

Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií

Ilkovičova 2, 842 16 Bratislava 4

Crowdex - systém pre manažment mikroúloh

Záverečná správa

Vedúci práce: Ing. Michal Kompan, PhD.

Členovia tímu: Bc. Dušan Cymorek, Bc. Peter Gašpar, Bc. Vladimír L'alík,
Bc. Michal Polko, Bc. Miroslav Šafárik (ZS), Bc. Slavomír Šárik,
Bc. Štefan Šmihla

Číslo tímu: Tím č. 5 IS-SI

Akademický rok: 2014/2015

1 Abstrakt

V tomto dokumente predstavujeme nový prístup k manažmentu a riešeniu mikroúloh, ktorý je založený na technológii Webu 2.0 a princípoch crowdsourcingu. Základom systému Crowdex je programovací jazyk Ruby spolu s aplikačným rámcom Ruby on Rails. Pre ukladanie dát slúži SQL databáza PostgreSQL v kombinácii s NoSQL riešeniami ElasticSearch a Redis pre zvýšenie výkonnosti aplikácie.

Cieľom systému Crowdex je zjednodušenie procesu manažmentu a urýchlenie realizácie malých a jednoduchých úloh – mikroúloh rôznym skupinám ľudí. Študenti na univerzitách, učitelia či výskumníci často potrebujú overiť svoje hypotézy alebo získať výsledky pre svoje záverečné práce podložené dostatočným počtom ľudí. Podobnému problému čelia aj firmy, ktoré potrebujú vykonať nejakú úlohu a vystačia si s externistom. Môže ísť napríklad o testovanie nového používateľského rozhrania aplikácie, prípadne vyplnenie dotazníka ohľadom spokojnosti s poskytovaním služieb.

Proces hľadania vhodných kandidátov na splnenie týchto úloh je neraz náročný a zdĺhavý. V súčasnosti prebieha prostredníctvom sociálnych sietí, diskusných fór, prípadne oslovovania svojich známych. Tento spôsob so sebou prináša aj rôzne problémy, či už ide o nedostatočný počet ľudí ochotných vykonať úlohu pre cudzieho človeka, ale tiež aj vhodnosť ľudí na konkrétnu úlohu. Väčšinou totiž nevieme vopred zistiť informácie o veku, dosiahnutom vzdelaní alebo znalostiach, čo môže viesť až k zahodeniu získaných výsledkov. V neposlednom rade je problémom aj nedostatočná motivácia, ktorá môže spôsobiť skreslenie výsledkov, napríklad pri rýchлом vyplnení dotazníka náhodnými odpoveďami.

Našou reakciou na spomenuté problémy je systém Crowdex, ktorý pokrýva väčšinu činností spojených s manažmentom mikroúloh, od ich vytvorenia, cez získanie účastníkov a realizáciu úloh, až po zber výsledkov a ich vyhodnotenie a odmenenie účastníkov. Okrem manažmentu mikroúloh ako takého, Crowdex kladie veľký dôraz na motiváciu a zabezpečenie kvalitných participantov pre zadávateľov úloh – a to všetko prostredníctvom integrovaného vzájomného hodnotenia zadávateľov i účastníkov úloh. Okrem hodnotenia je možné účastníkov motivovať aj odmenou vo forme kreditu – virtuálnej meny, ktorú je možné zameniť v integrovanom elektronickom obchode za rozličné odmeny od zadávateľov – či už vo forme malých reklamných predmetov, darčekových poukazov alebo iných hodnotnejších odmien.

Výsledkom našej práce je webová aplikácia, ktorá nielenže vo výraznej miere prispieva k uľahčeniu a urýchleniu manažmentu mikroúloh, skvalitneniu študentských i výskumných prác, ale taktiež povzbudzuje ľudí k vykonávaniu mikroúloh za atraktívne odmeny. Taktiež v sebe zahŕňa biznis potenciál vyplývajúci najmä z poskytovania platených prémiových funkcionalít, akými sú možnosť vytvorenia väčšieho počtu úloh, vyšší limit pre počet participantov, pokročilé štatistiky z dotazníkov či v budúcnosti plánované sledovanie emócií a pohľadu používateľa.

