

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

BEERMAPA.CZ - WEBOVÁ SLUŽBA PRO SDÍLENÍ A
POZNÁVÁNÍ OKOLNÍCH PIVNIC

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

ONDŘEJ SVOBODA

BRNO 2014



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

BEERMAPA.CZ - WEBOVÁ SLUŽBA PRO SDÍLENÍ A POZNÁVÁNÍ OKOLNÍCH PIVNIC

WEB SERVICE FOR SHARE AND EXPLORATION OF LOCAL PUBS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

ONDŘEJ SVOBODA

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. IGOR SZÖKE, Ph.D.

BRNO 2014

Abstrakt

Tato práce popisuje tvorbu webové aplikace k projektu Beermapa, který má za úkol zmapovat kvalitní podniky čepující dobré pivo. Projekt je veden jako Startup se snahou o vytvoření fungující firmy. Přidanou hodnotou webové aplikace je algoritmus pro doporučování nových podniků uživatelům, podle jejich předchozího hodnocení.

Abstract

This work describes create web application for Beermapa project which has goal exploring pubs, restaurants and all other kinds of similar businesses when you can order an excellent beer. Project is managed like Startup with effort to create a working company. The added value of a web application is an algorithm for recommending new pubs to users based on their previous evaluation.

Klíčová slova

Webová aplikace, doporučovací metody, Nette, Startup

Keywords

Web application, recommended method, Nette, Startup

Citace

Ondřej Svoboda: Beermapa.cz - webová služba pro sdílení a poznávání okolních pivnic, bakalářská práce, Brno, FIT VUT v Brně, 2014

Beermapa.cz - webová služba pro sdílení a poznávání okolních pivnic

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Igora Szökeho, Ph.D. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Ondřej Svoboda
30. července 2014

Poděkování

Chci poděkovat svému vedoucímu Ing. Igoru Szökemu, Ph.D. za vedení mé práce a rady, které mi poskytnul.

© Ondřej Svoboda, 2014.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

1 Úvod	3
2 Motivace práce	4
2.1 O Beermapě	4
2.2 Beermapa jako služba	4
2.3 Beermapa jako StartUp	5
2.4 Shrnutí	5
3 Návrh služby a webové aplikace	7
3.1 Validace nápadu	7
3.2 Analyzování podobných řešení	8
3.2.1 Pilsner Urquell Navigátor	8
3.2.2 Foursquare	9
3.2.3 Airbnb	10
3.2.4 Zhodnocení konkurence	10
3.3 Uživatelské role	10
3.4 Navržení stránek webové aplikace a jejich rozložení	11
3.4.1 Určení kvality podniku	13
3.4.2 ER Diagram	14
3.5 Návrh mobilní aplikace	14
4 Implementace a testování	16
4.1 Volba frameworku	16
4.1.1 Vývoj v Nette	16
4.1.2 Architektura Nette	17
4.1.3 Struktura balíku Nette	17
4.2 Implementace aplikace	18
4.2.1 Sbíráání bodů a profil ambasadora	18
4.2.2 Ambadorská aktivita a žebříčky	18
4.2.3 Aktuality	20
4.3 Výpočet hodnocení podniku	21
4.4 Testování aplikace	21
4.4.1 Testování uživatelského rozhraní a rozložení aplikace	22
4.4.2 Testování ambadorského dotazníku	22
4.5 Vyhodnocení implementace	23

5 Doporučování podniků	24
5.1 Doporučovací metody	24
5.1.1 Kolaborativní filtrování	24
5.1.2 Shlukování	25
5.1.3 Vyhledávací metody	25
5.1.4 Item-to-item kolaborativní filtrování	26
5.1.5 Zjištění preferencí pomocí lineární regrese	27
5.2 Návrh způsobu doporučování	28
5.3 Implementace doporučování	29
5.4 Testování doporučení	29
6 Závěr	31
A Obsah CD	33
B Tabulka rozdělení práce	34
C Plakátek k projektu	35

Kapitola 1

Úvod

Cílem této bakalářské práce je navrhnout a vytvořit webovou aplikaci pro službu Beermapa. Jedná se o vlastní reálný projekt, u vytvářené aplikace tak bude kladen důraz na výstupní kvalitu prvních verzí, které budou používány veřejností. Do budoucna je předpokládáno, že navržená a vytvořená aplikace bude rozšiřována a vylepšována i po skončení práce v rámci této bakalářské práce.

Beermapa je služba založená na webové aplikaci, která bude uživatele informovat o pivních podnicích splňující určitá pevně stanovená kritéria, která by měla správná pivnice splňovat. Dnes jsou podobné webové služby pro uživatele zcela běžné, poskytují jim pohodlí ovládání a mají nejširší možnosti, proto byla zvolena tato forma. Více je koncept Beermapy rozveden v kapitole 2. Po rozklíčování hlavních cílů bude v kapitole 3 proveden návrh Beermapy jako služby a samotné webové aplikace, pomocí které by tyto cíle měly být naplněny. Implementace a testování je uvedeno v kapitole 4 a poslední kapitola se věnuje algoritmům pro doporučování nových podniků uživatelům, na základě jejich hodnocení.

Kapitola 2

Motivace práce

V této kapitole bude, pro vytvoření lepšího povědomí o cílech Beermapy, přiblížen projekt Beermapa. Bude nastíněno jakým způsobem bude cílů dosaženo a jak je projekt veden.

2.1 O Beermapě

Beermapa je StartUp, na kterém pracuje v šestičlenný tým. Já se dvěma kamarády už od začátku loňského roku, kdy probíhalo spíše ujasňování konceptu. Reálné práce začaly až koncem roku 2013, kdy se tým postupně rozrostl o dalšího programátora webové aplikace, programátora mobilní aplikace a grafičku. Já programuji a zároveň mám na starost kompletní vývoj služby, kolegové spoluzakladatelé mají na starost tvorbu s komunity.

Beermapa se snaží řešit následující problém. Česká republika je pivní národ a podniků s dobrým pivem je u nás spousta, ale ještě více je těch méně kvalitních a ne vždy víme, kde takové podniky hledat. Je to projekt mapující podniky s dobrým čepovaným pivem daného města či regionu, které jsou charakteristické svoji výjimečností a unikátními vlastnostmi. Tyto podniky jsou důkladně vybírány na základě hodnocení v široké oblasti kritérií. Skloubením všeho výše zmíněného dochází k selekci „kvalitních“ podniků od těch „obyčejných“. Beermapa si tak dává za úkol přispět ke zkvalitnění pivní kultury na území ČR. Prostředkem má být vytvoření určitého měřítka, záruky kvality, pro uživatele neznalé prostředí. Beermapa je pro všechny, ale především ji využijí lidé nepříliš znalí lokální pivní scény, tedy studenti vysokých škol, kteří pochází z celé ČR, ale i podnikatelé na služebních cestách, turisté a další skupiny lidí. Pokud se rozhodnou večer někam zajít a užít si to v té správné atmosféře, mohou použít Beermapu, vybrat sami podle vlastních preferencí, a nespolehat se zda jim někdo dobře poradí. Zároveň i pro majitele podniků, skýtá projekt potenciál, nejen co se finanční a popularizační stránky týče a to především jako marketingový nástroj. Hlavním nástrojem Beermapy budou internetové stránky, ale i aplikace pro chytré telefony, která zvýší komfort a pohodlí uživatelů Beermapy. Beermapa má vysoký inovativní a geografický potenciál. Prvním zmapovaným místem bude Brno, ale do budoucna bude Beermapa rozšiřována do dalších měst a států. Podobné použitelné řešení nebylo nalezeno.

2.2 Beermapa jako služba

Abychom všeho zmíněného mohli dosáhnout, na následujících řádcích budou ujasněny cíle a bude stanoven způsob, jakým jich bude dosaženo.

Pro pohodlný výběr vhodného podniku je potřeba získávat informace o podnicích, zpracovávat je a poté přehledně zobrazovat uživatelům. Dále získávat informace o podnicích na základě hodnocení uživatelů Beermapy. Data jsou sesbírána, vyhodnocena a nakonec zobrazena v seznamu a mapě pro jednoduchou orientaci mezi doporučenými podniky.

Důležitým bodem bude práce s uživateli, pro získání co největšího množství dat a zobrazení kvalitních informací. Proto musí být navrženo přívětivé rozhraní webového dotazníku a využití dalších prostředků pro jejich motivaci k hodnocení.

2.3 Beermapa jako StartUp

Beermapa je vyvíjena a vedena jako StartUp. Což podle nejčastěji zmiňované definice od *Erica Riese*, který napsal knihu *The Lean Startup*, je proces inovace a je jedno v jak velkém měřítku, jestli student v garáži tvoří inovativní webovou službu nebo velká nadnárodní společnost v rámci jejich projektu vymýšlí inovativní produkt. Další známá definice je od *Steva Blank*a autora knih *The Startup Owner's Manual* a *The Four Steps to the Epiphany*, který StartUpem nazývá organizaci, hledající opakovatelný a škálovatelný obchodní model. Podle těchto dvou definic můžeme Beermapu označovat jako StartUp, protože je snaha najít fungující obchodní model i službu postupně inovovat. Další zajímavou knihou o StartUpech kniha *Running Lean* [3] od *Ashe Maurya* a je v podstatě kuchařkou pro začínající StartUpisty. Tato kniha jedním ze zdrojů informací pro budování Beermapy jako StartUpu. Předlohou nám tak je postup vycházející z lean přístupu, který je aplikovatelný v podstatě na řešení jakéhokoliv problému:

1. Nalezení zákazníků/uživatelů.
2. Definice a sdělení hypotézy problému.
3. Získání reakce.
4. Upravení hypotézy (eventuální iterace).
5. Neformální a formální definice minimálního produktu.
6. Jeho implementace.
7. Otestování a získání zpětné vazby (eventuální iterace).
8. Zhodnocení a nastínění dalšího postupu.

V našem případě to znamená především učení se nových věcí, hledání cest a zkoušení jejich úspěšnosti. Ať se jedná o získávání nových fanoušků facebookové stránky a měření, co jich přivede nejvíc (v našem případě to zatím byla vtipná videa) nebo například o získávání informací o podnicích, které je daleko těžší než bylo předpokládáno.

2.4 Shrnutí

K Beermapě je přistupováno jako ke studentskému projektu, ale zároveň by z ní měla být vytvořena reálná firma s fungujícím obchodním modelem. Aby byla úspěšná je potřeba vytvořit kvalitní webovou a mobilní aplikaci pro její pohodlné využívání a tedy i větší rozšíření. Kvalitní produkt nelze vytvořit bez zpětné vazby těch nejpovolanějších, což jsou v

tomto případě (potenciální) uživatelé. Proto během celého vývoje bude probíhat uživatelské testování, aby byly odstraňovány nepřesnosti a web byl co nejpoužitelnější a dokázal prorazit mezi širokou veřejností.

Kapitola 3

Návrh služby a webové aplikace

V kapitole 2 je definováno, co by měla Beermapa umět. Známe tak základní požadavky a cíle, které je potřeba při návrhu splnit. Protože je služba vytvářena od začátku je vhodné se poučit z dobrých a špatných stránek podobných řešení, podíváme se proto na konkurenci. Bude představeno rozvržení celé aplikace a stěžejní část pro vybrání kvalitních podniků - výpočet jejich hodnocení.

