

Informačný prieskum a hudba

Dávid Baranko

baranko1@uniba.sk

Informačný prieskum v hudbe je v súčasnosti neustále sa rozvíjajúca interdisciplinárna veda, ktorá poskytuje používateľovi možnosť jednoduchej a efektívnej identifikácie akejkoľvek neznámej hudby prostredníctvom rôznych systémov.

Cieľom príspevku je oboznámiť čitateľa s touto interdisciplinárnou vedou, priblížiť rôzne typy informácií v hudbe, problémy a výzvy, ktoré sa spájajú s vývojom systémov pre informačný prieskum v hudbe a priblížiť konkrétne systémy na identifikáciu hudby Shazam a SoundHound.

ZÁKLADNÁ TERMINOLÓGIA

Ako prvé je potrebné charakterizovať základné termíny, ktorými sa článok zaoberá. Medzi tieto termíny patrí najmä informačný prieskum a hudba.

Informačný prieskum možno charakterizovať, ako „komplex procesov smerujúcich k vyhľadaniu informácií zodpovedajúcich informačnej požiadavke z informačných zdrojov s výstupom ako špeciálnym spôsobom interpretovanými, štruktúrovanými a prezentovanými informáciami“ [1].

Hudba je „akékoľvek rytmické, melodické alebo harmonické zoskupenie zvukov, ktoré je špeciálne zložené a tvorí jeden celok s cieľom odovzdať správu, komunikovať alebo zabávať“ [2].

MUSIC INFORMATION RETRIEVAL (MIR)

Veda, ktorá sa zaoberá informačným prieskumom v hudbe, sa nazýva Music Information Retrieval (MIR). Túto vedu možno charakterizovať ako „interdisciplinárnu vedu, ktorá sa zaoberá získavaním informácií z hudby“ [3].

Medzi odbory, ktoré sa zaoberajú touto interdisciplinárnou vedou patrí napríklad muzikológia, kognitívna psychológia, lingvistiká. Taktiež sem patrí samozrejme IT a knižnično-informačná veda [3].

MIR A INFORMÁCIE V HUDBE

Ako možno pristupovať k MIR?

K informačnému prieskumu v hudbe možno pristupovať dvoma rôznymi spôsobmi. Spôsoby tohto prístupu sú: prístup založený na metadátach a prístup založený na obsahu [4].

■ Prístup založený na metadátach

Prístup založený na metadátach v MIR sa zaoberá informáciami, ktoré je možné vyjadriť v textovej podobe. Súčasný systém na informačný prieskum textových informácií nemá problém tieto informácie vyhľadať. Problémom v prístupe založenom na metadátach však môže byť vhodne definovať kategórie na popis hudby, pretože jeden umelec môže byť uvádzaný ako skladateľ a zároveň aj ako interpret a podobne [4]. Z týchto informácií vyplýva, že medzi informácie tohto prístupu založeného na metadátach patria napríklad údaje o interpretovi, skladateľovi, názov piesne alebo aj samotný text piesne a podobne. Knižnično-informačná veda sa zaoberá najmä týmto prístupom k MIR.

■ Prístup založený na obsahu

K informačnému prieskumu v hudbe možno pristupovať aj na základe samotného obsahu danej hudby. Patrí sem napríklad výška a farba tónu, rytmus, melódia a podobne. Jedným z problémov však je, že hudba existuje v dvoch formátoch: zvuk a noty. Ďalší problém predstavuje vysoká náročnosť extrakcie prvkov z hudby ako melódia, výška tónu a podobne [4].

VÝZVY A PROBLÉMY VO VÝVOJI MIR SYSTÉMOV

Vývoj systémov, ktoré sú schopné identifikovať a poskytnúť používateľovi informácie na základe vypočítania si určitej hudby, je zložitý proces a stretáva sa s niekoľkými problémami a výzvami.

