

Crowdsourcing posilnený sledovaním pohľadu

Jakub Šimko¹

¹Ústav Informatiky a Softvérového Inžinierstva,
Fakulta informatiky a informačných technológií,
Slovenská technická univerzita v Bratislave,
Ilkovičova 2, 842 16, Bratislava
jakub.simko@stuba.sk

Abstrakt. V tomto článku ponúkame víziu prepojenia a synergie dvoch výskumných (ale aj praktických) oblastí a to čerpania z davu (crowdsourcing) a sledovania pohľadu (gaze tracking). Zatiaľ čo praktiky čerpania z davu sú dnes široko nasadzované, sledovanie pohľadu bolo až doteraz významne kvantitatívne obmedzené cenou zariadení, ktoré toto sledovanie umožňovali. To sa však pomaly mení a otvára nový zdroj implicitnej spätnej väzby cennej pre úlohy čerpania z davu. V článku uvádzame teoretickú argumentáciu podporujúcu toto spojenie, ktoré zároveň demonštrujeme prieskumnou štúdiou úlohy tvorby experimentálnej vzorky pre zisťovanie významu homoným.

Kľúčová slova: čerpanie z davu, sledovanie pohľadu, spracovanie textu, rozlišovanie významu slov

1 Motivácia: viac implicitnej spätnej väzby do čerpania z davu

Ľudstvo v dnešnej dobe tvorí a uchováva enormné množstvo informácií. Prominentné postavenie v tomto ohľade má Web ako najväčší „korpus“ informačných zdrojov a ďalších služieb. Na množstvo dostupných informácií nadväzuje potreba ľudí sa v týchto informáciách rozumne orientovať, vyhľadávať a vo všeobecnosti napĺňať svoje potreby. K efektívnemu spracovávaniu veľkého množstva informácií neodmysliteľne patria automatické metódy, ktoré ako jediné dokážu obsah Webu dostatočne rýchlo sprostredkovať. Tradičné sú však aj problémy, na ktoré tieto metódy narážajú. Vyplývajú z povahy informačných priestorov (Webu), ktorých obsah je štruktúrou heterogénny a ťažko strojovo spracovateľný, predovšetkým z dôvodu prítomnosti prirodzeného jazyka a multimediálneho obsahu.

Vo výskume aj v aplikačnej praxi bolo vyvinutých množstvo automatických metód, schopných prekonať aj tieto prekážky. Ich fungovanie je však často obmedzené a to z dôvodu nedostatku vhodne štruktúrovaných informácií či znalostí, ktoré pre spracovanie „divokého“ Webu tieto metódy potrebujú. Či už ide o pravidlá spracovania obsahu (napr. pri sémantickej analýze viet v prirodzenom jazyku), alebo napríklad o tréningové vzorky pre algoritmy strojového učenia (napr. vzorky sémanticky anotovaných obrázkov pre tréningovanie počítačového videnia), veľmi často je na ich vytvorenie potrebná ľudská sila, ktorá je však drahá a nie vždy dostupná.

Celá plejáda prístupov sa preto zaoberá tým, ako čo najefektívnejšie zapojiť ľudí do tvorby explicitne vyjadrených informácií a znalostí na podporu metód automatického spracovania ťažko spracovateľných informačných zdrojov. Ide

o takzvané *ľudské počítanie* (angl. *human computation*), teda proces riešenia výpočtových úloh, ktoré sú len ťažko riešiteľné strojovo. Rozlišujeme dve skupiny prístupov [15]: (1) využitie expertov (teda malej skupiny vysokokvalifikovaných pracovníkov) a (2) *čerpanie z davu* (angl. *crowdsourcing*, čiže využitie širšej skupiny ľudí s neistou mierou kvalifikácie).

Využitie expertov je pri špecifických úlohách nezastupiteľné. Ich cena je však príliš vysoká a dostupnosť príliš nízka pre vytváranie rozsiahlych tréningových vzoriek [13]. Pre tieto je typické skôr využívanie väčších skupín laických ľudí pri čerpaní z davu, pri ktorom síce individuálne schopnosti prispievateľov nie sú také vysoké ako pri expertoch, no ako také kvalitatívne postačujú a sú dostupné v oveľa väčšom rozsahu. Pri automatickom spracovaní informačných zdrojov sú totiž často potrebné aj znalostné bázy obsahujúce „bežné“ poznatky (napr. že určitý obrázok obsahuje slnko, že automobil sa zvyčajne nezaobíde bez kolies). A práve takéto informácie, v primeranej kvalite, je schopná poskytnúť aj širšia skupina ľudí, vhodne organizovaná systémom čerpania z davu (napr. *Amazon Mechanical Turk*).