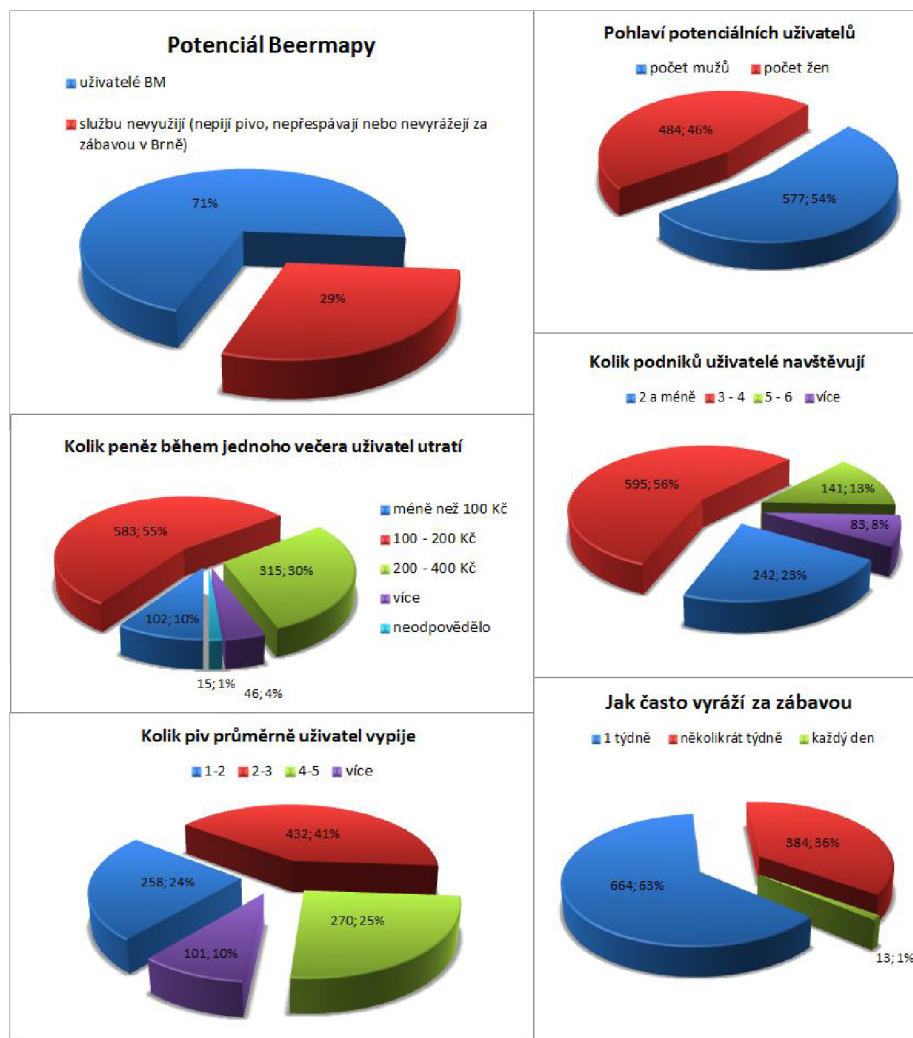
3.1 Validace nápadu

Na samotném začátku byl nápad na realizaci Beermapy validován pomocí internetového dotazníku, zjišťující potenciál Beermapy. Jako cílová skupina byli vybráni studenti všech brněnských vysokých škol. Oslovení pouze studentů dnes nepovažujeme za nejšťastnější, na druhou stranu se alespoň podařilo dosáhnout vyváženého poměru mezi studenty všech oborů a obou pohlaví. Především pomocí sociálních sítí se podařilo oslovit přesně 1500 respondentů. Bylo zjištěno, že 71% z dotázaných vyrazí alespoň jednou týdně za zábavou. Průměrný uživatel pak utratí mezi 100-200 Kč a vypije 2-3 piva. Výsledky jsou k vidění v grafech na obrázku 3.1.

V druhé fázi proběhla osobní komunikace se zhruba dvěma desítkami respondentů, pro zjištění co za funkce by na Beermapě ocenili. To odpovídá bodům 1-4 v lean přístupu uvedeném v seznamu 2.3. Některé z použitých otázek byly:

- Používal by jsi Beermapu?
- Proč by jsi ji používal?
- Co od takové služby očekáváš?
- Využil by jsi ji opakovaně? Proč?
- Jaká kritéria musí podle tebe splňovat dobrá pivnice?

Navržený model byl v podstatě ověřen, nebylo potřeba ho měnit, pouze byl doplněn o několik zajímavých tipů. Jedním z nich je informování uživatelů o aktuální obsazenosti v podnicích, která by měla být aktualizována uživateli pomocí mobilní aplikace, aby tak ve špičce nemuseli hledat volný podnik.



Obrázek 3.1: Grafy s výsledky průzkumu potenciálu Beermapy

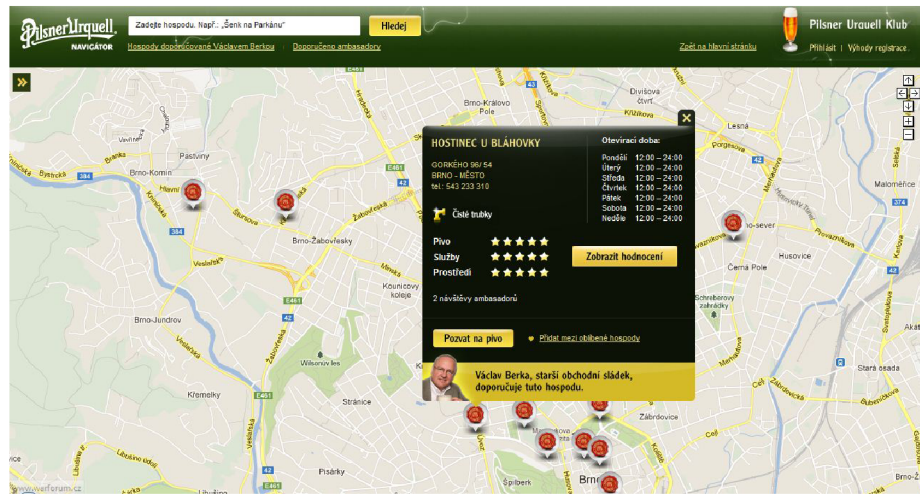
3.2 Analyzování podobných řešení

Souběžnou fází byl průzkum konkurence či nějakým způsobem podobných řešení. Zajímaly nás především řešení, která doporučují nějaká místa na základě jejich hodnocení a tyto místa pak doporučují uživatelům. Přímá konkurence nejsou pouze řešení doporučující konkrétně pivnice, proto jsou uvedeny řešení, z kterých se lze nejvíce poučit.

3.2.1 Pilsner Urquell Navigátor

Pěkné řešení doporučení pivnic, ale bohužel na stránky se už nelze dostat (analýza byla provedena před delší dobou). Pro pěkné provedení jsou zde uvedeny. Už od prvního pohledu profesionálně vytvořená webová aplikace, pod záštitou známého českého piva. Hospody zobrazené přehledně v Google mapách a po kliknutí na konkrétní hospodu zobrazení jejího profilu - přehledný s nejnútnejšími informacemi. Nechyběla možnost zobrazit hospody v seznamu a vyhledávat mezi nimi podle názvu. Co je škoda, ale není to překvapivé, je

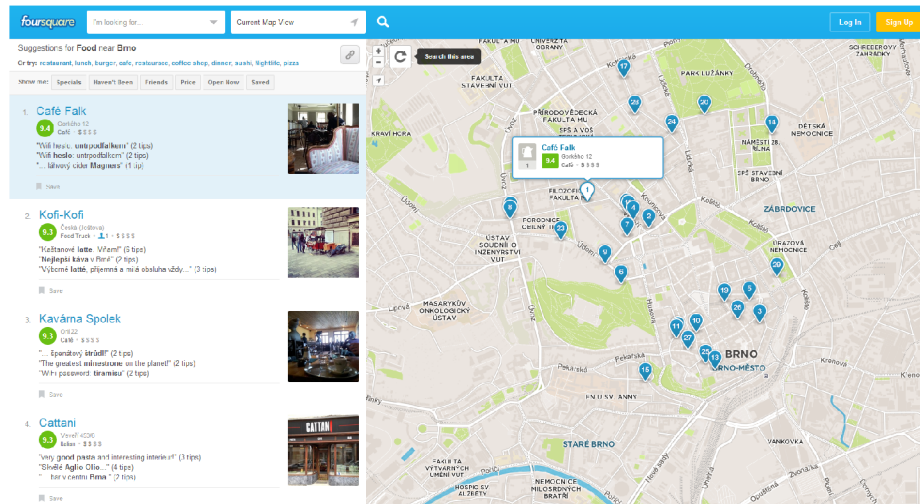
zobrazení hospod čepující pouze Pilsner Urquell.



Obrázek 3.2: Pilsner Urquell Navigátor

3.2.2 Foursquare

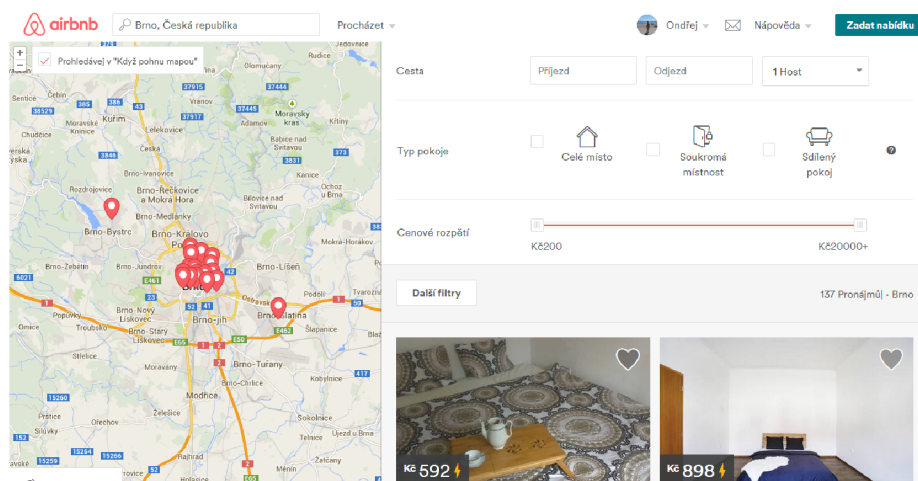
Foursquare je přímou konkurencí, stejně jako jemu podobná řešení (Google Places, Yelp aj.). Je to velmi známá a používaná služba s pěkným webem i aplikací. Rozdíl mezi Beermapou a Foursquare je v tom, že Beermapa se zaměřuje konkrétně na podniky s čepovaným pivem a jde tak více do hloubky. Chce uživateli dát možnost si vybrat, jestli chce hlavně výborné pivo nebo jestli se chce i dobře najít. Důvod je zmíněn v oddílu 2.1. Pokud se model osvědčí, nabízí se přenést ho na víno a kávu.



Obrázek 3.3: Foursquare

3.2.3 Airbnb

Řešením, které sice nepracuje s podniky, ale doporučuje uživatelům místa a je pro nás tedy jako předloha vhodné. Airbnb pomáhá uživatelům hledat ubytování po celém světě, podle jejich preferencí. Nabídku ubytování pak zobrazuje v přehledném seznamu doplněném mapkou. Stránka je jednak velmi přehledná, ale Airbnb je hlavně celosvětová služba se spoustou uživatelů a stránky má tak pravděpodobně postavené na základě spousty zkušeností. Na obrázku 3.4 je vidět nedávno redesignovaná verze stránky.



