

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

FIIT-5212-64327

Peter Dulačka

Získavanie sémantiky a metadát pomocou hier s účelom

Bakalárska práca

Študijný program: Informatika

Študijný odbor: 9.2.1 Informatika

Miesto vypracovania: Ústav Informatiky a Softvérového Inžinierstva, FIIT STU

Vedúci práce: Ing. Jakub Šimko

máj 2012

Anotácia

Slovenská Technická Univerzita v Bratislave
Fakulta Informatiky a Informačných Technológií

Študijný program: Informatika

Autor: Peter Dulačka

Bakalársky projekt: Získavanie sémantiky a metadát pomocou hier s účelom

Vedúci bakalárskeho projektu: Ing. Jakub Šimko

máj 2012

Informácie na Webe sú dostupné v rôznych formách a počítače ich dokážu spracovať (teda vedieť získať a použiť informácie o obsahu entity) len do určitej miery. Spracovanie vybraných informácií je podmienené iba ľudskou interakciou - niektoré veci je správne schopný posúdiť iba človek. Títo ľudia sú však takéto akcie ochotní vykonať len ak sú dostatočne motivovaní. Tento problém riešia výskumníci pomocou „hier s účelom“, kde hráči svojou aktivitou (často nevedome) dopĺňajú automatické prístupy a skvalitňujú tak spracované údaje.

Táto práca sa zaoberá spracovaním hudobných metadát a zameriava sa na validáciu existujúcich hudobných anotácií. Na Webe je množstvo užitočných anotácií zmiešaných spolu s neužitočnými, čo následne deformuje očakávané výsledky. Tieto anotácie sú zväčša vytvorené ľuďmi a preto títo ľudia nemajú žiadnu motiváciu ich bez odmeny po sebe navzájom overovať.

Práca opisuje „*CityLights*“, hru s účelom validovať hudobné anotácie. Hráči počúvajú skladby a dostávajú na výber množiny už existujúcich anotácií. Na základe hrajúcej skladby sa musia rozhodnúť, ktorá množina anotácií je patrí danej skladbe. S existujúcou možnosťou vyžiadať prídavné anotácie a nesprávnymi tipmi, ktoré hráči vykonávajú sme schopní určiť správnosť poskytnutých množín anotácií a odstrániť nesprávne. Správnosť hry a kvalitu získaného výstupu sme overili experimentmi.

Annotation

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

Degree Course: Informatics

Author: Peter Dulačka

Bachelor thesis: Acquiring Semantics and Metadata via Games with a Purpose

Supervisor: Ing. Jakub Šimko

May 2012

There is a lot of information available on the Web provided in different forms and machines cannot process them completely because of it. Processing of some kind of information is only possible with human interaction. However people need to be motivated to do such interactions. Researchers are attempting to solve this problem with “games with a purpose”. Players with their activity (mostly unknowingly) fill the gaps created by automatic approaches thus they improve the quality of processed data.

This paper deals with processing music metadata and focuses only on validating existing annotations instead of creating them. There is a lot of useful music metadata on the Web mixed with useless data in common sets of annotations which makes expected results worse. These annotations are created mostly by people therefor people have no motivation to validate it again with no reward given.

This paper describes “*CityLights*”, a game with a purpose for validating music annotations. Players listen to the song and they are presented sets of existing annotations. According to given song they have to decide which set of annotations belongs to the song. Thus the existing sets can get more accurate and provide better search and recommendation results.

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

Bratislava, máj 2012

podpis

Chcel by som hlavne poďakovať Ing. Jakubovi Šimkovi za odborné vedenie, nápady a podporu počas vypracovávania práce. Taktiež PeWe za spätnú väzbu a pomoc s realizáciou experimentu a všetkým, ktorý sa hru s účelom zahráli alebo akýmkoľvek iným spôsobom prispeli k skvalitneniu tejto práce.

Osobitné poďakovanie patrí rodine a priateľom za ustavičnú podporu.

Ďakujem.

OBSAH

1	Úvod	1
2	Získavanie metadát pomocou hier s účelom	5
2.1	Hry s účelom	5
2.2	Prehľad existujúcich hier	7
2.2.1	Hry s účelmi z rôznych oblastí	7
2.2.2	Hry s účelom anotovať hudbu	11
3	Získavanie metadát k hudbe	16
3.1	Dolovanie hudobných znalostí	16
3.1.1	Songs2See	17
3.1.2	GlobalMusic2One	18
3.2	Služby zamerané na získavanie hudobných anotácií	19
3.2.1	Pandora	19
3.2.2	LastFM	19
3.2.3	MusicBrainz	20
4	City Lights	21
4.1	Špecifikácia hry	21
4.2	Návrh aplikácie realizujúcej hru	24
4.2.1	Štruktúra aplikácie	24
4.2.2	Moduly realizujúce aplikáciu	25
4.2.3	Grafické prostredie	26
4.2.4	Herná logika	27
4.3	Realizácia hry s účelom	29
4.3.1	Zdroj údajov pre hru a spôsoby prezentácie	29
4.3.2	Aplikácia	30
4.3.3	Scenár herného kola	33
5	Experiment	35
5.1	Spôsob overovania získaných dát	35
5.2	Uzavretý experiment	35
5.3	Otvorený experiment	37
5.3.1	Výsledky živého experimentu	37
5.3.2	Výsledky najlepšieho simulovaného experimentu	38
6	Zhodnotenie	41
6.1	Možnosti rozšírenia hry	41
7	Bibliografia	43
PRÍLOHY		45
A	Technická dokumentácia	47
B	Používateľská príručka	51
C	Obsah priloženého CD	55
D	Článok z I-Semantics 2012	57
E	Plagát k článku z IIT.SRC 2012	63

1 Úvod

S nástupom Webu 2.0 sa niekoľkonásobne zväčšilo množstvo používateľmi vytváraných a zdieľaných informácií a zdrojov. Vyhľadávače dokážu do určitej miery rozoznať dôležitosť obsahu stránky a nový obsah je indexovaný a zobrazovaný rádovo v minútach až hodinách od jeho zverejnenia na Webe. Tento trend je znateľné zlepšenie oproti stavu spred niekoľkých rokov, kedy vyhľadávačom indexovanie trvalo rádovo dni až mesiace. Táto rýchlosť závisí od predpokladanej obľúbenosti indexovaného obsahu, ktorú vyhľadávače vedia určiť pomocou rôznych kritérií (obľúbenosť podobného obsahu, obľúbenosť zdrojovej stránky a pod.). Toto je možné aj vďaka tomu, že indexovaný obsah je správne anotovaný. Ideálny model pre vyhľadávanie nastane, keď každý používateľ pri pridani dát na sieť svoje dáta aj správne oanoťoval (odhliadnuc od automatickej anotácie údajov) – nereálna požiadavka, ktorú si v súčasnej dobe uvedomuje a akceptuje len malé množstvo používateľov.

Používatelia pridávajúci obsah sú často pri anotovaní nahrádzaní inými, niekedy viac a niekedy menej účinnými metódami na získavanie anotácií:

- **Expertmi.** Experti sú najlepšími vo svojej doménovej oblasti a sú schopní obsah anotovať najpresnejšie. Takéto anotácie dokážu od seba odlíšiť aj dve veľmi podobné entity – napríklad dva obrázky s kvetom, kedy laik vie opísať iba kvet všeobecne, no expert pozná kvet podrobne a vie vytvoriť presnejšiu anotáciu. Nanešťastie tento spôsob získavania anotácií je časovo najnáročnejší a nákladný a používa sa hlavne pri nutnosti získania čo najsprávnejších dát (tzv. referenčnej vzorky), ktoré sú neskôr použité na overenie správnosti dát získaných inými (zväčša menej presnými) prístupmi.
- **Získavanie dát masami** – „crowdsourcing“. Táto metóda je pomerne presná a jediným jej obmedzením je veľkosť mas. Získané anotácie sú ale presné iba do určitej miery a nie všetky anotácie je možné považovať za správne. Preto sa pri tejto metóde zavádza validácia anotácií, kedy sú považované za správne až keď rovnakú anotáciu vytvorí niekoľko na sebe nezávislých zdrojov. Nevýhodou tohto spôsobu je filtrácia prípadných expertných anotácií, ktoré sa pri takejto forme validácie tvária ako neoverené a nedostanú sa do finálnej množiny anotácií. Masy majú zvyčajne jednu z dvoch motivácií, prečo obsah anotovať:
 - **Osobný zámer.** V tomto prípade používatelia vytvárajú anotácie aby mohli dáta, s ktorými pracujú, sami jednoduchšie prehľadávať a kategorizovať. V prípade často používanej anotácie je táto povýšená a môže byť použitá pri globálnom vyhľadávaní (služby del.icio.us, LastFM).
 - **Odmena.** V prípade, že je ľuďom ponúknutá náhrada za čas strávený anotáciami, je pravdepodobnejšie že anotovať budú aj naďalej. Takýmito odmenami väčšinou bývajú nadstavbové služby aplikácií prípadne prístup k platenému obsahu.

- **Automatický prístupom.** Automatický prístup sa spolieha na už v minulosti vytvorené anotácie a podobnosť objektov. Dolovaním dát a inými rôznymi technikami sa snažia získať informácie o objekte a na základe nich ho automaticky anotovať. Avšak takto získané dáta taktiež trpia nepresnosťami a je potrebné ich validovať.

V praxi sa využívajú kombinované prístupy vyššie spomenutých metód – hlavne získavanie dát masami v kombinácii s automatickým prístupom, kedy používatelia môžu byť tvorcami anotácií alebo ich validátormi.

Na princípe získavania dát masami sú postavené aj hry s účelom („*games with a purpose*“ – *GWAP*). Tieto hry sa snažia nasmerovať masu na riešenie určitého problému (získanie anotácie, vytvorenie väzieb medzi entitami a pod.), ktorý nie je možné vyriešiť automatickým prístupom avšak pre človeka je jeho vyriešenie triviálne [16]. Najväčšie obmedzenie tejto metódy – veľkosť mas – sa snažia eliminovať efektom zábavy ako formou odmeny, ktorý udrží hráča dlhšie pri hre. Hráči samotní si môžu, ale nemusia byť vedomí účelu hry – niekedy je pre hráča viac motivujúce, keď vie, že svojou hrou niečomu prispieva. Ako príklad hry s účelom je možné uviesť „*ESP Game*“ od von Ahna, kde hráči hraním hry dokázali anotovať ľubovoľný obrázok. Dvaja náhodne spárovaní hráči sa museli zhodnúť na jednom z nimi vložených slov a jediná pomôcka, ktorú mali bol obrázok. Tým boli nútení vpisovať slová súvisiace s obrázkom a teda ho anotovali.

Naším cieľom je vytvoriť hru, ktorá spĺňa spomenuté požiadavky a dokáže generovať použiteľný výstup, ktorý automatický prístup tak presne generovať nedokáže. Problémom automatických prístupov je (1) nutnosť mať iniciálnu vzorku dát, na základe ktorej sa budú generovať nové dáta a (2) využívanie logických predikátov, ktoré nie vždy implikujú správne vytvorenie väzby a teda relatívne nízka úspešnosť generovaných dát.

Nami navrhovaná hra bude pracovať s množinami hudobných anotácií získaných z webu, ktoré boli z časti získané metódou získavania masami a z časti automatickými prístupmi. Oblasť, ktorá sa zaoberá získaním takýchto dát sa nazýva *Získavanie hudobných informácií* (z angl. *MIR* - „*Music Information Retrieval*“) a je založená na metadátach ako autor, názov skladby, text a pod. Hudobné metadáta sme na základe doterajšieho delenia definovanom pri získavaní hudobných informácií [19] rozdelili do troch podskupín:

- **Objektívne, získané automaticky.** Medzi takéto anotácie patria rytmus, takt alebo melódia, ktoré sa s určitou presnosťou dajú zo skladby získať automaticky pomocou podrobného rozboru skladby automaticky. Aj v prípade čiastočne nepresných anotácií je toto dobrý základ pre prvú kategorizáciu a bližšie určenie anotácií masami.
- **Objektívne, získané masami a expertmi.** Práve tu patria údaje ako autor, názov skladby, vydavateľ, hudobný štýl a pod. Tieto údaje musia byť vložené človekom, keďže žiaden automatický prístup nedokáže názov skladby alebo albumu získať priamo zo skladby.
- **Subjektívne, získané masami.** Tu patrí najväčšia (a najmenej presná) časť anotácií, ktoré súvisia s náladou skladby, časom kedy je vhodné skladbu počúvať, cieľovou skupinou poslucháčov a pod. Pocit zo skladby sa môže u rôznych ľudí líšiť a nie každá anotácia je

preto vhodná – no existuje väčšinová akceptácia, kedy je aj subjektívna anotácia braná za prípustnú.

Cieľom našej hry je validovať množiny anotácií (textových značiek – „tags“). Hráčovi prehrávame náhodnú skladbu a jeho cieľom je z ponúknutých množín anotácií vybrať tú najvhodnejšiu – jedna z množín obsahuje existujúce anotácie prehrávanej skladby a v ideálnom prípade hráč vyberie práve tú. Podľa hráčových akcií sme následne schopní určiť správnosť anotácií danej množiny. Po výbere má hráč ešte možnosť vo vybranej množine označiť anotáciu, ktorá sa ku skladbe nehodí. Každý z herných prípadov, ktorý môže nastať nám dá dôležité informácie:

- **Hráč vybral množinu obsahujúcu anotácie k prehrávanej skladbe.** Pri tejto možnosti nám hráč svojou voľbou implicitne oznámi, že množina je dostatočne presná na rozoznanie skladby. Po zvolení množiny má hráč ešte možnosť označiť niektorú z anotácií v prípade, že skladbu dostatočne presne neopisuje – pričom ostatné anotácie sú v poriadku.
- **Hráč vybral množinu obsahujúcu anotácie k inej skladbe.** Z tejto akcie vieme povedať, že nami ponúknuté anotácie k skladbe nie sú presné resp. vybraná množina ich opisuje presnejšie. Taktiež môžeme považovať vybranú množinu ako potenciálne anotácie k prehrávanej skladbe – a to bez ohľadu na to, že anotácie pôvodne patrili inej skladbe.

Týmito spôsobmi je možné správne označiť správnosť veľkej časti anotácií, ktoré automatickým prístupom nemôžu byť vytvorené a získavaním masami nie sú vytvorené dostatočne presne.

V kapitole 2 opisujeme princíp hier s účelom, existujúce hry a problémy pri ich vytváraní a spôsoby získavania a overovania metadát, ktoré sa doposiaľ využili. V kapitole 3 sa zaoberáme získavaním metadát z hudby – teda konkrétnou oblasťou výskumu – a problémami ako sú nízka úspešnosť automatických prístupov práve pri hudobných objektoch alebo využitím hudobných metadát na vzdelávacie účely. V kapitole 4 predstavujeme návrh aj realizáciu hry na overenie hudobných metadát, opisujeme problémy hry a diskutujeme vplyv hráčových akcií na hodnotenie správnosti hudobných anotácií. Kapitola 5 opisuje experiment, ktorý sme s hrou uskutočnili, predstavuje výsledky a problémy, na ktoré sme počas experimentu narazili. V kapitole 6 následne diskutujeme kvalitu výstupu hry, objavené problémy a ich možné riešenia.

2 Získavanie metadát pomocou hier s účelom

2.1 Hry s účelom

Hry s účelom ťažia z predpokladu, že každý deň sa obrovské množstvo ľudí na webe hrá rôzne hry a títo ľudia majú pre to rôzne motivácie: zábava, jednoduchá dostupnosť, sociálna interakcia, vzdelanie. Poskytnutím správnej motivácie časti hráčov sme schopní získať týchto hráčov pre náš účel. Hry boli vždy dobrým motivátorom, preto sa už dlhú dobu využívajú vo výučbe, tréningu a simuláciách (tzv. „*serious games*“) a to v širokom spektre oblastí: školstvo, zdravotníctvo, armáda, obchod a i. Hra s účelom je typ hry, kedy hráč (vedome alebo nevedome) pomáha riešiť problémy, ktoré počítače v dnešnej dobe dokážu vyriešiť len do určitej miery v závislosti od problému [16] – napr. opísať obsah obrázka, kde počítač vie zistiť iba jeho rozmery a použité farby, ale nedokáže presne určiť jeho objekty a prípadné väzby medzi objektmi.

Prvé hry s účelom vznikli iba pred niekoľkými rokmi a priekopníkom v tejto oblasti je Luis von Ahn, ktorému sa podarilo vytvoriť niekoľko úspešných projektov a zaviedol pojem „*hra s účelom*“ (angl. GWAP - Game With a Purpose). Najznámejšou hrou s účelom je *ESP Game*, v ktorej sa dvaja náhodne spárovaní hráči pokúšali zhodnúť na najvhodnejšom slove opisujúcom hrou ponúknutý obrázok – keďže jediná nápoveda pre oboch hráčov bol len spoločný obrázok, hádaním slov, ktoré ho čo najpresnejšie opisovali tak vytvárali jeho potenciálne anotácie. Veľká časť hier s účelom je vytvorená pre viac hráčov hrajúcich súčasne a spárovaných náhodne (kvôli väčšej zábave a súčasnej validácii anotácií).

Avšak práve kvôli dizajnu pre viac hráčov môže po spustení hry nastať „problém studeného štartu“ („*cold start problem*“, veľký problém väčšiny hier s účelom), kedy hra ešte nemá dostatočný počet stabilných hráčov, ktorých by bolo možné spárovať a nemá ani dostatočný počet odohratých hier, kedy by systém dokázal simulovať protihráča [10]. Von Ahn čiastočne rieši problém studeného štartu spoločným portálom pre všetky hry, kde využíva úspech niektorých hier na získanie hráčov pre nové hry. Ďalším spôsobom riešenia tohto problému je hranie simulátormi, ktoré sa tvária ako hráči. Aby boli simulátory čo najpresnejšie, hry priebežne ukladajú priebeh kola hraného živým hráčom a následne na jeho základe dokážu hru hrať čo najpresnejšie a najvierohodnejšie.

