

Smerom k automatickej detekcii problémov s použiteľnosťou prostredníctvom sledovania pohľadu

Martin Svrček a Mária Bieliková

Ústav informatiky, informačných systémov a softvérového inžinierstva,
Fakulta informatiky a informačných technológií, Slovenská technická univerzita v Bratislave
Ilkovičova 2, 842 16 Bratislava

{martin.svrcek, maria.bielikova}@stuba.sk

Abstrakt. Rozličné anomálie používateľského rozhrania aplikácií a webových stránok sú často nazývané problémy použiteľnosti. V súčasnosti existuje v tejto oblasti mnoho štúdií, ktoré sa zaoberajú detekciou problémov s použiteľnosťou softvéru. Avšak, existujúce metódy sa snažia automatizovať tento proces alebo jeho časti použitím výlučne používateľských aktivít vo forme klikov myši alebo vstupu z klávesnice pri kvantitatívnych štúdiách z online správania sa používateľov. Naším cieľom je analyzovať pohľad používateľa s cieľom odhalenia takých pohybov očí, ktoré môžu byť použité pre detekciu problémov s použiteľnosťou. Opisujeme používateľskú štúdiu so sledovaním pohľadu, ktorú sme vykonali. Porovnáваме metriky pohľadu medzi účastníkmi, ktorí pracovali s webovou stránkou bez zavedených problémov s použiteľnosťou a účastníkmi, ktorí pracovali s rovnakou webovou stránkou avšak so zavedenými problémami s použiteľnosťou. Cieľom bolo identifikovať tie metriky pohľadu, ktoré dokážu najlepšie odhaliť tieto problémy. Odhalené metriky môžeme použiť pre ďalšiu analýzu, kde sa chceme pozerat' na pohľad ako na sekvenciu pohybov očí. Takéto sekvencie sú v porovnaní so štandardnými prístupmi schopné odhaliť komplexnejšie vzory v dátach o používaní webu, čo umožní lepšiu identifikáciu problémov s použiteľnosťou.

Kľúčové slová: problémy použiteľnosti, používateľský zážitok, UX, sledovanie pohľadu, sekvencie pohybov očí, metriky pohľadu

1 Úvod

V súčasnosti je používateľský zážitok (angl. user experience, UX) veľmi dôležitou a populárnou oblasťou. Vyjadruje pocity používateľa pri používaní systému a zároveň významne ovplyvňuje samotnú použiteľnosť. V kontexte Webu existuje množstvo prác, ktoré sú súčasťou výskumu v oblasti použiteľnosti [3][7][8][9][11]. Z dôvodu širokého uplatnenia použiteľnosti v mnohých disciplínach neexistuje jedna všeobecná definícia vyhovujúca všetkým týmto oblastiam. Jednou z najvhodnejších definícií je práve ISO definícia použiteľnosti, ktorá hovorí, že použiteľnosť sú vnemy a odozvy osôb, ktoré vznikajú pri používaní produktu, systému alebo služby [11].

*J. Steinberger, M. Zíma, D. Fiala, M. Dostal, M. Nykl (eds.)
Data a znalosti 2017, Plzeň, 5. - 6. října 2017, pp. 209-214.*

Smerom k automatickej detekcii problémov s použiteľnosťou prostredníctvom sledovania pohľadu

Ak však nevieme používať produkt, ak nemôžeme nájsť položky, ktoré hľadáme alebo iba nevieme dosiahnuť naše ciele tak budú tieto vnemy skôr negatívne. V tomto prípade môžeme hovoriť o problémoch s použiteľnosťou, ktoré nám zabraňujú v dokončení našej aktivity alebo robia túto aktivitu náročnejšou a zdĺhavejšou. Samotný problém použiteľnosti je aspekt systému, ktorý spôsobuje, že tento systém je nepríjemný, neefektívny na použitie alebo dokonca úplne zabraňuje jeho použitiu v kontexte dosahovania našich cieľov [10].

