

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií
FIIT-5208-72255

Bc. Martin Mokrý

**IDENTIFIKÁCIA ZRUČNOSTI POUŽÍVATEĽA NA WEBE
NA ZÁKLADE VZOROV ZO SLEDOVANIA POHĽADU**

Diplomová práca

Študijný program: Informačné systémy
Študijný odbor: 9.2.6 Informačné systémy
Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového
inžinierstva, FIIT STU v Bratislave
Vedúci práce: Ing. Róbert Móro, PhD.

máj 2018

Zadanie diplomovej práce

Meno študenta: **Bc. Martin Mokrý**

Študijný program: Informačné systémy

Študijný odbor: Informačné systémy

Názov práce: **Identifikácia zručnosti používateľa na webe na základe vzorov zo sledovania pohľadu**

Samostatnou výskumnou a vývojovou činnosťou v rámci predmetov Diplomový projekt I, II, III vypracujte diplomovú prácu na tému, vyjadrenú vyššie uvedeným názvom tak, aby ste dosiahli tieto ciele:

Všeobecný cieľ:

Vypracovaním diplomovej práce preukážte, ako ste si osvojili metódy a postupy riešenia relatívne rozsiahlych projektov, schopnosť samostatne a tvorivo riešiť zložité úlohy aj výskumného charakteru v súlade so súčasnými metódami a postupmi študovaného odboru využívanými v príslušnej oblasti a schopnosť samostatne, tvorivo a kriticky pristupovať k analýze možných riešení a k tvorbe modelov.

Špecifický cieľ:

Vytvorte riešenie zodpovedajúce návrhu textu zadania, ktorý je prílohou tohto zadania. Návrh bližšie opisuje tému vyjadrenú názvom. Tento opis je záväzný, má však rámcový charakter, aby vznikol dostatočný priestor pre Vašu tvorivosť.

Riadte sa pokynmi Vášho vedúceho.

Pokiaľ v priebehu riešenia, opierajúc sa o hlbšie poznanie súčasného stavu v príslušnej oblasti, alebo o priebežné výsledky Vášho riešenia, alebo o iné závažné skutočnosti, dospejete spoločne s Vaším vedúcim k presvedčeniu, že niečo v texte zadania a/alebo v názve by sa malo zmeniť, navrhnete zmenu. Zmena je spravidla možná len pri dosiahnutí kontrolného bodu.

Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU Bratislava

Vedúci práce: **Ing. Róbert Móro**

Termíny odovzdania:

Podľa harmonogramu štúdia platného pre semester, v ktorom máte príslušný predmet (Diplomový projekt I, II, III) absolvovať podľa Vášho študijného plánu

Predmety odovzdania:

V každom predmete dokument podľa pokynov na www.fiit.stuba.sk v časti:
home > Informácie o > štúdiu > organizácia štúdia > diplomový projekt.

V Bratislave dňa 13. 2. 2017

**SLOVENSKÁ TECHNICKÁ UNIVERZITA
V BRATISLAVE**

Fakulta informatiky a informačných technológií
Ilkovičova 2, 842 16 Bratislava 4

1

prof. Ing. Pavol Návrat, PhD.
riaditeľ Ústavu informatiky, informačných systémov
a softvérového inžinierstva

Návrh zadania diplomovej práce

*Finálna verzia do diplomovej práce*¹

Študent:

Meno, priezvisko, tituly: Martin Mokrý, Bc.
Študijný program: Informačné systémy
Kontakt: martin.mokry@gmail.com

Výskumník:

Meno, priezvisko, tituly: Róbert Móro, Ing.

Projekt:

Názov: Identifikácia zručnosti používateľa na webe na základe vzorov zo sledovania pohľadu
Názov v angličtine: Identification of the user skill on the Web based on patterns in eyetracking data
Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU, Bratislava
Oblasť problematiky: identifikácia vzorov, zručnosť používateľa, sledovanie pohľadu, strojové učenie

Text návrhu zadania²

Dôležitým krokom pri spracovaní dát zo sledovania pohľadu je identifikácia opakujúcich sa vzorov (napr. v podobe postupnosti fixovaných oblastí záujmu), ktoré slúžia na vysvetlenie zaznamenatej interakcie, umožňujú porovnávať spôsob interakcie rôznych používateľov, hľadať medzi nimi podobnosti a na základe nich používateľov zhľukovať. Toto je využiteľné jednak pri vyhodnocovaní zaznamenatej interakcie (napr. z testovania použiteľnosti aplikácie), ale aj pri prispôbovaní rozhraní na identifikovanú situáciu (vzor). Identifikovaná situácia, ako napríklad systematické skenovanie stránky, môže vypovedať o neznalosti rozloženia elementov na stránke a na druhej strane vzory objavujúce sa opakovane môžu implikovať každodenné aktivity zručného používateľa.

Analyzujte existujúce prístupy identifikácie spoločných vzorov a metrík fixácií a sákad, ktoré sú charakteristické vzhľadom na zručnosť osoby na webe a prístupy využitia strojového učenia v oblasti dát extrahovaných zo sledovača pohľadu. Navrhnite a implementujte metódu umožňujúcu automatickú identifikáciu zručnosti používateľa na webovom sídle z dát získaných pomocou sledovača pohľadu s využitím metód strojového učenia. Navrhnutú metódu overte pomocou sledovania pohľadu nad zvoleným webovým sídlom.

¹ Vytlačiť obojstranne na jeden list papiera

² 150-200 slov (1200-1700 znakov), ktoré opisujú výskumný problém v kontexte súčasného stavu vrátane motivácie a smerov riešenia

Literatúra³

- Eraslan S., Harper, S., Yesilada Y.: Eye Tracking Scanpath Analysis Techniques on Web Pages: A Survey, Evaluation and Comparison, Journal of Eye Movement Research, 2016, 9(1), pp. 1-19.
- Michael J. C., Chathra H., Nicholas J. B., Chirag S.: Discrimination between tasks with user activity patterns during information search, Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, ACM, 2014, pp. 567-576.

Vyššie je uvedený návrh diplomového projektu, ktorý vypracoval(a) Bc. Martin Mokrý, konzultoval(a) a osvojil(a) si ho Ing. Róbert Móro a súhlasí, že bude takýto projekt viesť v prípade, že bude pridelený tomuto študentovi.

V Bratislave dňa 18.1.2017



Podpis študenta

Podpis výskumníka

Vyjadrenie garanta predmetov Diplomový projekt I, II, III

Návrh zadania schválený: áno / ~~nie~~⁴

Dňa: 13. 2. 2017



Podpis garanta predmetov

³ 2 vedecké zdroje, každý v samostatnej rubrike a s údajmi zodpovedajúcimi bibliografickým odkazom podľa normy STN ISO 690, ktoré sa viažu k téme zadania a preukazujú výskumnú povahu problému a jeho aktuálnosť (uvedte všetky potrebné údaje na identifikáciu zdroja, pričom uprednostnite vedecké príspevky v časopisoch a medzinárodných konferenciách)

⁴ Nehodiace sa prečiarknite

Čestné prehlásenie

Čestne prehlasujem, že táto diplomová práca bola vypracovaná samostatne, na základe konzultácií, uvedenej literatúry a programov.

V Bratislave, dňa 03.05.2018

podpis autora

Pod'akovanie

Veľmi rád by som sa pod'akoval vedúcemu mojej diplomovej práce Ing. Robovi Mórovi, PhD. za jeho rady, pripomienky, odborné vedenie a čas, ktoré mi poskytol pri spracovaní mojej diplomovej práce.

Taktiež sa by som sa chcel pod'akovať za pripomienky a spätnú väzbu skupine PeWe.

ANOTÁCIA

Slovenská technická univerzita v Bratislave

FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLOGIÍ

Študijný program: Informačné systémy

Autor: Bc. Martin Mokry

Diplomový projekt: Identifikácia zručnosti používateľa na webe na základe vzorov zo sledovania pohľadu

Vedúci diplomovej práce: Ing. Róbert Móro, PhD.

máj 2018

Téme rozlišovania nováčikov a expertov sa doteraz venovali výskumy v doménach vizualizácie a profesionálnej expertízy. Avšak určite je relevantná aj na webe, kam prichádzajú každodenne noví používatelia a usilujú sa efektívne dosiahnuť svoje ciele. V našej práci navrhujeme metódu pre automatickú identifikáciu nováčikov na konkrétnej stránke. Za týmto účelom využívame rôzne metriky sledovača pohľadu odvodené od fixácií, sakád, veľkosti zreničky, vzdialenosti hlavy od monitoru a sekvencií pohľadu na natrénovanie modelu strojového učenia.

S cieľom overiť nami navrhnutú metódu a preskúmať užitočnosť navrhnutých metrík sledovača pohľadu sme zrealizovali dva experimenty. Prvý experiment bol testovaný na jednoduchom e-shope a zúčastnilo sa ho 15 participantov. Výsledky vypovedajú o vysokej rozlišujúcej schopnosti metrík odvodených od veľkosti zreničky a vzdialenosti hlavy od monitoru. Druhého experimentu sa zúčastnilo 57 participantov, pričom boli testovaní na viac komplexnom e-shope s pokročilejšími filtrami. Výsledky tohto experimentu potvrdzujú vysokú rozlišovaciu schopnosť metrík odvodených od vzdialenosti hlavy od monitoru a identifikujú relatívne vysokú rozlišovaciu schopnosť metrík odvodených od sekvencií pohľadov v doméne identifikácie zručnosti používateľa na webe.

ANNOTATION

Slovak University of Technology Bratislava

FACULTY OF INFORMATICS AND INFORMATION TECHNOLOGIES

Degree Course: Information Systems

Author: Bc. Martin Mokry

Master's Thesis: Identification of the user skill on the Web based on patterns in eye tracking data

Supervisor: Ing. Róbert Móro, PhD.

2018, May

Distinguishing novices and experts was previously researched in domains of visualizations or professional expertise. However, it is relevant also on the Web, where new users coming to a web site can struggle to use it efficiently to fulfil their goals. In our work, we propose a method of automatic identification of novices on a specific web site. We use eye tracking for this purpose and derive various features based on fixations, saccades, pupil size, head distance and common scanpath, which serve as an input to a machine learning model.

In order to evaluate our proposed method and examine the usefulness of the derived eye tracking features, we carried out two group eyetracking studies. First one was conducted with 15 participants in simple e-shop. The results suggest that the most distinguishing features are the ones derived from pupil size and head distance. The second, more robust experiment, was conducted with 57 participants on a complex e-shop with advanced filters. The results of this experiment confirmed distinguishing ability of head distance features and identified relatively high distinguishing ability of scanpath features in domain of user skill identification on the Web.

Obsah

1	ÚVOD.....	1
2	SLEDOVANIE POHĽADU	3
2.1	Pohyby očí.....	3
2.2	Zaznamenávanie pohybov očí.....	4
3	ZRUČNOSŤ POUŽÍVATEĽA.....	5
3.1	Úrovne učenia	7
3.1.1	Definovanie úrovne zručnosti.....	8
3.2	Vplyv na rozdiely medzi expertmi a začiatočníkmi.....	9
3.3	Diskusia.....	9
4	EXISTUJÚCE METÓDY ANALÝZY POHĽADU	11
4.1	Sekvencia pohľadu	11
4.2	Používané metriky.....	12
4.3	Rekurentná kvantifikačná analýza	14
4.4	Analýza sekvencií pohľadu	17
4.4.1	Algoritmy na výpočet podobnosti sekvencií pohľadu.....	18
4.4.2	Algoritmy na identifikáciu spoločných sekvencií	19
4.4.3	Porovnanie algoritmov na identifikáciu spoločných sekvencií	24
4.5	Diskusia.....	26
5	AUTOMATICKÁ IDENTIFIKÁCIA ZRUČNOSTI POUŽÍVATEĽA NA WEBE.....	27
5.1	Zadefinovanie problému	27
5.2	Navrhované riešenie.....	28
5.3	Diskusia.....	31
6	OVERENIE ROZDIELOV V SEKVENCIÁCH POHĽADU	33
6.1	Návrh prvého experimentu.....	33
6.2	Priebeh experimentu.....	35
6.3	Vyhodnotenie prvého experimentu	35

6.3.1	Podobnosti sekvencií pohľadu.....	36
6.3.2	Podobnosti voči spoločnej sekvencií pohľadu	37
6.4	Tvorba modelov	39
6.4.1	Model zo základných atribútov	39
6.4.2	Model rekurentných metrík	43
6.4.3	Model spoločnej sekvencie pohľadu	44
6.4.4	Kombinované modely.....	46
6.5	Diskusia.....	46
7	OVERENIE ROZLIŠITEĽNOSTI ZRUČNOSTI	49
7.1	Návrh experimentu.....	49
7.2	Realizácia navrhnutého experimentu	51
7.2.1	Pilotné testovanie.....	51
7.2.2	Ostré testovanie	52
7.3	Opis nazbieraných dát	53
7.3.1	Rozdelenie participantov	55
7.4	Výber klasifikačného algoritmu	56
7.4.1	Identifikácia najvýraznejších atribútov	56
7.4.2	Porovnanie klasifikátorov.....	58
7.5	Tvorba modelov	58
7.5.1	Model zo základných atribútov	58
7.5.2	Model rekurentných metrík	59
7.5.3	Model spoločnej sekvencie pohľadu	60
7.5.4	Kombinované modely.....	60
7.5.5	Automatické generovanie a redukcia atribútov	61
7.6	Diskusia.....	61
8	ZÁVER.....	63
9	LITERATÚRA.....	65
	PRÍLOHA A – PROTOKOL PRVÉHO EXPERIMENTU.....	A-1
	PRÍLOHA B – PROTOKOL DRUHÉHO EXPERIMENTU	B-1
	PRÍLOHA C – VYHODNOTENIE PLÁNU PRÁCE POČAS DOBY RIEŠENIA.C-1	
	PRÍLOHA D – TECHNICKÁ DOKUMENTÁCIA	D-1

PRÍLOHA E – POUŽÍVATEĽSKÁ PRÍRUČKA	E-1
PRÍLOHA F – IIT.SRC ČLÁNOK	F-1
PRÍLOHA G – OBSAH PRILOŽENÉHO MÉDIA	G-1

1 Úvod

Používanie webových stránok je v dnešnej dobe každodennou činnosťou pre veľké množstvo ľudí. Táto činnosť môže byť zameraná na vyhľadávanie informácií, zábavu, nakupovanie, administratívnu činnosť a iné. V každom prípade, vždy budú existovať ľudia, ktorí sa učia používať nové rozhranie veľmi pomaly. Aj títo ľudia sú na dosiahnutie svojich cieľov občas nútení navštíviť stránku, ktorú nepoznajú alebo stránku, z ktorej typom (doménou) nie sú naučení pracovať. Takýto typ používateľov zvykne byť zo začiatku dezorientovaný a zmätený, čo nemusí byť nutne spôsobené len zlou použiteľnosťou stránky, ale napríklad aj nedostatkom informácií o tom, akým štýlom procesy na stránke fungujú. V tomto prípade by bolo vhodné tento typ používateľov usmerniť a naviesť k prvým krokom, ktoré zabezpečia, aby frustrovaný používateľ neodišiel ku konkurencii. Avšak, aby túto akciu bolo možné vykonať len v prípade návštevy používateľmi, ktorí naozaj potrebujú, je potrebné tieto dve skupiny používateľov rozlišovať automaticky.

Tento problém je možné definovať ako automatické rozlišovanie začiatočníkov a expertov v prostredí webového sídla. Rozlišovaniu expertov a začiatočníkov bolo venovaných niekoľko prác v oblastiach chirurgie [30], dermatológie [46], pilotovania lietadla [22] a vizualizácie [33], pričom analyzovali pohyby očí participantov pri vyhodnocovaní obrázkov. Tieto štúdie boli motivované snahou identifikovať, v čom robia začiatočníci chybu, a ako ich naviesť na správny spôsob analýzy znázornenej situácie.

V prípade webových sídel a rovnakej motivácie sa k riešeniu tohto problému zväčša pristupuje pomocou analýzy pohybov očí pri čítaní [40], čo je obmedzujúce a neaplikovateľné na všetky typy webových sídel. Preto plánujeme využiť metódy používané na analýzu obrázkov a vizualizácií spomenuté v predchádzajúcom odstavci. Tieto sú viac univerzálne, aj keď ich nedostatkom je, že niektoré zistenia v danej doméne nebudú pravdepodobne platiť v doméne webových sídel.

V našej práci sa podrobnejšie venujeme identifikácii zručnosti používateľa pri používaní webu, pričom naším cieľom je rozlišovať medzi výkonmi expertov a začiatočníkov na konkrétnom webe a doménou daného webu. Podobnú problematiku riešili aj iné práce, avšak väčšina sa sústredila buď na identifikáciu zručnosti používateľa vo vlastnom grafickom rozhraní [17], vo vizualizáciách [15,25,31] alebo pri analýze fotografií zo špecifickej domény [21,28,42], nie však na webe. Jedna práca sa sústredila na doménovú znalosť testovanú na webovej stránke [6], avšak znalosť rozpoznávali analýzou čítania a doménová znalosť pre nich značila, ako dobre človek pozná termíny z medicíny (čiže stránka bol len prostriedok na testovanie doménovej znalosti inej domény).

Nami navrhovaný prístup využíva vzory pohybov očí a rekurentné vlastnosti získané analýzou pohybov očí, ale aj bežne používané metriky fixácií a sakád. Vo väčšine výskumov sa totiž tieto metriky a reprezentácie pohybov očí analyzujú samostatne a my veríme, že ich spoločné použitie môže obsahovať viac informácií a tým byť aj viac rozlišujúce v prípade určovania úrovne zručnosti.

V kapitole 2 je popísané, čo sú pohyby očí, prečo sa merajú, a ako sa merajú. Kapitola 3 definuje pojmy spojené so zručnosťou používateľov, na aké úrovne sa delí, a aké faktory spôsobujú rozličnosť výkonov expertov a začiatočníkov. Kapitola 4 obsahuje prehľad používaných reprezentácií a metrík sledovania pohľadu, ako aj popis algoritmov na analýzu sekvencií pohľadu. Návrh riešenia sa nachádza v kapitole 5 a popis prvého experimentu na jeho overenie v kapitole 6. Návrh druhého, komplexnejšieho experimentu a popis zistení sú zhrnuté kapitole 7.

2 Sledovanie pohľadu

Aktuálne najpoužívanejším spôsobom, ako sa o skupine ľudí dozvedieť istý typ informácií, je požiadať ich o vyplnenie dotazníka. Toto však v prípade návštevy stránky nie je uskutočniteľné, a aj keby bolo, ľudia hodnotia vlastnú úroveň zručnosti subjektívne, kvôli čomu nebýva často hodnotenie zhodné so skutočnosťou. Toto tvrdenie má podporu vo psychologickú literatúru [36], v ktorej sa odporúča pýtať sa participantov na fakty ako vek a pohlavie, a nie na odhady niečoho abstraktného, ako napríklad už spomínaná úroveň zručnosti. Preto sa čoraz viac do popredia dostáva analýza dát získaných sledovaním chovania používateľov v prostredí webu. Typickými zdrojmi takýchto typov dát sú logy zo sledovania pohybu myškou a logy vykonaných akcií (kliknutie, skrolovanie). Tieto sledovania zabezpečujú informácie na úrovni interakcie. Čoraz viac populárnym sa však stáva využitie dát získaných sledovaním pohľadu [7], ktoré reprezentujú kognitívnu úroveň aktivity používateľa. Pohyby očí sú priamo spojené s duševným stavom človeka, čo ich robí veľmi atraktívnymi pre modelovanie správania pri vykonávaní hľadania a na rozdiel od ostatných sledovaní reprezentujú informácie, s ktorými nebolo manipulované či už podvedome alebo zámerne [6].

2.1 Pohyby očí

Pomocou vizuálnych vnemov získavame väčšinu informácií o okolitom prostredí. Tieto vnemy sú získavané prostredníctvom malej plochy nachádzajúcej sa v strede zreničky, ktorá sa nazýva ústredná jamka sietnice (angl. fovea). Táto dokáže vnímať obraz s vysokou ostrosťou len do uhlu 2° od svojho stredu, smerom k obvodu zreničky [10]. Vo vzdialenosti 5° dosahuje vnímaný vizuálny vnem už iba 50%-nú ostrosť oproti vnímanej ostrosti v strede zreničky [31]. Aby bolo možné vnímať väčšiu časť okolia, sú potrebné pohyby očí. Aj keď ide vo väčšine prípadov o podvedomú činnosť, ľudské oko vykoná priemerne 3-5 pohybov za sekundu. Tento druh pohybu očí patrí do jednej z dvoch kategórií posunu pozornosti (angl. attention shift) nazývanej zjavná pozornosť (angl. overt attention), a keď sa hovorí o sledovaní pohľadu, je väčšinou myslená táto kategória. Existuje aj druhý typ posunu pozornosti, ktorý sa nazýva skrytá pozornosť (angl. covert attention), ktorá popisuje posun pozornosti bez pohybu oka a je používaný najmä ako príprava na ďalšiu fixáciu [31].

Anatómia ľudského vizuálneho systému aj napriek obmedzeniu vizuálneho zorného poľa vyvoláva neustálu potrebu sústrediť sa na dôležité objekty a zároveň vnímať vizuálne vnemy okolia. Tieto dva protichodné pudy motivujú naše neustále prieskumné správanie. Komplexný dojem o našom okolí je postupne vyskladaný z viacerých ohnisk záujmu [31]. Ohniská záujmu sú miesta v scéne, na ktoré bol pohľad zafixovaný na určitý časový interval bez pohybu, pričom počas a tesne pred začatím vykonávania pohybov očí vzniká efekt nazývaný perцепčné potlačenie (angl. perceptual suppression), počas ktorého je znížené perцепčné povedomie (angl. perception awareness) a citlivosť na vizuálne vnemy.

Vyhodnocovanie zaznamenanej interakcie so stránkou sa najčastejšie vykonáva pomocou interpretácie zaznamenaných dát vo forme postupnosti sakád a fixácií. Sakády predstavujú krátke pohyby očí, medzi ktorými sa vyskytujú relatívne nehybné momenty – fixácie [10]. Postupnosti fixácií a sakád sa ďalej analyzujú a vytvárajú sa z nich ďalšie odvodené vlastnosti, hľadajú sa spoločné vzory na úrovni fixácií nad jednotlivými slovami (analýza čítania) alebo sa hľadajú spoločné vzory na úrovni vizuálnych elementov. Dĺžka trvania fixácie sa typicky pohybuje v rozmedzí 100 - 500 ms v závislosti od sledovaného materiálu. Pri čítaní je priemerná dĺžka fixácie 250 ms. Typická dĺžka sakády pri vnímaní scény je 40 - 50 ms, zatiaľ čo pri čítaní iba 30 ms [32].

2.2 Zaznamenávanie pohybov očí

Zaznamenávanie pohybov očí po obrazovke sa označuje ako sledovanie pohľadu (angl. eye-tracking) a je najčastejšie vykonávané pomocou zariadenia nazývaného sledovač pohľadu (angl. eye-tracker). Toto zariadenie zaznamenáva pohyby očí používateľa sledujúceho stimuly na obrazovke [9]. Na identifikáciu očí využíva infračervené svetlo, ktoré zabezpečuje, aby sa oči zobrazili na zaznamenanom obraze biele. Zariadenie ďalej vypočítava veľkosť a smer vektorov zo stredu zreničky do zaznamenaných malých odleskov na rohovke. Tieto vektory sú nezávislé od pozície hlavy a považované za vhodné a dostatočne všeobecné ukazovatele toho, kam sa používateľ pozerá. Pozícia hlavy však musí byť situovaná tak, aby bolo vždy možné zaznamenať stred zreničky aj odlesk na rohovke. Ďalším obmedzením je, že v miestnosti by nemal byť ďalší výrazný zdroj svetla, ktorý by mohol spôsobiť ďalší odlesk na rohovke, ktorý mohol sledovač pohľadu zmiestť [19].

Sledovač pohľadu zaznamenáva pohyby očí s rôznou frekvenciou v závislosti od typu pohybov, ktoré chceme rozpoznať. Bežne sa pohyby očí zaznamenávajú v rozmedzí 30-240Hz, čo je úplne postačujúce vzorkovanie na zaznamenávanie správania používateľa alebo na testovanie použiteľnosti. Avšak v prípade, že sa požaduje zachytávanie aj mikro-výkyvov fixácií (angl. micro-fluctuation), je vhodné použiť sledovač pohľadu so vzorkovaním väčším ako 1000 Hz [19]. Sledovanie pohľadu už v dnešnej dobe nepredstavuje príliš veľkú výzvu. Hardvér a softvér, ktorý sa stará o zachytávanie a interpretáciu pohybov očí, je štandardizovaný a funguje spoľahlivo, pričom jeho cena postupne klesá a stáva sa čoraz dostupnejší. Reálnou výzvou zostáva výber vhodnej analýzy a interpretácie nameraných údajov na odpovedanie riešenej úlohy [19]. Záujem o prácu so sledovačom pohľadu začal narastať po zverejnení hypotézy tvrdiacej, že to, kam sa človek pozerá, napovedá o tom, čomu venuje pozornosť, a o čom rozmýšľa [9], ktorá bola formulovaná Justom a Carpenterom [26]. Spojenie medzi rozhodovacím procesom človeka a vzormi pohľadov bolo potvrdené aj inými článkami [9]. V dnešnej dobe sa sledovanie pohľadu používa pri riešení problémov z rôznych domén, ako napr. interakcia človeka s počítačom (angl. human-computer interaction) alebo modelovanie používateľa [27]. Ako príklady sa dajú uviesť články, ktoré sa venovali identifikácii problémov použiteľnosti [9], identifikácii typu úlohy [7] alebo klasifikácii, či človek číta, alebo nie [5].

3 Zručnosť používateľa

Za výkon experta je považovaný konzistentne lepší výkon na sade úloh v danej doméne od výkonu bežného človeka v doméne. Je označovaný aj ako odbornosť (angl. expertise) [16]. Oxfordský slovník definuje zručnosť (angl. skill) ako: “Vedieť niečo spraviť. Schopnosť niečo dosiahnuť s presnosťou a istotou. Schopnosť vykonávať funkciu získanú alebo naučenú tréningom” [17]. Odbornosť preto predstavuje vysokú mieru zručnosti, ktorou sa odlišujú experti od začiatočníkov.

Zručnosť môžeme deliť na základe dimenzií, na ktorých sa zručnosť používateľa môže líšiť. Tieto dimenzie sa môžu líšiť na základe úrovne použitej abstrakcie. Podľa článku [17] existujú zručnosti:

- *doménovo nezávislé* (angl. general) – prenositeľné zručnosti, ktoré v IT doméne umožňujú použitie rovnakých zručností v rámci rôznych aplikácií (napr. kliknutie ľavým alebo pravým tlačidlom myši, použitie kroku späť, použitie skratiek na kopírovanie, uloženie a ďalšie)
- *doménovo závislé* (angl. domain-specific) – neprenositeľné zručnosti, na základe ktorých budeme používateľov rozlišovať (napr. aplikovanie filtra, vyhľadanie noviniek alebo relevantných informácií).

Nielson si zvolil nižšiu úroveň abstrakcie a rozlišuje dimenzie: zručnosť so systémom, všeobecná zručnosť práce na počítači a zručnosť v doméne vykonávanej úlohy [35].

Ľudia spracúvajú vizuálne vnemy dvomi spôsobmi, a to spôsobom oddola nahor (angl. bottom-up) – riadený vizuálnymi vlastnosťami prostredia a spôsobom zhora nadol (angl. top-down) – riadený kognitívnymi procesmi. Bolo identifikované, že na pohyby očí začiatočníkov majú väčší vplyv vizuálne vlastnosti prostredia, zatiaľ čo na pohyby očí expertov vplyv viac ich doménová znalosť [46]. V súvislosti s týmito zisteniami vznikli konkrétnejšie teórie týkajúce sa zopakovateľnosti lepšieho výkonu expertov pri plnení úloh v doméne vnímania vizuálnych vnemov a sú nimi [16]:

- *teória spojená s dlhodobou pracovnou pamäťou* – predpokladá kvalitatívne rozdiely štruktúr v pamäti a tvrdí, že expertnosť rozširuje kapacity spracovania informácií vytváraním získavacích štruktúr, ktoré expertom umožňujú rýchlo zakódovať informácie do dlhodobej pamäte aj k nim efektívne pristupovať.
 - Predpokladaný dôsledok teórie: rozdiely v rýchlosti spracovania informácií by mali mať vplyv na dĺžku trvania fixácií a spôsobiť kratšiu dĺžku trvania fixácií expertov.
- *hypotéza redukcie informácií* – predpokladá filtrovanie informácií pri ich spracovaní. Zaoberá sa tvrdením, že u expertov sú zdroje pozornosti selektívne pridelované informáciám relevantným na splnenie úlohy a nepodstatné informácie sú ignorované.

- Predpokladaný dôsledok teórie: Experti by mali mať viac fixácií a dlhšie fixácie nad regiónom relevantným pre splnenie úlohy a naopak by mali mať menej fixácií a kratšie fixácie na nepodstatné regióny.
- *holistický model vnímania obrazu* – predpokladá, že expertnosť mení organizáciu procesov vnímania periférnych oblastí tak, že umožňuje rýchly a plynulý prechod z fázy globálnej analýzy do dekompozície na menšiu granularitu a ďalej do hierarchickej štruktúry elementov.
 - Predpokladaný dôsledok teórie: experti by mali mať vďaka holistickej analýze dlhšie sakády a kratší čas potrebný na zafixovanie relevantného regiónu.

Všetky z týchto hypotéz boli v článku [16] overované v doméne vizualizácií a vyhodnotené ako pravdivé medzi skupinami používateľov expertov a začiatočníkov, avšak medzi skupinou začiatočníkov a skupinou pokročilých používateľov už neboli výsledky potvrdzujúce, čo ale nie je veľmi prekvapivé, keďže rozdiely v zručnosti medzi danými skupinami sú výrazne menšie. Z tohto dôvodu, by malo byť v našej práci lepšie definovať si len dve úrovne zručnosti a overiť si naše predpoklady na hraničných prípadoch.