Po úspechu von Ahna vznikali ďalšie hry, ktoré sa snažili získavať anotácie rôznymi spôsobmi, ale väčšina z nich nedosiahla potrebnú popularitu. Chybou veľkého množstva hier je nesprávna integrácia samotného účelu do rozhrania hry, čo spôsobí neatraktivnosť hry samotnej a jej nasledovný neúspech. Základnými prvkami úspechu je podľa von Ahna [1] využitie nasledovných prvkov:

- **Zvýšenie hráčovho potešenia z hry.**
 - **Zavedenie časového limitu.** Časový limit vytvára v hráčovi pocit výzvy, ktorú musí splniť. V prípade, že úlohu splní načas, obdrží druh bonusu - dobre naformulované ciele, ktoré sú zároveň výzvami zvyčajne vedú k lepším výkonom.
 - **Skóre.** Skóre je priama odozva hráčovi o jeho úspešnosti. Sumár skóre po každej hre ukáže hráčovi, v ktorých aspektoch je dobrý a v ktorých sa musí zlepšiť.
 - **Hodnosti.** Zavedenie hodností odlišuje hráčov z globálneho pohľadu na hru a motivuje dosiahnuť vyššiu hodnotu ako aktuálnu a odlíšiť sa. Je vhodné zobrazit' hráčovi, koľko bodov potrebuje na získanie ďalšej hodnoty.
 - **Tabuľka najvyšších skóre.** Ďalší z prístupov motivácie hráča podať v hre čo najlepší výkon – úroveň obtiažnosti dostať sa do tabuľky závisí od časovej periódy hier, z ktorých je tabuľka vytvorená. Využitie hodností ako atribútu zoradenia tabuľky môže byť vhodným motivačným prvkom pre hráčov – tí nevedia presný počet bodov potrebný na dosiahnutie lepšej pozície, avšak tušia bodový interval.
 - **Náhodnosť.** Hráči sú párovaní úplne náhodne, čo spolu s náhodnosťou vstupov prináša do hry nový rozmer, ktorý je výzvou ako pre začiatočníkov, tak pre expertov. Náhodnosť rovnako zvyšuje diverzitu jednotlivých hier.
- **Presnosť výstupu.**
 - **Náhodné testovanie hráčov.** Hráči sú v priebehu hry testovaní na už overené anotácie, ktoré by mali byť schopní označiť ako správne.
 - **Overenie výstupu.** Výstup z hry by nemal byť platný hneď. Dáta by sa mali zbierať a až keď anotáciu vytvorilo dostatočný počet hráčov, mala by byť označená za správnu. Taktiež je vhodné výstup validovať už počas hry, čo je možné správnym navrhnutím hry – hráči sa musia zhodnúť na výstupe aby získali body.
 - **Vytvorenie tabu anotácií.** Často opakujúce sa anotácie by mali byť vylúčené a blokové. Avšak v tomto prípade môže nastať problém ovplyvnením hráča samotným zoznamom zakázaných slov (kedy si vytvorí asociácie, ktoré by ho bez viditeľného zoznamu nenapadli) - prípadne vzniká možnosť výstupov vytvorených automatizovaných botom čisto na základe sprístupneného zoznamu ako bolo preukázané v [21].
- **Ostatné rady.**
 - **Nahraté hry.** Aby párovanie hráčov pri hrách vyžadujúcich zapojenie viac hráčov súčasne netrvalo dlho, je vhodné v prípade nemožnosti nájsť protihráča nasadiť, ktorý bude opakovať už predtým odohratú hru. V prípade správneho prevedenia si hráč nesmie uvedomiť, že nehrá proti skutočnému protivníkovi.
 - **Viac než dvaja hráči.** Čím viac hráčov hru hrá, tým súťaživejší hráči sú a dosahujú lepšie výsledky a výstupy.

Hry s účelom musia motivovať hráčov k hraniu, ale zároveň dbať na správnosť údajov, ktoré sa hraním získajú. Aby toto bolo možné, von Ahn zaviedol niekoľko modelov hier, pri ktorých sa určil cieľ, ktorý sa hráči snažia dosiahnuť. [1]:

- **Hry dosahujúce zhodu na výstupe.** V tomto type hier dostanú hráči rovnaký vstup (slovo, obrázok, pieseň..) a aby postúpili ďalej, musia sa zhodnúť na tom, čo po nich hra žiada – väčšinou priamo anotácie. Hry väčšinou končia pri prvej zhode hráčov. Okrem klasických hier, kedy sa dvaja hráči snažia nájsť zhodu boli predstavená aj hra KissKissBan [6], kde bol zavedený tretí hráč, ktorého úloha nebola spolupracovať, ale zabrániť prvým dvom hráčom nájsť zhodu. Rozširovanie existujúcich modelov správnym spôsobom môže zvyšovať zábavnosť hry a zároveň množstvo údajov získaných z jednej hry.
- **Hry s inverznými úlohami.** Pri týchto hrách majú hráči rozdielne úlohy, ktoré si navzájom striedajú. Úlohou prvého hráča je na základe vstupu vyprodukovať výstup (potenciálne anotácie) a úlohou druhého hráča je na základe poskytnutého výstupu uhádnuť, ktorý z jemu poskytnutých vstupov je ten správny. Tento druh hier motivuje hráčov ku konkrétnejším opisom vstupu, pretože neúplný opis by nebol dostatočný pre druhého hráča aby správne vybral jednu z možností.
- **Hry dosahujúce zhodu na vstupe.** Cieľom týchto hier je určiť, či hráči dostali na vstupe rovnaký alebo rozdielny objekt. Hráči sú nútení opisovať svoj vstup druhému hráčovi v dostatočnej miere na to, aby z oboch opisov mohli usúdiť zhodnosť vstupov.

Pri hrách s účelom môže dochádzať k anotáciám rôzneho obsahu: textov, obrázkov, multimédií, vytváraní väzieb medzi slovami, overovaní faktov. Keďže navrhnúť hru, ktorá by dokázala získavať údaje o všetkých spomenutých druhoch obsahu je náročné, každá hra sa sústreďí len na jeden konkrétny aspekt.

2.2 Prehľad existujúcich hier

2.2.1 Hry s účelmi z rôznych oblastí

2.2.1.1 ESP Game¹

ESP Game bola vôbec prvou hrou s účelom [2]. Hra bola vytvorená pre dvoch náhodne spárovaných hráčov, ktorých úlohou bolo zhodnúť sa na tom, čo sa nachádza na obrázku. Aby sa zabránilo opakovanému používaniu rovnakých slov rôznymi hráčmi, boli zavedené tabu slová, ktoré hráči nemôžu použiť. Ak sa chcú hráči zhodnúť a získať body, sú nútení opisovať práve to, čo sa nachádza na obrázku. Priebeh hry je zobrazený na obr. 2.1.

Hru v roku 2006 licencovala spoločnosť Google a v auguste 2006 ju spustila ako Google Image Labeler. Služba pomáhala anotovať obrázky indexované vyhľadávačom. V septembri 2011 Google službu zrušil. Pôvodná hra ESP Game funguje naďalej.

¹ <http://www.gwap.com/gwap/gamesPreview/espgame/>

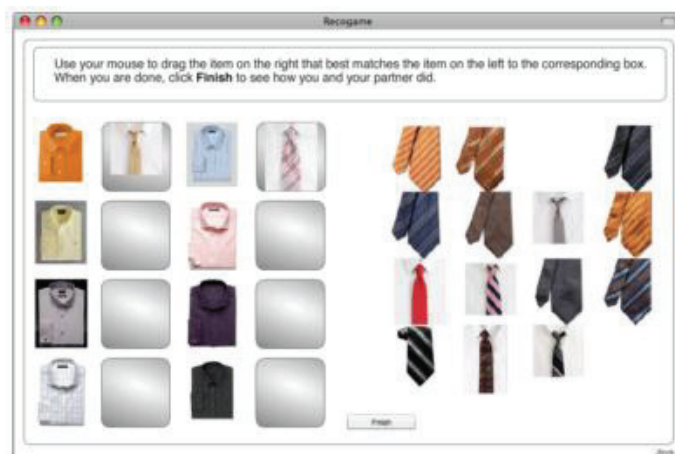


obr. 2.1 Priebeh hry ESP Game

2.2.1.2 Curator

Curator je hra, ktorá sa snaží v kolekcií objektov vyberať najvhodnejšie súvisiace objekty [20]. Na rozdiel od klasického prístupu, kedy by boli hráčovi ponúkané dvojice a hráč by mal rozhodnúť, či je dvojica vhodná alebo nie – takýto prístup by vyžadoval milióny kôl a stále by to nestačilo - ponúka hra množinu objektov a nechá hráča rozhodnúť, ktoré sa k sebe najviac hodia. Na obr. 2.2 je zobrazený priebeh hry, kedy hráč vyberá najvhodnejšie dvojice košele - kravata.

Hra je určená pre dvoch hráčov, ktorých úlohou je zhodnúť sa na čo najväčšom počte zvolených kombinácií. Aby boli hráči úspešní, nesmú uprednostňovať iba svoje preferencie, ale aj preferencie – náhodne spárovaného – spoluhráča, ktoré objavujú po každej hre keď hráči vidia voľby svojho spoluhráča. Ak sa dvojica objavila v niekoľkých na sebe nezávislých hrách, existuje medzi objektmi väzba. V prípade, že sa objekty spolu nachádzali vo veľkom množstve hier a neboli spolu nikdy spárované, je možné predpokladať, že objekty spolu nesúvisia vôbec alebo iba minimálne. Hra v súčasnosti nie je verejne dostupná.

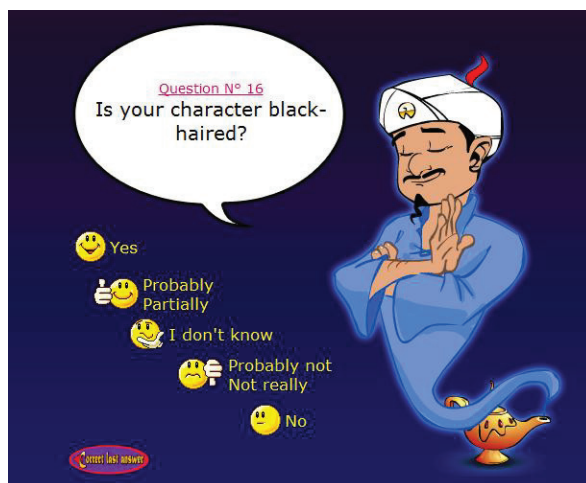


obr. 2.2 Priebeh hry Curator

2.2.1.3 Akinator²

Akinator je jediný komerčný projekt spomenutý v tejto práci a doposiaľ o ňom nebola napísaná žiadna odborná práca, aj keď v oblasti hier s účelom je to jeden z najúspešnejších projektov. Hra sa stala úspešnou hlavne vďaka svojej jednoduchosti. Je určená pre jedného hráča a bola vytvorená aj verzia pre mobilné zariadenia. Úlohou hráča v tomto prípade nie je rozmyšľať nad anotáciami, iba správne odpovedať na otázky, ktoré mu hra kladie. Hra sa na základe jeho odpovedí a odpovedí iných hráčov z predchádzajúcich hier snaží uhádnuť, na koho hráč myslí.

Hráč si na začiatku hry vyberie ľubovoľnú postavu (reálnu alebo fiktívnu) a jeho úlohou je len pravdivo odpovedať na otázky s výberom odpovede – tie sú vždy rovnaké a sú zobrazené na obr. 2.3. Odpoveďami si hra selektuje množinu vhodných možností a v prípade, že je pravdepodobnosť výberu dostatočne vysoká, prezentuje hráčovi svoj „tip“. Ak po ani 40 otázkach hra neuhádne, na koho hráč myslel, požiada hráča o zadanie mena, na ktoré myslel príp. vloženie otázok, ktoré s týmto menom súvisia. V hre nie je použité žiadne bodovanie a jedinou (a zjavne dostatočnou) odmenou hráčovi je prekvapenie zo správneho tipu hry.



obr. 2.3 Priebeh hry Akinator

Hra údaje získava výlučne zo vstupov hráčov a na overenie využíva spätnú väzbu – moment, kedy prezentuje hráčovi svoj tip. V prípade správneho tipu si uloží vzniknuté väzby na základe hráčových odpovedí a použije ich v nasledujúcich hrách. Po hre je taktiež možné zobrazit' si jej detail a vidieť, aké odpovede boli očakávané počas priebehu hry.

2.2.1.4 Little Search Game³

Cieľom Little Search Game [13, 14] je získať väzby medzi pojmami, ktoré sú podobné vzťahom vo folksonómiách a to za pomoci vyhľadávacích dopytov.

Úlohou hráča je upravovať vyhľadávací dopyt tak, aby získal čo najmenej výsledkov. Hra na začiatku hry ponúkne hráčovi slovo, ktoré má osamote vo vyhľadávači určitý počet výsledkov.

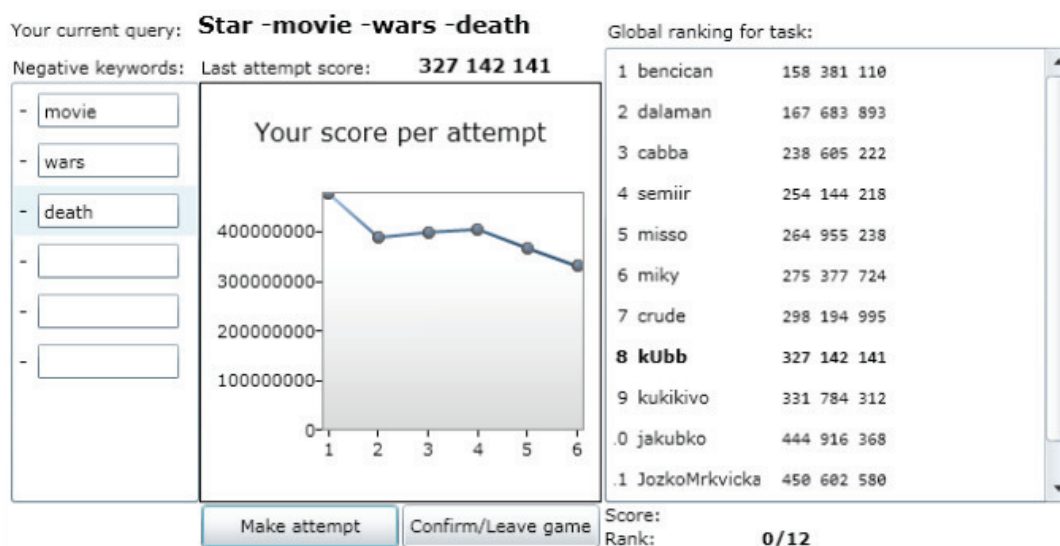
² <http://www.akinator.com>

³ <http://mirai.fiit.stuba.sk/LittleGoogleGame>

Medzi rozšírenú funkcionálnu vyhľadávačov patrí aj odstránenie výsledkov vyhľadávania, ktoré obsahujú iné slovo (napr. apple –fruit –big zobrazí výsledky obsahujúce slovo „apple“ a neobsahujúce slová „fruit“ a „big“). Hráč musí určiť maximálne 6 takýchto slov a získať tak čo najmenší počet výsledkov.

Keďže je hra iba pre jedného hráča, akákoľvek validácia získaných väzieb prebieha asynchrónne na základe početnosti vloženia danej väzby – hra tým pádom nepodlieha problému studeného štartu. Vďaka tomu, že hráči zadávajú rôznorodé slová a každý hráč dokáže k zadanému slovu vytvoriť rozličné asociácie, vytvárajú sa tak väzby, ktoré by automatizovaný prístup nedokázal rozoznať. Do úvahy sú brané aj väzby, ktoré nemali dosah na zníženie počtu výsledkov – za predpokladu, že hráč sa snaží zadávať len slová, ktoré nejakým spôsobom asociujú so zadaným.

Na základe prieskumu bolo zistené, že hra dokázala určiť 91% väzieb správne, čo je rádovo lepšie, než automatizované prístupy. Priebeh samotnej hry je zobrazený na obr. 2.4.



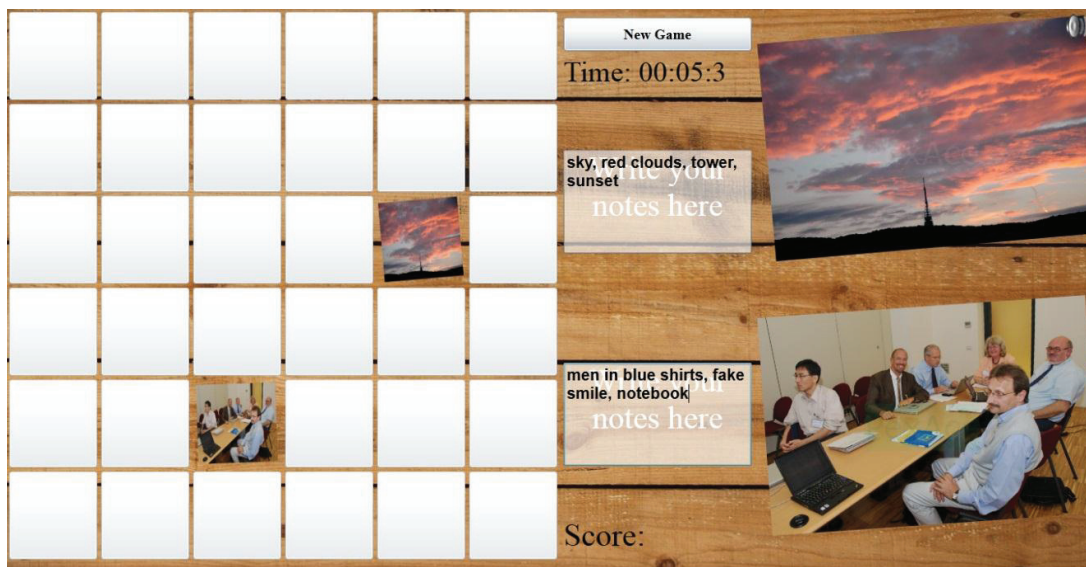
obr. 2.4 Priebeh hry Little Search Game pri zadanom slove „star“

2.2.1.5 PexAce⁴

PexAce je hra získavajúca anotácie k obrázkom pomocou upravenej verzie známej hry pexeso [11, 15], je teda pre jedného hráča. Cieľom hry je nájsť všetky páry obrázkov v balíku. Aby bolo možné získať anotácie, bodovanie hry bolo nastavené tak, aby bol hráč motivovaný otáčať čo najmenej dvojíc kariet a aby jeho výsledný čas bol čo najnižší. Hra umožňuje hráčovi po otočení dvojice písať k týmto dvom kartám poznámky, ktoré sú neskôr prístupné počas celého trvania hry. Počas písania poznámok čas hry stojí. Poznámky mu majú pomôcť zistiť, čo sa na danej karte nachádza bez toho, aby ju musel otočiť (a tým tak riskovať stratu bodov). Z týchto poznámok potom hra dokáže generovať anotácie. Hra je zobrazená na obr. 2.5.