Prístupy čerpania z davu sú dnes značne rozšírené vo výskume aj v aplikačnej praxi, majú však aj rezervy čo sa týka ich efektívnosti. Časť kvantitatívneho potenciálu davu (množiny prispievateľov zapojených do riešenia nejakej úlohy) musí byť často obetovaná na redundantné riešenie úloh (kontrolou zhody riešení tej istej inštancie úlohy viacerými riešiteľmi v dave), alebo na iné procesy kontroly na zabezpečenie kvality výstupov davu (prevencia chýb a spamu).

Dôvodom okrem iného je, že priebeh riešenia úloh a motivácia riešiteľov sú málo transparentné, nakoľko informačné systémy, zabezpečujúce celý proces, nedokážu zachytávať a vyhodnocovať potrebnú (implicitnú) *spätnú väzbu* svojich prispievateľov a na jej základe identifikovať neštandardné a toxické správanie. Okrem toho, adekvátna spätná väzba môže navyše priniesť aj zvýšenie prínosu už tak správnych riešení úloh. Možno napr. lepšie rozlíšiť, ktorý prispievateľ má na riešenie úlohy lepšie predpoklady [12] alebo či sa niektorí prispievatelia nesnažia o zneužitie systému [4]. Štandardne (bez spätnej väzby) totiž systémy čerpania z davu chápu svojich prispievateľov ako čierne skrinky: zadajú im úlohu a na konci riešenia očakávajú jej (správne) riešenie. Vďaka vedomosti o postupe riešenia, však pri špecifických úlohách vedú vyťažiť viac užitočných informácií [12].

Unikátnym zdrojom spätnej väzby pri interakcii človeka s počítačom (a teda aj pri riešení úloh ľudskeho počítania, resp. čerpania z davu), je *sledovanie pohľadu* (angl. *gaze tracking*). Abstrahujúc od technologickej realizácie, ide o *zaznamenávanie miest (na obrazovke), na ktoré sa človek pri interakcii s počítačom pozerá v konkrétnych časových okamihoch* [2]. Následne sú tieto informácie interpretované (čo už závisí od konkrétneho účelu), či už manuálne (typicky ľudským examinátorom počas vyhodnocovania experimentov [3]) alebo automaticky (napr. pri kvantitatívnom vyhodnocovaní záujmu používateľov o určité časti používateľského rozhrania). Dôležité je, že zaznamenaná sekvencia pohľadov v sebe implicitne nesie množstvo informácií o pozornosti, záujmoch či mentálnom stave používateľa [5], ktoré sa zároveň len veľmi ťažko dajú získať z iných „tradičných“ zdrojov implicitnej spätnej väzby (napr. sledovania pohybu myši, rolovania, záznamov webového servera).

Ako zdroj implicitnej spätnej väzby sa sledovanie pohľadu prirodzene ponúka aj do služieb čerpania z davu, no v tejto oblasti sa zatiaľ neuplatnilo, a to najmä z dôvodu vysokej ceny a nepatrného rozšírenia technológií sledovania pohľadu, čo je

v príkrom kontraste s princípom čerpania z davu, založenom na početných skupinách ľudí vybavených bežnými výpočtovými prostriedkami. Dostupnosť technológií sledovania pohľadu sa však postupne zvyšuje. V horizonte niekoľkých rokov môžeme byť konfrontovaní s ich rozšírením medzi bežných používateľov Webu (a teda potenciálnych participantov v čerpaní z davu). *Je preto namieste pripraviť sa na tento nástup už teraz a skúmať možnosti využitia sledovania pohľadu pri čerpaní z davu.*

Cieľom tohto článku je identifikovať východiská a možné smery výskumu metód spájajúcich čerpanie z davu a sledovanie pohľadu. Rozdelený je na dve časti. Prvá sa venuje analýze oblastí čerpania z davu a sledovania pohľadu a možností ich fúzie. Zameriava sa pritom na domény súvisiace so spracovaním informácií na Webe. Ako príklad možnej aplikácie sledovania pohľadu v typickej úlohe čerpania z davu, uvádzame v druhej časti článku prieskumnú štúdiu zaoberajúcu sa konkrétnou úlohou získavania vzorky pre určovania významu viacznačných pojmov textoch.

2 Smerom k synergii ľudského počítania a sledovania pohľadu

2.1 Čerpanie z davu

Čerpanie z davu nachádza uplatnenie pri veľkom spektre úloh, pri ktorých je nutné ľudské počítanie. Môže ísť napr. o vytváranie znalostných báz, prípravu tréningových vzoriek pre počítačové videnie, automatickú analýzu textu ale aj podporu občianskeho aktivizmu či mapovanie geografických lokalít [8]. Zároveň nadobúda mnoho foriem a existujúce prístupy možno klasifikovať z mnohých hľadísk [10], najmä podľa zloženia davu, spôsobu motivovania a verbovania prispievateľov, dekompozície a distribúcie úloh, zabezpečenia kvality výstupov či predchádzania zneužívaniu. Typickým scenárom čerpania z davu je využitie generickej platformy ako *Amazon Mechanical Turk* s využitím mikroplatieb ako spôsobu motivácie potenciálnych prispievateľov regrútovaných z bližšie neurčených používateľov služby.