Obrázek 3.4: Airbnb

3.2.4 Zhodnocení konkurence

Řešení, které přistupuje k definovanému problému jako Beermapa nebylo nalezeno, avšak přímá konkurence existuje. Tu chce Beermapa porážet hlavně díky zaměření konkrétně na podniky s čepovaným pivem a poskytnutí těch nejdůležitějších informací u takových podniků. Navigátor od Pilsneru, který se používal velice dobře, k čemuž přispívalo jeho přehledné zpracování, může být pěknou předlohou hlavně pro zobrazení informací na mapě. Výborně použitelné a zároveň pěkné řešení, zobrazení doporučovaných míst, nabízí Airbnb. Proto tohle je cesta, kterou se Beermapa bude ubírat.

3.3 Uživatelské role

Z výše uvedeného textu vyplývají některé uživatelské role, se kterými je tak v aplikaci potřeba počítat. **Běžný uživatel** se bude moci pohybovat po volně dostupných stránkách, vyhledávat podniky, přihlásit se a ohodnotit podnik (pouze celkovou známkou a krátkým textem). Je ten nejdůležitější, protože pro něho se Beermapa dělá.

Aby mohly být podniky vybírány poctivě, je potřeba mít podrobnější data o podnicích. Toto je důležitý úkol s určitou mírou zodpovědnosti a ne každý bude uživatel bude ochotný takové hodnocení provádět. Proto bude hodnocení pro vybrání podniků sbírat vybraná skupina uživatelů tzv. **ambasadoři**, kterým budou za hodnocení plynout výhody a zároveň bude transparentní, kdo podniky hodnotí. **Ambasador** je tak nejdůležitější pro samotné fungování Beermapy. V aplikaci může dělat vše, co běžný uživatel, ale na rozdíl od něj

vyplňuje propracovanější dotazník a vidí přehledy užitečné pro ambasadory - žebříčky s body, poslední ambasadorskou aktivitu a měl by mít možnost si zobrazit kandidátní podniky v mapě, aby mohl lépe naplánovat, který podnik vyrazí ohodnotit.

Potom je tu **podnik**, který může mít svůj vlastní účet pro zobrazení zpětné vazby na základě hodnocení uživatelů a ambasadorů a také přidávání různých informací o podniku, například akcí konaných v podniku, aktuálně naražených piv a vůbec správy svého profilu.

Poslední je **admin**, který může vše co ostatní a navíc spravuje uživatele, podniky, akce, píše aktuality a vidí veškerá hodnocení a známky podniků.

3.4 Navržení stránek webové aplikace a jejich rozložení

Webová aplikace byla navrhována na základě stanovených cílů a způsobu jejich dosažení, uvedených v oddíle 2.2. Podle záměru jednotlivých stránek, byly rozepsány a poté do stránky rozmístěny funkce, tento záměr plnicí. Pro všechny stránky byly poté vytvořeny wireframy, jako předloha pro pozdější grafický návrh. Vzniklo tak rozvržení stránek:

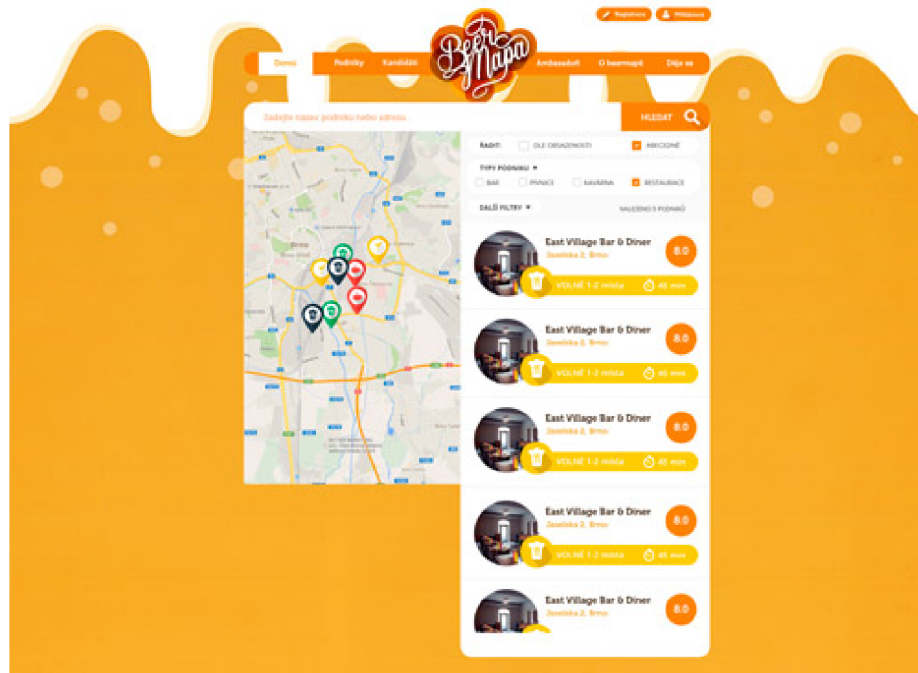
- **Úvodní stránka** - Obecně se používá pro informování uživatelů kam se dostali a jejich zorientování, co vše na stránkách naleznou. V našem případě to je lehký popis Beermapy a odkazování na všechny důležité stránky. To plní jednoduché horizontální menu zastupující hlavní navigaci na webu, které zároveň řeší situaci, kdy uživatel vstoupí na web z jiné než úvodní stránky. Dále odkazuje na poslední tři aktuality a úplně dole je patička s menu, které obsahuje stejné odkazy jako menu hlavní, odkazy na sociální sítě, odkaz na stránku *Kdo stojí za Beermapou?*, a na stránku s formulářem na zpětnou vazbu.
- **Doporučování podniků** - Stránka přehledně zobrazující vybrané podniky. Jak je uvedeno v oddíle 3.2.4 se shrnutím konkurence, zde se Beermapa inspiruje Airbnb (na obrázku 3.4), jejichž stránka splňuje ty nejnáročnější požadavky. Z ní je použito rozmístění mapy, filtru a výsledků hledání. Místo nalezených ubytování v naší stránce budou podniky pod sebou se zobrazenými základními informacemi. Na obrázku 3.5 je vidět jak vypadal původní návrh stránky a na obrázku 3.6 je vzhled vytvořený grafičkou.
- **Kandidáti** - Zobrazení kandidátních podniků sbírajících hlasy, aby bylo jasné, které z nich si uživatelé přejí na Beermapě nejvíce a mají je tak ambasadory vyrazit hodnotit jako první. Pro tento účel je nejúčinnějším řešením seznam s vyhledáváním. Zároveň umožňuje navržení nového kandidátního podniku. Pokud si stránku prohlíží přihlášený ambasadory, vidí u každého podniku možnost přidat nové hodnocení, které ho odkáže na dotazník pro hodnocení.
- **Ambasadoři** - Aby měl kdo kandidátní podniky hodnotit, je potřeba získávat nové ambasadory. K tomu slouží statická stránka s informacemi a možnost vložení e-mailu pro toho, kdo má zájem se jím stát. Bude sloužit i jako domovská stránka pro přihlášené ambasadory a nabízet jim například dotazník pro tisk.
- **O Beermapě** - Poslední stránka, která si zasloužila místo v hlavním menu, informuje uživatele podrobněji o čem Beermapa je.
- **Ambasadorský dotazník** Pro sbírání hodnocení podniků slouží formulář, který je přístupný pouze přihlášeným ambasadory ze stránky s kandidáty. Je rozložený na

6 částí, aby byl uživateli dáváno z psychologického hlediska a aby informace průběžně ukládali a nestalo se jim, že o ně během vyplňování přijdou. Jednotlivé části jsou: první dojem, pivo, obsluha, prostředí, kuchyně a celkový dojem. Koresponduje tak s určenými kritérii, které byly vybrány na základě průzkumu uvedeném v oddílu 3.1.

- **Profil** - Po přihlášení se uživatel dostane do svého profilu, kde vidí kolik nasbíral bodů za hodnocení a jak v minulosti hodnotil. Ambasadory vidí kolikáté jim patří místo mezi ostatními ambasadory a to jak v celkovém, tak třicetidenním horizontu. Vedle toho je zmíněný seznam všech uskutečněných hodnocení, které může otevřít a v druhém seznamu výpis všech akcí, za které získal body. Sbíráni bodů má mít motivační charakter. Profilová stránka byla ale hlavně navržena kvůli zobrazení provedených hodnocení. Do budoucna zde přibudou další prvky (například gamifikační).
- **Aktuality, akce, fotogalerie** - Protože snahou je vytvoření komunitního webu, bude web doplněn ještě dalšími podpůrnými stránkami: Aktuality z pivní scény s doporučeními kam vyrazit na pivo, různé akce konané v podnicích a fotogalerie z nich.



Obrázek 3.5: Původní návrh stránky s podniky



Obrázek 3.6: Vzhled stránky s podniky vytvořený grafičkou

3.4.1 Určení kvality podniku

Důležitým místem Beermapy je určení kvality podniku pomocí hodnocení, aby mohly být vybrány kvalitní podniky. Na základě sesbíraného většího množství bude určena hranice mezi lepšími podniky a těmi ostatními. Hodnotí se známkou 0-10 a odhadovaná hranice je mezi 7-8 body celkové známky. Na Beermapě pak budou pouze ty lepší podniky a to bez zobrazení konkrétní výsledné známky, ale pouze s přehledem, jak si podnik vede v konkrétních kritériích. To bude provedeno pomocí grafu, který bude ukazovat charakter podniku, což by mělo stačit pro vybrání si podniku podle svých preferencí. Tohle bude doplněno možností vyhledávat podniky podle preferencí a doporučování podniků podle uživatelem již provedených hodnocení.

Zveřejnění známky, byť je to u jiných služeb běžné, není u našeho modelu potřeba, protože budou zveřejněny pouze dobré podniky a bude tedy naplněn cíl doporučování kvalitních pivnic. Pokud uživatel půjde do kteréhokoliv doporučeného podniku, neprohloupí. Nebude si tak muset hledat mezi dobrými a špatnými podniky, ale pouze si vybere, podle toho co přesně hledá. V tom mu pomůže filtr ve vyhledávání, v kterém uživatel zvolí jaký typ podniku vyžaduje. Pro každé kritérium zadá jeho váhu.

Pro splnění tohoto požadavku - výběru podniku na základě preferencí - je třeba mít u podniku kromě výsledné známky, také známky jednotlivých kritérií. Kamenem úrazu by mohla být situace, kdy většina hodnocení podniku bude od uživatelů jedné vyhraněné skupiny, například studentů. Pak by mohla mít dobré hodnocení, ale nemusela by se líbit ostatním uživatelským skupinám. Z tohoto důvodu byly navrženy tři kategorie ambasadorů, do kterých se sami ambasadoři rozřazují podle jim nejbližšího popisu.

Protože ambasador hodnotí podnik na základě konkrétních věcí, hodnocení a výsledná známka, by měly být objektivní, a nedovolit tak ambasadorovi hodnotit podnik podle po-

vedeného nebo nepovedeného večera. Ale aby měl ambasador možnost ohodnotit podnik i na základě subjektivních pocitů, bude pro tento účel vytvořena zvláštní známka, říkájící jak se mu návštěva podniku líbila z jeho pohledu.

