

## ПРО ЗАСТОСУВАННЯ PAGERANK-ПОДІБНИХ МЕТОДИК ПРИ ПОБУДОВІ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ДЛЯ ВИБОРУ НАВЧАЛЬНИХ ДИСЦИПЛІН

Олецький О.В.

Національний Університет “Києво-Могилянська Академія”,  
м.Київ, вул. Сковороди, 2, oletsky@ukma.edu.ua

У 2016-17 рр. в Національному Університеті “Києво-Могилянська академія” була розроблена система автоматизованого запису студентів на вибіркові навчальні курси [1], яка працює через Веб і дозволяє студентам обирати, які вибіркові дисципліни вони хотіли б вивчати. Очевидною є доцільність набуття цієї системою рис рекомендаційної системи, яка могла б давати рекомендації кафедрам під час вибору дисциплін, які пропонуються до запису, а також студентам під час запису.

В [1] описані деякі підходи до вироблення таких рекомендацій, перш за все за рахунок розв’язання оптимізаційних задач, в загальному випадку багатокритеріальних. Суттєва проблема полягає в тому, що цільові функції та обмеження часто можуть бути задалегідь невідомими або відомими неточно. Тому видається доцільним шукати більш інтелектуалізовані, “м’які” підходи до вироблення рекомендацій на основі ранжування варіантів. В цьому контексті видається доцільним застосування методик, подібних до PageRank [2].

В рамках цього підходу розглядається граф станів, вузли якого в найпростішому випадку відповідають навчальним дисциплінам. Для цієї задачі розглядаються також складені стани, які враховують, чи вже вивчав студент дану дисципліну, чи ні. Деякі міркування щодо аналізу складених станів обговорювалися в [3]. На цьому графі вводяться також зв’язки, що показують, які дисципліни мають бути вивчені попередньо. Ми розглядаємо такі типи зв’язків:

- одинична залежність: для вивчення дисципліни А студент має попередньо вивчити дисципліну В;
- І-зв’язки: для вивчення А необхідно вивчити і В, і С;
- АБО-зв’язки: для вивчення А необхідно вивчити або В, або С.

Відповідно до ідеології PageRank, вектор  $p$  мір важливостей вузлів отримується як лівий власний вектор матриці перехідних імовірностей  $\Pi$ , тобто як розв’язок рівняння

$$p\Pi = p$$

При цьому матриця перехідних імовірностей будується на основі тих чи інших породжуючих моделей. Зокрема, як базова може розглядатися наступна породжуюча модель:

1. Агент починає навігацію з випадково вибраного вузла серед тих, які ще не розглядалися.
2. З імовірністю  $\alpha$  агент переходить з вузла  $X$  до одного з вузлів, від яких  $X$  залежить, а з імовірністю  $(1-\alpha)$  – до випадковим чином вибраного іншого вузла.
3. При аналізі залежностей розрахунок перехідних імовірностей визначається типом залежності:
  - за наявності одиничної залежності  $X$  від  $Y$  агент переходить до вузла  $Y$  з імовірністю  $\alpha$ ;
  - за наявності І-залежності  $X$  від  $Y$  та  $Z$  агент переходить до кожного з цих вузлів з імовірністю  $\alpha/2$ .
 Якщо ж вузол, скажімо,  $Y$ , уже розглядався, перехід до  $X$  відбувається з імовірністю  $0$ , а до  $Y$ -з імовірністю  $\alpha$ ;
  - за наявності АБО-залежності  $X$  від  $Y$  та  $Z$  агент переходить до кожного з цих вузлів з імовірністю  $\alpha/2$ . Але, на відміну від попереднього випадку, якщо хоча б один з вузлів  $Y, Z$  уже розглядався, перехід до будь-якого з них відбувається з імовірністю  $0$ .
4. Розглядаються ймовірності повернення на залежні вузли; розгляд цих ймовірностей має дуже велике значення для розрахунку мір їх важливості та відповідно – для їх ранжування.

Замість повністю випадкового переходу до іншого вузла, можна вводити нечіткі міри релевантності, які визначаються семантичними відстанями від вузлами. Зокрема, в [4] розглядалося нечітке правило “Якщо відстань СЕРЕДНЯ, релевантність ВИСОКА”. Дійсно, найбільш релевантні курси повинні бути не найбільш віддаленими, але й не найменш віддаленими – найменш віддалений вузол буде надто схожим на даний і існує значний ризик того, що відповідний матеріал буде вже знайомий агентові. Іншими словами, функція залежності міри релевантності від міри близькості має спочатку зростати, досягати максимуму в районі середніх значень мір близькості, і потім знову спадати.

В результаті реального запису можна оцінити стаціонарні розподіли ймовірності, які безпосередньо пов’язані з кількістю студентів, записаних на кожний курс; на основі цього можуть коригуватися параметри моделі. Крім того, можна певним чином формалізувати такі поняття, як “об’єктивна корисність курсу”, “індивідуальна корисність курсу”, “якість курсу”.