Stephen Downie vo svojej práci uvádza tieto výzvy a problémy sprevádzajúce vývoj MIR systémov. Jednou z týchto výziev je tzv. multiaspektová výzva. Informácie v hudbe podľa Downieho pozostávajú zo 7 aspektov: výška tónu, tempo, harmónia, timbre, editoriálny aspekt, textový a bibliografický aspekt. Tieto aspekty však nie sú striktne exkluzívne. Adagio napríklad môže patriť medzi aspekt tempa, a aj do editoriálneho aspektu v závislosti od kontextu. Multiaspektová výzva teda znamená vyriešiť problémy, ktoré nastávajú pri interakcii medzi jednotlivými aspektami hudobných informácií [5].

Medzi ďalšie výzvy ovplyvňujúce vývoj MIR systémov patria:

Multireprezentačné výzvy – budú sa MIR systémy zameriavať na textové informácie, audio informácie alebo oboje?

Multikultúrne výzvy – každá kultúra ma svoju špecifickú hudbu, ktorú by mali MIR systémy vedieť rozpoznať.

Multiempirické výzvy – vnímanie hudby je individuálne. Vývojári MIR systémov nevedia presne určiť cieľového používateľa a zámer, s akým chce identifikovať hudbu.

Multidisciplinárne výzvy – príliš veľa odborov zaoberajúcich sa vývojom MIR systémov. Vznikajú jazykové bariéry, komunikačné bariéry, rozličné názory a prístupy v rámci rozličných odborov a podobne [5].

MIR SYSTÉMY – POROVNANIE APLIKÁCIÍ SHAZAM A SOUNDHOUND

Nasledujúca kapitola sa venuje aplikáciám Shazam a SoundHound, ktoré dokážu efektívne identifikovať hudbu na základe jej vypočutia.

Aplikácia na identifikáciu hudby Shazam

Jedným zo spôsobov efektívnej identifikácie hudby na základne vypočutia si danej piesne je mobilná aplikácia Shazam. Táto aplikácia je pre každého používateľa operačných systémov Windows Phone, Android a iOS zadarmo. Túto bezplatnú a užitočnú aplikáciu teda môže vyskúšať každý používateľ smartfónov, tabletov alebo chytrých hodínok.

Z informácií uvedených v tejto práci možno tvrdiť, že Shazam je MIR systém fungujúci na prístupe k informáciám o hudbe založený na metadátach.

História aplikácie Shazam

Shazam sa objavil už v roku 2002, kedy fungoval viac ako služba než aplikácia. Všetko čo bolo potrebné na identifikáciu hudby bolo vytočenie príslušného čísla na mobilnom telefóne a podržanie telefónu blízko zdroja hudby. Používateľ následne dostal SMS správu s metadátami o danej piesni. V roku 2008 bol Shazam jednou z prvých aplikácií pre iPhone. Neskôr sa aplikácia postupne vylepšovala a v roku 2011 vedela napríklad už rozpoznávať aj televízne programy [6].

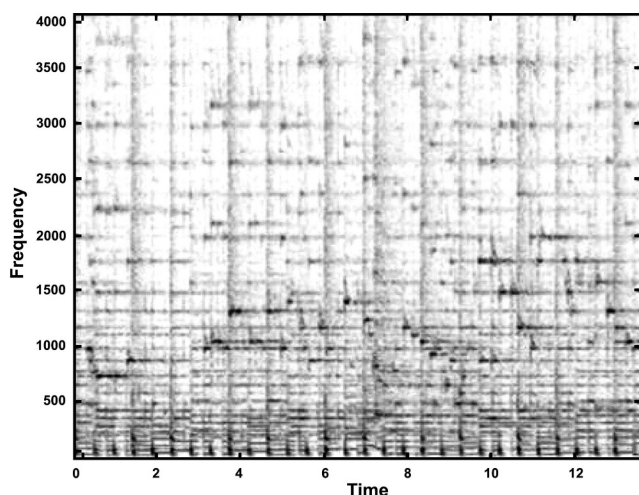
Ako funguje identifikácia hudby prostredníctvom aplikácie Shazam?

Ako môže každý používateľ získať takmer všetky informácie o neznámej hudbe, ktorú práve počuje? Jednoducho, používateľ stlačí príslušné tlačidlo v aplikácii a počká 10 sekúnd.