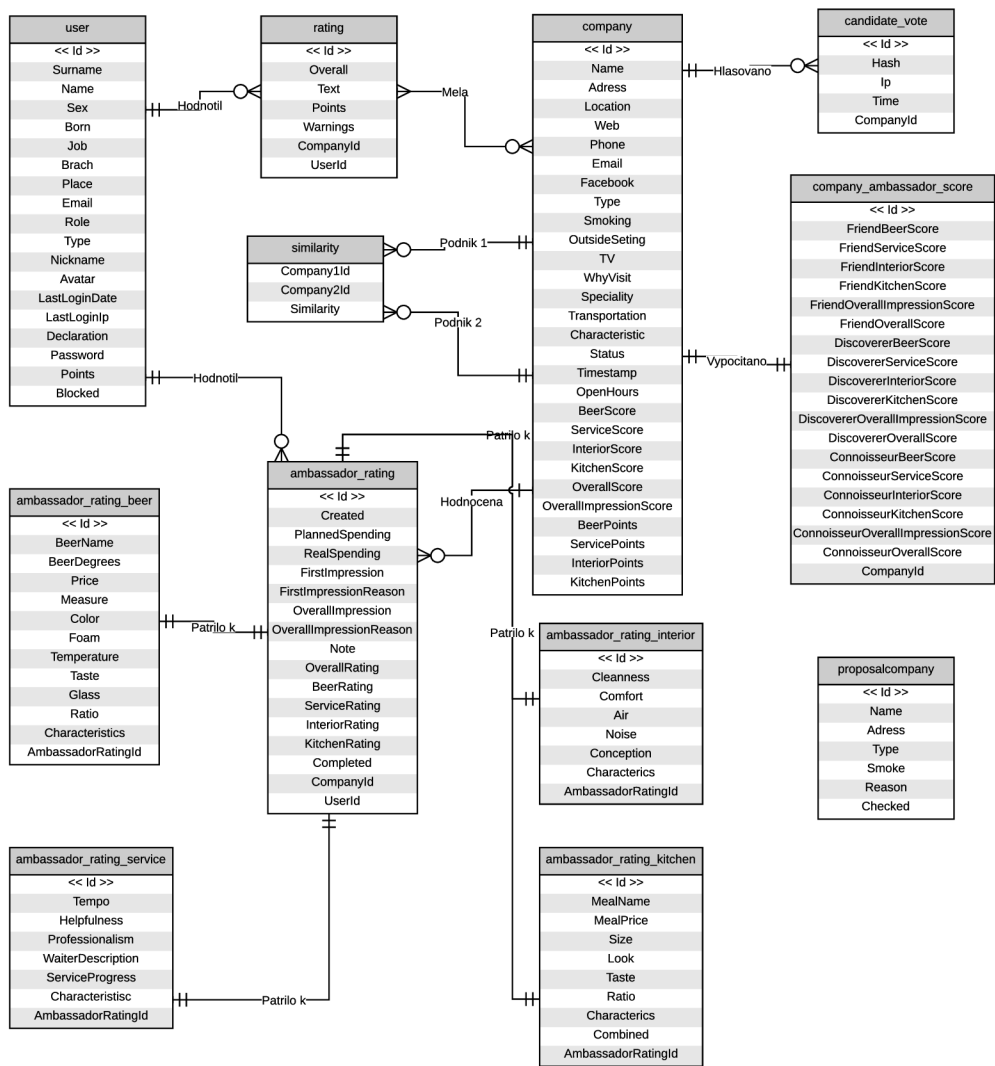
3.4.2 ER Diagram

Po navržení všech stránek a funkcí webové aplikace víme s jakými daty budeme pracovat, a která tak budeme muset uchovávat v databázi. Protože chceme uchovávat velké množství dat o podnicích a k tomu další data, jako jsou fotky z alb, různé akce, aktuality aj., na obrázku 3.7 uvádím pouze nejdůležitější část ER Diagramu, se stěžejními tabulkami. Nejdůležitější jsou *company* - podnik a *user* - uživatel.

U uživatele je uchováváno kromě běžných informací i poslední přihlášení v případě ukrazení účtu. Uživatel nabývá jednu ze čtyř uvedených rolí. Pokud je ambasador, tak může udělat rozšířené hodnocení podniku - *ambassador_rating*. To je rozdělené na hlavní část a pak do čtyřech dalších tabulek podle kritérií. Zde se uchovávají kompletní informace z hodnocení, ale pokud s nimi nebudeme potřebovat pracovat, bude stačit načíst jen jednu rozumně velkou tabulku. O každém podniku jsou uchovávány známky jednotlivých ambasadorských skupin v *company_ambassador_score*, spíše pro účely sledování jak jednotlivé skupiny hodnotí a aby při každém zobrazení nebylo potřeba spočítat známky znovu. U podniku jsou uchovány informace o něm, celková známka, známky za každé kritérium a ještě body, což je přepočítání známky pro zobrazení grafu podniku. *Status* určuje, jestli je podnik kandidát, schválený beermapový podnik nebo podnik, který z nějakého důvodu na Beermapu nemůže být zařazen. Pro navržení nových kandidátních podniků slouží *proposalcompany*. Až je kandidát zkontrolován a převeden na podnik, může v *candidate_vote* sbírat hlasy, kterými komunita říká, který se má ohodnotit co nejdříve. Abychom mohli doporučovat podniky, je potřeba počítat mezi nimi podobnost. Uchovává se jako podobnost vždy mezi dvěma podniky pomocí entity *similarity*.

3.5 Návrh mobilní aplikace

Spolu s analýzou a návrhem webové aplikace, byla mnou provedena také analýza konkurenčních řešení mobilních aplikací a návrh uživatelského rozhraní, ale protože tohle není předmětem této práce nebude to více rozebíráno. Návrh a implementaci API prováděl kolega programátor. Aplikaci vytváří třetí programátor.



Obrázek 3.7: Hlavní část ER diagramu v *Crown's foot* notaci

Kapitola 4

Implementace a testování

Před implementací je ještě potřeba zvolit vhodnou technologii, v čem bude aplikace vyvíjena. Vybrán byl framework Nette, v kapitole bude obsaženo proč spolu s lehkým úvodem do programování v Nette. Dále budou zmíněny některé zajímavé části zajímavější na implementaci, nebude zde zbytečně rozebíráno vytváření šablon, formulářů a dalších běžných prvků, kterých je v aplikaci spousta. Protože jsem aplikaci neprogramoval sám, zmíním pouze ty části, které byly vytvořeny mnou. Kompletní přehled rozdělení práce mezi mnou a kolegem je v příloze.

4.1 Volba frameworku

Jako vývojová platforma bylo vybráno Nette, které je vhodné pro vytvoření webových stránek pracujících s databází. Je to open-source webový framework, který zjednodušuje tvorbu stránek oproti PHP, jehož je nadstavbou. Není tak potřeba řešit některé věci do detailu, ale lze rovnou využít vestavěných funkcí. To se týká hlavně opakovaně používaných věcí. Dalším důvodem proč byl vybrán, jsou moje dřívější zkušenosti s ním. A nakonec to bylo výhodou i při hledání kolegy programátora. Kamarád Radek, má s Nette bohaté zkušenosti, měl chuť na Beermapě spolupracovat a tak jsme nemuseli shánět nikoho cizího.

Nette je framework českého původu, takže jsou dostupné zdroje v češtině na jeho internetových stránkách ¹ a je zde i fungující komunita, která je schopná poradit. V případě řešení chyb má Nette oproti PHP velkou výhodu a tou je tzv. *Laděňka* a chytrý logovací systém. Společně pak dokáží velmi zjednodušit hledání a odstraňování chyb při vývoji [1].

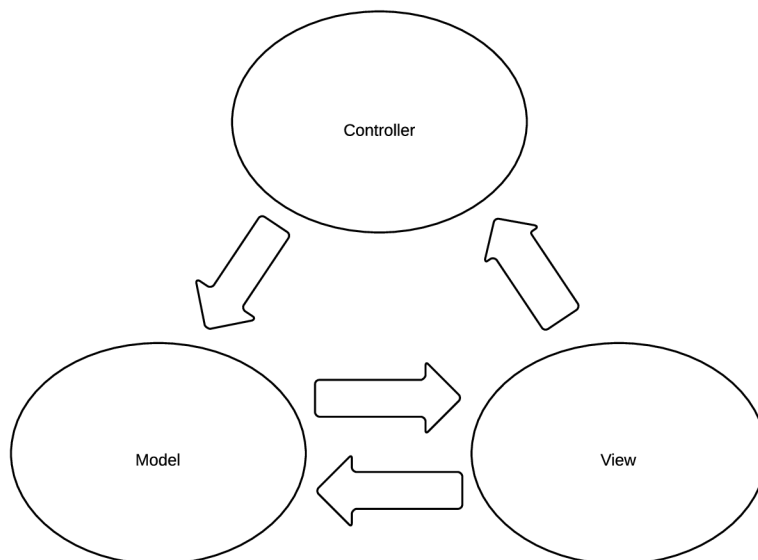
4.1.1 Vývoj v Nette

Pro zahájení vývoje v Nette je potřeba pouze ze stránek frameworku stáhnout balík, který obsahuje potřebné knihovny a strukturu projektu. Stránky tak vytváříte do již předpřipravené struktury, kterou celou spolu s knihovnami následně nahrajete na sever. V konfiguraci stačí nastavit databázi a vše funguje. Nette samozřejmě obsahuje další možnosti nastavení, ať je to třeba nastavení e-mailu administrátora pro zasílání upozornění na chybu na webu. Nette obsahuje vlastní vestavěné funkce, takže například o autentizaci uživatelů je dobře postaráno a není potřeba řešit vše od začátku. Stejně dobře jsou vyřešeny například formuláře.

¹<http://nette.org/>

4.1.2 Architektura Nette

Nette je založeno na softwarové architektuře **MVC** - model, view, controller. MVC odděluje práci do jednotlivých komponent, což má za výhodu, že modifikace v jedné z nich má minimální vliv na druhou. **Model** nám umožňuje pracovat s daty, konkrétně s databází, se kterou pak pracujeme pomocí pevně daného rozhraní. Komponenta **View**, neboli pohled, data pomocí šablon zobrazuje uživateli. A **Controller** má na starosti chování, reaguje tedy na události a promítá je do šablon nebo změnou v modelu.



Obrázek 4.1: Model-View-Controller

4.1.3 Struktura balíku Nette

Balík, který stáhnete na stránkách, obsahuje velké množství různých souborů, rozdělených do složek podle dané struktury. Zajímavá pro nás je především složka **app**, do ní se přidávají soubory, které mají na svědomí vzhled a funkčnost stránek. Druhá důležitá složka je **www**, také přístupná z kořenového adresáře balíku. Do ní se přistupuje z internetu, takže sem přidáváme soubory s kaskádovými styly, obrázky aj.

Pro nás nejdůležitější složky v **app** jsou:

- **models/**: Komponenta model, jak je popsáno výše, slouží k práci s daty. V složce models bývají umístěny datové modely, typicky ve formě tříd, zajišťující práci s databází.
- **templates/**: V templates jsou uloženy šablony, které se plní pomocí presenterů a dat získaných z databáze. Poté se vykreslí uživateli. Šablony jsou rozděleny po složkách podle presenterů, ke kterým patří, vyjma BasePresenteru, ze kterého všichni ostatní dědí. I u šablon je možné dědit, využívá se toho pro vytvoření základního rozložení

stránky pomocí `@layout.latte` a pro konkrétní stránky již specifikujeme jen určité bloky.