2 Prehľad riešenia

Crowdsourcing sa za posledné roky stal veľmi populárnou alternatívou ku zaužívaným spôsobom del'by práce. Spočíva v pridelení určitej činnosti alebo práce (všeobecne nazývanej aj „úloha“) davu, teda náhodnej skupine osôb. Zadávateľia úloh posudzujú získané výsledky a spravidla dav odmenia - konečný výber sa najčastejšie ponecháva na nich.

V súčasnosti existuje viacero crowdsourcingových platforiem, pričom každá sa špecializuje na rôznu oblasť a má tiež svoje výhody aj nevýhody. Najpopulárnejšou službou je *Amazon Mechanical Turk*¹, ktorá ponúka priestor pre úlohy objavovania znalostí, ktoré neumožňujú strojové spracovanie, a preto je potrebné zapojenie ľudskej inteligencie pri riešení týchto úloh. Napriek svojej popularite, *Mechanical Turk* nezahŕňa hodnotenie používateľov (a ani zadávateľov) alebo diskusie k úlohám. Ďalším často používaným riešením sa stal *Crowdfunder*² zameraný primárne na biznis zákazníkov, nakoľko použitie služby je pre bežných študentov nákladné a nevyhovujúce. Medzi ďalšie podobné systémy patria doménovo špecifické služby ako *TaskRabbit*³ (upratovanie, st'ahovanie) alebo *Gigwalk*⁴ (zameraný na plánovanie predaja). Spoločnou nevýhodou uvedených systémov je tiež nedostupnosť platforiem mimo určené krajiny (*Mechanical Turk* je napríklad určený len pre používateľov z USA) a z pohľadu Slovenska minimum participantov z našich geografických končín.

Uvedené nevýhody spolu so vždy existujúcim priestorom na zlepšenie nás motivovali ku vytvoreniu nového a inovatívneho prístupu pre komplexný **manažment mikroúloh**, ktorý je základom nášho systému **Crowdex**. Pri jeho vývoji sme sa snažili spojiť to najlepšie z dostupných riešení s našimi vlastnými novými nápadmi a vylepšeniami, a priniesť tak webovú aplikáciu, ktorá dokáže splniť nielen požiadavky študentov a biznis zákazníkov, ale taktiež zapojiť do hry aj širokú verejnosť. Systém Crowdex pokrýva väčšinu činností spojených s manažmentom mikroúloh, od ich vytvorenia, cez získanie účastníkov a realizáciu úloh, až po zber výsledkov a ich vyhodnotenie a odmenenie účastníkov úlohy.

Základom každého väčšieho informačného systému je správa používateľov. Prvý a zároveň najdôležitejší krok predstavuje **registrácia používateľa**. Štandardný interný systém prihlásenia cez Crowdex účet sme obohatili aj o možnosť použiť existujúce účty z iných služieb. Pre širokú verejnosť sme pripravili prihlásenie prostredníctvom sociálnej siete *Facebook*, s ktorým sa v súčasnosti stretávame takmer v každej webovej a mobilnej aplikácii. Svoju pozornosť sme však zamerali aj na zamestnancov a študentov STU, ktorí môžu použiť svoje prihlasovacie údaje z univerzitného systému (*AIS*). Študenti FIIT STU navyše získavajú ďalšiu výhodu, a to vo forme automatického vyplnenia odborných zručností vo svojom profile na základe úspešne absolvovaných predmetov. Tieto zručnosti vychádzajú priamo zo štruktúry odborných predmetov a zahŕňajú znalosti, ktoré študenti ich úspešným absolvovaním nadobudli. Využitie overených znalostí z prostredia univerzity považujeme v doméne správy mikroúloh za významnú a inovatívnu myšlienku.