⁴ <http://mirai.fiit.stuba.sk/pexace>

Pri hrách generujúcich anotácie k obrázkom je možné správnym zvolením množiny obrázkov donútiť hráča k presnejším anotáciám [17]. Tento prístup je pre hráčov prijateľnejší ako generovanie tabu slov. V prípade, že sú obrázky v množine z rôznych domén, stačila by hráčovi jednoduchá anotácia (napr. „kvet“), ale ak by bola množina obrázkov tvorená z obrázkov rovnakej domény (napr. zo samých kvetov), hráč by bol nútený zadať špecifickejšie charakteristiky („modrý“, „horec“ a pod.).



obr. 2.5 Priebeh hry PexAce; hracie pole (vľavo), poznámky (v strede) a zväčšený náhľad (vpravo)

2.2.2 Hry s účelom anotovať hudbu

2.2.2.1 Major Miner⁵

MajorMiner [9] je hra, ktorej hlavným cieľom bolo získať čo najväčšie množstvo a čo najpresnejšie anotácie k vybraným skladbám. Jej druhým cieľom bola zábavnosť hry tak pre nováčikov, ako aj pre skúsených hráčov. Na rozdiel od iných hudobných hier upustili od párovania hráčov a navrhli hru pre jedného hráča, kde sú anotácie validované asynchrónne – čím vyriešili aj problém studeného štartu.

Hlavnou proklamovanou devízou (čo sa neskôr ukázalo ako najväčší problém) má byť bodovanie hráčov. Aby sa docielil čo najväčší počet presných anotácií, bolo navrhnuté nasledovné bodovanie:

- 2 body pre hráča, ktorý vložil doposiaľ nevloženú anotáciu. Body nie sú pripočítané hneď, ale až po validovaní tejto anotácie iným hráčom.
- 1 bod pre hráča, ktorý zadal už predtým práve raz zadanú anotáciu. Súčasne s vložením sa anotácia zároveň validuje (to znamená, že stačia dvaja hráči, aby bola anotácia validná). Bod sa hráčovi pripočíta ihneď a rovnako sa pripočítajú aj 2 body hráčovi, ktorý anotáciu zadal ako prvý a body mu doposiaľ neboli pripočítané.

⁵ <http://www.majorminer.org>

- 0 bodov za každú vloženú anotáciu, ktorá už bola predtým validovaná.

Tento model bodovania mal motivovať hráčov zadávať nové anotácie pričom stále boli ohrození nutnosťou vloženia rovnakej anotácie ešte jedným hráčom – čo zabráňovalo náhodným anotáciám. Model ale demotivoval hráčov, keďže body sa dali ľahko získať len pri novo zaradených skladbách. Ďalším problémom bolo neatraktívne a zle navrhnuté prostredie, ktoré nemuselo byť rýchlo pochopiteľné a nebolo vždy dostatočne prehľadné. Priebeh hry je zobrazený na obr. 2.6.



obr. 2.6 Hlavná obrazovka hry MajorMiner

2.2.2.2 HerdIt⁶

HerdIt [3] je projekt, ktorý bol vyvinutý a implementovaný na sociálnej sieti Facebook⁷ a spomedzi všetkých hier s účelom najviac využíva možnosti a výhody sociálnych sietí.

Pri prvom vstupe do hry je hráč nasmerovaný priamo na začiatok bežiaccej hry, bez žiadneho nastavovania. Hráč je náhodne spárovaný s 9 hráčmi (prípadne so simulátorom hrajúcim predtým nahraté hry), pričom cieľom hry je získať čo najviac bodov a medzi ostatnými sa v rebríčku umiestniť čo najvyššie. Hráčom sa prehráva 10-20 sekúnd dlhá hudobná ukážka a ku každej ukážke dostanú inú úlohu, ako napr. určiť farbu hudby, určiť v dvojdimenzionálnom priestore chytľavosť skladby a čas dňa vhodný na prehrávanie skladby, či by hráč išiel na koncert kde hrať daná skladba a pod. Po dohraní skladby je všetkým hráčom položená otázka, či skladbu poznajú a na výber sú im ponúknuté 4 možnosti. Po dohraní je možné prihlásiť sa a vybrať si hudobnú doménovú oblasť, ktorá hráča zaujíma – hudobné ukážky budú len z tejto domény. Všetky nahraté body (pre každú doménovú oblasť osobitne) sa ukladajú a zvyšujú hráčovi jeho úroveň. Priebeh hry je zobrazený na obr. 2.7.

Bodovanie je postavené na základe zhody s ostatnými hráčmi. To znamená, že ak bola úloha vybrať farbu hudby a 5 hráči vybrali čiernu, 3 žltú a 2 modrú, tak 5 hráči dostanú najviac (každý po 50 bodov), 3 dostanú po 30 bodov a 2 po 20 bodov. Pri ostatných úlohách sa postupuje

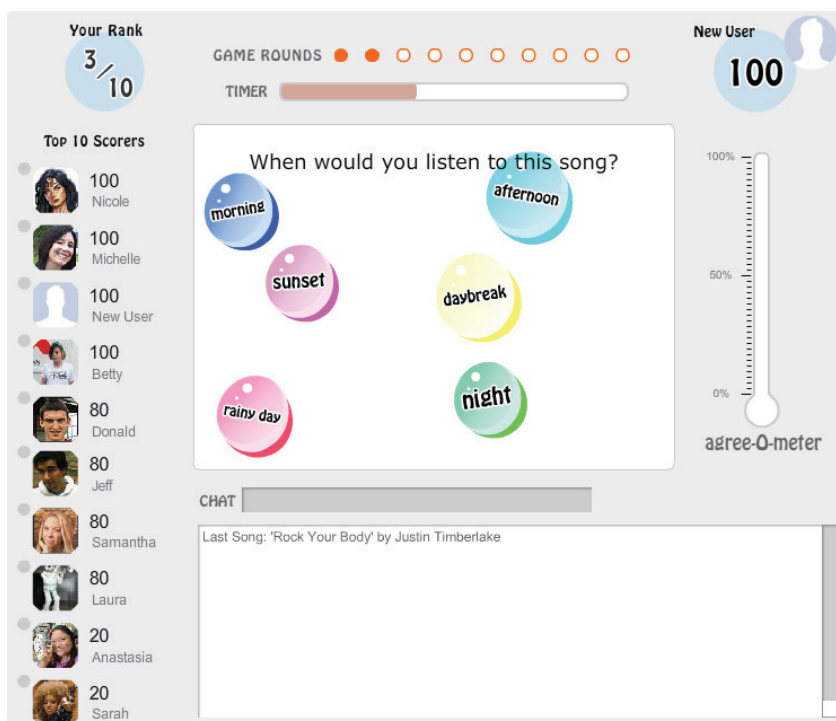
⁶ <http://apps.facebook.com/herd-it/>

⁷ <http://www.facebook.com>

analogicky (niekde s menšími potrebnými úpravami ako pri dvojdimenzionálnom priestore, kedy sa do úvahy správnosti berie rádius zvoleného bodu). Po dokončení úlohy a správnom odpovedaní na otázku práve hrajúcej skladby je hráč odmenený bonusovými 20 bodmi.

Hra je postavená na výbere ponúkaných možností, samotní hráči žiadne anotácie negenerujú. Cieľom hry bolo teda zaradiť skladby do vopred určených kategórií, ktoré je možné v budúcnosti použiť – toto zaradenie je presnejšie ako v prípade, kedy by sa o to snažili samotní hráči bez výberu možností. Overenie hry za pomoci prieskumu počas beta testovania⁸ prinieslo niekoľko zaujímavých výsledkov:

- Až 92,9% hráčov si uvedomovalo (úplne alebo čiastočné), že hrajú proti iným hráčom.
- Až 92,9% by sa pokúsilo (rozhodne alebo pravdepodobne) prekonať priateľove najvyššie skóre, ak by sa objavilo v profile priateľa.
- Nikto z hráčov v prieskume neuviedol, že by hra bola zlá alebo otrasná.
- 88,8% hráčov by hru ďalej odporučilo (rozhodne alebo pravdepodobne).



obr. 2.7 Hlavná obrazovka hry HerdIt

2.2.2.3 Tagatune⁹

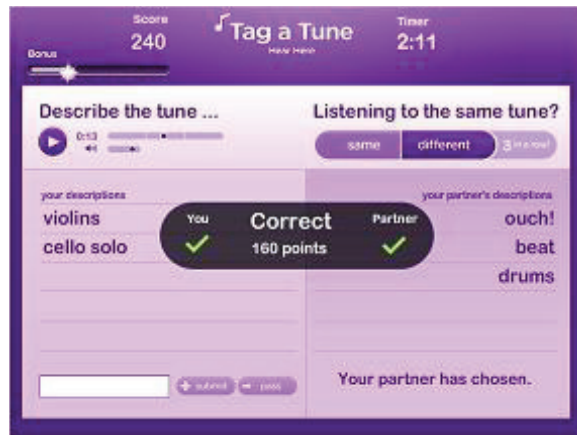
Tagatune je ďalšou hrou zo série hier od Luisa von Ahna [7]. Podobne ako pri ESP Game aj tu sa hráči musia zhodnúť, aby získali body. Avšak s malou, ale podstatnou zmenou.

⁸ Výpovedná hodnota výsledkov nemusí zodpovedať realite vzhľadom na nízku vzorku zúčastnených

⁹ <http://www.gwap.com/tagatune/>

Náhodne spárovanej dvojici hráčov je prezentovaná hudobná ukážka a oni musia rozhodnúť, či obaja počúvajú rovnakú ukážku. Správnosť rozhodnutia docielia pomocou spoločného chatu, do ktorého musia zadávať kľúčové slová – budúce anotácie – o práve počúvanej skladbe. V prípade, že majú dostatok informácií od spoluhráča, zvolia jednu z dvoch možností: rovnaká skladba / rozdielna skladba. Ak sa zhodli, z ich komunikácie sa stávajú potenciálne anotácie a hráči sú ohodnotení bodmi. Hra trvá časovo obmedzenú dobu a počas nej musia hráči stihnúť takto zhodnotiť čo najviac skladieb.

V prípade, že sa hráči nevedia dohodnúť v dostatočne krátkej dobe, majú možnosť preskočiť na ďalšiu skladbu. Na obr. 2.8 je zobrazená hra, kedy sa hráči zhodli, že počúvajú rozdielnú skladbu.



obr. 2.8 Hlavná obrazovka hry Tagatune

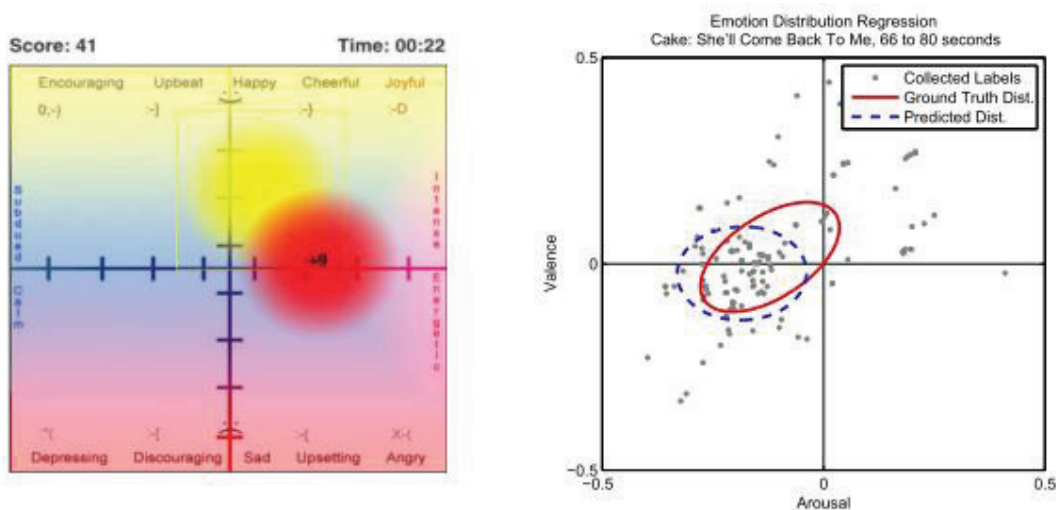
2.2.2.4 Moodswings¹⁰

Moodswings je hra, ktorej jediným cieľom je zbieranie informácií o nálade prehrávanej skladby [10]. Celá hra prebieha na dvojrozmernej ploche, kde na jednej osi je naznačená pozitívna/negatívna emócia a na druhej je naznačené nadšenie zo skladby. Oproti iným hrám získava Moodswings informácie počas celého prehrávania skladby, takže je schopná rozlišovať rôzne nálady v jednej skladbe.

Hra náhodne spáruje dvoch hráčov, ktorým pustí rovnakú skladbu. Úlohou hráčov je na dvojrozmernej ploche kurzorom určiť bod, ktorý zodpovedá aktuálnej nálade. Okolo kurzoru je vytvorený rádius, ktorý sleduje prekryv s rádiusom spoluhráča. Čím väčší tento prekryv je, tým viac bodov hráči získavajú. Hráči vidia umiestnenie kurzoru spoluhráča iba na krátku dobu raz za 3 sekundy, čo ich má naviesť ku korekciám a vyšším bodovým ziskom. Hráči obdržia bonusové body v prípade, že sa určitú dobu na ploche nehýbu. Jedna hra obsahuje hodnotenie piatich 20-30 sekundových ukážok. Priebeh hry je zobrazený na obr. 2.9 vľavo – farebné kruhy reprezentujú aktuálne pozície hráčov.

¹⁰ <http://schubert.ece.drexel.edu/moodswings>

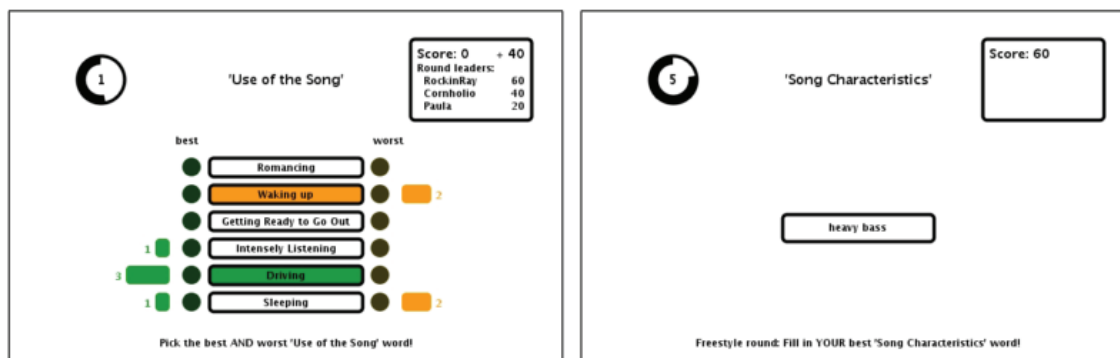
Zo všetkých získaných bodov v grafe sú analýzou zistené vrcholy, ktoré sa najviac prekrývali medzi oboma hráčmi, ktoré sú zoskupené v jednej oblasti a teda reálne zodpovedajú nálade pesničky – zobrazené na obr. 2.9 vpravo.



obr. 2.9 Priebeh hry Moodswings (vľavo) a analýza hrou získaných údajov (vpravo)

2.2.2.5 Listen Game

ListenGame je jednoduchá hra pre jedného hráča, ktorá pozostáva z dvoch častí [18]. V prvej hráči na základe prehrávanej ukážky vyberajú najviac a najmenej sediacu charakteristiku. V druhej časti sú hráči vyzvaní vložiť charakteristiku ku skladbe – táto charakteristika bude potom použitá u iných hráčov ako jedna z možností v prvej časti. Počas jedného kola dostane hráč 7 skladieb s výberom možnosti a 1 skladbu so zadáním charakteristiky. Problém studeného štartu hra rieši náhodným výberom slov z datasetu CAL500. Hra v súčasnosti nie je verejne dostupná. Priebeh hry je zobrazený na obr. 2.10.



obr. 2.10 Priebeh hry Listen Game; výber z možností (vľavo) a zadanie charakteristiky (vpravo)

3 Získavanie metadát k hudbe

3.1 Dolovanie hudobných znalostí

Donedávna boli ľudia pri objavovaní nových piesní odkázaní na rádia, televízie a obchody s hudobnými nosičmi. Typicky človek pri vyhľadávaní pozná úsek skladby (či už text alebo melódiu) a od tohto zakladá svoje vyhľadávanie. S rozmachom on-line hudobného priemyslu a vzniku obrovských databáz sa objavili nové možnosti, ako hudbu vyhľadávať: pomocou zvuku, textu piesne, žánru, štýlu alebo jej nálady [12]. Avšak tento spôsob nemusí byť nutne efektívny, keďže podobnosť opisov skladieb nemusí implikovať podobnosť samotných skladieb a rovnako rozdielne opisy nemusia automaticky implikovať štýlovo rozdielne skladby. Takýto prístup funguje pri menších databázach, avšak pri veľkých (viac ako 100 tisíc skladieb), ktoré sú spravované veľkým množstvom ľudí to fungovať prestáva.