- **presenters/**: Jak je patrné z 4.1, presentery slouží k propojení šablon, modelů a komponent. Ostatní vrstvy volají akce vytvořené v presenterech. Pokud budeme chtít například zobrazit podnik, vytvoříme si presenter *Company* a v něm renderovací metodu *renderZobrazPodnik*. Pokud máme vytvořenou odpovídající šablonu, můžeme odkazovat na podnik pomocí makra *Company:zobrazPodnik*.
- **config/**: Obsahuje soubor **config.neon** (v našem případě, kdy používáme starší balík Nette je přímo v adresáři *app*), v němž se nastavuje připojení k databázi a registraci služeb.

4.2 Implementace aplikace

Na základě navrženého ER diagramu, zobrazeného na obrázku 3.7, byla vytvořena databáze v *MySQL*. Poté vytvoření základní struktury stránek a propojení s databází, obstaral kolega programátor. Jedná se o: úvodní stránku, administraci, kandidátní podniky a statické stránky. Mnou byly vytvořeny další, spíše konkrétní části, které jsou uvedeny v tomto oddílu. Malé a běžné úpravy a implementace nejsou uvedeny. Výpočet hodnocení podniků je uveden v samostatném oddílů (4.3) a doporučování nových podniků v samostatné kapitole (5), protože jsou to důležité části a hlavně v případě doporučování i velmi rozsáhlá.

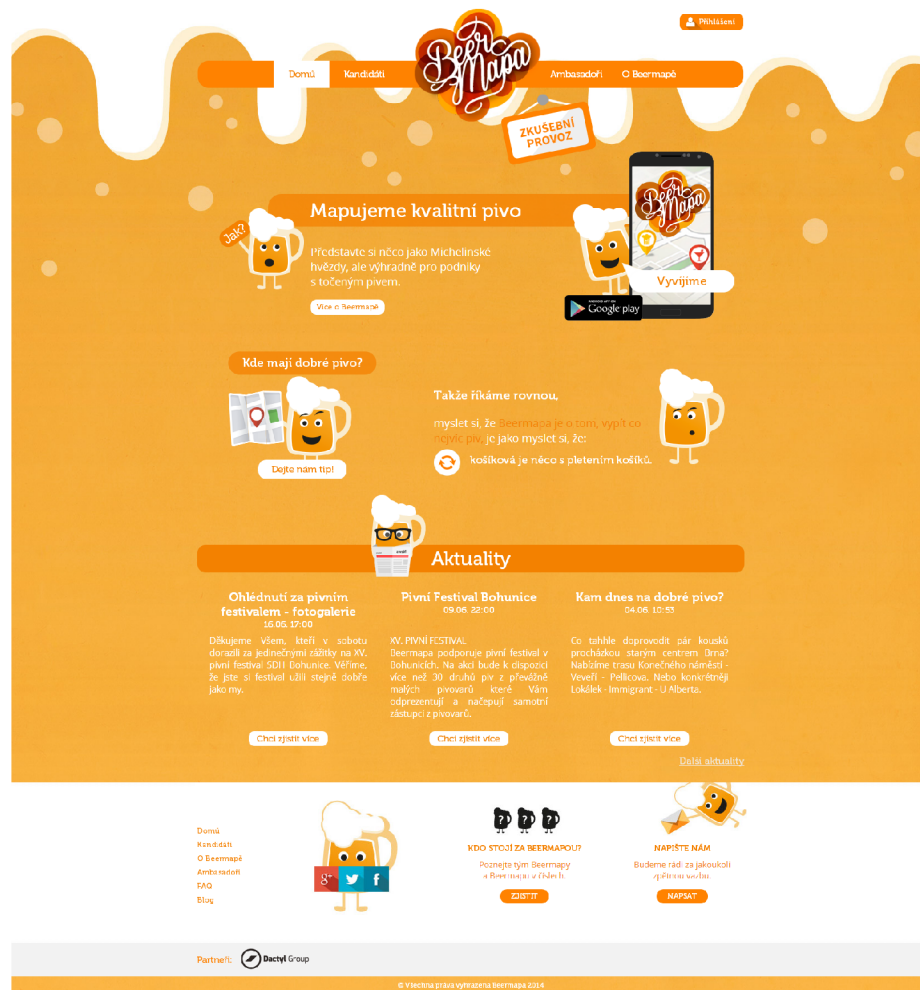
Podle vytvořených wireframů grafička navrhla vzhled jednotlivých stránek. Šablony jsou vytvářeny v jazyce html doplněném o Nette makra a stylovány pomocí kaskádových stylů. Šířka obsahu stránek byla stanovena na 990px, aby je bylo možné zobrazit i na menších displejích. Stránky nejsou navrhovány responsivně z důvodu naší nezkušenosti s responsivním designem a z toho plynoucí časovou úsporou vzhledem ke složitosti stránek. Byla vytvořena hlavní šablona stránek s navigací a patičkou, která je využívána pro všechny ostatní stránky. Jak vzhled webu vypadá je vidět na obrázku 4.2.

4.2.1 Sbíráni bodů a profil ambasadora

Ve svém profilu, který je vidět na obrázku 4.3, má ambasador již hotové hodnocení, nasbírané body za různé akce na webu (zatím jsou to pouze hotová hodnocení), kolik bodů získal celkem a za posledních 30 dní a jaká mu patří celková pozice a opět pozice za posledních 30 dní v porovnání s ostatními ambasadory. Každá odměna se ukládá po provedené akci do databáze a body s umístěními jsou počítány přímo z databáze pomocí SQL dotazů. Dokud nebude v databázi velké množství uživatelů a nasbíraných bodů, je to řešení plně vyhovující.

4.2.2 Ambadorská aktivita a žebříčky

Protože všechno bylo připraveno na sbírání hodnocení, tak aby mohly být vybrány ty lepší a zobrazeny v seznamu podniků, bylo potřeba hodnocení začít získávat. První várkou ambasadorů byli naši známí a lidé, které jsme poznali na pivním festivalu, kde jsme Beermapu představovali. Bohužel ani jedna skupina nezačala výrazněji hodnotit, celý proces se zpomalil a odsunul se vývoj samotné mapy s podniky a začala se řešit lepší motivace ambasadorů. Mnou byla navržena strategie na lepší zapojení ambasadorů, jejíž jednou částí je gamifikace



Obrázek 4.2: Současný vzhled úvodní stránky webové aplikace

2. Proto byly dodatečně navrženy a vytvořeny žebříčky ambasadorů se získanými body a přehled ambasadorské aktivity. Jak vypadají je vidět na obrázku 4.4. Dále pomocí služeb *Mailchimp*³ a *Mandrill*⁴, se kterými byla provedena synchronizace pomocí jejich API, jsme uživatelům, od nichž máme e-maily, začali posílat newsletter a e-maily v případě prvního hodnocení, pro získání zpětné vazby. Automatické e-maily na další akce budou přidány v nejbližší době. Budeme například zjišťovat, co je příčinou přihlášení uživatele jako ambasadora a jeho následná nulová aktivita. V rámci strategie byly navrženy další body, které realizujeme, ale zde je nebudu více rozebírat. Pro další postup Beermapy bylo zapotřebí zapojit více blízké kamarády, se kterými se nám podařilo obejít několik podniků a sesbírat alespoň 50 hodnocení.

²<http://cs.wikipedia.org/wiki/Gamifikace>

³Služba pro vytváření e-mailových kampaní s jejich měřením - <http://mailchimp.com/>

⁴Služba pro posílání mailů na základě akcí na webu a jejich měření - <https://www.mandrill.com/>



Obrázek 4.3: Profil ambasadora



Obrázek 4.4: Ambasadorské žebříčky a aktivita

4.2.3 Aktuality

Aby bylo možné uživatele informovat o novinkách a zároveň pomocí nich odkazovat ze sociálních sítí uživatele na web, je vytvořené jednoduché editační prostředí v administraci s možností nastavení budoucího času publikování aktuality, s možností přidání úvodní fotky a přidání alba fotografií. Vytváření a editování alb i s přidáváním nových fotek bylo vytvořeno také v administraci pomocí formulářů pro editaci a seznamů pro jejich přehled. Vždy tři poslední aktuality se zobrazují ve spodní části úvodní stránky, odkud se dá dostat i na jejich archiv. Do aktuality byly integrovány plugíny pro sdílení na sociálních sítích a facebookové komentáře.

4.3 Výpočet hodnocení podniku

Hodnocení podniku se skládá postupně ze subkritérií, kritérií a výsledné známky. Navíc je potřeba počítat se třemi ambasadorskými skupinami, uvedenými v 3.4.1. Abychom zamezili tomu, že se podnik bude líbit pouze jedné skupině uživatelů (například studentům), musí mít nějaké hodnocení od každé skupiny ambasadorů. Hodnocení je tak nejdříve vypočítáno pro každou skupinu zvlášť. V každé skupině je určen medián každého kritéria, aby se nezapočítávaly extrémní, tím vzniknou výsledná hodnocení každé skupiny. Nejdříve jsou načtena všechna hodnocení jedné skupiny, známky z jednotlivých kritérií rozděleny do polí. U spousty hodnocení není hodnocena kuchyně, toto se kontroluje podmínkou, aby nebyla hlášena chyba, že není co přiřadit. Pokud by nebyla u jednoho podniku hodnocena kuchyně vůbec, je to kontrolováno také, protože se změní poměr koeficientů pro výpočet výsledné známky. Znamky kritérií pro každou ambasadorskou skupinu jsou zprůměrovány do hodnocení výsledného.

Celková známka podniku je určena procentuálním poměrem z výsledných známek kritérií. Procentuální poměr byl určen na základě důležitosti uvedené uživateli v průzkumu 3.1 takto: pivo 50%, obsluha 20%, prostředí a kuchyně 15%.

V ambasadorském dotazníku dává ambasador jednotlivých subkritériím známky v rozmezí 0-10. Abychom dostali výslednou známku za jeho hodnocení, je výpočet známek jednotlivých kritérií proveden opět podle procentuálního rozdělení subkritérií a výsledná známka podniku v rámci hodnocení stejně tak.

Výpočet podle procentuálního rozložení je proveden následujícím postupem. Nejdříve je určen koeficient, kterým bude násoben počet procent připadající na kritérium. Tohle řešení je zvoleno z důvodu jednoduché změny procentuálního rozložení v případě potřeby. Na základě počtu subkritérií (nebo kritérií) je určen maximální počet bodů, který může podnik obdržet. Pokud máme čtyři kritéria je to $4 \cdot 10 = 40$ bodů. Jedno procento z toho je $40/100 = 0,4$ bodu a $0,4/10 = 0,04$ protože chceme získat výsledek opět v rozmezí do 10 bodů. Následně už jen zbývá udělat vážený průměr. Koeficient je vynásoben s procenty a známkou daného kritéria, hodnoty za všechny kritéria jsou sečteny a vyděleny počtem kritérií.

Pro sledování nových hodnocení a změny známek podniků, jsou implementovány přehledy formou tabulek v rámci všech podniků, konkrétního podniku, kde jsou zobrazeny známky i za jednotlivé ambasadorské skupiny a u každého ambasadora. Až bude určena hranice bodů nutných pro kvalitní podnik, který patří na Beermapu, bude přidáno automatické oznamování nového podniku splňujícího všechny nutné podmínky.

Protože máme zatím malé množství uživatelů, probíhá přepočítání známky podniku, při každém novém hodnocení. Do budoucna bude spouštěn *cronem* ⁵.

