

УДК 004.415.2

Білогуб Д.С.,  
Олещенко Л.М., к.т.н.

## **МОДЕЛЮВАННЯ РОЗМІЩЕННЯ НОВИХ ЗАКЛАДІВ ВІДПОЧИНКУ З ВИКОРИСТАННЯМ ДАНИХ GOOGLE MAPS ТА GOOGLE ANALYTICS**

Національний технічний університет України «КПІ імені Ігоря Сікорського»

[dasha.bilogub@gmail.com](mailto:dasha.bilogub@gmail.com),  
[oleshchenkoliubov@gmail.com](mailto:oleshchenkoliubov@gmail.com)

*Проаналізовано існуючі програмні рішення, їх недоліки для вирішення задачі визначення оптимального місця розташування майбутнього закладу відпочинку в межах міста. Запропоноване програмне рішення, яке дозволяє виконувати моделювання оптимального місця розташування майбутнього закладу відпочинку, використовуючи дані он-лайн сервісів Google Maps і Google Analytics*

**Ключові слова:** програмне забезпечення, ArcGIS Spatial Analyst, Google Maps, модель Хаффа, MySQL, ймовірність відвідування закладу, збір даних Google Analytics, аналіз веб-ресурсів, R, C#

### **Вступ та постановка проблеми**

Місце розташування закладів обслуговування та відпочинку визначає потенційну кількість клієнтів і оборот. Прогнозування обороту є центральною і найбільш складною процедурою при виборі місця розташування. Попит на послуги закладу має географічну спрямованість. Необхідно визначити максимальну відстань, на яку згоден переміщатись клієнт для відвідування закладу, розрахувати чисельність населення, яке проживає в межах кола з даними радіусом і визначити кількість конкурентів у даному регіоні. Отримані значення потім використовуються в якості основи для прогнозу обороту нового закладу.

Існують програмні рішення, що виконують аналіз за визначеними факторами та аналізують рельєф території, застосовуючи 3D-моделювання для більш наочного перегляду. Недоліками даних систем є те, що дані, що використовуються для аналізу, можуть бути застарілими (або користувачеві самому необхідно вводити дані для аналізу) та відсутність безкоштовних версій продуктів.

На даний час більшість замовлень послуг закладів відпочинку (замовлення квитків в кінотеатр тощо) здійснюється

через мережу Інтернет. Користувачі сайтів відповідних закладів залишають відгуки та оцінюють заклади, використовуючи он-лайн сервіси. Виникає можливість відслідковувати кількість відвідувачів сайту закладу та його популярність у мережі, порівнюваючи таким чином ймовірність відвідування самих закладів.

Таким чином, розроблюваний додаток має представляти собою прикладний програмний інтерфейс, за допомогою якого підприємці зможуть визначати оптимальне розміщення закладу використовуючи дані з мережі Інтернет. Також система має надавати можливість для звичайних користувачів переглядати інформацію про заклади, визначати найближчі заклади-конкуренти.

Для виконання поставленої задачі було розроблено наступні вимоги.

1. Система повинна оновлювати дані, отримуючи їх з Інтернету. Тобто, користувач не повинен вводити жодних даних, окрім місця розташування, яке необхідно проаналізувати. Всі дані система знаходить в мережі Інтернет;

2. Система повинна синхронізуватись з Google Maps;

3. Система повинна бути доступна для роботи на будь-якій ОС, незалежно від ПЗ комп'ютера;

4. Система повинна мати зручний інтерфейс. Тобто, користувач може працювати з програмою, не маючи спеціальних навиків для цього.

Для проектування нових розважальних центрів та закладів відпочинку необхідно виконати аналіз таких факторів:

1. площа об'єкту;

2. кількість конкурентів поблизу та їх популярність згідно он-лайн даних;

3. екологічні чинники (наявність шкідливих заводів поблизу, тощо);

4. можливість відвідувачів їздити з різних точок міста.