Nakoľko sú vyššie uvedené hypotézy pravdivé v doméne webových sídel, overené nie je. Avšak stránky podobne ako vizualizácie obsahujú relevantné aj zbytočné informácie, ktoré by mali experti rozlíšiť rýchlejšie a taktiež obsahujú štruktúry, ktoré by experti mali ľahšie dekomponovať. Z tohto pohľadu predpokladáme, že aspoň niektoré zo spomenutých hypotéz budú platiť aj v doméne webových stránok, ktorej sa venujeme.

Identifikovať metriky, ktoré sú rozlišujúce medzi rôznymi úrovňami zručnosti, sa snažilo viacero výskumov. Niektoré potvrdzujú dôsledky vyššie uvedených teórií, ale iné zistili opak. Príkladom je článok [2], v ktorom identifikovali, že experti majú dlhšie sakády aj trvania fixácií ako začiatočníci, čiže zhodné je iba tvrdenie o dĺžke sakád.

Zručnosť používateľa s webom narastá s narastajúcim počtom interakcií, ktoré na danom webe vykonal, čo priamo úmerne ovplyvňuje aj jeho oboznámenosť (angl. familiarity) a dôveru [48] v daný web. Všimnutie si elementu stránky, s ktorými používateľ nie je oboznámený, spôsobí, že sa používateľ pozastaví. Pozastavenie následne spôsobí pokles rýchlosti vykonania aktivity, kvôli ktorej na stránku prišiel [48]. Používatelia z tohto dôvodu preferujú stránky, s ktorými sú oboznámení (majú na nich hneď väčšiu úroveň zručnosti). Výška dôvery je závislá od sebavedomia človeka, reputácie stránky a oboznámenosti človeka s daným webom. Oboznámenosť a dôvera sú veľmi dôležité faktory najmä v doméne e-shopov, kde práve dôvera dosť často ovplyvňuje to, či si používateľ kúpi produkt na danej stránke alebo si nájde inú, z jeho pohľadu dôveryhodnejšiu [48].

Meranie oboznámenosti je o niečo ťažšie ako meranie dôvery. Podľa [48] sa ju viacero výskumov snažilo merať pomocou dotazníkov, čo je ale veľmi prácne, keďže je potrebné, aby používateľ vyplnil dotazník vždy, keď sa oboznámenosť zmení. Ďalším spôsobom je odvodiť oboznámenosť na základe správania pri hľadaní informácií –

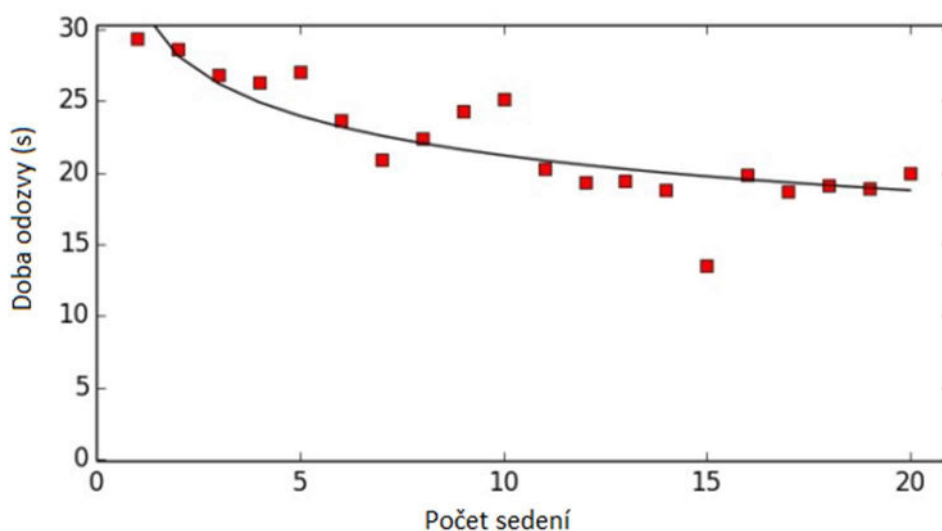
analýzou efektívnosti plnenia úlohy a času stráveného čítaním. Tretím spôsobom je inicializovať hodnotu oboznámenosti pri prvej návšteve na základe *predošlých skúseností* (angl. prior experience) s touto alebo podobnou službou a aktualizovať jej hodnotu pri každej návšteve používateľa na základe faktorov [48]:

- *opakované vystavenie* (angl. repeated exposure) – čím viac používateľ používa službu, tým viac je s ňou oboznámený,
- *úroveň spracovávania* (angl. level of processing) – spracovanie významu pridá viac na oboznámenosti ako vnemové spracovanie,
- *dĺžka trvania návštevy* (angl. study duration) – čím dlhšie používateľ strávi používaním služby, tým viac oboznámenosti nadobudne,
- *rýchlosť zabúdania* (angl. forgetting rate) – čím je dlhší časový interval medzi návštevami, tým väčší pokles oboznámenosti.

Tieto faktory a vzorce na ich výpočet sú podrobnejšie rozpísané v článku [48].

3.1 Úrovne učenia

Typickou metódou používanou v psychológii na zaznamenávanie zlepšenia výkonu človeka trénovaním je krivka učenia (angl. learning curve) [45]. Krivka učenia je matematická funkcia [33], ktorá dodržiava známy a jednoduchý model nadobúdania zručnosti - mocninová distribúcia precvičovania (angl. *The Power Law of Practices*). Tento model tvrdí, že logaritmus reakčného času špecifickej úlohy sa lineárne znižuje v závislosti od logaritmu počtu vykonaní úlohy [17]. Krivka učenia zaznamenáva vzťah medzi dĺžkou tréningu (napr. počet pokusov alebo sedení) a zmenou správania sa (Obr. 1). Využíva sa na modelovanie počiatočného stavu zručnosti používateľa, ako aj k vizualizácii, ako rýchlo používateľ nadobúda nové zručnosti. Môže byť použitá čisto iba na zobrazenie zlepšenia zručnosti v čase alebo použitá aj ďalej na predikciu či vybranie vhodnej úlohy na základe aktuálnej úrovne zručnosti.



Obr. 1. Príklad krivky učenia jedného používateľa (prebraté z [33]).

V prípade oboch, doménovo závislých aj doménovo nezávislých, sa zručnosti delia na základe úrovne (štádia) procesu nadobúdania zručností pri plnení úloh [17]:

- *úroveň založená na znalosti* – používatelia využívajú svoju duševnú kapacitu na riešenie nepreddefinovaného problému. Celý proces potrebujú mať naplánovaný od začiatku do konca. Celé postupy a pravidlá reagovania na podnety sa učia naspamäť,
- *úroveň založená na pravidlách* – používateľ si už uvedomuje stav systému a vie sa sám vedome rozhodnúť pre aplikovanie vhodného pravidla z konečného listu pravidiel,
- *úroveň založená na zručnosti* – používateľ nepotrebuje vynakladať takmer žiadnu duševnú námahu počas vykonávania úlohy. Vykonávanie akcií je podvedomé a takmer automatické.

Úroveň zručnosti používateľa nie je jednoduché kategorizovať a z tohto dôvodu vzniklo viacero spôsobov jej kategorizácie, ktoré sa líšia najmä úrovňou granularity, pričom každé delenie je vhodné na riešenie špecifického problému. Na účely našej práce by mala byť postačujúca nižšia úroveň granularity, čiže napr. začiatočník a expert.

3.1.1 Definovanie úrovne zručnosti

V tejto podkapitole si rozoberieme možné prístupy ako sa dá určiť, či je participant zručný alebo začiatočník. Veľká časť výskumov zaoberajúcich sa zručnosťou používateľa rozpoznávala zručnosť nie s webom ani aplikáciou, ale zručnosť vo vlastnom odbore (dermatológovia [46], piloti [22], neurochirurgovia [30] a biológia [6]). Títo tento problém nemuseli riešiť a stačilo im porovnať študentov so skúsenými odborníkmi.

Ďalšou skupinou sú výskumy, ktoré riešia zručnosť s aplikáciami na vizualizácie. Táto skupina sa často nezaoberala len zručnosťou, ale aj procesom učenia a predikciou ako rýchlo sa konkrétny človek naučí používať dané rozhranie. Títo k problému určovania pristupovali napríklad tak, že participantom dali na začiatku dotazníky, kognitívne testy a následne im začali dávať úlohy [45]. Ďalším prístupom je zaradenie participantov do zhlukov a nechať zhľuky označovať expertom [27]. Nakoniec je zaujímavou možnosťou rozlišovať participantov na základe percenta úspešne splnených úloh a času vykonávania konkrétnych úloh [33].

Pre porovnanie typické prístupy pre stanovenie oboznámenosti participanta sú použitie dotazníkov pred testovaním alebo použitie hodnotenia (1 až 5) – participant má určiť, nakoľko mu je dané slovo známe. Ďalšou možnosťou je použiť čas čítania, alebo efektívnosť hľadania (koľko informácií nájde za určitý čas) [48].

Oboznámenosť participanta úzko súvisí s jeho zručnosťou, takže to, že sa v tomto zozname vyskytujú aj nejaké zhody s možnosťami určovania zručnosti, len dokazuje, že ide o štandardné prístupy.

3.2 Vplyv na rozdiely medzi expertmi a začiatočníkmi

Rozdiely zaznamenané vo výkonoch expertov a začiatočníkov nezávisia iba od ich nadobudnutých zručností, ale aj od situácie a spôsobu podania informácií, s ktorými pracujú. Základnými faktormi ovplyvňujúcimi tieto rozdiely sú [16]:

- *počet nerelevantných informácií* – kapacita pracovnej pamäte je obmedzená a prítomnosť veľkého počtu nadbytočných informácií môže výrazne sťažiť proces plnenia úlohy. Vplyv tohto faktoru je významný najmä na začiatočníkov, ktorí na rozdiel od expertov nevedia rýchlo odfiltrovať nerelevantné informácie, a potrebujú viac času na spracovanie všetkých informácií. Príkladom sú realistické vizualizácie, ktoré na rozdiel od schematických obsahujú výrazne väčší počet informácií, ktoré nie sú nutne potrebné.
- *podpora generatívneho spracovania dôležitých materiálov* – zníženie rozdielov medzi expertmi a začiatočníkmi môže byť zabezpečené pomocou využitia vizuálneho aj verbálneho kanálu zároveň. Napríklad obrázok alebo textový popis nemusí byť pre začiatočníka dostatočne opisný v prípade, ak sú použité samostatne, avšak z viacerých výskumov vyplýva, že použitie viacerých kanálov zároveň má veľký prínos najmä v prípadoch používateľov neoboznámených s doménou.
- *efekt protichodnosti viacerých úrovní odbornosti (angl. expertise reversal effect)* – niektoré prístupy, ktoré uľahčujú prácu expertom, škodia začiatočníkom, a tak isto to funguje aj opačne. Ako príklad sa dá uviesť vizualizácia informácií na mape – experti vedia vyčítať dôležité informácie rýchlejšie čisto na základe vizualizácií bez pomocných textov, zatiaľ čo u začiatočníkov je to naopak.
- *zložitosť úlohy* – môžeme ju deliť do štyroch úrovní - úloha sledovania, detekčná úloha, rozhodovacia úloha alebo úloha riešenia problému. Každá úroveň má rôzny počet spôsobov riešenia, rôzny počet možných výsledkov a vyžaduje rôznu zložitosť koordinácie informačných zdrojov, s ktorými sa pracuje. Z tohto dôvodu je pochopiteľné, že boli identifikované aj rôzne závery pri analýze vzorov pohybov očí naprieč viacerými výskumami v tejto doméne.
- *čas na úlohu* – môže byť obmedzený alebo neobmedzený. Ak je čas neobmedzený, tak vykonanie úlohy trvá výrazne kratšie expertom, lebo začiatočníci majú nižšie sebavedomie a taktiež tendenciu prekontrolovať si celý postup procedúry. Pri obmedzenom čase si už začiatočníci vymedzia menej času na kontrolu a rozdiel v rýchlosti vykonávania úlohy je menší.

Ako možno vidieť, existuje mnoho faktorov, ktoré ovplyvňujú výkon používateľov pri plnení úloh, a je potrebné ich brať na vedomie pri vytváraní testovacích úloh a vytvoriť scenáre, ktorých výsledky budú ovplyvnené rovnakými externými faktormi.

3.3 Diskusia

V tejto kapitole boli zadefinované pojmy ako zručnosť, odbornosť a oboznámenosť. Zručnosť predstavuje schopnosť vykonávať funkciu získanú alebo naučenú tréningom.

Odbornosť vyjadruje vysokú úroveň zručnosti. Oboznámenosť má podobný význam ako zručnosť, avšak je viac spájaná s dôverou a formálnosťou správania. V rámci práce sme sa rozhodli skúmať zručnosť používateľov na webe. Tento problém je však vo väčšine prípadov skúmaný v doméne vizualizácií alebo analýzy obrázkov. Možným problémom, na ktorý môžeme naraziť, je prenositeľnosť existujúcich riešení do domény webu. Ako už bolo v tejto kapitole spomínané, niektoré zo zistení a formulovaných hypotéz nemusia v tejto doméne platiť.

Povedali sme si, že existuje doménovo závislá zručnosť a doménovo nezávislá zručnosť, a že v našej práci sa budeme zaoberať doménovo závislou. Definovali sme rôzne úrovne, na ktoré sa dá zručnosť používateľa deliť, a tiež, že sa dá zručnosť používateľa modelovať v čase pomocou krivky učenia. Pre náš problém sme sa však rozhodli zatiaľ pracovať s dvomi úrovňami zručnosti používateľov – začiatočník a expert.

Definovali sme rôzne prístupy, ako je možné participantov zaradiť do týchto dvoch úrovní. Typickými spôsobmi je použitie dotazníkov, použitie času plnenia úlohy, úspešnosti plnenia úlohy, koľko úloh stihne vyriešiť za časový interval a iné. Tieto by bolo vhodné aplikovať aj pri realizácii našich experimentov.

Nakoniec sme definovali faktory, ktoré ovplyvňujú rozdiely medzi expertami a začiatočníkmi. Medzi najdôležitejšie z nich považujeme zložitosť úlohy – v prípade zložitejších úloh by sa mohli lepšie prejavovať rozdiely medzi participantami (je možné ich použiť na zaraďovanie participantov do jednotlivých úrovní zručnosti). Ďalším dôležitým faktorom je čas na úlohu – nechceme čas na vykonanie úlohy obmedzovať, aby vznikali väčšie rozdiely v časoch plnenia úloh, čo spôsobí ľahšiu rozlíšiteľnosť zručnosti participantov.

4 Existujúce metódy analýzy pohľadu

V kapitole 2.1 boli zavedené pojmy fixácia a sakáda, pomocou ktorých sa modelujú sledované pohyby očí. Zaznamenané údaje sa však zvyčajne ďalej analyzujú a sú použité na výpočet odvodených metrík. Odvodené metriky majú väčšiu výpovednú hodnotu o tom, aké mal používateľ problémy s rozhraním (napr. testovanie použiteľnosti stránky), ako sa na stránke správal alebo aké sú jeho charakteristiky. Hodnoty odvodených metrík môžu byť ďalej jednotlivo interpretované expertom alebo použité na natrénovanie modelu, ktorý berie do úvahy viacero atribútov naraz.

Odvodené metriky sa zvyčajne delia na základe toho, od čoho sú odvodené na [9]:

- odvodené od sakád
- odvodené od fixácií
- odvodené od AOI
- odvodené od sekvencie pohľadu
- odvodené od zreničky
- odvodené inak

Príklady vhodných odvodených metrík patriacich do jednotlivých kategórií sa nachádzajú v kapitole 4.2. Avšak metriky v danom tvare sú použiteľné iba pre analýzu expertom. V prípade, že majú byť metriky použité neskôr na natrénovanie modelu, je ich vo väčšine prípadov potrebné agregovať tak, aby pre každý typ metriky zostala iba jedna reprezentatívna hodnota pre jedno meranie. Na agregovanie hodnôt sa používajú funkcie [4,17,31]:

- suma – sčítanie všetkých hodnôt danej metriky na vybranom intervale alebo meraní,
- priemer – spriemerovanie všetkých hodnôt danej metriky na vybranom intervale alebo meraní,
- štandardná odchýlka – výpočet štandardnej odchýlky zo všetkých hodnôt danej metriky na vybranom intervale alebo meraní,
- počet – spočítanie počtu výskytov všetkých hodnôt danej metriky na vybranom intervale alebo meraní,
- minimum – nájdenie najmenšej hodnoty danej metriky na vybranom intervale,
- maximum – nájdenie najväčšej hodnoty danej metriky na vybranom intervale,
- prvá – nájdenie prvej hodnoty danej metriky na vybranom intervale,
- posledná – nájdenie poslednej hodnoty danej metriky na vybranom intervale.

Keďže sa v práci usilujeme o automatickú identifikáciu zručnosti pomocou modelu, tak sa použitiu týchto agregáčnych funkcií nevyhneme.

4.1 Sekvencia pohľadu

Pred zavedením konkrétnych metrík v kapitole 4.2 je potrebné zadať pojmy ako sekvencia pohľadu (angl. scanpath) a miesto záujmu (angl. Areas of Interest, AOI), ktoré s niektorými typmi metrík veľmi úzko súvisia.

Sekvencia pohľadu je termín, ktorým sa pomenováva postupnosť pohybov očí tvorenú fixáciami a sakádami [10]. Sekvencia pohľadu sa zvyčajne vizualizuje ako číslovaná postupnosť kruhov symbolizujúcich fixácie. Fixácie sú prepojené čiarami symbolizujúcimi prechody medzi nimi – sakády (viď Obr. 2). Extrahované sekvencie pohľadov sa využívajú na rôzne účely, avšak vo väčšine prípadov sa používa ich reprezentácia v podobe reťazcov znakov, pričom jednotlivé znaky reprezentujú navštívené miesta záujmu (AOI) [10]. AOI môžu byť zvolené ručne alebo automaticky (analýzou zdrojového kódu stránky) a pokrývajú plochu vizuálnych elementov stránky, ktoré sú z hľadiska vykonávanej analýzy dôležité (červeno farbou orámované obdĺžniky na Obr. 2). Typickými riešeniami pre automatickú segmentáciu sú algoritmy VIPS [1] a BOM [38]. Nefungujú však úplne spoľahlivo. Preto by bolo pre našu prácu vhodnejšie použiť častejšie používaný prístup – voliť AOI ručne.



Obr. 2 Prevzaté z [7]. Ukážka webovej stránky Babylon, na ktorej sú vyznačené oblasti záujmu obdĺžnikmi s červeným orámovaním a sekvencia pohľadu ako postupnosť číslovaných kruhov prepojená čiarami symbolizujúcimi sakády.

Dôležitosť vizuálneho elementu stránky môže byť zvolená na základe relevantnosti elementu pri riešení konkrétnej úlohy, keďže úspešní riešitelia úloh venujú viac pozornosti relevantným informáciám na stránke [28]. Existujú však aj prístupy, ktoré netvorí reťazce znakov len z AOI, ale aj z uhlov sakád (absolútnych alebo relatívnych), dĺžky trvania fixácií alebo dĺžky sakád [20]. Definícia konkrétneho spôsobu výpočtu metrík pomocou týchto alternatívnych prístupov však chýba, preto sa nimi v práci zaoberať nebudeme.

4.2 Používané metriky

V tejto kapitole sú zavedené často využívané metriky v oblasti sledovania pohľadu, ktoré sú zaradené do jednotlivých kategórií na základe zdroja, od ktorého sú odvodené. Uvedené kategórie vznikli zlúčením delení zadefinovaných v článku od Ehmke a kol. [9] (metriky odvodené od fixácií sakád a sekvencií pohľadu) a v článku od Tokera a kol. [44] (metriky odvodené od AOI, od zaznamenávaných údajov o zreničke a odvodené inak).

Metriky odvodené od sakád

- *dĺžka sakády* [25,26] – vzdialenosť medzi dvomi následnými fixáciami tvoriacimi jednu sakádu,
- *relatívny uhol sakád* [25,26] – relatívny uhol medzi dvomi po sebe nasledujúcimi fixáciami (uhol medzi prechádzajúcou a aktuálnou sakádou),
- *absolútny uhol sakád* [25,26] – absolútny uhol medzi dvomi po sebe nasledujúcimi fixáciami (uhol medzi horizontálnou čiarou a sakádou),
- *počet sakád* [9] – počet sakád za časový interval alebo za čas plnenia úlohy (už agregovaná metrika).

Metriky odvodené od fixácií

- *frekvencia fixácií* (angl. fixation rate) [25,26] – počet fixácií za milisekundu,
- *počet fixácií* [15,25,26] – počet zaznamenaných fixácií počas stanoveného časového intervalu (už agregovaná metrika),
- *dĺžka trvania fixácie* [8,15,25,26] – dĺžka časového intervalu, počas ktorého bola zaznamenaná jedna fixácia,
- *čas do prvej fixácie* [8,15] – už agregovaná metrika,
- *pomer fixácií a sakád* [9] – pomer celkovej dĺžky trvania fixácií a sakád (už agregovaná metrika).

Metriky odvodené od zaznamenávaných údajoch zreničke

- *šírka zreničiek* [31,41] – priemer zreničky,
- *rýchlosť dilatácie zreničiek* [44] – rýchlosť rozšírenia zreničky,
- *šírka zreničky pri prvej a poslednej fixácii* [31,41] – zaznamenaný priemer zreničky pri prvej a poslednej fixácii plnenia úlohy.

Metriky odvodené od AOI

- *frekvencia fixácií na AOI* [31,41] – počet fixácií za milisekundu na konkrétnej AOI,
- *dĺžka trvania fixácie na AOI* [33] – dĺžka časového intervalu, počas ktorého bola zaznamenaná jedna fixácia na konkrétnej AOI,
- *čas do prvej a poslednej fixácie na AOI* [44] – už agregovaná metrika,
- *prechody medzi AOI* [31,41] – používajú sa agregácie tejto metriky ako počet a pomer prechodov medzi párami AOI,
- *počet fixácií na AOI* [44] – už agregovaná metrika,
- *suma dĺžok trvaní fixácií na AOI* [44] – už agregovaná metrika.

V prípade, že by sa AOI rozdelili do dvoch kategórií na základe toho, či sú pre riešenie úlohy relevantné, alebo nie, je možné odvodiť ešte ďalší typ metrík od vyššie uvedených metrík [16]:

- *počet fixácií na relevantnú/nadbytočnú AOI*,
- *dĺžka trvania fixácie na relevantnú / nadbytočnú AOI*,

- čas do prvej fixácie na relevantnú AOI.

Metriky odvodené od sekvencie pohľadu

- *dĺžka sekvencie pohľadu* [9] – suma dĺžok sakád sekvencie pohľadu (už agregovaná metrika),
- *dĺžka trvania sekvencie pohľadu* [9] – suma dĺžok trvania sakád a fixácií sekvencie pohľadu (už agregovaná metrika),
- *priestorová hustota sekvencie pohľadu* [9] – nad celým rozhraním sa vytvorí mriežka a priestorová hustota je vypočítaná ako pomer počtu buniek obsahujúcich aspoň jednu fixáciu k celkovému počtom buniek [21],
- *pravdepodobnosti prechodov medzi AOI* [9] – vypočítane pomocou matice prechodov (angl. transition matrix).

Metriky odvodené inak

- *vzdialenosť hlavy od monitoru* [44] – priemer vzdialeností pravého a ľavého oka od monitoru v centimetroch,
- *prvá a posledná vzdialenosť hlavy* [44] – priemer vzdialeností pravého a ľavého oka od monitoru v centimetroch pri prvej a poslednej fixácii plnenej úlohy.

Tieto metriky sú vo výskumoch často používané pri riešení klasifikačných problémov. V článku [44] sa dokázalo, že niektoré majú vysokú rozlišujúcu schopnosť medzi zručnými používateľmi a začiatočníkmi pri úlohách s vizualizáciami. Model využívajúceho tieto a ďalšie podobné atribúty v práci implementujeme a rozširujeme ho o ďalšie atribúty, ktoré ešte v súvislosti s týmito použité neboli (kapitola 4.3) a o nami vytvorené atribúty (kapitola 5).

4.3 Rekurentná kvantifikačná analýza

Rekurentná kvantifikačná analýza (angl. Recurrence quantification analysis, ďalej už iba ako RQA) je spôsob analýzy globálnych a lokálnych časových charakteristík fixácií, ktorý sa používa na charakterizovanie vzoru sekvencie pohľadu jedného pozorovateľa. Na charakterizovanie sekvencie pohľadu využíva vlastné metriky, ktoré zovšeobecňujú sekvencie pohľadu nad pozorovateľom alebo stimulom [2]. Tieto metriky sú vypočítavané nad grafom rekurentnosti (angl. recurrence plot) (viď Obr. 3). Fixácie sú považované za rekurentné, ak sú dostatočne blízko pri sebe. Blízkosť môže byť určená viacerými spôsobmi (napr. fixácie patria do jednej AOI, fixácie patria do jednej bunky v mriežke, alebo sú fixácie od seba vzdialené do určitého polomeru). V prípade častejšie používaného polomeru sa rekurencia dvoch fixácií definovaná ako [2]:

$$r_{ij} = \begin{cases} 1, & d(f_i, f_j) < \rho \\ 0, & \text{inak} \end{cases}, \quad (1)$$

Kde d je funkcia vzdialenosti medzi dvomi fixáciami a ρ stanovený polomer. Ak je hodnota r_{ij} rovná 1, tak je do grafu rekurencie na pozíciu i, j zaznačená bodka (viď Obr. 3).

V nasledujúcej časti sú vysvetlené štyri, podľa autorov v [2], najvhodnejšie RQA metriky na analýzu fixácií. Keďže graf rekurentnosti je symetrický po diagonále, stačí keď sa metriky vypočítavajú iba nad horným trojuholníkom.

Rekurencia (angl. recurrence) - je definovaná [2]:

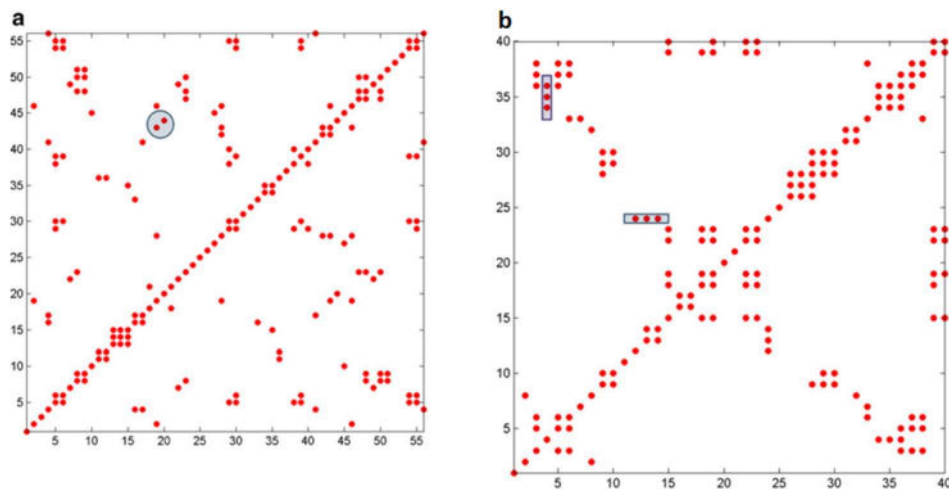
$$REC = 100 \frac{2R}{N(N-1)}, \quad (2)$$

kde R je suma všetkých rekurencií (všetkých bodiek) v hornom trojuholníku a N vyjadruje počet všetkých fixácií. Táto metrika vyjadruje, ako často sa ľudia znovu pozreli na časť obrazovky, na ktorú sa už v minulosti pozerali.

Determinizmus (angl. determinism) je definovaný ako [2]:

$$DET = 100 \frac{|D_L|}{R}, \quad (3)$$

kde D_L je suma všetkých bodov v hornom trojuholníku tvoriacich diagonály (viď Obr. 3a). Determinizmus meria, nakoľko sa opakujú vzory v sekvencii pohľadu. Hľadané body sú časti sekvencie pohľadu, ktoré sa zhodujú s časťou sekvencie pohľadu, ktorá nastala v minulosti (rovnaké s časťou hlavnej diagonály). Minimálny počet bodov tvoriacich diagonálu je 2.



Obr. 3. Prevzaté z [2]. Čísla na osiach reprezentujú jednotlivé fixácie. Obrázok (a) znázorňuje vzor, aký sa hľadá pri determinizme a obrázok (b) znázorňuje vzor, aký sa hľadá pri vrstvosť.

Vrstvosť (angl. laminarity) je definovaný ako [2]:

$$LAM = 100 \frac{|H_L| + |V_L|}{2R}, \quad (4)$$

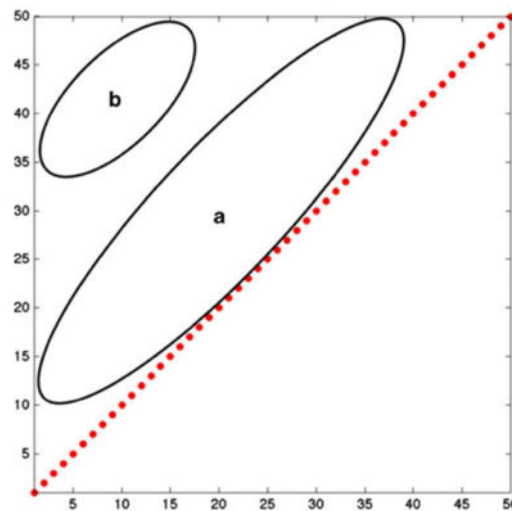
H_L je suma všetkých bodov v hornom trojuholníku tvoriacich horizontálne čiary a V_L je suma všetkých bodov v hornom trojuholníku tvoriacich vertikálne čiary (viď Obr. 3b). Identifikované horizontálne čiary hovoria o tom, kedy bola istá oblasť najskôr preskúmaná detailne a neskôr len zbežne. Naopak, vertikálne čiary vznikajú, keď sa najskôr na istú oblasť pozrieme zbežne (zaznamenaná sa len jedna fixácia) a až neskôr ju opäť preskúame,

ale už detailnejšie (zaznamená sa viac fixácií). Minimálny počet bodov tvoriacich čiary je 2.

