

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLOGIÍ**

FIIT-5220-64327

Bc. Peter Dulačka

**Hra s účelom ako
prostriedok získavania metadát
a objavovania autorít**

Diplomová práca

Študijný program: Softvérové inžinierstvo

Študijný odbor: 9.2.5 Softvérové inžinierstvo

Miesto vypracovania: Ústav Informatiky a Softvérového Inžinierstva, FIIT STU

Vedúci práce: Ing. Jakub Šimko, PhD.

máj 2014

Anotácia

Slovenská Technická Univerzita v Bratislave
Fakulta Informatiky a Informačných Technológií

Študijný program: Softvérové inžinierstvo

Autor: Bc. Peter Dulačka

Diplomový projekt: Získavanie sémantiky a metadát pomocou hier s účelom

Vedúci projektu: Ing. Jakub Šimko, PhD.

máj 2014

Nie všetky typy úloh je možné automatizovať, navyše človek ich vie vykonať omnoho jednoduchšie a rýchlejšie. Na tomto princípe funguje získavanie dát s využitím davu. Jedným zo spôsobov, ktorým môžeme dav motivovať vykonávať zadané úlohy sú tzv. hry s účelom. Súčasným problémom je efektívne využitie davu, kedy dáta získané od jedincov odlišujúcich sa od priemerného jednotlivca (pozitívne aj negatívne) nie sú brané do úvahy.

V práci sa zameriavame na vytvorenie hry s účelom, ktorá by dokázala s davom pracovať efektívne. Tento cieľ chceme dosiahnuť rozoznaním expertov v dave a ich povýšením nad priemerných jednotlivcov. Zachovali by sme tak súčasné využitie davu a vylepšili by sme ho o dáta získané od hráčov s vyššou odbornosťou. Expertov rozpoznávame na základe priameho porovnania hráčov, ktorí sú zároveň poslucháčmi vysielania nami vytvoreného internetového rádia a formou hry sú ich vedomosti overované voči faktom súvisiacim s práve prehrávanou skladbou – napr. názov skladby, interpret, album, časť textu prípadne iné rytmické alebo vizuálne prvky. Na základe navrhnutého modelu vyhľadávame expertov medzi hráčmi pre zadaný dopyt.

Expertov následne v rámci hry využívame na zrýchlenie a skvalitnenie procesu tvorby metadát v hrách s účelom. Porovnáваме kvalitu výstupu pri prístupe, kedy sú si všetci hráči rovnocenní a keď zvýhodníme expertov na základe ich expertného skóre.

Annotation

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies

Degree Course: Software engineering

Author: Bc. Peter Dulačka

Diploma thesis: Game with a Purpose as a Mean for Metadata Acquisition and Authority Discovery

Supervisor: Ing. Jakub Šimko, PhD.

May 2014

Not every task can be automated by machine and these kinds of tasks tend to be solvable by human easily and faster. This is also the main idea of crowdsourcing. One of the means of motivating crowd to do such tasks are Games with a Purpose (GWAP). Current problem of GWAPs is efficient utilization of the crowd, because data acquired from anomalous individuals are not considered into created dataset.

In this project we created GWAP being able to handle the crowd efficiently. We aim this goal by finding experts among individuals in crowd and promoting them above average individuals. The standard utilization of crowd remains and gets improved by expertise of educated players. Expert finding is based on direct comparison of players which are listeners of our custom-created internet radio. They're quizzed with facts related to song they are listening to - e.g. song title, artist, album, lyrics or rhythmical and visual elements of song. Then, based on given query, we search for experts among players meeting given boundaries.

Experts are utilized to improve and fasten the process of metadata generation in crowdsourcing games. We compare quality of metadata generated when all players are equal to each other and when some users are favored over others due to their expert score.

Čestne prehlasujem, že som diplomovú prácu vypracoval samostatne s použitím uvedenej literatúry pod odborným vedením Ing. Jakuba Šimka, PhD.

Bratislava, máj 2014

podpis

Chcel by som poďakovať Ing. Jakubovi Šimkovi, PhD. za odborné vedenie, nápady a podporu počas vypracovávania práce. Vďaka ide aj členom skupiny PeWe za spätnú väzbu a pomoc s realizáciou experimentu. Osobitná vďaka patrí všetkým hravým ľuďom, ktorí si zapli WoodstockFM, aspoň chvíľu za zahráli, zabavili a dali tejto práci zmysel.

Osobitné poďakovanie patrí rodine, blízkym a priateľom za to, že to so mnou vydržali.

Ďakujem.

Obsah

1. ÚVOD	1
2. ČERPANIE Z DAVU	3
2.1. Konvenčné prístupy	3
2.1.1. Amazon Mechanical Turk (MTurk).....	3
2.1.2. Yahoo Answers.....	5
2.1.3. Stack Overflow.....	6
2.1.4. LastFM	7
2.2. Hry s účelom.....	7
2.2.1. Little Search Game	10
2.2.2. Herd-It	11
2.2.3. Moodswings.....	11
2.2.4. City Lights.....	12
2.3. Diskusia.....	12
3. OBJAVOVANIE AUTORÍT	14
3.1. Typy autorít.....	15
3.1.1. Autority na webe.....	15
3.1.2. Autority v podnikových systémoch	16
3.1.3. Autority v hrách s účelom.....	16
3.2. Rozpoznávanie autorít	17
3.2.1. Metriky na hodnotenie autority používateľa v dave.....	17
3.2.2. Odbornosť experta na základe dopytu.....	19
3.2.3. Hodnotenie odbornosti využitím znalostí davu	20
3.2.4. Autorita reprezentovaná skupinou expertov	21
3.3. Diskusia.....	22
4. HRA S ÚČELOM S ROZPOZNÁVANÍM A VYUŽITÍM AUTORÍT	23
4.1. WoodstockFM.....	23
4.1.1. Herné mechanizmy	24
4.1.2. Realizácia	25
4.1.3. Herné prostredie.....	26
4.1.4. Architektúra	29
4.2. Objavovanie autorít	29
4.2.1. Podobnosť interpretov	29
4.2.2. Dáta o používateľoch získané z aplikácie	30
4.2.3. Utlizácia dát a objavovanie autorít	30
4.3. Diskusia	33
5. OVERENIE NAVRHNUTÝCH METÓD	34
5.1. Overenie vedomostí testom.....	34
5.2. Spätné overenie úrovne expertízy predikciou výsledkov nasledujúcich hier.....	36
5.2.1. Výpočet expertízy.....	37
5.2.2. Predikcia výsledkov	40
5.3. Overenie zlepšenia kvality generovania metadát davom v hre	42

6. ZHODNOTENIE.....	45
7. BIBLIOGRAFIA.....	47

1. Úvod

Množstvo informácií, ktoré sú v súčasnosti dostupné je obrovské, za čo môže hlavne rozšírenie blogov a sociálnych sietí na webe – Google v súčasnosti (máj 2014) indexuje niekoľko desiatok miliárd webov¹. S publikáciou veľkého množstva informácií vzniká problém s vyhľadávaním a využitím týchto informácií – je potrebné použiť dostupné metadáta k týmto dokumentom. Nanešťastie nie každý pri zverejnení dokumentu zverejní aj jeho metadáta. Dodatočné získavanie metadát pre zverejnené dokumenty nie je navyše rovnako jednoduché pre všetky typy dokumentov (novinový článok, film, skladba).

Na získavanie metadát sú štandardne využívané tieto prístupy:

- **Expertom v oblasti.** Expert je osoba, ktorá má znalosti z vybranej domény a tieto znalosti vie využiť pri dedukcii, riešení problému alebo porozumeníu situácie. Vďaka znalostiam experta je možné vytvoriť veľmi podrobné metadáta – z menovaných prístupov sú najvyššej kvality – avšak cena experta je vysoká vezmúc do úvahy časovú náročnosť vytvorenia správnych metadát a množstva metadát, ktoré v súčasnosti chýba. Aj preto sa experti využívajú len ako doplňujúci prístup k ostatným dvom.
- **Automatickým prístupom.** Umelá inteligencia a učiace sa algoritmy dokážu na základe kontextu a existujúcej vzorky dokumentov s kvalitnými metadátami (vytvorenými napr. expertmi) generovať metadáta aj pre iné - podobné dokumenty. Ich úspešnosť je závislá od typu spracovávaného dokumentu. Metóda je najrýchlejšia a najlacnejšia z menovaných.
- **Využitím znalostí davu – „crowdsourcing“.** Davom rozumieme náhodnú skupinu ľudí, ktorých spája rovnaký záujem, charakteristika, vlastnosť – vďaka spoločným záujmom sú schopní sústrediť sa na jednom mieste (reálnom aj virtuálnom) už s minimálnou snahou tretej strany. Pri využívaní znalostí davu jeho členovia riešia zadané úlohy. Počas jej riešenia a-priori nevieme, aká je odbornosť jednotlivcov v dave.

Znalosti priemerného používateľa by na vytvorenie použiteľných metadát nestačili, avšak kombinovaním veľkého množstva používateľov v dave nastávajú zjednodušene dva efekty: 1) kontradikciou medzi používateľmi sú vylúčené nesprávne generované dáta a 2) zhodou medzi používateľmi sú potvrdené správne dáta.

Využitie davu má pri správnom nastavení veľký potenciál, keďže dav dokáže generovať omnoho väčšie množstvo dát ako experti a je schopný jednoducho vyriešiť v súčasnosti strojovo ťažko riešiteľné úlohy (napr. anotovať hudbu, pomenovávať objekty v obrázkoch a pod.). Jednou z foriem využitia davu sú tzv. *hry s účelom* („*games with a purpose*“ – *GWAP*), ktoré využívajú zábavu ako motiváciu pre používateľa vykonávať úlohy. Navyše pri správnom návrhu hry hráč ani nemusí vedieť, že úlohu vykonáva.

Jedným zo súčasných problémov je efektivita využitia davu. Znalosti jednotlivcov v dave sa líšia a nemusia byť vždy priemerné. Keďže využitie davu závisí na využití zhody medzi používateľmi v dave, nastávajú dva efekty:

¹ <http://www.worldwidewebsize.com/>

- **Odstránenie nesprávnych informácií.** V prípade, že používateľ zadá nepravdivú informáciu, porovnaním s výstupom od ostatných používateľov je možné túto informáciu identifikovať a z výsledku vylúčiť. Nesprávna informácia by sa do výsledného výstupu dostala iba v prípade, že by ju za správnu považoval dav.
- **Odstránenie expertných informácií.** Ak sa v dave nachádza expert, ktorý na základe svojich znalostí vygeneruje špecifickejšie metadáta o dokumente ako zvyšok davu, sú tieto informácie z výsledného výstupu vylúčené – rovnako ako v predchádzajúcom prípade sa totiž vymykajú rozhodnutiu davu.

Napriek tomu, že využitie davu je schopné generovať aj expertné informácie, kvôli nemožnosti rozoznať expertné informácie od nesprávnych, nie sú expertné informácie brané do úvahy. Spomenuté dva efekty môžu viesť pri rozpoznávaní expertov až k paradoxom:

- **Expert dodá generickú informáciu porovnateľnú s informáciami generovanými davom.** Ak v súčasnosti ďalší používatelia s expertom nesúhlasia, metódy, ktoré sú založené na väčšinovej zhode označia expertnú informáciu ako chybnú. Javí sa totiž úplne rovnako ako ostatné chybné minoritné informácie. Tento efekt môže doceliť až to, že preváži chybný názor väčšiny v dave a názor experta je odignorovaný.
- **Expert dodá špecifickejšiu (odbornejšiu) informáciu v porovnaní s informáciami davu.** Odbornosť davu je ale nižšia a dav nie je schopný potvrdiť pravdivosť informácie dodanej expertom. V takom prípade je opäť expertná informácia vylúčená z datasetu.

Problém je možné vyriešiť rozpoznaním expertov v dave. Nimi generované informácie by nielenže neboli filtrované ako nesprávne, ale v porovnaní s ostatnými jedincami v dave by mohli mať nimi generované dáta väčšiu váhu.

Zakomponovaním rozpoznávania expertov v dave do hier s účelom by sme navyše mohli znásobiť rýchlosť a presnosť generovaných metadát. Mohlo by to viesť k odstráneniu spomenutých paradoxov a k zefektívneniu využitia davu – čo je vzhľadom na jeho potenciál dôležité. V tejto práci sa chceme zamerať na dva ciele:

- Rozpoznávanie expertov pre danú doménu v hrách s účelom.
- Vytváranie a validácia metadát k hudobným skladbám so zapojením znalostí expertov.

Nami navrhovaná hra je postavená na priamom porovnaní hráčov počúvajúcich internetové rádio – hra je súčasťou rádia. Pomocou postupov navrhnutých v našej predchádzajúcej hre City Lights [1] a ich modifikácií zabezpečujúcich využitie expertov opísaných v ďalších častiach tohto dokumentu a pomocou využitia samotného davu sa snažíme o rozpoznanie expertov priamo medzi hráčmi.

Na identifikáciu expertov, ktorá je hlavnou časťou tejto práce, využívame primárne modifikáciu existujúcich metód na rozpoznávanie autorít – tieto sú bližšie opísané v časti 3. *Objavovanie autorít*. Experti sú v súčasnosti najčastejšie rozpoznávaní v podnikových systémoch (hlavne na základe interakcie s dokumentmi) a fórach (hlavne na základe interakcie s ostatnými používateľmi). Úspešnosť týchto prístupov je nízka a pohybuje sa okolo 40-50%. Vybranú metódu sme museli upraviť aby korelovala s vlastnosťami hier s účelom a návrhom našej hry. Konkrétne úpravy a dosiahnuté vlastnosti opisujeme v časti 4. *Hra s účelom s rozpoznávaním a využitím autorít*. V závere navrhnutú metódu v niekoľkých experimentoch vyhodnocujeme.

2. Čerpanie z davu

Čerpanie z davu je jedným zo spôsobov získavania informácií, kedy sú využívané znalosti jednotlivcov v dav - tento spôsob získania informácií sa používa hlavne na webe. Oproti využitiu expertov je omnoho lacnejší, ale jeho výsledky silne závisia od typu a veľkosti davu. V porovnaní s automatizovanými prístupmi ponúka relevantné výsledky hneď od začiatku – automatické prístupy vyžadujú trénovaciu vzorku, na základe ktorej generujú ďalšie informácie.

Sila čerpania z davu je v tom, že dokáže veľmi jednoducho vyriešiť problémy, ktoré nie sú strojovo riešiteľné alebo dosahujú nízku úspešnosť. Dokonca aj dav pozostávajúci z malého množstva jednotlivcov dokáže vyplniť IQ test lepšie, ako 99% populácie [2].

Jedným z hlavných problémov pri čerpaní davu je jeho motivácia. Používatelia dobrovoľne nechcú riešiť zadané úlohy a vyžadujú určitú formu odmeny – najčastejšie finančnú kompenzáciu. Okrem financií sú v súčasnosti používatelia motivovaní zábavou pomocou tzv. hier s účelom [3]. Tie svoj „účel“ maskujú do návrhu hry a používateľ často ani nevie, že pomáha riešiť určitý problém.

2.1. Konvenčné prístupy

Keďže využitie znalostí más je najlepšie prevediteľné na Internete, všetky prezentované služby sú webové stránky alebo portály. Konvenčné prístupy môžeme rozdeliť na:

- **Portály bez finančnej odmeny.** Princíp portálu stavia na dobrovoľníctve používateľov, kedy je ich hlavná motivácia altruizmus a zdieľanie vedomostí. Rôzne druhy motivácie používateľov sú podrobne opísané v časti 3.1.1. Kladenie otázok a vyhodnocovanie ich správnosti prebieha samotnými používateľmi. Typy otázok sa líšia, avšak odpoveď väčšinou vyžaduje iba spojiť dostupné existujúce znalosti.
- **Portály s finančnou odmenou.** Úlohy zadávajú zvyčajne inštitúcie a organizácie, ktoré majú dostatočný rozpočet na odmeňovanie používateľov. Principiálne sa takéto portály zameriavajú na iné druhy otázok ako v predchádzajúcom prípade a namiesto získania ľubovoľnej informácie využívajú dav na získanie informácie, ktorá nie je strojovo generovateľná – využíva sa hlavná výhoda čerpania z davu.

V nasledujúcej časti sú opísané v súčasnosti (máj 2014) najznámejšie portály využívajúce dav na získanie informácií. Portály ako Wikipedia², ktoré síce využívajú dav, ale primárne slúžia na poskytovanie informácií - teda používatelia (dav) neriešia vopred zadanú úlohu, na ktorú hľadáme odpoveď - z tohto prehľadu vynechávame.

2.1.1. Amazon Mechanical Turk (MTurk)

Amazon Mechanical Turk³ je v súčasnosti najväčší portál poskytujúci pre používateľov možnosť distribuovať riešenie úlohy na dav (veľké množstvo jednotlivcov). Všetky úlohy pozostávajú z menších častí, za splnenie ktorých používateľ dostane odmenu – tzv. HIT („*human intelligence task*“). V súčasnosti sa priemerná odmena za HIT pohybuje od \$0.01 do \$0.10 v závislosti od času,

² <http://www.wikipedia.org/>

³ <https://www.mturk.com>

ktorý na svoje splnenie vyžaduje – ten je rádovo niekoľko minút. Aktívni používatelia dokážu zarobiť okolo \$2 za hodinu [4].

Demograficky je súčasné rozdelenie používateľov neznáme, avšak v marci 2008 tvorili 76% riešiacich používateľov obyvatelia USA a 8% obyvatelia Indie. V novembri 2009 bol tento pomer 56% z USA a 36% z Indie. Predpokladá sa, že väčšinu používateľov dnes tvoria India.

Názov Mechanical Turk má zaujímavé historické pozadie [5]. Koncom 60-tych rokov 18. storočia Wolfgang von Kempelen vybudoval prvý stroj, ktorý bol schopný poraziť človeka v šachu. Tento stroj nazval „Turk“. Počas svojich ciest po Európe s ním zožal obrovský úspech a porazil aj Napoleona a Benjamina Franklina. Bol to ale podvrh a vnútri stroja sedel človek – šachový majster, ktorý odohral všetky partie. Amazon ako poctu tomuto stroju pomenoval svoju službu Mechanical Turk.

Jednotlivci v dave riešiaci úlohy sa od seba navzájom líšia a pri vytváraní úlohy pre veľké množstvo jednotlivcov treba dbať na špecifické vlastnosti návrhu [2]:

- **Povolená chybovosť.** Niektoré úlohy vyžadujú od riešiteľa stopercentnú úspešnosť. Atribút je aplikovateľný až pri vyhodnocovaní a iba v prípade, že je možné overiť správnosť odpovede riešiteľa – napr. test s vopred danými odpoveďami.
- **Filtrovanie výsledkov.** Rozdielni používatelia vkladajú do riešenia úloh rôzne množstvo úsilia, čo sa môže odraziť aj na výsledných dátach. Ako efektívne riešenie sa ukázalo zavedenie reputácie medzi používateľmi a odmietnutie výstupov používateľov s reputáciou nižšou ako vopred danou.
- **Agregácia výsledkov.** V prípade, že HIT rieši niekoľko používateľov, nemusia byť do výsledku brané výstupy všetkých, ale iba najčastejšie vyskytujúca sa odpoveď. Agregácia je v tomto prípade dôležitý aspekt na zváženie, keďže ostatné odpovede sú „zahodené“, ale stále je potrebné ich riešiteľov finančne odmeniť.
- **Odmena za HIT.** Správna veľkosť odmeny za vyriešenie HITu je kľúčová. Vyššia odmena priamo nekoreluje s lepšimi výsledkami a ovplyvňuje iba čas na získanie potrebného počtu výstupov od používateľov. Vysoká odmena totiž môže na riešiteľa vytvárať psychologický tlak a znižovať jeho kognitívne schopnosti.

V kontexte rozpoznávania expertov a zníženia chybovosti davu môžeme sledovať niekoľko jednoduchých opatrení. Tieto opatrenia priamo ovplyvňujú riešiteľov úloh a tí sú si nimi vedomí. Avšak povolená chybovosť riešených úloh alebo reputácia používateľov príliš obmedzujú využitie jednotlivcov a môžu viesť k cielenému správaniu používateľov za účelom odstránenia týchto obmedzení a prístupu k lepšie hodnoteným úlohám. Pri navrhovaní hry s účelom, ktorá úroveň expertízy hráčov v doméne využíva by teda nemali hráči vedieť, čo všetko vplýva na ich úroveň expertízy (príp. že ich expertíza sa využíva).

V súčasnosti sa podľa [4] Mechanical Turk využíva aj v oblasti výskumu a to na rôzne oblasti: používateľské štúdie, anotovanie obrázkov, spracovanie prirodzeného jazyka, vyhodnotenie

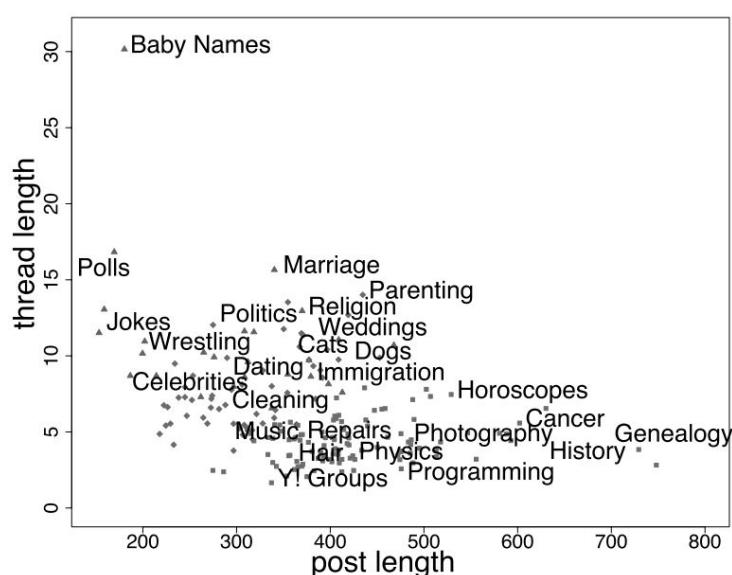
podobnosti a iné. Okrem Mechanical Turk patria medzi známe portály na využitie davu CrowdFlower⁴ alebo TaskCN⁵.