¹<http://www.mturk.com>

²<http://www.crowdfunder.com>

³<http://www.taskrabbit.com>

⁴<http://www.gigwalk.com>

V kontexte mikroúloh je dôležitá aj **motivácia používateľov** k riešeniu úloh. Existujúce systémy (ako napr. *Amazon Mechanical Turk*) ponúkajú za riešenie úloh peňažnú odmenu, no tá je poskytovaná len v minimálnom množstve (jednotky centov za jednu úlohu), čo pre používateľov často nepredstavuje dostatočnú motiváciu pre riešenie týchto úloh. V našom systéme sme zaviedli niekoľko motivačných prvkov. Jedným z nich je **kredit** - virtuálna mena, ktorú používatelia získavajú za riešenie úloh. Získaný kredit je následne možné vymeniť za rôzne odmeny v **integrovanom elektronickom obchode**. Môže ísť napríklad nielen o reklamné predmety, darčekové poukazy, ale aj o hodnotnejšie odmeny. Pri odmeňovaní participantov si zadávateľ môže zvoliť niekoľko spôsobov rozdelenia stanoveného rozpočtu úlohy a tiež určiť, ktorých participantov odmení. Problémom konkurenčných riešení je aj časté nevyplatenie sľúbenej odmeny zadávateľom úlohy. Aby sme predišli tejto skutočnosti, pri vytváraní úlohy je počet kreditov stanovených ako rozpočet úlohy odpočítaný z dostupného kreditu zadávateľa. V prípade, že nedôjde k vyplateniu odmeny do 7 dní od ukončenia úlohy, systém pristúpi k automatickému rovnomernému vyplateniu odmien všetkým participantom.

Častý problém, s ktorým sa stretávame v rámci manažmentu mikroúloh, tiež predstavuje **kvalita odovzdaných výsledkov** vzhľadom na nedostatočnú motiváciu a prípadnú neserióznosť používateľov. V našom systéme riešime tento problém prostredníctvom komunity. Zadávatelia a participantia sa môžu **vzájomne hodnotiť** za kvalitu svojej práce. Po odovzdaní úlohy vyzývame participantov, aby ohodnotili zadávateľa úlohy - napr. z dôvodu spokojnosti so samotnou úlohou alebo s vyplatením odmeny. Podobne aj zadávateľom úloh dávame možnosť ohodnotiť participantov na základe ich odovzdaných výsledkov. Na základe takéhoto hodnotenia si môžu zadávatelia úloh vybrať tých najvhodnejších participantov a participantia môžu prehodnotiť zapojenie sa do úloh vytvorených konkrétnym zadávateľom. Hodnotenie je zároveň dôležitá súčasť zabudovaného systému **odporúčaní úloh a participantov**.

Medzi časté funkcionality, s ktorými sa stretávame aj v iných systémoch, patrí **vyplňanie dotazníkov**. V našom systéme poskytujeme v rámci jednej úlohy možnosť vytvoriť viacero dotazníkov, čím umožňujeme rozdeliť na logické celky aj väčšie množstvo otázok a minimalizovať tak riziko, že bude používateľ frustrovaný rozsiahlosťou jedného dotazníka. V porovnaní s existujúcimi systémami pre prácu s dotazníkmi (ako napr. *Google Forms*) poskytujeme viacero pohľadov na výsledky dotazníkov. Základným pohľadom je prehľad odpovedí pre každého participanta samostatne. Pokročilejšiemu pohľadu zodpovedajú **agregované štatistiky**, ktoré poskytujú grafický prehľad o počte a percentuálnych podieloch zvolených možností v prípade otázok, v ktorých participant vyberá z poskytnutých možností. Špecifickým prípadom sú otázky s možnosťou výberu viacerých odpovedí, pri ktorých poskytujeme prehľad nielen podľa jednotlivých možností, ale aj podľa kombinácií, ktoré participantia zvolili. V prípade, že zadávateľ úlohy potrebuje ďalej elektronicky spracovať odpovede, poskytujeme mu možnosť **exportovania odpovedí** vo formátoch CSV alebo JSON. Oproti existujúcim systémom sme ho obohatili aj o údaje o participantoch, ktoré poskytli vo svojom profile, ako napr. vek alebo pohlavie. Takto môžeme poskytnúť kvalitnejšie údaje napr. pre sociologické prieskumy.