Stred masy opakovania (angl. center of recurrence mass) je definovaný ako [2]:

$$CORM = 100 \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N (j-i) r_{ij}}{(N-1)R}. \quad (5)$$

Vyjadruje, kde približne v čase sa body rekurencie nachádzajú. Ak sa rekurencie nachádzajú v čase navzájom ďalej od seba, hodnota CORM bude vysoká a naopak, ak sa nachádzajú v čase navzájom blízko, hodnota CORM bude nízka (viď Obr. 4).



Obr. 4. Prevzaté z [2]. CORM je vysoký keď sa rekurencie vyskytujú v čase od seba ďalej – elipsa b. Naopak CORM je nízky, keď sa rekurencie vyskytujú v čase blízko za sebou – elipsa a.

Celkovo sa dá teda povedať, že determinizmus a vrstvosť vypovedajú o podrobnejšej štruktúre sekvencie pohľadu v čase, zatiaľ čo rekurencia a CORM zachytávajú viac globálnu štruktúru sekvencie pohľadu [2]. Poskytujú teda iný pohľad na dáta zo sledovača pohľadu a dúfame, že ich pripojenie k bežne používaným atribútom môže automatickému rozlišovaniu zručnosti pomôcť.

V článku [29] však tvrdia, že dané metriky nie je vhodné nazývať rekurentné (angl. recurrence), ale za lepši názov považujú opätovne vyskytujúce (angl. reoccurrence). Dôvodom je, že neobsahujú krok vnárania časového oneskorenia (angl. Time-Delay Embedding, TDE), ktorý je pre rekurenciu typický a bez neho nie je splnený potrebný predpoklad, že je v grafe rekurentnosti zachovaný fázový priestor trajektórií. Rozdiel teda je, že rekurentná analýza hľadá cyklický priebeh dynamických systémov, čiže postupností opakujúcich sa dát. Analýza opätovného výskytu hľadá opakujúce sa prvky dát a matematicky nedokazuje zachovanie dynamických systémov. Vo svojej knižnici, ktorú vytvorili za účelom výpočtu rekurentných aj reokurentných metrík zadefinovali metriky:

- *rekurencia* – počet rekurentných buniek v hornom trojuholníku matice,

- *rekurentná miera (angl. Recurrence rate)* – sa počíta pre rekurenciu R a celkový počet buniek N ako:

$$recurrenceRate = \frac{R}{N * (N-1) * 0.5}, \quad (6)$$

- *priemerná rekurencia X a Y* – priemerná hodnota rekurencie na X -ovej a Y -ovej osi,
- *štandardná odchýlka rekurencie X a Y* – štandardná odchýlka hodnôt rekurencie na X -ovej a Y -ovej osi.

Výskum [46] skúmal, nakoľko sú vyššie zadefinované metriky užitočné pri rozlišovaní expertov a začiatočníkov v dermatológii pri analýze obrázkov. Pre všetky metriky vyšli rozdiely medzi kategóriami expertov a začiatočníkov signifikantne rozdielne. Na základe nimi zistených rozdielov sa dajú vyvodiť závery pre každú metriku:

- *rekurencia* – je u expertov nižšia ako začiatočníkov, takže experti robia spätné fixácie na už navštívené elementy menej často ako začiatočníci,
- *determinizmus* – je u expertov nižší ako u začiatočníkov, takže experti opakujú krátke postupnosti fixácií menej často ako začiatočníci,
- *vrstvosť* – je u expertov nižšia ako u začiatočníkov, takže experti majú menej opakovaných fixácií na konkrétnu oblasť ako začiatočníci,
- *CORM* – je u expertov vyšší ako u začiatočníkov, čiže experti robia spätné fixácie po dlhších intervaloch.

Aj keď tvrdenie, že na začiatku kapitoly definované metriky sú reokurentné môže byť pravdivé, tieto metriky sa ukázali ako rozlišujúce medzi expertmi a začiatočníkmi (v doméne dermatológie). Preto je ich určite vhodné aspoň vyskúšať. Keďže rekurentné metriky môžu byť definované aj inak [29], tak by bolo vhodné otestovať aj túto druhú implementáciu. V práci sme sa rozhodli použiť použiť rekurentné aj reokurentné metriky.

V článku [14] zisťovali, ako sú tieto metriky ovplyvnené zložitou obrázku a úlohou. Pre zložitú obrázku bolo identifikované, že determinizmus klesá so zvyšujúcou sa zložitou obrázku. V prípade vplyvu úlohy, bolo zistené, že metrika rekurentnosť ňou ovplyvnená nie je, avšak determinizmus je vyšší u používateľov, ktorý plnia úlohu, ako u používateľov, ktorý si obrázok len prezerajú. Je však otázne, ako budú tieto zistenia prenositeľné do domény elektronických obchodov.

4.4 Analýza sekvencií pohľadu

Ak sa neanalyzuje úloha s cieľom (angl. goal-directed), ale prieskumná úloha (angl. exploratory task), tak je za vhodnejší spôsob získania informácií považované použitie metód na analýzu sekvencií pohľadu, a nie metrick sledovania pohľadu. Medzi tieto metódy patria [11]:

- identifikácia vzorov,

- výpočet pravdepodobnosti prechodov medzi vizuálnymi elementmi,
- výpočet podobnosti sekvencií pohľadu,
- identifikácia spoločných sekvencií pohľadu.

Metódy na identifikáciu vzorov sa snažia nájsť často vyskytujúce sa vzory v rámci sekvencií pohľadu. Pokrývajú riešenie problémov od identifikácie konkrétneho vzoru (v sekvenciách pohľadu), až po detekciu všetkých vzorov (s určitým počtom výskytov) [11]. Patria medzi nich eyePatterns [47] a T-vzor [34]. V práci sa im však nebudeme venovať, lebo sú veľmi náchylné na malé odlišnosti vo vzoroch.

Výpočet pravdepodobnosti prechodov medzi vizuálnymi elementami nebude taktiež v práci ďalej rozoberaný, lebo nepracuje so vzormi (postupnosťou elementov v čase), ale len staticky popisuje pravdepodobnosti prechodov z jedného elementu do druhého [11].

Za vhodný spôsob analýzy sekvencií pohľadu považujeme použitie algoritmov na identifikáciu spoločných sekvencií pohľadu. Umožňujú totiž vytvárať trendujúce sekvencie pohľadu pre zadanú skupinu sekvencií. Nie sú náchylné na malé odlišnosti vo vzoroch a umožnia nám vytvoriť charakteristickú spoločnú sekvenciu pohľadu pre skupinu zručných participantov a pre skupinu začiatočníkov. Aby však bolo možné povedať, do ktorej skupiny zatiaľ nezaraďený participant patrí, je potrebné byť schopný vypočítať podobnosti medzi sekvenciami pohľadu.

Spôsoby fungovania metód na výpočet podobnosti sekvencií pohľadu a identifikáciu spoločných sekvencií pohľadu sú popísané v kapitolách 4.4.1 a 4.4.2.

4.4.1 Algoritmy na výpočet podobnosti sekvencií pohľadu

Aby sa dali sekvencie pohľadu nejakým spôsob porovnávať, je potrebné poznať spôsob, ako identifikovať, či sú podobné, alebo nie. Najštandardnejšie sa táto problematika rieši algoritmom založeným na porovnaní textových reťazcov (angl. string-edit).

Algoritmus založený na porovnaní textových reťazcov

Algoritmus vždy porovnáva dve sekvencie pohľadu. Na aplikovanie algoritmu je potrebné pracovať so sekvenciami pohľadu, ktoré sú vo formáte reťazcov znakov. Algoritmus transformuje jednu z porovnávaných sekvencií pohľadu na druhú. Túto transformáciu vykonáva za použitia troch typov operácií: *pridanie znaku*, *odstránenie znaku* a *nahradenie znaku*. Minimálny počet operácií, ktoré je potrebné vykonať pre transformáciu z jednej sekvencie pohľadu na druhú sa nazýva hodnota rozličnosti. Táto hodnota sa následne normalizuje dĺžkou dlhšej z dvojice sekvencií pohľadu tak, aby hodnota pre úplne rovnaké sekvencie pohľadu vyšla 0 a pre najviac rozdielne sekvencie vyšla 1 [15].

Aj keď ide o často používaný algoritmus, má stále niekoľko nedostatkov. V prvom rade neberie do úvahy časy trvania fixácií a taktiež neberie do úvahy pozíciu vizuálnych elementov na stránke [11]. Tieto problémy sa do istej miery podarilo zmierniť úpravami tohto algoritmu. Jedna z nich váhuje cenu operácie substitúcie na základe vzdialenosti týchto dvoch vizuálnych elementov od seba, a tým zapracováva do algoritmu vedomosť

o pozícii vizuálnych elementov [42]. Druhá úprava rieši problém trvania fixácií tak, že z dlhších fixácií vytvorí viacero kratších fixácií za sebou [15].

Tento algoritmus je štandard a vyhovuje našim potrebám. Možnosťou je ho ešte upraviť o spomínané váhovanie na základe vzájomných vzdialeností AOI od seba. V ďalšej časti opisujeme algoritmy na identifikáciu vzorov a algoritmy na identifikáciu spoločných sekvencií pohľadu. Predtým je však vhodné zadať, čo znamená, že je algoritmus redukčný. Redukčný algoritmus je algoritmus, ktorý v rámci svojho procesu stráca relevantné informácie jedným z nasledovných spôsobov [11]:

- algoritmus stráca spoločné vizuálne elementy kvôli ich nezhodujúcim sa pozíciám v rôznych sekvenciách pohľadu,
- algoritmus je netolerantný k malým odchýlkam medzi sekvenciami pohľadu.

4.4.2 Algoritmy na identifikáciu spoločných sekvencií

Na rozdiel od algoritmov na identifikáciu vzorov v sekvenciách pohľadu sa algoritmy na identifikáciu spoločných sekvencií pohľadu neusilujú o nájdenie vzorov, ktoré sa nachádzajú vo všetkých sekvenciách pohľadu, ale o identifikáciu jedného vzoru, ktorý je reprezentatívny pre väčšinu sekvencií pohľadu [11]. Ide teda o spoločnú sekvenciu pohľadu, ktorá je charakteristická pre určitú skupinu používateľov. V nasledujúcich podkapitolách sú popísané možné prístupy, ako sa dá spoločná sekvencia pohľadu vypočítať.

Pozične váhovaný model

Pozične váhovaný model (angl. Position based Weighted Models) analyzuje, ktorým vizuálnym elementom sa používateľ venuje na začiatku prehľadávania webovej stránky. V článkoch [23,38] aplikovali algoritmus rôzne. V [41] rozdelili stránku na segmenty pomocou mriežky 3x3, zatiaľ čo v [25] segmentovali stránku pomocou AOI. V oboch si však stanovili výslednú dĺžku sekvencie pohľadu, ktorú hľadajú, ako N . Následne zobrali z každej sekvencie pohľadu prvých N segmentov (návštev AOI) a za každý výskyt segmentu v sekvenciách pohľadu mu pričítali konkrétnu hodnotu (viď Obr. 5). V prípade článku [41] bolo stanovené $N = 3$, a ak sa segment nachádzal v sekvencii pohľadu na prvej pozícii, tak bola danému segmentu pripočítaná hodnota 1. Segmentu na druhej pozícii bola pričítaná hodnota 0.5 a tretiemu segmentu hodnota 0.25. Segmenty boli následne zoradené zostupne na základe celkového počtu bodov a za spoločnú sekvenciu pohľadu bola stanovená postupnosť troch segmentov s najväčším počtom bodov. Algoritmus nie je redukčný, avšak jeho nevýhodou je, že v spoločnej sekvencii pohľadu sa segmenty nemôžu opakovať. Opakovanie sa segmentov v sekvencii pohľadu je však bežné (spätne fixácie na už navštívené elementy už boli spomínané v kapitole 4.3).

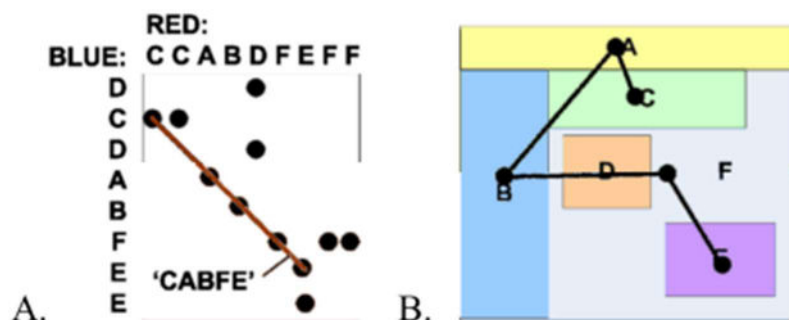
Prvé 3 vizuálne elementy sekvencie pohľadu	Vizuálne elementy						
	A	B	C	D	E	F	G
ABC	1	0.5	0.2	0	0	0	0
CAE	0.5	0	1	0	0.2	0	0
CAB	0.5	0.2	1	0	0	0	0
CAC	0.5	0	1+0.2	0	0	0	0
DAC	0.5	0	0.2	1	0	0	0
DCD	0	0	0.5	1+0.2	0	0	0
BDC	0	1	0.2	0.5	0	0	0
CAC	0.5	0	1+0.2	0	0	0	0
BCE	0	1	0.5	0	0.2	0	0
CAC	0.5	0	1+0.2	0	0	0	0
Body spolu:	4	2.7	7.2	2.7	0.4	0	0

Obr. 5. Preložené z [11].

Hierarchické zhľukovanie s vykresľovaním bodiek (angl. Dotplot)

Algoritmus vznikol pôvodne za účelom porovnania dvoch sekvencií v biológii [11]. Dotplot sám o sebe dokáže identifikovať spoločnú sekvenciu pohľadu naraz len z dvoch sekvencií pohľadu. Tieto sekvencie pohľadu sa zobrazia do grafu, pričom jedna sa zobrazí na horizontálnu os a druhá na vertikálnu a následne sa zakreslia bodky do buniek, v ktorých sa názvy AOI pre obe osi zhodujú [20] (viď Obr. 6a). Potom je v grafe nájdená najdlhšia lineárna čiara (lineárna čiara, ktorá obsahuje najväčší počet bodiek). Za spoločnú sekvenciu je označená postupnosť znakov reprezentujúca bodky v grafe tvoriace najdlhšiu nájdenú lineárnu čiaru (viď Obr. 6b, postupnosť znakov CABFE).

Aby bolo umožnené identifikovať spoločnú sekvenciu z viac než dvoch sekvencií pohľadu, tak sa výsledky dotplot algoritmu hierarchicky zhľukujú. To znamená, že sa vždy vyberie dvojica najpodobnejších sekvencií pohľadu (podobnosť sa určuje pomocou samotného dotplot algoritmu) a na túto dvojicu je aplikovaný dotplot algoritmus, ktorého výstupná sekvencia pohľadu sa pridá do zoznamu sekvencií pohľadu. Následne sa dve pôvodné sekvencie pohľadu zo zoznamu odstraňuje. Menej používaným prístupom je, že na najpodobnejšiu dvojicu nepoužije dotplot algoritmus, ale do zoznamu sekvencií pohľadu je pridaná jedna z dvojice podobných sekvencií pohľadu. Toto sa opakuje až pokiaľ nezostane jedna výsledná spoločná sekvencia pohľadu. Keďže aplikovaním dotplot algoritmu na dvojicu sekvencií pohľadu vznikne vždy kratšia sekvencia ako sú pôvodné sekvencie pohľadu, je algoritmus považovaný za redukčný.



Obr. 6. Prevzané z [20]. Obrázok A znázorňuje spôsob hľadania spoločnej sekvencie pohľadu (pôvodné sekvencie pohľadu sú označené na osiach). Obrázok B zobrazuje spoločnú sekvenciu pohľadu na stránke s pozadím zvýrazňujúcim jednotlivé AOI.

eMine

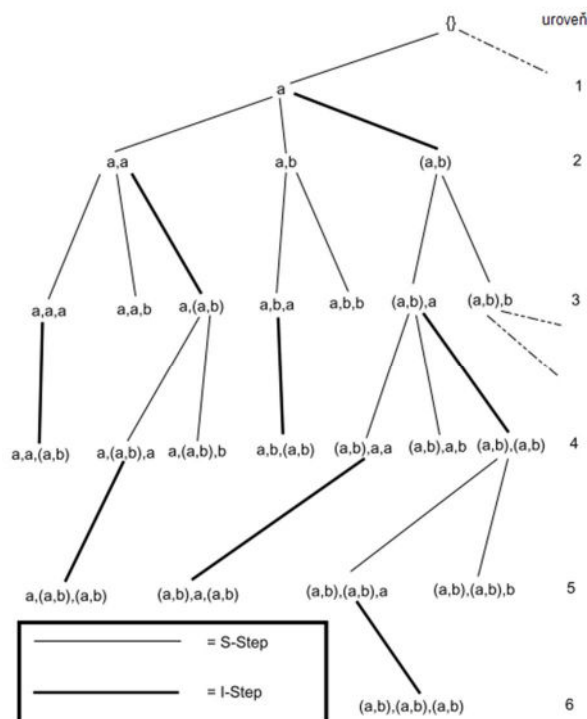
Snaží sa byť menej redukčný ako vyššie uvedené algoritmy a tým vytvoriť dlhšie spoločné sekvencie pohľadu. Vie pracovať s viac ako dvomi sekvenciami pohľadu naraz. Na začiatku má sekvencie pohľadu uložené v zozname. V tomto zozname hľadá dve sekvencie pohľadu, ktoré sú navzájom najpodobnejšie podľa algoritmu založenom na porovnaní textových reťazcov. Keď ich nájde, tak ich zo zoznamu odstráni a použije na vytvorenie ich spoločnej sekvencie pohľadu s využitím algoritmu na nájdenie najdlhšej spoločnej podpostupnosti (angl. Longest Common Subsequence). Spoločnú sekvenciu pohľadu pridá do zoznamu a proces sa opakuje, až dokým v zozname nezostane iba jedna sekvencia pohľadu. Výsledná sekvencia je zredukovaná (ak sa rovnaké znaky nachádzajú za sebou, tak sú nahradené iba jedným znakom napr. RNNNNN $\rightarrow$ RN) a následne je označená ako spoločná sekvencia pre sekvencie pôvodného zoznamu [10]. Opäť raz ide o hierarchické zhľukovanie ako v prípade Dotplot algoritmu. Kvôli tomu je algoritmu tiež redukčný a taktiež produkuje krátke spoločné sekvencie pohľadu.

SPAM

SPAM umožňuje nájsť aj viac spoločných sekvencií pohľadu ako jednu. Algoritmus bol pôvodne navrhnutý na hľadanie vzorov v transakciách zákazníkov, ktoré sú zapísané v databáze. Keď sa pospájajú transakcie jedného zákazníka za sebou vznikne reprezentácia, ktorá už je podobná sekvenciám pohľadu [3].

SPAM postupne vytvára lexikografický sekvenčný strom zobrazený na Obr. 7. Začína od prázdnej množiny, do ktorej postupne pridáva znaky a transakcie. Tieto sú pridávané dvomi rôznymi krokmi [3]:

- *S-krok* – na koniec sekvencie rodiča pridá novú transakciu s jedným prvkom (viď tenké čiary v Obr. 7),
- *I-krok* – pridá znak na koniec poslednej transakcie rodiča (viď hrubé čiary v Obr. 7).



Obr. 7. Preložené z [3]. Ukážka lexicografického sekenčného stromu.

Algoritmus prechádza a generuje prvky stromu prehľadávaním priestoru do hĺbky. Kedy prestať alebo ktoré prvky vynechať, určuje hodnota podpory sekvencie (angl. support). Podpora sekvencie vyjadruje frekvenciu výskytu sekvencie v databáze (v našom prípade percento prípadov, v ktorých sa vytvorená postupnosť nachádza v nameraných sekvenciách pohľadu). Ak je hodnota podpory sekvencie nižšia ako preddefinovaný prah, tak sa postupnosť do stromu nepridá. Algoritmus je optimalizovaný pomocou orezávania v krokoch S-krok aj I-krok. Ďalšou optimalizáciou je použitie bitovej mapy, ktorá obsahuje pre každý znak (z preddefinovanej množiny) hodnoty 0 alebo 1 pre každú postupnosť v databáze (pre každú nameranú sekvenciu pohľadu), čím zrýchľuje výpočet hodnoty podpory sekvencie [3]. Aj keď sa nejedná čisto o algoritmus na identifikovanie spoločnej sekvencie pohľadu (lebo spoločných sekvencií hľadá viac), tak by ho mohlo byť zaujímavé použiť tak, že by sme ako výsledok algoritmu brali najdlhšiu nájdenú spoločnú sekvenciu pohľadu.

STA

Celým názvom analýza trendujúcej sekvencie pohľadu (angl. Scanpath Trend Analysis, ďalej už iba ako STA). Nerieši výslovne problém spoločnej sekvencie pohľadu, keďže pojem *spoločný* je podľa oxfordského slovníka definovaný ako: “zdieľané, pochádzajúce alebo vykonané dvomi a viacerými ľuďmi, skupinami alebo vecami” [13]. Algoritmus sa však zameriava na hľadanie trendujúcej sekvencie pohľadu. *Trend* je podľa oxfordského slovníku definovaný ako: “Všeobecný smer, ktorým sa niečo vyvíja a mení” [13]. Z tohto dôvodu STA berie do úvahy vizuálne elementy navštívené všetkými používateľmi, ako aj vizuálne elementy navštívené väčšinou používateľov. Algoritmus dokáže pracovať s viac

ako dvomi sekvenciami pohľadu bez použitia hierarchického zhlukovania a taktiež sa vie vysporiadať s problémom vznikajúcimi, keď znaky v reťazci nie sú úplne v rovnakom poradí. Podobne ako eMine algoritmus sa snaží adresovať problém, že algoritmy sú vo väčšine prípadov redukčné a vytvárajú príliš krátke spoločné sekvencie pohľadu. STA algoritmus pozostáva z troch krokov [12]:

Predbežná fáza – prebieha v nej predspracovanie dát. Vizualné elementy sú pomenované a označené ako AOI. Následné sú dáta vyexportované zo softvéru na sledovanie pohľadu. Keďže sledovače pohľadu si sú pozíciami fixácií isté len do určitej miery, čo by mohlo pri ďalšej analýze spôsobiť problémy, tak sa na základe tejto miery rozširujú hranice AOI. V prípade, že sa fixácia nachádza v prekryve dvoch AOI, tak je zaradená do bližšej AOI. Blízkosť od AOI sa počíta ako vzdialenosť všetkých bodov danej AOI od fixácie. Predbežná fáza algoritmu je najpomalšou časťou algoritmu (má časovú zložitosť $O(n^6)$) najmä kvôli výpočtu blízkosti. Výstupom tejto fázy sú postupnosti AOI spolu s časmi trvania fixácií, pre každú sekvenciu pohľadu.

Prvý prechod – identifikuje trendujúce vizualné elementy. Najskôr sa identifikujú inštancie vizualných elementov. Inštancia vizualného elementu je definovaná ako jedna návšteva vizualného elementu, čiže viacero fixácií na ten istý vizualný element za sebou tvorí stále iba jednu inštitúciu. Aby boli inštancie rozlíšiteľné, sú číslované na základe dĺžky trvania fixácií (viď Príklad 1). Najdlhšia inštancia vizualného elementu je označená ako 1 a ďalšia v poradí ako 2 atď.

A2[200ms] C1[150ms] A1[56ms] A1 [130ms] A1 [70m]

Príklad 1. Druhý výskyt vizualného elementu A je označený ako 1, lebo $(56+130+70) > 200$.

Následne sa identifikujú trendujúce inštancie vizualné elementy. Keď sa názov vizualného elementu nachádza vo všetkých sekvenciách pohľadu, je vizualný element považovaný za trendujúci. Aby bolo možné brať do úvahy aj elementy, ktoré sa nevyskytujú vo všetkých sekvenciách pohľadu, vypočíta sa prah pre frekvenciu výskytu a dĺžku trvania inštancie vizualného elementu. Prah je stanovený ako minimálna hodnota frekvencie a trvania z vizualných elementov nachádzajúcich sa vo všetkých sekvenciách pohľadu. Za trendujúce elementy sú následne stanovené aj tie inštancie vizualných elementov, ktoré majú aspoň takú frekvenciu výskytu a aspoň takú dĺžku trvania ako stanovené prahy. Ostatné inštancie sú odstránené.

Druhý prechod – vytvorí trendujúcu sekvenciu pohľadu. Najskôr sú sekvencie pohľadu abstrahované – sú z nich odstránené údaje o dĺžke trvania a rovnaké vizualné elementy nachádzajúce sa za sebou sú nahradené jedným elementom (viď Príklad 2).

A2 C1 A1

Príklad 2. Abstrahovaná sekvencia pohľadu z príkladu v predchádzajúcej fáze. Zo sekvencie pohľadu sú odstránené údaje o dĺžke trvania a rovnaké vizualné elementy nachádzajúce sa za sebou sú nahradené jedným elementom.

Ďalej je pre každú inštanciu vizuálneho elementu vypočítaná jeho priorita:

$$\psi_i = 1 - P_i \cdot z, \text{ kde } z = \frac{\max_i - \min_i}{L-1}, \quad (6)$$

kde P_i je pozícia inštancie v sekvencii pohľadu, $\max_i = 1$, $\min_i = 0.1$ a L reprezentuje dĺžku sekvencie pohľadu. Potom sa pre každú inštanciu vypočítaná celková priorita:

$$\Psi = \sum_{i=1}^n \psi_i. \quad (7)$$

Nakoniec sú inštancie vizuálnych elementov zoradené zostupne (na základe celkovej priority) a v tomto poradí umiestnené do výslednej trendujúcej sekvencie pohľadu. O tom, či sa inštancia dostane do výslednej sekvencie, rozhoduje to, či má vyššiu hodnotu celkovej priority ako je minimálna hodnota celkovej priority inštancií, ktoré sa nachádzajú vo všetkých sekvenciách pohľadu.

V článku [23] bolo identifikované, že STA funguje lepšie, keď sú AOI zadefinované na vyššej úrovni granularity, teda viac AOI s menšou plochou. V prípade, ak sa na segmentáciu stránky používa automatický prístup, je preto vhodnejšie použiť na segmentáciu algoritmus VISP, a nie BOM.

4.4.3 Porovnanie algoritmov na identifikáciu spoločných sekvencií

Popísané algoritmy boli v článku [12] porovnané viacerými spôsobmi. Prvé porovnanie bolo v rámci toho, či spoločné sekvencie pohľadov jednotlivých algoritmov obsahujú elementy potrebné na splnenie úlohy v špecifickom poradí (v poradí potrebnom na splnenie úloh). Testovanie bolo vykonané na šiestich stránkach, pričom na každej sa plnila jedna úloha. Z výsledkov vyplýva, že elementy v špecifickom poradí obsahujú len spoločné sekvencie pohľadu vytvorené algoritmi STA a SPAM. Algoritmus eMine sa pomýlil len v jednom prípade zo šiestich a ostatné algoritmy vo väčšine prípadov správnu postupnosť vizuálnych elementov nevytvorili.

Keďže spoločná alebo trendujúca sekvencia pohľadu by mala byť podobná sekvenciám pohľadu, od ktorých bola odvodená, tak ďalším zaujímavým porovnaním je porovnanie podobnosti sekvencií pohľadu (algoritmami vytvorených voči pôvodným). Článok [12] testoval podobnosť sekvencií pohľadu na rovnakých dátach ako v predchádzajúcom prípade a došiel k nasledujúcim výsledkom: algoritmu eyePatterns sa nepodarilo identifikovať žiadne spoločné sekvencie pohľadu, zatiaľ čo Dotplot algoritmus mal problém len s niektorými. Algoritmus eMine dokázal identifikovať spoločné sekvencie pohľadu vždy, ale ich podobnosť s pôvodnými sekvenciami pohľadu bola nízka. Nakoniec algoritmy STA a SPAM dosahovali najvyššie podobnosti, avšak STA dosiahol vo väčšine prípadov väčšiu podobnosť.

Zo zistení tohto článku tým pádom vyplýva, že za najvhodnejší algoritmus považujú STA, o niečo horšie ale stále spoľahlivo fungujúce sú eMine, SPAM a Dotplot. EyePatterns vyšiel z výsledkov najhoršie.

Algoritmy na identifikáciu spoločnej sekvencie pohľadu sme porovnali aj na dátovej vzorke z Dňa otvorených dverí na FIIT STU, ktorý sa uskutočnil 16.12.2016. Dátová vzorka obsahuje záznamy sledovania pohľadu 54 participantov. Úlohou participantov bolo nájsť na stránke FIIT STU¹ minimálny počet bodov, aký je potrebný na prijatie do prvého ročníka bakalárskeho štúdia. Porovnávanými algoritmami boli Pozične váhovaný model (PBWM), Dotplot, STA a SPAM. Algoritmy Dotplot a PBWM sme implementovali v jazyku Python na základe článku [11], na algoritmus STA sme využili jeho implementáciu poskytnutú v článku [12] a na algoritmus SPAM bola použita knižica spmf². Algoritmy boli testované pri 5, 10 a 50 participantoch. Reprezentácie spoločných sekvencií pohľadu boli vytvárané z postupnosti znakov vytvorených na základe fixácií na AOI.

Tabuľka 1. Tabuľka zobrazujúca priemernú podobnosť spoločných sekvencií voči pôvodným sekvenciám pohľadu získaných z dátovej vzorky z DOD pomocou algoritmov PBWM, Dotplot, STA a SPAM .

algoritmus	počet participantov	Podobnosť voči spoločnej sekvencii pohľadu (%)
PBWM	5	17.7
	10	17.37
	50	22.25
Dotplot	5	2.69
	10	X
	50	X
STA	5	22.57
	10	15.96
	50	X
SPAM	5	16.26
	10	14.92
	50	15.46

Tabuľka 1 zobrazuje výsledky aplikovania algoritmov. Bunky označené znakom “X” sa algoritmom nepodarilo vypočítať. PBWM bolo spustené s parametrom N=3. Z tohto dôvodu algoritmus vrátil vo všetkých prípadoch sekvenciu pohľadu s dĺžkou 3, ktorá má pomerne vysokú podobnosť voči pôvodným sekvenciám.

Dotplot algoritmus nebol veľmi úspešný pri vytváraní spoločných sekvencií pohľadu. Podarilo sa mu to iba v prípade najnižšieho počtu participantov (5), aj to s dĺžkou jedného znaku, ktorý sekvenciu nevystihoval a dosiahol najnižšiu podobnosť voči pôvodným sekvenciám pohľadu 2.68%.