4.4 Testování aplikace

Protože vývoj probíhá neustále, testuje se také neustále. Testování probíhá výhradě osobně a pokud je to možné, tak při běžném používání. K tomuto účelu je zatím čerpáno z řady známých. V tomto oddílu uvedu některé konkrétní příklady testování.

⁵<http://cs.wikipedia.org/wiki/Cron>

4.4.1 Testování uživatelského rozhraní a rozložení aplikace

Po implementaci prvních stránek bylo provedeno jejich testování. Byly vytvořeny dvě sady mírně odlišných otázek. První pro ty, kteří Beermapu znali a druhá pro ty, kteří ji neznali. U toho, kdo ji neznal nás zajímalo, jestli po vstupu na stránky je schopen zjistit o co se jedná. U uživatelů, kteří už měli o Beermapě povědomí nás zajímaly konkrétnější věci, zda například dokáže přidat nový kandidátní podnik. Tímto testováním byly například zjištěny špatné informace na úvodní stránce a její nepřehledné rozložení, při kterém některé informace zanikali. Dále bylo zjištěno složité přidání kandidátního podniku, které se zobrazilo pouze při hledání pomocí filtru a nenalezení žádného podniku. Proto bylo upraveno rozložení a text na úvodní stránce a přidání kandidáta bylo vylepšeno přidáním viditelnějšího tlačítka a zároveň úpravou formuláře na jeho přidání.

4.4.2 Testování ambadorského dotazníku

Jak již bylo zmíněno, ambasadoři jsou snad nejdůležitější částí pro fungování Beermapy, proto je jim věnováno hodně pozornosti. Dotazník, pomocí kterého sbírají hodnocení podniků, proto musí být z uživatelského hlediska co nejpřívětivější. Dotazník byl vytvořen pomocí formulářů v líbivém designu, ale při testování v terénu (hodnocení podniků na mobilních telefonech) byla sesbírána řada podnětů, která ukázala, že dotazník je potřeba celý předělat. Některé z poznámek z testování i s návrhy na změnu:

- Špatné rozložení, nesedí s průběhem návštěvy - návrh nové struktury:
 - **Tutorial k dotazníku** - Nezabere místo na dotazníku a stačí ho mít při prvních vyplňování, takže ušetří místo a může být podrobnější.
 - **Začátek návštěvy** - Můžu vyplnit hned po příchodu (než mi donesou pivo), protože je to zřejmé okamžitě a nebudu se muset rozptylovat během návštěvy.
 - * První dojem
 - * Kdo nás obsluhuje
 - * Prostředí
 - **V průběhu návštěvy** - Je těžké zpětně ohodnotit subkritéria u všech piv (i když byly třeba jen 3), takže bude lepší dělat hodnocení v reálném čase.
 - * Piva
 - * Kuchyně
 - **Po návštěvě** - Třeba i druhý den, protože můžu (a lépe) zhodnotit návštěvu a můžu si zároveň více užít návštěvu. Zároveň pomůže ke kontrole a doplnění hodnocení, takže třeba budou poctivěji vyplněné.
 - * Hodnocení obsluhy
 - * Celkový dojem
- Přeskakovali úvodní text - Dát na zvláštní stránku (či zvláštní dokument) - tutorial. Stačí, když si uživatel přečte na webu nebo při prvním hodnocení.
- Opakovaně nevěděli adresu podniku
- Náповěda, co psát do pole u prvního dojmu - Například: „Je tu hrozně zakouřeno nebo usmívající se číšnice.“

- Popisky u skla jsou špatné (To dojde a 0,5l patří k míře)
- Celkový dojem podniku - Pomyslně rozdělit stupnici 0-10 na nedoporučuji a doporučuji, nevěděli, co je to za kritérium.
- Náповěda k charakteristice podniku - Těžké na vyplnění, nevěděli, co mají psát. Například: „Zkuste shrnout hodnocení jednotlivých kategorií: Slušné prostředí, fajn obsluha a dobré pivo.“ Může být v tutorialu.

Na základě těchto a dalších podnětů bych upraven tištěný dotazník, aby nebylo nutné předělávat ten webový několikrát dokola. Tištěný dotazník byl úspěšně otestován, opraveny ještě nějaké drobnosti a změny budou implementovány na web.

4.5 Vyhodnocení implementace

Byly implementovány všechny důležité části pro fungování Beermapy. V systému je 30 registrovaných ambasadorů, ale protože hodnotilo pouze 6 z nich, pozornost byla zaměřena více na ně a byla tak odsunuta tvorba samotné mapy podniků. V tomto ohledu je potřeba vytvořit jednodušší prostředí pro hodnocení - upravit webový dotazník, udělat ho responsivní pro mobilní telefony a zobrazit ambasadorům kandidátní podniky v mapě. Rozhodnutí nedělat web responsivní, z tohoto důvodu, dnes hodnotím jako nešťastné, ale těžko říct kolik času navíc by si to vyžádalo. Při implementaci se projevil dobrý návrh, ale vzhledem k přidávání dalších funkcí je potřeba neustále testovat a dbát na správný návrh, aby se příliš nezhoršila struktura kódu.

Kapitola 5

Doporučování podniků

Aby uživatelům bylo umožněno co nejjednodušší vyhledávání podniků, podle jejich preferencí, jsou použity algoritmy doporučující podniky, podobné těm, které již navštívili a hodnotili. Díky této funkci se stává Beermapa komplexním nástrojem při hledání ideálního podniku pro uživatele. Tato funkce je zároveň konkurenční výhodou. V kapitole bude popsáno jak takové systémy fungují, jaké jsou možnosti implementace, jaká implementace byla zvolena a s jakou úspěšností doporučuje.

5.1 Doporučovací metody

Pro doporučení předmětu, který by se nám mohl líbit a mohli bychom ho chtít koupit, se používají speciální metody a algoritmy. Dnes se s nimi můžeme setkat hlavně u e-shopů pod pojmem „Mohlo by se vám líbit“. Obzvláště oblíbené a využívané jsou u globálních prodejců, kterým pomáhají zvyšovat zisky i o desítky procent.

Doporučuje se vždy pro konkrétního zákazníka na základě jeho nákupů nebo hodnocení předmětů. Doporučení je provedeno pomocí hodnocení a nákupů ostatních zákazníků. Z toho vyplývá, že je potřeba mít dostatečné množství dat, jednak o konkrétním uživateli, abychom mohli poznat, co se mu líbí, a jednak mít dostatečné množství dalších uživatelů a dat o nich, aby bylo podle čeho doporučit. Požadavkem však je, aby kvalitní doporučení mohlo být vytvořeno i nad řídkými daty. V opačném případě, když je dat velké množství, nastává naopak problém s náročným výpočtem doporučení. Požadavkem je toto doporučení poskytovat v reálném čase a co nejrychleji, což není vždy jednoduché. Konkrétně u velkých e-shopů a služeb, jako jsou například Amazon a e-Bay, má výpočetní náročnost velký význam. U Amazonu, který má kolem 30 milionů zákazníků se výpočet doporučení v reálném čase nemůže protáhnout na víc než několik okamžiků [2]. Proto nás zajímá škálovatelnost doporučovacíh algoritmů.

Jsou tři společné přístupy k řešení doporučování: kolaborativní filtrování, shlukování a metody založené na vyhledávání [4] [7].

5.1.1 Kolaborativní filtrování

Kolaborativní filtrování vyhledává uživatele, kteří mají podobné návyky a doporučuje zboží podle nalezeného tzv. *nejbližšího souseda*. Z jeho předmětů vybere ten, který uživatel, kterému je doporučováno, ještě nezakoupil. V kolaborativním filtrování zákazníka reprezentuje N-dimenzionální vektor předmětů, které zakoupil nebo hodnotil. Pro vykompenzování

nejvíce nakupovaných položek, algoritmus typicky násobí vektor inverzním vektorem frekvence (počet uživatelů, kteří předmět zakoupili), takže se méně známé předměty dostanou do popředí. Zjišťovat podobnost vektorů můžeme různými způsoby. Běžně využívanou metodou je určení úhlu mezi nimi [2]:

$$\text{similarity}(\vec{A}, \vec{B}) = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \bullet \vec{B}}{\|\vec{A}\| * \|\vec{B}\|}$$

Následné vybrání konkrétního předmětu může být opět provedeno různými způsoby. Nejčastější technikou doporučení je podle počtu zakoupení každé z nich.

Při rozumném množství dat je relativně dobrá kvalita doporučení, ale velkou nevýhodou náročný výpočet hledání nejbližšího souseda, který probíhá online.

5.1.2 Shlukování

Shlukování přiřazuje uživatele či zákazníky do vytvořených uživatelských segmentů, podle průměrné podobnosti uživatelů v daném segmentu. Velký počet uživatelů je rozdělen na několik segmentů, ze začátku to bývá i náhodně, jen aby se pracovalo s menšími daty. Cílem je ale dosáhnout segmentů s podobnými uživateli, aby nový uživatel mohl být přiřazený do skupiny sobě podobných, proto jsou přístupy někdy kříženy a v rámci větších náhodných skupin jsou vytvořeny malé na základě podobnosti.

Výhodou tohoto přístupu je offline práce se segmenty a náročný výpočet tak může probíhat v předem určených časech a zařadit konkrétního uživatele do daného segmentu už je potom jednoduché. Opět může být provedeno pomocí určení úhlu mezi vektory. Samozřejmě, obecně je kvalita doporučení nižší, protože doporučujeme na základě podobnosti s velkým množstvím uživatelů. Pokud je použit křížený přístup, může se stát, že nám nejpodobnější uživatelé jsou v úplně jiných segmentech a neproběhne s nimi porovnání. Pokud bychom rozdělili uživatele do malých segmentů pouze podle podobnosti, shlukování by se výpočetní náročností blížilo kolaborativnímu filtrování a ztrácelo by tak svůj smysl.

5.1.3 Vyhledávací metody

Metody založené na vyhledávání, řeší doporučování pomocí hledání nejpodobnějších položek. Po nákupu nebo hodnocení vyhledávají další oblíbené položky stejného autora, umělce nebo na základě klíčových slov. Jestli tak uživatel nakoupí například knížku od 37signals - Restart, algoritmus mu doporučí knihu Práce na dálku, protože má stejného autora, ale může ještě přidat knížky ze stejného tématu (podle klíčových slov), například Startup za paketel. Pokud má uživatel více nákupů či hodnocení, přesnost samozřejmě stoupá a doporučení je kvalitní.

Problém nastává v případě velkého množství nákupů či hodnocení, protože by vyhledávací dotazy byly příliš komplikované. V takovém případě se vezme pouze část dat o uživateli, jenže to opět vede ke značnému snížení kvality doporučení. To se stává buď příliš obecným a doporučuje bestsellery nebo příliš úzce zaměřeným a doporučuje například pouze na základě autora. Jenže doporučení by mělo umožnit uživateli objevit pro něj neznámé, ale zároveň přesně podle jeho gusta a to v této chvíli přestává tento přístup plnit.