Vo svete MIR („Music Information Retrieval“ – získavanie hudobných informácií) stále neexistuje automatizovaný spôsob, ako získať úplne správne metadáta. Medzi ťažko získateľné údaje patria podľa [4]:

- **Farba tónu.** Rozoznanie nástroja v hudobnom telese.
- **Melódia.** Získanie melodickej a basovej linky.
- **Rytmus.** Aktuálny rytmus, priemerné tempo, hudobný takt.
- **Výška tónu.**
- **Harmónia.** Extrakcia akordov.
- **Tónina.** Sledovanie modulácií do rôznych tónin.
- **Text.** Identifikácia spevu a rozpoznávanie slov.
- **Nezápadná hudba.** Pochopenie konceptov východnej hudby.

Výskum ukázal, že automatické rozpoznávacie systémy majú pri rôznych úlohách rôzne úspešnosti, avšak všetky úspešnosti sú na v praxi nepoužiteľnej úrovni. Preto je potrebné využívať aj iné spôsoby získavania anotácií v kombinácii s automatickými prístupmi. V Tabuľka 1 je zobrazená úspešnosť automatického rozpoznávania rôznych atribútov hudobných skladieb.

Tabuľka 1 Úspešnosť rozpoznávania hudobných atribútov pomocou automatických systémov

Nálada	61,50%
Skladateľ klasickej hudby	53,72%
„Cover-song“ (prespievaná skladba)	52,00%
Autor	48,14%

Cieľ, ktorý chceme dosiahnuť je automatická anotácia akýchkoľvek skladieb. Takáto anotácia je dnes čiastočne využívaná na základe správania používateľov na konkrétnom webe – pri rôznych ľuďoch, ktorí si nezávisle na sebe vypočuli rovnaké skladby je možné predpokladať, že sú skladby podobné (aj na základe už existujúcich anotácií) a anotácie jednej skladby je možné použiť aj pri ostatných. Vďaka obrovským databázam je už získané veľké množstvo údajov

založených na správaní používateľov, ktoré pomáhajú uľahčiť kategorizáciu pri populárnych skladbách, avšak alternatívne skladby sú v tomto smere stále pozadu.

Okrem vyhľadávania presnej skladby je bežnou motiváciou človeka vytváranie zoznamov skladieb („*playlistov*“), objavovanie nových – podobných skladieb prípadne iba organizácia veľkých hudobných kolekcii. Z tohto dôvodu vznikajú rôzne služby, ktoré majú vyššie spomenuté ciele uľahčiť. Pri vyhľadávaní je potrebné rozlišovať medzi dvoma druhmi dopytov [8]:

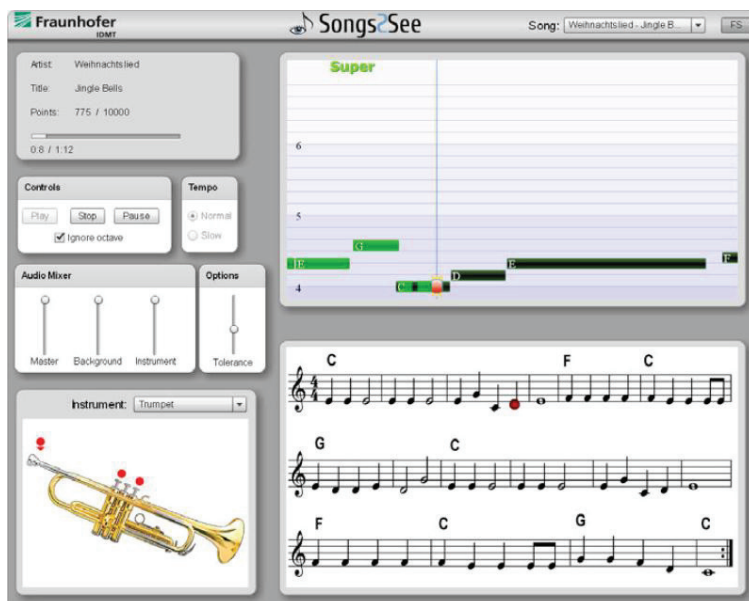
- **Exaktný dopyt.** Tento druh sa využíva najmä v relačných databázach a je veľmi jednoduchý na implementáciu. Vyhľadávanie prebieha na základe mena skladby alebo skladateľa prípadne iných dostupných informácií o skladbe. Príkladom takejto služby je Yahoo Audio Search¹¹ alebo hudobný obchod iTunes¹². Úspešnosť týchto systémov veľmi závisí od dát o skladbách, ktoré sú dostupné v systéme a bežne sa úspešnosť vyhľadávania pohybuje okolo 40% práve kvôli nedostatku meta-informácií. Túto situáciu sa snažia zmeniť služby LastFM a Pandora, ktoré sa zaoberajú odporúčaním hudby a získavaním dát o hudbe.
- **Dopyt na základe príkladu.** Pri tomto spôsobe vyhľadávania je systému poskytnutá časť skladby – buď reálne nahratá na mobilné zariadenie alebo zaspievaná používateľom. Systém sa v dostatočne rozumnom čase snaží nájsť podobnosť medzi skladbami v jeho databáze. Problémom pri tomto spôsobe môže byť priveľký šum v nahrávke alebo nepresný pokus človeka napodobniť skladbu. Navyše reprezentácia skladieb tak, aby ich bolo možné porovnávať je kapacitne a výpočtovo náročná.

3.1.1 Songs2See

Na Fraunhoferovom inštitúte sa pokúšali o projekty zamerané na získavanie hudobných dát a ich reálne využitie v praxi [5]. Prvým bola webová/desktopová aplikácia „*Songs2See*“, ktorá mala pomáhať študentom umeleckých škôl s ich hudobnými cvičeniami. Výhodou oproti iným, existujúcim projektom mala byť možnosť použiť vlastné, populárne skladby a viac tak motivovať študentov. Tieto skladby mali byť automaticky analyzované a rozpísané do notových zápisov jednotlivých partov. Úspešnosť hráča bola meraná presnosťou hry študenta oproti originálu. Prípadné chyby, ktoré mohol automat vytvoriť pri generácii dát mohli používatelia ručne opraviť. Na výpočet prepisu bolo použitých niekoľko kombinovaných algoritmov a výsledok bol uspokojivý, avšak získanie prepisu bolo výpočtovo príliš náročné. Samotná aplikácia je zobrazená na obr. 3.1.

¹¹ <http://new.music.yahoo.com>

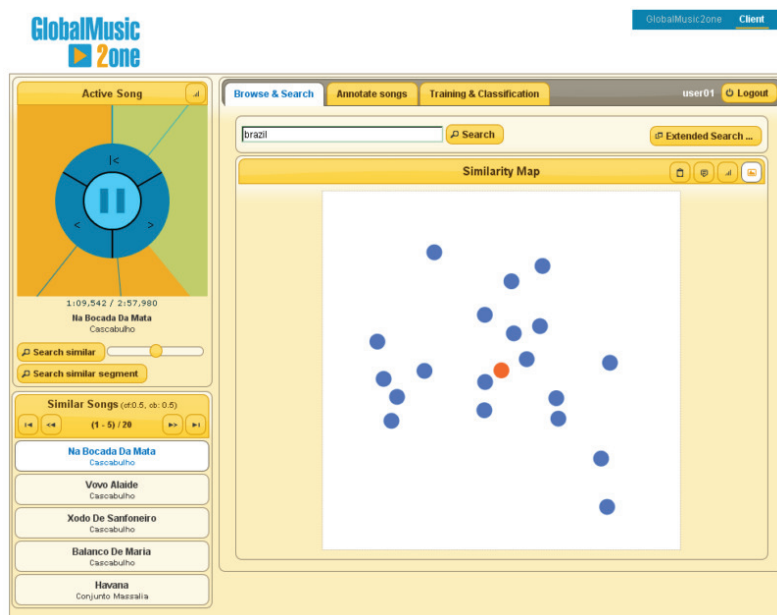
¹² <http://www.apple.com/itunes>



obr. 3.1 Výučbová aplikácia „Songs2See“

3.1.2 GlobalMusic2One

Druhou aplikáciou vyvíjanou na Fraunhoferovom inštitúte bola „*GlobalMusic2One*“ – sémantický hudobný vyhľadávač. Vytváranie hudobných anotácií z veľkej časti závisí na vstupoch od expertných používateľov. Preto vyvinuli anotačný nástroj, ktorý mal týmto používateľom napomáhať. Nástroj dokázal na rôznych úrovniach granularity na základe rytmu, tonality a zafarbenia skladby rozlíšiť jednotlivé časti pesničky a navrhnúť používateľovi notácie. Samotný vyhľadávač je zobrazený na obr. 3.2. Fungoval na základe zadanej anotácie a prehrával skladby čo najbližšie zadanému výrazu – skladby sú zobrazené v 2D priestore aby bola zvýraznená ich hudobná vzdialenosť od zadaného výrazu.



obr. 3.2 Sémantický hudobný vyhľadávač „GlobalMusic2One“

3.2 Služby zamerané na získavanie hudobných anotácií

Medzi webovými službami poskytujúcimi najväčšiu verejnú databázu hudobných metadát sú najznámejšie Pandora a LastFM. Tieto služby vystupujú ako on-line rádiá, kde používateľ zadá obľúbeného interpreta a služba mu spustí personalizované rádio. V súčasnosti nie je ani jedna služba na Slovensku prístupná zdarma. Ich nedostatky sa snaží dopĺňať open-sourcová databáza MusicBrainz. V nasledujúcej časti bude opísaný spôsob získavania anotácií jednotlivými službami a nevýhody s tým spojené.

3.2.1 Pandora

Pandora vznikla v roku 2005 a od začiatku sa vybrala cestou expertného získavania anotácií – má zaplatený tím hudobne vzdelaných ľudí, ktorí každý deň ručne vkladajú informácie o vypočutých skladbách. Používatelia tak majú zaručené presné zaradenie skladieb a kvalitu ich personalizovaného rádia. Avšak tento spôsob získavania údajov je nákladný (tím hudobníkov takto pracuje na plný úväzok) a financovanie z reklamy by hudobníkov a zmluvy s nahrávacími spoločnosťami nepokrylo. V súčasnosti je služba prístupná len občanom USA a to zadarmo, pre ostatné krajiny je Pandora zavretá (príjmy z reklamy v USA stačia na pokrytie nákladov na jej lokálnu prevádzku). Uzavretím pre USA Pandora zrušila verejný prístup aj k svojmu API a teda vytvoreným anotáciám.

Služba svoje anotácie rozširovala aj pomocou cieľených dotazníkov, kde dala masám na výber z niekoľkých možností (čím zabránili vkladaniu nesprávnych anotácií) a používatelia vybrali najviac hodiacu sa anotáciu. Tento spôsob bol len doplnkom platených expertných používateľov.

Používateľom ešte dodatočne ponúka možnosť explicitnej spätnej väzby, kedy môžu kliknutím na „Páči sa mi“ alebo „Nepáči sa mi“ ohodnotiť skladbu, ktorú im služba odporučila. Túto skladbu ohodnocujú na základe vhodnosti v ich práve prehrávanom personalizovanom rádiu, nie čisto na základe sympatií k skladbe.

3.2.2 LastFM

LastFM vystupuje ako hudobná sociálna sieť, ktorá okrem personalizovaného rádia obsahuje aj informácie o hudobných festivaloch (s pripojenými festivalovými rádiami), informácie o samotných hudobníkoch a pre tých, ktorí si nainštalujú poskytnutý program (nazvaný „*scrobbler*“) sa vytvárajú aj štatistiky počúvaných skladieb. Služba je pre občanov USA prístupná zdarma, pre zvyšok sveta je spoplatnená. Všetky anotácie sprístupňuje vývojárom zadarmo pomocou jednoduchého API¹³.

Služba získava svoje anotácie za pomoci nás (*crowdsourcingom*) a to cez už spomínaný *scrobbler*. Ten nainštaluje pluginy, ktoré sledujú hudbu prehrávanú v hudobných prehrávačoch a odosiela štatistické údaje na server. Priamo zo *scrobbleru* je tiež možné spustiť personalizované rádio, ktoré rovnako odosiela získané štatistiky. Na základe týchto štatistík si vie služba vytvoriť väzby medzi skladbami – čím populárnejšia skladba, tým rýchlejšie vytvorená a pevnejšia väzba.

¹³ <http://www.last.fm/api/radio>

Program skladbu zaráta až po určitom percente odohratej skladby, aby tak zabránil nesprávnym väzbám spôsobeným náhodným prehrávaním skladieb.

V scrobbleri je možné zobrazit' anotácie práve prehrávanej skladby a vložit' ďalšie. Používatelia túto službu využívajú na zlepšenie svojho personalizovaného rádia – a tým aj rádií ostatných. Populárne (často vkladané) anotácie sú potom zdieľané medzi skladbami so silnými väzbami. Nevýhodou takto získaných anotácií je šum, ktorý je vkladajú používatelia. Akokoľvek môžu byť anotácie presné, objavujú sa medzi nimi aj nepresné alebo nič nehovoriace anotácie. Tieto potom miestami vytvárajú silnejšie väzby u tých skladieb, ktorých množstvo anotácií je podpriemerné a deformujú tak samotnú skladbu a a skladby jej podobné v rámci okolia vrcholu grafu. Túto nevýhodu sa snažíme odstrániť v nami navrhovanej hre.

3.2.3 MusicBrainz

MusicBrainz¹⁴ je služba s otvoreným zdrojovým kódom, ktorá pôsobí ako relačná databáza hudobných informácií spravovaná iba samotnými používateľmi. Okrem základných údajov o skladbách (názov skladby, interpreta a albumu) uchováva aj jazyk, vydavateľa, odkazy na externé stránky, dĺžku skladby a pod. Všetky údaje uchovávané službou sú zobrazené na stránke služby spolu s odkazom na samotnú databázu. Kvôli otvorenosti nie je možné spravovať detaily do takej úrovne ako sa deje v LastFM a Pandora, avšak aj bez toho je táto služba dostatočne úspešná na to, aby bola využívaná pri automatickej anotácii skladieb bez akejkoľvek predchádzajúcej anotácie len na základe dát o skladbe, ktoré sú dostupné (dĺžka, veľkosť, bitrate a pod.).

¹⁴ <http://musicbrainz.org>

4 City Lights

4.1 Špecifikácia hry

Cieľom je vytvoriť webovú hru s účelom, ktorá je schopná validovať existujúce množiny anotácií a odstrániť z týchto množín také anotácie, ktoré neopisujú dostatočne dobre entitu, ktorej patria. Tento cieľ je zakomponovaný do hry tak, aby hráč o jeho plnení nemusel vedieť.

Úlohou hráča je na základe existujúcich anotácií (hra mu ponúkne niekoľko množín anotácií – jedna množina patrí jednej skladbe a práve jedna množina je správna) rozhodnúť, ktorá množina anotácií opisuje práve prehrávanú skladbu najlepšie. Správnosť jeho rozhodnutia je overená na základe zhodnosti práve prehrávanej skladby a zhodnosti skladby, ktorej patrí vybraná množina anotácií. Hráč môže po rozhodnutí explicitnými akciami vo zvolenej množine anotácií vylúčiť/ponechať tie, ktoré sa k skladbe podľa jeho názoru hodia.

Hráčovi sa na začiatku kola vygeneruje graf – tzv. „mapa skladieb“ – a cesta, ktorú musí touto mapou prejsť. Graf je generovaný na základe hráčových predošlých skúseností a príslušnej obtiažnosti – skúsenému hráčovi je vygenerovaný rozsiahlejší graf s väčším počtom (stále rádovo <10) krokov. Grafom je vždy štvorec, mení sa iba jeho rozmer, vrcholy grafu predstavujú rovnomernú sieť. Každému vrcholu grafu je priradená práve jedna skladba z množiny skladieb práve jednej doménovej oblasti (hudobného štýlu) a každej skladbe je náhodne vybraných n anotácií, ktoré budú použité v konkrétnom kole. Tieto anotácie sú vyberané polo-náhodne – hráčovi v rámci dosiahnutia dobrej hrateľnosti musia byť ponúknuté aj zmysluplné (overené) anotácie. Určité percento anotácií je vybraných z množiny validovaných anotácií danej skladby a tieto budú doplnené do potrebného počtu náhodne vybranou množinou nevalidovaných anotácií.

Po priradení skladieb a anotácií vrcholom sa vygeneruje v grafe cesta, ktorú sa bude snažiť hráč prejsť. Hráč sa môže v grafe pohybovať od 2 do 4 smerov (s možným rozšírením na viac smerov), v závislosti od obtiažnosti a aktuálnej pozície v grafe – situácia, kedy je možné ísť len jedným alebo žiadnym smerom nenastane. Po generovaní potrebných údajov sú hráčovi na začiatku kola prezentované a počas celého kola prístupné nasledujúce objekty:

- celý graf, aktuálny aktívny vrchol (vrátane počiatočného vrcholu na začiatku kola), vrcholy, ktoré sú v danom momente možnými cestami
- komponent na zobrazovanie anotácií vybraného vrcholu – vždy je možné zobrazovať iba anotácie vrcholov, ktoré sú možnými cestami z aktuálne aktívneho vrcholu
- hudobný prehrávač, ktorý ponúka možnosť posunúť a pozastaviť prehrávanie a ktorý nezobrazuje žiadne informácie o skladbe, ktoré by hráča ovplyvnili

Po označení vrcholu ako novej cesty na základe prehrávanej skladby a anotácií priradených danému vrcholu hráč zvolí svoju istotu, s ktorou vrchol označil. Čím vyššiu istotu zvolí, tým viac bodov do hry stavia. Hráčovi je následne daná spätná väzba o správnosti jeho voľby. V prípade korektného označenia vrcholu môže hráč v množine anotácií zvoleného vrcholu označiť tie anotácie, ktoré podľa neho neopisujú skladbu dostatočne presne. Ak sa hráč nevie rozhodnúť

z dôvodu veľkej podobnosti dvoch množín anotácií príp. zlej kvality všetkých množín, môže si za bodovú penalizáciu vyžiadať pridanie ďalších anotácií do množiny – to znamená, že nie všetky náhodne vybrané anotácie priradené jednotlivým vrcholom grafu vygenerované do hry sú zobrazované hneď od začiatku a nemusia byť v hre reálne použité.