¹ <http://www.fiit.stuba.sk/>

² <http://www.philippe-fournier-viger.com/spmf>

STA našiel na rozdiel od Dotplotu spoločnú sekvenciu aj pre 10 participantov, pričom podobnosť jeho spoločných sekvencií bola podobná hodnotám podobnosti PBWM. SPAM algoritmus, našiel spoločnú sekvenciu aj pre skupinu 50-tich participantov, avšak jeho podobnosti boli výrazne nižšie ako v prípade PBWM.

Zistenia z našej analýzy sú o niečo iné zisteniam z článku [12]. Veľa algoritmov zlyháva pri väčšom počte sekvencií pohľadu. Najlepšie zvládol veľké množstvo sekvencií pohľadu PBWM, ktorý spolu so SPAM algoritmom fungoval dobre pri všetkých testovaných počtoch sekvencií pohľadu. Dotplot vo svojej úlohe úplne zlyhal na dátach z DOD. Pravdepodobne sa nevie vysporiadať s vyššou varianciou v dátach. STA algoritmus prekonal všetky algoritmy pri malom počte sekvencií pohľadu, ale môžeme vidieť, že pri väčšom počte sekvencií pohľadu začína mať problém.

Pri testovaní modelu by bolo dobré použiť všetky tu testované algoritmy. Podobnosť sa dá vypočítať pre všetky spoločné sekvencie pohľadu a nakoniec sa môžu použiť tie, ktorých spoločná sekvencia pohľadu je rozlišujúca.

4.5 Diskusia

V tejto kapitole sme si definovali čo sú to sekvencie pohľadu, že sa vo väčšine prípadov reprezentujú v podobe reťazcov znakov, pričom jednotlivé znaky reprezentujú navštívené miesta záujmu. Taktiež, že pre naše účely je vhodnejšie voliť si AOI ručne, a nie automaticky, lebo automatické prístupy zatiaľ nie sú úplne spoľahlivé.

Definovali sme si bežne používané metriky zo sledovača pohľadu, ktoré sa často používajú pri určovaní zručnosti človeka. Tieto metriky používame na definovanie modelu, voči ktorému budeme ďalšie výsledky porovnávať, a ktorý budeme rozširovať o ďalšie atribúty. Ďalšie atribúty sú metriky, ktoré ešte v súvislosti so štandardnými použité neboli (RQA) a o nami vytvorené atribúty (odvodené od spoločnej sekvencie pohľadu).

Nami vytvorené atribúty predstavujú podobnosť sekvencií pohľadu voči spoločnej sekvencii pohľadu expertov a začiatočníkov. Algoritmy na výpočet spoločných sekvencií pohľadu boli v práci porovnané a zistili sme, že niektoré z nich majú problém nájsť spoločnú sekvenciu pohľadu pri veľkom počte pôvodných sekvencií pohľadu. Považujeme však za vhodné použiť pri testovaní modelu všetky a až po vyhodnotení výsledkov modelu sa rozhodnúť, ktoré algoritmy preferovať.

5 Automatická identifikácia zručnosti používateľa na webe

Naším cieľom je vytvoriť automatický model rozpoznávania úrovne zručnosti používateľa so stránkou na základe nameraných dát zo sledovača pohľadu. Automatické spracovanie týchto dát môže byť realizované hľadaním opakujúcich sa vzorov alebo vlastností fixácií, sakád, veľkosti zreničky a vzdialenosti hlavy od monitoru.

5.1 Zadefinovanie problému

Ľudia vykonávajú na stránke zväčša dva typy úloh [11]:

- *prehľadávanie* – preskúmavanie obsahu stránky, pre ktoré je vhodné použiť analýzu sekvencií pohľadu,
- *vyhľadávanie* – s cieľom riadená činnosť, ktorá sa lepšie analyzuje spracovaním metrík fixácií a sakád (napr. čas do splnenia úlohy).

Keďže dopredu nevieme povedať, za akým účelom prišiel používateľ na stránku, zameranie sa na identifikáciu zručnosti používateľa len pre jeden z vyššie uvedených typov úlohy by spôsobilo vysokú chybovosť pri druhom type úlohy. Problém nastáva pri spojenom použití oboch typov analýzy pohľadu, keďže ich reprezentácia je pre jednotlivé merania nasledovná:

- *sekvencia pohľadu* – reťazec znakov s variabilnou dĺžkou,
- *metriky fixácií a sakád* – zoznam desatinných čísel s konštantnou dĺžkou.

Doterajšie prístupy analyzujúce vzory z dát zo sledovača pohľadu pracovali s jednotlivými typmi samostatne [5,7,21,41]. Autori článku [22] sa snažili brať do úvahy už okrem vzoru aj dĺžku fixácie. Algoritmus rozdelil sekvenciu pohľadu do časových intervalov a vypočítal podobnosti medzi analyzovanými postupnosťami fixácií a ich natrénovanými zhlukmi fixácií pre každý interval. Avšak tento prístup predstavuje pridanie iba jednej metriky, ktorá nie je veľmi charakteristická pre sekvenciu pohľadu.

V našej práci sme sa rozhodli, že budeme participantov rozlišovať do dvoch úrovní zručnosti – začiatočníkov a expertov. Možné prístupy a rozdelenie participantov do týchto úrovní sú popísané v kapitole 3.1.1. Za najvhodnejší spôsob sme si vybrali použitie kombinácie metrík založených na vykonávaní úloh - času plnenia úlohy a počet úspešne vykonaných úloh a sebahodnotenia pomocou dotazníku. Prvým dôvodom je, že ide o prístup používaný pri určovaní úrovne zručnosti aj oboznámenosti. Druhým dôvodom je, že kategorizovanie participantov do týchto skupín je možné robiť automaticky. Tento prístup bol použitý v druhom experimente (pozri kapitolu 7). V prvom experimente bola zručnosť používateľov zisťovaná len pomocou dotazníka. Na základe kapitoly 3.1.1 ide o stále používaný prístup. Avšak po prvom experimente sme zistili, že v prípade dotazníku je ťažké definovať, kedy je participant zručný, a kedy nie.

5.2 Navrhované riešenie

Plánom bolo vybrať nerozlišujúcejšie atribúty pre každý typ atribútov a následne z týchto atribútov vytvoriť spoločný model. Definovali sme si tri typy atribútov:

- *základné atribúty* – typické metriky používané pri analyzovaní dát zo sledovača pohľadu (odvodené od fixácií, sakád, vzdialenosti participanta od monitoru). Z týchto metrík sa však použijú iba tie, ktoré majú vysokú frekvenciu výskytu, aby ich bolo možné extrahovať už z prvých sekúnd zaznamenávania dát,
- *rekurentné atribúty* – rekurentné a reokurentné metriky,
- *atribúty odvodené od spoločnej sekvencie pohľadu* – nami navrhnuté atribúty, ktorých výpočet je popísaný nižšie.

Výpočet atribútov odvodených od spoločnej sekvencie pohľadu

Pre každého participanta sa identifikuje jeho sekvencia pohľadu v rámci každej úlohy. Pre každú úlohu sa vypočítajú spoločné sekvencie pohľadu pre obe skupiny participantov samostatne. Spoločné sekvencie pohľadu sú počítané 4 rôznymi algoritmami na výpočet spoločnej sekvencie pohľadu - PBWM, Dotplot, SPAM, STA. Následne sa pre každého participanta (v rámci jednej úlohy) vypočíta podobnosť jeho sekvencie pohľadu voči spoločným sekvenciám pohľadu oboch skupín participantov. Vo výsledku z toho vyplynie 8 atribútov – pre každý zo 4 spomínaných algoritmov existujú dva atribúty a to:

- podobnosť sekvencie pohľadu participanta voči spoločnej sekvencií pohľadu **začiatočníkov**,
- podobnosť sekvencie pohľadu participanta voči spoločnej sekvencií pohľadu **expertov**.

Model

Celkový model využíva 58 atribútov. Z dôvodu zachovania prehľadnosti zobrazuje Jednotlivé kroky navrhovaného riešenia sa nachádzajú na obrázku Obr. 8. Navrhovaný model sa skladá z tréningovej a testovacej fázy, pričom proces predspracovania je v oboch fázach veľmi podobný. Kroky tréningovej fázy sú nasledovné:

1. Paralelný výpočet základných atribútov, RQA atribútov a sekvencií pohľadu.
2. Zo sekvencií pohľadu sa pre skupinu začiatočníkov a pre skupinu pokročilých participantov vypočítajú spoločné sekvencie pohľadu.
3. Vypočíta sa priemerná podobnosť sekvencií pohľadu jednotlivcov každej skupiny voči spoločným sekvenciám oboch skupín (výsledkom sú už metriky).
4. Pre každú skupinu metrík sa aplikuje výber metrík. Ten prebieha výpočtom korelácií s predikovaným atribútom a výberom najviac korelovaných. Na konci sú odstránené vysoko vzájomne korelované atribúty.
5. Posledný krok tréningovania, v ktorom sa najrozlišujúcejšie metriky spoločne podieľajú na natrénovaní klasifikátora.

Tabuľka 2 zredukovaný zoznam atribútov odvodených od pohľadu (celý zoznam sa nachádza v technickej dokumentácii) a kompletný zoznam zvyšných typov atribútov.

Jednotlivé kroky navrhovaného riešenia sa nachádzajú na obrázku Obr. 8. Navrhovaný model sa skladá z trénovacej a testovacej fázy, pričom proces predspracovania je v oboch fázach veľmi podobný. Kroky trénovacej fázy sú nasledovné:

6. Paralelný výpočet základných atribútov, RQA atribútov a sekvencií pohľadu.
7. Zo sekvencií pohľadu sa pre skupinu začiatočníkov a pre skupinu pokročilých participantov vypočítajú spoločné sekvencie pohľadu.
8. Vypočíta sa priemerná podobnosť sekvencií pohľadu jednotlivcov každej skupiny voči spoločným sekvenciám oboch skupín (výsledkom sú už metriky).
9. Pre každú skupinu metrík sa aplikuje výber metrík. Ten prebieha výpočtom korelácií s predikovaným atribútom a výberom najviac korelovaných. Na konci sú odstránené vysoko vzájomne korelované atribúty.
10. Posledný krok trénovania, v ktorom sa najrozlišujúcejšie metriky spoločne podieľajú na natrénovaní klasifikátora.

Tabuľka 2. Najvýznamnejšie atribúty použité v modeli.

Odvodené od pohľadu
priem. a štand. odchýlka trvania fixácie
priem., štand. odchýlka, min. a max. rýchlosti sakády
priem. a štand. odchýlka trvania sakády
priem. a štand. odchýlka absolútnych uhlov sakád
priem. a štand. odchýlka relatívnych uhlov sakád
priem. a štand. odchýlka dĺžky sakády
Odvodené od zreničky
priem., štand. odchýlka, min. a max. veľkosti zreničky
priem., štand. odchýlka, min. a max. rýchlosti zreničky
Odvodené od vzdialenosti hlavy
priem., štand. odchýlka, min. a max. vzdialenosti hlavy
Recurrentná kvantifikačná analýza
rekurencia, determinizmus,
vrstvovosť, CORM
rekurentnosť, rekurentná miera
priem. X, priem. Y rekurencie
Spoločná sekvencia pohľadu
STA, PBWM, SPAM a Dotplot podobnosti k skupine expertov a začiatočníkov

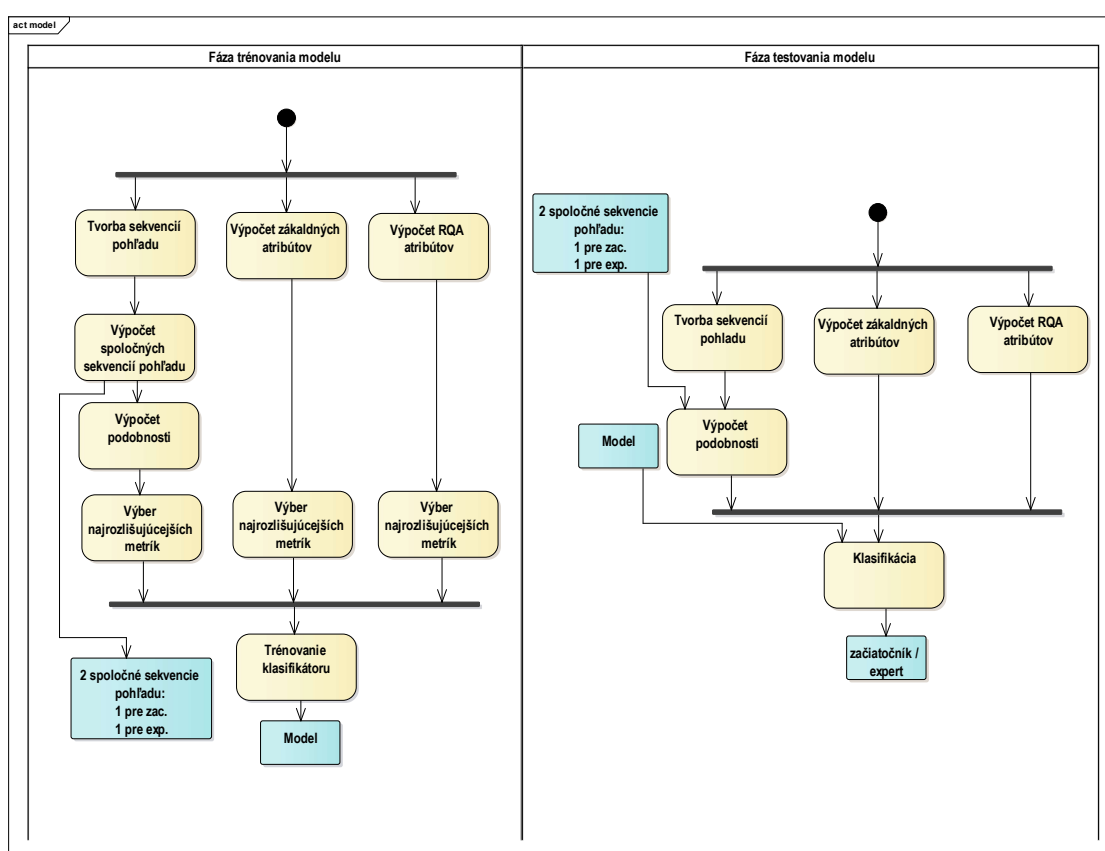
Testovacia fáza pozostáva z nasledovných krokov:

1. Paralelný výpočet základných atribútov, RQA atribútov a sekvencií pohľadu.
2. Vypočíta sa priemerná podobnosť sekvencií pohľadu jednotlivcov voči spoločnej sekvencii oboch skupín (výsledkom sú už metriky). Dáta jednotlivcov sú z

testovacej vzorky a použité spoločné sekvencie pohľadu pochádzajú z procesu tréningovania.

3. Metriky identifikované ako najrozlišujúcejšie vo fáze tréningovania sa vložia natréningovanému klasifikátoru ako vstup.

Na natréningovanie modelu, ktorý by bral do úvahy metriky fixácií a sakád a zároveň aj sekvencie pohľadu, navrhujeme transformovať sekvencie pohľadu do formátu metrík. Transformácie sú vykonané pomocou výpočtu hodnoty podobnosti medzi analyzovanou sekvenciou pohľadu a spoločnou sekvenciou pohľadu expertov a začiatočníkov získaných vo fáze tréningovania. Na výpočet podobnosti použijeme algoritmus založený na porovnaní textových reťazcov (angl. string-edit). Vypočítané podobnosti budú do modelu pridané priamo – dve metriky vo formáte desatinných čísiel.



Obr. 8. Jednotlivé kroky tréningovania a testovania modelu.

V našej metóde sme sa rozhodli použiť na výpočet spoločných sekvencií pohľadu algoritmy PBWM, Dotplot, STA a SPAM. Výpočet atribútov odvodených od spoločnej sekvencie pohľadu je však nezávislý od konkrétneho algoritmu a použiteľný aj s inými algoritmami na výpočet spoločnej sekvencie pohľadu. Metóda je taktiež nezávislá od použitého klasifikačného algoritmu. V práci používame klasifikačné algoritmy náhodný les, logistická regresia a algoritmus podporných vektorov (angl. Support vector machine, SVM). SVM je algoritmus založený na jadre, ktorý sa snaží pristupovať ku klasifikačnému problému hľadaním rozdeľujúcich nadrovin (angl. hyperplanes), čo najefektívnejším spôsobom [24]. Náhodný les je klasifikačný algoritmus, ktorý na rozhodovanie využíva

hlasovanie viacerých rozhodovacích stromov, pričom každý strom je natrénovaný na inej podmnožine dát [8]. Logistická regresia je algoritmus, ktorý sa používa na rozdelenie dát do dvoch tried. Je založený na prirodzenom logaritme pomeru šancí [37].

5.3 Diskusia

Ako hlavné prínosy nášho riešenia vidíme zavedenie tohto riešenia do domény webu a spracovanie sekvencií pohľadu do metrík vhodných pre algoritmus strojového učenia. Ako nedostatok tohto riešenia však stále vnímame, že pri vyhľadávacích úlohách neexistuje vždy iba jedna cesta na vyriešenie úlohy, a preto môžu vznikať veľké variácie aj medzi sekvenciami pohľadu participantov, ktorí sú približne rovnako zruční s danou stránkou.

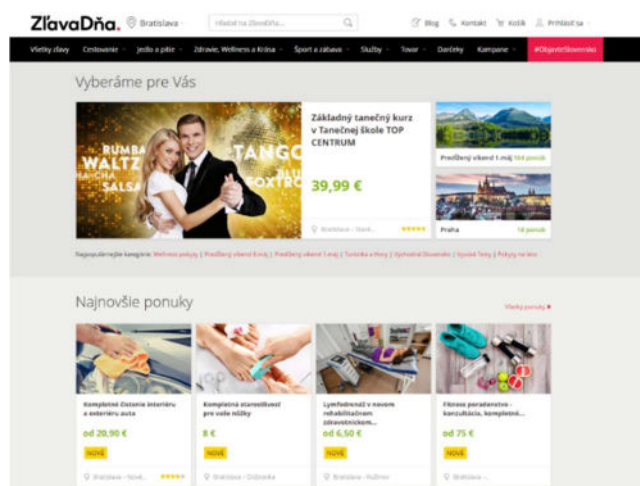
Možným rozšírením do budúcnosti by mohlo byť zavedenie identifikácie zručnosti na základe prvých sekúnd návštevy podstránky webového sídla. Motiváciou je, že ak má metóda byť použiteľná, tak je potrebné identifikovať zručnosť pred tým, ako používateľ prejde na inú podstránku sídla alebo sídlo opustí. Ďalšou možnosťou je, že by sa na začiatku v rámci skupiny začiatovníkov aj expertov identifikovali zhluky. Spoločná sekvencia pohľadu by sa vypočítala pre tieto zhluky a následne by sa aj podobnosti rátali voči spoločným sekvenciám zhlukov. Potom by pre každú skupinu mohlo vzniknúť toľko atribútov, koľko je zhlukov alebo by sa ako atribút použila najvyššia podobnosť k zhľuku v skupine.

6 Overenie rozdielov v sekvenciách pohľadu

V práci sme realizovali dva experimenty. Prvý experiment predstavoval kvantitatívnu štúdiu, v ktorej sme sa snažili odhaliť, či sú spoločné sekvencie pohľadu začiatočníkov na danej stránke odlišiteľne rozdielne od spoločnej sekvencie pohľadu skúsených používateľov (expertov). Podrobnejšie je rozpísaný v nasledujúcich podkapitolách. Návrh aj vyhodnotenie druhého experimentu sa nachádza v kapitole 7.

6.1 Návrh prvého experimentu

Experiment bol navrhnutý pre stránku Zľava Dňa³. Je to elektronický obchod na kúpu zľavových kupónov. Stránka má typickú štruktúru elektronického obchodu, s hlavným menu ponúkajúcim rôzne kategórie produktov, možnosť zobrazenia košíku a možnosť prihlásenia. Hlavná časť stránky zobrazuje produkty s ich názvom, ilustračným obrázkom a cenou (viď Obr. 9).



Obr. 9. Úkazka stránky ZľavaDňa.

Úlohy v experimente boli vytvorené tak, aby s nimi používateľ, ktorý stránku navštevuje pravidelne, nemal problém. Pozostávajú s bežných činností, ktoré sú typické pre túto typ stránky. Scenár priebehu experimentu sa nachádza v protokole prvého experimentu v prílohe (Príloha A – Protokol prvého experimentu). Konkrétne boli použité nasledovné úlohy:

1. Nachádzate sa na hlavnej stránke portálu ZľavaDňa a chcete skontrolovať najnovšie ponuky v okolí mesta Trenčín. Zistíte, aké sú 3 položky najlepšie spĺňajúce tieto podmienky. Názvy nájdených položiek zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.
2. Nachádzate sa na podstránke portálu ZľavaDňa v popise kupónu ponúkajúceho splav rieky Belá s videom, na čo by ste chceli ísť aj s ďalšími dvoma kamarátmi/kami. Skontrolujte, či je možné splavovať rieku v júni, a ak áno, pridajte 3 kupóny do košíka.

³ <https://www.zlavadna.sk/>

Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

3. Nachádzate sa na podstránke portálu ZľavaDňa v košíku a uvedomili ste si, že by rád išiel ešte jeden kamarát/ka. Pridajte do košíka ešte jeden kupón a prejdite do kroku vyplňania údajov. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.
4. Nachádzate sa na podstránke portálu ZľavaDňa v popise kupónu na návštevu aquaparku, kam by ste chceli s kamarátmi/kami ísť. Avšak pred tým, ako kupón objednáte, si chcete adresu aquaparku zapísať. Adresu aquaparku zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.
5. Nachádzate sa na podstránke portálu ZľavaDňa v popise kupónu na steak. Zistíte, aké je celkové a najnovšie hodnotenie tohto podniku. Obe hodnoty zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

Experiment bol vytvorený pomocou softvéru Tobii Pro Studio⁴. Pred každou úlohou sa participantovi zobrazila inštrukcia k danej úlohe. Prechody medzi úlohami (začatie a ukončenie úlohy), boli nastavené na tlačidlo F10. Úlohy neboli časovo obmedzené a v prípade neschopnosti vyriešenia úlohy bol participant inštruovaný, aby pokračoval na ďalšiu úlohu. Každý participant dostal odpovedový hárk, ktorý obsahoval kópiu inštrukcií ku každej úlohe, a taktiež miesto na zápis odpovedí na úlohy (na overenie, či úlohu splnil).

V experimente sme vyžadovali, aby participanti vyplnili aj dva dotazníky:

- *Dotazník na identifikovanie gramotnosti používateľov na webe*⁵ – bol prevzatý od Moniky Sanyovej [39]. Pozostával z 39 otázok a zisťoval, mieru webovej gramotnosti. Bolo potrebné ho vyplniť týždeň pred vykonaním experimentu, aby jeho vyplnenie neovplyvnilo výsledky experimentu.
- *Dotazník na identifikovanie doterajšej skúsenosti používateľa s webovým sídlom*⁶ – pozostával z troch otázok, ktoré identifikovali, ako často používateľ navštevuje danú stránku, ale aj stránky podobného charakteru. Znenie otázok dotazníku sa nachádza v prílohe Príloha A – Protokol prvého experimentu.

Na experiment boli pozvaní študenti a zamestnanci Fakulty informatiky a informačných technológií v Bratislave. U týchto účastníkov sa predpokladá vysoká webová gramotnosť, a preto by výsledky nemali byť veľmi ovplyvnené rôznymi úrovňami webovej gramotnosti. Aby toto tvrdenie bolo dokázateľné, používame vyššie spomenutý dotazník na test gramotnosti.

⁴ <https://www.tobiiipro.com/product-listing/tobii-pro-studio/>

⁵ <https://goo.gl/1Wgvk2>

⁶ https://crowdex.fiit.stuba.sk/experiment_management/56-oboznamenost-pouzivatela-s-webom

6.2 Priebeh experimentu

Experiment prebiehal vo Výskumnom centre používateľského zážitku a interakcií⁷ na Fakulte informatiky a informačných technológií v Bratislave, konkrétne v laboratóriu UX Group, ktoré umožňuje skupinových záznam až 20 používateľov naraz. Využili sme 15 počítačov, ktoré boli napojené na sledovač pohľadu Tobii X2-60 so 60Hz frekvenciou a využívali softvér Tobii Pro Studio.

Experiment bol súčasťou postupnosti troch experimentov, ktoré sa vykonávali priamo za sebou. Experimenty boli vytvorené ako súčasť jedného projektu v Tobii Pro Studio, pričom účastníci nemali medzi experimentami vytvorené prestávky. Prvý experiment testoval gramotnosť používateľov na webe a trval približne 25 min. Po ňom nasledoval náš experiment, ktorého dĺžka bola približne 10 min. Posledný bol 15-20 minút dlhý experiment, v ktorom účastníci klasifikovali obrázky do preddefinovaných kategórií.

Počas priebehu experimentu nastali dva problémy:

1. Dvomi účastníkmi sa počas experimentu spustila aktualizácia prehliadača Firefox, ktorá zapríčinila, že bolo experiment potrebné spustiť ešte raz. Účastníci následne pokračovali od časti, na ktorej skončili, avšak ich testovanie je zaznamenané ako dva záznamy.
2. Systém Crowdex⁸, v ktorom bol vytvorený *Dotazník na identifikovanie doterajšej skúsenosti používateľa s webovým sídlom*, neumožnil účastníkovi odovzdať dotazník, pravdepodobne z dôvodu nekompatibility systému Crowdexu s rozšírením do prehliadača od firmy Tobii, ktoré zabezpečuje komunikáciu so softvérom Tobii Pro Studio. Z tohto dôvodu boli účastníci požiadaní, aby odpovede na otázky zapísali na druhú stranu odpovedového hárku.

6.3 Vyhodnotenie prvého experimentu

Výsledky dotazníku na testovanie webovej gramotnosti potvrdzujú naše predpoklady, že väčšina účastníkov disponuje vysokou úrovňou webovej gramotnosti. Až 11 účastníkov z 15 dosahovalo najvyššiu úroveň webovej gramotnosti. Tria účastníci dosahovali druhú najvyššiu úroveň a jeden účastník dosahoval strednú úroveň, ktorá bola zároveň aj najnižšou dosiahnutou v celom experimente (viď Tabuľka 3).

Tabuľka 3. Popisuje počty účastníkov, ktorí dosahujú jednotlivé úrovne webovej gramotnosti.

Úroveň	Počet účastníkov
1	0
2	0
3	1
4	3
5	11

⁷ <https://www.uxi.sk>

⁸ <https://crowdex.fiit.stuba.sk>

Na základe vyhodnotenia v *dotazníku na identifikovanie doterajšej skúsenosti používateľa s webovým sídlom* sme zistili, že väčšina participantov stránku Zl'avaDňa nenavštevuje vôbec a 5 participantov navštevuje stránku menej ako raz za mesiac (viď Tabuľka 4). Jeden participant v dotazníku nevyplnil pravidelnosť, s akou navštevuje stránku, a preto bol z ďalších analýz dát vylúčený. Zo skutočnosti, že participantí sú zaradení do skupín najmenej pravidelných návštevníkov stránky, vyplýva, že rozdelenie participantov nebude vôbec rovnomerné a očakávané rozdiely medzi dvoma skupinami sú minimálne.

Tabuľka 4. Vyhodnotenie dotazníku zisťujúceho pravidelnosť návštev testovanej stránky participantami.

Pravidelnosť návštev stránky participantom	Počet participantov
stránku navštevuje	9
stránku navštevuje menej ako raz za mesiac	5
stránku navštevuje približne raz za mesiac	0
stránku navštevuje niekoľkokrát do mesiaca	0
stránku navštevuje niekoľkokrát do týždňa	0
nevyplnil	1

6.3.1 Podobnosti sekvencií pohľadu

Na začiatku procesu spracovania dát zo sledovača pohľadu sme vykonávali filtrovanie šumu dát. Toto bolo zabezpečené tak, že za fixáciu bol považovaný len taký segment fixačných dát, ktorý v sebe obsahoval aspoň jeden výskyt záznamu zo sledovača pohľadu, ktorý obsahoval validné dáta pre obe oči.

Analyzované sekvencie pohľadu participantov boli pre každého participanta v rámci každej úlohy zjednodušené na inštancie vizuálnych elementov rovnako, ako v implementácii algoritmu STA. Ovplyvnené boli sekvencie pohľadu, ktoré obsahovali postupnosti dvoch a viac rovnakých znakov za sebou. Tieto postupnosti boli zredukované na jeden znak (napr. AAABCCCA -> ABCA). Motiváciou za týmto úkonom bola snaha spraviť zo sekvencií pohľadu lepšie porovnateľné sekvenčné vzory.

Tabuľka 5 a Tabuľka 6 zobrazujú agregované štatistiky podobnosti sekvencií pohľadu pre skupinu participantov, ktorí stránku nenavštevujú a skupinu participantov, ktorí stránku navštevujú menej ako jedenkrát do mesiaca. Podobnosť bola vypočítaná pomocou algoritmu založenom na porovnaní textových reťazcov. Pre každého participanta bola vypočítaná podobnosť jeho sekvencie pohľadu voči ostatným participantom v skupine. Štatistiky boli vypočítané nad zoznamom týchto hodnôt.

Tabuľka 5. Štatistiky vypočítané nad hodnotami vzájomných podobnosti sekvencií pohľadu participantov, ktorí testovanú stránku nenavštevujú.

Participantí nenavštevujúci stránku				
Úloha	Minimálna podobnosť (%)	Maximálna podobnosť (%)	Priemerná vzájomná podobnosť (%)	Štandardná odchýlka podobnosti (%)
úloha 1	14.29	57.14	33.37	11.30
úloha 2	15.00	53.33	33.40	9.67
úloha 3	4.35	52.94	29.53	14.09
úloha 4	0.00	100.00	37.08	23.82
úloha 5	0.00	69.23	29.82	16.71

Tabuľka 6. Štatistiky vypočítané nad hodnotami vzájomných podobnosti sekvencií pohľadu participantov, ktorí testovanú stránku navštevujú menej ako raz za mesiac.