5.1.4 Item-to-item kolaborativní filtrování

Item-to-item je upravená varianta kolaborativního filtrování, která vyhledává podobnost mezi jednotlivými položkami v databázi, namísto mezi uživateli. Využívá ji například Amazon, protože umožňuje rychlé a kvalitní doporučení nad velkým množstvím dat. Zároveň vykazuje výbornou kvalitu nad řídkými daty. [5]

Pro hledání podobných produktů vytvoří algoritmus tabulku podobných položek. Šlo by vytvořit matici produkt-produkt ze všech předmětů v databázi, ale protože některé předměty nemají společné nakupující pro výpočet podobnosti, byl by to postup zbytečně paměťově náročný. Amazon proto používá následující algoritmus [2]:

```
1 foreach item  $I_1$  in product catalog do
2   foreach customer  $C$  who purchased  $I_1$  do
3     foreach item  $I_2$  purchased by customer  $C$  do
4       Record that a customer purchased  $I_1$  and  $I_2$ 
5     endforeach
6   foreach item  $I_2$  do
7     Compute the similarity between  $I_1$  and  $I_2$ 
8   endforeach
9 endforeach
10 endforeach
```

Takto vypočteme podobnost u předmětů, u kterých to lze, protože mají uživatele se společnými nákupy.

Samotná podobnost pak může být použita pro doporučení. Amazon v košíku, při procesu nákupu, zobrazuje produkty, které by se nám mohli hodit zakoupit s předmětem. Je to obdoba regálů u pokladen, kdy máme nutkání ještě něco přihodit do košíku. Tohle doporučení je vypočteno rychle a jednoduše tak, že se vezmou a zobrazí předměty s největší podobností k těm, které máme v košíku.

Důležitou částí item-to-item kolaborativního filtrování je výpočet podobnosti. Je několik možností, jak to udělat. Podle algoritmu uvedeného výše, je vidět, že vždy vypočítáváme podobnost pro jeden předmět se všemi ostatními, s kterými má společné hodnocení (uživatele, který ohodnotil obě tyto položky). Samotný výpočet podobnosti probíhá jako u klasického kolaborativního filtrování pomocí určení úhlu mezi vektory. Používá se výpočet kosinu [5]:

$$\text{sim}(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{A} \bullet \vec{B}}{\|\vec{A}\|_2 * \|\vec{B}\|_2}$$

Kde i a j jsou podniky, mezi kterými hledáme podobnost. Druhou možností je počítat korelaci mezi podniky [5]:

$$\text{sim}(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_i) (R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_i)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_j)^2}}$$

A vylepšenou variantou, vycházející z obou předchozích vzorců, je upravený výpočet pomocí kosinu, kde bereme v úvahu rozdíly mezi hodnocením jednotlivých uživatelů. Upravená kosinová podobnost kompenzuje tuto nevýhodu odečtením odpovídajícím uživatelským

průměrem [5]:

$$\text{sim}(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}}$$

Posledním a nejdůležitějším bodem je určení položek k doporučení. Výběr těch nepodobnějších se provede vybráním několika s nejvyšším predikovaným hodnocením. To se provádí už pro každého uživatele zvlášť. Na základě jeho předchozích hodnocení (nebo nákupů) se vezme předmět hodnocený i , ještě nehodnocený j , jejich podobnost a vynásobí se hodnocení předmětu i s podobností mezi předměty:

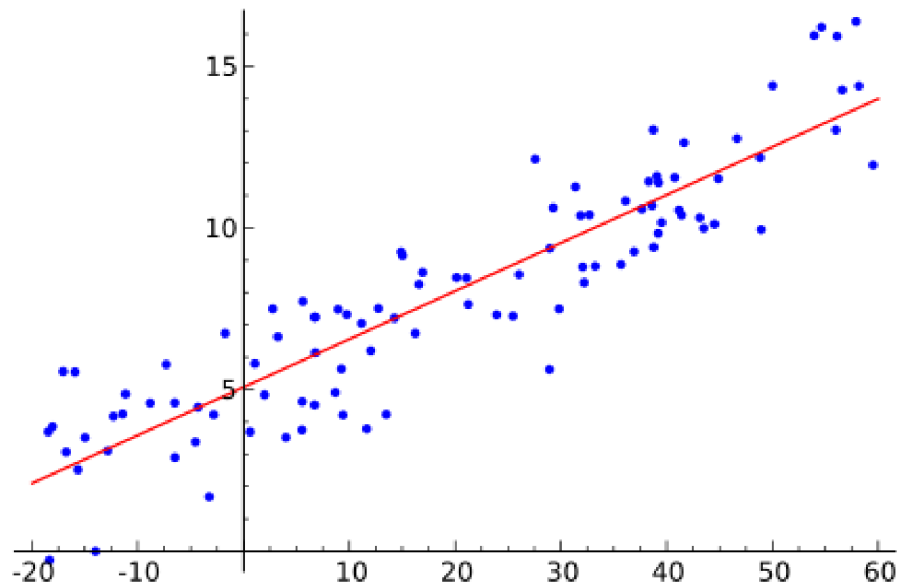
$$P_i = \text{sim}(i, j) * R_j$$

Tento přístup je vylepšený váženým součtem ze všech ohodnocených předmětů s vypočítanou podobností [5]:

$$P_i = \frac{\sum_{\text{allsimilaritem}, N} (s_{i,N} * R_{u,N})}{\sum_{\text{allsimilaritem}, N} (|s_{i,N}|)}$$

5.1.5 Zjištění preferencí pomocí lineární regrese

Další možností, určení doporučení, je získávat váhový vektor vlivu jednotlivých kategorií na subjektivní hodnocení uživatele. Tímto bychom zjistili, která kritéria jsou pro uživatele důležitější než jiná. Potom bychom mu mohli doporučit podnik vyčnívající v tomto směru. Lineární regrese aproximuje soubor hodnot v grafu pomocí přímky (obrázek 5.1). Pokud tedy máme dvě sady hodnot můžeme mezi nimi hledat vzájemnou závislost.



Obrázek 5.1: Lineární regrese (Převzato z Wikipedie)

Protože hledáme ideální proložení přímky body, vychází rovnice lineární regrese z rovnice přímky [8]:

$$y = a + bx + e$$

Parametr a značí posun přímky a parametr b její sklon. Ten nás zajímá především. Malé e je chyba regrese. V případě, že máme kritérií víc, výpočet bude pomocí vícenásobné regrese [6]:

$$y_1 = a + b_1x_{11} + \dots b_kx_{1k} + e_1$$

Tím bychom dostali váhový vektor určující důležitost jednotlivých kritérií pro daného uživatele. Vypočítané parametry nabývají hodnot v intervalu $< 0; 1 >$. Pokud bychom opět chtěli dostat známku, která by rozhodovala zda podnik bude doporučený, mohli bychom z vektoru udělat koeficienty pro násobení jednotlivých kritérií. To by bylo provedeno jednoduchým odečtením každého vypočítaného parametru od čísla 2. Po vynásobení získaného koeficientu s výslednou známkou kritéria bychom získali upravené hodnocení, které by muselo splnit určitou bodovou hranici, aby mohl být podnik doporučen. Základní princip je tedy v tom, že čím méně důležité kritérium a tím menší vypočtená váha, tím vyšší koeficient pro násobení.

Pokud bychom například získali váhu piva 0,5, což by znamenalo, že není pro uživatele moc důležité, výsledná známka kritéria, například 6,5 by se násobila koeficientem 1,5 a získali bychom tedy novou známku $1,5 * 6,5 = 9,75$. Pokud by byla hranice pro doporučení hodnota 8, podnik by byl doporučen.

Protože to ale nejsou ověřené postupy, bylo by potřeba domněnky řádně otestovat.

5.2 Návrh způsobu doporučování

Nejvhodnějším způsobem určování doporučení se tak jeví metoda kolaborativního filtrování *item-to-item*, která odstraňuje nevýhodu výpočtu klasického kolaborativního filtrování, kdy je potřeba velkého množství uživatelů, aby mohl být nalezen odpovídající *nejbližší soused*, ale má výhodu shlukování, u kterého hlavní náročný výpočet probíhá offline. Protože budeme mít ze začátku malé množství dat, všechny ostatní varianty by byly pravděpodobně příliš nepřesné.

Pro doporučování pomocí *item-to-item kolaborativního filtrování* potřebujeme spočítat podobnost mezi jednotlivými podniky, které mají společné hodnocení a poté podle hodnocení konkrétního uživatele určit doporučení jemu na míru. Pro oba body je potřeba určit, s jakou známkou bude počítáno. Pokud bychom vzali vypočítanou výslednou známku, která je relativně objektivní, neprozradilo by nám to, jak se uživateli podnik pocitově líbil. Mohli bychom vzít jednotlivá kritéria, která by byla víc vypovídající, ale nejlepší variantou se jeví známka celkového dojmu z podniku, která je subjektivní a uživatel do ní tak promítne vlastní preference a na rozdíl od jednotlivých kritérií ji můžeme použít tak jak je. Zde může nastat problém, pokud by uživatelé nehodnotili zodpovědně, ale spíše podle toho jak se v podniku opili. To naštěstí vyvrací tabulka 5.2, která porovnává subjektivní hodnocení a celkovou známku. Jako vzorek pro ukázkou bylo vybráno prvních deset hodnocení.

Průměrný rozdíl mezi všemi 50 hodnoceními, které máme hotové, je 0,89 bodu a korelace mezi těmito dvěma poli hodnot je 0,96, což ukazuje na vysokou závislost. Pokud by se toto číslo do budoucna změnilo, bylo by potřeba zvážit doporučení na základě kritérií.

Pro výpočet podobnosti byl zvolen třetí způsob - křížený přístup kosinu a korelace, který si z nich bere to dobré a měl by tak vykazovat nejlepší výsledky.

Id	Subjektivní hodnocení	Celková známka	Rozdíl
1	9	8,7	0,3
2	7	7,9	0,9
3	9	8,9	0,1
4	8	7,5	0,5
5	7	7,5	0,5
6	7	7,5	0,5
7	7	7,8	0,8
8	4	6,3	2,3
9	9	8,4	0,6
10	7	8,3	1,3

Tabulka 5.1: Rychlost algoritmů inteligentního řazení

5.3 Implementace doporučení

V databázi byla vytvořena tabulka *similarity*, do které jsou postupně přidávány podniky, které mají společné hodnotitele, takže pro ně může být určena podobnost. Hledání těchto vztahů je provedeno přesně pomocí algoritmu 10. SQL dotazy jsou hledány podniky s jejich hodnoceními a pomocí cyklů jsou nalezena společná hodnocení. Podobnost je potom iterativně vypočítána pomocí zmíněného vztahu křížícího kosinus a korelaci, ale pouze pro podniky, které mají alespoň dvě vazby. V případě jedné vazby byl výsledek buď -1 nebo 1, což jsou krajní hodnoty a značí silnou přímou nebo nepřímou závislost. Síla závislosti tak v případě jedné vazby neodpovídala.

Jak jsou jednou vytvořeny podobnosti, které jsou v rámci testování zatím spouštěny ručně v administraci, vypočítání podobností je už rychlé. Probíhá v reálném čase, při načítání stránky. Nejdříve se najdou uživatelem nehodnocené podniky a cyklem se pro každý podnik počítá predikce hodnocení. Prvně jsou vyhledány podobnosti, z nich se vezmou pouze ty, které mají hodnotu v intervalu $(0; 1 >$, protože pouze ty výpočet predikce zpřesňují [7]. To bylo ověřeno testováním. Pokud bychom brali celý interval $< -1; 1 >$ hodnoty by vůbec neodpovídaly a pokud bychom brali pouze vysokou podobnost například nad 0,8, ochuzovali bychom se o další zpřesnění predikované hodnoty. Což je patrné ze vzorce pro výpočet predikce. Nakonec je podmínkou vybráno hodnocení nad 6,5 a tyto podniky zobrazeny. Hodnota 6,5 je nastavena experimentálně, až přibude více hodnocení počítá se s jejím zvýšením.

