



Odporúčanie autorít na mikroblogoch

Diplomový projekt 2

Bc. Peter Korenek

Vedúci
Ing. Marián Šimko



Authority na sociálnom webe

- **„moc ovplyvňovať' myslenie, názor, alebo správanie jedinca“ [Merriam-Webster]**
- **Osoby, ktoré veľké množstvo ľudí považuje za vplyvné kvôli obsahu, ktorý produkujú**
- **Sú celebrity alebo firmy autoritami??**



Ciele práce

- **Identifikácia autorít na sociálnych sieťach**
- **Odstránenie nedostatkov existujúcich prístupov**
- **Využitie analýzy emócií pri identifikácii autorít**
- **Personalizované odporúčanie autorít používateľom**



Identifikácia autorít

- Prvý – iba počty nasledovníkov [Java a kol]
- Najznámejší - TwitterRank [Weng a kol] – HITS pre grafy s hranami follow v rámci témy
- TURank – Retweet, Follow, Post
- Welch M.J. – Mention, follow, retweet
- Najlepší - Pal A. a Counts S. – veľké množstvo charakteristík špecifických pre Twitter



Nedostatky súčasných prístupov

- **Ako authority identifikujú príliš veľa celebrit**
- **Pre nájdenie skutočných autorít nepostačujú iba vzťahy follow, retweet**
- **Neanalyzujú obsah toho čo píše mikroblogger**



Identifikácia autorít na sociálnych siet'ach

- **Náš prístup:**
 - Použitie všetkých známych charakteristík mikroblov – retweet, mention, follow, reply, hashtag ...
 - Porovnanie PageRank a klasifikácie
 - Pridanie väčšej váhy príspevkom, ktoré majú kladné hodnotenia (retweet a reply predstavujú pozitívne hodnotenia)



Detekcia emócií

- **Emotikony**
- **Pozitívne vs negatívne slová**
- **Veľké písmená**
- **Slangové slová**
- **Interpunkčné znamienka**
- **Zdvojovanie písmen**



Identifikácia emócií

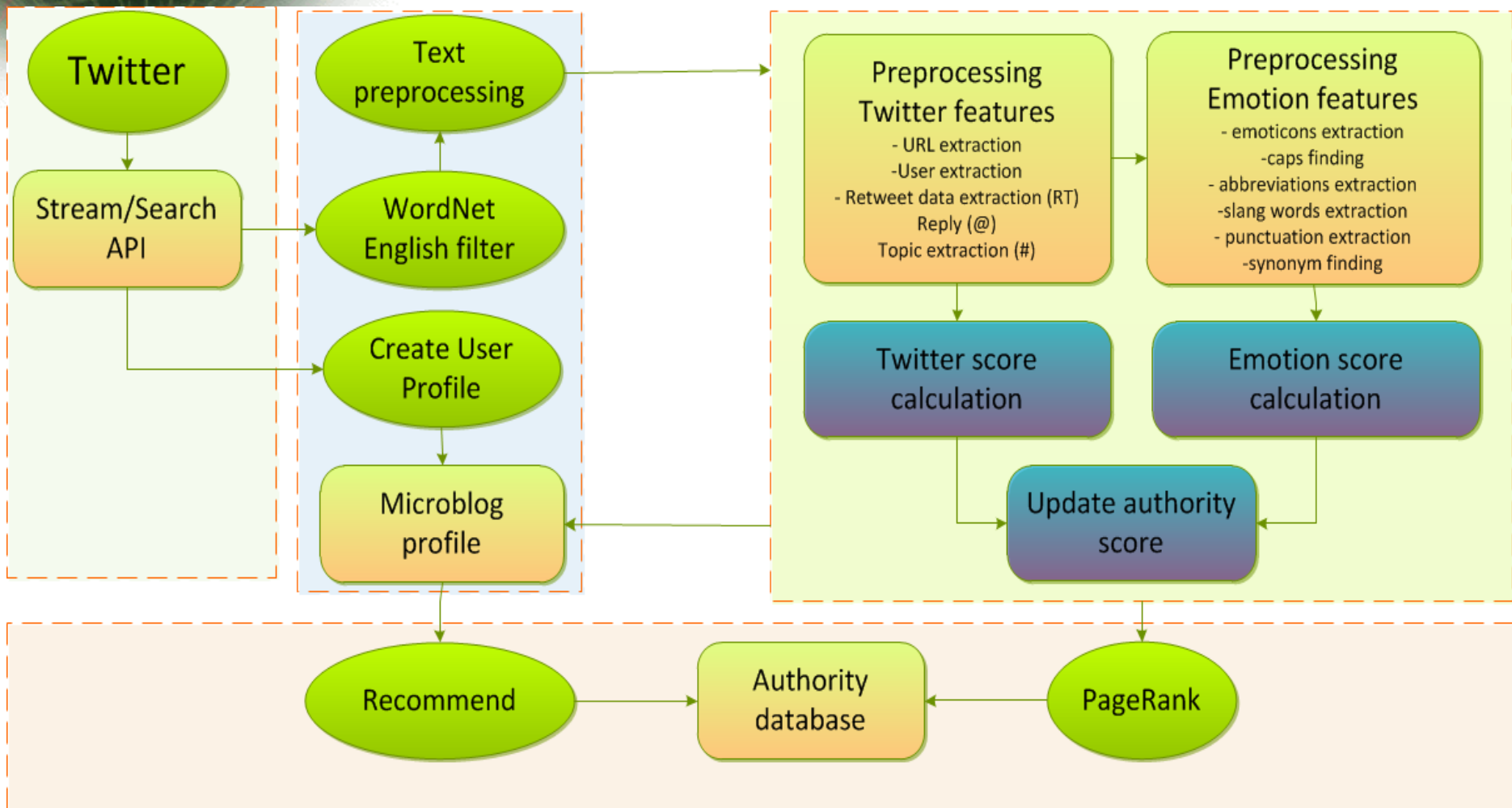
- **Bollen a kol. – Identifikácia vplyvu globálnych udalostí na emócie**
- **Barbosa a kol, - Kategorizovanie tweetov na pozitívne, neutrálne, negatívne**
- **Pak, Paroubek – Kategorizovanie nálad z tweetov, využitie lingvistickej analýzy**



Emocionálna identifikácia autorít

- **Identifikácia črt mikrobľgov v tweetoch**
- **Vytvorenie mikroblogovacieho modelu používateľa**
- **Emocionálny model používateľa**
- **Použitie upraveného PageRank**
- **Zvýšenie váhy tweetu v prípade, ak reply a retweet tweetu sú kladné/pozitívne**

Návrh systému pre odporúčanie





Problémy

- Použitie realtime / staticky
- Authority, ktoré píšú o negatívnych skutočnostiach – väčšinou negatívne reakcie
- Validácia identifikácie autorít – porovnanie so základnými metrikami (follow) / aplikáciami (TwitterRank)
- Validácia odporúčania?



Existujúce odporúčače

- **Twitter – Who to follow**
- **Bing – vyhľadávanie na základe dopytu**
- **Hannon, J. a kol. – Twittomender – odporúčanie na základe podobnosti dokumentov na mikroblogoch**



Resumé

- **Authority na mikrogloboch**
- **Identifikácia emócií v mikroblogochoch**
- **Využitie emócií pri určovaní autorít**
- **PageRank/Klasifikácia**
- **Model mikroblogera**
- **Personalizované odporúčanie autorít na mikroblogochoch**