontology Development 101: A Guide to Creating Your First Ontology.: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.
28. Davies J., Fensel D., Harmelen F. On- To-Knowledge: Content-Driven Knowledge-Management through Evolving Ontologies. – John Wiley & Sons, 2002. – 312 p.
29. Regression modeling strategies with applications to linear models logistic and ordinal regression and survival analysis: <https://www.pdfdrive.com/regression-modeling-strategies-with-applications-to-linear-models-logistic-and-ordinal-regression-and-survival-analysis-e157915101.html>
30. Linear regression analysis: <https://www.pdfdrive.com/linear-regression-analysis-e2065764.html>
31. Brachman R., Levesque H. Knowledge Representation and Reasoning. – Morgan Kaufmann Publishers, 2004. – 381 pp.