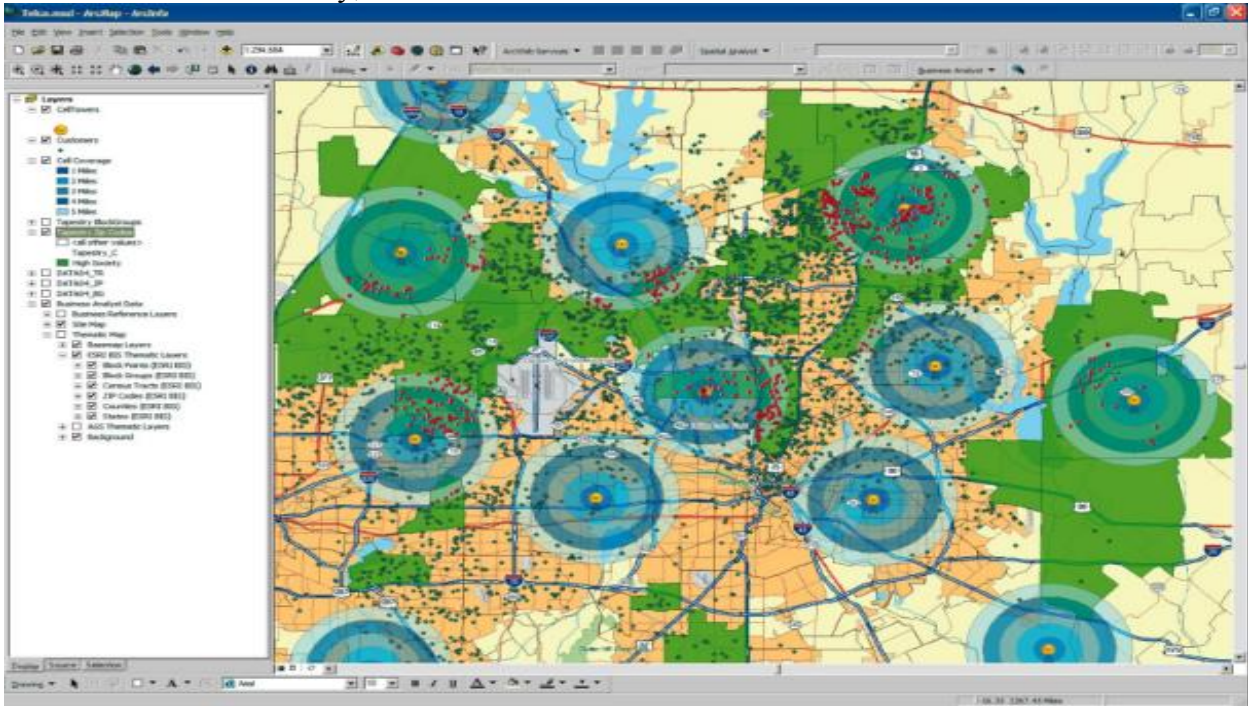


Рис.1. Результат визначення оптимального місця розташування закладу [1]

### **Огляд існуючих програмних рішень**

*ArcGIS Spatial Analyst* (рис. 1) пропонує широкий спектр потужних функцій просторового моделювання та аналізу. Можна створювати, запитувати та аналізувати растрові дані на базі клітин; виконувати інтегрований растровий / векторний аналіз; отримувати нову інформацію з існуючих даних; запит інформації за декількома шарами даних; і повністю інтегрувати клітинні растрові дані із традиційними джерелами векторних даних. Остання версія додатку інтегрована в середовище Python. *ArcGIS Spatial Analyst* інтегрує дані користувача, що дозволяє взаємодіяти між даними різних типів. Зображення, моделі висот та поверхні можуть бути об'єднані з автоматизованим проектуванням (САПР), векторними да-

ними, даними в Інтернеті та багатьма іншими форматами для забезпечення інтегрованого середовища аналізу. *ArcGIS Spatial Analyst* може створювати растрові дані з будь-якого джерела функцій будь-якої точки, лінії або багатокутника, таких як покриття *ArcInfo*, фігурні файли, бази даних геоданих, файли *CAD*, файли формату векторного продукту та теми *ArcGis*, створені за допомогою табличних даних. Крім того, дані в стандартному форматі можуть бути імпортовані, включаючи *TIFF*, *JPEG*, *BMP*, *USGS DEM*, *DTM*, *NIMA DTED*, загальні *ASCII*, *MrSID* та інші [1-2].

