

Personalizované odporúčanie využívajúce vizuálne stimuly

Peter Gašpar, Michal Kompan, Mária Bieliková

Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU v Bratislave
Ilkovičova 2, 842 16 Bratislava

{peter_gaspar, michal.kompan, maria.bielikova}@stuba.sk

Abstrakt. Analýza správania používateľa na Webe je v súčasnosti otvoreným výskumným problémom. Používatelia interagujú s webovými stránkami rôznymi spôsobmi a snahou výskumníkov je tieto interakcie správne interpretovať a použiť pre rôzne úlohy. Interakcie používateľov sú dôležitým vstupom pre odporúčacie systémy, ktoré sa využívajú ako jedna z metód pre zlepšenie používateľského zážitku na Webe. Obsahové odporúčanie využíva charakteristiky položiek pre generovanie personalizovaného odporúčania. Obrázky sa stali populárnym zdrojom informácií v mnohých doménach (napr. nakupovanie), kde ovplyvňujú to, ako sa používateľ rozhoduje. V našej práci využívame vizuálne črty extrahované z obrázkov. Naším cieľom je preskúmať zapojenie obrázkov a vizuálnych stimulov do procesu odporúčania so zámerom skvalitniť existujúce prístupy založené na obsahu.

Kľúčové slová: obsahové odporúčanie, analýza správania používateľa, extrakcia vizuálnych stimulov

1 Úvod

Štúdium správania používateľa na Webe sa v posledných rokoch stalo významným výskumným smerom. Mnohí výskumníci sa snažia analyzovať, ako ľudia interagujú s obsahom, aké sú ich preferencie a čo do najväčšej miery ovplyvňuje ich správanie počas rôznych úloh, ktoré vykonávajú na Webe. Poznanie používateľa a poznanie jeho správania otvára cestu pre skvalitňovanie prístupov na adaptáciu a personalizáciu v rôznych doménach.

Odporúčacie systémy vznikli ako jeden z nástrojov pre zlepšenie použiteľnosti. Ich hlavným cieľom je navrhovať používateľovi potenciálne zaujímavé a atraktívne položky. Ich hlavné využitie nachádzame najmä v procese rozhodovania človeka, a to najmä v prípade, keď je negatívne zasiahnutý problémom informačného preťaženia. S týmto problémom sa stretávame napríklad pri výbere filmu, pri nakupovaní v e-obchode alebo pri výbere hudby [3].

Obsahové odporúčanie využíva ako hlavný zdroj informácie extrahované z textu. Jedným z dôvodov je fakt, že text sa stal hlavným prostriedkom na šírenie informácií, a tak môže ovplyvňovať používateľov pri výbere položiek na Webe. Mnohé štúdie

*J. Steinberger, M. Zíma, D. Fiala, M. Dostal, M. Nykl (eds.)
Data a znalosti 2017, Plzeň, 5. - 6. října 2017, pp. 176-180.*

však ukázali, že aj obrázky môžu mať vplyv na rozhodovanie a môžu byť užitočné v procese odporúčania.

V tomto príspevku prezentujeme koncept dizertačnej práce. Vychádzame z myšlienky, že obrázky vplývajú na správanie používateľa na Webe, a preto ich zohľadnenie môže pomôcť pri návrhu systémov na odporúčanie obsahu. Naša prvotná analýza sa bližšie zamerala na obsahové odporúčanie obohatené o vizuálne črty. Naše hypotézy sme sa snažili podporiť experimentom, v ktorom sme porovnali dva rôzne prístupy k extrahovaniu vizuálnych črt z obrázkov.

2 Obsahové odporúčanie

Taxonómia personalizovaných odporúčacích systémov rozlišuje dva základné prístupy k odporúčaniam: kolaboratívne filtrovanie a obsahové odporúčanie. Kým *kolaboratívne filtrovanie* analyzuje podobnosť medzi správaním viacerých používateľov, *obsahové odporúčanie* skúma aktivitu každého používateľa samostatne a snaží sa mu odporúčať položky na základe tých, ktoré si obľúbil v minulosti. Pri výpočte podobnosti medzi položkami sa využívajú charakteristiky jednotlivých položiek – obsah [3]. Obsah vychádza z vlastností položiek, pričom tieto vlastnosti majú najčastejšie formát textu a závisia od domény, v ktorej sa pohybujeme – napr. v prípade filmov sú to žánre, herci či postavy.

S obsahovým odporúčaním sa najčastejšie stretávame práve v tých doménach, kde text tvorí jadro obsahu, napr. v doméne správ a filmov. Existujú však domény, v ktorých nemusia byť textové opisy postačujúce. Príkladom sú filmy (filmové plagáty), nakupovanie alebo umenie. V týchto prípadoch môže ovplyvňovať rozhodovanie aj výzor položiek.