2.1.2. Yahoo Answers

Yahoo Answers patrí medzi tzv. CQA weby („community question answering“) – weby zameriavajúce sa na odpovedanie používateľmi generovaných otázok. Štandardne možno otázky na tomto type portálov na základe úloh riešených na TREC konferencií rozdeliť do 3 kategórií [6]:

- **Faktoidné.** Najčastejší typ otázok, ktorý používatelia využívajú v prípade, keď vyhľadávače nedokážu nájsť presnú odpoveď alebo by vyhľadávanie pomocou vyhľadávača vyžadovalo niekoľko dopytov z rôznych oblastí.
- **Názorové.** Odpovede väčšinou pripomínajú skrátené blogy, kedy odpovedajúci vyjadrí svoj názor a podporuje ho argumentom. Otázky znejú „Čo si myslíte o...“ a buď sa sústreďujú na jeden predmet, alebo porovnávajú niekoľko vybraných predmetov.
- **Komplexné.** Otázky, ktoré nie je možné jednoducho zaradiť do predchádzajúcich dvoch skupín.

Na portáli sa vyskytujú otázky z rôznych (miestami vzdialených) kategórií a je zaujímavé pozorovať rozdielnu aktivitu používateľov naprieč kategóriami. Na obr. 2.1 je znázornené rozloženie dĺžky vlákna s otázkou a dĺžky jednej odpovede vlákna naprieč niektorými kategóriami [7]. Rovnaký výskum ukázal, že kvôli veľkému množstvu kategórií a širokému záberu portálu je menej ako 1% odpovedí možné považovať za expertné.



obr. 2.1 Rozloženie dĺžky príspevkov a dĺžky vlákien na Yahoo! Answers

Používatelia pri faktoidných otázkach nezvykli vystupovať v rámci jednej kategórie aj ako pýtajúci sa, aj ako odpovedajúci. Pri odborne zameraných kategóriách sa aktivita používateľov vyhradzovala práve na jednu kategóriu – entropia týchto používateľov bola nízka a odpovede týchto používateľov boli pýtajúcimi sa zvyčajne označované ako najlepšie. To pomohlo pri automatizovaných pokusoch o určenie najlepšej odpovede naprieč celým portálom. Z výskumu

⁴ <http://crowdflower.com/>

⁵ <http://www.taskcn.com/>

vyplynuli aj kategórie, kde používatelia odpovedali bez ohľadu na to, kde predtým kladli otázky (napr. kategória „rodina a vzťahy“).

Rast portálu a počet nových odpovedí sa v posledných mesiacoch spomaľuje. Najväčšia aktivita je v súčasnosti vykazovaná hodnotením existujúcich odpovedí. Jav môže byť spôsobený zodpovedaním veľkého množstva štandardných otázok, na ktoré portál dáva dostatočne dobré odpovede prípadne vznik špecializovaných portálov pre jednotlivé kategórie, ako napríklad TripAdvisor⁶, Stack Overflow pre programovanie resp. Stack eXchange⁷ sieť portálov rozdelených podľa záujmových oblastí. Rozšírenie špecializovaných portálov bolo nevyhnutné z dôvodu nízkej špecifickosti odpovedí na odborné otázky bez ohľadu na kategóriu [7].

Pre kontext využitia expertov v hrách teda najlepšie vychádza integrácia s faktoidnými otázkami – v hre napr. využitím kvízu. Zúčastnení hráči môžu preukazovať vedomosť odpovedaním na otázky, ktorých správnosť je overiteľná a a-priori majú takíto hráči vedomosti v rôznych kategóriách (doménach), vďaka čomu je možné využívať rôznu úroveň expertízy hráčov pre rôzne druhy domén.

2.1.3. Stack Overflow

Stack Overflow je CQA portál odborne zameraný na programovanie organizovaný výhradne pomocou používateľmi generovaných značiek a kľúčových slov. Odhaduje sa, že v súčasnosti existuje okolo 3000 značiek, ktoré kategorizujú obsah [8]. Generovanie nových značiek je obmedzené len pre skúsenejších používateľov na základe ich verejného hodnotenia.

Výhodou odborne zameraného fóra je, že celá komunita je zameraná na jednu špecifickú oblasť, vďaka čomu nie sú používatelia rozptyľovaní ostatnými nesúvisiacimi témami. Odpovede na otázky sú formulované konkrétne a sú hodnotené iba komunitou, ktorá problémovej oblasti rozumie. Používatelia majú taktiež možnosť hodnotiť otázky a vyzdvihnúť tie, na ktoré je pravdepodobne potrebná odpoveď. Hodnotenie samotných používateľov je založené na princípe reputácie, kedy každá pozitívna akcia používateľa zvyšuje jeho reputáciu.

Otázky formulované na Stack Overflow sa líšia od otázok na všeobecnejšom Yahoo Answers. Kvôli odbornosti neobsahujú samotnú otázku, ale aj stručný úvod do konkrétnej oblasti. Ak otázka neobsahuje dostatočný opis súvisiacich technológií, používatelia ju označia ako nedostatočnú a systém ju zablokuje. Podľa [9] je možné otázku rozdeliť na dve dimenzie:

- **Téma otázky.** Začiatok otázky väčšinou prezentuje technológiu alebo konštrukciu, s ktorou používateľ pracuje a až následne je položená samotná otázka. Takéto otázky sa týkajú API, návrhových princípov, algoritmov alebo porovnania dvoch a viac technológií.
- **Dôvod polozenia otázky.** Používateľ Stack Overflow kladie otázku z týchto dôvodov:
 - *Hľadanie chyby, revízia.* Spojené s problémami pri vývoji, chybami počas behu programu alebo nevhodným dizajnom.
 - *Hľadanie možnosti riešenia.* Otázka indikuje nedostatok vedomostí pýtajúceho sa a väčšinou sa kladie pri zisťovaní, či je niečo možné/dostupné v danej technológii.
 - *Ako to spraviť.* Hľadanie konkrétneho riešenia (kusu kódu) pre daný problém.

⁶ <http://www.tripadvisor.com/>

⁷ <http://stackexchange.com/>

- *Hľadanie iného riešenia.* Keď súčasné riešenie funguje, avšak z ľubovoľných príčin je potrebné nájsť iné riešenie.

Ku každému typu otázky sa viaže iný druh riešenia: diskusia, konkrétny zdrojový kód, slovný opis riešenia a otázku je nutné formulovať tak, aby používateľ dostal ten typ odpovede, ktorý očakáva.

Rozdeľovanie portálov na špecificky zamerané weby (v rámci siete Stack eXchange) naznačuje, že faktoidne orientovaná aplikácia (v našom prípade hra) by sa mala zameriavať na konkrétne domény a vyhýbať sa všeobecnosti. Špecifickosť jednoduchšie zacieli očakávanú skupinu hráčov, ktorých skupinová úroveň expertízy pre konkrétnu doménu môže byť vyššia, ako pri úplne náhodnom (všeobecnom) dave. Čerpanie dát z takéhoto davu dáva predpoklad na kvalitnejšie používateľmi-generované artefakty.

2.1.4. LastFM

LastFM je webová služba poskytujúca prehrávanie hudby za paušálny mesačný poplatok. Oproti ostatným podobným službám sa odlišuje v tom, že dovoľuje ľuďom verejne anotovať jednotlivé skladby, ktoré im služba hrá. Na základe týchto anotácií vie následne prehrávať žánrovo (anotačne) zamerané skladby a vytvárať tak personalizované rádiá.

Analýza týchto anotácií v našej predchádzajúcej práci ukázala, že poskytované anotácie sú objektívne, ale aj veľmi subjektívne [1] – niektorí používatelia používali svoju prezývku ako anotáciu aby skladbu vedeli neskôr dohľadať. To bol aj prvotný impulz prečo vôbec existujúce anotácie validovať.

Služba poskytuje API rozhranie, cez ktoré je možné dopytovať sa na jednotlivé anotácie. Jednu z vrátených hodnôt pre každú anotáciu je aj tzv. *počet*. Domnievame sa, že tento počet predstavuje počet ľudí (zhora ohraničený 100), ktorí danú anotáciu vložili do systému. Pri mnohých anotáciách (hlavne pri populárnych skladbách) je tento počet 0. Dedukciou a na základe analýzy týchto anotácií sa domnievame, že tieto patria príbuzným skladbám a interný algoritmus tieto anotácie vyhodnotil ako možných kandidátov aj pre danú skladbu.

LastFM teda postavila svoj produkt na dátach získaných davom - okrem anotácií je veľmi dôležitá aj história počutých skladieb jednotlivých používateľov, ich podobnosť a na základe toho generované odporúčania. Navyše do získavania týchto dát zapojilo len zlomok svojich používateľov – tých, čo používajú špeciálny softvér a ktorí to robia z čistého altruizmu. Oproti službám ako Pandora, ktoré majú svoj produkt postavený výhradne na generovaní dát expertmi a internými algoritmami je spomenutý proces inovatívny a učebnicovým príkladom využitia davu.

2.2. Hry s účelom

Hry s účelom sú špeciálnou kategóriou čerpania z davu, ktoré ako motiváciu poskytujú hráčom zábavu a sociálnu interakciu. Rovnako ako pri konvenčných prístupoch pri čerpaní z davu, aj hry s účelom sa snažia riešiť problém, ktorý má pri automatických prístupoch nízku úspešnosť. Hry s účelom vychádzajú z predpokladu, že ľudia na webe strávia veľké množstvo času hraním, ktorý by sa dal využiť na splnenie vybranej úlohy. Namiesto jednoduchého splnenia úlohy (napr. za finančnú odmenu) je ale táto úloha integrovaná do samotnej hry a hráči svojou aktivitou prispievajú k jej vyriešeniu. Dobre navrhnutá hra s účelom hráčovi priamo neprezentuje svoj účel a hráč ani nemusí vedieť, že rieši vybranú úlohu.

Hry s účelom sa často zamieňajú so hrami, ktoré nie sú primárne zamerané na zábavu – tzv. „*serious games*“. Tento typ hier môže byť použitý v školstve ako učebná pomôcka, v armáde a letectve ako typ simulátora, v zdravotníctve pri liečbe ochorení a pod.

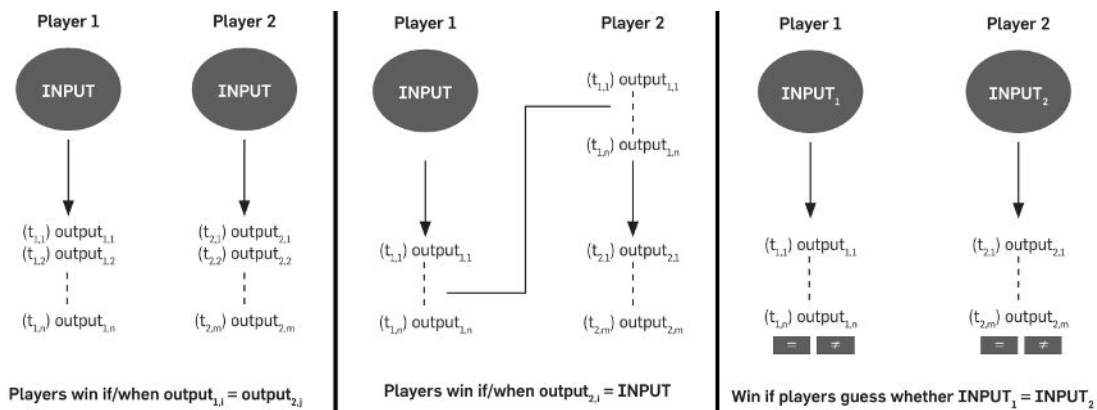
Cieľom hier s účelom je vytvárať dáta a metadáta k dokumentom. Účely hier sa líšia: anotácia obrázkov v hre ESP Game [3], určovanie pozície objektov v obrázkoch hrou Peekaboom [10], určovanie nálady skladieb v čase v Moodswings [11], ale aj vytváranie folksonómii v Little Search Game [12].

Keďže hráči nemusia vedieť o účele hry, môžu neúmyselne generovať aj nesprávne dáta. Aby sa zabezpečila kvalita výstupu z hry, prezentuje sa niekoľko prístupov ako výstup hráčov overovať [13] – jednotlivé prístupy sú graficky znázornené na obr. 2.2:

- **Overovanie výstupom.** Náhodne spárovaným hráčom je prezentovaný rovnaký dokument (obrázok, skladba) a ich úlohou je na základe zadania generovať dáta – napr. „čo vidíte na obrázku?“. Hráči svoje výstupy navzájom nevidia. Keď sa hráči zhodnú na niektorom zo svojich výstupov, postupujú ďalej. Vďaka tejto zhode je s veľkou určitosťou možné prehlásiť, že výstup, na ktorom sa hráči zhodli je správnou odpoveďou na zadanú úlohu k prezentovanému dokumentu.
- **Overovanie vstupom.** Náhodne spárovaným hráčom je prezentovaný dokument (môže a nemusí byť rovnaký). Hráči majú znovu generovať dáta, aby opísali im prezentovaný dokument. Na rozdiel od predchádzajúcej metódy, hráči svoje dáta navzájom vidia a na základe týchto dát sa majú rozhodnúť, či hra obom prezentovala rovnaký alebo rozdielny dokument.
- **Overovanie inverznými rolami.** Úlohy hráčov sú v tomto type hier rozdielne a striedajú sa. Jeden z hráčov na základe prezentovaného dokumentu musí vygenerovať čo najpodrobnejšie metadáta. Druhý hráč potom dostane niekoľko dokumentov (jeden z nich je rovnaký, aký videl prvý hráč) a na základe dát vygenerovaných prvým hráčom musí uhádnuť, ktorý dokument videl prvý hráč. Tento typ úloh je vhodný na generovanie podrobných dát, keďže herný mechanizmus k tomuto hráčov priamo motivuje.

Veľkým problémom hier s účelom je „problém studeného štartu“ („*cold start problem*“), kedy hra neobsahuje dáta, na ktorých by bolo možné stavať a ani základnú bázu hráčov, ktorí by dáta generovali [11]. Vyskytuje sa hlavne pri hrách, ktoré vyžadujú interakciu viac hráčov súčasne. Problém je čiastočne riešiteľný jedným z dvoch spôsobov [3]:

- **Predhrané hry.** Keď hra nedokáže spárovať hráča s iným hráčom, je možné použiť v už odohranú hru a proti reálnemu hráčovi „postaviť“ akcie iného hráča, ktoré boli vykonané v minulosti. Toto neovplyvňuje výstupy hry, keďže aj takáto zhoda medzi dvoma hráčmi je stále validná.
- **Využitie botov.** Okrem predhraných hier je možné použiť automatického bota, ktorý bude predstierať protihráča. Ak je bot navrhnutý dobre, skutočný hráč nemusí rozoznať, že proti nemu stojí počítač. Výstupy z takýchto hier ale nemusia byť vždy správne a záleží na implementácii samotného bota.



obr. 2.2 Overovanie výstupov v hrách s účelom. (vľavo - overovanie výstupom, v strede - overovanie inverznými rolami, vpravo - overovanie vstupom)

Okrem studeného štartu je častým problémom hier podvádžanie, proti ktorému neexistuje univerzálne riešenie. Podvádzanie je možné obmedziť náhodným párovaním hráčov (pri hre viacerých hráčov; jedným z pravidiel môže byť rozdielny IP rozsah) alebo testovaním hráčov na hre už známe fakty [14] – tzv. „bootstrapping“ – a vylúčenie výstupov tých hráčov, ktorí takýmto testom neprejdú. V tomto prípade ide ale hlavne o prevenciu pred hráčmi, ktorí generujú náhodné dáta a o zachovanie čistoty dát. Prevencia proti štandardnému typu podvádžania, ktoré neovplyvňuje získané dáta, je závislá od konkrétneho návrhu hry.

Jedným z najväčších problémov využívania davu je uprednostňovanie populárneho názoru (vo všeobecnosti očakávaného-štandardného výstupu) voči nepravdepodobným výstupom [15]. Dôsledkom tohto problému je neschopnosť davu generovať expertné artefakty, keďže navonok vystupujú rovnako ako nesprávne artefakty – dav nie je schopný so svojimi znalosťami overiť ich správnosť.

Napriek spomenutým problémom a napriek tomu, že aj renomovaný časopis *Newsweek*⁸ označil slovné spojenie „znalosť davu“ za oxymoron sú stále hry s účelom jednou z najpokročilejších metód na využívanie davu v súčasnosti [16].

Pri návrhu hry je dobré dbať na niekoľko faktorov [14]:

- **Identifikácia úloh v kontexte vytvárania sémantiky.** Nie všetky úlohy môžu byť do hry integrované tak, aby to neovplyvnilo hrateľnosť. Je dôležité ešte na začiatku určiť, ktoré úlohy je do hry potrebné nutne integrovať a ktoré by bolo dobre mať integrované.
- **Návrh herných scenárov.** Po určení úloh je tieto potrebné zakomponovať do herných scenárov tak, aby čo najmenej narúšali vplyv na hru.
- **Návrh rozhraní.** Hra musí ponúkať pútavé a jednoduché rozhranie pre hráča.
- **Identifikácia korpusu dát.** Nie všetky dáta vyžadujú rovnaké úsilie na zakomponovanie do hry a prípadnú aktualizáciu. Čím je jednoduchšie, tým lepšie (napr. YouTube videá, články z Wikipédie a pod.).

⁸ <http://www.newsweek.com/user-generated-content-out-84203>

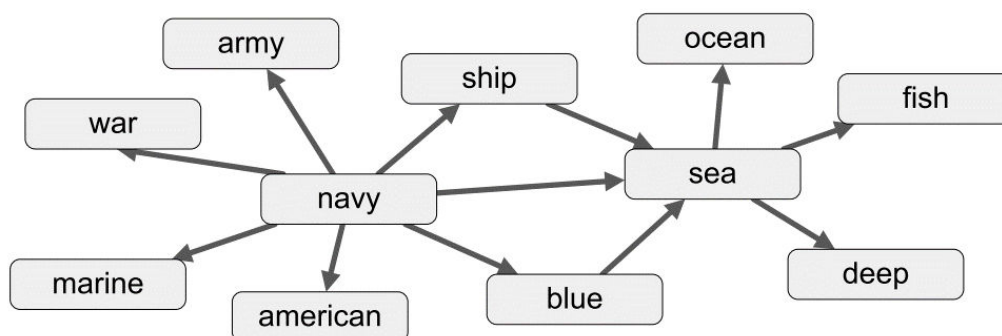
- **Prevenia pred podvádzaním.** Ako je naznačené v predchádzajúcej časti, možnosti podvádžania závisia od návrhu konkrétnej hry. Je dobré identifikovať tieto možnosti ešte na začiatku a myslieť na ne pri návrhu.
- **Odstránenie štandardných nástrah.** Hráči sa môžu ukliknúť alebo nemusia zvládať niektorú z úloh. Je dobré identifikovať takéto miesta a ak ich nie je možné úplne odstrániť, je dobré ignorovať výstupy z takto získaných miest pri konkrétnych hráčoch.
- **Podpora návratu hráča.** Hra by mala motivovať hráča hrať aj v budúcnosti a to aj v prípade jeho súčasnej neaktivity – rebríčkami, notifikáciami, odmenami a pod.
- **Efektívna distribúcia práce a využitie davu.** Prezentácia konkrétnych úloh by nemala byť náhodná a mala by motivovať hráčov k čo najskoršiemu dokončeniu už začatých riešení. Tiež je dobré dbať na veľkosť vzorky úloh, aby jeden hráč nedostával opakovane rovnakú úlohu.

V nasledujúcej časti sú opísané rôzne druhy hier s účelom s ich charakteristickými vlastnosťami. Zameriavame sa hlavne na hry vytvorené na Ústave informatiky a softvérového inžinierstva a na hry zaoberajúce sa získavaním metadát k hudbe.

2.2.1. Little Search Game⁹

Little Search Game [12][17] je hra pre jedného hráča, ktorej účelom je vytváranie väzieb medzi slovami – folksonómii – za pomoci webového vyhľadávania. Vyhľadávače štandardne podporujú odstránenie výsledkov dopytu, ktoré obsahujú zadané slovo – definovanie výsledkov na odstránenie štandardne prebieha pomocou prepínača „-“. Hra je založená na formulovaní takýchto dopytov pre vyhľadávače na základe zadaného slova a snahe dosiahnuť čo najmenší počet vrátených výsledkov resp. čo najväčší počet odstránených výsledkov.

Hráč napr. pre slovo „*apple*“, pre ktoré vyhľadávač vráti N výsledkov, musí zadať 6 ďalších slov, ktoré filtrujú nájdenú množinu výsledkov, napr. „*-fruit*“ alebo „*-big*“. Hráči sú sami motivovaní zadávať termy so silnou väzbou na zadané slovo, keďže takéto termy majú tendenciu odstrániť najväčšie množstvo vrátených výsledkov z pôvodného dopytu. Príklad siete vytvorených väzieb medzi slovami pomocou hry je zobrazený na obr. 2.3.



obr. 2.3 Sieť termov vytvorená pomocou Little Search Game

Vďaka návrhu hry je hra univerzálne použiteľná na ľubovoľný jazyk a je závislá iba od schopností vyhľadávačov. Vyhľadávače zároveň poskytujú jediný potrebný dataset a hra netrpí problémom studeného štartu. Možnosť hry pre jedného hráča výrazne obmedzuje možnosť podvádzať.

⁹ <http://mirai.fiit.stuba.sk/LittleGoogleGame>

Výsledky ukazujú, že hra dokáže generovať konkrétnejšie folksonómie a navyše vďaka návrhu hry je možné priamo špecifikovať cieľové koreňové slovo. Hra zároveň dokáže generovať väzby, ktoré nie sú priamo rozoznateľné štatistickou analýzou jazykového korpusu [12].