*ArcGIS* для сервера дозволяє працювати безпосередньо із збереженими просторовими даними керованими комерційними базами даних, які підтримують просторові типи. *ArcGIS for Server Basic*

Edition, дозволяє увімкнути функції для читання лише для своїх даних.

Ці службові функції дозволяють відображати та запитувати базу даних про інформацію на карті з веб-переглядачів та мобільних пристроїв. В додатку підтримуються наступні бази даних:

- ALTIBASE®;
- Amazon RDS для Microsoft SQL Server;
- Amazon RDS для PostgreSQL;
- Dameng;
- IBM® DB2®;
- IBM Informix® Dynamic Server;
- IBM Netezza®;
- Microsoft® SQL Server®;
- База даних Microsoft Azure® SQL;
- Oracle®;
- PostgreSQL;
- SAP HANA®;
- Teradata®.

Всі бази даних доступні лише для читання в ArcGIS for Server Basic edition і доступні в режимі читання-запис в ArcGIS for Server Standard and Advanced editions (крім Netezza, який є лише для читання). Мобільна версія додатку – відсутня.

Незважаючи на переваги, даний додаток має ряд недоліків. З точки зору програмної розробки:

- програма вимагає потужного ПЗ від комп'ютера, що є проблематичним для звичайних користувачів;
- введення даних вручну, що зазвичай невідомі для користувача.

З точки зору користувача:

- має незрозумілий інтерфейс;
- потребує певних навичок для коректної роботи з програмою.

Додаток ArcGis Spatial Analyst пропонує даний алгоритм у вигляді скрипту, написаного на мові R.

### Модель Хаффа

Модель Хаффа – це модель просторової взаємодії, яка обчислює ймовірність на основі гравітації споживачів при кожному місці походження, що патронує кожний заклад у наборі даних закладу. З цих ймовірностей потенціал продажу може бути розрахований для кожного місця походження на основі наявного доходу, населення або інших змінних.

У моделі Д. Хаффа для оцінки областей роздрібною торгівлі споживачі мають ряд альтернативних купуючих можливостей. Торгові області є ймовірнісними, а не детермінованими, кожна торгова точка має деяку ймовірність постійного відвідування, корисність пункту роздрібною торгівлі залежить від його розміру ( $S$ ) і відстані до нього ( $D$ ). Ймовірність  $P_{ij}$  – відвідування споживачем деякого закладу рівна відношенню його корисності ( $U_{ij}$ ) до суми корисності усіх закладів, що розглядаються споживачем (модель мультипликативної взаємодії, MCI):

$$P_{ij} = \frac{U_{ij}}{\sum_{k=1}^n U_{ik}} = \frac{S_j^{\alpha} \cdot D_{ij}^{\beta}}{\sum_{k=1}^n S_k^{\alpha} \cdot D_{ik}^{\beta}}, \quad (1)$$

де  $P_{ij}$  – ймовірність відвідування споживачем з міста  $i$ , пункту  $j$ ;  $n$  – набір конкуруючих пунктів (або міста) у регіоні,  $U_{ij}$  – корисність пункту  $j$  для індивіда в  $i$ ,  $S_j$  – площа закладу  $j$ ,  $D_{ij}$  – відстань між споживачем в  $i$  та пунктом  $j$ ,  $\alpha$  і  $\beta$  – параметри чутливості (згідно закону Рейлі:  $\alpha = 1$  і  $\beta = -2$ ).

У загальному вираз для ймовірності вибору споживачем закладу має вигляд:

$$P_{ij} = \frac{\prod_{k=1}^q x_{kij}^{\beta_k}}{\sum_{k=1}^n \prod_{k=1}^q x_{kij}^{\beta_k}}, \quad (2)$$

де  $x_{kij}$  –  $k$ -та змінна, що описує точку  $j$  у ситуації вибору  $i$ ,  $\beta_k$  – показник чутливості функції корисності відносно  $k$ -ї змінної (або рівень еластичності по цій змінній),  $q$  – число змінних функції корисності.