V práci [8] boli obrázky použité na odporúčanie tovaru. Autori modelovali preferencie používateľa na základe črt, ktoré bolo možné identifikovať aj v obrázkoch (napr. farba tovaru). Zhou et al. [9] študovali odporúčač oblečenia, ktorý pre vstupný obrázok (reprezentujúci oblečenie) vedel vygenerovať podobné obrázky, pričom sa zohľadňovali viaceré črty (napr. tvar, farba). Zaujímavým problémom je však aj porovnanie obsahových a vizuálnych črt a ich významu pri odporúčaní.

2.1 Akvizícia vizuálnych črt

Pri získavaní vizuálnych črt z obrázkov môžeme rozlišovať tri základné prístupy: obrázkové deskriptory, farebné charakteristiky a sémantické črty získané z neuronových sietí.

Obrázkové deskriptory sa snažia extrahovať kľúčové body v obrázkoch, čím vznikajú spravidla binárne vektory aplikovateľné na rôzne úlohy, a to obzvlášť výpočet podobnosti obrázkov či rozpoznávanie objektov na obrázkoch. Medzi najznámejšie deskriptory patria SIFT, SURF a BRIEF.

Špecifickým druhom obrázkových deskriptorov sú *farebné charakteristiky*, pri ktorých pracujeme na úrovni jednotlivých pixelov obrázka a analyzujeme jeho farebné zložky. Výstupom tejto analýzy sú črty, ktoré dokážu charakterizovať vizuál

obrázka na základe farieb. Literatúra pritom najčastejšie pracuje s farebným modelom HSV (Hue-Saturation-Value), ktorý považujeme za blízky ľudskému vnímaniu farieb [6]. Tu môžeme skúmať nielen farbu, ale aj ďalšie charakteristiky, akými sú príjemnosť, dominancia a nabudenie vychádzajúce zo psychologických modelov.

Na opis obsahu obrázku využívame *sémantické črty*, pričom v tejto oblasti sa najčastejšie stretávame s konvolučnými neurónovými sieťami. V posledných rokoch vznikli viaceré prístupy a architektúry (napr. Inception-v3 [5] a VGG19 [4]), ktoré umožnili získať črty vhodné pre úlohu porovnania podobnosti obrázkov.

3 Porovnanie obsahových a vizuálnych črt v doméne filmov

V našom prvom experimente sme sa zamerali na analýzu filmových žánrov s využitím dátovej množiny MovieLens, ktorá obsahuje filmy a k nim prislúchajúce žánre [1]. Naším cieľom bolo porovnať vizuálne črty s obsahovými a zároveň porovnať dva druhy vizuálnych črt. Pre každý film sme pomocou The Movie Database API zvolili hlavný obrázok (plagát filmu), z ktorého sme extrahovali farebné charakteristiky a *sémantické črty*. Jednotlivé plagáty sme následne usporiadali do skupín podľa toho, do akého žánru prislúchali filmy. V rámci týchto skupín sme následne zisťovali priemernú kosínusovú vzdialenosť a varianciu medzi všetkými vektormi črt. Keďže niektorým filmom boli priradené viaceré žánre, umiestnili sme ich aj do viacerých skupín (napr. film majúci žánre komédia a romantický sa nachádzal aj v skupine filmov so žánrom komédia, ale taktiež aj v skupine filmov so žánrom romantický).

Na extrakciu farebných charakteristík sme využili farebný model HSV, pričom sme analyzovali jas, sýtosť a odtieň farieb, ale taktiež aj príjemnosť, dominanciu a nabudenie.

Pre získanie *sémantických črt* sme využili konvolučnú neurónovú sieť Inception-v3 [5] natrénovanú na dátovej množine ImageNet¹. Inception-v3 bola natrénovaná na úlohu predikcie kľúčových slov opisujúcich obsah obrázka.

Vychádzali sme pritom z práce [2], kde autori využili predposlednú vrstvu tejto natrénovanej siete. Pomocou siete sa pokúsili predikovať kľúčové slová z inej domény (na ktorej sieť nebola natrénovaná), po natrénovaní extrahovali predposlednú vrstvu siete a s použitím SVM klasifikovali produkty v e-obchode (ktoré sa v pôvodnej dátovej množine ImageNet nenachádzajú).

Tento princíp sme uplatnili aj my a nechali sme konvolučnú neurónovú sieť Inception-v3 predpovedať kľúčové slová pre filmové plagáty. Tie nás však nezaujímali, a preto predposledná vrstva predstavovala vektor črt reprezentujúci film.

Tabuľka 1 zobrazuje výsledky našej analýzy pre vybrané filmové žánre. Pri farebných charakteristikách bola pre niektoré žánre priemerná kosínusová vzdialenosť významne nižšia, ako pre iné žánre. Napríklad, ak porovnáme detský žánr s westernom, vidíme výraznejšie rozdiely. Vo všeobecnosti bola kosínusová vzdialenosť relatívne vyššia pri *sémantických črtách*, a tu sme už podobné rozdiely, ako pri farebných charakteristikách nespozorovali.