Participantí navštevujúci menej ako raz za mesiac				
Úloha	Minimálna podobnosť (%)	Maximálna podobnosť (%)	Priemerná vzájomná podobnosť (%)	Štandardná odchýlka podobnosti (%)
úloha 1	14.29	57.14	30.95	13.11
úloha 2	20.00	100.00	43.93	23.83
úloha 3	20.00	40.91	30.13	6.55
úloha 4	5.26	47.06	27.74	15.28
úloha 5	0.00	42.11	22.04	13.07

Z podobnosti identifikovaných vo vyššie uvedených tabuľkách nie je možné dokázať, že sekvencie porovnávaných skupín participantov sú rozlišujúce. V prípade dvoch úloh je priemerná podobnosť sekvencií pohľadu vyššia pre skupinu participantov, ktorí stránku navštevujú, avšak v prípade troch úloh je vyššia podobnosť pre skupinu nenavštevujúcu stránku. Podobné tvrdenia platia aj pre zvyšné štatistiky.

6.3.2 Podobnosti voči spoločnej sekvencií pohľadu

Keďže podobnosti identifikované v predchádzajúcej kapitole neboli veľmi rozlišujúce, rozhodli sme sa identifikovať priemernú hodnotu podobnosti sekvencií pohľadu participantov voči spoločnej sekvencií pohľadu. Na dosiahnutie týchto výsledkov bolo potrebné implementovať algoritmus na identifikáciu spoločnej sekvencie pohľadu. Aby ale meranie nebolo ovplyvnené nedostatkami konkrétneho algoritmu, rozhodli sme sa použiť

implementácie viacerých algoritmov a porovnať nimi dosahované výsledky. Algoritmy PBWN a Dotplot sme implementovali na základe popisu v článku [11].

STA algoritmus implementovaný autormi článku [12] je verejne dostupný a bol do projektu zapracovaný s jednou úpravou. Poskytovaná implementácia STA algoritmu totiž identifikuje AOI, do ktorej fixácia patrí tak, že určí vzdialenosť každého bodu AOI od bodu fixácie, tieto vzdialenosti agreguje a určí najbližšiu AOI, do ktorej fixáciu zaradí. Toto je jeden z dôvodov, prečo má algoritmus časovú zložitosť $O(n^6)$. Pre zrýchlenie algoritmu sme túto podmienku upravili tak, že sa použijú iba rohové body AOI a nie všetky, čo algoritmus zjednoduší a zachová si funkčnosť vo väčšine prípadov (problémy môžu vzniknúť pri prekryve AOI, ale také sme nepoužívali).

Tabuľka 7. Priemerná podobnosť sekvencií pohľadu participantov konkrétnej skupiny od spoločnej sekvencie pohľadu tej istej skupiny participantov.

Priemerná podobnosť spoločných sekvencií pohľadu od nameraných (%)				
úloha	algoritmus	Všetci participant	Participant nenavštevujúci stránku	Participant navštevujúci menej ako raz za mesiac
úloha 1	PBWM	25.72	24.30	27.86
	Dotplot	-	-	13.33
	STA	10.51	23.94	13.33
	SPAM	28.96	25.53	32.86
úloha 2	PBWM	13.57	11.42	15.55
	Dotplot	-	-	32.25
	STA	20.33	28.74	40.31
	SPAM	20.33	15.69	47.42
úloha 3	PBWM	20.57	16.31	19.55
	Dotplot	10.62	11.92	8.18
	STA	10.62	11.92	24.55
	SPAM	23.53	21.62	30.23
úloha 4	PBWM	17.90	22.61	15.23
	Dotplot	-	-	-
	STA	-	12.23	9.08
	SPAM	25.73	33.36	19.23
úloha 5	PBWM	24.25	31.41	13.36
	Dotplot	-	-	-
	STA	-	-	-
	SPAM	16.46	31.41	22.22

Posledným použitým algoritmom na výpočet spoločnej sekvencie pohľadu bol SPAM algoritmus, ktorého implementácia je poskytnutá vo forme JAR knižnice⁹, na ktorú sme implementovali adaptér.

Výsledky aplikovania jednotlivých algoritmov zobrazuje Tabuľka 7. Číselné hodnoty vyjadrujú priemernú podobnosť sekvencií pohľadu participantov danej skupiny voči spoločnej sekvencii pohľadu danej skupiny participantov. Skupiny participantov sú: všetci participant, nenavštevujúci participant a participant navštevujúci webové sídlo menej ako raz za mesiac.

Z výsledkov nie je možné určiť skupinu participantov, ktorá by dosahovala výrazne vyššie hodnoty podobnosti. Toto môže byť spôsobené tým, že rozdiel medzi zručnosťou participantov v skupinách je minimálny (identifikované v *dotazníku na identifikovanie doterajšej skúsenosti používateľa s webovým sídlom*). Druhým dôvodom by mohla byť existencia viacerých ciest, ktorými sa daná úloha riešiť, čo môže spôsobiť výrazne rozdiely aj u skupiny skúsenejších používateľov. Ďalším dôvodom by mohli byť situácie, aké sme identifikovali v prípade druhej úlohy, pri ktorej niektorí participant nesplnili prvú časť úlohy (skontrolovať dátum) a rovno prešli k druhej časti.

6.4 Tvorba modelov

Plánovanou stratégiou bolo vytvoriť čo najlepšie modely pre každý typ atribútov a následne vytvoriť spoločný model zostavený z atribútov, ktoré boli identifikované ako najvhodnejšie v predchádzajúcich modeloch. Popis jednotlivých typov atribútov sa nachádza v kapitole 5.

6.4.1 Model zo základných atribútov

Model bol trénovaný a testovaný nad dátami všetkých úloh. Z dát bol odstránený participant tester10, lebo nevyplnil dotazník a tester18, ktorému chýbajú dáta až do takej miery, že pri prvej úlohe neboli žiadne validné dáta zaznamenané.

Základné atribúty boli vypočítané s využitím EMDAT knižnice [44], ktorá poskytuje výpočet veľkého množstva metrík odvodených od fixácií, sakád, AOI, vzdialenosti od monitoru, ale aj metrík odvodených od klikov kláves a myšky. Pre naše účely sme knižnicu upravili tak, aby bola schopná spracovať s jedným vstupným tsv súborom pre všetkých participantov. Ďalej sme obmedzili metriky, ktoré má počítať, len na metriky odvodené od fixácií, sakád, zreničky a od vzdialenosti od monitoru. Nakoniec bola za účelom pohodlného používania pomocou príkazového riadku pridaná možnosť volať knižnicu s nami definovanými argumentami (taskID, participantsToProcess). Keďže EMDAT bola implementovaná pre nižšiu verziu jazyku python, akú používame my, bolo potrebné taktiež implementovať adaptér pre túto knižnicu.

⁹ <http://www.philippe-fournier-viger.com/spmf>

Keď sme dosiahli stav, v ktorom sme mali atribúty pre každého participanta vo všetkých úlohách vypočítané, mohli sme identifikovať metriky, ktoré sú pre určenie zručnosti najcharakteristickejšie. Na identifikovanie týchto atribútov sme použili koreláciu všetkých metrík voči predikovanému atribútu, pričom metriky závislé od celkovej dĺžky plnenia úloh boli ignorované. Absolútne hodnoty korelácií zobrazuje Tabuľka 8.

Tabuľka 8. Zobrazuje atribúty najviac korelované s predikovaným atribútom. Atribúty, ktoré sú závislé od celkovej dĺžky plnenia úlohy, sú odstránené.

Absolútna hodnota korelácie	Názov metriky
0.3500	priem. trvanie sakády
0.3074	priem. dĺžka sakády
0.3038	štand. odchýlka trvania sakády
0.2843	koncová vzdialenosť hlavy
0.2843	min. vzdialenosť hlavy
0.2713	max. trvanie sakády.
0.2470	štand. odchýlka veľkosti zreničky

Analyzované dáta pozostávali z 13 participantov, pričom 9 patrili skupiny participantov, ktorí webové sídlo Zľava Dňa nenavštevujú a 4 patrili do skupiny, ktorá navštevuje webové sídlo menej ako raz za mesiac. Keďže participant plnil 5 úloh a každé riešenie úlohy berieme ako jednu vzorku, tak sa dostávame k celkovému počtu vzoriek 65.

Ďalším krokom bolo vytvorenie modelu pre automatické rozlišovanie zručnosti na webe. Pre túto úlohu sme vybrali SVM najmä kvôli jeho robustnosti a schopnosti klasifikovať aj lineárne neseparovateľné dáta. Na overenie výsledkov modelu bolo použité náhodné vzorkovanie, avšak vždy bolo dodržané pravidlo, podľa ktorého sa dáta jedného participanta môžu nachádzať len v jednej z týchto množín. To zabezpečuje, aby sa model netestoval na dátach participanta, na ktorom sa aj trénoval. Výsledky vyhodnotenia modelu zobrazuje Tabuľka 9.

Tabuľka 9. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelu natrénovaného na všetkých atribútoch. Presnosť a pokrytie boli vypočítavané pre triedu expertov.

	Správnosť	Presnosť	Úplnosť
Základný model	0.63	0.63	1.0

Z výsledkov vyhodnotenia vyplýva, že model funguje trochu lepšie ako náhodne. Avšak model predpovedal pre všetky testovacie vzorky, že sa jedná o zručných participantov, takže fungoval rovnako ako triviálny väčšinový klasifikátor (angl. majority classifier). V ďalších podkapitolách sú popísané kroky, ktorými sme sa snažili zlepšiť výsledky modelu.

Normalizácia a odstránenie neúspešných úloh

Prvým pokusom o zlepšenie bola normalizácia dát. Na dáta sme aplikovali z-score normalizáciu tak, aby nebola ovplyvnená testovacími dátami. Čiže v každej iterácii krížovej validácie boli pre každý atribút vypočítané agregované hodnoty (priemer, min, max) pre trénovaciu množinu a následne boli tieto agregované hodnoty použité na normalizáciu trénovacej aj testovacej množiny.

Druhým pokusom o vylepšenie bolo odstrániť dáta, v ktorých participanti úlohu nesplnili. Vyhodnotením nahrávok z plnenia úloh sme zistili, že tester24 nesplnil druhú a piatu úlohu a tester 21 a 23 nesplnili druhú úlohu. Dôvodom prečo sme sa pokúsili tieto dáta nebrať do úvahy je, že pri nesplnených úlohách nebolo jasné či daný participanti úlohu pochopili. Druhým dôvodom je, že tí čo úlohu nesplnili, zavádzali príliš veľkú mieru variability, s čím sme sa momentálne pri tak malom počte participantov nevedeli vysporiadať. Taktiež sa jednalo o zručných participantov, ktorých odstránením sme dosiahli vyváženejšiu dátovú vzorku. Menovaní participanti boli následne odstránení zo všetkých úloh. Keby sme participantov neodstránili zo všetkých úloh, tak by sa proces rozdeľovania participantov do trénovacej a testovacej množiny počas krížovej validácie výrazne skomplikoval.

V prípade tohto merania bolo použitých 6 participantov, ktorí webové sídlo Zľava Dňa nenavštevujú a 4 participanti, ktoré dané webové sídlo navštevujú menej ako raz za mesiac. Spolu bolo analyzovaných 50 záznamov. Vyhodnotenie výsledkov oboch modelov zobrazuje tabuľka Tabuľka 10.

Tabuľka 10. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelov voči triede expertov. Presnosť a pokrytie boli vypočítavané pre triedu expertov.

	Správnosť	Presnosť	Úplnosť
Základný model po normalizácii	0.37	0.2	0.4
Základný model po odstránení neúspešných úloh	0.18	0.05	0.2
Základný model po odstránení neúspešných úloh a normalizácii	0.59	0.79	0.55

Výsledky vypovedajú o tom, že normalizácia úspešnosť modelu zhoršuje (buď predpovedá iba zručných participantov alebo iba začiatočníkov). Vo väčšine prípadov však predikuje začiatočníkov, kvôli čomu správnosť modelu výrazne poklesla. Aj keď odstránenie participantov, ktorí nesplnili úlohu dva, modelu hneď nepomohlo (klasifikoval všetky vzorky ako začiatočníkov), tak aplikovanie normalizácie zvýšilo úspešnosť modelu na vyššiu ako náhodnú, pričom model už predikoval obe triedy. Základný model po odstránení neúspešných úloh a normalizácií budeme považovať za referenčný model. Dôvodom je, že je očistený od zašumených dát a pracuje s veľkou škálou bežne používaných metrík odvodených zo sledovača pohľadu (od fixácií, sakád, veľkosti zreničky, vzdialenosti hlavy od monitoru).

Výpočet nových korelácií

Keďže odstránenie určitých participantov pomohlo správnosti modelu, rozhodli sme sa vypočítať korelácie pre takto osekanú dátovú vzorku nanovo. Atribúty najviac korelovaných atribútov obsahuje Tabuľka 11.

Tabuľka 11. Zobrazuje metriky najviac korelované s predikovaným atribútom. Metriky, ktoré sú závislé od celkovej dĺžky plnenia úlohy sú odstránené.

Absolútna hodnota korelácie	Názov metriky
0.8261	priem. veľkosť zreničky
0.6592	koncová veľkosť zreničky
0.6550	max. veľkosť zreničky
0.6295	začiatočná veľkosť zreničky
0.6224	priem. vzdialenosť hlavy
0.6195	štand. odchýlka veľkosti zreničky
0.6054	koncová vzdialenosť hlavy

Atribút *enddistance* sa nachádza v oboch zoznamoch, a keďže v aktuálnejšom má viac ako dvojnásobnú hodnotu korelácie, môže byť dôvodom, prečo došlo k zlepšeniu správania modelu. Nové najviac korelované atribúty majú hodnotu absolútnej korelácie o 30-50% vyššiu ako korelácia dát všetkých použiteľných participantov. Preto budeme participantov *tester21*, *tester23*, *tester24* ďalej považovať za odľahlé hodnoty.

Najviac korelované atribúty s predikovanou hodnotou sú odvodené od veľkosti zreničky a vzdialenosti hlavy od monitoru. Tieto výsledky nie sú veľmi prekvapivé, keďže veľkosť zreničky priamo súvisí s kognitívnou záťažou a participant, ktorí stránku vidia prvýkrát, musia na splnenie úlohy vynaložiť viac kognitívneho úsilia. Taktiež to potvrdzuje výsledky článku [44], v ktorom zistili, že metriky odvodené od veľkosti zreničky a od vzdialenosti hlavy od monitoru majú pozitívny vplyv na modely rozlišujúce fázu nadobúdania zručnosti, v ktorej sa participant nachádza.

Ako ďalší krok sme overili správanie modelu, ktorý používa atribúty identifikované v tejto kapitole (Tabuľka 11) a na tréning a testovanie využíva 50 záznamov identifikovaných v podkapitole riešiacej odstránenie neúspešných úloh. Výsledky zobrazuje Tabuľka 12.

Tabuľka 12. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelov vytvorených nad najviac korelovanými atribútmi. Presnosť a pokrytie boli vypočítavané pre triedu expertov.

PupHeadDist = Model s atribútmi odvodenými od zreničky a vzdialenosti hlavy od monitoru.

	Správnosť	Presnosť	Úplnosť
PupHeadDist	0.75	0.83	0.78
PupHeadDist + normalizácia	0.7	0.75	0.62

Výsledky (viď Tabuľka 12) sú o niečo lepšie ako referenčný model. Správnosť modelu stúpala o 6 percentuálnych bodov, presnosť o 4 percentuálnych bodov, pokrytie o 23 percentuálnych bodov. Zaujímavým zistením je taktiež fakt, že normalizácia dát sa pri týchto typoch dát prejavuje negatívne.

Redukcia vzájomne korelovaných atribútov

Keďže atribúty identifikované v predchádzajúcej kapitole sú agregácie len nad metrikami odvodenými od veľkosti zreničky a vzdialenosti hlavy od monitoru, predpokladali sme, že sú medzi nimi vysoké vzájomné korelácie. Z tohto dôvodu sme sa rozhodli z vysoko vzájomne korelovaných atribútov použiť len tie, ktoré majú vyššiu hodnotu korelácie voči predikovanému atribútu. Po aplikovaní prahu 0,5 na výšku vzájomných korelácií zostali len 2 atribúty *meanpupilsiz*e a *meandistance*. Čiže pre každý typ atribútov zostal jeden najvýraznejší. Pre výsledné atribúty bol natrénovaný model s aplikovaním normalizácie (Tabuľka 13).

Tabuľka 13. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelu s redukovaným počtom atribútov na základe vzájomných hodnôt korelácií. Presnosť a pokrytie boli vypočítavané pre triedu expertov.

	Správnosť	Presnosť	Úplnosť
PupHeadDist + reduk. att. + normalizácia	0.73	0.78	0.66
PupHeadDist + reduk. att.	0.75	0.82	0.78

Aj napriek redukcii atribútov zostala správnosť modelu rovnaká. Tento model budeme ďalej označovať ako *redukovaný model základných atribútov* (RMZA).

6.4.2 Model rekurentných metrík

Pre korektné porovnanie výsledkov bol model pre rekurentné metriky testovaný na rovnako zredukovaných dátach ako RMZA. Spolu bolo analyzovaných 50 záznamov. Dáta boli získané z piatich úloh, pričom pre každú úlohu bolo použitých 10 participantov (6 participantov ešte webové sídlo nenavštívilo, 4 participantov navštevujú webové sídlo menej ako raz za mesiac).

Rekurentné metriky boli získané použitím knižnice vytvorenej autormi článku [29]. Ich implementácia¹⁰ poskytovala rekurentné aj reokurentné metriky, pričom spolu sa jedná o 11 atribútov. Korelácie jednotlivých atribútov s predikovaným atribútom zobrazuje Tabuľka 14.

Korelácie s predikovaným atribútom sú výrazne nižšie ako v prípade základných metrík. Toto zistenie sa prejavilo aj na úspešnosti modelu, ktorý využíval dané atribúty. Úspešnosť viacerých nastavení modelu zobrazuje Tabuľka 15. Tieto metriky sa na dátach z tohto experimentu neprejavili ako vhodné na určovanie zručnosti používateľa na webe.

¹⁰ <http://geovis.cis.rit.edu/dataProcessing.html>

Tabuľka 14. Absolútne hodnoty korelácií rekurentných atribútov voči predikovanému atribútu.

Absolútna hodnota korelácie	Názov metriky
0.0990	rekurentná miera
0.0934	stred masy opakovania (CORM)
0.0651	reokurentná miera
0.0580	reokurencia
0.0283	priemerná rekurencia X
0.0127	vrstvosť
0.01	štandardná odchýlka rekurencie Y
0.0099	determinizmus
0.0057	štandardná odchýlka rekurencie X
0.0000	priemerná rekurencia Y
0.0000	rekurencia

Korelácie s predikovaným atribútom sú výrazne nižšie ako v prípade základných metrík. Toto zistenie sa prejavilo aj na úspešnosti modelu, ktorý využíval dané atribúty. Úspešnosť viacerých nastavení modelu zobrazuje Tabuľka 15. Tieto metriky sa na dátach z tohto experimentu neprejavili ako vhodné na určovanie zručnosti používateľa na webe.

Tabuľka 15. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelov trénovaných na rekurentných atribútoch. Presnosť a pokrytie boli vypočítavané pre triedu expertov. Za najviac korelované atribúty sú považované tie, ktoré majú koreláciu s predikovaným atribútom vyššiu ako 0.05.

	Správnosť	Presnosť	Pokrytie
Všetky atribúty	0.41	0.75	0.39
Všetky atribúty + Normalizácia	0.34	0.64	0.20
Najviac korelované atribúty	0.34	0.685	0.14
Najviac korelované atribúty + Normalizácia	0.57	0.69	0.60

6.4.3 Model spoločnej sekvencie pohľadu

Rôzne nastavenia modelu boli testované na rovnako zredukovaných dátach ako v prípade modelu základných metrík a modelu rekurentných metrík.

Implementácia jednotlivých algoritmov je popísaná v kapitole 6.3.2 a spôsob použitia spoločných sekvencií pohľadu na výpočet atribútov sa nachádza v kapitole 5.2. Absolútne hodnoty korelácií atribútov s predikovaným atribútom popisuje Tabuľka 16. Podľa výsledkov sa dá usúdiť, že najviac rozlišujúce spoločné sekvencie pohľadu vytvára algoritmus SPAM, keďže korelácie jeho výsledkov sa umiestnili najvyššie s hodnotami 0.1162 a 0.0878. Celkovo sú však hodnoty korelácií nízke (podobné hodnoty ako v prípade rekurentných atribútov).

Tabuľka 16. Zobrazuje absolútne hodnoty korelácií metrík odvodených od spoločnej sekvencie pohľadu voči predikovanému atribútu. Skupina 0 je skupina participantov, ktorý webové sídlo ešte nenavštívili a skupina 1 sú participant, ktorí webové sídlo navštevujú menej ako raz za mesiac.

Absolútna hodnota korelácie	Názov metriky	Popis metriky
0.1162	SPAM_similarityTo1	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 1 (vypočítané pomocou SPAM)
0.0878	SPAM_similarityTo0	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 0 (vypočítané pomocou SPAM)
0.0567	PBWM_similarityTo0	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 0 (vypočítané pomocou PBWM)
0.0270	STA_similarityTo0	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 0 (vypočítané pomocou STA)
0.0226	PBWM_similarityTo1	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 1 (vypočítané pomocou PBWM)
0.0156	STA_similarityTo1	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 1 (vypočítané pomocou STA)
0.0073	DOTPLOT_similarityTo1	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 1 (vypočítané pomocou DOTPLOT)
0.0032	DOTPLOT_similarityTo0	podobnosť sekvencie pohľadu voči spoločnej sekvencií pohľadu skupiny 0 (vypočítané pomocou DOTPLOT)

Výsledky úspešnosti rôznych nastavení modelu zobrazuje Tabuľka 17. Najlepšia správnosť modelu bola dosiahnutá pri použití najviac korelovaných atribútov. Správnosť tohto modelu je len o 2% nižšia ako úspešnosť RMZA, čím prejavuje vysoký potenciál pre rozlišovanie skupín používateľov s rôznymi úrovňami zručnosti

Tabuľka 17. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelov trénovaných na atribútoch odvodených od spoločnej sekvencie pohľadu. Presnosť a pokrytie boli vypočítavané pre triedu expertov. Za najviac korelované atribúty sú považované tie, ktoré majú koreláciu s predikovaným atribútom vyššiu ako 0.05.

	Správnosť	Presnosť	Úplnosť
Všetky atribúty	0.72	0.71	0.88
Všetky atribúty + Normalizácia	0.71	0.71	0.92
Najviac korelované atribúty	0.73	0.72	0.88
Najviac korelované atribúty + Normalizácia	0.69	0.731	0.75

6.4.4 Kombinované modely

Aj keď samostatné modely založené na atribútoch odvodených od RQA a spoločnej sekvencie pohľadu fungovali s náhodnou úspešnosťou, rozhodli sme sa najviac korelované atribúty použiť v spoločnom modeli. Ako základný model, ktorý sa bude rozširovať, sme si zvolili redukovaný model základných atribútov (RMZA). Postupne sme do modelu pridávali atribúty odvodené od RQA a spoločnej sekvencie pohľadu. Výsledky zobrazuje Tabuľka 18.

Tabuľka 18. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) kombinovaných modelov. Presnosť a pokrytie boli vypočítavané pre triedu expertov.

	Správnosť	Presnosť	Úplnosť
RMZA	0.75	0.82	0.78
RMZA + RQA atribúty	0.62	0.076	0.57
RMZA + atribúty spol. sek. pohľadu	0.75	0.83	0.74
RMZA + RQA atribúty + atribúty spol. sek. pohľadu	0.61	0.70	0.57

Rozšírenie modelu nemalo v žiadnom z prípadov pozitívny efekt na správnosť modelu, aj keď pri rozšírení RMZA o atribúty odvodené od spol. sek. pohľadu žiadnym spôsobom úspešnosť modelu neovplyvnilo.

6.5 Diskusia

Prvým zaujímavým zistením pri analýze základných atribútov odvodených z dát zo sledovača pohľadu je, že odstránenie vzoriek, v ktorých participanti nespĺnili úlohu (vzoriek, ktoré do dát zavádzali väčšiu variabilitu), naštartovalo rozlišovaciu schopnosť modelu medzi skupinami začiatocníkov a zručných participantov. Druhým zaujímavým zistením je fakt, že atribúty odvodené od veľkosti zreničky a od vzdialenosti hlavy od monitoru sa javia ako najviac rozlišujúce. Toto zistenie potvrdzuje nedávny výskum [44], kde im pri rozlišovaní zručnosti vyšli podobné zistenia. Hlavným dopadom tohto zistenia je, že takýto model by mohol byť prenositeľný medzi stránkami, keďže tieto atribúty nie sú na rozdiel od sekvencií pohľadu viazané výlučne k jednému rozhraniu. Posledným zistením je fakt, že model využívajúci atribúty spol. sek. pohľadu dosahuje v istej konfigurácii správnosť len o 2% nižšiu ako RMZA.

Výsledky rozšírenia modelu základných atribútov vypovedajú o tom, že RQA atribúty môžu model zhoršovať a atribúty odvodené od sekvencií pohľadu správnosť modelu nemenia. Avšak zistenia sú namerané na príliš malej dátovej vzorke, aby sme ich mohli tvrdiť s istotou, a preto sme sa tieto zistenia pokúsili overiť v ďalšom experimente.

Navrhnutý experiment mal niekoľko nedostatkov. Najväčším nedostatkom prvého experimentu je nevyváženosť analyzovaných vzoriek, keďže väčšina participantov dosahuje veľmi nízku úroveň zručnosti s testovanou stránkou. Vieme však pri nej aspoň zaručiť, že participanti mali približne rovnakú úroveň webovej gramotnosti, čiže namerané

rozdiely by mali byť spôsobené najmä doménovo závislou zručnosťou používateľov. Ďalej sa experimentu zúčastnilo pomerne málo participantov (len 15). Experiment bol taktiež vykonaný na veľmi jednoduchom rozhraní (väčšina nováčikov splnila všetky úlohy). Vznikali tiež otázky či je určovanie zručnosti len na základe sebahodnotenia v dotazníku postačujúcou metrikou zručnosti. Aby sme adresovali tieto nedostatky, realizovali sme druhý experiment opísaný v nasledujúcej kapitole.

7 Overenie rozlíšiteľnosti zručnosti

Na základe skúseností a zistení z prvého experimentu opísaných v predchádzajúcej kapitole sme sa rozhodli zrealizovať aj druhý experiment. Prvotným cieľom experimentu bolo overiť nasledovné hypotézy:

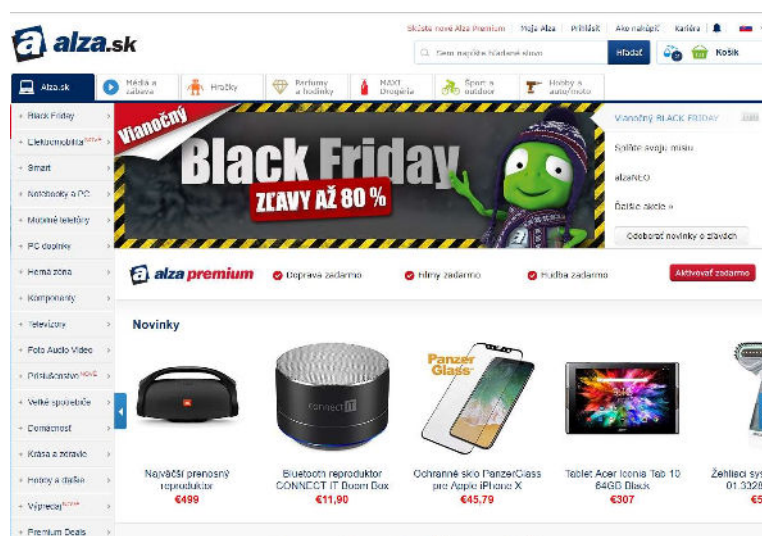
1. Zručných používateľov stránky je možné automaticky odlíšiť od menej zručných na základe metrík odvodených zo sledovania pohľadu (odvodené od fixácií, sakád, veľkosti zreničky atď.).
2. Atribúty odvodené od spoločnej sekvencie pohľadu majú pozitívny vplyv na úspešnosť modelu pre automatické rozlišovanie zručnosti používateľa s webom.

Druhotným cieľom ďalšieho experimentu bolo overiť predchádzajúce zistenia na väčšej a lepšie vyváženej vzorke dát.

7.1 Návrh experimentu

Za testovanú doménu sme si vybrali Alza.sk¹¹. Jedná sa o internetový obchod s elektronikou. Používateľské rozhranie je o niečo komplikovanejšie ako v prípade Zľavy Dňa (viac kategórií, porovnávanie produktov, dovoz na predajne, viď Obr. 10). Testované webové sídlo sme menili z dvoch dôvodov:

- Alza.sk má zložitejšie používateľské rozhranie, čo by mohlo zväčšiť rozdiely medzi zručnými participantami a začiatocníkmi.
- predpokladáme, že sídlo Alza.sk má viac zručných participantov ako ZľavaDňa.sk (v predchádzajúcom experimente bol problém nájsť zručných používateľov).



Obr. 10. Ukážka stránky Alza.sk.

Počas experimentu sú participantom zadané dva typy úloh - *rozlišovacie* a *obyčajné*. Rozlišovacie sú ťažšie úlohy, ktoré slúžia na určenie, či je participant zručný na webovom

¹¹ <https://www.alza.sk>

sídle, alebo nie, a obyčajné úlohy slúžia na overenie našich hypotéz. Celkový popis experimentu sa nachádza v prílohe (viď Príloha B – Protokol druhého experimentu).

Rozdelenie používateľov do skupiny zručných a začiatočníkov je vykonané na základe ich rozdelenia s využitím informácií z dotazníku 1 (viď nižšie) a ich časov a úspešnosti plnenia rozlišujúcich úloh. Podrobnejšie popis rozdelenia participantov sa nachádza v kapitole 7.3.1. Na začiatku a konci testovania bolo participantom zobrazených 9 kalibračných obrázkov, aby bolo možné neskôr určiť presnosť sledovača pohľadu.