Використовується також функція корисності мультиноміального виду:

$$P_{ij} = \frac{\exp(\sum_k \beta_k \cdot x_{kij})}{\sum_{j=1}^n \exp(\sum_k \beta_k \cdot x_{kij})}, \quad (3)$$

Для оцінки коефіцієнтів  $\beta_k$  необхідно знати фактори  $x_{kij}$  та ймовірності вибору



$P_{ij}$ . Значення цих змінних прив'язані до певних «ситуацій вибору», тобто ситуацій, для яких передбачаються однакові умови вибору. Як правило, чинники, що розмежовують множину усіх ситуацій на штучні групи або «ситуації вибору», одночасно входять у рівняння (2-3) як регресори.

Кожну ситуацію вибору асоціюють з відповідними ймовірностями  $P_{ij}$  та змінними  $x_{kij}$ . Оскільки дійсні значення ймовірності не відомі, вони замінюються усередненими по кожній ситуації оцінками розподілу переваг споживачів. У випадку, коли множина ситуацій вибору співпадає з множиною респондентів, ймовірності оцінюються через розподіл числа візитів, здійснених даним споживачем.

Після отримання масиву підготовлених для аналізу даних проводиться розрахунок оцінок часток  $P$  і параметрів споживачів. З цими даними аналітик оцінює модель (3) методами найменших квадратів або максимальної правдоподібності [3].

### **Система моделювання з використанням сервісів Google Maps та Google Analytics**

В якості рішення всіх вище перелічених недоліків існуючих рішень пропонується додаток для обчислення оптималь-

ного місця розташування закладів відпочинку, в основі якого лежить модель Хаффа (рис. 2).

Враховуючи те, що основним недоліком існуючих аналогів є те, що користувачеві необхідно вручну вводити дані, створений додаток повинен сам отримувати всю необхідну інформацію з мережі Інтернет. З боку користувача необхідно введення таких параметрів, як тип закладу та радіус обслуговування.

В якості факторів, що впливатимуть на аналіз, було обрано наступні.

Площа закладів конкурентів. Дана інформація отримується за допомогою Google Maps. Урахування площі конкуруючого закладу є доцільним, оскільки чим більший розмір об'єкту – тим ширшу сферу послуг він може запропонувати. І як результат, тим більше клієнтів приваблює заклад. Тобто, за допомогою площі закладу можна зробити ймовірнісну оцінку його популярності.

Відстань від житлових будинків до торгової точки, відстань від торгової точки до закладів конкурентів. Гіпотетично більші райони мають більшу торговельну зону (ареал) і це, відповідно, може притягувати споживачів.

