

ІНТЕРАКТИВНИЙ ПОШУК РЕЛЕВАНТНОЇ ІНФОРМАЦІЇ З ВИКОРИСТАННЯМ RL-НАВЧАННЯ

Удовенко Сергій¹, Чала Лариса²

¹Харківський національний економічний університет імені С. Кузнеця

²Харківський національний університет радіоелектроніки

Анотація

В доповіді розглядається підхід до побудови інтерактивної процедури пошуку релевантної інформації, що передбачає використання машинного навчання з підкріпленням для вибору інструментів підтримки пошуку та модифікації запитів користувача з метою підвищення якості результатів пошуку. Здійснено програмну реалізацію запропонованої інтерактивної процедури.

Abstract

The paper looks at how to build of relevant information interactive search procedure that involves using machine learning with reinforcement to select search support tools and modify user queries to improve the quality of search results. Software implementation of the proposed interactive procedure has been implemented.

Вступ

Пошук інформації в ресурсах мережі Інтернет та у бібліотеках електронних документів іноді розглядається як значною мірою досліджений напрям. Сучасні інформаційно-пошукові системи (ІПС) дійсно вирішують основну проблему доступу до документів, але, в той же час, існує чимало конкретних випадків дослідження, коли основні інструменти ІПС не є достатніми. Зокрема, коли користувачі не володіють експертними знаннями про саму пошукову систему (відсутність системного досвіду) або коли їх запити є нечіткими (розпливчате визначення потреб користувача), результати пошуку є часто незадовільні. Такі ж труднощі виникають також у випадках, коли користувачі мають точні уявлення про свої потреби, але не про корпус чи лексичний каталог, який використовується в ньому для опису домену (неадекватність словникового запасу користувача). Це стимулює розвиток нових підходів в галузі інтерактивного пошуку інформації [1]. Основним завданням ІПС є пошук та і відбір з усіх доступних системі інформаційних ресурсів релевантної інформації, тобто такої, яка найбільш відповідає інформаційним потребам, сформульованим у запиті. Концепція інтерактивності полягає у створенні динамічних систем, здатних підвищити ефективність процедури взаємодії користувача та ІПС під час сеансу пошуку. У доповіді розглядається підхід до побудови такої інтерактивної процедури, що передбачає її навчання шляхом використання машинного навчання з підкріпленням (reinforcement learning – RL) для вибору інструментів підтримки пошуку та модифікації запитів користувача в реальному часі з метою підвищення якості результатів пошуку.

Основна частина

Інтеграція додаткових інструментів поліпшення пошуку в ІПС (додаткових модулів – ДМ) має базуватися на гнучкій архітектурі, яка дозволяла б різним інструментам пропонувати зміни на всій сторінці результатів, щоб не обмежувати їх повторним плануванням. Для опису взаємодії користувачів з кожним із ДМ, що реалізують різні підходи для модифікації запитів, пропонується модель взаємодії з користувачем, що на кожному кроці генерує спостереження $x(t) = (a, U, t, \Delta)$, де використовуються такі елементи: a – дія, здійснена користувачем, U – це сукупність аналізованих інформаційних одиниць (документів або фрагментів документу), t – час,

що минув з початку сеансу, Δ – тривалість дії a . Ці спостереження дозволять оновити допоміжні елементи пошуку.

Інтерактивний пошук релевантних результатів може здійснюватися в реальному часі за допомогою спеціального модуля СМ, що базується на використанні ідеї RL-навчання. Такий підхід передбачає відновлення двох циклів у процесі діалогу між системою та користувачем, присутніх у моделі пошуку інформації. Цикл запиту відповідає ритму, накладеному діалогом «запит / відповідь» між користувачем та системою, що зазвичай реалізується у більшості ІПС. Цикл взаємодії з користувачем враховує всі взаємодії користувачів, зокрема, кліки на кожен результат пошуку. Це дозволяє здійснювати динамічне оновлення інформації за допомогою суджень про релевантність результатів. Модуль СМ в запропонованій схемі RL-навчання відіграє роль агенту, який вирішує, які дії (тобто застосування комбінації допоміжних інструментів пошуку (ДП)) використовувати в будь-який час за допомогою навчання з підкріпленням. Користувач може бути асоційований із середовищем системи (джерелом корисної інформації для побудови моделі рішення), що реагує на прийняті рішення. Агент має за мету аналізувати елементи відповіді, надані ІПС з використанням ДП, та надавати допомогу користувачу або безпосередньо, відповідаючи на його потреби в інформації, або дозволяючи йому далі просунутися в пошуку. Стан навколишнього середовища S буде описуватися опосередковано через взаємодії, які буде мати користувач із системою, та через певні елементи, які можуть надходити від ІПС, зокрема, даних по корпусу документів та набору результатів для кожного нового запиту. Дії, доступні агенту, можуть бути згруповані в окремі категорії. Будемо вважати, що рішення буде прийнято в наборі доступних агенту дій A розмірності N (включно дію «нічого не міняти»). Оцінка N впливає на складність навчання з підкріпленням. Позначимо: m – кількість категорій ДП, а k_i – кількість конкуруючих можливостей для категорії i ; отже маємо $N = \prod_{i=1}^m (k_i + 1)$. Зауважимо, деякі з ДП не завжди можуть запропонувати відповідь (або оновлення неможливо) на черговому такті інтерактивного пошуку.