Hlavnou úlohou štúdií použiteľnosti webových stránok alebo aplikácií je detegovať tieto problémy. Podobné štúdie zvyčajne vykonávajú UX experti, ktorí odhaľujú tieto problémy ich manuálnym vyhodnocovaním, čo je často pomerne zdĺhavý proces.

Cieľom viacerých prác v tejto oblasti je automatizovať tento proces. V súčasnosti už existuje viacero podobných prístupov avšak tieto používajú iba záznamy v podobe klikov myši alebo vstupu z klávesnice [4][6][5][7][12]. Avšak my veríme, že tento proces detekcie problémov s použiteľnosťou môže byť zlepšený zavedením metrik sledovania pohľadu čo dokazujú aj niektoré práce v tejto oblasti [2].

Naším cieľom je v prvom kroku analyzovať pohľad používateľa a odhaliť charakteristické pohyby očí v kontexte metrik pohľadu, ktoré dokážu odhaľovať konkrétne problémy s použiteľnosťou a zlepšiť tak proces ich identifikácie. Ako súčasť našej práce sme vytvorili zoznam 53 problémov s použiteľnosťou na základe dostupnej literatúry, keďže takto ucelený zoznam nebol dostupný.

2 Automatická identifikácia problémov s použiteľnosťou

Kliknutia myši, skrolovanie alebo vstup z klávesnice dokážu poskytnúť množstvo dôležitých informácií pre identifikáciu anomálií v správaní sa používateľa. Avšak, ak hovoríme o vizuálnej pozornosti, tak neexistuje žiadna lepšia metóda ako sledovanie pohľadu [12]. Z dostupnej literatúry je možné vidieť, že existuje viacero rôznych prístupov k automatickej detekcii problémov s použiteľnosťou využívajúce reverzné inžinierstvo, strom úloh alebo práce vychádzajúce z pachov zdrojového kódu.

V týchto prácach vidíme rozdiel aj v kontexte sledovaných problémov s použiteľnosťou, kedy sa každá z prác zameriava na odlišné problémy. V prvej štúdií [6] sa autori zamerali na automatickú detekciu 16 problémov s použiteľnosťou ako sú napríklad: element nie je samo opisný, zavádzajúci odkaz, stránka bez odozvy, vzdialený obsah.

Ďalšie práce využívali odlišné prístupy pre detekciu. Snažili sa napríklad detegovať problémy s použiteľnosťou na základe pachov zdrojového kódu [1]. Pri ich práci využívali celkovo 6 problémov s použiteľnosťou, ktoré boli viazané na implementačné aspekty systému. Výhodou takýchto prístupov vychádzajúcich priamo zo zdrojového kódu je určite lepšie prepojenie problémov s použiteľnosťou a samotnej implementácie. Toto nám umožňuje a zjednodušuje automatickú refaktorizáciu odhalených problémov.

Niektoré z prác na druhej strane vychádzali z konceptu reverzného inžinierstva [12], kedy bol základom pre detekciu znovu zdrojový kód a rovnako aj rôzne modely pomáhajúce samotnej detekcii. Ďalším veľmi zaujímavým prístupom je využitie

stromov úloh [7]. V rámci tohto prístupu sú jednotlivé akcie používateľov transformované do stromu úloh. V ďalšom kroku je tento strom skenovaný s cieľom odhalenia určitých vzorov prepojených na problémy s použiteľnosťou.

Okrem podobných prístupov využívajúcich klasické akcie používateľov je veľmi zaujímavé aj sledovanie pohľadu [2]. V rámci tohto prístupu autori vychádzajú z predpokladu, že experti svojimi odhadmi poskytujú iba určitý prehľad a nevedia určiť presné prepojenie medzi metrikami a problémami s použiteľnosťou.