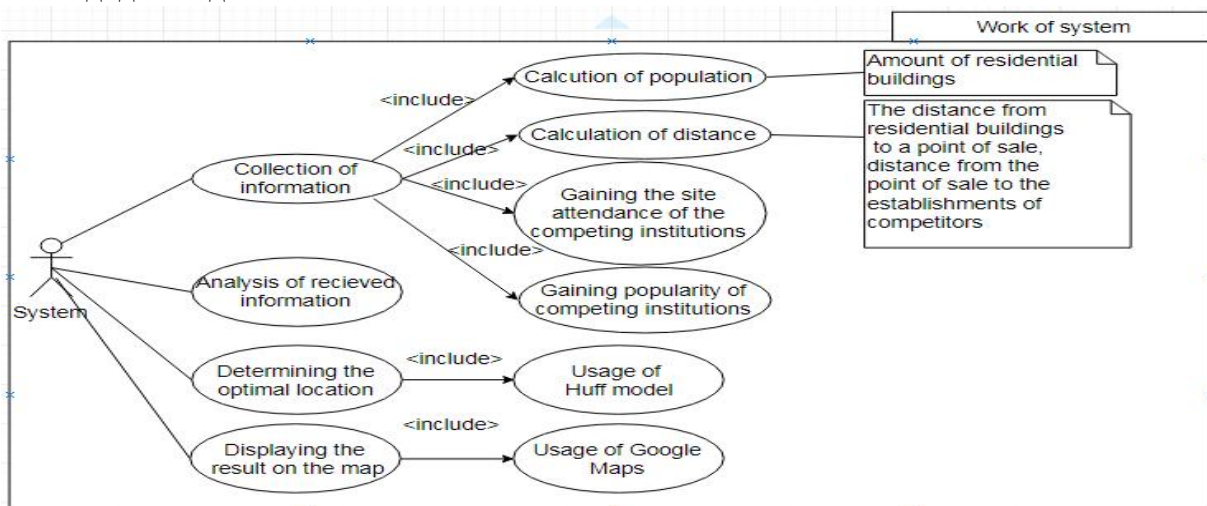


Рис. 2. Uml-діаграма роботи розроблюваної системи моделювання оптимального розміщення закладів відпочинку

Популярність конкуруючих закладів. Розуміється оцінка від відвідувачів,

яку можна отримати за допомогою Google Maps. Так як збір даних відбувається че-

рез мережу Інтернет, враховується оцінка від користувачів сайтів відповідних закладів. Сервіс Google Maps пропонує користувачам залишати коментар та оцінку для будь-якого об'єкту.

Відвідуваність сайту закладу. Враховуючи те, що в основному кожен заклад має свій сайт, враховується фактор «відвідуваність сайту закладу». Збір даної інформації виконується за допомогою Google Analytics.

Додаток розробляється на мові С#, оскільки С# – це повнофункціональна об'єктно-орієнтована мова, яка підтримує інкапсуляцію, наслідування та поліморфізм. Вона має підтримку компонентів, надійна та стійка завдяки використанню «збірника сміття», оброблення винятків, безпеки типів.

Для створення графічних інтерфейсів за допомогою платформи .Net застосовуються різні технології – Windows Forms, WPF. Проте більш простою та зручною платформою залишається Windows Forms або форми.

Блок-схема роботи алгоритму моделі Хаффа у даній програмній реалізації зображена на рис. 3.

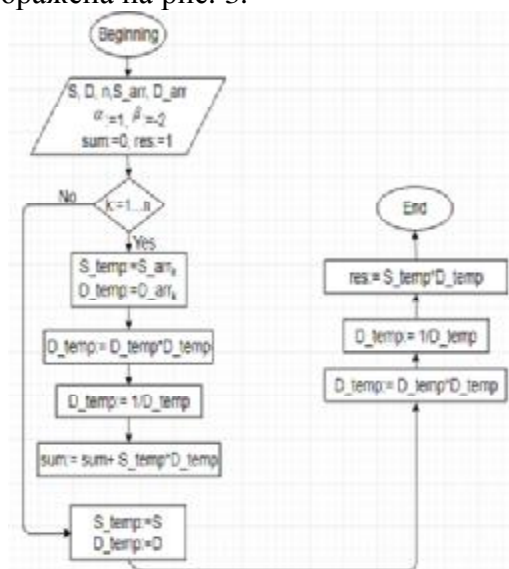


Рис. 3. Блок-схема роботи алгоритму моделі Хаффа

Завдяки використанню мови С# досягається прискорення роботи алгоритму, оскільки R має певні недоліки у порівнянні з С#: більшість функцій R значно перевантажені. У R перевантаження реалізується за допомогою однієї функції з багатьма необов'язковими названими параметрами; багато R-команд приділяють мало уваги керуванню пам'яттю, і тому R може дуже швидко споживати всю наявну пам'ять. Це може бути обмеженням при виконанні вилучення даних [4].

В якості бази даних використовується MySQL, оскільки MySQL характеризується високою швидкістю, стійкістю і простотою використання. Можливості сервера MySQL полягають у підтримці необмеженої кількості користувачів, що одночасно працюють із БД; кількість рядків у таблицях може досягати 50 млн; висока швидкість виконання команд; наявності простої і ефективної системи безпеки.

Всі дані зберігаються у форматі JSON, що забезпечує швидкий обмін даними невеликим об'ємом даних між клієнтськими браузерами та веб-службами AJAX.

В якості карти, що використовується в додатку, обрано Google Maps. Було обрано саме цей варіант карт, оскільки існують готові бібліотеки для роботи з картами, написані на С#.