¹ <http://www.image-net.org>

Tabuľka 1. Porovnanie priemernej vzdialenosti (avg_D) a variancie (σ) pre farebné a sémantické črty v rámci jednotlivých filmových žánrov.

Žáner	Farebná charakteristika		Sémantické črty		Počet filmov
	avg_D	σ	avg_D	σ	
Detský	0.09	0.06	0.35	0.07	1 139
Animovaný	0.10	0.07	0.40	0.09	1 027
Dobrodružný	0.11	0.06	0.36	0.07	2 329
Komediálny	0.13	0.07	0.37	0.06	8 374
Fantazijný	0.13	0.08	0.37	0.06	1 412
Mysteriózny	0.17	0.09	0.37	0.07	1 514
Dramatický	0.15	0.09	0.38	0.06	13 344
Dokumentárny	0.21	0.10	0.40	0.07	2 471
Kriminálny	0.22	0.11	0.36	0.06	2 939
Vojnový	0.23	0.13	0.37	0.07	1 194
Westernový	0.33	0.15	0.34	0.07	676

Takéto zistenie nám môže naznačovať, že farebné charakteristiky sú dobrým indikátorom toho, do akého žánru patrí film. K týmto zisteniam sme dospeli aj manuálnou analýzou plagátov, kde sme odhalili, že napríklad plagáty v rámci žánru *Detský* sú si podobné na základe dominantnej farby, ktorá sa pri mnohých plagátoch opakuje. Naopak, samotný obsah (detekovateľný pomocou sémantických črt) sa pri týchto plagátoch líši.

Vo všeobecnosti, výsledky pre sémantické črty naznačujú, že nie sú vhodné na získanie informácie o žánri, avšak by mohli byť vhodné pre akvizíciu ďalších (latentných) charakteristík filmov.

4 Záver

V našej práci skúmame, ako možno využiť vizuálne stimuly extrahované z obrázkov pre skvalitnenie existujúcich prístupov k obsahovému odporúčaniam. V prvotnej fáze sme sa zamerali na experiment, v ktorom sme sa snažili zistiť, či vizuálne črty vedú nahraďovať obsahovú črtu. Využili sme na to rozsiahlu dátovú množinu z domény filmov. Výsledky naznačujú, že vizuálne črty majú potenciál pre ďalšie využitie v budúcej práci, avšak je potrebné preskúmať ich aplikáciu v konkrétnom prístupe odporúčaniam. Plánujeme preto navrhnuť hybridný Top-N odporúčač, ktorý by s pomocou vizuálnych stimulov vedel skvalitniť vygenerované odporúčania.

Literatúra

1. Harper, F. M., Konstan, J. A.: The MovieLens Datasets: History and Context. ACM Trans. Interact. Intell. Syst. 5 (2015) 19:1 19:19.

2. Kernix Lab: Image classification with a pre-trained deep neural network. https://www.kernix.com/blog/image-classification-with-a-pre-trained-deep-neural-network_p11 [navštívené 25.8.2017]. (2016).
3. Ricci, F., Rokach, L., and Shapira, B.: Recommender Systems Handbook. Springer US, New York, NY, USA, 2nd edition, (2015).
4. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. In: Proc. of The International Conference on Learning Representations (2015).
5. Szegedy, Ch., et al.: Rethinking the Inception Architecture for Computer Vision. CoRR (2015).
6. Tkalcic, M., Tasic, J. F.: Colour spaces: perceptual, historical and applicational background. In: Proc. of The IEEE Region 8 EUROCON 2003. Computer as a Tool., (2003).
7. Valdez, P., Mehrabian, A.: Effects of color on emotions. Journal of experimental psychology: General (1994), 394.
8. Yu, L., Fangjian, H., Shaobing, H., Yiwen, L.: A content-based goods image recommendation system. Multimedia Tools and Applications (2017) 1 15.
9. Zhou, Z., Xu, Y., Zhou, J., Zhang, L.: Interactive Image Search for Clothing Recommendation. In: Proceedings of the 2016 ACM on Multimedia Conference, ACM, New York, NY, USA (2016), 754 756.

Pod'akovanie: Tento článok vznikol vďaka čiastočnej podpore projektov *Informačné správanie sa človeka v digitálnom priestore* Agentúry na podporu výskumu a rozvoja, projekt č. APVV-15-0508 a *Inovatívne metódy výučby informatiky vo veľkých skupinách s podporou online vzdelávania* Kultúrnej a edukačnej grantovej agentúry, grant č. KEGA 028STU-4/2017.

Annotation:

Personalized Recommendation Enhanced by Visual Stimuli

User behavior analysis on the Web is still an open research problem. Users interact with Web pages in a various ways and researchers are trying to correctly interpret this behavior. User interactions are an input for the recommender systems that are employed as one of the methods intended to enhance user experience on the Web. Content-based recommendation uses item properties to generate personalized recommendations. Images have become a popular source of information in many domains (e.g., online shopping) where they influence user decision process. In our work, we analyze visual features extracted from images. Our goal is to study how images can be incorporated into the recommendation process.