Problémom generických platforiem v takomto nastavení je však už spomenutý nedostatok spätnej väzby od prispievateľov. Prispievatelia síce dodávajú svoje riešenia zadaných úloh (pričom ich kvalita sa dá overiť redundantným riešením a kontrolou zhody výstupov), no priebeh samotného riešenia je netransparentný, a sledovať a interpretovať sa dá len v obmedzenej miere. Možno napríklad zaznamenávať vykonané akcie používateľa (kliky, rolovanie, stlačenia kláves) v rozhraní riešenia úlohy či merať čas medzi nimi. Na správny odhad kognitívnych procesov prispievateľov to však málokedy stačí. Vieme napríklad odsledovať, či sa prispievateľ „zasekol“ pri určitej otázke v dotazníku (pred jej vyplnením ubehol od poslednej akcie dlhý čas), nevieme však povedať, či to bolo z dôvodu únavy, subjektívnej náročnosti, alebo si prispievateľ len dal prestávku.

Viacere štúdie nám ukazujú, ako môže viac implicitnej spätnej väzby od prispievateľov pomôcť k lepšej efektívnosti procesu čerpania z davu. Chen a kol. vo svojej štúdií [4] napríklad detegovali spameroch v systémoch komunitného odpovedania otázok (*community question answering – CQA*). Viacerí používatelia týchto systémov sa snažia vo svojich odpovediach podsúvať odporúčania na špecifické komerčné produkty často v rozpore so záujmami komunity. Autori ich

detekciu (a tým zvyšovanie kredibility komunity) opierajú o sledovanie špecifického správania takýchto používateľov: používanie istých vetných konštrukcií, opakované príspevky či umelé dvíhanie počtu hlasov pri odpovediach.

V inej štúdií [12] Rzeszotarski a Kittur predstavili všeobecnú metódu sledovania priebehu riešenia úloh na úrovni používateľských akcií nízkej úrovne (kliky, pohyby myši, rolovanie, stláčanie kláves). Autori natrénovali modely, ktoré predpovedali kvalitu riešení zadaných úloh (ako presnosť, zrozumiteľnosť). Svoje riešenie pritom overovali na typických úlohách čerpania z davu: značkovaní obrázkov, klasifikácie slovných druhov a určovaní zrozumiteľnosti textov.

Celkovo je teda jedným z východísk našej argumentácie to, že *sledovanie prispievateľov počas riešenia úloh prináša zefektívňovanie čerpania z davu*, no zároveň aj konštatovanie, že *takáto spätná väzba v súčasnosti využívaná len slabo*.

2.2 Sledovanie pohľadu

Zaznamenávanie pohybu očí (ang. eye-tracking) resp. sledovanie (zaznamenávanie) pohľadu na základe tohto pohybu (*gaze-tracking*) nie je novou disciplínou. Jej korene siahajú až do 19. storočia pričom automatizované prístupy sa začali objavovať už v 40. rokoch 20. storočia. Z technologického pohľadu ide o rôznorodú doménu zahŕňajúcu prístupy založené na vizuálnej analýze obrazu tváre (očí) pozorovateľa, mechanickom meraní pohybov očí (využívajúcich napríklad špeciálne kontaktné šošovky s vyčnievajúcimi tyčinkami indikujúcimi smer pohľadu) či meraní elektrického potenciálu generovaného pohybom očí (s použitím špeciálnych kontaktných šošoviek a elektród pripojených k telu pozorovateľa) [2].

V minulosti bolo sledovanie pohľadu využívané predovšetkým v laboratórnych podmienkach v experimentoch s malým rozsahom. Dôvodom bola okrem ceny zariadení, ich slabá prenosnosť, obmedzenia pre použitie v určitom prostredí či náročnosť použitia. Dnes je najrozšírenejším spôsobom zaznamenávania pohybu očí analýza odrazov infračervených lúčov od oka [2]. Ide o pomerne neinvazívnu, z pohľadu používateľa ľahko použiteľnú a zároveň dostatočne presnú metódu, ktorú dnes využíva väčšina komerčne dostupných technológií sledovania pohľadu. Tomuto prístupu sa zároveň podarilo nastaviť trend zvyšovania svojej dostupnosti². Okrem toho sa zlepšuje aj softvér obsluhujúci tieto zariadenia – analytické nástroje a programové rozhrania umožňujú čoraz jednoduchší a rýchlejší vývoj aplikácií využívajúcich sledovanie pohľadu. Viacerí výrobcovia už dokonca avizovali úmysel dostať svoje zariadenia medzi širšiu skupinu používateľov, napríklad ako súčasť osobných výpočtových zariadení. Celkovo teda možno v blízkej budúcnosti predpokladať rozšírenie technológií sledovania pohľadu, s novými možnosťami ich aplikácií.