Kolo pokračuje prehrávaním ďalších skladieb až do momentu, kedy sa hráč dostane na koniec trasy. V neskorších častiach hry môže byť v rámci zvýšenia obtiažnosti spätná väzba poskytnutá až po niekoľkých označených vrcholoch – nie hneď po jednom. Každá akcia hráča, ktorá akýmkoľvek spôsobom prispieva k validácii anotácií je ohodnotená bodmi. Konkrétny spôsob akým je hráč bodovo ohodnocovaný nie je hráčovi prezentovaný – je mu prezentovaná schéma stávkovania, kedy hráč vie, že za stavené body pri správnom tipe dostane dvojnásobok, avšak pri nesprávnom tipe stavené body stráca. Hra pri vyhodnocovaní musí brať do úvahy hráčovu istotu, ktorú uviedol pri svojej stávke (ako veľkosť svojej stávky) a počet možných smerov, z ktorých hráč vyberal – váha množiny anotácií zvolenej z 4 možností musí byť väčšia ako váha množiny zvolenej z 2 možností. Na základe váhy sa anotáciám rovnomerne zvýši ich globálne skóre. Po prekročení zvoleného skóre sa anotácia stáva validnou.

Bodové ohodnotenie hráča (teda počiatkové skóre, možné stávky a bonusy) je navrhnuté tak, aby spadnúť na nulu bolo možné až v priebehu niekoľkých skladieb – teda zabrániť rýchlej frustrácii hráča hneď na začiatku a aby výsledné skóre hry najlepších hráčov bolo aspoň dvojnásobné oproti počiatkovému skóre kola.

Doména, z ktorej sa hráčovi vyberajú skladby je zvolená samotným hráčom pri registrácii, v neskorších častiach hry sa mu sprístupnia ďalšie hudobné štýly, teda do jedného kola hry bude generátor schopný kombinovať skladby z niekoľkých štýlov súčasne. Registráciou sa hráčovi vytvorí profil, v ktorom bude na základe svojho celkového skóre vidieť svoj postup v hre – na základe tohto skóre sa určuje obtiažnosť kola pri jeho generovaní. Hráčovi sa s rastúcim skóre budú odomykať rôzne vylepšenia a bonusy, ktoré mu uľahčia hru (možnosť preskočiť skladbu bez penalizácie, možnosť vyžiadať si dodatočné anotácie do množiny zadarmo, možno vložiť do hry skladbu, ktorú pozná a pod.). Všetky akcie hráča sú kvôli budúcemu vyhodnoteniu zaznamenané.

Zmena skóre jednotlivých anotácií sa realizuje na základe hráčových akcií. Toto skóre bude v nasledujúcich matematických vyjadreniach označené ako *support*. Jednotlivé akcie obsahujú konštanty, ktorých hodnota ovplyvňuje váhu jednotlivých akcií oproti ostatným. Medzi statické konštanty, ktoré vyplývajú z hry a nie je možné ich upravovať patria:

- **Počet aktívnych ciest aktuálneho vrcholu grafu - *nbr*.** Tento počet ovplyvňuje pravdepodobnosť, s akou mohol hráč uhádnuť správnu odpoveď iba na základe tipu. Počet ciest je závislý od vygenerovanej cesty, kedy táto môže viesť okrajom grafu alebo okolo už navštívených vrcholov grafu.
- **Počet hráčových pokusov na danej skladbe - *attmp*.** Každý pokus na konkrétnej skladbe okrem posledného pokusu je klasifikovaný ako nesprávny. Čím viac nesprávnych pokusov hráč spraví, tým horšie by malo byť skóre anotácie a tým menší vplyv by mali mať hráčove akcie na prípadné zlepšenie skóre.

- **Počet hráčom explicitne vyradených anotácií z danej množiny - *marked*.** Pri tejto konštante je dôležitá jej nulovosť resp. nenulovosť, pretože na základe jej hodnoty sa berie do úvahy zmena skóre anotácie, ktorá nastane pri explicitnom označovaní – v prípade, že hráč nevyradí ani jednu anotáciu, skóre žiadnej anotácie z množiny sa pri akcií vyradovania nezmení.
- **Istota hráčovho tipu - *conf*.** Hráč pri každom svojom tipe určuje, koľko bodov chce staviť. Na základe tejto stávky sa vypočíta istota tipu, ktorá vplyva na zmenu skóre anotácie. Základná istota stávky, od ktorej sa odvíja výpočet ostatných hodnôt istôt je 1 a je adekvátne počtu stavených bodov 100. Za každých ďalších 100 bodov istota narastie o 0.5. Istota c sa teda počíta z počtu stavených bodov pts daného pokusu hráča nasledovne:

$$conf = \frac{pts}{200} + 0.5$$

Okrem vyššie spomenutých konštánt sme definovali nasledovné dynamické konštanty, ktoré je možné upravovať a zlepšovať tak úspešnosť validácie vyplývajúcej z hry:

- **Hráčom vybratá množina anotácií je správna - *ans_c*.** Konštantu sa aplikuje v prípade zvolenia správnej množiny anotácií k práve prehrávanej skladbe – správneho tipu.
- **Hráčom vybratá množina anotácií je nesprávna - *ans_w*.** Konštantu sa aplikuje pri každom nesprávnom hráčovom pokuse.
- **Hráč pri explicitnej spätnej väzbe nevyradil anotáciu, ale vyradil minimálne jednu inú anotáciu z danej množiny - *tag_c*.** Konštantu mení skóre anotácie len v prípade, že je hodnota konštanty *marked* väčšia ako nula – teda hráč vyradil aspoň jednu anotáciu.
- **Hráč pri explicitnej spätnej väzbe vyradil anotáciu - *tag_w*.** Konštantu sa využije v prípade manuálneho vyradenia anotácie hráčom.

S použitím definovaných konštánt je možné vyjadriť zmenu skóre anotácie pri jednotlivých akciách hráča nasledovne:

- a) Hráč neuhádol správnu množinu anotácií pre hrajúcu skladbu a pri pokuse uviedol istotu *conf*.** Pri nesprávnom tipe sa do úvahy berie hodnota konštanty ans_w , istota $conf$ s akou hráč staval body a počet možných ciest z daného vrcholu grafu.

$$\Delta support = - \left(\frac{ans_w * conf * 3}{nbr} \right)$$

- b) Hráč k -tým pokusom v poradí (hodnota *attmp*) uhádol správnu množinu anotácií s istotou *conf*.** Pri správnom tipe je kladne zohľadňovaná istota hráča $conf$, počet možných ciest z daného vrcholu nbr a konštantu ans_c , záporne je zohľadnený počet pokusov na danom vrchole (pri viac ako 2 pokusoch je zmena skóre anotácie takmer zanedbateľná).

$$\Delta support = \left(\frac{ans_c * conf * nbr}{2^{attmp}} \right)$$

- c) **Hráč explicitne nevyradil danú anotáciu z množiny, ale vyradil minimálne jednu inú anotáciu patriacu do rovnakej množiny.** Pri explicitnej väzbe nie je zohľadňovaná žiadna zo statických konštánt a zmena skóre anotácie závisí iba od hodnoty tag_c . V prípade, že je vyradená aspoň jedna iná anotácia, všetkým ostatným neoznačeným anotáciám sa skóre

$$\Delta support = tag_c$$

- d) **Hráč explicitne vyradil anotáciu z množiny.**

$$\Delta support = -tag_w$$

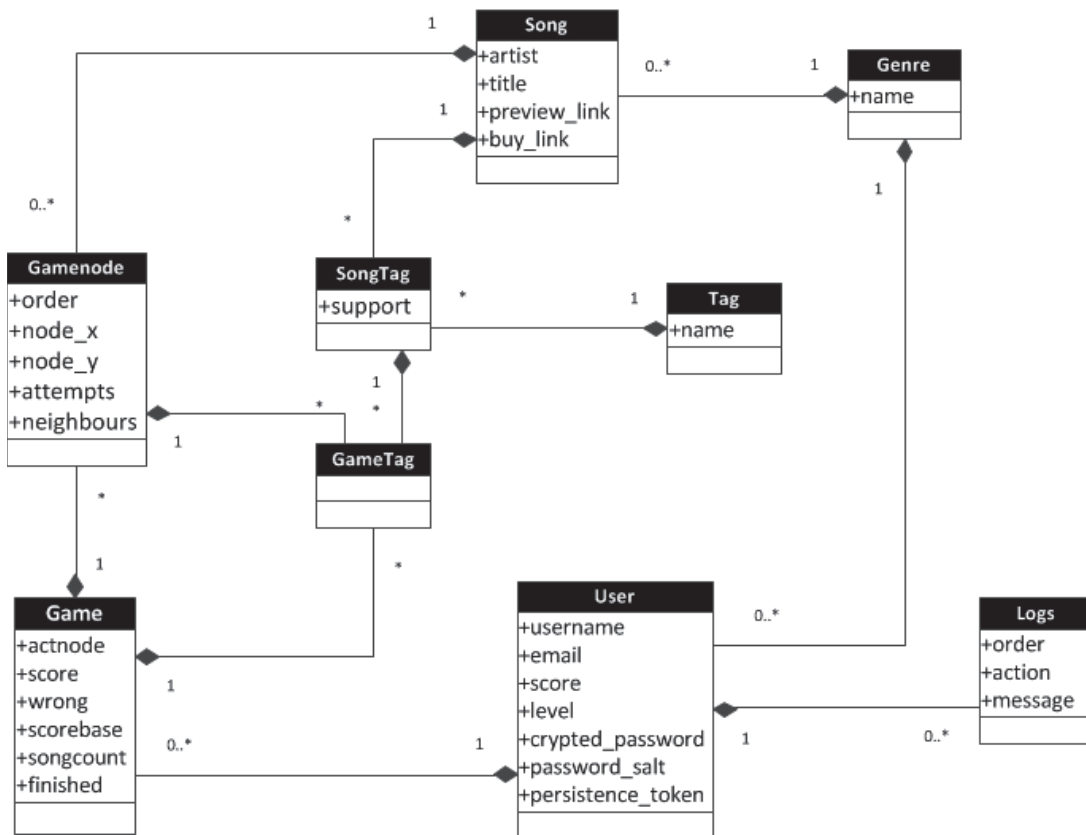
Výsledná hodnota skóre jednotlivých anotácií je súčtom všetkých výskytov konkrétnej anotácie naprieč všetkými hrami. Vďaka faktu, že anotácie sa do každej hry generujú polo-náhodne a striedajú sa, je hra schopná aj napriek tomu že pracuje hlavne s množinami anotácií (okrem explicitného vyradenia anotácie hráčom) určiť správnosť/nesprávnosť pre jednotlivé anotácie.

4.2 Návrh aplikácie realizujúcej hru

4.2.1 Štruktúra aplikácie

Na splnenie špecifikovaných požiadaviek bol navrhnutý dátový model prezentovaný na obr. 4.1 a v ňom identifikované nasledovné entity:

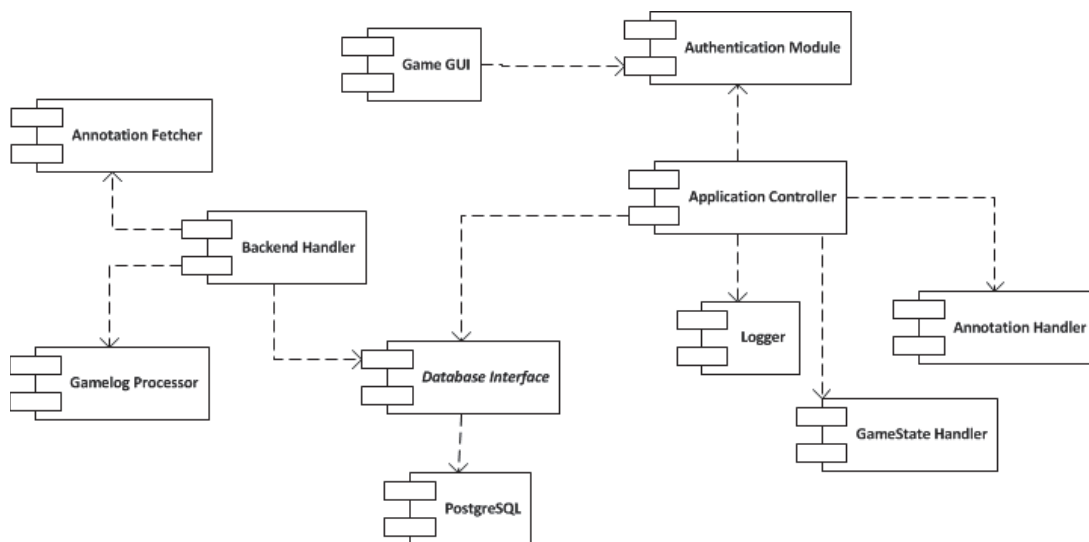
- **Users.** Profil používateľa a jeho doterajší celkový výsledok. Využíva sa pri generovaní nového kola na určenie obtiažnosti.
- **Genres.** Hudobné štýly dostupné v hre. Štýly boli vybraté na základe čo najdôslednejšieho pokrytia hudby za posledných 80 rokov.
- **Songs.** Údaje o skladbách. Okrem základných metadát skladby obsahuje aj odkaz na ukážku skladby a odkaz na stránku, kde si je možné skladbu zakúpiť (podmienka sprístupnenia ukážok).
- **Tags.** Anotácie prebraté z webu k príslušným skladbám. Jedna anotácia môže patriť aj viacerým skladbám.
- **Gametags.** Anotácie vybraté pre konkrétnu hru a vrchol grafu. Podľa hráčovej interakcie sú tieto kladne alebo záporne ohodnocované. Z ohodnotenia týchto herných anotácií sa vypočíta celkové skóre anotácie pre vybranú skladbu.
- **Gamenodes.** Vrcholy grafu konkrétnej hry so zadanými súradnicami a poradím v trase, ktorú má hráč prejsť.
- **Games.** Samotná hra. Ukladá sa hráčove skóre, doba hrania v sekundách a počet skladieb (dĺžka trasy grafu) daného kola.
- **Gamelogs.** Akcie hráča vykonané počas hry. Na základe týchto je na konci kola počítané výsledné skóre hráča.



obr. 4.1 Logický model údajov

4.2.2 Moduly realizujúce aplikáciu

Aplikácia je rozdelená do niekoľkých na sebe nezávislých modulov. Jednotlivé časti systému sú rozvrhnuté tak, aby dokázali fungovať samostatne a komunikovali spolu pomocou svojich rozhraní. Štruktúra je zobrazená na diagrame komponentov na obr. 4.2.



obr. 4.2 Štruktúra aplikácie

Game GUI. Modul riadi vykresľovanie herných prvkov v prehliadači, spracováva odpovede aplikačného servera a aplikuje ich do hry a validuje korektnosť hráčom vykonaných akcií.

Authentication Module. Stará sa o overenie hráčov pri prihlásení a overuje požiadavky na server vznikajúce počas hry. Žiadna neoverená požiadavka (okrem požiadaviek z ukážkovej hry) nie je na serveri spracovávaná.

Application Controller. Zabezpečuje správu žánrov, skladieb a používateľov (teda entít, ktoré nepodliehajú častým priamym zmenám).

Annotation Handler. Stará sa o polo-náhodný výber anotácií do generovanej hry a taktiež o zmenu skóre jednotlivých anotácií počas hry.

GameState Handler. Riadi generovanie hry (výber skladieb, anotácií, vytváranie cesty).

Logger. Zaznamenáva všetky akcie hráča počas hrania na neskoršie použitie.

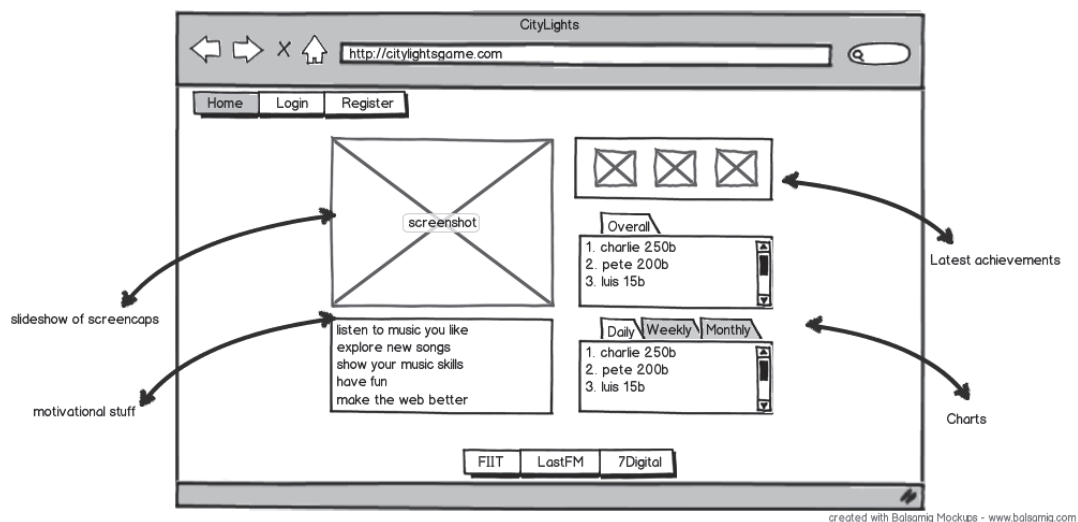
Backend Handler. Stará sa o správu modulov, ktoré priamo nesúvisia s hrou a nie sú ustavične využívané. Jednotlivé volania vykonáva až na externý podnet.

Gamelog Processor. Modul spracováva nazbierané logy a simuluje hry s pozmenenými konštantami za účelom získania presnejšieho výstupu z hry.

Annotation Fetcher. Za pomoci využitia externých služieb zabezpečuje výber interpretov, skladieb, s nimi súvisiacich anotácií a ich zakomponovanie do hry.