2.2.2. Herd-It

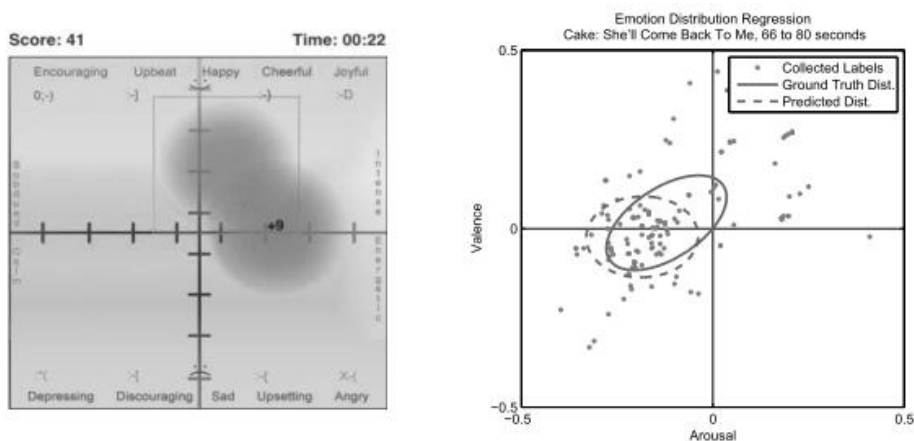
Hra postavená na sociálnej sieti Facebook využíva sociálne prvky ako prídavný zdroj motivácie. Je navrhnutá pomocou minihier, kde každá minihra získava iný druh dát: náladu skladby, chytlavosť skladby, farbu skladby, cieľovú skupinu skladby, časť dňa, kedy je vhodné skladbu prehrať, adjektíva opisujúce skladbu a iné. Hra teda plne využíva potenciál davu na získanie dát, ktoré by automatizovanými prístupmi neboli možné. Herd-It ponúka možnosť výberu hudobného žánru, s ktorým chce hráč hrať. V súčasnosti (máj 2013) hra nie je dostupná.

Hráč hrá niekoľko minihier proti rovnakým súperom, ktorí sú buď jeho priatelia alebo náhodne vybraní hráči. Každá minihra trvá 10-20 sekúnd a obsahuje inú hudobnú ukážku. V prípade, že nie je pripojený dostatočný počet hráčov, hra využíva ako súperov botov. Bodovanie je založené na zhode s ostatnými hráčmi – čím viac ostatných hráčov zvolilo rovnaký výstup, tým lepšie je hráčove hodnotenie (princíp zhody na výstupe). Hráči výstupy negenerujú, vždy sú im ponúknuté možnosti, z ktorých si majú vybrať – teda skladby sa selektujú do vopred určených kategórií.

2.2.3. Moodswings¹⁰

Moodswings je hra s účelom zbierať subjektívnu náladu skladby v čase – počas celého trvania skladby [11]. Hra je navrhnutá pre dvoch hráčov a úlohou hráčov je umiestňovať kurzor myši v dvojdimenzionálnom priestore tak, aby čo vykreslená kružnica čo najviac zodpovedala aktuálnemu stavu skladby – zhoda na výstupe. Priebeh hry a analýza dát je zobrazená na obr. 2.4.

Hráči štandardne nevidia umiestnenia kurzora spoluhráča (ten sa objavuje náhodne raz za niekoľko sekúnd) a sú odmeňovaní za čo najväčšiu zhodu s umiestnením kurzora spoluhráča. Rádus kurzoru sa počas jedného kola neustále znižuje a motivuje hráča určovať náladu skladby čo najpresnejšie. Aby sa predišlo podvádžaniu vo forme statického kurzoru, hra núti hráča pohnúť s kurzorom, aby bolo možné vykonať hodnotenie prekryvu s druhým hráčom. Získané dáta sú potom analyzované a nálada skladby je určená z prekryvu rádiusu kurzorov.



obr. 2.4 Priebeh hry Moodswings (vľavo) a analýza zozbieraných dát (vpravo)

¹⁰ <http://music.ece.drexel.edu/mssp>

2.2.4. City Lights¹¹

Hra sa na rozdiel od ostatných hier, ktorých primárny účel je vytvárať dáta, zameriava primárne na overovanie existujúcich dát – hudobných anotácií získaných pomocou davu na serveri LastFM [1]. Na overovanie využíva implicitnú aj explicitnú spätnú väzbu, ktorú hráč pomocou hrania hry generuje.

Hráč počuje niekoľko sekundovú ukážku skladby vybraného žánru a dostáva na výber niekoľko množín hudobných anotácií. Jedna z množín patrí podľa LastFM skladbe, ktorú hráč počuje. Na základe toho, čo hráč počuje, sa má rozhodnúť, ktorá z ponúknutých množín patrí hranej skladbe. Aby bola hra zábavnejšia a zároveň bolo možné získať presnejšie dáta (istotu hráča), hra zavádza možnosť staviť body na svoj tip – pri správnom tipe je stávka zdvojnásobená (za predpokladu, že je to prvá stávka) a pri nesprávnom tipe hráč stráca stavené body. Tento model redukuje generovanie náhodných dát, kedy by hráči tipovaním hľadali správnu odpoveď. Návrh hry pre jedného hráča zároveň odstránil problémy so studeným štartom a výrazne obmedzil možnosti podvádžania.

Hra na overovanie metadát využíva dve techniky:

- **Implicitná spätná väzba.** Pomocou analýzy hráčových stávok na konkrétne množiny hudobných anotácií a správnosti týchto stávok je možné určiť kvalitu ponúknutej množiny anotácií pre danú skladbu. Aj napriek tomu, že hráč primárne pracuje s množinami anotácií, vďaka ich náhodnému výberu do množín sa tieto v množinách striedajú a po odohraní niekoľkých hier je možné určovať aj kvalitu jednotlivých anotácií.
- **Explicitná spätná väzba.** Po uhádnutí správnej množiny anotácií hra ponúka hráčovi možnosť vyradiť z množiny anotácie, ktoré podľa neho k prehrávanej skladbe nepatria. Táto väzba podporuje implicitnú analýzu a skvalitňuje získané dáta.

Experiment ukázal, že hra veľmi presne (0% chybovosť) a zároveň rýchlo (rádovo niekoľko hier) dokáže určiť nesprávne hudobné anotácie k daným skladbám a relatívne presne (~33% chybovosť) dokáže určiť správne anotácie. Zároveň okrem validácie existujúcich anotácií ponúka priestor na sekundárne vytváranie anotácií vďaka nesprávnym hráčovým pokusom.

2.3. Diskusia

V tejto časti sme prezentovali dva principiálne odlišnú prístupy čerpania z davu zamerané na prácu s používateľmi a získavania dát. Keďže sa v ďalšej časti zaoberáme objavovaním autorít a teda priamou prácou s používateľmi, vyplývajú z analýzy čerpania davu nasledujúce dôsledky:

Potreba motivácie. V rámci nami navrhovanej hre musíme riešiť dva druhy motivácie: (1) motiváciu riešiť časovo náročnejšie úlohy a (2) motiváciu vracať sa k hre aj v budúcnosti. V oboch prípadoch je náš cieľ postavený na získaní čo najväčšieho množstva metadát od jedného používateľa. Tento cieľ môžeme riešiť typom hry, kde sú hráči priamo konfrontovaní s ostatnými hráčmi. Porovnanie s ostatnými hráčmi na úrovni jednotlivých úloh môže hráčov dostatočne motivovať k vyriešeniu čo najväčšieho množstva úloh – aj tých časovo náročnejších.

¹¹ <http://citylights.rootpd.com>

Odmeňovanie. Hráči môžu byť odmeňovaní na základe obtiažnosti úlohy. Aby bola ich motivácia dostatočná, môže odmeňovanie voči obtiažnosti úlohy rásť nelineárne (napr. kvadraticky).

Určenie hranice chybovosti. Keďže hráči v rámci úloh generujú dáta, je pravdepodobné, že časť z nich nie je správna. Čím vyššia odmena je za úlohu, tým väčšie je riziko, že hráč vygeneroval hoci aj zlé dáta len za účelom získania odmeny. Pri takomto type úloh musíme sledovať globálnu chybovosť v rámci úlohy, chybovosť hráča na danom type úlohy a chybovosť hráča celkovo. Na základe týchto parametrov sa potom môžeme rozhodovať kedy je chybovosť natoľko veľká, že hráčove výsledky nie sú brané do úvahy.

Overenie používateľmi vygenerovaných artefaktov. Aby sme zistili, či sú hráčmi generované dáta správne, musíme tomu prispôbiť návrh hry. Kvôli problému studeného startu sa nemôžeme spoliehať na priame overovanie na výstupe a správnosť dát môžeme vyhodnocovať spätne. V prípade, že rovnakú úlohu bude riešiť niekoľko hráčov súčasne, bude zaujímavé sledovať získané dáta a môžeme sa pokúsiť ich vyhodnocovať agregovane príp. určovať správnosť získaných informácií už na základe takto súčasne získaných dát a nevyhodnocovať dáta spätne. Oba prístupy neskôr môžeme porovnať.

3. Objavovanie autorít

Autorita (expert) je podľa Bransford et al. [18] človek, ktorý uchováva znalosti v takej forme, ktorá mu umožňuje rozoznať základnú kostru problému a jeho príbuzné dimenzie a pomocou ktorých je schopný nájsť efektívnejšie riešenie problému než laici. V tejto práci sa zameriavame na autority v oblasti hudobných znalostí, ich objavovanie a využitie na skvalitnenie hudobných metadát.

Rozoznávanie expertov je vo všeobecnosti potrebné z dvoch dôvodov [19]:

- **Expert ako zdroj informácií.** Autorita vystupuje ako komplementárny alebo suplementárny prvok k prístupným zdrojom (napr. dokumenty). Využitie autority v takomto prípade sa dá rozdeliť na niekoľko účelov:
 - *Špecifikácia.* Keď používateľ nepozná/nevie čo potrebuje a autorita pomôže tento cieľ identifikovať.
 - *Efektivita.* Keď používateľ nechce informáciu zdĺhavo vyhľadávať a využije znalosti experta na ušetrenie času.
 - *Interpretácia.* Keď používateľ nemá záujem/nedokáže pochopiť informácií a expert ju podá zjednodušene.
 - *Prístup k nezdokumentovaným informáciám.* Nie všetky informácie sú dostupné v katalógoch a na webe. Vtedy je potrebné obrátiť sa na experta.
 - *Socializácia.* Keď používateľ vyžaduje ľudskú interakciu pri hľadaní informácie.
- **Expert ako účastník s úlohou.** Využíva sa napr. pri prezentáciách, konzultáciách, rozhovoroch a pod., kedy autorita zastáva, obhajuje alebo prezentuje určitý názor z doménovej oblasti, ktorej sa venuje.

V minulosti bolo možné autority rozoznávať len v rámci menšieho okruhu ľudí v konkrétnej záujmovej doméne (kvôli málo rozšírenému prístupu na Internet). Experti boli združovaní v expertných databázach a katalógoch rozdelených podľa domény: ProfNet, ExpertNet, Virginia Tech Enterprise Database a iné. Tieto databázy boli využívané hlavne v podnikovej oblasti, kde pomáhali zefektívňovať prácu rádových zamestnancov. Databázy museli byť vytvárané manuálne, buď človekom na to určeným alebo dobrovoľne samotnými expertmi, ktorí do databázy pridali svoje zručnosti. Nanešťastie databázy trpeli neúplnosťou a ich prínos nedosiahol plánovaný dosah.

V kontexte práce vychádza potreba hľadania autority z potreby vyhľadávania relevantných informácií – expert je braný ako zdroj informácií. Vyhľadávanie na webe zvyčajne prebieha jedným z dvoch štandardných spôsobov:

- **Vyhľadávanie pomocou kľúčových slov, vyhľadávanie v katalógu.** Vracia výsledky rôznej kvality a laický používateľ je zvyčajne nútený sám ohodnotiť kvalitu nájdených výsledkov len na základe svojich skúseností. Vyhľadávače zoraďujú výsledky na základe interných hodnotiacich algoritmov („PageRank“, „sRank“) a je pravdepodobné, že nájdené dokumenty na začiatku zoznamu sú kvalitnejšie – relevantnejšie – ako výsledky umiestnené nižšie.

- **Položenie otázky v ľudskej reči.** Okrem vyhľadávania sú štandardom aj fóra, kde sa používateľ na svoj problém môže spýtať (*Yahoo Answers*¹², *Stack Overflow*¹³), avšak aj tieto trpia problémom nutnosti vedieť ohodnotiť kvalitu odpovede iných používateľov. Oproti vyhľadávaniu pomocou kľúčových slov prinášajú možnosť spýtať sa na špecifický problém a dostať konkrétnu odpoveď od ostatných používateľov.

Využitie hľadania odpovede položením špecifickej otázky so sebou prináša niekoľko problémov. Keďže odpoveď nie je hľadaná strojovo, nie je možné deterministicky určiť kvalitu získanej odpovede na základe vopred daného algoritmu. Kvalita odpovede je najčastejšie hodnotená používateľom, ktorý otázku položil, prípadne používateľmi, ktorí na otázku/odpoveď narazili. Získanie odpovede a zároveň aj ohodnotenie kvality odpovedí sú formou získavania dát a metadát pomocou davu.

Objavovanie autorít ako také je samostatná oblasť štúdia, avšak jej výsledky nie sú prezentované samostatne, ale v rámci vyhľadávacích systémov, odporúčacích systémov alebo informačných systémov. V tejto práci sú výsledky integrované do hry s účelom, kde sú autority využívané na tvorbu a validáciu hudobných metadát.

3.1. Typy autorít

Pri využívaní autorít ako takých nie je dôležité, v akom prostredí bola táto autorita objavená: či už 1) na webe alebo 2) v podnikovom prostredí. Toto rozdelenie je ale dôležité pri rozpoznaní autorít, keďže toto sa v jednotlivých prostrediach diametrálne odlišuje. Nami navrhovaná metóda je využívaná v hre s účelom využívajúcej dav na webe, preto pre zvyšok práce je dôležitá hlavne časť 3.1.1, avšak pre porovnanie uvádzame aj spôsob rozpoznania a využitia autorít v podnikovom prostredí (3.1.2).

3.1.1. Autority na webe

Každý používateľ webu za sebou necháva stopu – dáta, podľa ktorých je možné určiť jeho charakteristiky: pohlavie, vek, doménu, ktorej sa venuje a iné. Presné určenie domény používateľa a jeho odbornosti v nej je dôležité pri hodnotení jeho aktivít a ich vplyvu na ostatných používateľov.

Najčastejším miestom, kde je potrebné využiť odbornosť používateľa, sú doménovo zamerané internetové fóra (Java fórum) alebo tzv. stránky otázok a odpovedí („*Q&A sites*” – Stack Overflow). Princíp oboch typov stránok je podobný. Používatelia na tomto type webov sa podľa vypracovanej prípadovej štúdie delia na tri skupiny [20]:

- pýtajúcich sa: ~64%
- odpovedajúcich: ~32%
- používatelia patriacich do oboch predchádzajúcich skupín: ~4%

Zo všetkých používateľov je možné označiť ako autoritu 0,5% - 0,7% z nich. V priemere strávia aktívni používatelia takýchto služieb ročne pri odpovedaní otázok okolo 100 hodín. Motivácií ľudí odpovedať ostatným na ich otázky je niekoľko [21]:

- altruizmus (nesebeckosť).

¹² <http://answers.yahoo.com/>

¹³ <http://stackoverflow.com/>

- potreba podporiť komunitu.
- výhody získané od poskytovateľa služby.
- snaha o zabezpečenie získania pomoci v prípade potreby.
- priame získanie vedomostí pri tvorení odpovede na otázku.

Táto motivácia je pre používateľov postačujúca, keďže viac ako 92% otázok položených na službe Stack Overflow má medián dĺžky získania odpovede okolo 11 minút [22]. To zároveň stavia služby založené na kladení otázok na porovnateľnú úroveň s prehľadávaním stránok zameraných na zdieľanie obsahu (blogy, články a pod.).

3.1.2. Autority v podnikových systémoch

Podnikové systémy sa v kontexte využitia autorít líšia od webu v tom, že využívajú špecifické znalosti expertov a poznajú doménu/zaradenie jednotlivých expertov v systéme. Experti sú využívaní hlavne na špecifikáciu a interpretáciu neúplných informácií – konzultáciou s expertom sa zvyšuje efektívnosť práce a teda aj rýchlosť vývoja produktu. Softvérový inžinier strávi v priemere komunikáciou s expertom až 16% času [23] – tento čas nezahŕňa všetku komunikáciu inžiniera pri práci.

Potreba vyhľadávať expertov sa zväčšila globalizáciou a členením tímov na veľké vzdialenosti [24], avšak problém s vyhľadaním správneho človeka existuje v rovnakej miere aj medzi 30 metrov vzdialenými kancelármi. Spôsoby, ktoré sú využívané na odhalenie expertov sú [24]:

- **Sledovanie zmien v dokumentoch.** Čím väčšia bola zmena od vybraného používateľa a čím menej času od tejto zmeny prešlo, tým väčšia je pravdepodobnosť že pozná správanie danej časti systému a môže byť považovaný za experta.
- **Heuristické prístupy.** Systémy sledovali zmeny používateľov naprieč systémom a pri vyhľadávaní experta brali do úvahy počet zmien v dokumentoch, vzdialenosť experta od žiadateľa v rámci organizácie (v kontexte pracovného zaradenia na základe predchádzajúcej analýzy sociálnych väzieb) príp. ďalšie podnikovo-špecifické parametre. V prípade, že systém nedokázal nájsť expertov, je možné manuálne znížiť hraničné hodnoty pre heuristický algoritmus.
- **Expertné databázy.** Ich najväčší problém je neúplnosť, keďže nie je možné obsiahnuť a udržať aktuálne konkrétne znalosti všetkých zamestnancov. Znalosti sú väčšinou reprezentované všeobecne, čo pri špecifických dopytoch vnáša do výsledkov ďalšie nepresnosti a tento model sa preto nezvykne používať.

Najpoužívanejším prístupom sú rôzne variácie heuristických prístupov špecifických pre jednotlivé organizácie. Jedným zo všeobecnejších prístupov je sledovanie väzieb expertov na dokumenty, ktorý je opísané v časti 3.2.2.

3.1.3. Autority v hrách s účelom

Väčšina prístupov v hrách s účelom využívajú autority len na vytváranie kontrolnej vzorky dát príp. manuálnu kontrolu výstupov z hry. Autority sú vybrané tvorcom aplikácie.

V súčasnosti je nám známy len jeden projekt, ktorý sa zaoberá optimalizáciou využitia autorít v hrách s účelom s cieľom skvalitniť metadáta. Jedná sa o modifikáciu hry ESP Game [3] – hra, v ktorej sú dvaja hráči náhodne spárovaní a ich úlohou je zhodnúť sa v tom, čo vidia na spoločnom

obrázku. Nepriamo tak anotujú obrázok a vďaka zhode oboch hráčov v reálnom čase je zabezpečené aj overenie správnosti anotácie.

Modifikácia je postavená na párovaní na základe schopnosti hráčov (*CAM* – „*capability aligned matching*“). Pod párovaním rozumieme: (1) inteligentné párovanie samotných hráčov na základe ich schopností a zároveň aj (2) párovanie obtiažnosti anotovaných obrázkov so schopnosťami hráčov [25]. V práci boli navrhnuté dva modely hráčov, kedy jeden pozná schopnosti hráča a druhý ich musí odhadovať.

Navrhovaný model je postavený na ekonomickom princípe, kedy server vypočítava hodnotu odmeny za nové anotácie k danému obrázku a hráči si vyberajú obrázky na základe výšky odmeny. Odmena je počítaná na základe vektora obsahujúceho už existujúce anotácie, ich kvalitu a predchádzajúce pokusy. Keď sa hráči zhodnú na anotácií, server vypočíta odmenu pre hráčov na základe vygenerovanej anotácie – teda hráči nemusia dostať plnú sľubovanú odmenu, zvyšok zostáva serveru. Hráči si teda na základe výšky vybranej odmeny sami určujú obtiažnosť hry a tým implicitne aj svoju expertízu.

3.2. Rozpoznávanie autorít

Rozpoznanie experta všeobecne pozostáva z dvoch úloh [26]:

- rozpoznanie a identifikácia oblastí záujmu používateľa
- ohodnotenie odbornosti používateľa v rozpoznaných oblastiach záujmu

Identifikácia záujmov je možná len na základe používateľom zadaných dát – či už formou testu kde sám určí oblasti záujmu alebo analýzou jeho aktivity. Ohodnotenie odbornosti môže prebiehať automaticky na základe vybraných metrík alebo ostatnými používateľmi – davom. Prístupy na ohodnotenie odbornosti sú opísané v nasledujúcej časti.

3.2.1. Metriky na hodnotenie autority používateľa v dave

Metriky hodnotenia odbornosti používateľov stavajú na skúsenostiach získaných pri hodnotení dokumentov objavených počas indexovania webu a na algoritmoch, ktoré sa pri tomto hodnotení používajú. Na základe správania sa používateľov v rámci Java fóra bolo porovnaných a vyhodnotených niekoľko najpoužívanejších prístupov na automatické hodnotenie odbornosti [27]:

- jednoduché štatistické prístupy. (počty odpovedí)
- InDegree
- Z-skóre
- ExpertiseRank
- HITS

Priame porovnanie ukázalo [27], že každá z metrík je vhodná na ohodnotenie určitej stránky odbornosti používateľa a neexistuje univerzálna metrika. Jednoduché štatistické prístupy sa ukázali ako porovnateľne dobré s pokročilými iteratívnymi metrikami. Porovnanie taktiež ukázalo, že použitie automatického ohodnotenia odbornosti pomocou jednotlivých metrík s využitím sociálnych sietí môže byť za predpokladu správneho nastavenia siete označené ako dôveryhodné aj bez ďalšej nutnosti overovania používateľmi.