Tradičnými oblasťami využitia sledovania pohľadu sa stali najmä marketing a návrh rozhraní človeka s počítačom. Pre tieto oblasti je významná potreba spätnej väzby konzumentov resp. používateľov (napr. pre zisťovanie efektívnosti pôsobenia prvkov reklamy či odhaľovanie problematických scenárov v pri používaní rozhraní). Využívajú sa pritom najmä kvalitatívne štúdie (napr. [3]) s menším počtom

² <http://www.economist.com/news/technology-quarterly/21567195-computer-interfaces-ability-determine-location-persons-gaze>

účastníkov (reflektujúc práve kvantitatívne obmedzenia technológií sledovania pohľadu spomenuté vyššie), pričom sledovanie pohľadu je len jedným zo zdrojov spätnej väzby participantov popri rečovej interakcii s experimentátormi, zadávaní textu, analýze rozhodnutí a podobne. Nevyhnutná je pritom manuálna evaluácia (dostatočne skúseným) experimentátorom.

Aj manuálne vyhodnocovanie v experimentoch malého rozsahu však vyžaduje automatické spracovanie veľkého množstva údajov o pohľade a ich vhodnú (agregovanú) vizualizáciu. Typickými technikami sú teplotné mapy (*heat maps*), vyjadrujúce záujem používateľov o určité súčasti skúmaného predmetu/rozhrania či grafy pohľadu (*gaze plots*), zobrazujúce postupnosti *fixácií* (momentov, kedy používateľ venoval bez pohnutia oka pozornosť istej oblasti skúmaného predmetu) a *sakád* (prechodov medzi fixáciami). Za pomoci vizualizácií potom správanie participantov interpretuje examinátor. Keďže nejde vždy o jednoduchú úlohu, časť výskumníkov sa venuje tvorbe špecializovaných vizualizácií napr. pre odhaľovanie problematických vzorov správania používateľov vo výučbových systémoch [11], či na webových stránkach [9].

V budúcom využití sledovania pohľadu pri čerpaní z davu však musíme stavať na automatických – teda menej tradičných prístupoch vyhodnocovania údajov o pohľade, ktorým sa v minulosti tiež venovali viacerí výskumníci. Žiaden z existujúcich prístupov nebol navrhovaný ako úloha čerpania z davu, mnohé prístupy však k nemu majú v rôznych aspektoch blízko. Spektrum (problémových domén) týchto metód je široké, prirodzene však často smerujú k používateľsko-centrickým (*user-centric*) aplikáciám snažiacim sa odhadovať zámer, záujmy či kontext používateľa a následne túto informáciu využiť v jeho prospech (napr. personalizáciou), alebo prospech vlastníka aplikácie (napr. marketingovým využitím).

Významná pozornosť je venovaná čítaniu textu, predovšetkým otázke, ako čo najpresnejšie zachytiť priebeh čítania textu človekom v čase. Následne sa s touto informáciou ďalej pracuje podľa potreby aplikácie. Keďže základným výstupom zariadenia na sledovanie pohľadu je v princípe časová postupnosť dvojrozmerných súradníc odhadnutých bodov pohľadu na obrazovke, je potrebné túto transformovať na tvar zmyslupnejší z hľadiska analýzy priebehu čítania textu: napríklad postupnosť dĺžok *fixácií* na jednotlivé slová. To však nie je triviálna úloha, keďže výstupy sledovania pohľadu nemusia byť dostatočne presné či správne kalibrované. Týmto problémom sa vo svojej práci zaoberal napr. Pascual Martinez-Gomez [5], ktorý si na pomoc zobral informácie o samotnom obsahu textu (pozíciach stredov slov, ktorým by mali zodpovedať *fixácie*, dĺžkach slov, ktorým by mali zodpovedať dĺžky *fixácií*).

Ak už máme informáciu o priebehu čítania textu, možno ju ďalej využiť v kvantitatívnom spracovaní. Martinez-Gomez a kol. sa vo svojej ďalšej práci [6] venovali identifikácii častí textov z rôznych dôvodov náročných na čítanie (identifikovali miesta ktoré museli čitatelia čítať opakovane, či na ktorých sa často „zasekli“). V inej práci [18] sa zasa Xu a kol. zaoberali tvorbou sumarizácií textov. Na základe informácií o čítaní textu vypočítali agregovanú mieru „intenzity“ pohľadu na každé slovo dokumentu a prešírili túto intenzitu na ostatné slová prostredníctvom sémantickej podobnosti. Následne intenzity agregovali pre vety dokumentu, zoradili ich a z najvyššie umiestnených vytvorili sumarizáciu.