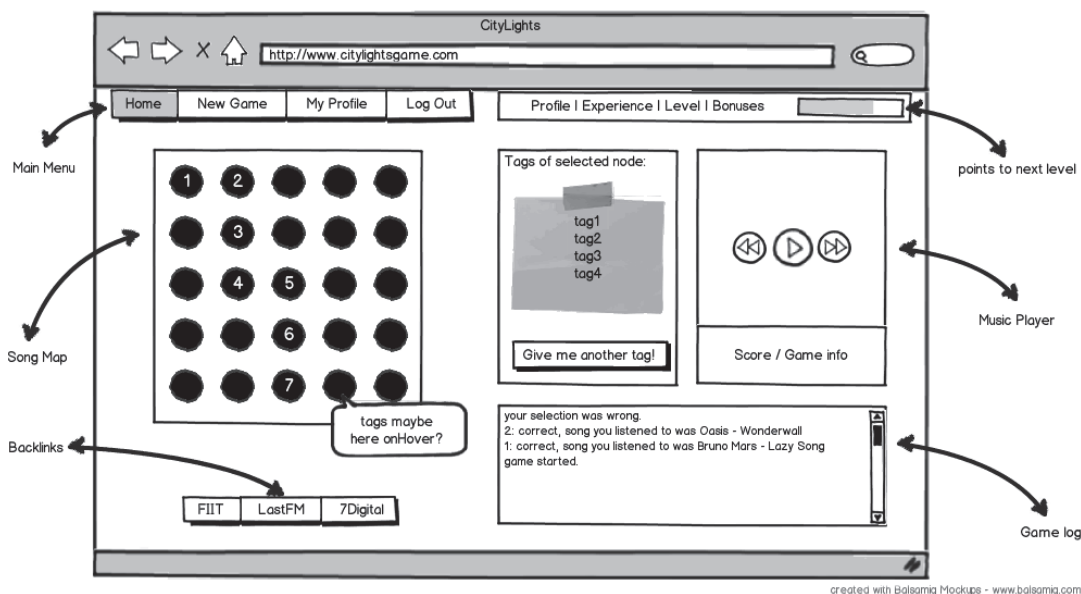
4.2.3 Grafické prostredie

Úvodná obrazovka hry musí zostať čo najjednoduchšia a pritom musí spĺňať funkciu sprievodcu hrou pre nových návštevníkov. Bude obsahovať obrázky z hry a stručný (iba niekoľkokrokový) návod ako hru hrať. Okraje obrazovky budú využité na zobrazenie rebríčkov najlepších hráčov: denný, týždenný, mesačný (ktoré budú medzi sebou v časových intervaloch prepínať) a celkový (ktorý bude viditeľný stále). Návrh je znázornený na obr. 4.3.



obr. 4.3 Návrh úvodnej obrazovky hry

Najdôležitejšia bude obrazovka samotnej hry, kde sme identifikovali nasledovné komponenty: mapa skladieb, zoznam anotácií, hudobný prehrávač a informačná tabuľa. Rozhodli sme sa vytvoriť dva návrhy: (1) všetky ovládacie komponenty umiestniť k sebe a mapu skladieb umiestniť vedľa nich (dvojstĺpcový) a (2) ovládacie prvky umiestniť po bokoch hracej plochy (trojstĺpcový; použitý v implementácii). Návrh konkrétneho rozmiestnenia komponentov pre dvojstĺpcové rozdelenie je zobrazený na obr. 4.4. Na obrázku je naznačená aj sekundárna možnosť zobrazovania anotácií k vybranému vrcholu grafu a to cez tzv. „*tooltip*“, ktorý by sa zobrazoval v prípade, keď by hráč myšou ukázal (nestláčal) na jeden z vrcholov.



obr. 4.4 Návrh grafického prostredia hry

4.2.4 Herná logika

Na začiatku nového kola hry aplikácia vytvára graf. Podľa skúsenosti hráča (jeho levelu a celkového skóre) sa zvolí počet skladieb s , ktoré budú tvoriť trasu grafu. Veľkosť grafu musí byť dostatočne veľká na to, aby mal hráč vždy čo najviac možných ciest a zároveň nie zbytočne priveľká, čo by mohlo spomaľovať vytváranie hry. Na základe tejto požiadavky je rozmer grafu vypočítaný podľa nasledovného vzorca a je následne zaokrúhlený nadol.

$$d = \frac{s - 3}{2} + 4 \downarrow$$

Postupom hrou hráč navrhuje svoje celkové skóre, tzv. level. Na jeho základe sú určované veľkosti hracích plôch a následne bonusy jednotlivých kôl. Výpočet levelu musí byť zvolený tak, aby pri prvých hrách rástol rýchlo a motivoval hráča k hraníu a jeho rast sa počtom odohraných hier spomaľoval. Hranica prechodu do nasledujúceho levelu na základe výšky hráčovho skóre b a výšky aktuálneho levelu l je nasledovná:

$$b = \left[(l + 1) + \left(\frac{(l + 1) * (l) * (l - 1)}{6} \right) \right] * 420$$

Počas hry je nutné hráča ohodnocovať za jeho akcie – v princípe čím prínosnejšia akcia, tým viac bodov by hráč mal dostať. Hráč na začiatku hry začína s bodovým stavom b_s , ktorý je určený podľa počtu skladiieb v kole. Počas hry hráč pri voľbe vrcholu grafu prejaví svoju istotu stávkou bodov, ktoré na svoj tip vsádza. Na výber má stávky 100 a 200 bodov. V prípade, že zvolil správny vrchol, je stavená suma pripočítaná, v opačnom prípade odpočítaná. Ak hráč prešiel celú trasu bez jediného zaváhania resp. s maximálne jedným zaváhaním, obdrží bonusy. Veľkosť týchto bonusov je priamoúmerná počtu skladiieb v danej hre (zohľadnené v parametre „s“) a zohľadňuje skúsenosť hráča – teda čas, hraním sa strávil – zohľadnené v parametri „l“. Konkrétne bonusy sú vypočítané podľa nasledovných vzorcov:

$$b_{max} = s * 200 * (l * 0,5 + 0,5)$$

$$b_{max-1} = s * 100 * (l * 0,5 + 0,5)$$

Hráč obdrží bonusové body aj v prípade, že po zvolení vrcholu označí vo vybratej množine konkrétne anotácie ako nesprávne. Za označenie aspoň jednej anotácie ako nesprávnej dostane hráč 15 bodov. Označenie viac ako jednej anotácie už ďalej nie je zohľadnené ako prevencia proti zneužívaniu hráčmi. Ak si hráč vyžiada počas hry rozšíriť niektorú množinu anotácií, je mu odrátaných 10 bodov – jedno vyžiadanie pridá jednu anotáciu do zvolenej množiny. Ak hráč dosiahne hranicu 0 bodov, kolo predčasne končí. Celkové bodové skóre hráča b_c , ktoré je udržiavané v jeho profile a z ktorého sa vypočítava level hráča l . Ten je vypočítaný ako súčet váženého skóre b_v nahratého v každej hre. Vážené skóre závisí od počtu skladiieb s v danom kole a hráčom nahratého skóre b_h . Spomenuté bodové stavy sa počítajú nasledovne:

$$b_v = b_h - (s * 65)$$

$$b_c = \sum b_v$$

Aby nenastala pri generovaní trasy situácia, kedy by mal hráč počas svojej trasy na výber iba jednu možnosť prípadne situácia, že sa dostane pred koncom kola do slepej uličky, je potrebné cestu vytvárať pred začatím hry. Na začiatku generovania sa na základe hráčovho celkového skóre a levelu zvolí rozmer grafu a počet smerov, ktorými sa hráč môže vybrať. Hra zvolí jeden vrchol náhodne a nastaví ho ako iniciálny pre dané kolo. Pre vybraný bod zistí koľkými možnými smermi sa môže hráč vybrať a náhodne jeden z týchto smerov vyberie. Takto postupuje až kým nie je trasa dlhá potrebný počet hrateľných vrcholov. V prípade, že je možný pohyb iba jedným smerom, generátor sa automaticky posunie daným smerom a opakuje procedúru – v tomto prípade vrchol, v ktorom sa vyskytla takáto možnosť nie je označený ako hrateľný. V prípade, že sa generátor dostane do slepej uličky, je celá doposiaľ vygenerovaná trasa zahodená a generovanie je spustené od začiatku. Postupnosť vrcholov v trase aj s počtom aktívnych susedných vrcholov (potenciálnych ciest) je uložená do databázy.

4.3 Realizácia hry s účelom

4.3.1 Zdroj údajov pre hru a spôsoby prezentácie

Ešte pred začatím implementácie bolo potrebné vybrať vhodnú množinu anotácií, s ktorou budeme pracovať. Z dostupných možností boli najvhodnejšie a najprístupnejšie anotácie serveru LastFM, ktorý ku každej populárnej skladbe pomocou verejne dostupného API¹⁵ ponúka až do 100 používateľmi vytvorených anotácií. Dostupné API obsahuje aj metódu na odporúčanie podobných skladieb, ktorá bola využitá pri získavaní iniciálnej množiny skladieb a príslušných anotácií. Služba LastFM (príp. akákoľvek iná hudobná sociálna sieť) môže byť v budúcnosti využitá na získanie hráčových obľúbených skladieb a teda väčšiu personalizáciu hry smerom k hráčom. V súčasnej implementácii hráč túto možnosť nemá.

Na klientskej časti je hra oklieštená na čo najmenšie množstvo údajov, aby si nimi hráč nebol schopný pomôcť k správnej trase a vyššiemu bodovému zisku. Klientovi sú odoslané iba súradnice začiatočného vrcholu, prvá skladba trasy, počet skladieb kola, rozmer grafu a anotácie všetkých vrcholov v grafe – tie v skutočnosti hráč všetky na začiatku nepotrebuje, ale žiadnym spôsobom ho nemôžu počas hry ovplyvniť. Aby hra zistila, či hráč vybral správny vrchol grafu, musí odoslať súradnice zvoleného vrcholu na server. Server požiadavku overí (či kolo naozaj patrí žiadajúcemu hráčovi, či sa hráč testuje práve dostupný vrchol a pod.) a ak hráč zvolil správny vrchol zaeviduje a obratom odošle odpoveď s prípadnými ďalšími dátami. Na konci kola hra požiada server o vyrátanie konečného bodového stavu kola, ktoré následne zobrazí hráčovi. Všetky výpočty sú realizované na serveri ako ochrana pred zneužitím hráčom.

Aby bolo možné hráčovi vybrané skladby prehrať, rozhodli sme sa využiť hudobné ukážky serveru 7Digital¹⁶, ktorého databáza skladieb je jedna z najväčších na webe a ukážky sú za určitých podmienok dostupné zdarma. Ukážky sú poskytované vo formáte mp3 a trvajú od 30 do 60 sekúnd, čo plne vyhovuje podmienkam hry. Kvôli licenčným podmienkam servera sme boli nútení zaradiť odkazy na stránky 7Digital, kde je možné si prehrávané skladby zakúpiť. Tieto odkazy sú zobrazované až po odohratí skladby a správnom určení cesty, aby hráč nebol ovplyvnený názvom skladby alebo interpreta.

Na prehrávanie samotných skladieb bol pôvodne zvažovaný `<audio>` tag jazyka HTML5 ako najlepšia voľba vďaka natívnej podpore prehliadačov, avšak kvôli rozdielnej podpore jednotlivých formátov v prehliadačoch bolo potrebné iné riešenie. Prehliadače totiž štandardne podporujú iba dva vybrané formáty (každý prehliadač si volí svoje dva). U niektorých pri jeho podpore zavážila otvorenosť (OGG), u niektorých rozšírenosť (MP3). Avšak ani jeden z dnes najdostupnejších formátov nie je v súčasnosti podporovaný naprieč všetkými prehliadačmi. Tie podporujú hudobné formáty spustiteľné *audio* tagom v jazyku HTML5 nasledovne:

¹⁵ <http://www.last.fm/api/show?service=320>

¹⁶ <http://www.7digital.com>

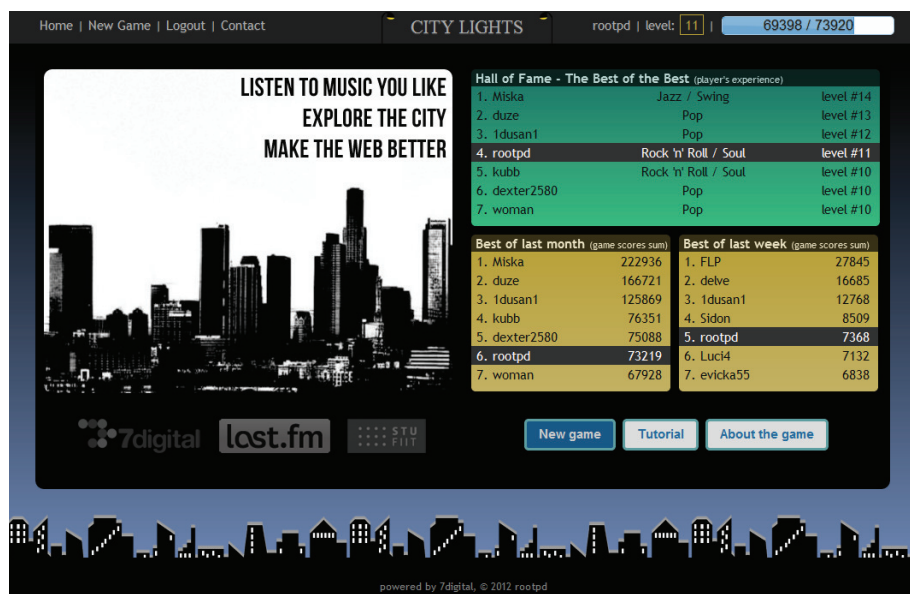
		OGG	MP3	AAC	WAV
	Internet Explorer	✗	✓	✓	✗
	Mozilla Firefox	✓	✗	✗	✓
	Google Chrome	✓	✓	✓	✓
	Safari	✗	✓	✓	✓
	Opera	✓	✗	✗	✓

obr. 4.5 Súčasná podpora hudobných formátov HTML5 audio tagu v prehliadačoch

Ako hudobný prehrávač bolo nakoniec zvolené otvorené riešenie SoundManager 2¹⁷, ktoré nedostatky v HTML5 dopĺňa prehrávačom vo Flashi. Prehrávač ponúka dobre ovládateľné rozhranie s množstvom grafických rozšírení, ktoré ale do projektu nebudú zapracované aby hráča nerozptyľovali. Na akúkoľvek ďalšiu interakciu bol použitý JavaScript v kombinácii s knižnicou jQuery a dostupnými zásuvnými modulmi („*pluginmi*“), ktoré boli využité na grafické efekty v hre a jednoduchú manipuláciu s DOM („*document object model*“) elementmi v kóde.

4.3.2 Aplikácia

Hra¹⁸ je implementovaná na serverovej časti v jazyku Ruby nad frameworkom Ruby on Rails a na klientskej časti (v internetovom prehliadači) v jazykoch HTML, CSS a JavaScript. Na grafické prvky bola použitá knižnica jQuery a hra je optimalizovaná pre všetky aktuálne verzie najpoužívanejších internetových prehliadačov. Na uchovávanie dát je využívaná *PostgreSQL* databáza a ako webserver bol zvolený *nginx*. Celý priebeh kola je priebežne ukladaný a všetky výpočty prebiehajú priamo na serveri aby sa zabránilo podvrhnutiu dát hráčmi. Na správne prehrávanie hudby je potrebné mať nainštalovaný *Flash* prípadne používať prehliadač podporujúci HTML5 audio tag.



obr. 4.6 Hlavná obrazovka hry City Lights v prehliadači Google Chrome

¹⁷ <http://www.schillmania.com/projects/soundmanager2/>

¹⁸ hra je dostupná na adrese <http://bit.ly/city-lights>

Nasadené sú 4 hudobné žánre pokrývajúce druhú polovicu 20. Storočia:

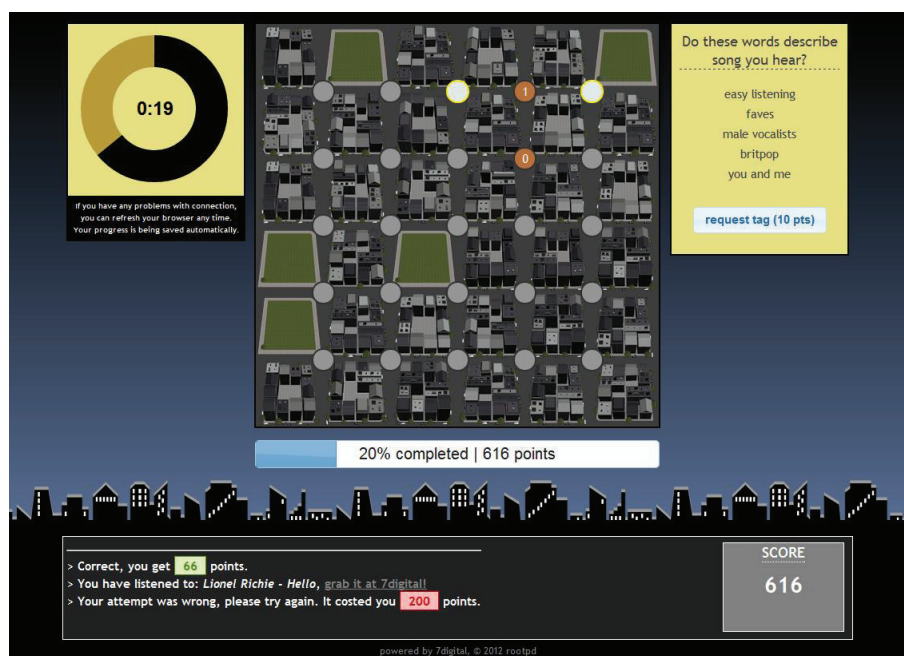
- *Pop* (271 skladieb, 26037 anotácií)
- *Rock* (267 skladieb, 26154 anotácií)
- *Jazz* (236 skladieb, 22314 anotácií)
- *Rock 'n' Roll / Blues* (253 skladieb, 22190 anotácií)

Každá skladba obsahuje minimálne 30 anotácií. Jednotlivé experimenty boli uskutočňované na menšej vzorke skladieb a anotácií - podrobne opísané v kapitole 5.