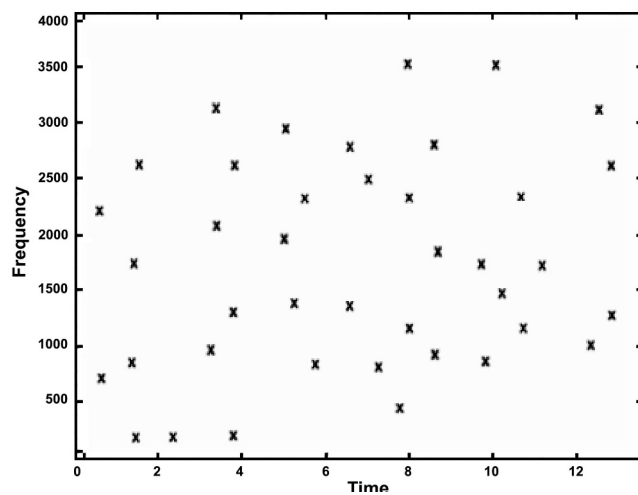
Počas týchto 10 sekúnd sa vytvorí určitá vzorka danej piesne, ktorú aplikácia počuje. Tá sa následne odošle do databázy spoločnosti Shazam, kde sa porovná so vzorkami piesni, ktoré má Shazam uložené na svojich serveroch. Ak sa nájde zhoda medzi vytvorenou vzorkou vypočutej hudby a vzorkou z databázy Shazam, metadáta o príslušnej hudbe sú odoslané naspäť používateľovi [7].

Ako sa však vytvorí daná vzorka z vypočutej piesne? Shazam si na základe vypočutej hudby vytvorí Spektrogram. Je to graf času, frekvencie a intenzity [8].

Zdroj: <http://gizmodo.com/5647458/how-shazam-works-to-identify-nearly-every-song-you-throw-at-it>



Graf č. 1 – Spektrogram



Graf č. 2 – Zjednodušený spektrogram

Ako
v y -

plyva z grafu, na osi X sa nachádza čas a na osi Y sa nachádza frekvencia. Jednotlivé body v spektrograme predstavujú intenzitu daného zvuku.

Tieto body intenzity predstavujú hlasitosť daného zvuku v určitom čase. Čím viac je tento bod tmavší, tým viac je daný zvuk v určitom čase hlasnejší oproti svetlejším bodom. V rámci šetrenia miesta databáza Shazamu uchováva zo všetkých piesní iba tie najhlasnejšie zvuky v danom čase a v danej frekvencii. Preto sa z počiatočného spektrogramu následne vytvorí zjednodu-

šený spektrogram, v ktorom budú iba body s najväčšou intenzitou zvuku [8]. (s. 32)

Zdroj: <http://gizmodo.com/5647458/how-shazam-works-to-identify-nearly-every-song-you-throw-at-it>

Jednoducho povedané, identifikácia hudby prostredníctvom aplikácie Shazam funguje na princípe porovnania zvukov s najvyššou intenzitou v danom čase medzi vytvorenou vzorkou so vzorkami v databáze spoločnosti Shazam.

Test aplikácie Shazam

Shazam bol otestovaný na 12 rôznych skladbách. Nižšie uvedená tabuľka obsahuje informácie o skladbách, ktoré boli použité pri testovaní aplikácie a informácie o ich úspešnej identifikácii. Dané piesne boli vybrané tak, aby bol celkový výber pestrý a rôznorodý pre test.

Interpret	Pieseň	Žáner	Krajina pôvodu	Identifikácia
PSY	Gangnam Style	Pop	Kórea	Áno
Billy Talent	Red Flag	Punk rock	Kanada	Áno
Eva Máziková	Iná žena	Pop	SK	Nie
Percival	Sargon	Folk	Poľsko	Nie
The Prototypes	Pale Blue Dot	Drum and Bass	Veľká Británia	Áno (na 4. pokus)
KOAN Sound & Asa	Sanctuary	Chillstep	Veľká Británia	Áno
Cannibal Corpse	Hammer Smashed Face	Death metal	USA	Áno (na 5. pokus)
Ludovico Einaudi	Nightbook	Vážna hudba	Taliansko	Áno
Desmod	Hemeroidy	Rock	Sk	Áno
Jay Rock	Hood Gone Love It	Rap	USA	Áno
Chiki liki tu-a	Láska moja de si	Alternative	SK	Áno
The Amity Affliction	Pittsburgh	Metalcore	Austrália	Áno