3.2.1.1. Jednoduché štatistické prístupy

Jedným z najjednoduchších a často používaných prístupov je počet odpovedí používateľa na otázky ostatných používateľov. Problémom tejto metriky je nerozlišovanie odbornosti samotnej odpovede a obtiažnosti zadanej otázky. Taktiež neberie do úvahy možný SPAM, ktorý sa na fórach vyskytuje. Podobná metrika riešiaci spomenuté problémy je sčítanie počtu ľudí, ktorým používateľ svojou odpoveďou pomohol. Tým sa do určitej miery zabezpečí vyššie hodnotenie pre používateľa, ktorý odpovedá na náročnejšie otázky.

3.2.1.2. InDegree metrika

Zovšeobecňuje jednoduché štatistické prístupy a prezentuje ich vo forme grafu. Metrika samotná meria autoritu vrcholu (používateľa) počtom hrán, ktoré do daného vrcholu smerujú. Pri použití hrán je možné brať do úvahy aj váhu jednotlivých hrán – výhradne pri použití takejto metriky je to pre výpočtovú hodnotu nevyhnutné [20].

3.2.1.3. Z-skóre

Metrika stavia na predpoklade, že kladenie otázok značí nevedomosť – opak odpovedania na otázky. Ak používateľ odpovedal na rovnaký počet otázok aký položil, Z-skóre je nulové. Ak odpovedal na viac otázok ako položil, skóre je kladné. Skóre sa počíta v dvoch variáciách: 1) pre počet položených/zodpovedaných otázok a 2) pre počet používateľov, ktorým meraný používateľ odpovedal a počet používateľov, ktorí meranému používateľovi odpovedali. Skóre sa počíta podľa nasledovného vzorca, kde a je počet odpovedí a q je počet otázok (pre druhý variant analogicky):

$$z = \frac{a - q}{\sqrt{a + q}}$$

3.2.1.4. Algoritmus „ExpertiseRank“

Predchádzajúce metriky do určitej miery trpeli problémom, kedy používateľ odpovedal vo veľkom množstve na jednoduché otázky. Metriky ho označili za experta v doméne. Na základe metriky „PageRank“ [28], ktorá hodnotí relevanciu dokumentov na základe relevancie webov, ktoré na daný dokument odkazujú, vytvorili v [27] metriku založenú na 1) počte ľudí, ktorým používateľ pomohol a 2) odbornosti ľudí, ktorým používateľ pomohol. Algoritmus na výpočet odbornosti je ($C(U_i)$ – počet ľudí, ktorým používateľ pomohol, d – činiteľ tlmenia / „damping factor“):

$$ER(A) = (1 - d) + d (ER(U_1)/C(U_1) + \dots + ER(U_n)/C(U_n))$$

Metrika počíta odbornosť v iteráciách a vychádza z predpokladu, že ak používateľ B pomohol používateľovi A a používateľ C pomohol používateľovi B, používateľ C ma z troch spomenutých najvyššiu odbornosť, keďže pomohol používateľovi, ktorého miera odbornosti už bola na vyššej úrovni ako miera odbornosti iných používateľov úrovni. Metrika využíva grafy a šírenie odbornosti používateľa v grafe.

3.2.1.5. HITS algoritmus

Podobne ako PageRank/ExpertiseRank, aj HITS metrika [27] využíva iteratívny prístup na získanie odbornosti používateľa. Stavia na algoritme HITS [29], ktorý sa na cieľový dokument/používateľa pozerá z dvoch kontextov (opis algoritmu je tvorený z pohľadu hodnotenia odbornosti):

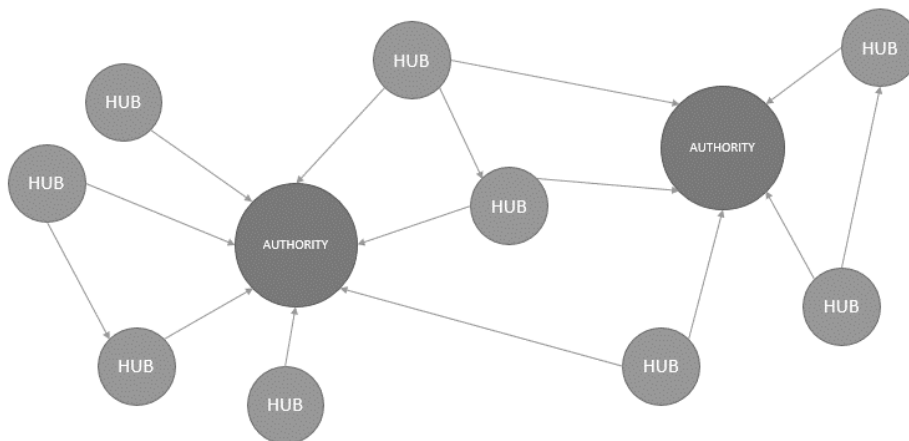
- 1) Ako zdroj pozornosti („hub“). Počet expertných používateľov, ktorí pomohli vybranému používateľovi

2) *Ako autoritu.* Počet expertných používateľov, ktorým vybraný používateľ pomohol.

HITS algoritmus stavia na grafe homogénnych vrcholov - kde vrcholy môžu predstavovať jednotlivé weby prípadne používateľov – a hranách, ktoré predstavujú orientované prepojenia medzi týmito entitami – napr. vzájomne odkazovanie medzi webmi alebo interakcia niekoľkých používateľov. Typ vrcholov je ale vždy homogénny.

Na je znázornený graf, s ktorým môže HITS algoritmus pracovať. V kontexte objavovania autorít môžeme orientáciu hrán interpretovať spôsobom, kedy si „vrchol *hub* nechť pomôcť od vrcholu *authority*“. Na základe tejto interpretácie je teda z grafu možné odčítať, že vrcholy označené ako autority pomohli výrazne väčšiemu počtu používateľov a môžu byť považované za autority. Keby niektorý z *hub* vrcholov vyžadoval pomoc od veľkého počtu používateľov, jeho *hub* skóre by bolo väčšie ako u ostatných vrcholov a bol by teda skutočne väčší „zdroj pozornosti“.

pozn.: Na obrázku je ako menovka vrcholu zobrazená iba jedna z hodnôt „hub“, „authority“ (tá väčšia). V skutočnosti sa pre každý vrchol počítajú obe tieto hodnoty.



obr. 3.1 Graf vrcholov vstupujúci do HITS algoritmu

3.2.1.6. Iné metriky

Väčšina ostatných používaných metrik zakladá na vyššie spomenutých základných metrikách [30]. Algoritmus SALSA spája výhody algoritmov PageRank a HITS. Algoritmus *HubAvg* mení algoritmus HITS tak, že váhu centra pozornosti („*hub*“) nastaví na priemer váh ostatných centier pozornosti, na ktoré pozorovaný vrchol grafu smeruje.

Ďalšie pokročilé techniky [30] sú hlavne zamerané na analýzu dokumentov a odkazov na ne na webe – BFS („*Breath-First-Search*“) algoritmus alebo Bayesovské algoritmy – a nebudú v tejto práci ďalej analyzované z dôvodu nízkeho využitia pri objavovaní autorít.

3.2.2. Odbornosť experta na základe dopytu

Jedným zo spôsobov na rozpoznanie autority je prístup navrhnutý v [31]. Využíva navrhnuté modely dokumentov a kandidátov na autoritu. Prístup uvažuje prácu s heterogénnym repozitárom dokumentov, pre ktorý je zadaný dopyt na základe ktorého určuje odbornosť tzv. kandidáta. Dopyt sa skladá z niekoľkých termov a nie je nutné aby vrátený dokument úplne vyhovoval zadaným termom. Model stavia na otázke: „Aká je pravdepodobnosť p , že kandidát ca je expertom pre tematický dopyt q “? Po aplikovaní Bayesovej vety na konkrétny model je pravdepodobnosť, že používateľ je expert pre daný dopyt nasledovná:

$$p(ca|q) = \frac{p(q|ca) * p(ca)}{p(q)}$$

Na základe existujúcich modelov vytvorených pri výskume dolovania informácií („*Information Retrieval*“) boli vytvorené dva modely [31]: 1) model zameraný na dokumenty a 2) model zameraný na kandidátov na autoritu. Experiment preukázal, že model zameraný na dokumenty má výrazne lepšie výsledky a preto sa v tejto časti budeme ďalej zaoberať iba ním.

Model pracuje s váženou väzbou a medzi entitami dokument (d) – kandidát (ca). Dokument je vybraný z množiny dokumentov D . Pravdepodobnosť, že dokument je naviazaný na kandidáta je definovaná nasledujúcim vzorcom. Myšlienkou za prevodom váženej väzby $a(d,ca)$ na pravdepodobnosť je normalizácia váhy medzi kandidátmi, ktorí majú väzbu na vybraný dokument.

$$p(d|ca) = \frac{a(d, ca)}{\sum_{d' \in D} a(d', ca)}$$

Rozoznávanie expertov pri modeli zameranom na dokumenty uvažuje, že kandidát na autoritu je vopred daný. Vyhľadávanie prebieha nasledovne:

- Z kolekcie dokumentov D je vybraný dokument s väzbou na vopred určeného kandidáta – na základe $p(d|ca)$.
- Z dokumentu sa odvodí jednoduchý vyhladený model a vypočíta sa pravdepodobnosť $p(q|d)$ - že dopyt q bude sedieť na vybraný dokument na základe termov dopytu t . Vyhladenie modelu dokumentu prebieha metódou Jelinek-Mercer, ktorá zahŕňa lineárnu interpoláciu maximálneho pravdepodobnostného modelu dokumentu a modelu kolekcie dokumentov [32] – ich vplyv nastavuje konštantou λ .
- Vypočíta sa váha väzby $p(q|ca)$ medzi dopytom a kandidátom na experta

$$p(q|d) = \prod_{t \in q} ((1 - \lambda) * p(t|d) + \lambda * p(t))$$

$$p(q|ca) = \sum_d p(q|d) * p(d|ca)$$

Výhodou spomenutého modelu je to, že pri vytvorení a ohodnotení väzieb medzi dokumentom a kandidátom je implementácia modelu do vyhľadávacieho systém jednoduchá. Model mal po vrátení 10 kandidátov hodnotenie metriky *precision* ~ 0.316 a strednú hodnotu metriky *precision* na celej kolekcií výsledkov ~ 0.188. Experimenty prebiehali na datasete konferencie TREC¹⁴ a aj napriek relatívne slabej úspešnosti by sa model umiestnil v TOP 5 navrhnutých prístupov na vyhľadávanie expertov.

3.2.3. Hodnotenie odbornosti využitím znalostí davu

Sociálne siete umožňujúce verejne zdieľanie sú veľmi dobrým zdrojom na objavovanie autorít. V takýchto sieťach sa používatelia obvykle delia ľudí, čo sledujú a sieť používajú na získanie

¹⁴ <http://trec.nist.gov/>

informácií a na ľudí, ktorí tieto informácie zdieľajú – táto menšia skupina je zaujímavá v kontexte objavovania autorít.

Na sociálnej sieti Twitter¹⁵ sa na objavovanie autorít využívajú okrem štandardných možností ako počet ľudí, ktorí vybraného používateľa sledujú aj systémovo-špecifické prvky ako je počet zmienok o konkrétnom používateľovi alebo znovuzdieľanie jeho správ [33]. Autorita v kontexte zadanej domény je v takomto prípade rozoznaná najlepšie vtedy, keď distribúcia záujmových tém prvej skupiny (sledujúci používatelia) je čo najväčšia a keď je počet správ možných autorít čo najviac zameraný na vybranú tému. Twitter sám vyvinul službu „*Twitter WTF (Who To Follow)*“¹⁶, kde na základe používateľových záujmov odporúča rozoznané autority v danej oblasti.

Tradičné prístupy rozoznania autorít v záujmovej oblasti konkrétneho používateľa teda spoliehajú na čo najpresnejšie rozoznanie záujmov používateľa. Záujmy môžu byť rozoznané na základe používateľovho profilu – teda sám používateľ zadá, čo ho zaujíma – alebo na základe analýzy jeho aktivity na sociálnej sieti.

Twitter ponúka okrem možnosti zdieľať informácie aj možnosť vytvárať si zoznamy používateľov. Takéto zoznamy sú pre používateľa vhodné, ak chce oddeliť príspevky z rôznych doménových oblastí. Zoznam vytvorený jedným používateľom už ponúka relatívne dobré rozdelenie sledovaných autorít do domén na základe názvov zoznamov a spoločných vlastností autorít v zozname. Analýzou všetkých verejných zoznamov na sociálnej sieti je ale možné veľmi presne určiť autority v konkrétnych doménových oblastiach [34]. Pri veľkom množstve používateľov sa rozdiely v jednotlivých zoznamoch odignorujú a dav je schopný nielen určiť autority (účty opakovane objavujúce sa v zoznamoch), ale aj ich zaradiť do špecifickej domény a priradiť im konkrétne metadáta – napr. účet *BarrackObama* dostal na základe zaradenia do zoznamov značku „politik“. Oproti iným prístupom, ktoré využívajú viacero systémovo-špecifických funkcií (vrátane *Twitter WTF*) je prístup založený výhradne na zoznamoch vytvorených dobrovoľne davom na rozoznanie autorít porovnateľne dobrý.

3.2.4. Autorita reprezentovaná skupinou expertov

Nie každý expert má úplne vedomosti vo svojej doménovej oblasti. Uchovávanie úplných informácií o používateľovej odbornosti by bolo neefektívne. V prípade zhromažďovania informácií je možné využiť na jeden problém znalosti skupiny expertov. Z hľadiska uchovávania informácií o expertoch by to znamenalo uchovávanie generickejších, ale stále dostatočne presných metadát o používateľoch – napr. namiesto uchovania informácie o znalosti hud. štýlu „*British indie rock*“ by sa uchovala informácia o znalosti štýlu „*indie rock*“ („*indie*“).

Za predpokladu, že sme schopní relatívne dobre rozoznať znalosti používateľov, prispelo by využitie tejto techniky pri dolovaní informácií pomocou davu k jeho efektívnejšiemu využitiu. Nebolo by nutné spoliehať sa výhradne na expertov, ale bolo by možné zapojiť semi-expertov s čiastočnými vedomosťami v doméne. Na získanie/potvrdenie informácie by sa teda využívalo niekoľko odborníkov zároveň, čo prináša problém určenia „koľko odborníkov“ je na čo najpresnejšie určenie potrebných – „*Jury Selection Problem*“ [35].

¹⁵ <https://twitter.com>

¹⁶ https://twitter.com/who_to_follow/suggestions

Každý expert prináša so svojou odbornosťou aj určitú mieru chybovosti – čím väčšia je jeho odbornosť, tým menšia by mala byť pravdepodobnosť jeho chybovosti. Pri výbere expertov na získanie odpovede je teda potrebné uvažovať aj pravdepodobnosť chybovosti vybranej skupiny – „*Jury Error Rate*“. Na základe štandardných pravidiel pri volení v skupine [36], aj skupina expertov sa mýli vtedy, keď sa mýli aspoň polovica zo zvolených expertov na riešenie vybraného problému. Pravdepodobnosť chybovosti celej skupiny sa vypočíta súčinom pravdepodobností chybovosti jednotlivcov. Príklad pravdepodobnosti chybovosti po zaradení jednotlivcov do skupín je zobrazený v tab. 3.1.

tab. 3.1 Pravdepodobnosť chybovosti skupiny pri rôznej kombinácii expertov

vybraní experti	individuálna chybovosť	chybovosť skupiny
C	0.2	0.2
A	0.1	0.1
C,D,E	0.2,0.2,0.3	0.174
A,B,C	0.1,0.2,0.2	0.072
A,B,C,D,E	0.1,0.2,0.2,0.3,0.3	0.0703
A,B,C,D,E,F,G	0.1,0.2,0.2,0.3,0.3,0.4,0.4	0.0805
A,B,C,F,G	0.1,0.2,0.2,0.4,0.4	0.104

Na základe tab. 3.1 je vidieť, že vytváranie expertných skupín znižuje pravdepodobnosť chybovosti pri určení odpovede, avšak nie nutne vždy. Rovnako nie je nižšia pravdepodobnosť chybovosti zaručená väčším počtom expertov v skupine. Keďže výpočty môžu byť pre väčšie množstvo kombinácií zdĺhavé, je nutné vopred stanoviť maximálne množstvo jednotlivcov v skupine a maximálnu prípustnú chybovosť skupiny. Z tabuľky tiež vyplýva, že pri rozhodovaní pomocou hlasovania musí byť počet expertov v skupine nepárny.

Na výpočet pravdepodobnosti chybovosti jednotlivca je v [35] navrhnutý algoritmus postavený na grafe $G(N,E)$ využívajúci interakcie medzi používateľmi sociálnej siete Twitter postavený na algoritmoch PageRank a HITS opísaných v časti 3.2.1.

3.3. Diskusia

Opísali sme prístupy na hodnotenie expertízy používateľov. Prezentovali sme prístupy za pomoci využitia davu a tiež využitie skupiny expertov na riešenie problému. Ako ukázali výsledky niekoľkých experimentov, odhaľovanie expertov je stále otvorený problém s relatívne nízkou úspešnosťou. Tento problém riešia hlavne prístupy v oblasti dolovania informácií (*IR*). Keďže automatické prístupy majú nízku úspešnosť, vidíme potenciál na využitie davu a hier s účelom, ktoré práve v takýchto prípadoch dokážu mať omnoho vyššiu úspešnosť.

Prezentovali sme aj utilizáciu expertov v hre s účelom, kedy odmena pre hráčov korelovala na základe obtiažnosti úlohy vypočítanej serverom. Hráči teda výberom úloh s vyššou náročnosťou a ich úspešným vyriešením implicitne prezentovali svoju expertízu. Tento fakt je pre nás dôležitý pri odmeňovaní hráčov v nami navrhnutej hre. V ďalšej časti práce sa zaoberáme návrhom hry s účelom, ktorá opísané prístupy využíva.

4. Hra s účelom s rozpoznávaním a využitím autorít

V rámci analýzy sme uviedli problém nízkeho využitia autorít pri získavaní dát davom – všeobecne sa využitie expertov separuje od davu z dôvodu kvality dát. Keby sa nám podarilo rozpoznať a využiť expertov aj v rámci davu, vedeli by sme skvalitniť a zrýchliť aj tento spôsob získavania informácií.

Našou primárnou doménou využitia davu sú hry s účelom. Hra dokáže poslúžiť davu ako motivátor a nie je nutné hráčov motivovať dodatočne. V našej minulej hre City Lights sme sa zaoberali overovaním hudobných metadát. Pri vyhodnocovaní sme si uvedomili, že množstvo správnych anotácií dav označil ako nesprávne (resp. neoznačil nijak) aj napriek tomu, že expertní používatelia tieto anotácie rozoznali. Tento expertný potenciál, ktorý sa v dave nachádza, by dokázal získavanie informácií pomocou davu urýchliť a spresniť – zamerali sme sa preto na rozpoznávanie a využitie expertov v rámci vlastnej hry s účelom.

Vytvorili sme vlastné internetové rádio, do ktorého sme implementovali jednoduché hry na overenie znalostí hráčov. Hry vždy súvisia so skladbou, ktorá práve v rádiu hrá. Rádio je postavené na 4 hudobných štýloch (doménach), ktoré sme navrhli tak, aby sa čo najmenej prekrývali. Tieto hudobné štýly následne využívame na získanie a overenie domény záujmu hráča a jeho expertízy v rámci domény.

Minihry pozostávajú z jednoduchých úloh, ktoré majú overiť vedomosti a hudobné znalosti hráča. Štandardne trvá vyriešenie úlohy niekoľko sekúnd. V rámci skladby môže hráč vyriešiť obmedzený počet úloh. Vybrané úlohy ale môže riešiť niekoľko krát, vždy s náhodne pripraveným vstupom. Správnosť všetkých odpovedí je overiteľná v reálnom čase vďaka externému datasetu.

Hráči súperia proti sebe navzájom na rôznych úrovniach, primárne na úrovni jednej skladby, ktorá predstavuje kolo. Cieľom hráčov je počas kola získať čo najviac bodov správnym riešením úloh (odpovedaním na otázky), za čo obdržia tzv. žetón, ktorý môžu použiť na zmenu playlistu rádia.

Na základe akcií hráčov a ich správnych odpovedí navrhujeme metódu na spätné rozpoznanie expertov medzi hráčmi. Metóda je postavená na algoritme HITS a rozpoznáva expertízu hráčov v konkrétnej doméne. Navrhnutú metódu neskôr overíme a vyhodnotíme v časti 5.

4.1. WoodstockFM

Nami navrhnuté internetové rádio je postavené na vlastnom playliste – nepreberáme playlist ani skladby iných rádií. Playlist rádia je rozdelený do štyroch hudobných žánrov - uvedení interpreti slúžia ako kostra playlistu rádia a môžu byť automatickými odporúčačmi a skladbami zadanými používateľmi doplnení o ďalších interpretov:

- *Populárna americká hudba (hip-hop, popmusic):* Beyoncé, Rihanna, Lily Allen, Eminem, Jay-Z, Lady Gaga, Miley Cyrus, Katy Perry
- *Rocková a metalová hudba:* Led Zeppelin, Deep Purple, Black Sabbath, Kiss, The Who, Jimi Hendrix, Queen, David Bowie, The Doors, Metallica, Ozzy Osbourne, Pink Floyd, Pearl Jam Misfits, Sex Pistols, Nirvana

- *Alternatívna hudba a elektronika*: Prodigy, Chemical Brothers, Moby, Fatboy Slim, Daft Punk, Gorillaz, Arctic Monkeys, Kasabian, Coldplay, Florence and the Machines, Radiohead
- *Jazzová hudba, soundtracková hudba a staré skladby*: Louis Armstrong, Miles Davis, Duke Ellington, Django Reinhardt, Ella Fitzgerald, Hans Zimmer

V rámci všetkých hudobných štýlov sme na základe konzultácie s expertmi určili základnú vzorku skladieb, ktorá v rádiu hráva. Vďaka tomuto rozdeleniu sme neskôr schopní určovať úroveň expertízy hráča pre konkrétnu doménu, čo má vplyv na presnosť určenej expertízy. Vzorka skladieb taktiež slúži na natréňovanie Bayesovského klasifikátora, ktorý priradzuje ďalších interpretov do jednej z štyroch definovaných domén.