Hlavná obrazovka hry obsahuje informačný panel na pripútanie pozornosti, rebríčky (celkový, mesačný a týždenný) ako motivačný prvok a odkazy na novú hru, ukážkovú hru a informácie o hre (v prípade, že hráč nie je prihlásený je odkaz na novú hru nahradený odkazom na prihlásenie). Samotná obrazovka je zobrazená na obr. 4.6.

Ukážková hra je oproti štandardnej rozšírená grafické prvky navádzajúce hráča na vykonanie správnej akcie a sprievodných textov, ktoré vysvetľujú hráčovi dôvod vykonania danej akcie. Taktiež obsahuje dodatočné grafické efekty na zvýraznenie zmeny niektorých prvkov. Okrem ukážkovej hry nie je žiadnu inú hru možné odohrať viac krát, každá hra je teda unikátna.

Herná obrazovka obsahuje vopred navrhnuté prvky. Ako hracia plocha bolo zvolené mesto, v ktorom si hráč riešením hudobných hádaniek osvetľuje svoju cestu. Hudobný prehrávač a zoznam anotácií sú rozmiestnené po bokoch mesta a záznam z hry je umiestnený pod mestom. V záznamoch hráč môže vidieť históriu svojich tipov a s tým spojenú zmenu jeho bodového stavu. Po odohratí skladby sa v zázname objaví aj názov interpreta a názov skladby. Z licenčných dôvodov je do záznamu pridaný odkaz na službu 7Digital poskytujúcu hudobné ukážky. Na obr. 4.7 je zobrazená hracia obrazovka po uhádnutí správnej cesty, kedy hráč môže explicitne vyradiť anotácie, ktoré ho zmýlili alebo mu nepomohli.

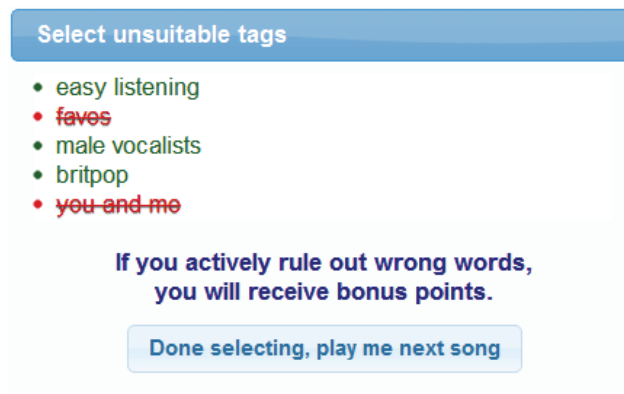


obr. 4.7 Hracia obrazovka hry City Lights

Všetky akcie po načítaní hry sa vykonávajú asynchrónne a preto takto nebudú ďalej explicitne označované. Po spustení hry sa načíta hudobný prehrávač a po spustení hry hráčom sa automaticky spustí prehrávanie skladby. Hráčovi sa po ukázaní myšou na niektorý z aktívnych vrcholov grafu (možné cesty) zobrazia anotácie patriace danému vrcholu – jeden vrchol reprezentuje práve jednu skladbu a anotácie, ktoré sa na ňom zobrazujú patria skladbe, ktorá mu patrí. Kliknutím na vrchol hráč vyvolá akciu overenia svojej voľby, kedy sú na server odoslané súradnice označeného vrcholu, poradie skladby v hre, identifikátor hry a hráč je overený voči serveru. Prehliadač túto požiadavku vykonáva cez JavaScript a odpoveď vyžiada ako JSON (*JavaScript Object Notation*). Server na požiadavku môže odpovedať dvoma spôsobmi:

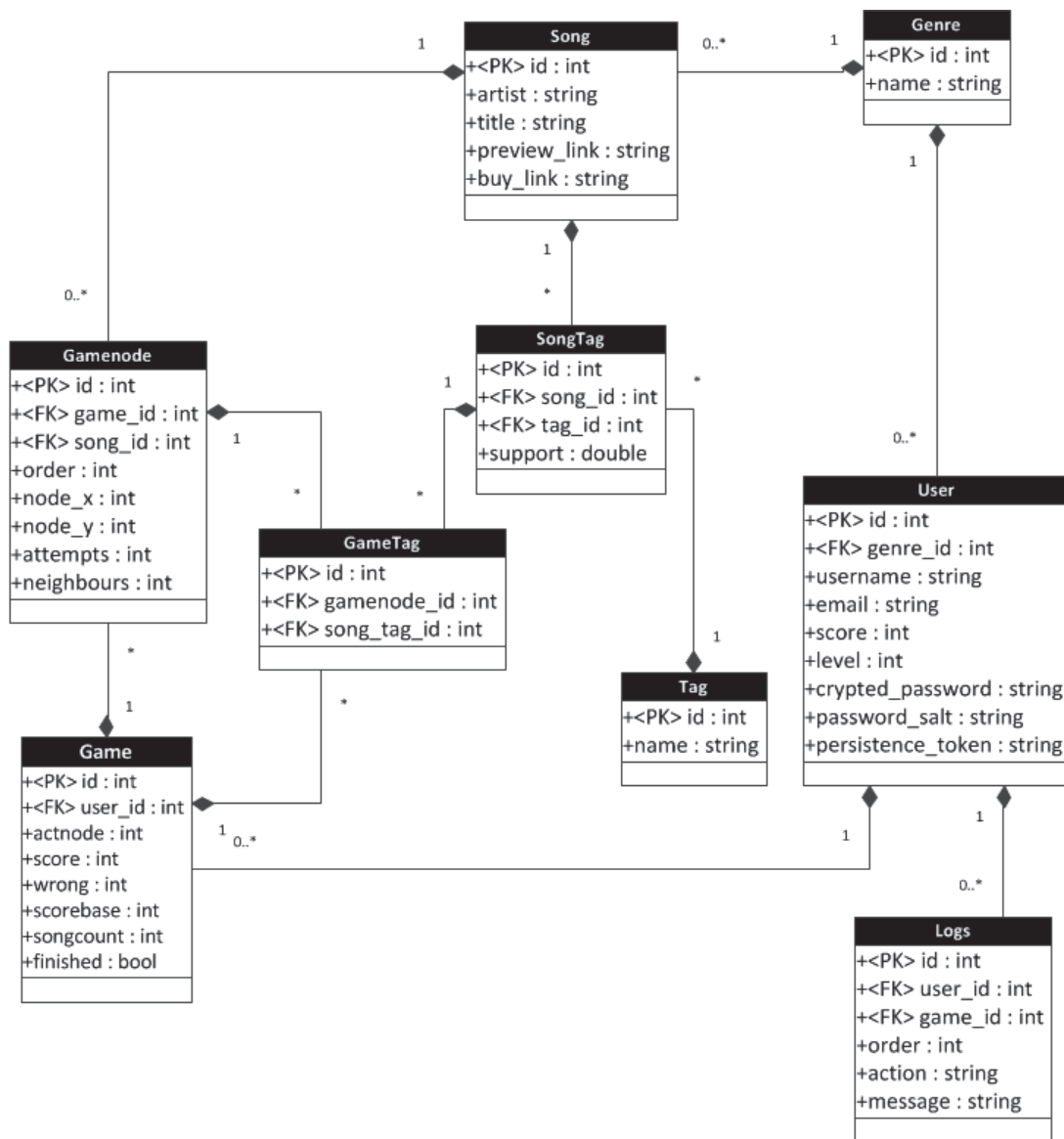
- **False.** V tomto prípade sa na serveri uloží záznam o nesprávnej voľbe hráča a v prehliadači sa prepočíta jeho aktuálne skóre. Výsledné skóre sa počíta z dostupných záznamov uložených na serveri počas priebehu hry.
- **True.** Ak hráč označil správny vrchol, prehliadač prefarbí vrcholy grafu a server okrem návratovej hodnoty vráti aj informácie o ďalšej prehrávanej skladbe, ktorú prehliadač načíta a automaticky začne prehrávať.

Po overení správnosti má hráč možnosť explicitne vyradiť anotácie, ktorú mu nepomohli alebo si myslí, že do množiny nepatria. Toto rozhodnutie mu uľahčuje skladba stále hrajúca na pozadí (aj po správnom rozhodnutí). Hráč má možnosť ľubovoľne meniť rozhodnutia o vyradení až do odoslania a načítania ďalšej skladby. Vyradňovacie okno je zobrazené na obr. 4.8.



obr. 4.8 Vyradňovacie okno po úspešnom tipe hráča

Na obr. 4.9 je zobrazený súčasný fyzický model hry. Jednotlivé entity boli stručne opísané v podkapitole Návrh.



obr. 4.9 Fyzický model údajov

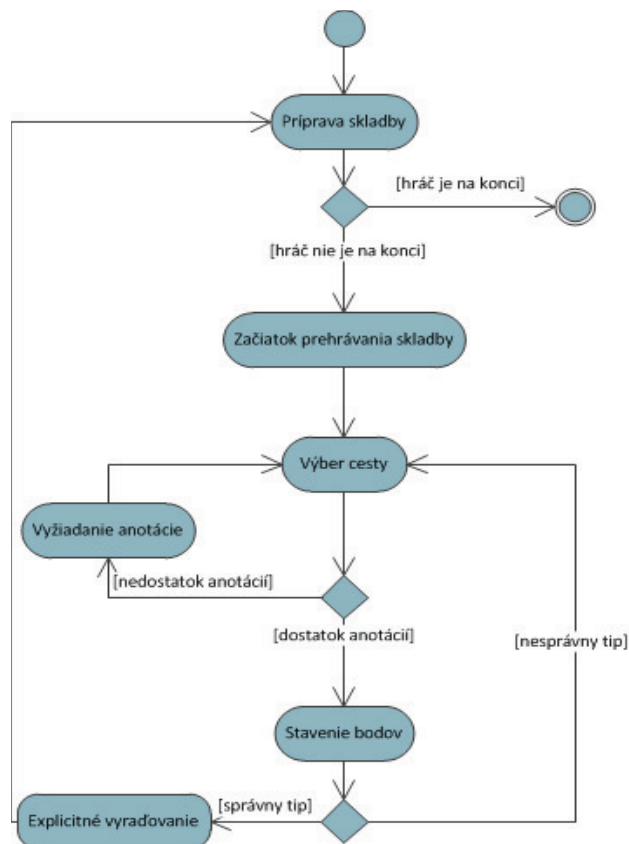
4.3.3 Scenár herného kola

Hráč má pred sebou vždy jeden vrchol označený ako aktuálny (križovatka, na ktorej sa nachádza) a susedné vrcholy – možné cesty – označené ako aktívne, vždy minimálne dva. Pri tejto splnenej podmienke je možné hlavný scenár hry opísať nasledovne:

- 1) Hráčovi sa vygeneruje cesta, kde má vyznačený počiatočný križovatku a možné smery cesty. Každý križovatke je priradená práve jedna skladba a k nej množina anotácií.
- 2) Spustí sa prehrávanie hudobnej ukážky, kedy hráč môže skúmať množiny anotácií na jednotlivých aktívnych križovatkách. Len na základe týchto anotácií a ich súvisu s práve hranou skladbou sa musí rozhodnúť, ktorý smer je správny. Každá križovatka obsahuje po vygenerovaní hry konštantný počet anotácií, avšak hráč si môže vyžiadať doplnenie ďalších anotácií za bodový poplatok.

- 3) Po zvolení križovatky si hráč musí zvoliť výšku stávky (tým stanoví istotu svojho tipu).
- 4) Po správnom tipe má hráč možnosť explicitne vyradiť tie anotácie, ktoré mu nijakým spôsobom nepomohli alebo ho zmýlili.
- 5) V prípade, že hráč dosiahol poslednú križovatku vo vygenerovanej ceste, hra končí. V opačnom prípade hra pokračuje ďalšou skladbou.

Všetky možnosti hráča a ich postupnosť počas kola sú zobrazené na obr. 4.10.



obr. 4.10 Opis herného kola diagramom aktivít

5 Experiment

Cieľom hry bolo z verejných databáz odstrániť tie metadáta, ktoré sú príliš subjektívne príp. nezmyselné resp. validovať jednotlivé anotácie a overiť ich správnosť. Experiment prebiehal v dvoch fázach. V prvej bol uzavretý a kontrolovaný, kde cieľom bolo overiť správnosť hry na malej vzorke hráčov a odhaliť možné nedostatky. V druhej fáze bol otvorený a nekontrolovaný, kde bola hra zverejnená a voľne dostupná na webe.

5.1 Spôsob overovania získaných dát

Keďže okrem služby LastFM neexistuje služba s tak rozsiahlou voľne dostupnou databázou anotácií, využili sme na porovnanie výsledkov hier referenčnú množinu anotácií ručne vyhodnotenú expertmi. Služby, ktoré hudobné anotácie používajú a sú verejne prístupné ich používajú väčšinou len pre hudobné žánre a úplne ignorujú subjektívne alebo menej používané anotácie a nebolo ich možné použiť pri vyhodnocovaní. V Tabuľka 2 sú porovnané anotácie najpoužívanejších internetových obchodov¹⁹ pre skladbu „Wonderwall“ od skupiny „Oasis“ (na porovnanie, LastFM ponúka ku každej skladbe do 100 anotácií – záleží na obľúbenosti skladby):

Tabuľka 2 Porovnanie anotácií najznámejších hudobných internetových obchodov

iTunes	Alternative, Music, Rock, British Invasion, Pop, Britpop, Hard Rock, Rock & Roll, Adult Alternative
Google Play	Alternative, Indie, '00s Alternative, Brit Pop, Brit Rock
7Digital	Rock

Pri uzavretom experimente bol na vyhodnotenie použitých anotácií využitý jeden expert, pri otvorenom experimente traja experti (do úvahy sa bral názor väčšiny). Experti pri vyhodnocovaní nemali navzájom prístup k svojim výsledkom a ich úloha bola definovaná nasledovne: „Na základe poskytnutej hudobnej ukážky, názvu interpreta a názvu skladby určíte, či je nasledovná anotácia pre danú hudobnú ukážku pravdivá“.

5.2 Uzavretý experiment

Experiment s hrou prebiehal v dvoch fázach. V prvej fáze bol experiment uzavretý, hráči boli priamo inštruovaní o pravidlách a možnostiach hry a hrali sa hru pod dozorom. Taktiež boli oboznámení o koncepte hier s účelom. Cieľom experimentu bolo overiť si správnosť hodnôt konštant na výpočet hodnoty kvality anotácie a nájsť prípadné nedostatky v hernej logike. V tejto časti sa experimentu zúčastnilo 5 hráčov, ktorí sa zahráli 5 hier (túto množinu hier mal každý hráč totožnú vrátane skladieb a anotácií v nich použitých). Prvá hra bola ukážková a nezapočítavala sa

¹⁹ Obchod Google Play je v súčasnosti (máj 2012) prístupný iba pre občanov USA

do výsledku experimentu. Každá použitá množina anotácií obsahovala práve 4 anotácie bez možnosti rozšíriť si množinu o ďalšie anotácie. Hráči boli motivovaní k najlepšiemu výsledku malou odmenou pre najlepšieho z nich. Počas experimentu bola hraničná hodnota overenia správnosti anotácie stanovená na 1.5 (v priemere traja až štyria hráči museli vykonať také akcie, ktoré nedegradovali hodnotu kvality anotácie – uhádnuť množinu čo najskôr, nevyradiť explicitne anotáciu a pod.) a hraničná hodnota vyradenia anotácie z množiny na -1.0 (aspoň dvaja hráči z piatich museli anotáciu označiť ako nesprávnu). S týmito nastaveniami boli výsledky uzavretého experimentu nasledovné:

Tabuľka 3 Štatistika úspešnosti hry pri uzavretom experimente

	Vyradené	Nerozhodnuté	Správne
1. hra	6	16	2
2. hra	6	12	6
3. hra	5	14	5
4. hra	7	14	3

Výsledky experimentu ukázali, že už pri 20 odohratých hrách (aj keď striktne kontrolovaných) je hra schopná vyradiť anotácie typu: *elotmbgmegamixx*, *test*, *nice*. Naopak ako správne označila anotácie ako *female vocalists*, *love*, *british*. Tieto výsledky boli pred otvoreným experimentom naozaj sľubné.

Tabuľka 4 Ohodnotenie jednotlivých vybraných množín anotácií

	hráč 1	hráč 2	hráč 3	hráč 4	hráč 5	skóre
art rock	✗	✗	✓	✗	✓	-0.5
punk-pop	✗	✗	✓	✗	✓	-1.5
pop rock	✓	✓	✓	✓	✓	1.5
popular	✗	✓	✓	✓	✓	0.5
makes my heart melt	✓	✗	✓	✗	✓	-1.0
male vocalists	✓	✓	✓	✓	✓	1.0
uk	✓	✓	✓	✓	✓	1.0
easy listening	✓	✗	✓	✓	✓	0.0
00s	✗	✗	✓	✗	✗	-2.0
male vocal	✓	✓	✓	✓	✓	2.0
catchy	✓	✗	✓	✓	✓	1.0
chill	✓	✓	✓	✓	✓	2.0

Experiment avšak odhalil aj nedostatky v bodovom ohodnotení hráčov a množine použitých anotácií v hre – príliš veľká obtiažnosť dokázala odradiť hráčov, ktorí potom nepoužívali vsádzanie bodov zamýšľaným spôsobom, ale z frustrácie stavili vždy maximum. Taktiež bolo potrebné vytvoriť ukážkovú hru (*tutorial*) ako hru hrať pre hráčov, ktorí sa s hrou stretnú po prvý krát, keďže aj pri riadenom experimente mali hráči miestami problém pochopiť princíp hry. Avšak po jeho pochopení ho opisovali ho ako „*skutočne jednoduchý*“. Tieto zmeny boli pred neriadeným otvoreným experimentom do hry úspešne zapracované.