Vo vykonanom experimente [2] boli zozbierané dáta z pozorovania ako aj zo sledovania pohľadu od 19 účastníkov a zisťovali sa korelácie medzi týmito dvoma zdrojmi dát. Autori na konci poskytujú sumarizáciu sledovaných problémov s použiteľnosťou a s nimi súvisiacich metrik sledovania pohľadu. Avšak aj na základe ich výstupov je zrejmé, že kombinácia rôznych vzorov pohľadu musí byť skúmaná v kontexte správania sa používateľov viac do hĺbky.

3 Potenciál sledovania pohľadu pre detekciu problémov s použiteľnosťou

V rámci cieľu našej práce bolo potrebné preskúmať potenciál sledovania pohľadu pre detekciu problémov s použiteľnosťou a preto sme sa rozhodli vykonať experiment so sledovaním pohľadu. Experimentu sa zúčastnilo 10 ľudí. Išlo o 8 mužov a 2 ženy. Všetci účastníci boli vysokoškolskí študenti. Pre účely testovania sme využili webovú stránku obsahujúcu novinové články. Na tejto stránke mali účastníci za úlohu vykonať nasledujúce tri úlohy:

- Úloha 1: Registrácia sa do systému
- Úloha 2: Nájdenie špecifického článku
- Úloha 3: Aktualizácia informácií v profile

Počas celého experimentu sme zaznamenávali ich pohľad prostredníctvom zariadenia na sledovanie pohľadu s frekvenciou 300 Hz, ktorý je dostupný vo Výskumnom centre používateľského zážitku a interakcie (<http://uxi.sk>) na Fakulte informatiky a informačných technológií STU v Bratislave.

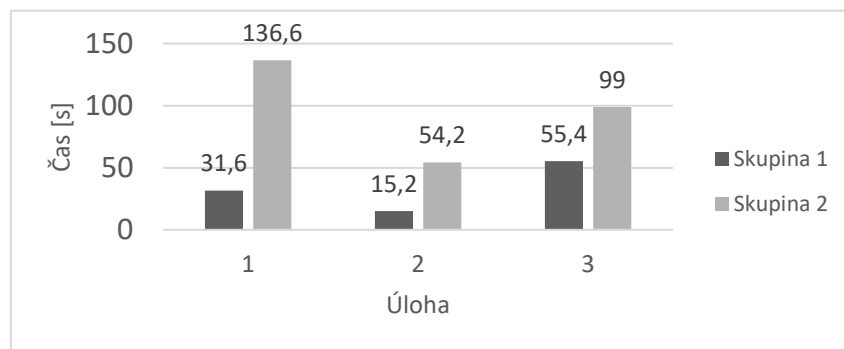
Pre účely porovnania výsledkov vykonania úloh sme sa rozhodli rozdeliť účastníkov experimentu do dvoch skupín. Skupina 1 pracovala s webovou stránkou, ktoré neobsahovala problémy s použiteľnosťou. Skupina 2 pracovala s webovou stránkou, ktoré obsahovala problémy s použiteľnosťou. V rámci nášho experimentu sme vybrali 8 problémov, a tieto sme manuálne zaviedli do našej webovej stránky:

- Dlhý registračný formulár
- Krátke zadávacie pole
- Nedôležité zadávacie pole
- Navigácia obsahujúca príliš veľa položiek
- Mätúce názvy položiek navigácie
- Zlá pôvodná hodnota
- Chýbajúce vysvetlenie chýb

Smerom k automatickej detekcii problémov s použiteľnosťou prostredníctvom sledovania pohľadu