4.1.1. Herné mechanizmy

4.1.1.1. Vedomostné úlohy

Hráči môžu počas hrania skladby odpovedať na vedomostné otázky súvisiace s danou skladbou. Otázky sú rôznej obtiažnosti a týkajú sa samotnej skladby, interpreta a albumu, z ktorého skladba je. Po zodpovedaní aspoň 3 primárnych úloh sa hráčovi sprístupnia sekundárne úlohy, ktoré môže riešiť aj viac krát v rámci jednej skladby. Po skončení skladby sa hra zresetuje a hráč začína ďalšie kolo od nuly.

Hráč odpoveďami na zadané otázky nepriamo prejavuje záujem o doménu, z ktorej skladba je. Rozpoznanie domény záujmu hráča je prvým krokom pri vyhľadávaní expertov - keď poznáme doménu hráča, môžeme začať študovať úroveň jeho expertízy v danej doméne. Keďže sú otázky vedomostné, jedným zo spôsobov ako k problému pristupovať by bolo túto úlohu vyriešiť dichotomicky a pracovať s počtom správnych odpovedí hráča. Ďalším spôsobom riešenia je brať do úvahy obtiažnosť otázky a na jej základe prikladať hráčovým odpoveďiam dôležitosť. Metóda prezentovaná neskôr v tejto časti využíva na spresnenie výpočtu aj obtiažnosť úlohy.

Použité otázky môžeme rozdeliť do týchto kategórií

- *názvy*: albumu, skladby, interpreta
- *roky*: vznik a zánik skupiny, vydanie albumu
- *miesta*: mesto a krajina pôvodu interpreta

Okrem otázok s možnosťami sú hráčovi na základe jeho čiastočnej odpovede ponúknuté možné správne riešenia (tzv. „*fulltext autocomplete*“) z dôvodu eliminovania chýb preklepmi a mierne odlišných názvov. Pre prípady kedy sa používateľ rozhodne nevyužívať ponúknuté možné odpovede a zadávať odpovede manuálne, zaviedli sme aj vyhodnocovanie na základe Levenshteinovej vzdialenosti, ktorá slúži ako ochrana proti preklepom. V prípade číselnej odpovede sme zaviedli pre každý typ úlohy toleranciu, ktorá hráčovi odpoveď uzná, ale s nižším bodovým ziskom.

4.1.1.2. Úlohy na overenie hudobnej zručnosti

Úlohy na overenie hudobnej zručnosti sú primárne určené na zistenie globálnej úrovne hudobnej expertízy hráča – v tomto type úloh sa čiastočne vynecháva zisťovanie domény záujmu z dôvodu ich generickosti (aj keď stále záleží od konkrétnej hry). Úlohy napovedia o hráčovom hudobnom vzdelaní prípadne hudobnom sluchu.

Použitú úlohu môžeme rozdeliť do týchto kategórií:

- *rytmické*: určenie tempa skladby, zhoda s ostatnými hráčmi
- *sluchové*: určenie tóniny skladby
- *vedomostné*: určenie správnych anotácií pre skladbu, určenie podobných skladieb a interpretov; práve tento typ úloh môže byť využitý na tvorbu a validáciu metadát k hudbe pri prípadnom overení navrhnutých prístupov

Úlohy na overenie hudobnej zručnosti sú na rozdiel od vedomostných úloh viac zábavným prvkom hry a informácie získané ich hraním slúžia len na doplnenie už získaných dát z vedomostných minihier.

4.1.2. Realizácia

4.1.2.1. Externé metadáta

Hra je postavená na overovaní hráčových odpovedí voči faktom. Tieto fakty je potrebné získať a predspracovať z dôveryhodných zdrojov. V tejto časti prezentujeme zdroje, ktoré sme použili a druh dát, ktorý sme z nich extrahovali.

4.1.2.2. Echonest

Služba existuje ako agregátor veľkej časti ostatných služieb poskytujúcich hudobné metadáta (7Digital, Spotify, MusicBrainz a iné). Dáta ponúka v predspracovanej forme, ale zároveň ponúka odkaz na originálny zdroj, kde je možné nájsť dodatočné metadáta. Aplikácia primárne pracuje s dátami:

- *Umelec / skladateľ*: krajina pôvodu, rok vzniku a zániku hudobnej aktivity, miesta koncertov, popularita, recenzie a biografie
- *Skladba*: názov, dĺžka, tempo, tónina, text, špecifické metadáta poskytnuté službou (popularita, tanečnosť, živosť, akustickosť a pod.)
- *Album*: názov, rok vydania, vydavateľ, počet skladieb, dĺžka albumu

Spomínané koeficienty (tanečnosť, popularita) vypočítané službou Echonest môžeme využiť pri nepriamom usudzovaní expertízy hráča. V prípade málo populárnej kapely je menej pravdepodobné, že hráč bude poznať odpovede na otázky. Avšak v prípade, že ich pozná, jeho úroveň expertízy v oblasti (a aj voči ostatným hráčom) môže byť omnoho vyššia ako pri všeobecne populárnych interpretoch – túto tému ďalej rozoberáme v časti 4.2.1.

V prípade nedostatku dát získavame dodatočné metadáta na základe poskytnutých ID záznamov priamo v externých službách (LastFM a 7Digital) a nie sme závislí výhradne od jednej služby.

4.1.2.3. Grooveshark

Hudbu rádia zabezpečuje spolupráca so serverom Grooveshark. Playlist rádia je obmedzený na skladby, ktoré pre región strednej a východnej Európy služba poskytuje. Zdrojové skladby sú z veľkej časti nahrávané samotnými používateľmi – z tohto dôvodu sa občas stane, že namiesto očakávanej skladby služba vráti síce správnu skladbu, ale prespíevanú iným interpretom, príp. karaoké verziu skladby.

Z licenčných dôvodov služba neposkytuje úplne všetky používateľmi nahrané skladby – chýba napr. skupina „The Beatles“, ktorá svoje albumy neposkytuje žiadnej on-line službe. Na tieto obmedzenia pri vyhodnocovaní experimentov prihliadame.

4.1.3. Herné prostredie

Aplikácia je rozdelená na dva základné pohľady: pohľad pasívneho poslucháča a pohľad hráča. Pohľad pasívneho poslucháča poskytuje čisté prostredie so základnými informáciami o prehrávanej skladbe. Pohľad hráča následne rozširuje toto prostredie o herné prvky.

4.1.3.1. Rádio (pohľad poslucháča)

Pasívny pohľad neponúka žiadne interaktívne prvky. Poslucháč má k dispozícii menu na prepínanie medzi pohľadmi a nastavenia profilu. V hlavnej časti obrazovky má k dispozícii iba fotografiu interpreta. Žiadne ďalšie informácie nie sú poskytnuté z dôvodu prevencie proti podvodníkom následne zneužívajúcim tieto informácie v hernom režime. Poslucháč má možnosť vyjadriť preferenciu nad hranou skladbou jednoduchým „páči sa mi“ / „nepáči sa mi“ označením. Súčasná verzia pohľadu poslucháča je zobrazená na obr. 4.1 a širšie opísaná v prílohe tejto práce.



obr. 4.1 WoodstockFM - pohľad poslucháča

4.1.3.2. Herný režim (pohľad hráča)

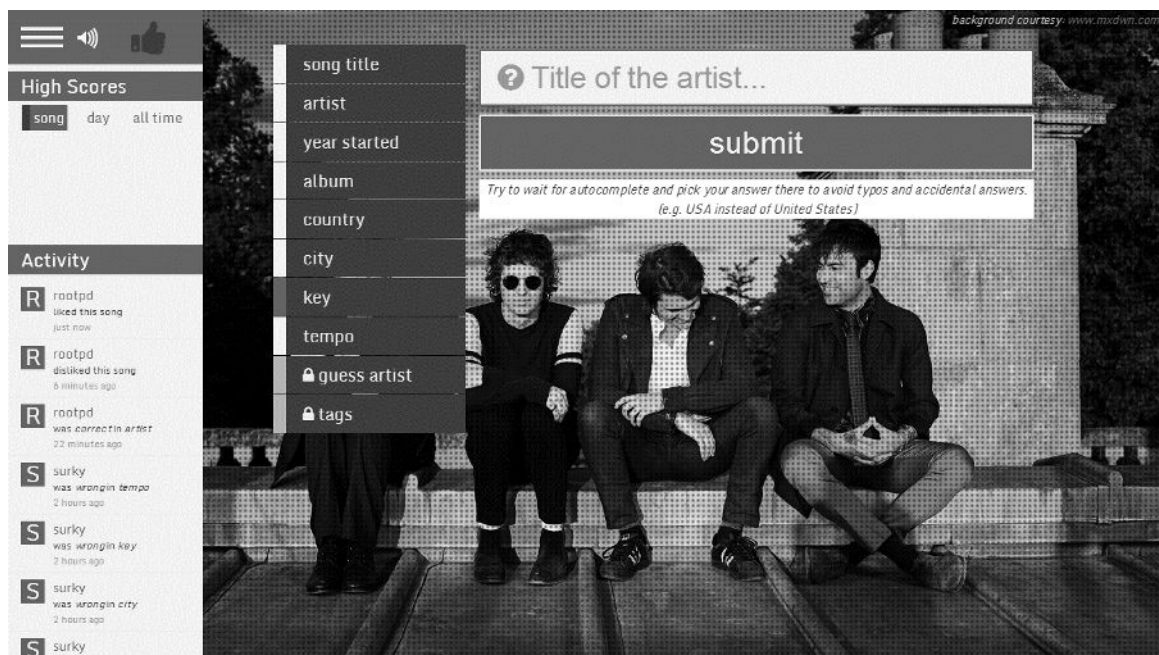
Režim je možné prepnúť voľbou v menu aplikácie. Poslucháč sa môže prepínať medzi herným režimom a rádiom bez obmedzenia. Hráčovi sa uprostred obrazovky zobrazí nástenka s informáciou o počte pripojených hráčov, jeho bodovom stave a dostupnými minihrami. Súčasná verzia je zobrazená na obr. 4.2 a širšie opísaná v prílohe tejto práce.

Na začiatku kola majú hráči na výber k dispozícii vybrané minihry, do ktorých sa môžu zapojiť – nie vždy sú totiž k dispozícii všetky metadáta, ktoré slúžia ako prerekvizity k minihram. Každý hráč môže riešiť všetky minihry – hra nefunguje na princípe „uchmatávania“ úloh z dôvodu získania čo najväčšieho objemu dát, ktoré môžeme vyhodnocovať. Minihry majú časové obmedzenie, v ktorom je nutné ich splniť. Tieto návrhové rozhodnutia do hry prinášajú:

- *Súťaživosť.* Hráči môžu porovnávať svoje odpovede v rámci konkrétnych minihier a v ľavom paneli vidieť, ktorí v poradí odpovedali, koľko hráčov odpovedalo správne a celkovú úspešnosť hráčov na danej skladbe. Všetky minihry sú hráčom dostupné naraz, avšak na poradí riešených úloh nezáleží.

- *Obmedzenie.* Nie všetky minihry sú dostupné od začiatku a aby sa hráčovi sprístupnili, je nutné aby odpovedal na definovaný počet úloh. Sprístupnenie týchto úloh výrazne zvyšuje jeho šancu na výhru, keďže je spravidla možné tieto úlohy riešiť v rámci skladby aj opakovane.

Rozhranie zobrazené na obr. 4.2 má za úlohu ešte viac podporiť vyššie spomenuté prvky. Hráči vždy vidia, ktorý z nich práve odpovedal na niektorú z úloh a aká bola jeho úspešnosť. Aby sme hráčov podporili riešiť čo najviac úloh, je bodovanie postavené súčte bodov za jednotlivé minihry, pričom náročnejšie minihry majú spravidla väčšiu bodovú odmenu. Po kliknutí na hráčovo meno v ľavom paneli sa zobrazia jeho detailné štatistiky a jeho úspešnosť od registrácie. Aktuálne skóre kola, dňa a sieň slávy je stále viditeľná a dostupná v ľavom paneli.

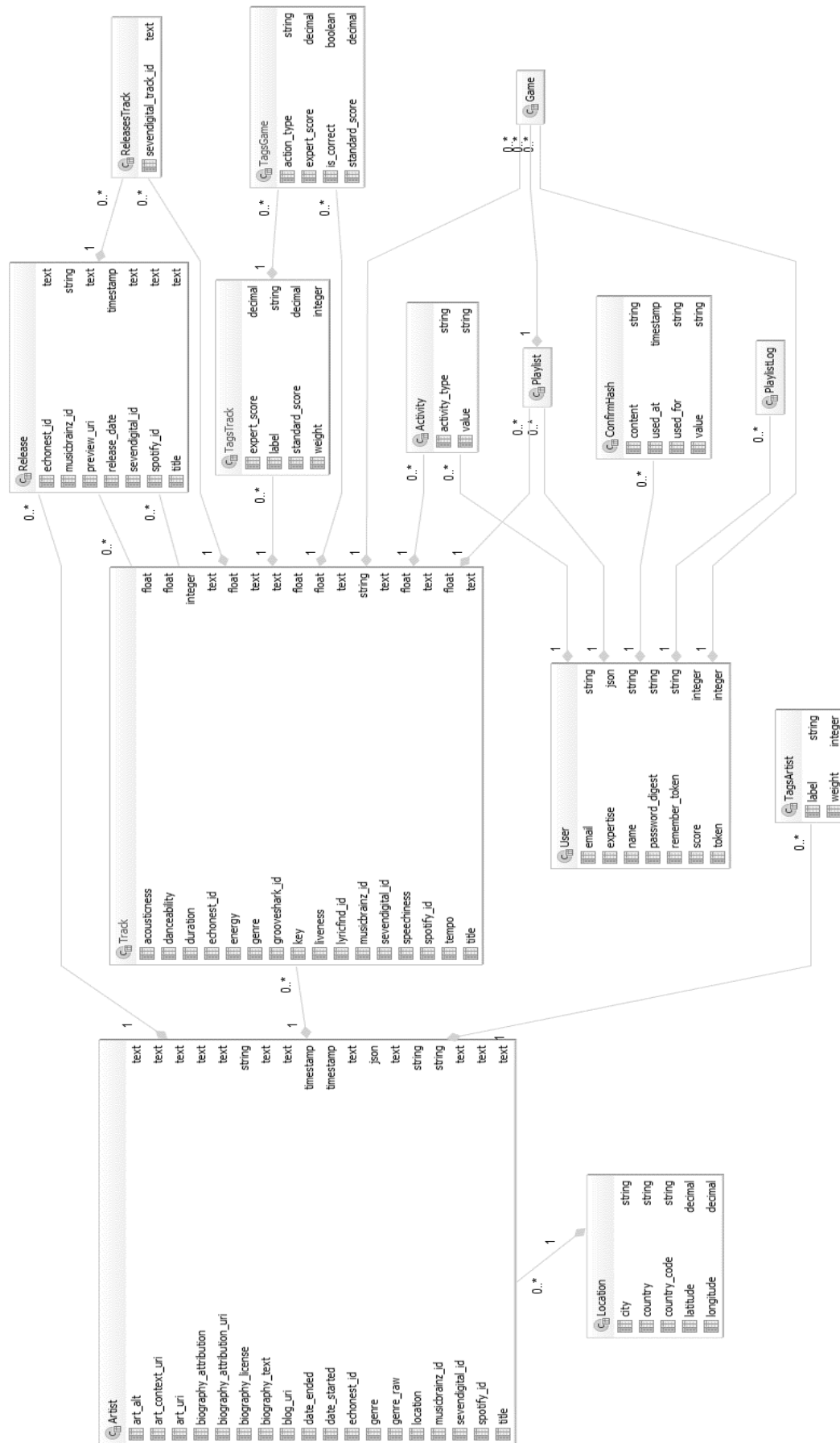


obr. 4.2 Návrh pohľadu hráča

4.1.3.3. Rebríčky a odmeňovanie (pohľad hráča)

Dôležitou súčasťou hry a zároveň objavovania expertov ostatnými hráčmi je prístup k rebríčkam a štatistikám. Štandardne majú hráči dostupný rebríček jednotlivých frakcií pre dané kolo. Po zobrazení detailu hráča majú ale prístupné podrobné štatistiky hráča vrátane anonymizovaných štatistík, koľko krát bol hráčovi udelený žolík.

Okrem číselných štatistík dostávajú hráči po dosiahnutí vopred určených limitov odznaky. Tieto slúžia ako (1) nečíselný indikátor možnej expertízy pre ostatných hráčov a (2) motivátor pre samotných hráčov hrať ďalej a dosiahnuť ďalšiu úroveň limitu. Hráči v súčasnosti vopred nevedia kedy a za čo dostanú odznak, ani aká veľká je odmena za jeho získanie - či už bodová alebo vo forme žetónov do playlistu.

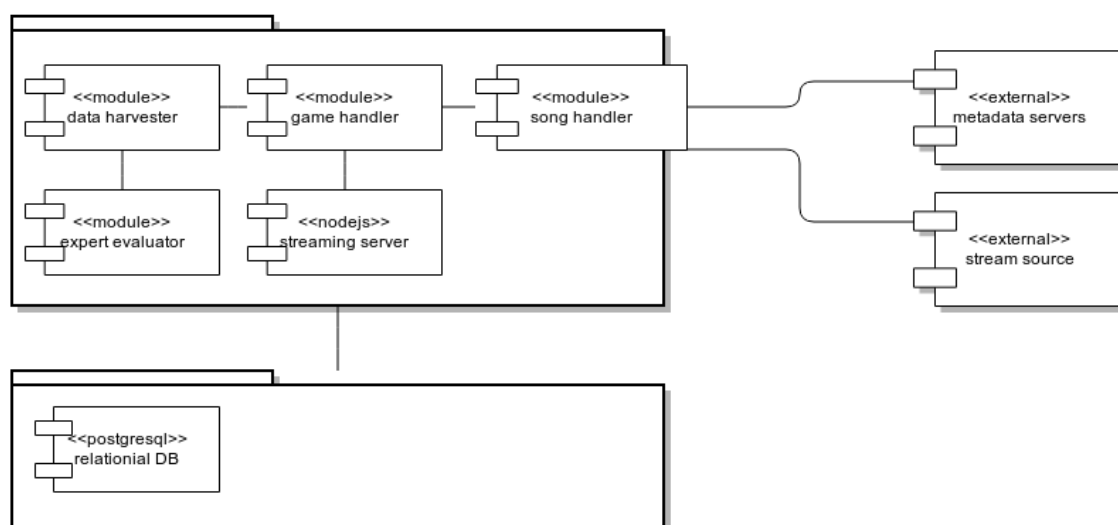


obr. 4.3 Diagram nejdůležitějších tříd systému

4.1.4. Architektúra

Aplikácia je primárne rozdelená na dve hlavné časti: (1) webovú aplikáciu pre používateľov, ktorá sa stará o zobrazovanie minihier, štatistík, nastavenia aplikácie a všetky neinteraktívne prvky a (2) serverovú aplikáciu, ktorá sa stará o servis samotného rádia a komunikáciu s používateľmi v reálnom čase. Aplikácie využívajú spoločnú databázu. Na obr. 4.3 je zobrazený diagram najdôležitejších častí systému.

Modul pre vyhodnocovanie expertízy je oddelený od herného modulu, keďže vyhodnocovanie expertízy priamo nekoreluje s výsledkami v hre. Taktiež nám umožňuje a posteriori vyhodnotenie zozbieraných dát. Do úvahy tak môžeme brať nie len aktuálne herné výsledky hráča, ale aj dáta o počúvanosti skladieb, ktoré môžu byť získané až po odohraní vyhodnocovanej hry. Prehľad architektúry projektu je zobrazený na obr. 4.4 a širší opis jednotlivých častí je podrobne opísaný v prílohe projektu.



obr. 4.4 Architektúra aplikácie woodstockFM

4.2. Objavovanie autorít

4.2.1. Podobnosť interpretov

Nami vytvorené 4 skupiny hudobných žánrov nie sú od seba striktne oddelené a interpreti sa viac či menej svojou tvorbou v hudobných štýloch prekrývajú – tento prekryv sme sa výberom hudobných žánrov snažili minimalizovať. Pri vyhodnocovaní dát sa teda nemôžeme spoliehať na kategorizáciu interpreta do jedného z menovaných žánrov. Z tohto dôvodu pre každého interpreta počítame *vzdialenosť od hudobného žánru* - tento výpočet pri vybranom interpretovi robíme pre všetky 4 nami definované skupiny.

Na získanie podobnosti sme sa rozhodli využiť už overené anotácie o interpretoch z externého datasetu. Jednotlivé hudobné žánre (skupiny žánrov) prezentované na začiatku kapitoly môžeme transformovať na charakteristickú skupinu anotácií, ktorá ich reprezentuje.

Tieto použijeme na natrénovanie Bayesovského klasifikátora, ktorý posluží na zaradovanie v budúcnosti pridaných interpretov do jednotlivých hudobných žánrov. Klasifikátor je počítaný

nasledovne (C – trieda, hudobný žáner, do ktorého klasifikujeme; p – pravdepodobnosť výskytu; F – klasifikátor, v našom prípade anotácie:

$$p(C|F_1, \dots, F_n) = \frac{p(C) * p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)}$$

Podobnosť primárne využívame pri skladbách od interpretov, ktorí neboli zafinovaní v rámci jednotlivých žánrov – buď ich do systému pridal automatický odporúčač, alebo jeden z používateľov výmenou za jeden žetón.