5.3 Otvorený experiment

Do hry bolo vložených 100 skladieb s rôznymi hudobnými žánrov (rock, pop, hip-hop). Ku každej skladbe sme stiahli 40 najlepších anotácií z LastFM a najlepších 10 odstránili z dôvodu zachovania optimálnej obtiažnosti hry. Chceli sme sa vyhnúť situácií, kedy by hráči boli schopní uhádnuť správnu cestu vždy na prvý pokus (nízkej obtiažnosti) a situácií, kedy by boli hráči frustrovaní zo zlých anotácií (vysokej obtiažnosti). Interval, ktorý sme zvolili pre experiment sa javil ako optimálny – po vytvorení referenčnej množiny expertmi sa ukázalo, že približne 50% stiahnutých anotácií je nesprávnych. V experimente bolo v konečnom súčte použitých 3000 anotácií.

Hra bola následne nasadená na web. V záujme budúceho spracovania získaných dát z akcií hráčov sme sa rozhodli vyžadovať na hranie registráciu a prihlásenie. Aby sme neodradili potenciálnych hráčov, rozhodli sme sa ponechať ukážkovú hru s návodom na hranie prístupnú aj bez prihlásenia. Výsledky získané z hrania ukážkovej hry boli na konci experimentu ignorované. Experiment prebiehal 10 súvislých dní a hra bola šírená na sociálnych sieťach a emailom. Ukladali sme každú hráčovú akciu, ktorá by akokoľvek mohla prispieť k lepšiemu (hoci aj budúcemu) vyhodnoteniu anotácií, ktoré sa dostali do hry počas živého experimentu²⁰:

- Hráčom zvolená množina anotácií, hodnota stávky a správnosť tipu
- Explicitne vyradené anotácie
- Otvorenie rozhodovacieho okna (bez stavenia)
- Ukázanie kurzorom na niektorú z možných ciest (bez jej výberu)
- Vyžiadanie ďalšej anotácie
- Čas každej z vyššie spomenutých akcií

Počas živého experimentu boli použité konštanty s nasledujúcimi hodnotami (význam jednotlivých konštánt bol podrobne opísaný v kapitole 4.1):

ans_c	ans_w	tag_c	tag_w
0.05	0.10	0.35	0.60

Údaje získané experimentom sme po jeho ukončení využili na dodatočné vylepšenie použitých konštánt. Pomocou uložených záznamov sme boli schopní reprodukovat' každú hru s rôznymi konštantami a tak sledovať rôzne výsledky pri rôznej kombinácii konštánt. Pri tejto metodike sme boli ohraničení množstvom anotácií, ktoré sa objavili v živom experimente a čase, kedy sa anotácie dostali do hry. V prípade, že simulácia validovala správnosť/nesprávnosť anotácie rýchlejšie ako živý experiment, daná anotácia už nebola ďalej v simulácii použitá (teda kopirovalo sa správanie živého experimentu).

5.3.1 Výsledky živého experimentu

Počas experimentu sa zaregistrovalo 78 hráčov, ktorí odohrali 875 hier (z nich 731 bolo dokončených – to, či hra bola dokončená nijak neovplyvňovalo vyhodnocovanie kvality anotácií).

²⁰ Nie všetky spomenuté akcie boli použité pri súčasnom vyhodnotení experimentu

Hráči odohrali 4933 križovatiek – skladieb. Do hry sa z 3000 anotácií dostalo 1492. Každá anotácia, ktorá sa dostala do hry získala v priemere 17.75 implicitných akcií (správny/nesprávny tip hráča) a 5.29 explicitných akcií (vyradenie/ponechanie anotácie v množine po správnom tipe). Počet akcií hráča na jednu anotáciu bol dostatočne veľký, aby sa vylúčilo náhodné overenie/odstránenie anotácie. Výsledky živého experimentu boli nasledovné:

- **487 overených** anotácií (73% určených správne)
- **0 vyradených** anotácií

Správnosť určenia hrou bola overovaná porovnaním s referenčnou množinou vytvorenou expertmi. Hra overila/vyradila 33% z anotácií, ktoré sa dostali aspoň raz do hry. Dôvodom nízkeho počtu overených anotácií a nulového počtu vyradených anotácií boli nízko nastavené konštanty (resp. vysoko nastavená hranica overenia) a nízka obtiažnosť hry. Aj keď hodnota konštant bola rovnaká ako pri uzavretom experimente (kde sa ukázala ako postačujúca), hranica overenia bola stanovená až na 5 resp. -5 z dôvodu zabránenia nesprávnych náhodných rozhodnutí hrou. Nízka obtiažnosť hry bola spôsobená diverzitou vybraných skladieb v hre – nízky počet skladieb a ich relatívna rozmanitosť spôsobila, že hráči boli schopní vybrať správnu možnosť veľmi rýchlo už na základe veľmi odlišných množín anotácií od práve hrajúcej skladby - popová skladba mala okrem svojej množiny anotácií na výber anotácie vianočnej a hip-hopovej skladby. To spôsobilo, že sa hráči nemýlili, často stavili vyšší počet bodov a v menšom množstve vyradili anotácie explicitne.

5.3.2 Výsledky najlepšieho simulovaného experimentu

Dáta zozbierané počas živého experimentu sme spracovali a pripravili na simulácie. Hodnoty dynamických konštant opísaných v časti 4.1 sme testovali na rôznych kombináciách spôsobom *všetky možné kombinácie*. Konkrétne hodnoty testovaných konštant boli nasledovné (konštanty boli podrobne opísané v kapitole 4.1):

ans_c	0.10,0.20,0.30,0.40
ans_w	0.10,0.20,0.30,0.40,0.50
tag_c	0.30,0.60,0.90
tag_w	0.60,0.90,1.20,1.50

Výsledky simulácií boli rôzne, dochádzalo k zlepšeniu aj zhoršeniu oproti živému experimentu. Ako najlepšie boli vyhodnotené kombinácie konštant:

$$ans_c = 0.2, ans_w = 0.3, tag_c = 0.9, tag_w = 0.9.$$

Z hodnôt konštant je vidieť vyššiu váhu explicitnej spätnej väzby hráča a mierne zvýšenú váhu pri hráčovom pomýlení sa oproti správnomu tipu. Výsledky simulácie pri takto nastavených konštantách sú:

- **736 overených** anotácií (66% určených správne)
- **39 vyradených** anotácií (100% určených správne)

Dôvodom nižšej úspešnosti pri overení správnych anotácií je rozličné chápanie niektorých anotácií hráčmi a expertmi. Experti vyhodnocovali anotácie striktne objektívne, hráči počas hrania pri tipovaní a explicitnom vyradovaní dávali aj na svoje subjektívne pocity. V prípade, že anotácia obsahovala názov inej skladby interpreta, od ktorého bola prehrávaná skladba, experti túto anotáciu vyhodnotili ako nesprávnu. Naopak hráčom indikovala nepriame spojenie s prehrávanou skladbou a anotáciu vyhodnotili ako správnu. Tieto nepriame väzby (a ich záporné vyhodnotenie expertmi) spôsobili nižšiu úspešnosť hry pri overovaní anotácií – pritom využitie hry aj na objavovanie nepriamych väzieb (tak ako sa dialo v Little Search Game²¹) môže byť zaujímavý problém pre výskum.

Optimalizácia dynamických konštánt simuláciou teda zlepšila počet rozhodnutých anotácií až o 59 % oproti živému experimentu. Pomer správne overených anotácií sa mierne zhoršil (avšak na vine môže byť už vyššie spomínaný rozdiel pri vyhodnocovaní hráčmi a expertmi), objavili sa ale aj vyradené anotácie, ktoré hra určila totožne ako experti a mala 100% úspešnosť.

²¹ <http://mirai.fiit.stuba.sk/LittleGoogleGame>

6 Zhodnotenie

V tejto práci sme opísali problémy existujúce pri získavaní anotácií pre obsah na webe a možné riešenia týchto problémov. Zamerali sme sa na oblasť hudby a opísali spôsoby získavania hudobných metadát. Taktiež sme predstavili služby, ktoré na webe pracujú so získanými hudobnými metadátami za účelom sprostredkovania personalizovaného obsahu.

Jedným z riešení získavania metadát sú tzv. hry s účelom, kedy hráči pomáhajú počas hry získavať inak ťažko získateľné dáta. Opísali a zhodnotili sme klady a zápory existujúcich hier a na základe získaných poznatkov sme navrhli hru s účelom, ktorá má za úlohu validovať už existujúce a vyradiť nevhodné hudobné anotácie získané z webu. Hra sa teda od doteraz predstavených riešení odlišila tým, že nevytvára žiadne metadáta resp. jedinou úlohou hráča je „vybrať si“. To prinieslo niektoré problémy, ktoré doposiaľ neboli pri realizáciách hier s účelom diskutované a riešené.

Hru sme implementovali, nasadili a vykonali dva experimenty: uzavretý a otvorený. Oba experimenty priniesli špecifické výsledky, ktoré ukázali že hra má potenciál overovať hudobné metadáta presnejšie ako automatické prístupy. Uzavretý experiment poukázal na hraničné možnosti hry, kedy so správne nastavenými konštantami na určenie správnosti anotácie dokáže hra už po niekoľkých odohraniach anotácie určiť jej správnosť resp. nesprávnosť. Otvorený experiment poukázal na problémy pri voľbe skladieb, kedy potreba pokryť čo najväčšie hudobné spektrum a prilákať čo najviac ľudí paradoxne zhoršila výsledky, pretože diverzita skladieb neúmerne znížila obtiažnosť hry.

Na základe otvoreného experimentu sme vykonali aj dodatočné simulácie a hľadali najvhodnejšiu kombináciu konštánt na vyhodnotenie kvality anotácie. Vďaka simuláciám sa nám podarilo overiť o 59% viac anotácií ako pri živom experimente pri podobnej úspešnosti presnosti. Presnosť pri vyradení anotácií pritom zostala 100%.

Hra teda bola schopná priblížiť sa svojmu účelu a to overovaniu metadát a zároveň si našla svojich priaznivcov, ktorí ju hrali aj po skončení experimentu. Splnila teda oba ciele, ktoré sme stanovili na začiatku (validovať metadáta, prilákať stabilných hráčov). Hra v budúcnosti bude vylepšovaná o ďalšie prvky na skvalitnenie hrateľnosti a spresnenie hodnotenia anotácií, na ktoré sme počas jej realizácie narazili.

Práca bola na študentskej vedeckej konferencii IIT.SRC 2012 ocenená v kategórii „*BEST PAPER in Bachelor degree program*“.

6.1 Možnosti rozšírenia hry

Nie všetky nápady a možnosti ako hru sprístupniť boli zrealizované. Niektoré z nich zlepšujú prístupnosť hry, niektoré hrateľnosť a iné možnosti vyhodnocovania. Napriek tomu, že hra dokázala splniť svoj účel, je na mieste prezentovať ďalšie vylepšenia, ktoré by dokázali zlepšiť pocit z hry a pomohli vytvoreniu stabilnej hráčskej základne.

Aj keď je hra iba pre jedného hráča, existujúce hry s účelom ukázali že zábava je dôležitým faktorom každej hry a možnosť hrať spolu s iným hráčom (prípadne proti nemu) tomuto faktoru výrazne prispieva. Najväčším problémom pri súčasnom porovnávaní hráčov je unikátna hra generovaná každému hráčovi a nemožnosť sa priamo porovnať s iným hráčom. Aby sme zachovali princíp hry a zároveň sprístupnili možnosť vedieť sa priamo porovnať s inými hráčmi, predstavujeme možnosť hry na pozvánku. Hráč by po vygenerovaní svojej hry mohol odoslať iným hráčom pozvánku s unikátnym odkazom. Po navštívení odkazu iným hráčom by sa mu vygenerovala totožná hra s hrou zdrojového hráča. Navyše by sa na konci hry okrem štatistik kola uviedol aj rebríček všetkých hráčov, ktorí takéto identické kolo odohrali. Toto sprístupnenie však prináša so sebou riziko podvádzania, kedy by hráč dokázal odohrať jednu hru viac krát a zlepšoval tak svoje najlepšie skóre. Preto by bol tento typ hier bodovaný iným spôsobom prípadne vôbec.

Ďalším dôležitým aspektom je prístupnosť. Ako sa ukázalo, nutnosť registrácie (aj keď rýchlej a jednoduchej) je pre niektorých hráčov obmedzujúca. Preto navrhujeme použiť na registráciu a prihlasovanie niektorú zo služieb ako *OpenID*, *LastFM*, *Google*, *Twitter* a pod. (prípadne kombináciu). V prípade použitia prihlasovania cez LastFM prezentujeme možnosť analyzovania hráčových hudobných štatistik a prispôbenie jeho hlavného hudobného štýlu (alebo priamo prehrávaných skladieb) jeho preferenciám.

Ako bolo už spomenuté (a v nasledujúcej kapitole bude opísané podrobne), hra vyhodnocuje kvalitu anotácií na základe hráčových akcií. V súčasnej implementácii má akcia každého hráča rovnakú váhu. Prezentujeme možnosť rozlíšenia skúsenosti jednotlivých hráčov a zvýhodnenie akcií hráčov rozoznaných ako expertov. Títo totiž pracujú s anotáciami presnejšie a dokázali by urýchliť overenie jednotlivých anotácií. Taktiež by sa zaviedla možnosť zníženia váhy akcií iných hráčov, ktorých výsledky by sa výrazne odlišovali od priemeru a svojimi akciami by znehodnocovali a spomaľovali overovanie.

Aby sme motivovali hráčov hrať hru čo najviac, je potrebné ich k tomu motivovať aj v neskorších štádiách hry. Ako takúto motiváciu môžeme použiť:

- odznaky za dosiahnutie výsledku,
- zavedenie motivácie striedať hudobné žánre vďaka separátnemu bodovaniu pre každý žánr (a následné odmenenie za dosiahnutie určitého skóre v danom žánri),
- možnosť vložiť do hry obľúbenú skladbu,
- vygenerovanie „nekonečnej“ hry, kedy by cesta, ktorú musí hráč prejsť nebola obmedzená svojou dĺžkou, ale hráč by hru hral až pokiaľ by nenarazil na slepú ulicu – tento mód by bol spojený s osobitným motivačným bodovaním, kedy by sa kolo skončilo pri prvej chybe hráča (resp. počet chýb by bol stanoviteľný), avšak každá ďalšia križovatka by násobila hráčove stávky dvojnásobne.
- možnosť preskočenia jednej skladby počas kola (hráč by nestratil body a ani nárok na získanie bonusu za bezchybný prechod kola)

7 Bibliografia

- [1] von Ahn, L. and Dabbish, L. 2008. Designing games with a purpose. *Communications of the ACM*. 51, 8 (Aug. 2008), 57.
- [2] von Ahn, L. and Dabbish, L. 2004. Labeling images with a computer game. *Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, New York, USA, 2004), 319–326.
- [3] Barrington, L. et al. 2009. User-centered design of a social game to tag music. *Proceedings of the ACM SIGKDD Workshop on Human Computation - HCOMP '09*. (2009), 7.
- [4] Casey, M.A. et al. 2008. Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE*. 96, 4 (Apr. 2008), 668-696.
- [5] Dittmar, C. et al. 2011. Two Applied Research Projects in Music Information Retrieval at Fraunhofer IDMT. (2011), 259-272.
- [6] Ho, C.-ju et al. KissKissBan : A Competitive Human Computation Game for Image Annotation Categories and Subject Descriptors. *Memory*. 1-4.
- [7] Law, E.L.M. et al. 2003. Tagatune: A game for music and sound annotation. (2003), 361–364.
- [8] Magistrali, M. et al. 2010. Tonal MIR: a music retrieval engine based on semantic web technologies. *Proceedings of the 6th International Conference on Semantic Systems* (2010), 1–5.
- [9] Mandel, M.I. and Ellis, D.P.W. 2008. A web-based game for collecting music metadata. *Journal of New Music Research*. 37, 2 (2008), 151–165.
- [10] Morton, B.G. et al. 2010. Improving music emotion labeling using human computation. *Proceedings of the ACM SIGKDD Workshop on Human Computation*. (2010), 45-48.
- [11] Nagy, B. 2011. Acquisition of Semantic Metadata via Interactive Games. *Proceedings of IIT.SRC, vol. 1*. (2011), 9-15.
- [12] Shao, B. et al. 2008. Quantify music artist similarity based on style and mood. *Proceeding of the 10th ACM workshop on Web information and data management - WIDM '08*. (2008), 119.
- [13] Šimko, J. et al. 2011. Little search game: term network acquisition via a human computation game. *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*. (2011), 57-61.
- [14] Šimko, J. et al. 2011. Semantics Discovery via Human Computation Games. *International Journal on Semantic Web and Information Systems*. 7, 3 (2011), 23.

- [15] Šimko, J. and Bielíková, M. 2011. Games with a Purpose: User Generated Valid Metadata for Personal Archives. *SMAP '11 Proceedings of the 2011 Sixth International Workshop on Semantic Media Adaptation and Personalization* (Washington, DC, USA, 2011).
- [16] Siorpaes, K. and Hepp, M. 2008. Games with a Purpose for the Semantic Web Games with a Purpose for the Semantic Web. *Computer*. 23, 3 (2008).
- [17] Steinmayr, B. et al. 2011. Karido: A GWAP for Telling Artworks Apart. *pms.ifl.lmu.de*. (2011), 193-200.
- [18] Turnbull, D. et al. 2007. A game-based approach for collecting semantic annotations of music. *International Symposium on Music Information Retrieval* (2007), 2-5.
- [19] Turnbull, D. 2005. Automatic Music Annotation. *Department of Computer Science and Engineering, University of California, San Diego, CA*. (2005), 19.
- [20] Walsh, G. and Golbeck, J. 2010. Curator : A Game with a Purpose for Collection Recommendation. *Human Factors*. (2010), 2079-2082.
- [21] Weber, I. and Robertson, S. 2008. Rethinking the ESP game. *Proceedings of the 27th*. (2008).