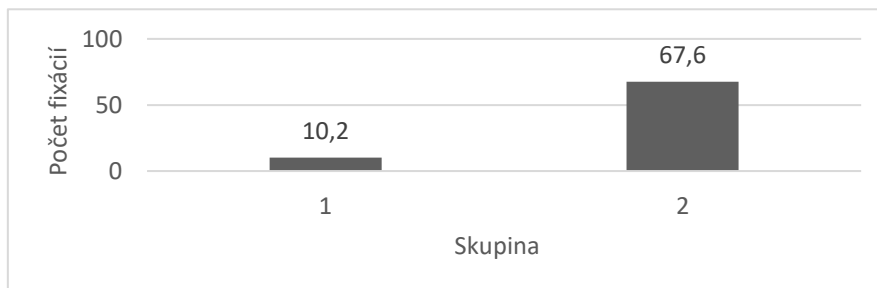
— Chýbajúca klientská validácia

Týchto 8 problémov sme vybrali s ohľadom na možnosť porovnania správania sa používateľov pri odlišných problémoch s použiteľnosťou. Našou hypotézou bolo, že existujú konkrétne metriky pohľadu, ktoré dokážu identifikovať konkrétne problémy s použiteľnosťou. Na základe analýzy získaných dát z experimentov sme odhalili viaceré odlišnosti. Na obrázku 1 je vidieť, že výrazný rozdiel bol pri všetkých úlohách v rámci dĺžky trvania úloh. Pri úlohe 1 bol tento rozdiel najväčší z dôvodu náročného formulára na vyplnenie.



Obrázok 1: Priemerný čas trvania úloh

Rovnaký vzor však môžeme pozorovať aj na metrikách pohľadu. Napríklad pri úlohe 2 (Obrázok 02) sa ukázal veľký rozdiel medzi počtom fixácií pohľadu, kde skupina so zavedenými problémami s použiteľnosťou potrebovala sledovať položky navigácie viackrát (priemerne 10.2 fixácií pre skupinu 1 vs. 67.6 fixácií pre skupinu 2). Avšak metrika dĺžky fixácií sa neukázala ako vhodná pre identifikovanie problému súvisiaceho s navigáciou, keďže rozdiel medzi dĺžkou fixácií bol iba minimálny.



Obrázok 2: Počet fixácií pri úlohe 2

Podobné výsledky sme dosiahli aj pri ostatných problémoch, a preto môžeme usudzovať, že existuje predpoklad, že rôzne problémy s použiteľnosťou je možné odhaliť

odlišnými metrikami. Avšak pre kvalitnejšiu validáciu týchto dosiahnutých výsledkov by sme potrebovali vykonať experiment s väčším počtom účastníkov.

4 Zhodnotenie

Prítomnosť zvolených problémov s použiteľnosťou v našom experimente bola overovaná na základe troch metrik (celková doba fixácie, počet fixácií a čas splnenia úlohy). Experimenty ukázali, že samotná dĺžka fixácií nie je vhodná pre odhaľovanie zvolených problémov. Avšak tieto výsledky sú iba prvými výstupmi v rámci našej práce. V kontexte ďalšej práce sa ukázalo ako odlišné sú jednotlivé problémy, a preto sa chceme zamerať na konkrétnu problematiku (napr. navigácia) a preskúmať iba problémy súvisiace s touto oblasťou.

Z našich pozorovaní ako aj z pozorovaní iných prác [2] je zrejmé, že samotné metricky pohľadu a ich vplyv musí byť skúmaný viac do hĺbky. Rovnako si uvedomujeme, že problémy s použiteľnosťou nie sú prepojené iba na jednu jednoduchú metriku pohľadu ale na špecifickejšie sekvencie týchto vzorov. Tu sa dostávame k sekvenciám pohybu očí (angl. scanpath) a ich rôznym interpretáciám.

V rámci tejto problematiky existuje viacero možných reprezentácií týchto sekvencií (počet fixácií, dĺžka fixácie, sakády, atď.). Práve experiment, ktorý sme vykonali nám môže poskytnúť základné informácie potrebné pre ďalší výskum v kontexte voľby vhodnej reprezentácie pre konkrétny problém s použiteľnosťou.