4.2.2. Dáta o používateľoch získané z aplikácie

Dáta získané z hry rozdelíme na tie získané minihier (objektívne, vieme overiť ich správnosť) a na tie získané implicitne a explicitne za pomoci davu (empirické). Keďže empirické rozhodovanie davu je postavené na faktoroch prezentovaných hrou, sú tieto dáta považované len ako rozšírenie metódy prezentovanej nad objektívne získanými dátami.

V rámci jednej skladby má každý hráč na výber rovnakú množinu minihier. Minihry sa líšia svojou obťažnosťou spôsobom sprístupnenia hráčovi. Z hráčových akcií zaznamenávame:

- Poradie a čas začatia minihry
- Dĺžku riešenia minihry
- Úspešnosť vyriešenia minihry – v prípade neúspešného dokončenia minihry aj dôvod (nesprávna odpoveď)

Vyhodnocovanie expertízy je postavené výhradne na správnosti odpovede, avšak odmeňovanie hráčov a sprístupňovanie minihier môže byť obmedzované na základe poradia, času začatia alebo dĺžky riešenia jednotlivých úloh.

4.2.3. Utilizácia dát a objavovanie autorít

Z hier a zo záznamov počúvanosti skladieb pre každého používateľa máme k dispozícii dve skupiny dát potrebné na odvodenie expertízy hráčov:

- Hudobné preferencie používateľov – ich domény záujmu.
- Hry, ktoré sa pokúsili riešiť a ich úspešnosť – ich odbornosť v rámci nami definovanej domény.

Výpočet expertízy používateľa sme sa rozhodli postaviť na HITS algoritme, ktorý mal zo spomenutých algoritmov v kapitole 3.2.1 najlepšiu úspešnosť. Algoritmus je ale postavený na grafe obsahujúcom len jeden typ vrcholov – napr. na URL pri vyhľadávateľoch prípadne na používateľoch v rámci podnikových expertných systémoch.

Jeho výpočet je iteratívny a konverguje k výslednej hodnote. Výstupom algoritmu sú dve hodnoty opisujúce objekt: (1) *authority* opisujúca expertízu a závisiaca od *hub* hodnôt objektov, na ktoré odkazuje a (2) *hub* opisujúca popularitu, ktorá závisí od autority objektov smerujúce na daný objekt.

$$auth = \sum_{i=1}^n hub(i)$$

$$hub = \sum_{i=1}^n auth(i)$$

Pri určovaní expertov v hrách s účelom máme dataset postavený na dvoch typoch dát: (1) na používateľoch, ktorých úroveň expertízy chceme ohodnotiť a (2) na minihrách (všeob. úlohách), ktoré slúžia ako prostriedok ohodnotenia. Aby sme mohli využiť HITS algoritmus, je potrebné nájsť prepojenie z jednotypového použitia na dvojtýpové. V rámci špecifikácie definujeme väzbu medzi hráčom a úlohou, ktorú rieši:

Definícia. Nech *používateľ* je objekt určovania expertízy a nech *úloha* je prostriedkom na určenie expertízy. Nech *väzba* je relácia medzi používateľom a úlohou. Nech väzba medzi používateľom a úlohou je vytvorená vtedy a len vtedy, keď sa používateľ zapojí do riešenia úlohy a úlohu vyrieši správne. Potom expertíza *auth* používateľa je určená na základe vytvorených väzieb nasledovne:

- 1) Hodnoty *authority* aj *hub* pre všetky objekty inicializujeme na 1.
- 2) Hodnoty *authority* používateľov U vypočítame podľa vzorca exp. 1.
- 3) Hodnoty *hub* úloh T vypočítame podľa vzorca exp. 2.
- 4) Znormalizujeme všetky hodnoty tak, že sa predelia maximálnou hodnotou daného typu v rámci objektu – *authority* oddelene od *hub* (už zahrnuté v zmienených výrazoch).
- 5) Pokračujeme v ďalšej iterácii od kroku 2.

$$auth = \frac{\sum_{i=1}^n \frac{1}{hub(i)}}{hub_{max}}; hub_{max} \in T: hub_{max} > h \forall h \in T$$

exp. 1 Iterácia výpočtu *authority* používateľa

$$hub = \frac{\sum_{i=1}^n auth(i)}{auth_{max}}; auth_{max} \in U: auth_{max} > a \forall a \in U$$

exp. 2 Iterácia výpočtu *hub* úlohy

Obe hodnoty, ktoré algoritmus pre každý objekt vracia využijeme, avšak každú z nich len v jednom type dát (zvyšné hodnoty ignorujeme a nie je nutné ich počítať, keďže nezasahujú do nami požadovaných hodnôt):

- *Authority.* Hodnota je počítaná iba pre hráčov. Jej hodnota sa rovná normalizovanému súčtu hodnôt *hub* všetkých úloh, ktoré vyriešili správne a reprezentuje úroveň expertízy používateľa v intervale $<0;1>$ - čím vyššia hodnota, tým vyššia obtiažnosť..
- *Hub.* Hodnota je počítaná iba pre úlohy. Jej hodnota sa rovná normalizovanému súčtu hodnôt *authority* používateľov, ktorí úlohu vyriešili správne a reprezentuje úroveň obtiažnosti na intervale $<0;1>$ - čím nižšia hodnota, tým vyššia obtiažnosť.

Pre výpočet expertízy je potrebné mať vopred určenú referenčnú vzorku interpretov pre dané hudobné žánre, z ktorých sa následne extrahujú verejne dostupné anotácie a natrénuje klasifikátor – v princípe stačí mať doménu zadefinovanú množinou dobre opisujúcich anotácií. Pri výpočte následne klasifikátor rozhoduje, do ktorej hudobnej domény skladba (interpret) patria. Samotný výpočet v prípade našej hry a internetového rádia prebieha v nasledovných krokoch:

- 1) Pre vybranú skladbu sa získajú všetky minulé odohrané skladby a im prislúchajúce hry. *Úloha* je v tomto prípade definovaná unikátnou dvojicou *skladba-tyt hry a používateľ* hráč, ktorý úlohu plnil. Takáto definícia umožňuje algoritmu určovať rôzne obtiažnosti rovnakého typu úloh na rôznych skladbách – teda uhádnutie názvu jednej skladby môže algoritmus ohodnotiť inou obtiažnosťou ako názvu inej skladby rovnakého hud. žánra. Pre všetky úspešne vyriešené hry sa vytvorí väzba používateľ – úloha.
- 2) Pomocou navrhutej metódy sa iterovaním vypočíta globálna hodnota *authority* a *hub* pre hudobný žánr. Pre náš scenár je dôležitá hodnota *hub* na objektoch typu *úloha*, ktorá predstavuje obtiažnosť jednotlivých úloh pre vybraný žánr. Objekty typu *používateľ* sa v tomto momente zahodia.

```

get_previous_games.each do |game|
  player = players[game['user_id'].to_i] || Player.new
  tasks[game['task_key'].to_sym] ||= Task.new
  player.set_correct_task tasks[game['task_key'].to_sym]
  players[game['user_id'].to_i] = player
end

n.times do
  iterate_authority tasks.values, players
  iterate_hub tasks.values, players
end

```

- 3) Pre vybraného používateľa sa získajú všetky minulé hry (*úlohy*), ktoré patria pod rovnaký hudobný žánr ako skladba, s ktorou sa pracuje a na ktoré odpovedal úspešne. Tieto *úlohy* tvoria podmnožinu *úloh* z kroku 1. Vytvorí sa väzby medzi používateľom a týmito *úlohami*. **Jedinou iteráciou hodnôt *authority*** metóda vráti *authority* používateľa na základe reálnej obtiažnosti úloh, ktoré vyriešil správne.

```

player = Player.new
get_previous_user_affected_games.each do |game|
  player.set_correct_task tasks[game['task_key'].to_sym]
end

# just once, no further iterations
iterate_authority tasks.values, [player]
return player.score

```

V prípade našej hry a internetového rádia počítame expertízu jednotlivo pre každú úlohu na skladbe – do úvahy teda berieme úlohy, ktoré používateľ správne vyriešil na konkrétnej skladbe (spadajúcej pod niektorý z definovaných hudobných žánrov). Úroveň expertízy E_d v doméne potom počítame ako vážený priemer expertíz vypočítaných na úlohách skladieb spadajúcich pod daný hudobný žánr (doménu).

$$E_d = \frac{\sum_{task} auth(task)}{task_count}$$

4.3. Diskusia

Expertíza používateľa v doméne je určovaná na základe schopnosti riešenia úloh, ktorých správnosť je overiteľná. Na výpočet tejto odbornosti používame upravenú verziu algoritmu HITS pre dvojtypový graf dát. Metóda prináša nasledujúce výhody:

- Nie je nutné manuálne určovať číselnú obtiažnosť úloh. Ich obtiažnosť je vyhodnotená samotným algoritmom na základe množstva interakcie, ktoré hráči vygenerujú. Čím menej hráči interagujú a čím menej odpovedajú správne, tým je úloha náročnejšia.
- Metóda nie je staticky viazaná na doménu. Keďže výpočty prebiehajú na nižšej úrovni (v našom prípade na úrovni interpreta), doména expertízy môže byť určená ľubovoľne. Jediné, čo je potrebné je vypočítať podobnosť objektov nižšej úrovne (interpretov) k vybranej doméne. Táto podobnosť je potom použitá ako váha pri výpočte priemernej expertízy.
- Metóda nie je obmedzená iba na hry. Úroveň expertízy je možné vypočítať na ľubovoľnom dataseťe postavenom na princípe riešenia úloh používateľom.

Spolu s metódou prichádza aj niekoľko obmedzení:

- Metóda je postavená na vytváraní väzieb medzi používateľmi a správne vyriešenými úlohami. Je nutné vopred poznať správne riešenie úlohy a v prípade, že používateľ toto riešenie pozná, je nutné vedieť správne jeho odpoveď/interakciu spracovať.
- Overiteľná správnosť láka pri hrách k podvádaniu. Tomu sa snažíme zamedziť časovými limitmi a neponúknutím interakcie medzi hráčmi (kvôli zdieľaniu informácií), avšak bez riadeného experimentu a sledovania hráča nevieme určiť, akým spôsobom sa hráč k správnej odpovedi dopracoval..
- Doménu, voči ktorej expertízu určujeme, je potrebné vopred zdefinovať. V rámci tejto definície je potrebné určiť aj kritérium podobnosti, na základe ktorého sa počíta podobnosť parciálnych výpočtov expertíz. V našom prípade je potrebné určiť základných interpretov domény a získať overené anotácie, voči ktorým porovnáваме podobnosť interpretov s doménou.

Napriek spomenutým nevýhodám a obmedzeniam je metóda dôležitým spojovníkom medzi oblasťou objavovania autorít a hrami s účelom resp. s rozpoznávaním autorít na základe riešenia úloh všeobecne.

5. Overenie navrhnutých metód

Prostredie. Aplikácia bola nasadená na dobu neurčitú a správanie používateľov bolo monitorované len na základe záznamov získaných počas ich hrania. Žiadne priamo kontrolované používanie aplikácie primárne neplánujeme, keďže chceme zabezpečiť získanie čo najväčšieho množstva dát a zapojenia čo najviac používateľov v širokom časovom rozpätí.

5.1. Overenie vedomostí testom

Hypotéza. Navrhnutou metódou na objavovanie expertov v hrách s účelom postavenou na riešení úloh dokážeme rozpoznávať expertov v dave lepšie, ako súčasné prístupy.

Proces. Vybranej vzorke 6 účastníkov sme odoslali 20 náhodne vybraných hudobných ukážok a dotazník – ukážky boli anonymizované a nebolo možné zistiť z názvu súboru alebo ID3 anotácie o akú skladbu sa jedná. Jeden z účastníkov bol nami vopred označený ako expert na základe jeho súčasnej špecializácie. Účastníci dostali vedomostné otázky zo všetkých 4 hudobných domén, ktoré hra ponúka. Okrem toho mali ohodnotiť svoje hudobné znalosti a zadať svoje hudobné preferencie (množstvom nie obmedzený zoznam obľúbených interpretov). V nasledujúcej tabuľke sú vymenované použité hudobné ukážky:

Robbie Williams - Go Gentle Rihanna - Diamonds Jay Z- New York(feat Alicia Keys) Lady Gaga - Paparazzi Lyrics Katy Perry - Firework	The Who - My Generation Queen - The Show Must Go On Nirvana - Come As You Are Pink Floyd - Wish you were here David Bowie - Life On Mars
Radiohead - Creep Daft Punk - Around The World Kasabian - Days Are Forgotten The Prodigy - Breathe Florence + The Machine - Shake It Out	Gladiator Soundtrack - Now We Are Free Louis Armstrong - When You're Smiling Duke Ellington - Take the A Train Ella Fitzgerald - Summertime Charleston - Django Reinhardt

Pre každú z týchto skladieb mali účastníci vyplniť nasledujúce údaje – nemohli pri tom používať internet ani externú pomoc, nemali na výber zo žiadnych možností. V rámci tabuľky uvádzame aj výsledné percentuálne pokúsenie sa odpovedať na otázku (bez ohľadu na správnosť odpovede).

Tabuľka 1 Miera interakcie s jednotlivými typmi úloh v dotazníku.

Názov skladby	60,8%
Interpret	74,2%
Album	5,83%
Tónina	2,50%
Rok vydania	55,00%
Rok začiatku vystupovania interpreta	61,67%
Odkiaľ interpret pochádza	75,83%
3 skladby od tohto interpreta	50,83%

3 podobní interpreti	39,17%
Hudobný žáner	76,67%

Po prvotnom vyhodnotení interakcie s otázkami a správnosti odpovedí sme sa z experimentu rozhodli vylúčiť takmer polovicu údajov. Ich použitie by deformovalo výsledné expertné skóre účastníkov, ktoré sme nechceli normalizovať z dôvodu ponechania informácie miery nevedomosti účastníkov. Jednalo sa konkrétne o:

- Album: nízka interakcia, čiastočná úspešnosť
- Tónina: nízka interakcia, nulová úspešnosť
- 3 skladby od tohto interpreta: priemerná interakcia, veľmi nízka úspešnosť
- 3 podobní interpreti: podpriemerná interakcia, veľmi nízka úspešnosť
- Hudobný žáner: vysoká interakcia, avšak nízka úspešnosť (nevhodný typ otázky na dopĺňanie)

Výsledky. Údaje, ktoré neboli vylúčené a boli vyhodnotené ako správne boli vložené na vstup algoritmu na počítanie expertného skóre. Pre každý z hudobných žánrov si účastníci viedli nasledovne:

Tabuľka 2 Expertné skóre účastníkov pre jednotlivé hudobné žánre.

Kategória	#1 (expert)	#2	#3	#4	#5	#6	priemer
pop	0,5132	0,2812	0,2852	0,3217	0,3073	0,3034	0,3353
rock	0,4477	0,2155	0,3483	0,3328	0,2778	0,3835	0,3342
alternative	0,5278	0,1553	0,2798	0,3478	0,0225	0,2451	0,2631
old / soundtrack	0,2441	0,0497	0,1838	0,0000	0,0256	0,1307	0,1056
	0,4332	0,1754	0,2743	0,2506	0,1583	0,2657	

Z výsledkov je zrejmé, že ani nami označený expert nebol schopný správne odpovedať na viac ako polovicu otázok (pozn.: skóre nereprezentuje podiel správnych a všetkých otázok, tento pomer je ale na jeho základe možné odhadnúť). Napriek tomu je jasne oddeliteľný od ostatných účastníkov – hudobných laikov.

Aby sme overili správnosť výstupu, dostali vybraní účastníci validačnú úlohu. Účastníci si mali vypočítať ukážku (v tomto prípade vedeli, o akú skladbu ide) a určiť správnosť 3 anotácií k danej skladbe. Tieto anotácie sme prebrali z LastFM a mali ich overené aj 3 expertmi v rámci nášho zlatého štandardu z minulých experimentov. Anotácia bola určená ako správna ak sa aspoň 2 experti zhodli na jej správnosti. V rámci vyhodnotenia sme brali do úvahy percentuálnu zhodu s expertmi pri validácii jednotlivých anotácií skladby (*FP* – „false positives“, *FN* – „false negatives“, *FPx* – „false positives“ iba voči anotáciám, kde panovala úplná zhoda u expertov, *FNx* – „false negatives“ iba voči anotáciám, kde panovala úplná zhoda expertov. Výsledky overovacej úlohy sú premietnuté v nasledujúcej tabuľke:

Tabuľka 3 Skóre validačnej úlohy jednotlivých účastníkov na určenie správnosti expertného skóre.

alternative	skóre	FP	FN	FPx	FNx	autorita
#4	0,83	0,17	0,17	0,11	0,22	0,35
#5	0,70	0,25	0,33	0,11	0,22	0,02
pop	skóre	FP	FN	FPx	FNx	autorita

#2	0,57	0,47	0,36	0,17	0,67	0,28
#5	0,67	0,53	0,09	0,50	0,33	0,31
#4	0,73	0,47	0,00	0,17	0,00	0,32
rock	skóre	FP	FN	FPx	FNx	autorita
#5	0,67	0,31	0,41	0,00	0,33	0,28
#4	0,80	0,38	0,12	0,00	0,00	0,33

Vyhodnotenie. Porovnanie vypočítaného expertného skóre a validačnej úlohy prinieslo niekoľko zaujímavých zistení:

- *Menšie rozdiely medzi účastníkmi vo validačnej úlohe.* Aj napriek tomu, že expertné skóre predurčovalo účastníkov na väčšie rozdiely, dichotomické určovanie správnosti anotácií dopadlo u hráčov percentuálne veľmi podobne. Výsledky sa začali líšiť až pri sledovaní ďalších faktorov, ako bolo porovnanie iba voči anotáciám, kde u expertov panovala úplná zhoda. Účastníci s vyšším expertným skóre mali na týchto úlohách lepšiu úspešnosť (menšiu chybovosť).
- *Potreba rozlišovať typ úlohy.* Dotazník a teda aj expertné skóre boli generované na základe priamych vstupov účastníkov, bez možnosti voľby odpovede. Tento fakt výrazne vplýval aj na obtiažnosť úloh a nutnosť ignorovať niektoré typy úloh. Naopak vo validačnej fáze mali účastníci možnosť výberu – a to dichotomicky. Toto výrazne zjednodušilo obtiažnosť úlohy a objektívne zmazalo rozdiely medzi expertným skóre hráčov. Tieto rozdiely sa ale vo výsledku prejavili v type chýb, ktoré účastníci pri validácii vygenerovali. Potreba rozlišovať úlohy sa výrazne prejavila pre hud. žáner „*alternative*“, kde účastník s takmer nulovým skóre konkuroval inému účastníkovi.
- *Miera interakcie s jednotlivými typmi úloh.* V Tabuľka 1 sme zobrazili mieru interakcie hráčov s jednotlivými úlohami. Táto miera nám dáva do budúcnosti obraz, s akým typom úloh môžeme pracovať a pomôže nám správne nastaviť obtiažnosť celej hry.

Samotný experiment považujeme za čiastočne úspešný, keďže nám poskytol cenné údaje, na ktorých môžeme postaviť ďalšie experimenty a ktoré nám pomôžu vykonať v rádiu/hernom module také zmeny, ktoré pomôžu spresniť v budúcnosti vygenerované expertné skóre.

5.2. Spätné overenie úrovne expertízy predikciou výsledkov nasledujúcich hier

Hypotéza 1. Navrhnutou metódou na objavovanie expertov postavenou na riešení úloh dokážeme rozpoznávať expertov v dave lepšie, ako súčasné prístupy.

Proces. Počas dvojtyždňového nekontrolovaného experimentu sme zbierali dáta z odohraných hier v našom hudobnom rádiu. Hráči neboli priamo inštruovaní riešiť hudobné úlohy, avšak vedeli o možnom vplyve na playlist rádia v prípade, že tieto hry budú popri počúvaní hrať.

Po skončení skladby aplikácia skontrolovala, ktorí hráči odpovedali aspoň na jednu otázku (bez ohľadu na úspešnosť odpovede) a pre týchto hráčov vypočítala na základe ich dovtedajších výsledkov úroveň ich expertízy v doméne odohranej skladby – do tohto výpočtu neboli zaradené výsledky práve odohranej skladby; aplikácia počítala úroveň expertízy hráčov pred začatím

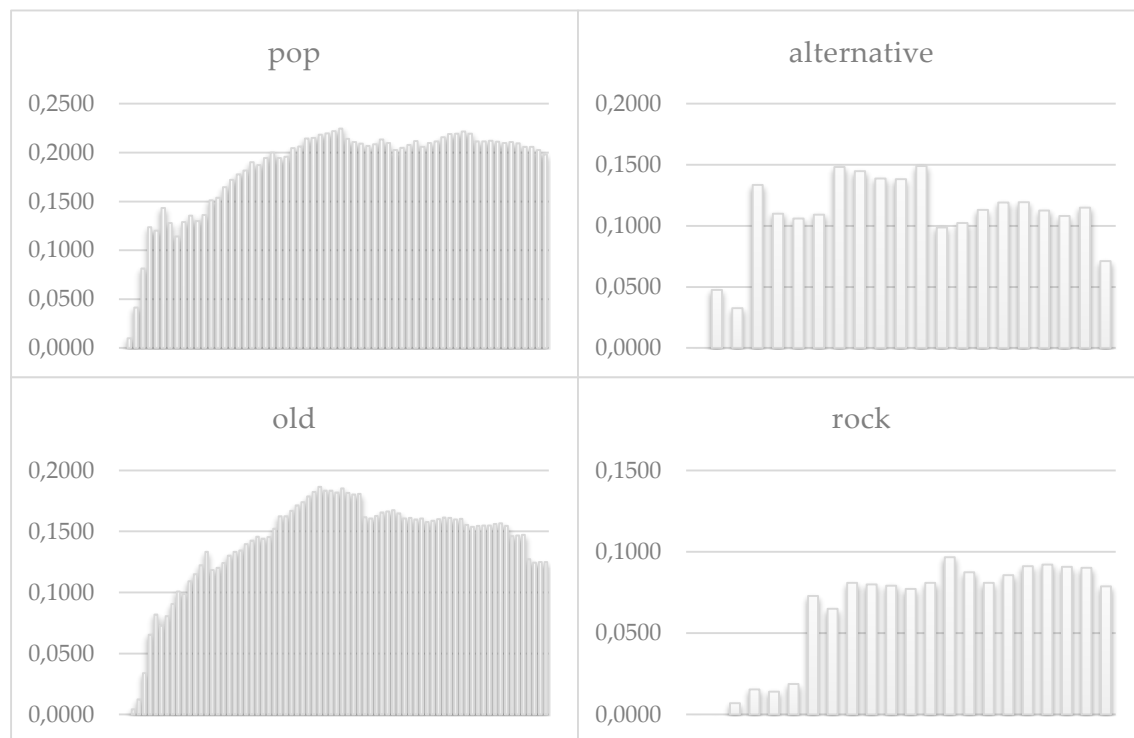
skladby. V prípade, že pre konkrétneho hráča neboli v tom čase dostupné minulé dáta (odohral svoju prvú hru), úroveň expertízy sa pre neho nepočítala.