Vhodnou doménou pre uplatnenie sledovania pohľadu je Web. Buscher a kol. vo svojej práci [1] využívajú sledovanie pohľadu na získavanie informácií o

krátkodobom kontexte (záujmoch) používateľa pre rozširovanie vyhľadávacích dopytov. Ich metóda vyberá z „výsekov“ (*snippets*) vo výsledkoch vyhľadávania kľúčové slová pomocou extrakcie metrikou TF-IDF, avšak len z tých textov, ktoré si používateľ prezrel (pričom berie do úvahy aj čas fixácií ako aj opakované čítanie).

Predmetom sledovania pohľadu môže byť tiež získavanie informácií o dlhodobých záujmoch používateľa Webu, typicky využívaných pri odporúčaní. Napr. Xu a kol. vo svojej práci [17] sledovali používateľov pri interakcii s Webom nejaký čas. Zároveň sledovali ich pohľad pri sledovaní videa, textov či reklamných plôch a odvodzovali z toho záujmy používateľov, ktorým na základe toho ďalej odporúčali.

Okrem zisťovania záujmov používateľov, predstavuje sledovanie pohľadu aj možnosť zisťovať priamo mentálny stav používateľa (ako uvádza McDonald „kognitívny stav človeka sa premieta aj do pohybu jeho očí a pohľadu“ [7]). Praktický význam to môže mať napr. v technológiách podporených vzdelávaní vo využití empatického softvérového agenta, schopného detegovať a v reálnom čase reagovať na problémové stavy sledovaného študenta vo vzdelávacom systéme [16].

2.3 Fúzia do novej disciplíny

Množstvo aplikácií sledovania pohľadu ako zdroja spätnej väzby priam nabáda na jeho využitie aj v scenároch čerpania z davu. Mnohé tieto aplikácie sú až nápadne podobné typickým úlohám ľudského počítania. Napr. spomínaná sumarizácia textov sa často typicky musí vykonávať ručne pri vytváraní tréningových vzoriek pre algoritmy strojového učenia, ktoré túto sumarizáciu majú vykonávať. Pomocou sledovania pohľadu z nej však možno vytážiť informáciu o vhodných sumarizačných vetách rýchlejšie a pohodlnejšie (v porovnaní s explicitným zadávaním týchto viet človekom). Podobná je situácia pri ďalších typických úlohách ako značkovanie multimédií. Napr. pri tradičnej úlohe značkovania obrázkov konceptami by sme vedeli navyše vytážiť aj presné umiestnenie týchto konceptov na obrázkoch na základe údajov o pohľade. Mnoho ďalších úloh, pri ktorých sa dnes údaje o pohľade používateľov vyhodnocujú manuálne, napr. diagnostika formy či obsahu rozhraní (webových, výučbových, herných) či analýza marketingových materiálov, si možno predstaviť s väčším množstvom používateľov a s automatickým vyhodnocovaním zaznamenaných údajov. Optimistický pohľad teda vyzerá tak, že nasadeniu sledovania pohľadu do čerpania z davu chýbajú len dostupnejšie technológie.

Nie je to však celkom pravda – na takýto stav by sme totiž mali byť pripravení ako výskumníci a aj ako tvorcovia praktických aplikácií. Už teraz by sme mali rozvíjať metódy čerpania z davu s využitím sledovania pohľadu. Dnes, prirodzene, neexistuje adekvátna samostatná výskumná disciplína na konceptualizáciu, opis či návrh takýchto metód. Jej formovanie bude mať zmysel až keď naberieme dostatok skúseností s konkrétnymi prvými aplikáciami. Už teraz však možno vytýčiť možné smery výskumu v tejto oblasti a uvažovať niektoré jeho hľadiská:

Spôsob návrhu úloh. Hoci nie je vylúčené, že samotná technológia sledovania pohľadu podnieti vznik celkom nových metód, je prirodzené, že inovácie vyvstanú skôr modifikovaním existujúcich metód čerpania z davu. *Každý návrhár úloh pre čerpanie z davu by sa mal zamyslieť, ako by informácia o pohľade jeho prispievateľov mohla pomôcť k skvalitneniu výstupu celého procesu.*

Aplikačná doména. Otvorené sú zatiaľ všetky možnosti doterajšieho pôsobenia čerpania z davu, pri ktorých nie sú technické obmedzenia využitia zariadení sledovania pohľadu. Špeciálnu pozornosť by sme však mali venovať tým úlohám, pri ktorých sú prispievatelia nútení robiť nejaký typ rozhodovania na základe vizuálnych vnemov (rozpoznávanie obrazu, orientácia na webe či vo výsledkoch vyhľadávania) alebo trávajú dlhší čas analýzou nejakej vizuálnej informácie (čítanie textu).