Sekvencie pohybov očí poskytujú pohľad na dáta zo sledovania pohľadu spolu s kontextom, v ktorom sa vykonali. Preto práve odhaľovanie problémov s použiteľnosťou môže byť výrazne obohatené, keďže tieto problémy môžu byť často súčasťou určitého reťazca aktivít. Veríme, že odhaľovanie týchto reťazcov a vzorov aktivít zlepši presnosť identifikácie problémov s použiteľnosťou softvéru.

Literatúra

1. Almeida, D., Campos, J. C., Saraiva, J., & Silva, J. C. (2015). Towards catalog of usability smells. *ACM Symposium on Applied Computing* (s. 175-181). ACM.
2. Ehmke, C., & Wilson, S. (2007). Identifying web usability problems from eye-tracking data. *People and Computers: HCI... but not as we know it. 1*, s. 119-128. British Computer Society.
3. Falkowska, J., Sobiecki, J., & Pietrzak, M. (2016). Eye tracking Usability Testing enhanced with EEG Analysis. *International Conference of Design, User Experience, and Usability* (s. 399-411). Springer International Publishing.
4. Ganesh, S. G., Sharma, T., & Suryanarayana, G. (2013). Towards a Principle-based Classification of Structural Design Smells. *Journal of Object Technology*, 1-1.
5. Grigera, J., Garrido, A., & Rivero, J. M. (2014). A tool for detecting bad usability smells in an automatic way. *International Conference on Web Engineering*. Springer International Publishing.
6. Grigera, J., Garrido, A., Rivero, J. M., & Rossi, G. (2017). Automatic detection of usability smells in web applications. *International Journal of Human-Computer Studies*, 129-148.

Smerom k automatickej detekcii problémov s použiteľnosťou prostredníctvom sledovania pohľadu

7. Harms, P., & Grabowski, J. (2014). Usage-based automatic detection of usability smells. *Human-Centred Software Engineering* (s. 217-234). Springer Berlin Heidelberg.
8. Hertzum, M. (2016). Usability testin: too early? too much talking? too many problems? *Journal of Usability Studies*, 83-88.
9. Hlaváč, P. (2016). Impact of characteristics of individuals on evaluating the quantitative studies. *UMAP*.
10. Lavery, D., Cockton, G., & Atkinson, M. P. (1997). Comparison of evaluation methods using structured usability problem reports. *Behaviour & Information Technology*, 246-266.
11. Law, E. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. *Human factors in computing systems* (s. 719-728). ACM.
12. Silva, J. C., Campos, J. C., Saraiva, J., & Silva, J. L. (2014). An approach for graphical user interface external bad smells detection. In *New Perspectives in Information Systems and Technologies (Zv. 2, s. 199-205)*. Springer International Publishing.
13. Xu, P., Sugano, Y., & Bulling, A. (2016). Spatio-temporal modeling and prediction of visual attention in graphical user interfaces. *Human Factors in Computing Systems* (s. 3299-3310). ACM.

Pod'akovanie: Tento článok vznikol vďaka čiastočnej podpore projektu APVV-15-0508 a vďaka podpore v rámci projektu „Rozvoj výskumnej infraštruktúry STU“, projekt č. 003STU-2-3/2016 zo zdrojov štátneho rozpočtu prostredníctvom dotácie z Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky.

Annotation:

Towards Automatic Detection of Usability Problems Using Eye-Tracking

Various anomalies of the user interface design of applications and web pages are also called usability problems. There is a number of studies in this field, which deal with detection of so called problems. However, existing methods try to automate this detection process using solely user activity in the form of clicks or keyboard input. Our aim is to analyze gaze of the users to determine the eye movement metrics that can be used in future work to detect usability problems. We performed user study with eye-tracking. First group of participants worked with a web page that did not contain usability problems and the second group worked with the same web page but with usability problems. We compared each gaze metrics between these two groups in order to identify gaze characteristics which can reveal usability problems. These characteristics can be than used as a start point for next analysis, where we want to look at the gaze as a sequence of eye movement. These sequences are called scanpaths and in comparison with standard approach, they can reveal more complex patterns of user behavior in order to identification of usability problems.