Наведена породжуюча модель носить досить базовий характер. При її розвитку необхідно врахувати наступні міркування:

- окремо повинна розглядатися ситуація, коли певні курси є нормативними, тобто обов'язковими до вивчення;

- залежності між курсами також можуть мати нечіткий характер і вказувати, наприклад, наскільки важливим і необхідним для вивчення певного курсу є деякий інший;

- при формуванні перехідних імовірностей варто враховувати міри засвоєння курсів, які неминуче будуть мати нечіткий характер і можуть описуватися на основі теорії нечітких множин;

- суттєвий вплив на формування перехідних імовірностей може мати міра задоволеності студента від проходження курсу; врахування цього і попереднього факторів наближає модель до навчання з підкріпленням та до марковських моделей прийняття рішень [5, 6];

- наведена модель повинна надавати перевагу “інтегрованим” курсам, вивчення яких не вимагає попереднього вивчення інших курсів, а натомість ці курси включають певні елементи з них; в залежності від ситуації це може бути як добре, так і погано;

- очевидно є доцільність комбінування описаної методики з часом можливого перебування в окремих вузлах, що змістовно відповідає часові, який потрібний для вивчення курсу [3];

- крім аналізу на рівні цілісних курсів, можна проводити аналіз мір важливостей окремих модулів та понять як вузлів тематичної онтології; ця обставина може мати значення при динамічному формуванні самих курсів.

Описана методика може бути застосована не тільки для локальних університетських локальних систем, а й для масових відкритих онлайн-курсів (МООС), суттєвою проблемою для яких може виявитися дублювання навчальних курсів та навчальних матеріалів.

#### ЛІТЕРАТУРА:

1. Горборуков В.В. Сервіс автоматизованого запису на вибіркові навчальні дисципліни в НаУКМА та можливі напрямки його інтелектуалізації // Горборуков В.В., Олецький О.В. - Наукові записки НаУКМА. Комп'ютерні науки. Т.190. – К., 2016. – С.52-56.
2. Маннинг К.Д. Введение в информационный поиск. – Маннинг К.Д., Рагханан П., Шютце Х. М.: ООО «И.Д. Вильямс», 2011. – 528 с.
3. Олецький О.В. Моделювання часу перебування у вузлах тематичного порталу на основі процесів випадкового блукання на множині складених станів // Олецький О.В. Теоретичні та прикладні аспекти побудови програмних систем. Матеріали міжнародної конференції ТАAPSD'2015. Київ, 23 – 26 листопада 2015 р. – С.145-149
4. Олецький О.В. Про підхід до автоматичного формування рекомендацій для відвідувачів веб-порталу на основі теорії нечітких множин // Олецький О.В. - Наукові записки НаУКМА. Комп'ютерні науки. – К., 2015. – С.37-40
5. Олецький О.В. Про застосування марковських процесів прийняття рішень для автоматизованого добору навчальних матеріалів у системах blended learning. //Наукові записки НаУКМА. Комп'ютерні науки. – К., 2013. – С.115-118
6. Олецький О.В. Про деякі підходи до моделювання поведінки відвідувачів тематичного порталу на основі марковських процесів прийняття рішень. // Теоретичні та прикладні аспекти побудови програмних систем. Матеріали міжнародної конференції ТАAPSD'2013. Ялта. 25 травня – 2 червня 2013 р. – С.100-104

## ТЕХНОЛОГІЇ ФОРМУВАННЯ РЕКОМЕНДАЦІЙ В МОБІЛЬНИХ РЕКОМЕНДАЦІЙНИХ СИСТЕМАХ

Пасічник В.В.\*, Кунанець Н.Е.\*\*, Артеменко О.І.\*\*\*

\*Національний університет «Львівська політехніка», м. Львів, vpasichnyk@gmail.com,

\*\*Національний університет «Львівська політехніка», м. Львів, nek.lviv@gmail.com

\*\*\*Приватний вищий навчальний заклад «Буковинський університет», м. Чернівці, olga.hapon@gmail.com

Рекомендаційні системи як окремі класи систем підтримки прийняття рішень з'явилися на ринку інформаційних технологій в середині двохтисячних років з розвитком сектору інтернет-торгівлі та надання послуг на основі інтернет-технологій. Розвиток інтелектуальних технологій та поширення гаджетів, зорієнтованих на негайне задоволення різноманітних запитів свого користувача з врахуванням його смаків, обставин та потреб спричинив появу нового виду рекомендаційних систем – мобільних. Основними відмінними рисами таких технологій є:

- *Персоналізація* – гаджет, як правило, належить одному власнику і це дає змогу накопичувати та аналізувати контент та контекстні дані саме цього користувача.