Okrem štandardného riešenia úloh v prostredí našej webovej aplikácie poskytujeme aj REST-ové aplikačné rozhranie (**API**). To využívame najmä pre samotné dokončenie úloh, vďaka čomu je náš systém ľahko integrovateľný aj s inými externými systémami.

3 Realizácia riešenia

Vstupnú bránu do nášho systému pre prihláseného používateľa tvoria rôzne aktuálne a štatistické informácie v podobe interaktívnych grafov, ktoré mu poskytnú rýchly prehľad o aktivitách v celom systéme. **Používateľské rozhranie** systému je plne **lokalizované do dvoch jazykov** - slovenčiny a angličtiny - s možnosťou jednoduchého rozšírenia o ďalšie jazyky. Vďaka tomu je portál prístupný nielen domácim, ale taktiež i zahraničným študentom a širokej verejnosti.

Neodmysliteľnou súčasťou dnešných webových aplikácií je **profil používateľa**. Okrem základných informácií o pohlaví, veku či záujmoch sa nachádzajú v profile aj informácie o jeho znalostiach, hodnotení a vytvorených a vyriešených úlohách. Tieto informácie môžu poslúžiť zadávateľom pri výbere participantov pre riešenie ich úloh. Vybraných používateľov môžu zadávateľia pozvať na riešenie úlohy prostredníctvom **integrovaných pozvánok**. Pozvánky je možné zaslať aj ľuďom, ktorí nie sú registrovaní v našom systéme, a to prostredníctvom emailu. Tento email zároveň umožňuje prejsť procesom zrýchlenej registrácie, vďaka čomu vieme ľahko získať nových používateľov. Nezabudli sme pri tom ani na možné obavy o súkromie, preto umožňujeme zúčastniť sa na úlohe aj ako **anonymný participant**.

Proces schvaľovania nových participantov sme sa snažili zadávateľovi uľahčiť aj pomocou definovania **podmienok účasti** na riešení úlohy, ktoré musí participant spĺňať. Okrem veku, pohlavia či dosiahnutého vzdelania, náš systém umožňuje aj výber povinných odborných znalostí. Podmienky účasti umožňujú zadávateľovi úlohy získať presne takých participantov, ktorých potrebuje pre vykonanie svojej úlohy. Medzi ďalšie pomocné prvky patrí zabudované **odporúčanie participantov**. To zohľadňuje nastavenia a parametre každej úlohy a tiež predchádzajúcu aktivitu používateľa v systéme - hodnotenia a vyplatené odmeny.

Vďaka pokročilému systému **notifikácií** používatelia nezmeškajú v Crowdexe významné udalosti (medzi ktoré patrí napríklad aj pridanie novej úlohy alebo pridelenie odmeny používateľovi, ktorý dokončil úlohu). Používateľ si môže podrobne nastaviť zobrazovanie jednotlivých typov notifikácií a určiť, ktoré z nich sa budú zasielať na jeho emailovú adresu.

Systém Crowdex je možné spravovať pomocou **rozhrania pre administrátorov**, ktoré zahŕňa komplexnú správu používateľov, úloh, obchodu a objednávok. Rozhranie tiež obsahuje aj **vlastný modul pre manažment chýb**, prostredníctvom ktorého dokážeme identifikovať a spravovať vzniknuté chyby počas prevádzky systému. V neposlednom rade administrátor prostredníctvom tohto rozhrania dokáže okamžite vyriešiť nahlásený nevhodný obsah od používateľov.

Pod kapotou systému Crowdex sa nachádza populárny aplikačný rámec *Ruby on Rails*. Výstupom je preto webová aplikácia založená na štandardných **technológiách** ako *HTML*, *CSS*, *JavaScript* a rámcoch *Bootstrap* a *jQuery*. V oblasti dizajnu sme sa od papierových návrhov dopracovali až ku jednotnému a ucelenému používateľskému rozhraniu vychádzajúcemu z princípov a pravidiel *Material dizajnu*⁵. Ako primárne dátové úložisko sme použili SQL databázu *PostgreSQL* v spojení s NoSQL riešeniami *ElasticSearch* pre rýchle textové vyhľadávanie a *Redis* ako rýchlu vyrovnávaciu pamäť.