Špecifickosť prístupu. Mnoho budúcich prístupov bude vyvinutých len pre špecifické úlohy a väčšina bude ťažko využiteľné inde. Určite emergujú aj niektoré širšie využiteľné princípy (ako v každej disciplíne), ktoré však teraz predpovedať nevieme. Od začiatku sa však výskumníci môžu zamerať aj na niekoľko skutočne univerzálne využiteľných princíпов. V čerpaní z davu je napríklad často vhodné určovať, či bol príslušný prispievateľ počas riešenia úlohy v stave nervozity, diskomfortu, sústredenia či uvoľnenosti keďže tieto stavy môžu do značnej miery ovplyvňovať výslednú kvalitu prispievateľovho riešenia. No a práve údaje o pohľade resp. pohybe očí túto informáciu nesú a budú k dispozícii [7].

Diagnostický vs. ovládací systém. Podľa Hansena a Jiho [2] možno systémy využívajúce sledovanie pohľadu rozdeliť do dvoch skupín: (1) diagnostické (také pri ktorých sledovanie pohľadu slúži len na zber informácií, ktoré spätne správanie aplikácie neovplyvňujú, a ak, tak len nepriamo) a (2) ovládacie (pri ktorých používateľ pohľadom aplikáciu priamo riadi). V kontexte ich využitia pri čerpaní z davu možno v tejto súvislosti uvažovať nad vylepšeniami ovládania používateľských rozhraní prispievateľov (napríklad adaptívnym ovládaním rýchlosti prehrávania sekvencie obrázkov podľa toho, ako ich prispievateľ stíha pohľadom spracúvať), skôr však možno očakávať zavádzanie diagnostických aspektov.

V každom prípade bude potrebné rozpracovať a (zatiaľ v laboratórnych podmienkach) overiť prvé koncepty metód čerpania z davu s podporou sledovania pohľadu. V našom prípade sa chceme v budúcnosti zamerať najmä na scenáre užitočné pri inteligentnom spracovaní informácií na Webe.

3 Prípadová štúdia: rozlišovanie významu slov

S cieľom demonštrovať potenciál využitia sledovania pohľadu pri čerpaní z davu sme sa rozhodli uskutočniť malú prieskumnú kvalitatívnu štúdiu v doméne rozlišovania významu homoným, ako tradičného problému riešeného v automatickom spracovaní prirodzeného jazyka [13, 14]. Dnešné metódy realizujúce disambiguáciu sú väčšinou založené na strojovom učení. Vstupom pre ne býva okrem slova, ktorého význam treba rozlíšiť, vektor slov či text, v kontexte ktorého sa slovo nachádza. Výstupom je následne určenie významovej kategórie. Pri tvorbe tréningových vzoriek je toto určenie úlohou pre človeka v procese čerpania z davu [14].

Cieľom nášho experimentu bolo ukázať, že *sledovaním pohľadu ľudí počas čítania kontextového textu pri rozhodovaní o význame určitého slova odhalíme pridané informácie, ktoré umožnia tréningovú vzorku vytvoriť lepšie*. Apriori sme postulovali, že o význame slova budú v kontexte rozhodovať určité slová „s vysokou rozlišovacíou schopnosťou“ (teda také ktoré sa s predmetným slovom sémanticky spájajú) a stanovili hypotézu, že práve na ne spočinie pohľad prispievateľa ako

posledný pred tým, než vysloví svoje rozhodnutie o priradení významu. V tréningovom vektore by tým pádom takéto slová mohli dostať vyššiu váhu. Okrem toho sme predpokladali, že a posteriori manuálnou analýzou zaznamenaného pohľadu odhalíme ďalšie využiteľné vzorce správania.

Experimentu sa zúčastnilo 5 účastníkov. Každý dostal rovnakú úlohu: určiť význam toho istého slova („*president*“) v 10 príkladoch textov (v rozsahu niekoľkých viet) s jeho výskytom. Možné významy boli 3 (šéf firmy, hlava štátu, prezident USA). Všetky príklady boli náhodne prevzaté z datasetu *SemEval2*, používaného práve na overovanie disambiguácie [14]. Zadaná úloha aj opisy možných významov mali účastníci k dispozícii vo forme prezentácie a svoje rozhodnutia hlásili prítomnému experimentátorovi. Ich pohľad bol po celý čas sledovaný zariadením *Tobii X2-30*³.