Po vypočítaní úrovni expertízy pre všetkých hráčov, ktorí odohrali na skladbe aspoň jednu minihru sme porovnávali skutočné výsledky hráčov na skladbe (na základe ich nahratého skóre) a predikované výsledky hráčov (na základe výšky úrovne ich expertízy). Porovnávali sme celkové poradie v rebríčkoch a taktiež jednotlivé dvojice hráčov navzájom.

Výsledky. Do experimentu sa zapojilo 104 registrovaných používateľov, z ktorých 72 odpovedalo aspoň na jednu otázku v hernom režime. 32 hráčov odpovedalo na viac ako 30 otázok. Spolu hráči odpovedali na viac ako 8700 otázok a v 590 prípadoch explicitne označili, či sa im skladba páči alebo nie (372 pozitívnych proti 218 negatívnym). Počas experimentu odznelo v rádiu viac ako 4078 skladieb, z ktorých takmer 500 bolo pridaných samotnými používateľmi.

5.2.1. Výpočet expertízy

Sledovali sme zmenu úrovne expertízy používateľov v čase – chceli sme zistiť, koľko skladieb je potrebných aby sa expertíza ustálila. Na obr. 5.1 je zobrazený vývoj expertízy najaktívnejšieho hráča celého experimentu v rôznych doménach.



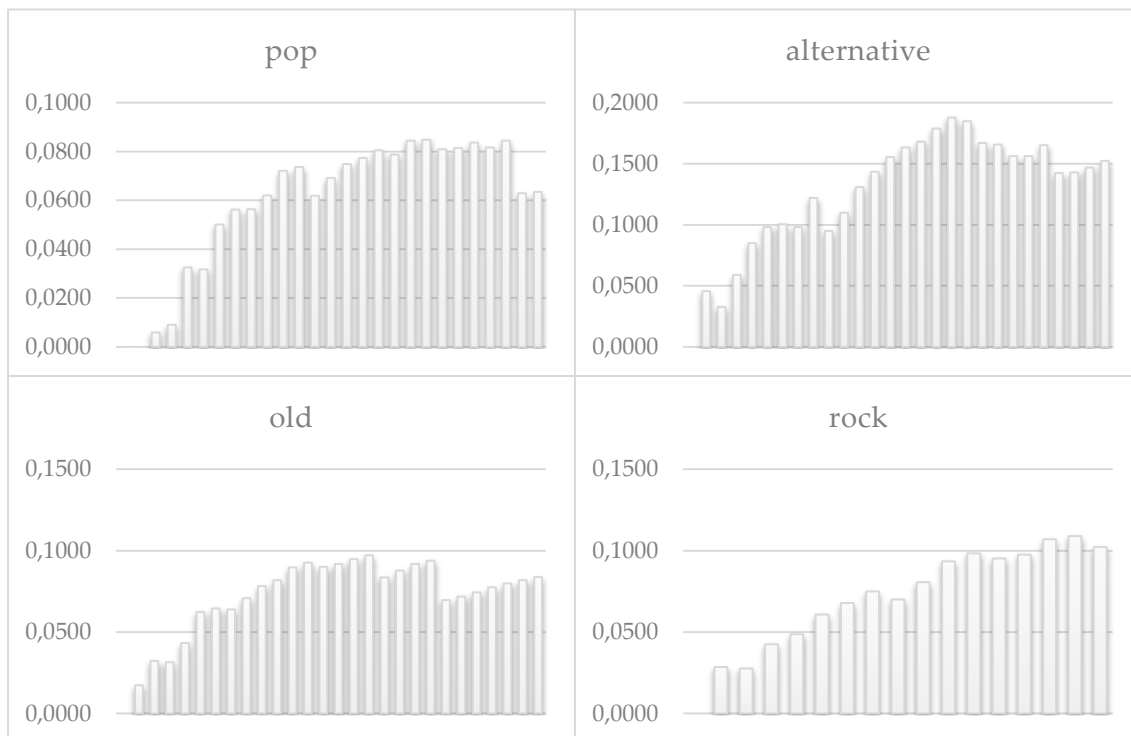
obr. 5.1 Vývoj expertízy hráča v 4 rôznych hudobných doménach

V prvých dvoch prípadoch je zreteľné, že úroveň expertízy rastie rýchlo, postupne sa ustáľuje a blíži sa očakávanej hodnote. Z analýzy dát usudzujeme, že toto ustálenie nastáva približne po 10 skladbách a potom sa mení minimálne. Priemerná zmena úrovne expertízy zaznamenaná po 10. skladbe a posledným získaným záznamom jednotlivých hráčov bola 0,04. Mierna zmena hodnôt aj po ustálení bola spôsobená zmenou obtiažnosti jednotlivých úloh v čase.

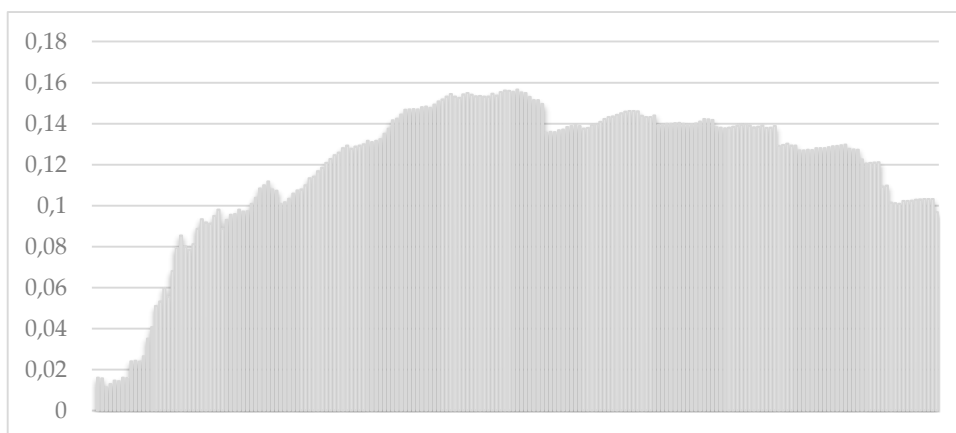
Na obr. 5.2 sú grafy vývoja úrovne expertízy pre najaktívnejších hráčov každého hudobného žánru (v prípade, že to bol rovnaký hráč ako na , boli použité údaje druhého najaktívnejšieho hráča). Dostatok dát dáva lepšiu predstavu o rôznom vývoji expertízy v rôznych doménach. V kombinácii

s predchádzajúcim setom grafov je vidieť, že ustálenie expertízy pre hudobný žáner „old“ nastáva omnoho neskôr (v zmysle počtu odohratých skladieb) ako pri hudobnom žánri „alternative“.

obr. 5.2 Vývoj expertízy najaktívnejších hráčov v konkrétnych hudobných žánroch

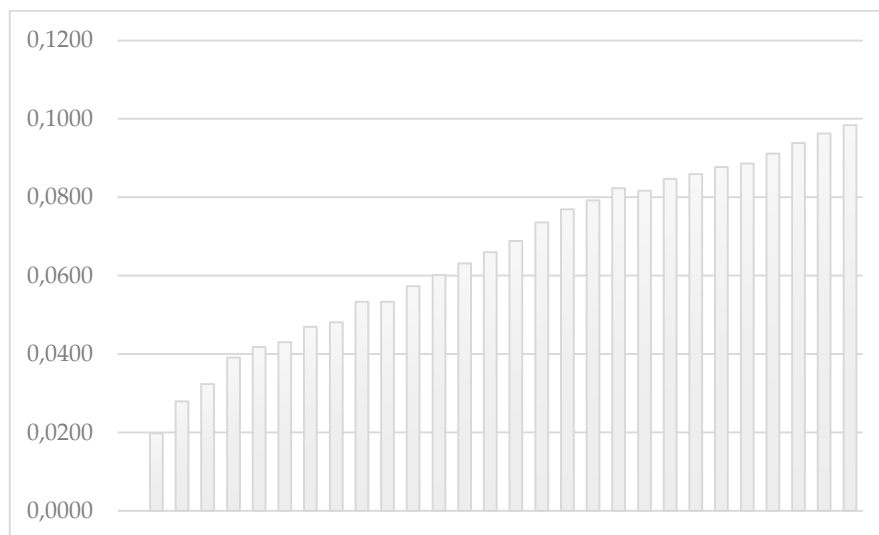


Aby sme poukázali na dôležitosť výpočtu expertízy v konkrétnej doméne, na obr. 5.3 je zobrazený vývoj expertízy najaktívnejšieho hráča bez ohľadu na doménu. V porovnaní vývojom expertízy totožného hráča podľa domény na obr. 5.1 je vidieť, že zavedenie domén výrazne mení interpretáciu a presnosť výsledkov – v tomto prípade sa úplne stratil rozdiel úrovne expertízy medzi žánrom *rock* a ostatnými žánrami. Predpokladáme, že výpočet expertízy na nižšej úrovni ako sú hudobné žánre by rovnakým spôsobom zlepšil úspešnosť určenia expertízy. Z dôvodu malého množstva dát na takúto vyhodnotenie ale nie je možné potvrdiť resp. vyvrátiť túto hypotézu. Porovnanie úspešnosti metódy počítajúcej úroveň expertízy na základe hudobných žánrov a globálne prezentujeme v časti 5.2.2.



obr. 5.3 Vývoj expertízy v čase najaktívnejšieho hráča bez ohľadu na hudobný žáner

Vďaka sledovaniu expertízy v čase bolo možné sledovať neštandardný vývoj expertízy niektorých hráčov. Na obr. 5.4 je zobrazený vývoj hráča, ktorý priznal výrazné používanie vyhľadávačov a nedostatkov hry, ktoré mu odpovedanie na faktoidné otázky uľahčovali. Oproti štandardnému správaniu (rýchlejši rast a ustálenie s miernou kulmináciou) vývoj jeho expertízy lineárne rástol. Keďže sme v tomto prípade mali dostatočné množstvo dát, môžeme prehlásiť že sledovaním vývoja expertízy v čase je naša metóda schopná odhaľovať podozrivé správanie hráčov.



obr. 5.4 Vývoj expertízy hráča - podvodníka - v čase

Metóda nám zároveň ako vedľajší produkt vygenerovala *hub* skóre pre jednotlivé úlohy – ich obtiažnosť. V tab. 5.1 sú uvedené obtiažnosti riešených úloh pre jednotlivé hudobné žánre a obtiažnosti úloh globálne (bez ohľadu na hud. žáner). Môžeme sledovať zmenu obtiažnosti rovnakých úloh naprieč hudobnými žánrami – takáto zmena je sledovateľná v ľubovoľnej použitej doméne. Keďže obtiažnosti boli počítané až na úroveň skladby, výsledná hodnota pre hudobný žáner je aritmetický priemer obtiažnosti skladieb patriacich pod daný hudobný žáner.

tab. 5.1 HUB skóre (obtiažnosť) jednotlivých úloh podľa hudobného žánru - priemerné hodnoty

typ minihry	pop	rock	alternative	old	globálne
interpret	0,074	0,050	0,068	0,044	0,045
mesto interpreta	0,073	0,044	0,059	0,042	0,041
krajina	0,072	0,049	0,066	0,044	0,045
tónina	0,054	0,038	0,029	0,036	0,032
album	0,057	0,039	0,062	0,037	0,034
názov skladby	0,069	0,048	0,065	0,044	0,042
tempo	0,031	0,031	0,037	0,020	0,022
rok začiatku interpreta	0,062	0,042	0,060	0,038	0,037

V tab. 5.2 sú zobrazené maximálne namerané hodnoty pre jednotlivé hudobné žánre. Globálne maximum nemusí nutne dosahovať jedno z maxím pre hudobný žáner, keďže do výpočtu *hub* skóre bolo naraz zahrnutých viac skladieb, ktoré na seba vplývali inak, ako pri oddelenom výpočte pre jednotlivé žánre.

tab. 5.2 HUB skóre (obtiažnosť) jednotlivých úloh podľa hudobného žánru - maximá

typ minihry	pop	rock	alternative	old	globálne
interpret	0,377	0,130	0,224	0,142	0,248

mesto interpreta	0,226	0,106	0,156	0,106	0,142
krajina	0,312	0,137	0,192	0,142	0,197
tónina	0,087	0,062	0,043	0,088	0,091
album	0,187	0,063	0,205	0,090	0,150
názov skladby	0,312	0,130	0,181	0,149	0,193
tempo	0,052	0,052	0,058	0,025	0,040
rok začiatku interpreta	0,298	0,088	0,143	0,054	0,186

Zhodnotenie. Výpočet expertízy v doméne sa ustalaťuje s počtom odohratých skladieb. Pre každú doménu je počet skladieb, po ktorých sa hodnota začína ustalaťovať, iný. Expertíza má tendenciu jemne kulminovať aj po ustálení z dôvodu učenia sa hráčov a nájdenia systému v hre.

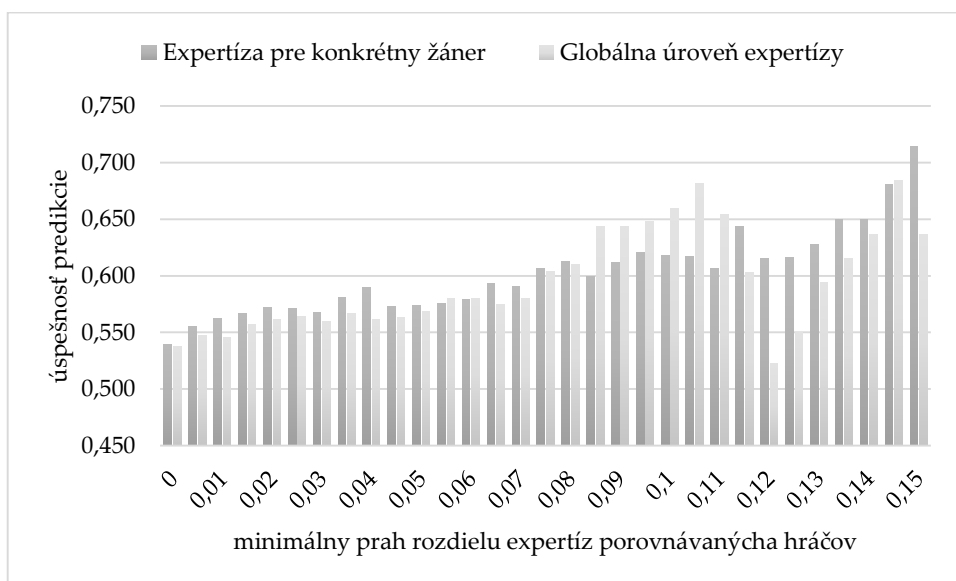
5.2.2. Predikcia výsledkov

Na predikciu je možné použiť dva modely:

- Predikovať výsledok (poradie všetkých zúčastnených hráčov) nasledujúcej hry.
- Predikovať úspešnejšieho z dvojice (kombinácia dvojíc všetkých zúčastnených hráčov) nasledujúcej hry.

V prvom modeli by sa pri vyhodnocovaní jednalo o silnejšiu predikciu – predikovať konkrétne poradie viac hráčov je náchylnejšie na chybu, keďže nesprávne určenie poradia jedného (a teda v skutočnosti dvoch) hráčov označuje celú predikciu za nesprávnu a to aj v prípade, ak bolo poradie ostatných hráčov určené správne. Na ďalšie vyhodnocovanie bol teda použitý model porovnávania dvojíc zúčastnených hráčov.

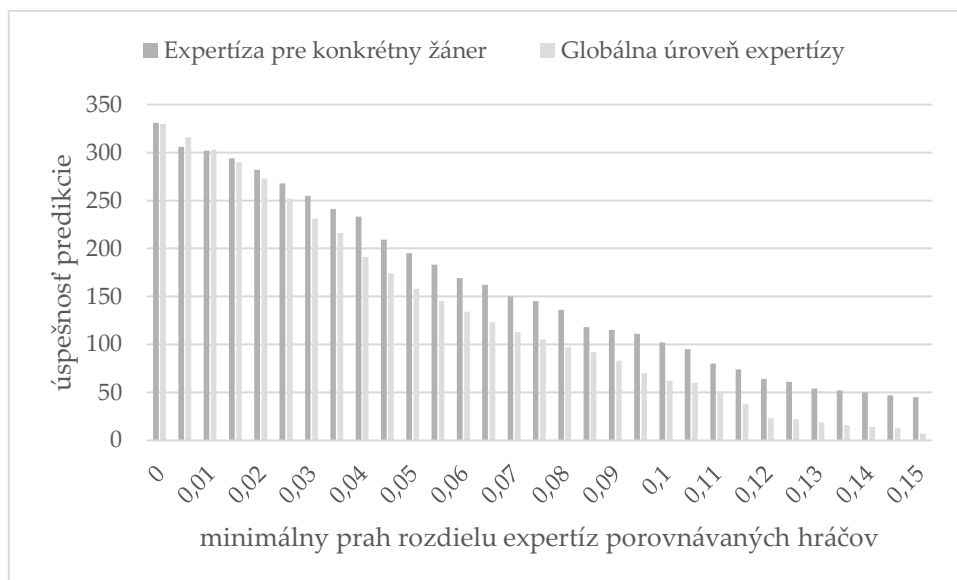
Rovnako ako v časti 5.2.1 aj tu sme našu metódu podrobili porovnaniu s výpočtom úrovne expertízy, ktorý nebral do úvahy jednotlivé hudobné žánre. Do porovnávania sme zapojili ďalší atribút na spresnenie výsledkov: minimálny prah rozdielu expertíz porovnávania hráčov. Teda bol prah nastavený na 0,05 a rozdiel expertíz hráčov bol 0,02, nebol výsledok týchto hráčov predikovaný. Výsledky sú prezentované v grafe na obr. 5.5.



obr. 5.5 Úspešnosť predikcie umiestnenia dvojíc hráčov na základe expertíz vypočítaných našou metódou

Predikcia expertízy na základe výpočtu pre konkrétny žáner z celkového hľadiska vychádza lepšie. Medzi prahmi 0,08-0,11 síce vyšla predikcia pre globálnu expertízu lepšie, náhle však potom padla.

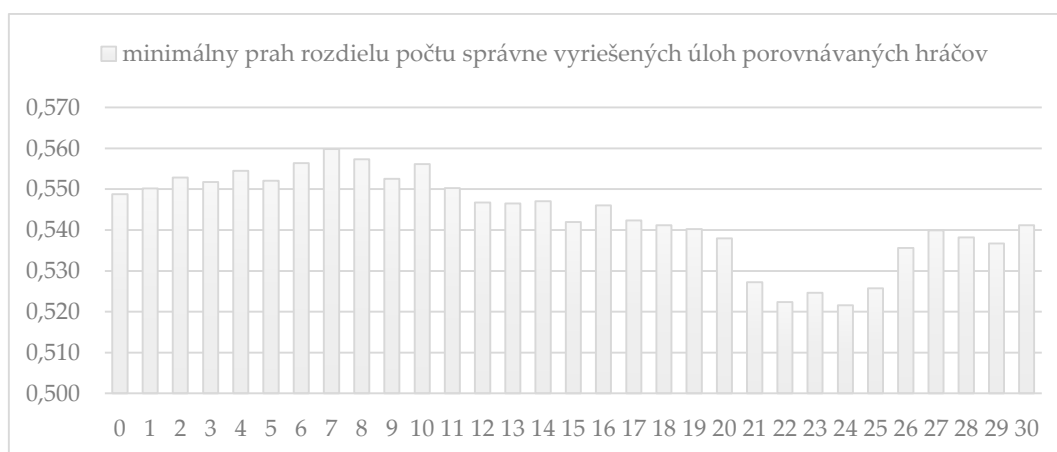
Tieto výkyvy sú spôsobené počtom dvojíc, ktoré sa do porovnávania dostali. Podľa porovnania hodnôt na obr. 5.3 a obr. 5.1 je zrejmé, že rozdiely medzi expertízami hráčov počítaných globálne sú menšie ako rozdiely v rámci jednotlivých žánrov. Ak sme teda definovali potrebný prah rozdielu na 0,1, do predikcie výsledkov založenej na globálnych expertízach sa reálne dostalo menej dvojíc ako do predikcie výsledkov založenej na expertízach pre konkrétne žánre. Túto skutočnosť prezentujeme aj na obr. 5.6.