⁵<http://www.google.com/design/>

Zdrojový kód sme medzi sebou zdieľali pomocou verziovacieho systému *git* spolu s nástrojom *Github* pre manažment repozitárov a prehliadku zdrojového kódu. Na automatické vyhodnocovanie kvality zdrojového kódu, ako aj pachov v kóde, sa nám osvedčila služba *Code Climate*⁶. Pre zabezpečenie kontinuálnej integrácie a vykonávania integračných testov sme si zvolili nástroj *Travis CI*⁷. Webový server spolu s aplikáciou je neustále monitorovaný službou *New Relic*, vďaka ktorej dokážeme pružne reagovať na výpadok funkčnosti aplikácie a taktiež odhaliť prípadne úzke hrdlo vo výkone aplikácie.

Registráciu používateľa a používateľské účty sme okrem klasických opatrení (ako napr. captcha) **zabezpečili aj proti rôznym pokusom o útoky**. Používateľovi povolíme pre jednu emailovú adresu v priebehu piatich minút zadať nesprávne heslo najviac päťkrát, následne sa jeho účet dočasne zablokuje. Pre tento typ zabezpečenia sme využili rýchlosť úložiska *Redis*, ktoré zastrešuje réžiu pamätania si počtov prihlásení v danom časovom okamihu. Používateľ si môže taktiež prispôbiť nastavenia zabezpečenia, čo zahŕňa najmä spravovanie aktívnych zariadení a prihlásení k účtu.

Systém Crowdex bol počas obdobia vývoja niekoľkokrát podrobený **penetračnému testovaniu** a taktiež testu úmyselného modifikovania HTTP požiadaviek s dôrazom nato, aby každú jednu akciu v systéme mohol vykonať len oprávnený používateľ. Systém Crowdex bol viackrát podrobený aj testovaniu voči útokom zameraným na znepriístupnenie systému za použitia penetračných nástrojov *Siege*, *AB*, *Hulk*. Nedostatky sme odstránili modulom *Rack::Attack*, ktorý limituje množstvo vyslaných požiadaviek za sekundu pre každú unikátnu IP adresu. Systém sme testovali aj proti najčastejším typom útokov ako napr. *SQL Injection*, *Cross-Site Scripting*, *Cross-Site Request Forgery* a *Path-Traversal* a ďalšie. Vďaka týmto testom sme mohli prispôbiť bezpečnostné nastavenia na serveri a zvýšiť tak odolnosť systému voči rôznym typom útokov.

V priebehu vývoja systému sme **spolupracovali s reálnymi používateľmi**, čo nám pomohlo získať **spätnú väzbu**, ktorá sa pretavila do zlepšenia systému. Počas letného semestra sme systém odovzdali na **testovanie** našim známym, priateľom, rodinným príslušníkom i spolužiakom na fakulte. Systém mali možnosť vidieť tiež kolegovia z Fakulty sociálnych a ekonomických vied, ktorí prejavili záujem o jeho využitie pri svojom výskume. Naš systém sme podrobili aj kvalitatívnemu overeniu prostredníctvom testovania použiteľnosti. Oslovili sme pritom našich potenciálnych používateľov a spoločne s nimi absolvovali test v priestoroch laboratória na **testovanie používateľského zážitku** (UX-LAB FIIT). Prostredníctvom dostupných nástrojov na sledovanie pohľadu a získanej spätnej väzby sme identifikovali rôzne nedostatky a možnosti ďalšieho vylepšenia systému. Takáto forma overenia nám umožnila opraviť chyby, ktoré sa neprejavili pri vývoji, a skvalitniť tak celkový dojem z práce v našom systéme. Systém sa tiež používal v rámci **diplomového projektu** v experimente testovania použiteľnosti, pri ktorom slúžil na získavanie informácií o participantoch pred testovaním a následne spätnej väzbe po skončení testovania. Systém Crowdex sa tu osvedčil ako spoľahlivý nástroj pri organizovaní experimentu s väčším počtom účastníkov.