На першому етапі досліджень обмежимо набір усіх доступних дій лише однією категорією допомоги для пошуку, а саме, модифікації запитів, що може інтегрувати 4 різних підходи: М1 (локальний), який використовує неявну модель повернення релевантності; М2 (сліпий), який використовує модель повернення сліпої актуальності, М3 (метод граней), який відповідає представленню граней пошуку, розрахованих за списком результатів, та М4 (орфографічний), який пропонує орфографічну корекцію. Скорочення набору дій до $N = 4$ дозволяє суттєво обмежити складність проблеми навчання.

Розглянемо функцію поточної винагороди $r(s, a, s')$, де s' – стан, до якого переходить система зі стану s під впливом дії a , яка відповідає сигналу підкріплення з навколишнього середовища, мета якого – дозволити агенту судити про якість його останньої дії по відношенню до його довгострокової мети. Метою базового алгоритму навчання з підкріпленням є оптимізація довгострокового накопичення цієї винагороди [1]. У контексті проблеми інтерактивної модифікації запитів, реалізація винагороди може сприяти покроковому поліпшенню результатів пошуку за запитами користувача. В цьому разі можна просто використовувати середню частоту вибору запиту для кожної комбінації $\{s, a, s'\}$. Одним із завдань системи пошуку інформації є надання задовільних документів у верхній частині списку результатів. Це сприяє траєкторіям, коли користувач взаємодіє з документами, розміщеними вгорі списку результатів. Однак оскільки користувачі, природно, прагнуть взаємодіяти з цими самими документами, ми будемо використовувати різницю між рангом поточного документа та середнім рангом кількості документів, які зазвичай обробляються користувачами. Крім того, для оцінювання

взаємодії між користувачем та СМ може бути використано ентропійну міру невизначеності поточних результатів пошуку [2].

У контексті алгоритмів навчання з підкріпленням важливою є можливість дати достовірні оцінки винагород, які можна досягти при кожному рішенні системи. Отже, потрібно протестувати систему на реальних сесіях, щоб на кожному етапі порівняти прогнози системи з фактично отриманою винагородою. Тому для тестування запропонованого підходу доцільно обчислювати абсолютну помилку передбачення еволюції функції виграшу, задану наступним чином:

$$\varepsilon = \sum_{t=0}^T |r^*(s(t), a(t), s(t+1)) - r(t+1)|,$$

де $r^*(s(t), a(t), s(t+1))$ – прогнозування винагороди системи в стані $s(t)$; $a(t)$ – дія, обрана в момент часу t ; $s(t+1)$ – стан системи в момент часу $t+1$; $r(t+1)$ – фактичне значення отриманої винагороди.

Для отримання більш достовірної оцінки середньої системної помилки під час тестування доцільно провести перехресне оцінювання за всіма доступними даними. Таким чином, серед N сесій взаємодії, які будуть проведені з довідковою системою, ми виберемо конкретний сеанс, який буде використовуватися для тестування, інші сесії використовуються для попереднього навчання системи. Цей процес буде повторюватися стільки разів, скільки буде сеансів, тобто N . Таким чином, ми будемо змінювати кількість m сеансів, які використовуються для навчання, від 1 до $N-1$, і будемо порівнювати ефективність для кожного випадку. Кількість способів вибору m сеансів навчання дорівнює $(N-1)!/m!(N-1-m)!$ (Можливості) та навчання на різних сесіях може потенційно впливати на ефективність при однаковій кількості навчальних сесій. Отже, ми будемо оцінювати середні значення для кожного значення m , отримані для k різних виборів.

На основі запропонованого підходу розроблено програмний модуль (СМ) інтерактивного пошуку релевантної інформації за запитам користувачів з використанням нейромережевої реалізації модифікованого SARSA-алгоритму навчання з підкріпленням [3].

Проведений інтерактивний експеримент підтвердив працездатність запропонованого підходу. Модуль СМ здатний аналізувати зміни винагороди та покращувати їх оцінку за новими послідовностями. Крім того, RL-навчання дозволяє правильно розрізняти різні дії та їх вплив на навколишнє середовище. Перспективним є подальше удосконалення обчислювальних процедур СМ з проведенням репрезентативних експериментальних досліджень для завдань пошуку релевантних тексто-графічних документів за гібридними запитам.

Список використаних джерел

1. Dupont Gérard. A step toward an adaptive composition of query suggestion approaches / G. Dupont, A. Requier, S. Lecourtier, B. Grilhères S. Brunessaux // Proceedings of the third Symposium on Information interaction in context. – 2010. – P. 271-276.
2. Rachdi Mohamed. Semantic enrichment of queries with generic and specific terms in the definition sentences / M. Rachdi, H. Ben Lahmer, H. Labriji // Journal of Theoretical and Applied Information Technology. – 2014. – No.3, | Vol. 59 P. 33-41.
3. Удовенко Сергій Григорович. Модифікована процедура нейромережевого навчання з підкріпленням в задачах прогнозування / С. Г. Удовенко, Л. Е. Чала, В.Л. Шергін // Матеріали VIII Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології» (ІУСТ- ОДЕСА -2019) (23-25 вересня, 2019, Одеса, Україна) – 2019. – С. 180-182.