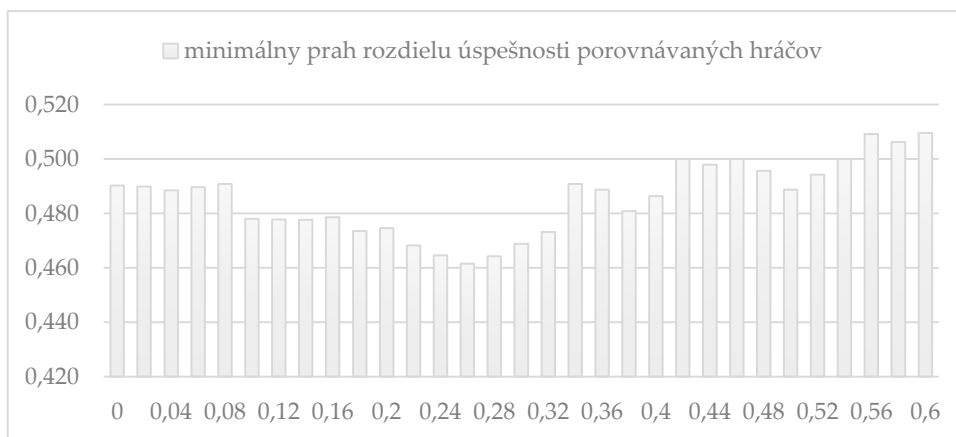
obr. 5.6 Porovnanie počtu dvojíc, ktoré boli použité pri predikcii úspešnosti

Práce zaoberajúce sa rozpoznávaním expertov štandardne vyhodnocujú svoju úspešnosť rozpoznávaním konkrétnych expertov – priamym určením, kto je expert. Keďže takéto určovanie nie je cieľom našej práce, ako referenčné metriky voči ktorým porovnáваме našu metódu sme zvolili počet správnych odpovedí hráča a pomer správnych a nesprávnych odpovedí hráča.

Referenčné metriky sme vyhodnocovali v rámci žánra – napr. pre vybranú skladbu sme brali do úvahy dovtedajší počet správnych odpovedí hráča na danom žánri. Výsledky týchto dvoch metrik prezentujeme na obr. 5.7 a obr. 5.8.



obr. 5.7 Úspešnosť predikcie umiestnenia dvojíc hráčov na základe expertíz vypočítaných referenčnou metrikou – počtom správnych odpovedí



obr. 5.8 Úspešnosť predikcie umiestnenia dvojíc hráčov na základe expertíz vypočítaných referenčnou metrikou – úspešnosťou vyriešených úloh (počet správnych / počet nesprávnych)

Prezentované výsledky referenčných metód ukazujú, že ich použitie generuje výsledky s úspešnosťou 50% - čo je porovnateľné so štatistickým normálnym rozdelením. Oproti našej metóde z nej nie je možné vyčítať žiadne vzory správania.

Zhodnotenie. V tejto časti sme prezentovali úspešnosť metódy pri predikcii výsledkov. Porovnali sme úspešnosť predikcie ak metóda brala do úvahy aj hudobné žánre (domény) a ak ich do úvahy nebrala. Úspešnosť metódy v závislosti od nastavenia parametrov osciluje medzi 55%-70%. Metóda je stabilnejšia ak berie do úvahy aj žánre a navyše je schopná pracovať s väčším množstvom dát. V prípade ignorovania žánrov boli výsledky metódy pre rôznych hráčov omnoho podobnejšie a nebolo možné ich porovnať. V závere sme porovnali našu metódu s jednoduchými referenčnými metrikami, ktorých úspešnosť oscilovala okolo 50%.

5.3. Overenie zlepšenia kvality generovania metadát davom v hre

Hypotéza. Rozpoznaním expertov a zvýhodnením nimi generovaných informácií vieme zrýchliť a skvalitniť proces tvorby metadát za pomoci davu v hrách s účelom.

Proces. Do herného módu rádia sme implementovali hru s podobným scenárom, ktorý sme použili aj v hre s účelom City Lights [1]. Hráči k vybranej skladbe dostali niekoľko množín anotácií skladieb – jedna z nich patrí skladbe, ktorá v rádiu práve hrá. Ich úlohou je vybrať tú správnu. Po výbere množiny sme hráčovi dali možnosť označiť anotácie, ktoré ho presvedčili vybrať si túto možnosť – tu je rozdiel oproti hre City Lights, kde hráč označoval anotácie, ktoré považuje za zlé.

Anotácie náhodne rotujú a vďaka tomu môže byť úspešnosť hráčov pri výbere (a následná explicitná väzba) prenesená na určenie správnosti jednotlivých anotácií v rámci množín. Skóre jednotlivých anotácií bolo počítané dvoma spôsobmi: (1) výhradne na základe akcií hráčov a (2) kombináciou akcií hráčov a ich úrovne expertízy.

Výhradne na základe akcií hráčov. Postup bol podobný ako pri vyhodnocovaní hry City Lights. Akcie hráčov vplývali na zmenu skóre jednotlivých anotácií. Keďže City Lights malo relatívne nízku úspešnosť pri overovaní správnych anotácií, rozhodli sme sa zmeniť scenár pri explicitnej spätnej väzbe a hráči mali vybrať anotácie, ktoré považujú za správne.

Experimentálne sme overovali optimálne nastavenie váh jednotlivých akcií hráča (implicitné a explicitné akcie nad anotáciou) a nastavili sme ich nasledovne:

Typ akcie hráča	Váha akcie
implicitná správna	0,15
implicitná nesprávna	-0,30
explicitná správna	0,90
explicitná nesprávna	-0,60

Z dôvodu rozdielneho zamerania minihry na overenie anotácií voči hre City Lights nebolo možné smerovať dizajn hry na čo najrýchlejšie vyhodnotenie anotácií – počet skutočne overených anotácií nie je až taký veľký ako pri hre City Lights, avšak výrazne vzrástol počet anotácií zapojených do procesu. Vďaka tomu vieme, či je vplyv expertov značný už po jednom-dvoch zapojeniach anotácií do procesu.

Kombináciou akcií hráčov a úrovne expertízy. Na začiatku sme určili referenčné hodnoty premennej prepájajúcej výpočet bez zapojenia experta a s jeho zapojením – teda expertné skóre, pri ktorom sa skóre anotácie nezmení. Jej hodnota bola vypočítaná ako priemer expertíz všetkých hráčov pre jednotlivé žánre. Do tohto prímeru boli započítavané aj nulové hodnoty hráčov, ktorí s daným hudobným žánrom (doménou) neinteragovali:

Hudobný žánr	Priemerná hodnota expertízy
pop	0,03950
jazzová hudba a skladby z filmov	0,03293
alternatívne	0,04855
rock	0,04158

Ak teda hráč s priemernou expertízou v danom žánri vykonal ľubovoľnú akciu, váha akcie a teda aj výsledná zmena anotácie bola totožná ako pri ohodnotení bez zapojenia expertízy.

Pri vyhodnocovaní implicitných akcií sme veľkosť zmeny zhora obmedzili, keďže metóda bez tohto obmedzenia generovala na základe implicitných akcií s priveľkou váhou nepresné výsledky. Horný limit implicitných akcií vyhodnocovaných s expertízou bol nastavený na výšku váh implicitných akcií pri vyhodnocovaní bez expertízy.

Výsledky. Do vyhodnotenia bolo polo-náhodne vybraných 200 anotácií. Jediné obmedzenie pri výbere bola nutnosť rozdielu vypočítaného skóre oboma spomenutými spôsobmi aspoň 0,4 – tento rozdiel reálne predstavuje aspoň dve implicitné alebo jednu explicitnú akciu hráča, čo považujeme za dostatočne signifikantné. Kandidátov spĺňajúcich túto podmienku bolo 769, nespĺňajúcich 6084. Správnosť náhodne vybratých anotácií bola manuálne overená - 110 anotácií bolo správnych. 90 nesprávnych. Jednotlivé metódy na výpočet skóre anotácie sme porovnali v rôznych smeroch opísaných v nasledujúcej časti kapitoly.

Validované anotácie. Keďže počet hráčov, ktorí ovplyvnili konkrétnu anotáciu bol výrazne obmedzený, ani jedna z anotácií ohodnotených výhradne na základe akcií hráča nepresiahla hranicu potvrdenia/vyradenia určenú v hre City Lights ($abs(skóre) > 4$). Po zapojení úrovne expertízy túto hranicu prekročila iba jedna anotácia, tá bola ohodnotená správne – vyradená.

Trend skóre. 151 z 200 anotácií bolo bližšie k správnej hranici pri použití štandardného ohodnotenia – teda ak bola anotácia manuálne vyhodnotená ako správna, jej skóre bolo kladné; analogicky pre nesprávne. Po použití ohodnotenia používajúceho expertízu tento počet klesol na 149 anotácií. Z globálneho hľadiska ale použitie expertného skóre vplývalo na množinu anotácií kladne:

- Až pre 147 z 200 anotácií (73,5%) znamenalo zlepšenie (spresnenie) skóre v priemere o 0,8349 bodu. Toto zlepšenie bolo viditeľné už po 2 vstupoch anotácie do procesu hodnotenia akciami hráča.
- Pre zostávajúcich 53 anotácií (26,5%) nastal vďaka využitiu expertov pokles resp. vzdialenie sa hranici, ktorú sa snažíme dosiahnuť. V priemere toto zhoršenie nastalo o 0,8877 bodu.

Anotácie, na ktorých prebiehalo hodnotenie (ako už bolo vyššie spomenuté), boli vybrané polo-náhodne a podmienkou na ich výber bol rozdiel jednotlivých ohodnotení aspoň 0,4 bodu čo zabezpečilo odstránenie šumu a malých rozdielov medzi skóre. V priemere naprieč všetkými 200 sledovanými anotáciami sledujeme pri výpočte skóre s použitím expertízy zlepšenie skóre o 0,3784 bodu.

Pokles a zhoršenie skóre v niektorých prípadoch bol prirodzeným výstupom metódy, kedy implicitné aj explicitné váhy akcií niektorých hráčov boli na základe ich expertízy znížené. Toto zníženie považujeme ako prijateľnú cenu za celkové zlepšenie metódy. V ostatných prípadoch metóda bohužiaľ trpela buď kvôli nedostatku dát pre konkrétneho používateľa (čo sa odrazilo v nepresnom ohodnotení expertízy), nejednoznačnou správnosťou anotácie, ktorú nevieme odstrániť (viacslovné názvy, kedy len časť anotácie nie je správna) alebo jednoducho chybovosťou používateľov.

Zhodnotenie. Overenie zlepšenia a zrýchlenia procesu prebehlo na polo-náhodnej vzorke 200 anotácií. Zapojenie expertov do metódy v priemere zlepšilo ohodnotenie anotácií – pri predvolenom nastavení sa ich skóre v priemere priblížilo o 10% k očakávanej hranici potvrdenia/vyradenia anotácie.

Časť nepresností a zhoršenia skóre bola spôsobená nedostatočným množstvom dát pre niektorých používateľov, v iných prípadoch metóda jednoducho znížila relevanciu hráčových akcií a výsledné skóre sa vzdialilo od očakávaného prahu. Napriek tomu považujeme zlepšenie procesu validácie našou metódou za nezanedbateľné.

6. Zhodnotenie

V práci sme opísali spôsoby získavania informácií. Opísali sme konkrétne možnosti využitia davu na získanie informácií a na príkladoch sme prezentovali problémy, ktoré s využitím davu súvisia. Prezentovali sme hry s účelom, dôvod ich vzniku a ich využitie. Ďalej sme sa zaoberali rozpoznávaním expertov v dave a integráciou tohto rozpoznávania do hry s účelom – primárnym dôvodom integrácie je zrýchlenie a spresnenie procesu generovania a validovania artefaktov v hre s účelom.

Navrhli sme hru s účelom s využitím rozpoznaných expertov, ktorú sme integrovali do nami vytvoreného internetového rádia. Rádio je postavené na 4 nami definovaných hudobných žánroch (doménach). Hra je postavená na sústave faktoidných minihier, ktoré súvisia s práve hranou skladbou a ich správnosť je overiteľná – experti majú byť rozpoznaní na základe týchto minihier. Cieľom hráča je byť najlepší na jednej skladbe, za čo je odmenený žetónom, ktorý môže využiť na priame ovplyvnenie playlistu. Hra ako taká ma na podporu efektu zábavy implementované rebríčky, odznaky pre hráčov za splnenie rôznych úloh a prúd aktualizácií obnovovaný v reálnom čase, aby hráči videli úspešnosť ostatných hráčov.

Účel našej hry s účelom je (ako v našom minulom projekte) validovať hudobné anotácie zo služby LastFM. Do tejto validácie sú integrovaní aj rozpoznaní experti. Na rozpoznávanie expertov sme navrhli metódu založenú na algoritme HITS. Túto metódu sme mierne upravili a zaviedli pomocné postupy (napr. zaradenie interpreta do nami definovanej domény), ktoré boli na jej využitie potrebné.

V práci sme navrhnutú metódu na rozpoznávanie expertov (v 2 experimentoch) a zároveň aj vplyv využitia expertov na proces tvorby metadát (1 experiment) overili. Úspešnosť rozpoznania expertov sa v závislosti od nastavenia metódy pohybovala v najlepšom prípade okolo 70%, čo je v oblasti rozpoznávania expertov výborné. Proces validácie anotácií sa v priemere zlepšil – konkrétne hodnoty neuvádzame, keďže výrazne závisia od nastavenia výpočtu a bez kontextu sú neinterpretovateľné; čitateľovi odporúčame kapitolu 5.3.

Hru a internetové rádio sme mali v pláne rozšíriť o ďalšie (hlavne hru podporujúce) prvky, do rozsahu tohto projektu sa už bohužiaľ nevošli. V záujme motivácie ďalšieho výskumu uvádzame tie najslabšie:

- *Určovanie expertov davom samotným.* Jedna z našich neimplementovaných hypotéz bola, že dav je na základe objektívnych signálov (hráčom pridané skladby, úspešnosť hráča v hrách) spomedzi seba schopný autonómne zvoliť experta resp. ohodnotiť úroveň expertízy časti hráčskej základne.
- *Rozdelenie hráčov do frakcií.* Aby sme podporili predchádzajúcu hypotézu a zároveň pridali do hry viac súťaživosti, chceli sme hráčov rozdeliť do medzi sebou súperiacich frakcií. Zachovala by sa súťaživosť, ktorá by vznikla snahou poraziť niekoho a hra by získala tímovú spoluprácu, ktorá slúži ako jeden z motivátorov prečo sa vôbec hrať.
- *Zapojenie hráčových hudobných preferencií do výpočtu úrovne expertízy a integrácia so službami tretích strán.* Určovanie expertízy sa skladá z dvoch krokov: určenia doménových preferencií hráča a výpočtu úrovne expertízy v danej doméne. Dáta získané hrou a zároveň

dáta existujúce v externých službách (napr. LastFM) by mohli byť použité v oboch týchto bodoch.

- *Strojovo riadení protivníci.* Jedným z problémov hry bola nedostupnosť základnej hráčskej základne. Často sa stalo, že bol súčasne pripojený iba jeden hráč, a nemal s kým súperiť. Minimálne tento faktor môže byť nahraditeľný umelou inteligenciou alebo použitím v minulosti odohraných hier. Vytvorenie rámca na strojovo riadených protivníkov pre určité typy hier s účelom by výrazne posunulo ich vývoj vpred.

Napriek našej snahe sme neboli schopní realizovať všetky naše hypotézy a plány, ktoré sme s projektom mali. Výskum v oblasti hier s účelom a objavovania autorít je široká oblasť a dúfame, že náš prínos do nej bude motivovať k ďalšiemu výskumu.

7. Bibliografia

- [1] P. Dulačka, J. Šimko, and M. Bieliková, “Validation of music metadata via game with a purpose,” in *Proceedings of the 8th International Conference on Semantic Systems - I-SEMANTICS '12*, 2012, p. 177.
- [2] M. Kosinski and Y. Bachrach, “Crowd IQ: Measuring the intelligence of crowdsourcing platforms,” *Proceeding WebSci '12 Proc. 3rd Annu. ACM Web Sci. Conf.*, pp. 151–160, 2012.
- [3] L. von Ahn and L. Dabbish, “Labeling images with a computer game,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2004, vol. 6, no. 1, pp. 319–326.
- [4] J. Ross, L. Irani, and M. Silberman, “Who are the crowdworkers?: shifting demographics in mechanical turk,” *Proceeding CHI EA '10 CHI '10 Ext. Abstr. Hum. Factors Comput. Syst.*, pp. 2863–2872, 2010.
- [5] J. Howe, “The Rise of Crowdsourcing,” *Wired Mag.*, no. 14, pp. 1–5, 2006.
- [6] Y. Liu and E. Agichtein, “On the evolution of the yahoo! answers QA community,” *Proceeding SIGIR '08 Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, p. 737, 2008.
- [7] L. Adamic and J. Zhang, “Knowledge sharing and yahoo answers: everyone knows something,” *Proceeding WWW '08 Proc. 17th Int. Conf. World Wide Web*, pp. 665–674, 2008.
- [8] F. Riahi, Z. Zolaktaf, M. Shafiei, and E. Milios, “Finding expert users in community question answering,” *Proc. 21st Int. Conf. companion World Wide Web - WWW '12 Companion*, no. i, p. 791, 2012.
- [9] S. Nasehi and J. Sillito, “What makes a good code example?: A study of programming Q&A in StackOverflow,” *28th IEEE Int. Conf. Softw. Maint.*, pp. 25–34, 2012.
- [10] L. Von Ahn, R. Liu, and M. Blum, “Peekaboom: a game for locating objects in images,” *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, pp. 55–64, 2006.
- [11] B. G. Morton, J. A. Speck, and E. M. Schmidt, “Improving music emotion labeling using human computation,” *Proc. ACM SIGKDD Work. Hum. Comput.*, pp. 45–48, 2010.
- [12] J. Simko, M. Tvarožek, and M. Bieliková, “Little search game: term network acquisition via a human computation game,” *Proc. 22nd ACM Conf. Hypertext hypermedia*, pp. 57–61, 2011.
- [13] L. von Ahn and L. Dabbish, “Designing games with a purpose,” *Commun. ACM*, vol. 51, no. 8, p. 57, Aug. 2008.
- [14] K. Siorpaes and M. Hepp, “Games with a Purpose for the Semantic Web Games with a Purpose for the Semantic Web,” *Computer (Long. Beach. Calif.)*, vol. 23, no. 3, 2008.

- [15] D. Roman, “Crowdsourcing and the question of expertise,” *Commun. ACM*, vol. 52, no. 12, p. 12, Dec. 2009.
- [16] J. Šimko, M. Tvarožek, and M. Bieliková, “Human computation: Image metadata acquisition based on a single-player annotation game,” *Int. J. Human- ...*, vol. 71, no. 10, pp. 933–945, Oct. 2013.
- [17] J. Simko, M. Tvarožek, and M. Bieliková, “Semantics Discovery via Human Computation Games,” *Int. J. Semant. Web Inf. Syst.*, vol. 7, no. 3, p. 23, 2011.
- [18] J. D. Bransford, A. L. Brown, and R. R. Cocking, *How People Learn: Brain, Mind, Experiencem, and School*. National Academy Press, 2004.
- [19] D. Yimam-Seid and A. Kobsa, “Expert-finding systems for organizations: Problem and domain analysis and the DEMOIR approach,” *J. Organ. Comput. ...*, no. 1997, 2003.
- [20] M. Bouguessa, B. Dumoulin, and S. Wang, “Identifying authoritative actors in question-answering forums: the case of yahoo! answers,” *Proc. 14th ACM ...*, pp. 866–874, 2008.
- [21] K. R. Lakhani and E. Von Hippel, “How open source software works : ‘ free ’ user-to-user assistance,” vol. 32, no. July 2002, pp. 923–943, 2003.
- [22] L. Mamykina, B. Manoim, M. Mittal, G. Hripcsak, and B. Hartmann, “Design lessons from the fastest q&a site in the west,” *Proc. 2011 Annu. Conf. Hum. factors Comput. Syst. - CHI '11*, p. 2857, 2011.
- [23] T. J. Allen, *Managing the Flow of Technology*. MIT Press, 1977, p. 320.
- [24] A. Mockus and J. D. Herbsleb, “Expertise Browser: a quantitative approach to identifying expertise,” *Proc. 24th Int. Conf. Softw. Eng. ICSE 2002*, pp. 503–512, 2002.
- [25] C. Chiou and J. Hsu, “Capability-aligned matching: Improving quality of games of a purpose,” in *Proceeding AAMAS '11 The 10th International Conference on Autonomous Agents and Multiagent Systems*, 2011, no. Aamas, pp. 2–6.
- [26] K. Balog and M. De Rijke, “Determining Expert Profiles (With an Application to Expert Finding),” pp. 2657–2662.
- [27] J. Zhang, M. Ackerman, and L. Adamic, “Expertise networks in online communities: structure and algorithms,” *Proc. 16th Int. ...*, pp. 221–230, 2007.
- [28] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank citation ranking: bringing order to the web.,” pp. 1–17, 1999.
- [29] J. M. Kleinberg, “Authoritative sources in a hyperlinked environment,” *J. ACM*, vol. 46, no. 5, pp. 604–632, Sep. 1999.
- [30] A. Borodin and G. Roberts, “Link analysis ranking: algorithms, theory, and experiments,” *ACM Trans. ...*, vol. 5, no. 1, pp. 231–297, Feb. 2005.
- [31] K. Balog, L. Azzopardi, and M. De Rijke, “Formal models for expert finding in enterprise corpora,” ... *29th Annu. Int. ACM ...*, pp. 43–50, 2006.

- [32] C. Zhai and J. Lafferty, “A study of smoothing methods for language models applied to information retrieval,” *ACM Trans. Inf. Syst.*, vol. 22, no. 2, pp. 179–214, Apr. 2004.
- [33] H. Kwak, C. Lee, H. Park, and S. Moon, “What is Twitter , a Social Network or a News Media ? Categories and Subject Descriptors,” pp. 591–600, 2010.
- [34] S. Ghosh, N. Sharma, and F. Benevenuto, “Cognos: crowdsourcing search for topic experts in microblogs,” *Proc. 35th ...*, no. 1, pp. 575–584, 2012.
- [35] C. Cao, J. She, Y. Tong, and L. Chen, “Whom to ask? Jury selection for decision making tasks on micro-blog services,” *Proc. VLDB Endow.*, vol. 5, no. 11, pp. 1495–1506, 2012.
- [36] E. David and K. Jon, “Networks, Crowds, and Markets: Reasoning About a Highly Connected World,” Jul. 2010.