Spolu účastníci uskutočnili 50 určení významov, pričom individuálne určili význam nesprávne v 5 prípadoch. Tri z týchto chýb pripadli na jeden z desiatich príkladov (ukázal sa byť ako pomerne náročný). Pomocou väčšinového hlasovania bola teda dosiahnutá úspešnosť 90%, čo potvrdzuje experimenty Snowa a kol. ktorí overovali použitie čerpania z davu na prípravu vzoriek pre úlohy disambiguácie [14].

Čo sa týka zapojenia sledovania pohľadu, v 54% prípadov účastníci pred svojím rozhodnutím spočinuli zrakom na takom slove (spojení slov) v texte, ktoré sme (pred experimentom) nezávislo od toho označili za slovo s veľkou rozlišovaciu schopnosťou vzhľadom na určenie významu skúmaného slova. Za moment rozhodnutia sme považovali buď vyslovenie rozhodnutia účastníkom alebo zmenu štýlu čítania, kedy sa podrobné a pomalé čítanie textu menilo na rýchle skenovanie (ktorým si účastníci potvrdzovali svoju voľbu). Považujeme tak našu a priori hypotézu za čiastočne potvrdenú, nakoľko vo ostatných prípadoch sa účastníci rozhodovali prevažne (36%) až po prečítaní celého textu.

A posteriori analýza záznamov o pohľade ukázala niekoľko ďalších skutočností. Účastníci používali rôzne stratégie čítania textov, traja čítali text od začiatku až do konca (a čítanie prípadne prerušili, ak sa už rozhodli na základe nejakého charakteristického slova), dvaja dali prednosť (rýchlejšiemu) skenovaniu slov a viet v bezprostrednom okolí slova, ktorému určovali význam. K čítaniu celého textu sa potom uchýľovali len v „ťažších“ prípadoch, kedy im rýchly prieskum okolia slova nepostačoval. Pre všetkých účastníkov bolo tiež charakteristické opakované čítanie celých vzoriek alebo ich častí v prípadoch neistoty, pričom najmarkantnejšie to bolo práve v jedinom zle určenom prípade (ktorý bol zo všetkých najťažší). Ide síce o pomerne priamočiare, no zároveň v praxi použiteľné zistenie. Odhaľovanie takýchto a podobných vzorov (neistého) správania by sa pri tvorbe tréningovej vzorky davom dalo využiť pri váhovaní hlasov jednotlivých prispievateľov s priaznivým efektom pri vyhodnocovaní sporných prípadov.

4 Budúcnosť (zhrnutie)

Cieľom tohto príspevku bolo poukázať na reálnu možnosť budúcej výhodnej fúzie dvoch výskumných a aplikačných oblastí: čerpania z davu a sledovania pohľadu

³

<http://www.tobii.com/en/eye-tracking-research/global/>

abližšie predstaviť výhody tohto spojenia ako aj niektoré aspekty, cez ktoré sa na budúce prístupy v novej oblasti možno pozerat'. Cieľom bolo tiež vyzvať výkumníkov už teraz v tejto oblasti pôsobiť a overovať svoje koncepty aspoň laboratórne.

Za kľúčový považujeme predovšetkým fakt, že systémy čerpania z davu môžu benefitovať z pridanej implicitnej spätnej väzby, ktorú sú schopné dodávať práve zariadenia na sledovanie pohľadu, ktoré sa stávajú cenovo dostupnejšími. Dobré možnosti využitia sledovania pohľadu vidíme predovšetkým v tých úlohách čerpania z davu, pri ktorých prispievatelia analyzujú vizuálne informácie a rozhodujú sa na ich základe: úlohy rozpoznávania obrazu či čítania a značkovania textu. Ako „najschodnejšiu cestu“ k tvorbe nových metód predpokladáme transformáciu existujúcich prístupov čerpania z davu. Za samostatnú skupinu prístupov hodnú pozornosti tiež považujeme všeobecne použiteľné určovanie mentálneho stavu prispievateľov (nervozity, únavy, sústredenia, atď.) na základe údajov o ich pohľade.

Možno samozrejme očakávať emergenciu ďalších aspektov systémov čerpania z davu podporených sledovaním pohľadu (zoznam považujeme za otvorený). Taktiež čakáme otvorenie ďalších výskumných, technologických a spoločenských otázok (napr. ako budovať infraštruktúru schopnú spracúvať údaje o pohľade v reálnom čase pre dostatočnú adaptívnosť procesu, alebo ako budú ľudia vnímať dlhodobú prácu so sledovaním pohľadu). Ich konceptualizácia je však úlohou pre budúce štúdie.