Tabuľka č. 1 – Test aplikácie Shazam

Na základe vykonaného testu aplikácie Shazam možno tvrdiť, že aplikácia funguje relatívne spoľahlivo, efektívne a nemá žiadne väčšie problémy s identifikáciou známejších piesní aj so Slovenska. Taktiež jej nevedil hluk v pozadí počas identifikácie piesní.

Nastali však dva prípady, kedy aplikácia nevedela identifikovať danú hudbu. V prípade týchto piesní môžeme hovoriť o chýbajúcich informáciách v databáze Shazam o týchto piesňach, nakoľko nie sú príliš známe.

Pri dvoch piesňach The Prototypes – Pale Blue Dot a Cannibal Corpse – Hammer Smashed Face však nastal problém s identifikáciou. V prípade týchto piesní sa jednalo o rýchlu a relatívne chaotickú hudbu a Shazam ich vedel identifikovať až v časti, kde znela jasná a čistá melódia špecifická pre danú pieseň, ktorú už vedel systém bez problému zachytiť a identifikovať.

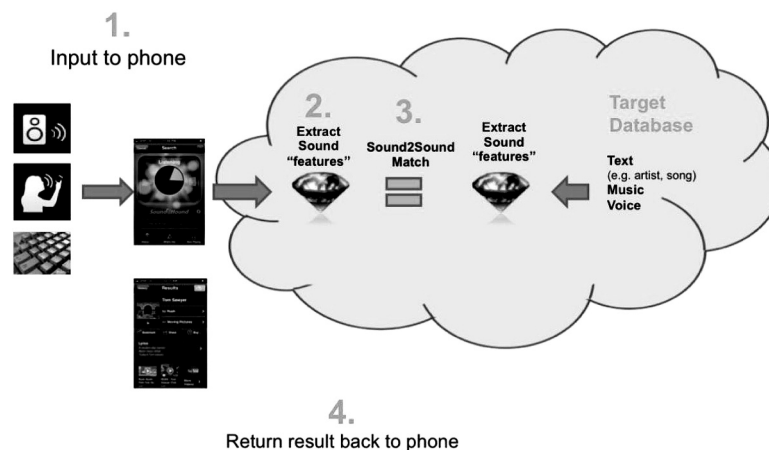
Aplikácia na identifikáciu hudby SoundHound

SoundHound predstavuje alternatívu k aplikácii Shazam. Je rovnakou mobilnou aplikáciou dostupnou zadarmo na všetky v súčasnosti najpoužívanejšie mobilné operačné systémy.

Ako funguje identifikácia hudby prostredníctvom aplikácie Shazam?

Táto aplikácia funguje na základe technológie nazvanej Sound2Sound (S2S) vyvinutou spoločnosťou SoundHound. Identifikácia hudby prebieha na základe extrakcie informácií z vypočítanej hudby a konverzie týchto informácií do tzv. Kryštálovej reprezentácie. Tento Vstupný kryštál je následne porovnávaný s databázou Cieľových kryštálov. Cieľové kryštály sú automaticky generované z rôznych hudobných formátov a zahrňujú hudobné a textové informácie [9].

Identifikácia hudby teda funguje na podobnom princípe ako Shazam. Aplikácie si extrahuje potrebné informácie z vypočítanej hudby a následne ich porovná s databázou, ktorá obsahuje taktiež extrahované informácie z dostupnej hudby, ktorú má k dispozícii spoločnosť Shazam. Na nižšie uvedenom obrázku je znázornený spôsob, akým SoundHound identifikuje hudbu.