Rozlišovacie úlohy

1. Nachádzate sa na hlavnej stránke portálu Alza. Zistíte stav objednávky s číslom 169446698. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
2. Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zistíte do kedy by bol produkt doručený na predajňu Bratislava - Petržalka, keby si ho kúpili teraz. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
3. Nachádzate sa na hlavnej stránke portálu Alza. Nájdite najlacnejšiu ponuku rozbalených produktov (produkty vo výpredaji). Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Obyčajné úlohy

1. Nachádzate sa na podstránke so zoznamom digitálnych fotoaparátov. Nájdite fotoaparát nad 200 eur s najnižšou cenou. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
2. Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zistíte či daný fotoaparát poskytuje funkciu detekcie tváre. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
3. Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zobrazte si porovnanie tohto fotoaparátu s fotoaparátom Canon PowerShot G9 X Black pomocou alza nástroja na porovnanie. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
4. Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zapište dve pozitíva tohto produktu, ktoré produkt má na základe hodnotenia. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.
5. Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Skontrolujte, či je možné produkt vrátiť do 14 dní. Ak áno, tak pridajte produkt do košíka. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Rovnako ako v prípade prvého experimentu, je experiment vytvorený v softvéri Tobii Pro Studio. Úlohy nie sú časovo obmedzené. Ak participant úlohu nevie splniť, je inštruovaný, aby pokračoval na ďalšiu úlohu. Všetky prechody medzi inštrukciami a úlohami boli

nastavené na tlačidlo F10. Každý participant dostal odpoveďový hárok v papierovej podobe. Hárok obsahoval kópiu otázok, a priestor na vyplnenie úloh.

V experimente sme použili 3 dotazníky:

1. *Dotazník na identifikovanie oboznámenosti používateľa s doménou Alza.sk* – identifikoval, ako často používateľ navštevuje danú stránku, ako často navštevuje podobné stránky, a koľko nákupov na danej stránke v poslednej dobe spravil. Bol participantom predložený na konci experimentu. Jeho účelom bolo poskytnutie informácií o oboznámenosti participanta, ktorá sa spolu s informáciami o rozlišujúcich úlohách dá použiť na určenie zručnosti participanta. Dôvodom, prečo sme si vybrali oboznámenosť, a nie zručnosť je, že oboznámenosť sa dá jednoduchšie identifikovať ako zručnosť (na základe dotazníka). Zároveň je priamo úmerná zručnosti používateľa, takže hodnota zručnosti je veľmi blízka hodnote oboznámenosti.
2. *Dotazník na identifikovanie oboznámenosti používateľa s internetovými obchodmi vo všeobecnosti* – bol použitý na identifikovanie spôsobu, akým participant pristupuje k riešeniu typických problémov v doméne internetových obchodov. Participantovi bol predložený bezprostredne po dokončení úloh a slúži na získanie lepšieho popisu dát.
3. *Dotazník na identifikovanie gramotnosti používateľov na webe* – zisťoval mieru webovej gramotnosti. Rovnaký dotazník ako v prvom experimente. Participantovi bol zaslaný týždeň pred realizáciou experimentu a slúži na získanie lepšieho popisu dát. Jeho možným použitím je overenie, či existuje nejaká závislosť medzi gramotnosťou a zručnosťou človeka na webovom sídle.

7.2 Realizácia navrhnutého experimentu

Pre vyššie navrhnutý experiment sa vykonalo pilotné tetovanie, na základe ktorého sa identifikovali a opravili nedostatky návrhu. Následne sme zrealizovali ostré testovanie, ktoré prebiehalo v 5 piatich sedeniach a zúčastnilo sa ho spolu 57 participantov.

7.2.1 Pilotné testovanie

Pilotné testovania prebiehali vo Výskumnom centre používateľského zážitku a interakcií¹² na Fakulte informatiky a informačných technológií v Bratislave. Prvé dve testovania boli vykonané v laboratóriu UX Group, ktoré je zamerané podrobné zaznamenávanie pohľadu jedného používateľa. Testovanie bolo vykonané na počítači využívajúcom Tobii Pro Studio a napojenom na Tobii TX-300 so 300Hz frekvenciou.

Prvý participant bol na základe *dotazníku na identifikovanie oboznámenosti používateľa s doménou Alza.sk* so stránkou alza.sk oboznámený. Testovanie trvalo 20 minút, pričom rozlišovacie aj obyčajné úlohy zvládol bez problémov. Druhý participant na stránke Alza.sk ešte nebol a jeho testovanie trvalo 25 minút. Obyčajné úlohy zvládol, aj keď nie tak rýchlo ako oboznámený participant. Z troch rozlišovacích úloh nesplnil jednu (zo zadania

¹² <https://www.uxi.sk>

nepochopil, čo má robiť). Dve úlohy splnil, avšak pri jednej z nich bola cesta k riešeniu príliš komplikovaná.

Počas týchto pilotných testovanií sme identifikovali nasledovné nedostatky:

- V inštrukciách na papieri nebula možnosť zapisovať odpovede. Zistili sme, že ju musíme pridať, aby sme boli schopní vyhodnocovať splnenie a nesplnenie úlohy.
- Pred každým testovaním je potrebné prečistiť históriu prehliadača - pri druhom testovaní bolo zapamätané porovnanie produktov.
- Niekedy spustenie úlohy po inštrukcii v Tobii Studiu trvá niekoľko sekúnd. Do inštrukcií je potrebné pridať pokyn, aby v tomto čase participanti na nič neklikali.
- Niektoré obyčajné úlohy by bolo dobré zmeniť:
 - Úloha 4 - Zapísanie nevýhod produktu zmeniť na výhody (výhody sú jednoslovné, nevýhody celé vety).
 - Úloha 2 - výraz “súčasťou balenia” sa obom participantom mýlil s príslušenstvom, tu by bolo vhodné nájsť v popise inú informáciu, ktorú hľadať.
- Posledné nedostatky sa týkali nedostatočných úvodných inštrukcií a preklepu v dotazníku.

Druhé pilotné testovanie bolo vykonané v laboratóriu UX Group, ktoré umožňuje skupinových záznam až 20 používateľov naraz. Experimentu sa zúčastnili štyria participanti, pričom sa v tejto skupine nachádzali zruční požívatelia aj začiatčníci. Testovanie bolo vykonané s upravenými úlohami na základe prvých pilotných testov a s pridanými vlastnými kalibračnými obrazovkami. Experiment trval spolu s inštrukciami pol hodinu. Počas experimentu sme odhalili nasledovné nedostatky:

- Pred experimentom je potrebné participantov upozorniť, aby si na telefónoch vypli zvonenia.
- Pri kontrolovaní úloh treba skontrolovať aj produkty, ktoré sa vyhľadávajú (už neexistovala 100% zhoda v názve, iba podobný názov).
- Pred experimentom skontrolovať aj hardware – jednému participantovi nešlo skrolovať pomocou kolieska na myške.

7.2.2 Ostré testovanie

Ostré testovanie prebiehalo podobne ako druhá časť pilotného testovania v laboratóriu UX Group, ktoré umožňuje skupinový záznam až 20 používateľov naraz. Experimentu sa zúčastnilo 57 participantov, ktorých vek sa pohyboval v rozmedzí 17-49 rokov a pochádzali z rôznych študijných a profesijných oblastí (študenti stredných škôl, informatiky, medicíny, učiteľstva a pracujúce osoby). Testovanie bolo realizované v piatich sedeniach, pričom väčšina bola realizovaná v rovnakom čase (z dôvodu zachovania rovnakého osvetlenia miestnosti a podobnej úrovne únavy participantov). Na vykonanie testovania boli použité počítače, ktoré boli napojené na sledovač pohľadu Tobii X2-60 so 60Hz frekvenciou a využívali softvér Tobii Pro Studio.

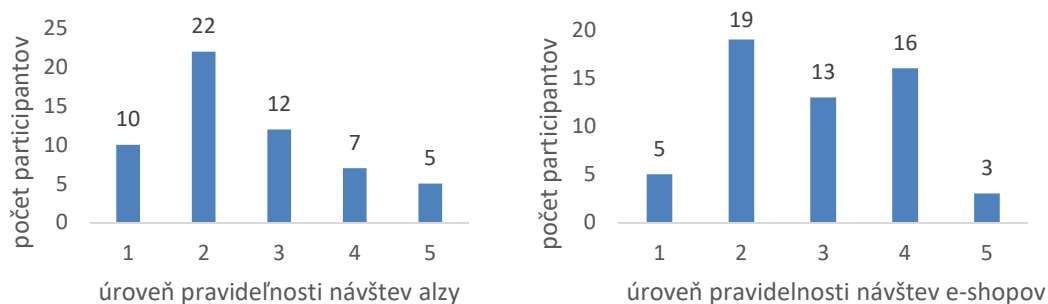
Priemerná dĺžka vykonávania experimentu sa pohybovala v rozmedzí 15-20 minút, v závislosti od zručnosti používateľa. Participanti tieto úlohy vykonávali súčasne, avšak poradie úloh bolo náhodné (použitá technika vyvažovania, angl. counterbalancing), aby sa minimalizovali vplyvy poradia úloh na ich úspešnosť a rýchlosť plnenia.

Počas experimentu nastalo niekoľko problémov:

1. U siedmych participantov nastal problém s filtrom. V týchto prípadoch sa omylom spustilo testovanie v staršej verzii prehliadaču Mozilla Firefox, ktorý daný filter nepodporoval. Táto chyba ovplyvnila plnenie úloh 3 a 4. Keďže v prípade úlohy 3 sa vyhodnocovali dáta sledovača pohľadu len do času prvého zobrazenia filtra, tak v konečnom dôsledku bola ovplyvnená len úloha 4, pri ktorej boli tieto inštancie označené ako nesplnená úloha.
2. Z dôvodu náhodného poradia zobrazovania úloh sa niektorí participanti pomýlili pri vyplňovaní úloh do odpoved'ového hárku a buď výsledok úlohy zapísali do zlého poľa alebo zhodnotili, že úloha sa na papieri nenachádza a vynechali ju. Tieto úlohy boli tiež označené ako nesplnené.
3. V prípade stredoškolákov len traja z 10-tich vyplnili tretí dotazník, ktorý im bol poslaný po testovaní.

7.3 Opis nazbieraných dát

Všetci participanti vyplnili *dotazník na identifikovanie oboznámenosti používateľa s doménou Alza.sk* a *dotazník na identifikovanie oboznámenosti používateľa s internetovými obchodmi vo všeobecnosti*. Najdôležitejšie výsledky sú zobrazené na Obr. 11. Prvý histogram zobrazuje sebahodnotenie participantov na to, ako často navštevujú stránku Alza.sk. V každej úrovni pravidelnosti návštev sa nachádza aspoň 5 participantov, pričom graf dodržiava tvar Gaussovho rozloženia, ktoré je mierne vychýlene doľava. Z toho vyplýva, že experimentu sa zúčastnilo viac ľudí, ktorí Alzu nenavštevujú, ako tých, ktorí sú pravidelní zákazníci. V prípade pravidelnosti návštev e-shopov (Obr. 11, histogram vpravo), je vychýlenie doľava, len minimálne, čiže pravidelných návštevníkov e-shopov je približne rovnako ako tých, ktorí e-shopy navštevujú len málo.

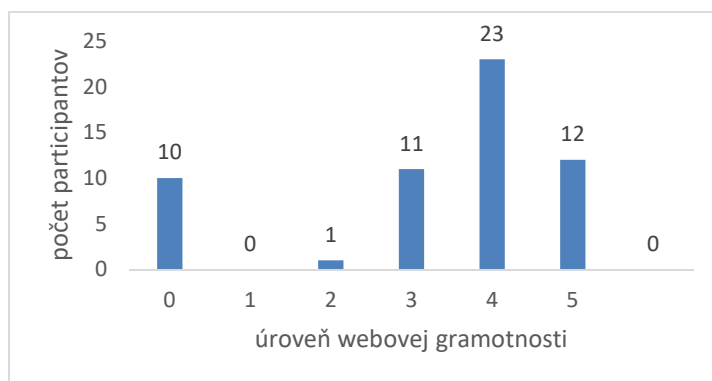


Obr. 11. Vyhodnotenie pravidelnosti návštev Alzy (vľavo) a e-shopov (vpravo) na základe prvého dotazníku. X-ová os vyjadruje pravidelnosť návštev (1 = nenavštevujem, 5 = navštevujem niekoľkokrát do týždňa). Y-ová os vyjadruje počet participantov.

Dotazníku na identifikovanie gramotnosti používateľov na webe sa zúčastnilo 47 ľudí zo všetkých 57. Úrovně webovej gramotnosti zúčastnených zobrazuje Obr. 12. Tvar grafu predstavuje Gaussove rozloženie vysoko vychýlené doprava. Väčšina účastníkov sa dá preto považovať za webovo zručných. Zaradenie participantov do jednotlivých úrovni gramotnosti bolo vykonané na základe definície v práci [39]. Každému participantovi bola vypočítaná váhovaná percentuálna úspešnosť a následne bol zaradený do úrovne použitím prahov (Tabuľka 19).

Tabuľka 19. Spôsob určovania úrovne webovej gramotnosti prevzatý z [39].

Úspešnosť v dotazníku [%]	0-15	15-35	35-65	65-85	85-100
Úroveň webovej gramotnosti	1	2	3	4	5



Obr. 12. Zobrazuje počet participantov patriacich do jednotlivých úrovni webovej gramotnosti. Do úrovne 0 sú zaradení participant, ktorí sa dotazníku nezúčastnili.

Pre všetkých participantov bola meraná správnosť a presnosť sledovača pohľadu pomocou 9-tich kalibračných obrazoviek na začiatku a na konci experimentu. Hodnoty správnosti a presnosti boli vypočítané pomocou metód popísaných v špecifikácii testovania Tobii Studia [43]. Priemerná nameraná hodnota správnosti bola 1,99° a priemerná hodnota presnosti 0,13°. Hodnota správnosti sledovača pohľadu je tým pádom výrazne vyššia ako v špecifikácii zariadenia (správnosť 0,4° a presnosť 0,34° [43]). To znamená, že rozdiely medzi miestom, kam sa človek pozerá, a ktoré je zaznamenané, budú väčšie. Na druhej strane hodnota presnosti je nižšia ako v špecifikácii, čo je pozitívne.

Úroveň obtiažnosti jednotlivých úloh sa dá do určitej miery odvodiť od počtu neúspešných riešení jednotlivých úloh (Tabuľka 20). Vysoký počet neúspešných riešiteľov úlohy 2 je spôsobený najmä tým, že komplikované úspešné riešenie úlohy - participant zistili požadovanú informáciu až pri vyplňaní objednávky - nebolo uznané (najmä kvôli tomu, že išlo o rozlišovaciu úlohu). V úlohe 4 vzniklo niekoľko neúspešných riešení kvôli problémom s prehliadačom Mozilla Firefox (viď podkapitulu 7.2.2). Po zvážení týchto okolností sa dajú za najviac obtiažné úlohy považovať úloha 2 a 5.

Tabuľka 20. Počet neúspešných riešení pre jednotlivé úlohy.

Úloha	Počet nesplnených úloh
úloha1	5
úloha2	9
úloha3	11
úloha4	11
úloha5	12
úloha6	5
úloha7	5
úloha8	6

Vzhľadom na zistenia z predchádzajúceho experimentu (v ktorom sme identifikovali, že použitie neúspešných úloh pre dátovú analýzu nie je vhodné), sme sa rozhodli tieto dáta z ďalších analýz vylúčiť. Konkrétnymi dôvodmi sú:

- medzi nesplnenými úlohami sú prípady, v ktorých ľudia nepochopili samotné zadanie úlohy, a preto ani nemali veľkú šancu dôjsť k úspešnému riešeniu,
- sú medzi nimi prípady, keď participant riešil úlohu len do 10 sekúnd a preskočil ju,
- sú medzi nimi opačné prípady, v ktorých participanti riešia úlohy 3-5 minút,
- z celkového pohľadu prinášajú tieto riešenia do dát príliš veľkú mieru variability.

Na rozdiel od predchádzajúceho experimentu nie sú z dát odstránené všetky riešenia participanta, ak má medzi nimi aspoň jedno neúspešné riešenie, ale len konkrétne inštancie riešení.

7.3.1 Rozdelenie participantov

Na rozdelenie participantov do skupiny expertov a začiatočníkov sme sa rozhodli použiť informácie nepoužívané v procese tréningu a testovania modelu. Medzi použiteľné informácie patria dáta z dotazníkov a opisné dáta o úlohách (úspešnosť plnenia úloh a časy plnenia úloh). Najskôr sme sa snažili rozdeliť participantov iba podľa úloh alebo iba podľa odpovedí dotazníkov. Keď sme ale testovali, nakoľko si tieto dve rozdelenia zodpovedajú, tak sme zistili, že pri skupine expertov určených dotazníkom a skupine začiatočníkov určených úlohami vzniká výrazný prekryv (prekryv vznikol aj keď boli experti určený úlohami a začiatočníci dotazníkom). Tieto prekryvy súvisia najmä s tým, že korelácie medzi opisnými dátami o úlohách a hodnotením v dotazníku nie sú príliš vysoké (viď Tabuľka 21). Korelácie sú však ľahko vysvetliteľné – čím kratší čas na úlohu, tým je človek viac expert, a ak človek úlohu splnil, je väčšia pravdepodobnosť, že je expert.

Tabuľka 21. Korelácie opisných dát úloh voči hodnoteniu návštevnosti Alzy z dotazníku.

uloha1 úspešnosť	0.0274
uloha1 čas	-0.3055
uloha2 úspešnosť	-0.1217
uloha2 čas	-0.1578
uloha3 úspešnosť	0.03656
uloha3 čas	-0.0106
počet úspešných úloh	-0.0506

Z tohto dôvodu sme sa rozhodli použiť lineárnu kombináciu informácií z dotazníku a informácií o úlohách. Vytvorili sme skórovaciu funkciu:

$$\text{celkové skóre} = \text{skóre úloh} * \text{skóre dotazníku}, \quad (7)$$

kde skóre dotazníku je hodnota od 1 po 5 (vyjadrujúca ako často človek navštevuje Alzu - 1 = nenavštevujem, 5 = navštevujem niekoľkokrát do týždňa). Skóre úloh je vypočítané ako suma skóre pre jednotlivé úlohy. Skóre pre úlohu je určené nasledovne:

- ak participant splnil úlohu a zároveň bol rýchlejší ako väčšina - 2 body,
- ak participant splnil úlohu a zároveň bol pomalší ako väčšina - 1 bod,
- ak participant nesplnil úlohu – 0 bodov.

Keďže za dotazník je možné získať 1 až 5 bodov a za úlohy 0 až 6 bodov, tak váhy vplyvov jednotlivých zložiek sú približne rovnaké. Po aplikovaní skórovacej funkcie sme zadelili participantov do skupiny expertov, ak mali skóre väčšie ako 75% ostatných participantov a do skupiny začiatocníkov, ak mali skóre menšie ako 25% ostatných participantov. Týmto rozdelením sme získali 13 začiatocníkov a 9 expertov, v ktorých už nevznikali prekryvy ako v predchádzajúcom prípade.

7.4 Výber klasifikačného algoritmu

V tejto kapitole sú identifikované najvýraznejšie atribúty modelu a porovnané klasifikátory SVM, náhodný les a logistická regresia.

7.4.1 Identifikácia najvýraznejších atribútov

Najvýznamnejšie atribúty sme hľadali nad množinou dát identifikovanou v kapitole 7.3.1. Všetky používané atribúty sú nezávislé od celkovej dĺžky plnenia úlohy. Korelácie medzi atribútmi a predikovanou hodnotou vypočítane pomocou Spearmanovho algoritmu zobrazuje Tabuľka 22.

Tabuľka 22. Hodnoty korelácie atribútov k predikovanému atribútu, ktorých absolútna hodnota je vyššia ako 0,25.

Hodnota korelácie	Názov metriky
0.5	min. vzdialenosť hlavy
0.5	koncová vzdialenosť hlavy
0.49	priem. vzdialenosť hlavy
0.41	začiatočná vzdialenosť hlavy
-0.35	reokurencia
0.35	min. rýchlosť sakády
0.31	STA_podobnosť_k_1
-0.29	rekurencia
-0.29	štand. odchýlka X rekurencie
0.29	priem. rýchlosť sakády
-0.29	priem. X rekurencie
-0.28	štand. odchýlka Y rekurencie
-0.28	priem. Y rekurencie
0.27	štand. odchýlka vzdialenosti hlavy

Z výsledkov vyplýva, že najvyššiu hodnotu korelácie dosahujú atribúty odvodené od vzdialenosti hlavy od monitoru, pričom zo zvyšujúcou sa vzdialenosťou rastie aj zručnosť používateľa. Tento výsledok z časti potvrdzuje zistenia z predchádzajúceho experimentu, kde sa identifikovali za najvýznamnejšie atribúty tie, ktoré sú odvodené od veľkosti zreničky a vzdialenosti hlavy od monitoru. Preto je však prekvapivá absencia atribútov odvodených od veľkosti zreničky. Medzi najrozlišujúcejšími atribútmi sú tiež vysoko zastúpené atribúty rekurencie. Zo zápornej korelácie môžeme usúdiť, že čím menej vykonávajú ľudia spätné fixácie na oblasti, ktoré už videli, tým viac sú zruční. Menej zastúpenými sú atribúty odvodené od spoločnej sekvencie pohľadu a od sakád.

Najviac korelované atribúty sme porovnávali s atribútmi identifikovanými ako najdôležitejšie algoritmom náhodný les. Medzi identifikovanými atribútmi bola vysoká miera prekryvu, pričom pri aplikovaní 10-násobnej krížovej validácie a náhodného lesu sme zistili, že poradie najvýraznejších atribútov sa vôbec nemení, dochádza iba k jemným výkyvom hodnôt váh. Z tohto dôvodu veríme, že identifikované atribúty môžu byť považované za najrozlišujúcejšie. Pre ďalšie potvrdenie sme identifikovali aj váhy atribútov pri natrénovaní klasifikátoru logistickej regresie. Tam už vznikali trochu väčšie rozdiely voči zoznamu korelácií. V rovnako dlhom zozname ako pri koreláciách pribudli na začiatku atribúty odvodené od zreničky. Do zoznamu pribudli taktiež atribúty odvodené od uhlov medzi sakádami. Tieto atribúty však nenahradili najviac korelované atribúty, iba ich variácie agregácií. Atribúty odvodené od vzdialenosti hlavy od monitoru, rekurencie a spoločnej sekvencie pohľadu sa v zozname stále nachádzali akurát, už nie vo väčšine svojich agregovaných podôb. Keďže sa tieto atribúty nachádzajú aj medzi váhami logistickej regresie, tak sme sa ďalej rozhodli použiť atribúty identifikované koreláciou ako najrozlišujúcejšie.

7.4.2 Porovnanie klasifikátorov

Ďalší krok pozostával z identifikácie klasifikačného algoritmu, ktorý si vie najlepšie poradiť s nami riešeným problémom, na ktorom môžeme následne testovať rôzne konfigurácie modelu. Porovnávanými algoritmami boli algoritmus podporných vektorov (SVM), náhodný les a logistická regresia. Priemerné výsledky klasifikačných algoritmov po 10-násobnej krížovej validácii zobrazuje Tabuľka 23.

Tabuľka 23. Správnosť, presnosť a úplnosť modelu využívajúceho všetky atribúty spriemerované po 10-násobnej krížovej validácii. Presnosť a pokrytie boli vypočítavané pre triedu začiatocníkov.

	Správnosť	Presnosť	Úplnosť
SVM	0.58	0.58	0.80
Náhodný les	0.58	0.59	0.68
Logistická regresia	0.65	0.66	0.63

Z troch algoritmov fungoval najhoršie algoritmus SVM, ktorý predikoval vo všetkých iteráciách krížovej validácie jednu a tú istú triedu. Keďže sa jednalo o triedu začiatocníkov, tak má model klamlivo vysokú úplnosť. O niečo lepšie fungoval náhodný les, ktorý dosahoval podobnú správnosť a presnosť ako SVM, avšak vždy bol schopný predikovať obe triedy. Za najvhodnejší klasifikačný algoritmus môžeme považovať logistickú regresiu, ktorá má oproti ostatným výrazne vyššiu správnosť, pričom podobne ako náhodný les nepreferuje žiadnu triedu.

7.5 Tvorba modelov

V tejto sekcii sme porovnali úspešnosť modelu pre každý typ atribútov a vytvorili spoločný model zostavený z atribútov, ktoré boli identifikované ako najvhodnejšie v predchádzajúcich modeloch, za účelom identifikácie vplyvov jednotlivých typov atribútov na úspešnosť modelu. Na natrénovanie všetkých modelov boli použité dáta participantov identifikovaných v kapitole 7.3.1.

7.5.1 Model zo základných atribútov

Základné metriky boli podobne ako v prípade prvého experimentu vypočítané s využitím upravenej EMDAT knižnice [44] – podrobnejší popis sa nachádza v kapitole 6.4.1.

Pre vypočítané hodnoty atribútov bol natrénovaný model nad všetkými atribútmi a nad najviac korelovanými atribútmi (najviac korelované sú tie, ktorých abs. hodnota korelácie je vyššia ako 0.25). Pre vyhodnotenie klasifikačnej úlohy bola použitá 10-násobná krížová validácia, pričom rozdeľovanie bolo vykonané v rámci používateľov, nie úloh tak, aby nemohla vzniknúť situácia, v ktorej sa model trénuje a zároveň overuje na dátach toho istého participanta. Najrozlišujúcejšie identifikované atribúty sú: *min. vzdialenosť hlavy*, *koncová vzdialenosť hlavy*, *priem. vzdialenosť hlavy*, *zač. vzdialenosť hlavy*, *min. rýchlosť*

sakády, priem. rýchlosť sakády, štand. odchýlka vzdialenosti hlavy. Výsledky týchto modelov aj s obmenami využívania normalizácie zobrazuje Tabuľka 24.

Tabuľka 24. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelu po 10-násobnej krížovej validácii. Presnosť a úplnosť boli počítané pre skupinu začiatočníkov.

Model základných atribútov	Správnosť	Presnosť	Úplnosť
všetky atribúty	0.67	0.67	0.70
všetky atribúty + normalizácia	0.70	0.69	0.67
atribúty s kor. > 0.25	0.72	0.73	0.71
atribúty s kor. > 0.25 + normalizácia	0.72	0.60	0.64

Z výsledkov vyhodnotenia vyplýva, že všetky alternatívy majú správnosť výrazne vyššiu ako náhoda, pričom najhoršou alternatívou je použitie všetkých atribútov bez predspracovania. Požitie normalizácie je vhodné iba v prípade použitia všetkých atribútov, lebo normalizácia najviac korelovaných atribútov pomáha modelu lepšie predikovať expertov na úkor začiatočníkov. Najlepšou variantnou je použitie najviac korelovaných atribútov bez predspracovania, kde model dosahuje správnosť 72%. Model využívajúci základné atribúty s koreláciou vyššou ako 0.25, bude ďalej označovaný ako MZA (model základných atribútov).

7.5.2 Model rekurentných metrík

Rekurentné metriky boli získané podobne ako v prípade prvého experimentu použitím knižnice vytvorenej autormi článku [29]. Na základe korelácie s predikovanou hodnotou boli v kapitole 7.4.1 za najrozlišujúcejšie vybraté *reokurencia*, *rekurencia*, *štand. odchýlka rekurencie X*, *priem. rekurencie X*, *štand. odchýlka rekurencie Y*, *priem. rekurencie Y*. Hodnoty úspešnosti natrénovaných variant modelov zobrazuje Tabuľka 25. Všetky výsledky boli získané krížovou validáciou definovanou v kapitole 7.5.1.

Tabuľka 25. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelu po 10-násobnej krížovej validácii. Presnosť a úplnosť boli počítané pre skupinu začiatočníkov.

Model rekurentných atribútov	Správnosť	Presnosť	Úplnosť
všetky atribúty	0.55	0.69	0.49
všetky atribúty + normalizácia	0.54	0.57	0.44
atribúty s kor. > 0.25	0.57	0.63	0.48
atribúty s kor. > 0.25 + normalizácia	0.57	0.68	0.61

Aj napriek tomu, že z tohto typu atribútov bolo identifikovaných pomerne veľa vysoko korelovaných atribútov, tak správnosť modelu dosahuje len o niečo lepšie výsledky ako náhoda. Táto skutočnosť môže byť spôsobená vysokou vzájomnou koreláciou rozlišujúcich atribútov. Najlepšou z alternatív je model využívajúci rozlišujúce atribúty a normalizáciu.

7.5.3 Model spoločnej sekvencie pohľadu

Implementácia algoritmov na výpočet spoločnej sekvencie pohľadu je popísaná v kapitole 6.3.2 a spôsob výpočtu metrík odvodených od spoločných sekvencií pohľadu je uvedený v kapitole 5.2. Z celkového počtu 8-mich atribútov sa medzi najrozlišujúce dostal v kapitole 7.4.1 iba jeden: *STA podobnosť k 1* (podobnosť sekvencií pohľadu k spoločnej sekvencií expertov, identifikovanej algoritmom STA). Úspešnosť variantov klasifikačných modelov zobrazuje Tabuľka 26.

Tabuľka 26. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelu po 10-násobnej krížovej validácii. Presnosť a úplnosť boli počítané pre skupinu začiatovníkov.

Model atribútov odvodených od spoloč. sekvencie pohľadu	Správnosť	Presnosť	Úplnosť
všetky atribúty	0.60	0.64	0.74
všetky atribúty + normalizácia	0.60	0.59	0.64
atribúty s kor. > 0.25	0.56	0.63	0.62
atribúty s kor. > 0.25 + normalizácia	0.56	0.57	0.73

V prípade atribútov odvodených od spoločnej sekvencie pohľadu sa javí ako vhodnejšie použitie všetkých 8 atribútov, keďže redukcia počtu atribútov spôsobuje zníženie správnosti modelov. Zaujímavým pozorovaním je však, že model dosahuje lepšie hodnoty správnosti ako model rekurentných metrík, čo potvrdzuje rozlišovaciu schopnosť nami navrhnutých atribútov.

7.5.4 Kombinované modely

Po vyhodnotení jednotlivých typov atribútov samostatne sme začali kombinovať typy atribútov do modelov. Toto bolo vykonané za účelom identifikácie prínosu konkrétnych typov k úspešnosti modelu. Keďže doteraz najlepšiu správnosť dosahoval MZA, tak sme rozširovali tento model. Na vytvorenie modelov boli použité len najrozlišujúcejšie atribúty identifikované v kapitole 7.4.1, bez aplikovania normalizácie. Správnosť, presnosť a úspešnosť kombinovaných modelov zobrazuje Tabuľka 27.