PodĎakovanie. Táto publikácia vznikla vďaka čiastočnej podpore projektov VG1/0675/11, APVV 0208-10 a podpore MŠVVaŠ SR v rámci OP Výskum a vývoj pre projekt: Univerzitný vedecký park STU Bratislava, ITMS 26240220084, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

Referencie

1. Buscher, G., Dengel, A., van Elst, L. Query expansion using gaze-based feedback on the subdocument level. *Proc. int. c. on Research and development in inf. retrieval (SIGIR '08)*. ACM, New York, 387-394, 2008.
2. Hansen, D. W., Ji, Q. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 3, 478-500. 2010.
3. Hidayah, N., Zain, M., Hanis, F., Razak, A., Jaafar, A., Zulkipli, P. H. Eye tracking in educational games environment: evaluating user interface design through eye tracking patterns. *Proc. of the 2nd int. c. on Visual informatics: sustaining research and innovations*, Springer, Berlin, 64-73, 2011.
4. Chen, C., Wu, K., Srinivasan, V., Bharadwaj, R. K. The best answers? think twice: online detection of commercial campaigns in the CQA forums. *Proc. of the 2013 IEEE/ACM int. c. on Advances in Social Networks Analysis and Mining (ASONAM '13)*. ACM, New York, 458-465, 2013.
5. Martínez-Gómez, P. Quantitative analysis and inference on gaze data using natural language processing techniques. *In Proc. of the 2012 ACM int. c. on Intelligent User Interfaces (IUI '12)*. ACM, New York, 389-392, 2012.
6. Martínez-Gómez, P., Hara, T., Chen, C., Tomita, K., Kano, Y., Aizawa, A. Synthesizing image representations of linguistic and topological features for

- predicting areas of attention. *Proc. of the 12th Pacific Rim int. c. on Trends in Artificial Intel. (PRICAI'12)*, Springer, Berlin, 312-323, 2012.
7. McDonald, S. A., Shillcock, R. C. Eye movements reveal the on-line computation of lexical probabilities. *Psychological Science*. 14, 6, 648-652, 2003.
 8. Michelucci, P. *Handbook of Human Computation*. Springer. 2013.
 9. Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., Newman, J. K. The determinants of web page viewing behavior: an eye-tracking study. *Proc. o. t. '04 s. on Eye tracking res. & app.* ACM, New York, 147-154, 2004.
 10. Quinn, A. J., Bederson B. B. Human computation: a survey and taxonomy of a growing field. *Proc. of the 2011 annual conf. on Human factors in computing systems (CHI '11)*. ACM, New York, 1403-1412, 2011.
 11. Rakoczi, G., Pohl, M. Visualisation and Analysis of Multiuser Gaze Data: Eye Tracking Usability Studies in the Special Context of E-learning. *Proc. of the 2012 IEEE 12th Int. c. on advanced learning technologies (ICALT '12)*. IEEE Computer Society, Washington, DC, 738-739, 2012.
 12. Rzeszotarski, J. M., Kittur, A. Instrumenting the crowd: using implicit behavioral measures to predict task performance. *Proc. of the 24th an. ACM s. on User interface software and tech. (UIST '11)*. ACM, New York, 13-22, 2011.
 13. Sabou, M., Bontcheva, K., Scharl, A. Crowdsourcing research opportunities: lessons from natural language processing. *Proc. of the 12th int. conf. on Knowledge Management and Knowledge Technologies*. ACM, New York, 2012.
 14. Snow, R., O'Connor, B., Jurafsky, D., Ng, A. Y. Cheap and fast--but is it good?: evaluating non-expert annotations for natural language tasks. *Proc. of the c. on Empirical Methods in NLP*. ACL, Stroudsburg, PA, USA, 254-263, 2008.
 15. Šimko, J., Bieliková, M. *Semantic Acquisition Games*, Springer, (2014).
 16. Wang, H., Chignell, M., Ishizuka, M. Empathic tutoring software agents using real-time eye tracking. *Proc. of the 2006 s. on Eye tracking research & applications (ETRA '06)*. ACM, New York, 73-78, 2006.
 17. Xu, S., Jiang, H., Lau, F. C. M. Personalized online document, image and video recommendation via commodity eye-tracking. *Proc. of the 2008 ACM c. on Recommender systems*. ACM, New York, 83-90, 2008.
 18. Xu, S., Jiang, H., Lau, F. C. M. User-oriented document summarization through vision-based eye-tracking. *Proc. of the 14th int. c. on Intelligent user interfaces (IUI '09)*. ACM, New York, 7-16, 2009.

Annotation:

Crowdsourcing reinforced by gaze tracking

In this paper, we propose a vision of synergy of two research (but also applicational) areas of *crowdsourcing* and *gaze tracking*. While the practice of crowdsourcing is nowadays widespread, the gaze tracking has been quantitatively bounded by high prices of technologies. However, this is slowly changing and in gaze tracking we gradually acquire new source of implicit user feedback, valuable in crowdsourcing tasks. We offer theoretical arguments supporting the fusion of crowdsourcing and gaze tracking and for demonstration present a small explorative study of crowdsourced acquisition of training set usable for word sense disambiguation models.