Obrázok č. 1 – Vizualizácia identifikácie hudby

Zdroj: <https://www1.soundhound.com/index.php?action=s.sound2sound>

Test aplikácie SoundHound a porovnanie so Shazam-om

SoundHound bol podrobený rovnakému testu ako Shazam. Test prebehol identicky ako v prípade aplikácie Shazam, vid'. Tabuľka č. 1 – Test aplikácie Shazam. SoundHound nevedel rozpoznať rovnaké piesne ako Shazam a taktiež mal rovnaké menšie problémy s identifikáciou piesní The Prototypes – Pale Blue Dot a Cannibal Corpse – Hammer Smashed Face. V tomto prípade vedel SoundHound identifikovať menované piesne rovnako ako Shazam až v tej časti, ktorá obsahovala určitú melódiu špecifickú pre danú pieseň. Keďže sa pri týchto piesňach jedná o rýchlu a relatívne chaotickú hudbu je to však pochopiteľné.

Na základe porovnania týchto dvoch aplikácií možno vyvodiť skutočnosť, že obe aplikácie sú na vysokej a takmer podobnej úrovni a predstavujú efektívny a rýchly spôsob identifikácie hudby pre každého používateľa.

Záver

Na základe informácií a testov v článku je možné tvrdiť, že súčasné systémy na informačný prieskum v hudbe sú na relatívne vysokej úrovni a nemajú väčšie problémy s identifikáciou známejších piesní aj v hlučnejšom prostredí. Stále je tu však veľa problémov a výziev, ktoré treba vyriešiť. Keďže sa táto veda neustále rozvíja a zaoberá sa ňou veľa odborov, je iba otázkou času, kedy sa tieto problémy prekonajú.

Zoznam použitej literatúry

1. STEINEROVÁ, Jela. Teória informačného prieskumu. Bratislava: SITK-CVTI, 1996. 262 s. ISBN 80-85165-58-9
2. Music. *OnMusic Dictionary* [online]. 2013 [cit. 2015-03-31]. Dostupné na: <http://dictionary.onmusic.org/terms/2278-music>
3. TYPKE, Rainer. *Music Retrieval based on Melodic Similarity* [online]. Utrecht University, 2007 [cit. 2015-03-31]. ISBN 978-90-393-4441-5. Dostupné na: <http://dSPACE.library.uu.nl/handle/1874/19776>
4. WIERING, Frans. Can Humans Benefit from Music Information Retrieval? In: *AMR'06 Proceedings of the 4th international conference on Adaptive multimedia retrieval: user, context, and feedback* [online]. Berlin, 2007 [cit. 2015-05-10]. ISBN 978-3-540-71544-3. Dostupné na: <http://www.cs.bu.edu/~snyder/cs591/Handouts/AMRWiering.pdf>
5. DOWNIE, J. Stephen. Music information retrieval. *Annual Review of Information Science and Technology* 37 [online]. Medford, NJ: Information Today, 2003, s. 295-340 [cit. 2015-06-03]. Dostupné na: http://www.music-ir.org/archive/downie_mir_arist37.pdf
6. Our mission is to help people recognize and engage with the world around them. *Shazam* [online]. Shazam Entertainment [cit. 2015-05-15]. Dostupné na: <http://www.shazam.com/company>
7. How Shazam Works To Identify (Nearly) Every Song You Throw At It. JACOBS, Bryan. *GIZMODO* [online]. 2010 [cit. 2015-06-03]. Dostupné na: <http://gizmodo.com/5647458/how-shazam-works-to-identify-nearly-every-song-you-throw-at-it>
8. How does Shazam work to recognize a song? SURDU, Nicolae. *So, you code?* [online]. [cit. 2015-06-03]. Dostupné na: <http://www.soyoucode.com/2011/how-does-shazam-recognize-song>
9. SOUND2SOUND (S2S) SEARCH SCIENCE. *SoundHound: Instant Music Search and Discovery* [online]. SoundHound Inc. [cit. 2015-06-03]. Dostupné na: <https://www1.soundhound.com/index.php?action=s.sound2sound>

Dávid Baranko

baranko1@uniba.sk

(Univerzita Komenského v Bratislave, Filozofická fakulta, Katedra knižničnej a informačnej vedy)