Для збору статистики сайту обрано сервіс Google Analytics, оскільки він надає актуальні дані про дії користувачів на сайті, про те, як вони перейшли на нього та скільки відвідувачів було за певний проміжок часу. З точки зору розробника, цей сервіс зручний завдяки тому, що в С# є готові бібліотеки для роботи з даними утилітами. Google Analytics використовується для збору даних про дії користувачів з веб-сайтів, мобільних додатків, а також реального і цифрового оточень. Щоб відстежити сайт, Google Analytics використовує JavaScript, який збирає інформацію. При розміщенні цього скрипту на сторінці, він починає збір різноманітної інформації про дії користувача на сайті [5-7].

JavaScript може збирати інформацію безпосередньо з сайту, наприклад, URL-сторінок, які переглядає користувач. Скрипт може отримувати інформацію про

JavaScript може збирати інформацію безпосередньо з сайту, наприклад, URL-сторінок, які переглядає користувач. Скрипт може отримувати інформацію про

браузер користувача, наприклад, налаштування браузера, його назву, а також назву пристрою і ОС, які використовувалися для входу на сайт. JavaScript скрипт може передавати інформацію про джерела, що містять посилання на даний сайт, з яких користувач потрапив на сайт.

Всі ці дані пакуються і відсилаються на сервіси Google в чергу на обробку.

Для відстеження кількості замовлень, виконаних на сайті, об'єкт для продажу налаштовують таким чином, щоб він надсилав пакет даних кожен раз, коли покупка була здійснена. Такий пакет може містити в собі інформацію про місце знаходження магазину, артикул, дату покупки тощо.

### **Висновки**

Аналізуючи існуючий ринок програмних продуктів можна зробити висновок, що їх багато, проте далеко не всі виконують функції, необхідні для пересічних користувачів, зокрема, приватних підприємців. Основними недоліком даних програм є те, що користувачеві необхідно вручну вводити статистичні дані та те, що не враховуються різні чинники, що можуть впливати на розміщення закладів відпочинку. Основними перевагами є те, що дані додатки дозволяють відобразити 3D-моделі та те, що аналіз виконується не лише за тими параметрами, що згадуються вище, а й урахуванням рельєфу.

Тож, рішенням даної проблеми є створення програмного забезпечення, що отримуватиме всю необхідну інформацію з використанням даних мережі Інтернет. Аналіз даних виконується за допомогою ймовірнісної моделі Хаффа, що дозволяє визначити відносні пропорції, в яких

споживачі розподіляються між зонами різних закладів відпочинку.

### **Список літератури**

1. Опорний конспект лекцій з дисципліни «ГІС в задачах комп'ютерного моніторингу» // [Електронний ресурс] – Режим доступу: <http://kruzhan.vk.vntu.edu.ua/file/49d34849e12b4257705624e19cce3bb0.pdf>
2. ArcGIS Spatial Analyst: Advanced GIS Spatial Analysis // [Електронний ресурс] – Режим доступу: [https://www.esri.com/library/whitepapers/pdfs/arcgis\\_spatial\\_analyst.pdf](https://www.esri.com/library/whitepapers/pdfs/arcgis_spatial_analyst.pdf)
3. Олещенко Л. М. Моделювання оптимального місця розташування зони роздрібної торгівлі засобами ARCGIS. // Телекомунікаційні та інформаційні технології. – 2015. – №3. – С. 57-63.
4. Data Tools: R and RStudio // [Електронний ресурс] – Режим доступу: <http://analyticstrainings.com/?p=101>
5. SE Ranking URL // [Електронний ресурс] – Режим доступу: <https://seranking.ru/blog/seo/3-sposoba-prosmotra-poseshhaemosti-chuzhogo-sajta/>
6. How Does Google Maps Work? // [Електронний ресурс] – Режим доступу: <http://www.makeuseof.com/tag/technology-explained-google-maps-work/>
7. Google Analytics // [Електронний ресурс] – Режим доступу: <https://blog.aweb.ua/osnovy-raboty-s-veb-analitikoj-kak-rabotaet-google-analytics/>

Статтю подано до редакції 27.09.2017