Tabuľka 27. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) kombinovaných modelov po 10-násobnej krížovej validácii. Presnosť a úplnosť boli počítané pre skupinu začiatovníkov.

Kombinované modely	Správnosť	Presnosť	Úplnosť
MZA	0.72	0.73	0.71
MZA + RQA atribúty	0.69	0.73	0.66
MZA + atribúty spol. sek. pohľadu	0.77	0.69	0.75
MZA + RQA atribúty + atribúty spol. sek. pohľadu	0.69	0.70	0.68

Z porovnania výsledkov vyplýva, že RQA zhoršujú správnosť modelu. Toto zistenie potvrdzuje výsledky identifikované v prvom experimente, kde pridanie RQA atribútov

spôsobovalo rovnaký efekt. Na druhej strane atribúty spol. sekvencie pohľadu zvyšujú správnosť modelu. Konkrétne sa jedná o jeden atribút: *STA_podobnosť_k_1*. Tento typ atribútov na dátach z prvého experimentu nemal na model základných atribútov žiaden efekt (nezvyšovali ani neznižovali správnosť modelu). Teraz však môžeme zhodnotiť, že atribúty odvodené od spoločnej sekvencie pohľadu pozitívne vplyvajú na správnosť modelu.

7.5.5 Automatické generovanie a redukcia atribútov

Model využívajúci všetky definované typy atribútov sme sa usilovali vylepšiť najskôr redukciou vysoko vzájomne korelovaných atribútov (ďalej už iba redukcia) a následne automatickým generovaním atribútov (angl. automatic feature engineering, ďalej už iba generovanie). Redukcia s aplikovaním prahu na vzájomnú koreláciu 0.5 ponechala iba atribúty: *min. vzdialenosť hlavy*, *reokurencia*, *priem. rýchlosť sakády*, *stand. odchýlka vzdialenosti hlavy*. Výsledky modelu využívajúceho tieto atribúty zobrazuje Tabuľka 28 na prvých dvoch riadkoch. Pri generovaní a následnej redukcii musel byť pozitívny prah 0.6, aby po redukcii nezostal len jeden najvýraznejší atribút. Výslednými atribútmi boli: *min. vzdialenosť hlavy * koncová vzdialenosť hlavy*, *priem. rýchlosť sakády * STA_podobnosť_k_1*, *stand. odchýlka vzdialenosti hlavy * reokurencia*. Úspešnosť modelu s týmito atribútmi zobrazuje Tabuľka 28 na posledných dvoch riadkoch.

Tabuľka 28. Zobrazuje priemerné hodnoty správnosti (accuracy), presnosti (precision) a úplnosti (recall) modelov po 10-násobnej krížovej validácii. Presnosť a úplnosť boli počítané pre skupinu začiatočníkov. Pri redukcii je použitý prah na vzájomnú koreláciu 0.5 a pri generovaní + redukcii je použitý prah 0.6.

Model s generovanými a redukovanými atribútmi	Správnosť	Presnosť	Úplnosť
redukcia	0.78	0.78	0.82
redukcia + normalizácia	0.75	0.73	0.68
generovanie + redukcia	0.63	0.71	0.52
generovanie + redukcia + normalizácia	0.72	0.70	0.77

Na základe výsledkov môžeme usúdiť, že je vhodnejšie použiť iba redukciu atribútov a neaplikovať pred ňou generovanie. Redukciou atribútov bola dosiahnutá o trochu vyššia správnosť ako doterajšie maximum získané pomocou *MZA + atribútmi spol. sek. pohľadu*.

7.6 Diskusia

Druhý experiment bol navrhnutý tak, aby adresoval nedostatky prvého experimentu. Stránka Alzy predstavovala komplikovanejšie rozhranie, na ktorom bolo možné vytvoriť väčšie množstvo úloh na rôznych úrovniach zložitosti. Tiež sa nám podarilo dosiahnuť vyváženejšiu dátovú vzorku, ktorá je len mierne vychýlená na stranu začiatočníkov a vďaka pomerne veľkému počtu participantov sme boli schopní porovnávať hraničné skupiny participantov (hraničné prípady z pohľadu zručnosti s danou stránkou).

Taktiež sa nám podarilo dosiahnuť vhodnejší spôsob rozdeľovania participantov do skupín expertov a začiatočníkov. Na rozdiel od prvého experimentu, v ktorom bolo na stanovenie

zručnosti požitie iba sebahodnotenie, sme do procesu rozdeľovania pridali aj vplyv plnenia rozlišujúcich úloh a použili participantov pod prvým a nad tretím kvartilom skóre zručnosti (podrobnejšie vysvetlené v kapitole 7.3.1). Motiváciou použiť len hraničné kvartily bolo identifikovať na zjednodušenej situácii, či je automatické rozlišovanie začiatočníkov od expertov možné. V reálnom svete však neexistujú len hraničné prípady, a preto by bolo vhodné v ďalšej práci zistiť ako by automatické rozlišovanie fungovalo v prípade troch skupín: začiatočníci, skúsený a experti. Ďalšou alternatívou by bolo porovnať skupinu začiatočníkov voči zvyšku participantov.

V procese identifikácie najrozlišujúcejších atribútov sme zistili, že medzi najvýraznejšie patria atribúty odvodené od vzdialenosti hlavy od monitoru. Toto zistenie podobne ako v prvom experimente potvrdzuje výsledky článku Tokera [44], v ktorom identifikovali za rozlišujúce atribúty tie, ktoré sú odvodené od veľkosti zreničky a vzdialenosti hlavy od monitoru. Na rozdiel od prvého experimentu, kde sa prejavili obidva tieto typy atribútov, sa v druhom experimente prejavil iba typ odvodený od vzdialenosti hlavy od monitoru. Aj tak sa však jedná o zaujímavé zistenie, lebo riešenia založené na ľubovoľnom z týchto typov atribútov sú ľahko prenositeľné na inú stránku.

Poslednou fázou boli identifikovať vplyv jednotlivých typov atribútov na úspešnosť modelu. Model využívajúci najrozlišujúcejšie základné atribúty (MZA) dosahoval 72% správnosť, čo potvrdzuje prvú hypotézu, v ktorej sme tvrdili, že zručných participantov je možné odlišiť od expertov na základe bežne používaných metrík sledovača pohľadu. Teda metódy aplikované na identifikáciu zručnosti v doméne vizualizácií a profesijnej expertnosti sú vo veľkej miere prenositeľné aj na doménu webu.

Aj keď sa medzi rozlišujúcimi atribútmi nachádzalo pomerne veľa RQA atribútov, tak rozšírenie MZA modelu o tieto atribúty, spôsobilo zníženie správnosti modelu. Toto môže byť spôsobené vysokou vzájomnou koreláciou týchto atribútov. Na druhej strane, rozšírenie modelu o atribúty odvodené od spoločnej sekvencie pohľadu, pomohlo správnosti modelu vzrásť na 77%. Táto skutočnosť potvrdzuje našu druhú hypotézu, v ktorej sme tvrdili, že nami navrhnuté atribúty odvodené od spoločnej sekvencie pohľadu budú mať pozitívny efekt na správnosť modelu.

Nakoniec sme sa usilovali zlepšiť úspešnosť modelu aplikovaním automatického generovania atribútov a automatickej redukcie atribútov. V tomto prípade sme identifikovali, že je vhodnejšie použiť iba automatickú redukciu atribútov, ktorá jemne zlepšuje správnosť modelu na 78%.

8 Záver

V práci sme zanalyzovali pojmy zručnosť, odbornosť a oboznámenosť – na aké úrovne sa delia, a čo ich ovplyvňuje. V prípade zručnosti sme zanalyzovali aj jej efekt na namerané dáta zo sledovania pohľadu a taktiež, aké faktory môžu spôsobovať väčšie rozdiely medzi začiatočníkmi a expertmi. Keďže zručnosti používateľa na webe sa výskumy takmer vôbec nevenujú (obe analýzy boli prevedené v oblasti vizualizácií a obrázkov), tak nás zaujímala prenositeľnosť týchto poznatkov do domény webových stránok.

Ďalej sme sa zaoberali identifikáciou vhodných metrík z dát zo sledovača pohľadu. Identifikovali sme najčastejšie používané metriky extrahované z fixácií, sakád, oblastí záujmu, sekvencií pohľadu, zreničky alebo iným spôsobom a ich možné agregácie. Taktiež sme identifikovali potrebu zredukovať tieto metriky na tie, ktoré majú vysokú frekvenciu výskytu, aby ich bolo možné zaznamenať už v skorej fáze záznamu.

V praktickej časti analýzy sme sa venovali porovnaniu algoritmov, ktoré slúžia na identifikáciu spoločných sekvencií pohľadu. Porovnali sme tri algoritmy: PBWM, Dotplot a STA. Algoritmy boli porovnané nad dátovou vzorkou z DOD na FIIT spomenutou na konci kapitoly 4.4.3. Algoritmus STA vyšiel z testovania výrazne lepšie ako druhé dva – dosahoval až 3x väčšiu podobnosť voči pôvodným sekvenciám pohľadu ako zvyšné dva algoritmy.

Ako ďalšie sme navrhli a zrealizovali experiment, v ktorom sme sa snažili odhaliť, či sú spoločné sekvencie pohľadu začiatočníkov na danej stránke odlišiteľne rozdielne od spoločnej sekvencie pohľadu skúsených používateľov (expertov). Experiment bol realizovaný v doméne e-shopu Zľava Dňa a zúčastnilo sa ho 15 účastníkov. Namerané výsledky a ich zhodnotenie sa nachádza v kapitole 6.

Rozdiely medzi začiatočníkmi a skúsenými participantami sa na základe sekvencií pohľadu nepodarilo rozlíšiť. Toto môže byť spôsobené vysokou nevyváženosťou dátovej vzorky a tým, že viacero úloh bolo možné riešiť viacerými cestami. Zaujímavým zistením však je, že za vysoko rozlíšiteľné atribúty v týchto dátach sme identifikovali atribúty odvodené od veľkosti zreničky a od vzdialenosti hlavy od monitoru. Keďže ide o atribúty, ktoré nie sú priamo závislé od konkrétneho rozhrania, tak by toto riešenie mohlo byť prenositeľné aj na iné e-shopy.

Za účelom overenia výsledkov z prvého experimentu sme zorganizovali druhý experiment, ktorý adresoval problémy prvého. Experiment bol vykonaný s výrazne vyšším počtom participantov, ktorých rozloženie sa podobalo jemne vychýlenému normálnemu rozloženiu. Ďalej bol experiment vykonaný na komplikovanejšom rozhraní a definovanie zručnosti používateľa bolo formalizované. Na základe výsledkov môžeme potvrdiť prenositeľnosť riešení z domén vizualizácie a profesijnej expertnosti do domény webu. Taktiež sa potvrdili zistenia z prechádzajúceho experimentu, týkajúce sa vysokej rozlišovacej schopnosti vlastností odvodených od vzdialenosti hlavy od monitoru, avšak nepotvrdila sa rozlišujúca schopnosť veľkosti zreničky. Na základe správnosti

kombinovaných modelov sa nám podarilo dokázať, že nami navrhnuté atribúty odvodené od spoločnej sekvencie pohľadu majú pozitívny efekt na správnosť modelu. Tým pádom môžeme tvrdiť, že obe definované hypotézy z kapitoly 7 sa podarilo overiť.

Medzi hlavné obmedzenia modelu patrí fakt, že model je trénovaný na celej dôležitej časti interakcie pri riešení modelu. Ako ďalší krok by bolo preto vhodné testovať, nakoľko si vie model poradiť s kratšími nahrávkami (prvých pár sekúnd interakcie). Model nie je taktiež ľahko prenositeľný do inej domény ako celok. Časť modelu, založená na atribútoch nezávisiacich od AOI, sa dá preniesť ľahko a sama o sebe má rozlišovaciu schopnosť (viď 7.5.1). Avšak, aby bola prenositeľná aj zvyšná časť, bolo by potrebné do procesu predspracovania dát zaviesť aj krok automatickej segmentácie obrazu do AOI.

Ak by bolo potrebné model aplikovať na väčšie skupiny participantov, by bolo vhodné dopracovať predvýber najcharakteristickejších participantov, pre ktorých by sa počítali spoločné sekvencie pohľadu. Ako bolo totiž identifikované v kapitole 4.4.3, algoritmy na výpočet spoločnej sekvencie pohľadu majú problém pracovať s veľkým počtom sekvencií pohľadu a často v týchto situáciách nie sú schopné identifikovať žiadnu.

9 Literatúra

- [1] M. E. Akpınar and Y. Yeşilada: “Vision based page segmentation algorithm: Extended and perceived success,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8295 LNCS, pp. 238–252, 2013.
- [2] N. C. Anderson, W. F. Bischof, K. E. W. Laidlaw, E. F. Risko, and A. Kingstone: “Recurrence quantification analysis of eye movements,” *Behav. Res. Methods*, vol. 45, no. 3, pp. 842–856, 2013.
- [3] J. Ayres, J. Gehrke, T. Yiu, and J. Flannick: “Sequential pattern mining using a bitmap representation,” *Proc. eighth ACM SIGKDD Int. Conf. Knowl. Discov. data Min.*, pp. 429–435, 2002.
- [4] R. Bednarik, S. Eivazi, and H. Vrzakova: “A computational approach for prediction of problem-solving behavior using support vector machines and eye-tracking data,” in *Eye Gaze in Intelligent User Interfaces*, 2013, pp. 111–134.
- [5] R. Biedert, J. Hees, A. Dengel, and G. Buscher: “A Robust Realtime Reading-Skimming Classifier,” in *ETRA '12: Proceedings of the Symposium on Eye Tracking Research and Applications*, 2012, pp. 123–130.
- [6] M. J. Cole, J. Gwizdka, C. Liu, N. J. Belkin, and X. Zhang: “Inferring user knowledge level from eye movement patterns,” *Inf. Process. Manag.*, vol. 49, no. 5, pp. 1075–1091, 2013.
- [7] M. J. Cole, C. Hendaheva, N. J. Belkin, and C. Shah: “Discrimination between tasks with user activity patterns during information search,” *SIGIR '14 Proc. 37th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 567–576, 2014.
- [8] R. Díaz-Uriarte and S. Alvarez de Andrés: “Gene selection and classification of microarray data using random forest,” *BMC Bioinformatics*, vol. 7, pp. 1–13, 2006.
- [9] C. Ehmke and S. Wilson: “Identifying web usability problems from eye-tracking data,” in *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 1 - BSC-HCI 2007*, 2007, pp. 119–128.
- [10] S. Eraslan, Y. Yesilada, and S. Harper: “Identifying patterns in eyetracking scanpaths in terms of visual elements of web pages,” in *Proceedings of the 14th International Conference on Web Engineering - ICWE 2014*, 2014, vol. LNCS 8541, pp. 163–180.
- [11] S. Eraslan, Y. Yesilada, and S. Harper: “Eye Tracking Scanpath Analysis Techniques on Web Pages : A Survey , Evaluation and Comparison,” *J. Eye Mov. Res.*, vol. 9, no. 1, pp. 1–19, 2016.
- [12] S. Eraslan, Y. Yesilada, and S. Harper: “Scanpath Trend Analysis on Web Pages: Clustering Eye Tracking Scanpaths,” *ACM Trans. Web*, vol. 10, no. 4, pp. 1–35, Nov. 2016.
- [13] S. Eraslan, Y. Yesilada, S. Harper, and A. Davies: “What is Trending in Eye Tracking Scanpaths on Web Pages? Eye Tracking on the Web Trending Scanpath,” in *Proceedings of Measuring Behavior 2016*, 2016, pp. 360–362.
- [14] S. P. Farnand and J. B. Pelz: “Recurrence Metrics for Eye Movements in Perceptual Experiments,” *J. Eye Mov. Res.*, vol. 9, pp. 1–12, 2016.
- [15] R. Foerster and W. Schneider: “Functionally sequenced scanpath similarity method (FuncSim): Comparing and evaluating scanpath similarity based on a task’s inherent sequence of functional,” *J. Eye Mov. Res.*, vol. 6, no. 5, pp. 1–22, 2013.
- [16] A. Gegenfurtner, E. Lehtinen, and R. Säljö: “Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in

- Professional Domains,” *Educ. Psychol. Rev.*, vol. 23, no. 4, pp. 523–552, 2011.
- [17] A. Ghazarian and S. M. Noorhosseini: “Automatic detection of users’ skill levels using high-frequency user interface events,” *User Model. User-adapt. Interact.*, vol. 20, no. 2, pp. 109–146, 2010.
 - [18] J. Goldberg and J. Helfman: “Visual scanpath representation,” *Proc. 2010 Symp. Eye-Tracking Res. Appl. -ETRA ’10*, pp. 203–210, 2010.
 - [19] J. H. Goldberg and J. I. Helfman: “Eye Tracking on Visualizations: Progressive Extraction of Scanning Strategies,” in *Handbook of Human Centric Visualization*, 2014, pp. 337–372.
 - [20] J. H. Goldberg and J. I. Helfman: “Scanpath clustering and aggregation,” in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA ’10*, 2010, p. 227.
 - [21] J. H. Goldberg and X. Kotval: “Computer interface evaluation using eye movements : Methods and constructs,” *Int. J. Ind. Ergon.*, vol. 24, no. November 2015, pp. 631–645, 1999.
 - [22] T. Grindinger, A. T. Duchowski, and M. Sawyer: “Group-wise similarity and classification of aggregate scanpaths,” in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA ’10*, 2010, pp. 101–104.
 - [23] S. Harper: “Trends in Eye Tracking Scanpaths : Segmentation Effect ?,” *Proc. 27th ACM Conf. Hypertext Soc. Media*, pp. 15–25, 2016.
 - [24] M. Hofmann: “Support Vector Machines—Kernels and the Kernel Trick,” *Univ. Bamb.*, pp. 1–16, 2006.
 - [25] J. Holsanova, H. Rahm, and K. Holmqvist: “Entry points and reading paths on newspaper spreads: comparing a semiotic analysis with eye-tracking measurements,” *Vis. Commun.*, vol. 5, no. 1, pp. 65–93, 2006.
 - [26] M. a Just and P. a Carpenter: “A theory of reading: from eye fixations to comprehension,” *Psychol. Rev.*, vol. 87, no. 4, pp. 329–354, 1980.
 - [27] S. Kardan and C. Conati: “Comparing and combining eye gaze and interface actions for determining user learning with an interactive simulation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7899 LNCS, pp. 215–227, 2013.
 - [28] S. Kardan and C. Conati: “Exploring Gaze Data for Determining User Learning with an Interactive Simulation,” *UMAP’12 Proc. 20th Int. Conf. User Model. Adapt. Pers.*, no. 1, pp. 126–138, 2012.
 - [29] T. P. Keane, N. D. Cahill, and J. B. Pelz: “Eye-movement sequence statistics and hypothesis-testing with classical recurrence analysis,” in *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA ’14*, 2014, pp. 143–150.
 - [30] T. C. Kübler, S. Eivazi, and E. Kasneci: “Automated visual scanpath analysis reveals the expertise level of micro-neurosurgeons,” in *Joint MICCAI’15 Workshop on Interventional Microscopy*, 2015, pp. 1–8.
 - [31] T. C. Kübler, C. Rothe, U. Schiefer, W. Rosenstiel, and E. Kasneci: “SubsMatch 2.0: Scanpath comparison and classification based on subsequence frequencies,” *Behav. Res. Methods*, 2016.
 - [32] M. L. Lai *et al.*: “A review of using eye-tracking technology in exploring learning from 2000 to 2012,” *Educ. Res. Rev.*, vol. 10, no. 88, pp. 90–115, 2013.
 - [33] S. Lallé, D. Toker, C. Conati, and G. Carenini: “Prediction of Users’ Learning Curves for Adaptation while Using an Information Visualization,” in *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI ’15*, 2015, pp. 357–368.

- [34] M. S. Magnusson: "Discovering hidden time patterns in behavior: T-patterns and their detection.," *Behav. Res. Methods. Instrum. Comput.*, vol. 32, no. 1, pp. 93–110, 2000.
- [35] J. Nielsen: "Usability Engineering," *Morgan Kaufmann Pietquin O Beaufort R*, vol. 44, no. 1/2002, p. 362, 1993.
- [36] R. E. Nisbett and T. D. Wilson: "Telling more than we can know: Verbal reports on mental processes," *Psychological Review*, vol. 84, no. 3. pp. 231–259, 1977.
- [37] C. Peng, K. L. Lee, and G. M. Ingersoll: "An introduction to logistic regression analysis and reporting," *J. Educ. Res.*, vol. 96, no. September/October 2002, pp. 3–14, 2002.
- [38] A. Sanoja and S. Gancarski: "Block-o-Matic: a Web Page Segmentation Tool and its Evaluation," *29ème journées "Base de données avancées", BDA '13*, 2013.
- [39] M. Sanyová: "Odhad HCI gramotnosti používateľa s využitím sledovania pohľadu," STU, 2017.
- [40] Stolińska A; Andrzejewska M; Błasiak W; Pęczkowski P; Rosiek R; Rożek B; Sajka M; Wcisło D; "Analysis of saccadic eye movements of experts and novices when solving text tasks," *New Technol. Sci. Educ.*, no. December, pp. 21–29, 2014.
- [41] A. Sutcliffe and A. Namoun: "Predicting user attention in complex web pages.," *Behav. Inf. Technol.*, vol. 31, no. 7, pp. 679–695, 2012.
- [42] H. Takeuchi and Y. Habuchi: "A quantitative method for analyzing scan path data obtained by eye tracker," in *Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007*, 2007, no. March, pp. 283–286.
- [43] Tobii Technology: "Accuracy and precision test method for remote eye trackers," *Test Specif.*, pp. 1–28, 2011.
- [44] D. Toker, S. Lallé, and C. Conati: "Pupillometry and Head Distance to the Screen to Predict Skill Acquisition During Information Visualization Tasks," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces - IUI '17*, 2017, pp. 221–231.
- [45] D. Toker, B. Steichen, M. Gingerich, C. Conati, and G. Carenini: "Towards Facilitating User Skill Acquisition - Identifying Untrained Visualization Users through Eye Tracking," *Iui '14*, pp. 105–114, 2014.
- [46] P. Vaidyanathan, J. Pelz, C. Alm, P. Shi, and A. Haake: "Recurrence quantification analysis reveals eye-movement behavior differences between experts and novices," in *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14*, 2014, pp. 303–306.
- [47] J. M. West, A. R. Haake, E. P. Rozanski, and K. S. Karn: "eyePatterns: Software for Identifying Patterns and Similarities Across Fixation Sequences," *Proc. Eye Track. Res. Appl. Symp. ETRA*, pp. 149–154, 2006.
- [48] J. Zhang and A. a. Ghorbani: "Familiarity and trust: Measuring familiarity with a web site," *Proc. 2nd Annu. Conf. Privacy, Trust Secur.*, no. January, pp. 23–28, 2004.

Príloha A – Protokol prvého experimentu

Protokol experimentu

Kontext experimentu

Meno a priezvisko Martin Mokry

Supervisor Ing. Róbert Móro

Názov projektu Identifikácia skúsenosti používateľa na základe vzorov v dátach zo sledovača pohľadu

Názov projektu v AJ Identification of the user skill on the Web based on patterns in eyetracking data

Kľúčové slová identifikácia vzorov, zručnosť používateľa, sledovanie pohľadu, strojové učenie

Stručný opis projektu

Dôležitým krokom pri spracovaní dát zo sledovania pohľadu je identifikácia opakujúcich sa vzorov (napr. v podobe postupností fixovaných oblastí záujmu), ktoré slúžia na vysvetlenie zaznamenatej interakcie, umožňujú porovnávať spôsob interakcie rôznych používateľov, hľadať medzi nimi podobnosti a na základe nich používateľov zhľukovať. Toto je využiteľné jednak pri vyhodnocovaní zaznamenatej interakcie (napr. z testovania použiteľnosti aplikácie), ale aj pri prispôbovaní rozhraní na identifikovanú situáciu (vzor). Identifikovaná situácia, ako napríklad systematické skenovanie stránky, môže vypovedať o neznalosti rozloženia elementov na stránke a na druhej strane vzory objavujúce sa opakovane môžu implikovať každodenné aktivity zručného používateľa.

V práci sa zaoberáme automatickou identifikáciou vzorov v sekvenciách pohľadu z dát sledovania pohľadu, ktoré sú závislé od doménovej zručnosti používateľa na webe. Prvá fáza pozostáva z identifikácie vhodných vlastností fixácií a sakád (tieto by mali mať vysokú frekvenciu výskytov a nemali by byť závislé od typu vykonávanej úlohy), analýzy metód na vytváranie spoločných alebo trendujúcich sekvencií pohľadu a analýzy kvantitatívnych spôsobov vyhodnocovania podobnosti sekvencií pohľadu.

Cieľom druhej fázy projektu je s využitím strojového učenia implementovať metódu schopnú automaticky identifikovať doménovú zručnosť používateľa.

Stručný opis projektu v AJ

Identification of repeating patterns from eye tracking data (patterns like sequences of fixations, saccades or areas of interest) is considered to be an important step in eye tracking analysis. Its main focuses are: explanation of recorded interaction, comparison of ways, in which different users interact or clustering of user based on their similarity. This can be used in evaluation of recorded interaction (UX testing) or customization of graphical interface based on identified situation (pattern). This situation e.g. systematic scanning of web page can imply unfamiliarity of user with specific web page. On the other hand, reoccurrences of a similar situation can imply activities of skilled user.

In our work we focus on automatic identification of patterns in scanpaths in eye tracking data, related to domain skill level of user. First phase consists of identifying proper features of fixations and saccades (features not task-related and with high frequency of occurrences) and analysis of methods for creating commons scanpaths and quantification of similarities between scanpaths. Goal of second phase is to

implement machine learning model capable of automatic identification of domain skill level of user using just first few seconds of interaction.

Príprava experimentu

Cieľ experimentu Cieľom experimentu je získať vzory sekvencií pohľadu od troch typov používateľov:

1. tých čo webovú stránku navštevujú pravidelne,
2. tých čo webovú stránku navštevujú zriedkavo
3. a tých, čo webovú stránku nenavštevujú.

Hypotézy

Zručných používateľov stránky je možné odlíšiť od nezručných na základe vzorov v sekvenciách pohľadu.

Parametre experimentu NA

Použitý dataset NA

Účastníci

Prvá fáza experimentu bude vykonávaná s 10 až 20 účastníkmi. Všetci z nich budú študenti alebo zamestnanci Fakulty informatiky a informačných technológií v Bratislave, čiže účastníci u ktorých sa predpokladá vysoká počítačová gramotnosť.

Sledované metriky

Sledované metriky prvej fázy experimentu budú:

- binárna úspešnosť plnenia úloh (áno / nie),
- postupnosti fixácií na jednotlivé oblasti záujmu (slúžiace na tvorbu trendujúcej sekvencie fixácií), (STA, prehľad)
- postupnosti fixácií na úrovni súradníc x,y (slúžiace na výpočet rekurentných vlastností:
 - opakovanie (angl. recurrence),
 - determinizmus (angl. determinism),
 - vrstvosť (angl. laminarity),
 - stred masy opakovania (angl. center of recurrence mass))
- vlastností fixácií (dĺžka trvania fixácie, čas do prvej fixácie, počet fixácií) (prehľad)
- a vlastností sakád (relatívny/absolútny uhol medzi sakádami, dĺžka sakády)(prehľad).

Scenár experimentu

Vyplnenie dotazníku 1

Týždeň pred vykonávaním experimentu je účastníkovi predložený dotazník zisťujúci jeho gramotnosť používania webu (dotazník) prevzatý od Moniky Sanyovej. Účastníci dotazník vyplní a odošle.

Prvá fáza

Na začiatku experimentu je účastníkovi vysvetlený účel tohto experimentu. Je informovaný o tom, že všetko čo spraví bude nahrávané a je taktiež vyžiadaný jeho súhlas so spracovaním týchto údajov.

Kalibrácia

Účastník je informovaný o tom ako kalibrácia prebieha, upozornený, aby po kalibrácii už nenechal svoju polohu. Sú mu vysvetlené základné inštrukcie týkajúce sa stránky, ktorá sa mu zobrazí, ako označiť úlohu za splnenú a ako spustiť nasledujúcu úlohu. Účastník je tiež požiadaný, aby sa počas vykonávania úloh na nič nepýtal a vykonával úlohy podľa vlastného uváženia. Vykoná sa kalibrácia a zapne sa inštrukcia k prvej úlohe.

Úlohy

Účastník rieši všetky úlohy v rámci Tobii Studia. Pred vykonaním každej úlohy sa mu na jej začiatku zobrazia pokyny, ktorých potvrdením prejde k vykonávaniu danej úlohy. Účastník rieši úlohu až kým ju považuje za ukončenú.

Jednotlivé úlohy:

1. Nachádzate sa na hlavnej stránke portálu Zlavadňa a chcete skontrolovať najnovšie ponuky v okolí mesta Trenčín. Zistíte, aké sú 3 položky najlepšie spĺňajúce tieto podmienky. Názvy nájdených položiek zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

2. Nachádzate sa na podstránke portálu Zlavadňa v popise kupónu ponúkajúceho splav rieky Belá s videom, na čo by ste chceli ísť aj s ďalšími dvoma kamarátmi/kami. Skontrolujte, či je možné splavovať rieku v júni, a ak áno, pridajte 3 kupóny do košíka. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

3. Nachádzate sa na podstránke portálu Zlavadňa v košíku a uvedomili ste si, že by rád išiel ešte jeden kamarát/ka. Pridajte do košíka ešte jeden kupón a prejdite do kroku vyplňania údajov. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

4. Nachádzate sa na podstránke portálu Zlavadňa v popise kupónu na návštevu aquaparku, kam by ste chceli s kamarátmi/kami ísť. Avšak pred tým, ako kupón objednáte, si chcete adresu aquaparku zapísať. Adresu aquaparku zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

5. Nachádzate sa na podstránke portálu Zlavadňa v popise kupónu na steak. Zistíte, aké je celkové a najnovšie hodnotenie tohto podniku. Obe hodnoty zapíšete do odpovedového hárku. Pre začatie úlohy stlačte F10. V prípade, že úlohu neviete vyriešiť, stlačte F10 pre pokračovanie.

