

Побудова концептуальних сіток для аналізу тематичної близькості геопросторових об'єктів

Ірина Хміль

Кафедра соціальних комунікацій та інформаційної діяльності, Національний університет "Львівська політехніка", УКРАЇНА, м. Львів, вул. С. Бандери, 12,
E-mail: khmil.iryana@gmail.com

Abstract – Methods of forming conceptual grid of geospatial objects for further analysis of thematic proximity of the object are considered. A comparative analysis of similar and related services that help to automatically extract semantic core of large volumes of information is described.

Ключові слова – концептуальні сітки, геопросторовий об'єкт, концепт, дескриптор.

I. Вступ

Для аналізу інформаційного образу геопросторових об'єктів, які відображаються у наповненні віртуальних спільнот, можна використати декілька підходів. Одним з них є побудова концептуальної сітки, яка буде відображати тематичну близькість між двома геопросторовими об'єктами.

II. Складові концептуальної сітки

Під час аналізу зв'язку головного геопросторового об'єкту та інших об'єктів, які згадуються в масиві релевантної інформації ми можемо розглядати географічну і тематичну близькість. Географічну близькість встановлюємо на основі відстані об'єктів у реальному світі. Але така близькість не може відобразити інформаційного впливу близького об'єкту на аналізований, тому використовуємо тематичну близькість, під якою розуміємо зв'язок двох геопросторових об'єктів через слова-дескриптори – описові лексичні одиниці, що відображають певну характеристику об'єкта.

Для отримання складових, з яких буде формуватись концептуальна сітка потрібно виконати наступні кроки:

- 1) відібрати і екстрагувати всю доступну релевантну основному геопросторовому об'єкту (концепту) інформацію;
- 2) визначити множину інших геопросторових об'єктів (концептів другого порядку), що наявні у відібраній релевантній множині інформації;
- 3) екстрагувати множину дескрипторів, що описують основний аналізований геопросторовий об'єкт;

4) визначити, до яких інших геопросторових об'єктів, що згадуються в масиві відібраної релевантної інформації, є дотичними відібрані раніше дескриптори.

Дескриптори і концепти другого порядку можна отримати за допомогою методів, що базуються на аналізі положення терміну в тексті, а потім накладаючи додаткове обмеження на їх частоту появи у релевантному масиві: частотні міри, використання відносних частотних параметрів, визначення відношення «сигнал-шум» тощо [1].

Таким чином, фактично ми здійснюємо: індексування обраного масиву тексту, надання виділеним дескрипторам абсолютного і надалі відносного вагового коефіцієнта, що буде відображати важливість дескриптора, аналіз і розподіл дескрипторів на класи, а також визначення відношень між головним концептом і концептами другого порядку через дескриптор, що їх поєднує.

Після проходження цих кроків відношення «концепт-дескриптор» можна зобразити за допомогою сітки. При цьому будь-який концепт другого порядку може мати зв'язок з будь-яким дескриптором. Графічно відстань між концептом і дескриптором визначає міру їх спорідненості, а саме відображення їх спільної появи у визначеному обмеженому масиві тексту (речення, абзац, допис, дискусія тощо). Вона і буде відображати частоту появи визначеної пари концепту і дескриптора.

В залежності від поставлених завдань ми можемо побудувати сітку з множиною дескрипторів і концептів третього і далі порядку і віддалятися від головного геопросторового об'єкта, головного концепту, таким чином, маючи можливість акцентувати увагу та аналізувати конкретний вузол сітки. Для цього можемо використовувати різні методи і технології, одна з яких – fisheye [2]. При використанні цього методу представлення даних аналізований концепт стає головним, всі інші через зв'язки з дескрипторами стають концептами другого і далі порядків.

Висновок

Розглянуто кроки підготовки матеріалу для побудови концептуальних сіток для аналізу і відображення тематичної близькості геопросторових об'єктів.

Література

1. Солтон Дж. Динамические библиотечно-информационные системы. – М., Мир, 1979. – 557с.
2. Sarkar Manojit, Brown H. Marc, Graphical fisheye views. – Communications of the ACM 37 (12), 1994. – p.73 – 83.