⁶<http://www.codeclimate.com>

⁷<http://www.travis-ci.com>

4 Zhrnutie

Výsledok našej práce počas dvoch semestrov predstavuje systém **Crowdex**, ktorý v mnohých smeroch obohacuje a skvalitňuje existujúce prístupy manažmentu a realizácie mikroúloh. Systém uľahčuje a zrýchľuje činnosti spojené so spomenutými procesmi a prispieva k skvalitneniu študentských a výskumných prác, ktorých častým problémom je získanie dostatočného počtu vhodných účastníkov na overenie navrhnutých rozhraní, prístupov alebo hypotéz. Podobne môže naša aplikácia pomôcť aj firmám pri realizácii malých úloh, na ktoré nepotrebujú zamestnávať bežného zamestnanca. Systém Crowdex je založený na technológii Webu 2.0 a princípoch crowdsourcingu. Základom nášho systému je programovací jazyk *Ruby* spolu s aplikačným rámcom *Ruby on Rails*. Pre ukladanie dát slúži SQL databáza *PostgreSQL* v kombinácii s NoSQL riešeniami *ElasticSearch* a *Redis* pre zvýšenie výkonnosti aplikácie.

Systém Crowdex sme od začiatku vyvíjali s úmyslom jeho nasadenia v praxi, kde bude môcť pomôcť ľuďom v **manažmente a vykonávaní mikroúloh**. Tento cieľ sa umocnil najmä v závere letného semestra, kedy sme sa rozhodli v práci na našom systéme pokračovať aj nad rámec predmetu Tímový projekt.

Nové myšlienky, nápady a návrhy na zlepšenie sme získali najmä vd'aka **spätnej väzbe od používateľov**. Veľká časť týchto návrhov sa už pretavila do podoby a fungovania systému Crowdex. Ďalším veľkým impulzom bolo **testovanie používateľského zážitku (UX)** v novovybudovanom UX laboratóriu na fakulte, ktoré sme realizovali v ostatných dňoch. Počas testovania sme mali možnosť sledovať prácu používateľov v homogénnom prostredí v reálnom čase. Výstupom bolo niekoľko priamych návrhov na zlepšenie od používateľov ale i našich vlastných postrehov a nápadov, ktoré aktuálne zapracúvame do systému. Ide napríklad o redizajn rozhrania manažmentu úlohy, ktorý bude realizovaný s ohľadom na moderné prvky používateľských rozhraní. Hlavnou inšpiráciou sú aktívne dlaždice, ktoré nahradia v súčasnosti používané karty. Práve na základe UX testovania sme si všimli, že rozdelenie manažmentu úloh do kariet nie je pre používateľov prehľadné a občas aj máťúce. Ďalšou zmenou bude rozdelenie riešenia úlohy do systematickej postupnosti krokov, ktorou bude používateľ vedený priamo systémom. Ak teda úloha bude obsahovať inštrukcie, odovzdanie súborov a vyplnenie dotazníka, používateľ bude musieť postupne prejsť týmito krokmi pred odovzdaním úlohy. Týmto sa redukuje zahltenie používateľa množstvom informácií, a tiež sa zamedzí javu, kedy používatelia zabudli vyriešiť niektorú zo súčastí úlohy, prípadne zabudli ukončiť jej riešenie.

Okrem týchto zlepšení plánujeme tiež **rozšíriť existujúcu funkcionality** systému Crowdex. Zmeny sa dotknú nielen rozšírenia funkčnosti dotazníkov pridaním nových typov otázok a pokročilých štatistík, ale i možnosti zvoliť si, aké moduly chce zadávateľ použiť vo svojej úlohe. V spolupráci s UX laboratóriom na fakulte by sme tiež chceli poskytnúť zadávateľom úloh možnosť **sledovania pohľadu a emócií** používateľov pri riešení špecifických úloh, napríklad pri testovaní používateľského rozhrania aplikácií a vyplňaní dotazníkov. Zozbierané dáta by pritom mali k dispozícii v spracovanej forme, ktorá by umožnila identifikovať, na aké prvky sa používateľ pozerá, aký čas strávil pozeraním na niektorú časť aplikácie, či aké emócie prežíval počas vyplňania otázok v dotazníku.