6. Nakoniec vyplňte krátky dotazník "Oboznámenosť používateľa s webom" v systéme crowdex. Pre začatie úlohy stlačte F10

Po splnení úloh je zastavené zaznamenávanie a účastník je požiadaný o vyplnenie dvoch dotazníkov:

Vyplnenie dotazníku 2

Účastníkovi je predložený dotazník zisťujúci jeho doterajšiu skúsenosť a zručnosť s testovanou stránkou <https://www.zlavadna.sk/> (dotazník). Účastník dotazník vyplní a odošle.

Otázky z dotazníku 2:

1. Zadaťte svoje použité prihlasovacie meno zo stránky sunny (vo formáte testerXX@gmail.com):
 - a)
- 2) Zlavu dňa:
 - a) nenavštevujem
 - b) navštevujem menej ako raz za mesiac
 - c) navštevujem približne raz za mesiac
 - d) navštevujem niekoľkokrát do mesiaca
 - e) navštevujem niekoľkokrát do týždňa
- 3) Iné stránky podobného typu:
 - a) nenavštevujem

- b) navštevujem menej ako raz za mesiac
 - c) navštevujem približne raz za mesiac
 - d) navštevujem niekoľkokrát do mesiaca
 - e) navštevujem niekoľkokrát do týždňa
- 4) Ak navštevujete iné stránky podobného typu, tak ktoré?
- a)

Na konci je účastník usmernený na pokračovanie na nasledujúcom experimente.

Príloha B – Protokol druhého experimentu

Protokol experimentu

Kontext experimentu

Meno a priezvisko Martin Mokrý

Supervisor Ing. Róbert Móro, PhD.

Názov projektu Identifikácia oboznámenosti používateľa s webovým sídlom na základe vzorov v dátach zo sledovača pohľadu

Názov projektu v AJ: Identification of the user familiarity with Web domain, based on patterns in eyetracking data

Kľúčové slová Uved'te aspoň tri kľúčové slová, ktoré charakterizujú povahu vášho projektu, metódy a aj experimentu.

Stručný opis projektu

Dôležitým krokom pri spracovaní dát zo sledovania pohľadu je identifikácia opakujúcich sa vzorov (napr. v podobe postupnosti fixovaných oblastí záujmu), ktoré slúžia na vysvetlenie zaznamenatej interakcie, umožňujú porovnávať spôsob interakcie rôznych používateľov, hľadať medzi nimi podobnosti a na základe nich používateľov zhľukovať. Toto je využiteľné jednak pri vyhodnocovaní zaznamenatej interakcie (napr. z testovania použiteľnosti aplikácie), ale aj pri prispôbovaní rozhraní na identifikovanú situáciu (vzor). Identifikovaná situácia, ako napríklad systematické skenovanie stránky, môže vypovedať o neznalosti rozloženia elementov na stránke a na druhej strane vzory objavujúce sa opakovane môžu implikovať každodenné aktivity zručného používateľa.

V práci sa zaoberáme automatickou identifikáciou vzorov v sekvenciách pohľadu z dát sledovania pohľadu, ktoré sú závislé od oboznámenosti používateľa s webovým sídlom. Prvá fáza pozostáva z identifikácie vhodných atribútov fixácií, sakád, zreničky a vzdialenosti hlavy od monitora (tieto by mali mať vysokú frekvenciu výskytov a nemali by byť závislé od typu vykonávanej úlohy), analýzy metód na vytváranie spoločných alebo trendujúcich sekvencií pohľadu a analýzy kvantitatívnych spôsobov vyhodnocovania podobnosti sekvencií pohľadu.

Cieľom druhej fázy projektu je s využitím strojového učenia implementovať metódu schopnú automaticky identifikovať oboznámenosť používateľa s webovou doménou.

Stručný opis projektu v AJ

Identification of repeating patterns from eye tracking data (patterns like sequences of fixations, saccades or areas of interest) is considered to be an important step in eye tracking analysis. Its main focuses are: explanation of recorded interaction, comparison of ways, in which different users interact or clustering of user based on their similarity. This can be used in evaluation of recorded interaction (UX testing) or customization of graphical interface based on identified situation (pattern). This situation e.g. systematic scanning of web page can imply unfamiliarity of user with specific web page. On the other hand, reoccurrences of a similar situation can imply activities of skilled user.

In our work we focus on automatic identification of patterns in scanpaths in eye tracking data, related to level of user's familiarity. First phase consists of identifying proper features of fixations, saccades, pupils and head distance (features not task-related and with high frequency of occurrences) and analysis of methods for creating common scanpaths and quantification of similarities between scanpaths.

Goal of second phase is to implement machine learning model for automatic identification of user familiarity with e-shop. Model will take as input basic eye-tracking features and will be extended to use also recurrence, reoccurrence metrics and metrics calculated by similarity of scanpaths to common scanpath.

Príprava experimentu

Cieľ experimentu Cieľom experimentu je získať dáta zo sledovača pohľadu pre:

- používateľov s vysokou úrovňou oboznámenosti s domenou alza.sk
- používateľov s nízkou úrovňou oboznámenosti s domenou alza.sk

, pri plnení netriviálnych úloh na doméne alza.sk.

Hypotézy

1. Oboznámených používateľov stránky je možné automaticky odlišiť od menej oboznámených na základe metrik odvodených zo sledovania pohľadu (odvodené od fixácií, sakád, veľkosti zreničky atď.).
2. Atribúty odvodené od spoločnej sekvencie pohľadu majú pozitívny vplyv na úspešnosť modelu pre automatické rozlišovanie oboznámenosti používateľa s webom.

Parametre experimentu NA

Použitý dataset NA

Účastníci

TBA

Predikovaná metrika

Počas experimentu budú participantovi zadané 2 typy úloh - Rozlišovacie a obyčajné. Rozlišovacie sú ťažšie úlohy, ktoré budú slúžiť na určenie či je participant oboznámený s webovým sídlom alebo nie a obyčajné úlohy budú slúžiť na overenie hypotéz.

Oboznámený participant - za oboznámeného používateľa budeme považovať participanta, ktorý úspešne splní minimálne 2 z troch rozlišovacích úloh a bude pri ich plnení rýchlejší ako väčšina participantov (percentil času plnenia úlohy menší než 50%).

Začiatočník – Za začiatočníka je považovaný participant, ktorý úspešne splní menej ako dve rozlišovacie úlohy alebo bude pri plnení všetkých úspešne splnených úloh pomalší ako väčšina (percentil času plnenia úlohy väčší než 50%).

Sledované metriky

Sledované metriky prvej fázy experimentu budú:

- binárna úspešnosť plnenia úloh (áno / nie),
- postupnosti fixácií na jednotlivé oblasti záujmu (slúžiace na tvorbu trendujúcej sekvencie fixácií), (STA, prehľad)
- postupnosti fixácií na úrovni súradníc x,y (slúžiace na výpočet rekurentných vlastností:
 - opakovanie (angl. recurrence),
 - determinizmus (angl. determinism),
 - vrstvosť (angl. laminarity),
 - stred masy opakovania (angl. center of recurrence mass)
- atribúty odvodené od fixácií (dĺžka trvania fixácie, čas do prvej fixácie, počet fixácií) (prehľad)
- atribúty odvodené od sakád (relatívny/absolútny uhol medzi sakádami, dĺžka sakády)(prehľad).
- atribúty odvodené od veľkosti zreničky (šírka zreničky, rýchlosť dilitácie zreničky) (prehľad)

- atribúty odvodené od vzdialenosti hlavy od monitoru (vzdialenosť hlavy od monitoru - na začiatku a konci úlohy) (prehľad)

Scenár experimentu

Vyplnenie dotazníku 1

Potenciálni účastníci dostanú dotazník doma pred experimentom. Dotazník zisťuje doterajšiu oboznámenosť s testovaným webovým sídlom (link). Na základe jeho výsledkov sa rozhodneme, ktorí potenciálni účastníci budú pozvaní na experiment.

Prvá fáza

Na začiatku experimentu je účastníkovi vysvetlený účel tohto experimentu. Je informovaný o tom, že všetko čo spraví bude nahrávané a je taktiež vyžiadaný jeho súhlas so spracovaním týchto údajov.

Kalibrácia

Účastník je informovaný o tom ako kalibrácia prebieha, upozornený, aby po kalibrácii už nenechal svoju polohu. Sú mu vysvetlené základné inštrukcie týkajúce sa stránky, ktorá sa mu zobrazí, ako označiť úlohu za splnenú, ako spustiť nasledujúcu úlohu a že na vyriešenie úloh nie je potrebné ani žiadané aby bol prihlásený. Účastník je tiež požiadaný, aby sa počas vykonávania úloh na nič nepýtal a vykonával úlohy podľa vlastného uváženia. Vykoná sa kalibrácia a zapne sa inštrukcia k prvej úlohe.

Úlohy

Účastník rieši všetky úlohy v rámci Tobii Studia. Pred vykonaním každej úlohy sa mu na jej začiatku zobrazia pokyny, ktorých potvrdením prejde k vykonávaniu danej úlohy. Účastník rieši úlohu až kým ju považuje za ukončenú.

Úlohy sú rozdelené do dvoch skupín:

- *Rozlišujúce úlohy* - jedná sa o úlohy, na základe ktorých budeme identifikovať oboznámenosť používateľa s testovaným webovým sídlom.
- *Obyčajné úlohy* - menej komplikované úlohy, ktoré budú slúžiť na overenie hypotéz

Jednotlivé úlohy:

Rozlišovacie úlohy

1)

Začiatok: hlavná stránka

Úloha: Nachádzate sa na hlavnej stránke portálu Alza. Zistite stav objednávky s číslom 169446698. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Stav objednávky:.....

2)

(<https://www.alza.sk/sony-cybershot-dsc-w800?dq=601733>)

Začiatok: Stránka s detailom konkrétneho fotoaparátu

Úloha: Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zistite do kedy by bol produkt doručený na predajňu Bratislava - Petržalka, keby si ho kúpili teraz. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Produkt bude doručený:.....

3)

Začiatok: hlavná stránka

Úloha: Nachádzate sa na hlavnej stránke portálu Alza. Nájdite najlacnejšiu ponuku rozbalených produktov (produkty vo výpredaji). Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Názov najlacnejšej ponuky je:.....

Obyčajné úlohy

4)

Začiatok: Stránka so zoznamom fotoaparátov (<https://www.alza.sk/digitalne-foto/18843129.htm>)

Úloha: Nachádzate sa na podstránke so zoznamom digitálnych fotoaparátov. Nájdite fotoaparát nad 200 eur s najnižšou cenou. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Názov fotoaparátu je:.....

5)

Začiatok: Stránka s detailom konkrétneho fotoaparátu (<https://www.alza.sk/sony-cybershot-dsc-w800?dq=601733>)

Úloha: Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zistite či daný fotoaparát poskytuje funkciu detekcie tváre. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Súčasťou balenia je:.....

6)

Začiatok: Stránka s detailom konkrétneho fotoaparátu

(<https://www.alza.sk/canon-eos-m10?dq=3753223>)

Úloha: Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zobrazte si porovnanie tohto fotoaparátu s fotoaparátom Canon PowerShot G9 X Black pomocou alza nástroja na porovnanie. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Porovnanie sa podarilo zobrazit': áno / nie

7)

Začiatok: Popis produktu (<https://www.alza.sk/sony-cybershot-dsc-w800?dq=601733>)

Úloha: Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Zapište dve pozitíva tohto produktu, ktoré produkt má na základe hodnotenia. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Negatíva sú:.....

8)

Začiatok: Detail produktu (<https://www.alza.sk/sony-alpha-5000?dq=600282>)

Úloha: Nachádzate sa na podstránke s detailom konkrétneho fotoaparátu. Skontrolujte, či je možné produkt vrátiť do 14 dní. Ak áno, tak pridajte produkt do košíka. Pre začatie úlohy stlačte F10. V prípade, že ste úlohu splnili alebo ju neviete vyriešiť, stlačte F10 pre pokračovanie.

Produkt je možné do 14 dní vrátiť: áno / nie

Po splnení úloh je zastavené zaznamenávanie a účastník je požiadaný o vyplnenie dvoch dotazníkov doma:

Vyplnenie dotazníku 2

Účastníkovi je dotazník predložený bezprostredne po splnení úloh. Dotazník identifikuje účastníkovu oboznámenosť s prostredím eshopov (link). Účastník dotazník vyplní a odošle.

Vyplnenie dotazníku 3

Týždeň po vykonávaní experimentu bude účastníkovi predložený dotazník zisťujúci jeho gramotnosť používania webu (dotazník) prevzatý od Moniky Sanyovej. Účastník dotazník vyplní a odošle.

Príloha C – Vyhodnotenie plánu práce počas doby riešenia

Plnenie plánu DP2

Týždeň	Popis
Pred semestrom	analýza dát z experimentu
	implementácia extrakcie metrík
	implementácia algoritmov na získanie spoločnej sekvencie pohľadu
	implementácia transformácie sekvencií pohľadu do metrík
	testovanie obmien klasifikátorov
1-2	vylepšovanie o váhovanie a využitie viacerých intervalov
3-4	návrh experimentu
5-6	pilotné testovanie
7-10	vyhodnotenie experimentu
10-12	spísanie výsledkov experimentu

Celý plán na tento semester nepodarilo splniť. Ak by sa ho však podarilo splniť, tak by bola diplomová práca takmer celá hotová. Míľníky sme identifikovali dobre, ale podcenili sme úsilie, ktoré je potrebné na ich dosiahnutie. Pred semestrom sa podarilo splniť len prvé dva riadky tabuľky. Z tohto dôvodu vznikol posun a ten sa ťahal až do konca semestra. Skončili sme na konci 6-teho týždňa, teda pilotným testovaním druhého experimentu. Plán na ďalší semester pokračoval tam, kde sme v tomto semestri skončili. Napriek nestihnutiu celého aktuálneho plánu je potrebné už len vykonať ďalší experiment, vyhodnotiť a spísať výsledky, čo veríme, že počas budúceho semestra zvládneme.

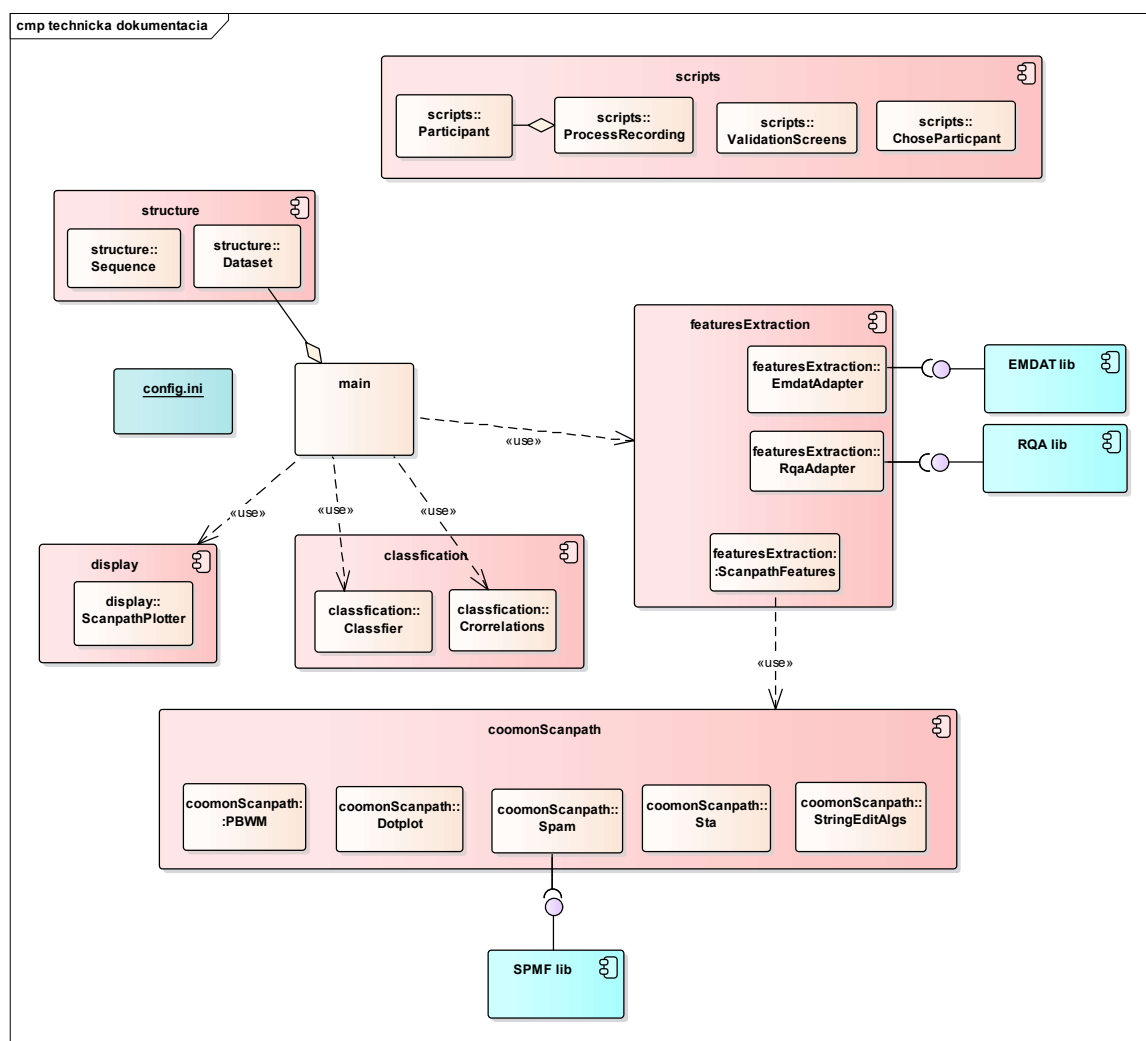
Plnenie plánu DP3

Týždeň	Popis
Pred Semestrom	kvantitatívny experiment s 1. skupinou participantov
	písanie článku na IIT.SRC
1	kvantitatívny experiment s 2. skupinou participantov
2-5	vyhodnotenie experimentu
6-7	spísanie výsledkov experimentu
8	tvorba posteru na IIT.SRC
9-12	úprava a finalizácia dokumentu

Plán na letný semester sa nám podarilo dodržať takmer na 100%. Nezrovnalosti vznikli len v týždňoch, ktoré sme definovali pre písanie IIT.SRC článku a tvorbu posteru. Obe boli posunuté o týždeň neskôr. Inak sme plán od začiatku dodržali. Prvé ostré testovania boli vykonané už pred prvým experimentom. Z odhadovaného počtu 40 participantov sa nám podarilo získať 57. Výsledky sa nám podarilo spísať už na konci 8-meho týždňa. Preto hodnotíme, že plán na letný semester bol navrhnutý lepšie a objektívnejšie ako plán na zimný semester.

Príloha D – Technická dokumentácia

Na overenie v práci definovaných hypotéz sme implementovali štruktúru skriptov, ktorá predstavuje konzolovú aplikáciu, ktorú je možné použiť aj pomocou Jupyter notebookov. Aplikácia je napísaná v jazyku python 3, pričom využíva závislosti na štandardné vedecké knižnice ako numpy, pandas, sklearn a seaborn. Štruktúru komponentov aplikácie zobrácuje *Obrázok 1* a popis jednotlivých komponentov sa nachádza nižšie.



Obrázok 1. Diagram komponentov implementovaného riešenia.

Komponent structure

Obsahuje základné dátové typy. Trieda *Dataset* v sebe obsahuje uložené základné informácie pre skupinu participantov. Medzi hlavné patria oblasti záujmu a pre každého participanta definované fixácie vo formate: časová pečiatka, čas trvania, X, Y, názov stránky. *Sequence* je trieda, ktorá z dátovej vzorky vytvára sekvencie pohľadu.

Komponent featuresExtraction

Obsahuje triedy, ktoré vypočítavajú atribúty použité ďalej v modeli. *EmdatAdapter* slúži na získanie základných atribútov, pričom na ich výpočet používa upravenú externú knižnicu EMDAT¹³. *RqaAdapter* slúži na výpočet rekurentných atribútov pomocou knižnice vytvorenej autormi článku [29]. *ScanpathFeatures* vypočítava atribúty odvodené od spoločnej sekvencie pohľadu (spoločné sekvencie pohľadu sú identifikované komponentom *commonScanpath*).

Komponent commonScanpath

Triedy *PBWM*, *Dotplot*, *Spam* a *Sta* sa používajú na výpočet spoločnej sekvencie pohľadu rôznymi algoritmami. Trieda *Spam* predstavuje adaptér na knižnicu SPMF. Podobnosť sekvencií pohľadu je vyhodnocovaná pomocou algoritmu založenom na porovnaní textových reťazcov implementovanom v triede *StringEditAlgs*.

Komponent classification

Obsahuje dve hlavné triedy: *Classification* a *Correlations*. *Correlations* sa stará o výpočet korelácií voči predikovanej hodnote, vzájomných korelácií a o odfiltrovanie vysoko vzájomne korelovaných atribútov. Trieda *Classifier* obsahuje obmeny klasifikátorov náhodný les, SVM a log. regresia, ktoré sa dajú aplikovať na predspracované dáta pomocou modulu *featuresExtraction*.

Komponent display

Skladá sa len z jednej triedy *ScanpathPlotter*, ktorá sa stará o vizualizáciu sekvencií pohľadu konkrétneho účastníka.

Komponent scripts

Je nezávislý od všetkých ostatných komponentov a obsahuje pomocnú funkcionality. Trieda *ProcessRecording* sa používa na vybratie relevantných dát z kompletných exportov účastníkov. Vyžíva na to časové intervaly definované v triede *Participant*. Následne až export vybraných dát je ďalej požívaný v hlavnom skripte. Trieda *ValidationScreens* sa požíva na výpočet správnosti a presnosti sledovača pohľadu. *ChoseParticipant* sa požíval na rozdelenie účastníkov do skupiny expertov a začiatočníkov na základe počtu splnených úloh a času plnenia úlohy (neskôr sa však prešlo na vzorec definovaný v kapitole 7.3.1).

Pomocnou triedou, ktorá nie je súčasťou žiadneho komponentu, ale slúži na pospájanie ostatných komponentov je *main*. Táto v nasledovnom poradí volá komponenty: *structure* (načítanie dát), *feature extraction*, *classification*. Ďalším samostatne stojacím súborom je *config.ini*, predstavuje konfiguračný súbor, ktorého nastavenia sú podrobnejšie rozpísané v používateľskej príručke.

¹³ <https://github.com/Groosling/EMDAT>

Externé knižnice

RQA lib [29] je jedinou použitou externou knižnicou, ktorá je tiež napísaná v jazyku python 3 a bolo ju možné použiť volaním metód. Požitie ostatných knižníc je o niečo komplikované a vyžaduje si to komunikáciu pomocou súborov.

Upravená verzia *EMDAT* knižnice vyžaduje :

- Súbor *TobiiV3_sample.seg* – je možné zadať viacero segmentov, ak by človek potreboval vypočítať atribúty pre viacero časových rozmedzí. Ak nie, tak ho treba nastaviť na: *problem1 S1 1 99999999*.
- Súbor *pupilSizes.tsv* definuje pre každého účastníka jeho priemernú veľkosť zreničky vo formáte: *tester08 2.786452702702704*.
- Súbor *[X]_Data_Export.tsv*, kde X sa nahrádza číslom úlohy. Súbor obsahuje export tobii studia z relevantných časových intervalov.
- Volanie z príkazového riadku s argumentami oddelenými medzerami. Prvý argument je číslo úlohy a druhý argument je zoznam účastníkov oddelený medzerami. Príklad: *7 tester07 tester08 tester30 tester43 tester44*.

EMDAT vypočítava nasledujúci zoznam atribútov:

- Odvozené od pohľadu: *numsegments, length, numsamples, numfixations, fixationrate, meanabspathangles, meanfixationduration, meanpathdistance, meanrepathangles, stddevabspathangles, stddevfixationduration, stddevpathdistance, stddevrepathangles, eyemovementvelocity, abspathanglesrate, repathanglesrate, sumabspathangles, sumfixationduration, sumpathdistance, sumrepathanglesnumsaccades, sumsaccadedistance, meansaccadedistance, stddevsaccadedistance, longestsaccadedistance, sumsaccadeduration, meansaccadeduration, stddevsaccadeduration, longestsaccadeduration, meansaccadespeed, stddevsaccadespeed, minsaccadespeed, maxsaccadespeed, fixationsaccadetimeratio*
- Odvozené od veľkosti zreničky: *meanpupilsize, stddevpupilsize, maxpupilsize, minpupilsize, startpupilsize, endpupilsize, meanpupilvelocity, stddevpupilvelocity, maxpupilvelocity, minpupilvelocity*
- Odvozené od vzdialenosti hlavy od monitoru: *meandistance, stddevdistance, maxdistance, mindistance, startdistance, enddistance*

Poslednou je SPMF knižnica na aplikovanie algoritmu SPAM. Pre túto sa za behu programu generuje vstupný súbor, ktorý obsahuje sekvencie pohľadu vo formáte: *0 -1 1 -1 0 -1 2 -2*, kde AOI sú identifikované číslami, -1 je oddelovač AOI a -2 je koniec sekvencie pohľadu.

Príloha E – Používateľská príručka

V prípade potreby práce s dátami z druhého experimentu nie je potrebné robiť žiadne predspracovanie (predspracované dáta sa už nachádzajú na priloženom médiu). Avšak pri práci s vlastnými dátami je pred spustením hlavného skriptu potrebné dáta predspracovať a povytvárať potrebné súbory definované v technickej dokumentácii. Kroky potrebné na vytvorenie vyžadovaných súborov s dátami zo sledovača pohľadu sú nasledovné:

1. Vyexportuj z Tobii Studia dáta každého účastníka ako jeden súbor.
2. Definuj v skripte Participant časy kalibračnej obrazovky a plnenia jednotlivých úloh pre každého účastníka.
3. Pusti skript Process recording, ktorý vygeneruje súbor *pupilSizes.tsv* a súbory *[X]_Data_Export.tsv*.
4. Pre každú úlohu definuj súbor s AOI, kde jedna AOI má formát:
Identifier, X-from_coordinate, X-offset, Y-from_coordinate, Y-offset, ShortID
5. Všetky vytvorené súbory nakopíruj do dátového priečinku hlavného skriptu.

Požítie notebookov

Skripty je možné použiť aj pomocou Jupyter notebookov. V tomto prípade treba vytvoriť notebook v koreňovom priečinku projektu a použiť daný notebook namiesto main skriptu. V prípade potreby je možné preťažiť ľubovoľnú triedu a odstrániť jej import, čím sa zabezpečí že sa použije nová implementácia. Ukážky takýchto notebookov sa nachádzajú na priloženom médiu. Pred časťami kódu je vždy popísané, čo treba v konfigurácii nastaviť inak voči prednastaveným hodnotám. Súbory sa nachádzajú v koreňovom adresári, pričom názvy notebookov použitých na vyber klasifikačného modelu začínajú znakom “1”, notebooky použité na porovnanie rôznych typov atribútov začínajú názov s “2” a notebooky s automatickým generovaním a redukciou atribútov začínajú znakom “3”. Medzi notebookmi nie sú žiadne závislosti a je ich možné púšťať v ľubovoľnom poradí.

Konfigurácia

V nasledovnej časti sú popísané podstatné nastavenia, ktoré je potrebné skontrolovať a upraviť podľa potreby pred spustením hlavného skriptu :

[sequence]

maxAoi – maximálny počet znakov v prvotnej sekvencii pohľadu.

maxFinalScanpathLength - maximálny počet znakov vo finálnej sekvencii pohľadu (po zredukovaní odstránením po sebe idúcich rovnakých znakov).

normalizeSimilarity – 0/1 – či normalizovať podobnosť medzi sekvenciami pohľadu.

[display]

minFixDurationThreshold - minimálna dĺžka fixácie, aby sa zobrazila vo vizualizácii sekvencií pohľadu.

[spam]

lib – cesta ku knižnici spmf.jar.

inputFilePath - skade bude knižnica čítať sekvencie pohľadu.

outputFilePath - kam knižnica zapíše spoločnú sekvenciu pohľadu.

[EMDAT]

python27Path – cesta k pythonu 2.7.

[RQA]

clusteringDistanceThreshold, phaseSpaceClusteringThreshold, numTimeDelaySamples

timeDelayValue - sú podrobnejšie vysvetlené v dokumentácii [29].

[run]

scanpathFilePaths – zoznam predspracovaných exportov z Tobii Studia.

aoiFilePaths – zoznam ciest k súborom definujúcim oblasti záujmu.

websiteNames – zoznam mien analyzovaných stránok v každej úlohe.

backgroundPageImages – zoznam ciest k obrázkom, na ktoré sú vykresľované sekvencie pohľadu.

[correlations]

interAttributesCorrThreshold - ak je vzájomná korelácia väčšia, ako daný prah, tak je zo zoznamu odstránený atribút s nižšou koreláciou k predikovanému atribútu.

[classification]

useCsv - 0/1. Ak je nastavený na 0, tak sa počítajú všetky atribúty nanovo. Ak je 1, tak sa použijú naposledy vypočítané atribúty.

csvPath – cesta k súborom, kde sú uložené naposledy vypočítané atribúty.

columnNames – zoznam atribútov, ktoré sa majú použiť.

[delete]

1, 2, 3, 4, 5, atď - počet podľa počtu úloh. Pre každú úlohu je možné definovať zoznam participantov, ktorých nepoužiť.

[participants]

group1, group2 = zoznamy participantov patriacich do jednotlivých skupín . V našom prípade skupina začiatočníkov a skupina expertov.

Príloha F – IIT.SRC článok

Príloha G – Obsah priloženého média

Priložené médium obsahuje zdrojové súbore použité pri analýze dát, exporty dát z Tobii Studia a samotný dokument. Súbory sú rozložené do adresárov s nasledujúcou štruktúrou:

- *\DP_features_and_scanpaths* – zdrojové súbory projektu,
 - *\data* – dáta z druhého experiment vo formáte .tsv + všetky ostatné potrebné pomocné vstupné súbory.
- *\dokument* – aktuálna verzia elektronickej formy tohto dokumentu.