5.4 Testování doporučení

Jelikož máme k dispozici pouze 50 hodnocení od 6 uživatelů, probíhalo testování doporučení ručně, odebráním několika ambadorských hodnocení, následným zobrazením doporučení i s predikovanou známkou (která uživatelům viditelná nebude) a její porovnání se známkou jakou uživatel reálně udělil. Nemohlo být odebráno velké množství dat, v případě odstranění společných hodnocení by chyběly vazby pro výpočet podobnosti. Proto bylo náhodně odstraněno pouze 7 hodnocení, u podniků s nejvyšším počtem hodnocení. Jak se změnilo hodnocení je k vidění v tabulce 5.2.

Po dalších dvou odebraných odebraných hodnocení, už nebyla zobrazena žádná ověřitelná doporučení. I při tak malém vzorku dat byla průměrná odchylka doporučení 0,76

Uživatel	Podnik	Subjektivní hodnocení	Predikce
Daniel	Výčep na stojáka Jakubák	9	10
Ondřej	Pivní opice	8	7
Tomáš	U třech čertů	9	9
Tomáš	East Village Bar & Diner	9	8
Jiří	Výčep na stojáka Pekanda	9	8
Radek	Výčep na stojáka Pekanda	9	Nedoporučilo
Michal	Zelená kočka Pivárium	7	7,6

Tabulka 5.2: Rychlost algoritmů inteligentního řazení

bodů. Vzhledem k tomu, že nepotřebujeme predikovat známku, ale hlavně doporučit podniky, které se budou pravděpodobně líbit, tedy ty s nejvyšším hodnocením, byla přesnost dobrá. Korelace mezi reálnými hodnoceními a predikovanými hodnotami je 0,58, což není špatná, ale ani výborná hodnota. Ta se však pravděpodobně bude zvyšovat s přibývajícím hodnoceními.

Kapitola 6

Závěr

Zadáním této práce bylo navrhnout a vytvořit webovou službu pro poznávání okolních pivnic, tak aby pomocí sesbíraných hodnocení bylo možné nabídnout tipy uživatelům, kam zajít. Vytvořenou webovou aplikaci pak otestovat a reflektovat změny. A v poslední řadě implementovat algoritmus na doporučování podniků. Postupnými kroky tak byla vytvořena funkční webová aplikace, při jejíž vývoji byla snaha poctivě splnit všechny kroky správného vývoje webové aplikace. Vývoj probíhal v iteracích, kdy v každém dalším kroku byly opraveny chyby a přidány nové funkce, tak jak je popsán proces Lean StartUpu. Po celou dobu vývoje probíhalo testování s potenciálními uživateli, aby byla služba zdokonalována.

Aplikace již delší dobu běží na <http://www.beermapa.cz>. Protože se ale doposud nepodařilo nasbírat dostatečné množství hodnocení, i přes vyvinutou snahu získáváním ambasadourů mezi známými a na pivním festivalu, nejde jednoznačně určit, které podniky jsou kvalitní, a proto nebyl zatím vytvořen přehled pivnic. Protože se ale sami snažíme hodnocení získávat, bude vytvoření přehledu pivnic jedním z nejbližších úkolů, aby mohl být přidáván podnik po podniku a Beermapa tak přitáhla větší pozornost veřejnosti. Aktuálně je největší snahou více motivovat ambasadory k hodnocení podniků.

V rámci práce byl navržen a implementován algoritmus na doporučování podniků na základě předchozího hodnocení. Algoritmus byl otestován a i na malém vzorku 50 hodnocení vykazoval dobré výsledky. Kvalita jeho doporučení bude testována i nadále a v případě potřeby bude algoritmus upraven. Jedním z dalších úkolů je otestování doporučování podniků na základě lineární regrese a porovnání výsledků s implementovaným item-to-item kolaborativním filtrováním.

Práce tak splnila svůj účel, posunula Beermapu výrazně dále, tak aby v nejbližší době mohly být zveřejněny první kvalitní pivnice, což je hlavním cílem. Během práce jsem se naučil od základu navrhovat a vytvořit webovou službu a hlavně vedení práce tak, aby se neustále posouvala dále, čehož si cením nejvíc.

Literatura

- [1] GRUDL, D., KOLEKTIV VÝVOJÁŘŮ. *Dokumentace webového frameworku Nette* [online]. 2008- [cit. 2013-01-17]. Dostupné z: <<http://doc.nette.org/cs/>>.
- [2] LINDEN, G., SMITH, B. a YORK, J. Amazon.com recommendations: item-to-item collaborative filtering. *Internet Computing, IEEE*. Jan 2003, roč. 7, č. 1. S. 76–80. ISSN 1089-7801.
- [3] MAURYA, A. *Running Lean: Iterate from Plan A to a Plan That Works*. [b.m.]: O'Reilly Media, 2012. ISBN ISBN 978-1449305178.
- [4] SARWAR, B., KARYPIS, G., KONSTAN, J. et al. Analysis of Recommendation Algorithms for e-Commerce. In *Proceedings of the 2Nd ACM Conference on Electronic Commerce*. New York, NY, USA: ACM, 2000. S. 158–167. EC '00. Dostupné z: <<http://doi.acm.org/10.1145/352871.352887>>. ISBN 1-58113-272-7.
- [5] SARWAR, B., KARYPIS, G., KONSTAN, J. et al. Item-based Collaborative Filtering Recommendation Algorithms. In *Proceedings of the 10th International Conference on World Wide Web*. New York, NY, USA: ACM, 2001. S. 285–295. WWW '01. Dostupné z: <<http://doi.acm.org/10.1145/371920.372071>>. ISBN 1-58113-348-0.
- [6] UC SANTA BARBARA, DEPARTMENT OF GEOGRAPHY. *Multiple Regression in Matrix Form* [online]. [cit. 2014-07-22]. Dostupné z: <http://www.geog.ucsb.edu/~joel/g210_w07/lecture_notes/lect09/oh07_09_2.html>.
- [7] WANG, J., SARWAR, B. a SUNDARESAN, N. Utilizing Related Products for Post-purchase Recommendation in e-Commerce. In *Proceedings of the Fifth ACM Conference on Recommender Systems*. New York, NY, USA: ACM, 2011. S. 329–332. RecSys '11. Dostupné z: <<http://doi.acm.org/10.1145/2043932.2043995>>. ISBN 978-1-4503-0683-6.
- [8] YALE UNIVERSITY, DEPARTMENT OF STATISTICS. *Linear Regression* [online]. 1997-1998 [cit. 2014-07-22]. Dostupné z: <<http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm>>.

Dodatek A

Obsah CD

Seznam souborů na přiloženém cd včetně popisu:

- **xsvobo33_bp.pdf** - tato bakalářská práce ve formátu pdf
- **xsvobo33_bp_zdroj.zip** - latexové zdrojové kódy této bakalářské práce
- **xsvobo33_beermapa.zip** - zdrojové kódy webové aplikace spolu s databází
- **xsvobo33_plakatek.pdf** - plakát o projektu
- **xsvobo33_readme.txt** - návod a popis jednotlivých archivů

Dodatek B

Tabulka rozdělení práce

Kdo	Co
Já	Návrh webu a jeho stránek
	Návrh a tvorba databáze
	Testování
	Aktivity - vytváření i zobrazení
	Profil ambasadora
	Ambadorské tabulky a aktivita
	Počítání známek podniků
	Počítání podobností a doporučování podniků
	Sbírání e-mailů, synchronizace s Mailchimpem
	Automatické maily pro zpětnou vazbu přes Mandrill
Kolega Radek	Úvodní stránka a šablona webu
	Základ administrace
	Kandidátní podniky
	Statické stránky
	API pro mobilní aplikaci

Dodatek C

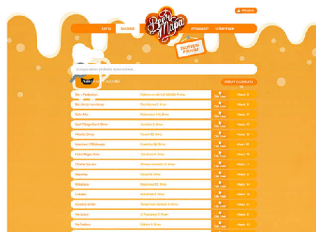
Plakátek k projektu



StartUp jako diplomka **Beermapa.cz**

Webová služba pro sdílení a poznávání okolních pivnic

Česká republika je pivní národ a podnik s dobrým pivem je u nás spousta, ale ještě více je těch méně kvalitních a ne vždy víme, kde takové podniky hledat. Proto chceme mapovat podniky s dobrým čepovaným pivem daného města či regionu, které jsou charakteristické svoji výjimečností a unikátními vlastnostmi. Tyto podniky jsou důkladně vybírány na základě hodnocení v široké oblasti kritérií. Skloubením všeho výše zmíněného dochází k selekci "kvalitních" podniků od těch "obyčejných". Beermapa je pro všechny, ale především ji využijí lidé nepřítli znalí lokální pivní scény, tedy studenti vysokých škol, kteří pocházejí z celé ČR, ale i podnikatele na služebních cestách a turisté. Pokud se rozhodnou večer někam zajít a užít si to v té správné atmosféře, mohou použít Beermapu, vybrat sami podle vlastních preferencí, a nespolehat se zda jim někdo dobře poradí. Hlavním nástrojem Beermapy budou internetové stránky, která je předmětem této práce, ale i aplikace pro chytré telefony, která zvýší komfort a pohodlí uživatele Beermapy.



Aktuálně probíhá sbírání dat o podnicích. Data o podnicích, sbírá vybraná skupina uživatelů, tzv. ambasadorů. Ambasadori ale ještě není dostatek, a proto budeme rádi, když se připojíte i Vy. Pokud nechcete být ambasadori, můžete ke kvalitnější pivní scéně pomoci alespoň doporučením podniku kandidujícího na to stát se podnikem beermapovým. Kandidáti poté sbírají hlasy a ty, u uživatelů nejoblíbenější, jsou ambasadori hodnoceni jako první. Ambasadori hodnotí podle předpřipravených dotazníků obsahujících kritéria, která by správná pivnice měla splňovat. Ke každému pak přiřazují známku v rozsahu 0-10. Na základě hodnocení je vypočítána celková známka podniku, podle které bude v budoucnu určeno, který podnik je kvalitní a bude tak doporučen všem uživatelům.

Aby uživatelé mohli objevovat ty správné kvalitní podniky, přidanou hodnotou je doporučování podniků na základě podobnosti. Všichni to známe z internetových obchodů pod pojmem „Mohl by se vám líbit“. Základní přístup, který hledá vám nejpodobnějšího uživatele a podle něj doporučí. Amazon vylepšil a vyhledává předměty, které jsou si podobné. Tento způsob se jmenuje Item-to-item kolaborační filtrování. V jednoduchosti funguje tak, že se najdou všechny dvojice podniků, které mají společné hodnotitele. Na základě známek udělených těmito dvěma podnikům se určí jejich podobnost. Pak už stačí vzít nehodnocené podniky jednoho uživatele, vzít spočítané podobnosti s těmi podniky, které uživatel hodnotil a vynásobit je udělenými známkami. Tak dostaneme předpověď uděleného hodnocení.

Snad i vám nějaký ten podnik budeme moct doporučit. Ať už to není jedno.



www.beermapa.cz

Autor: Ondřej Svoboda
ssvob33@stud.fit.vutbr.cz
Vedoucí Ing. Igor Szóke, Ph.D.
FIT VUT
2014

viz. prilohy.tex