Automated user profiling

SCS 4001

Kasun Gamlath
xxxxxxxx

Motivation

- User profiling is useful in many contexts
- Expansion of social media usage
- Nature of the social media to encourage the sharing
- Vastly available personal data on public
- User profiling is easier than ever

Problem (When done in manually)

- Data is scattered
- Time consuming
- Tedious
- Repetetive
- Collisions of available data

Solution - Automating the process

- Use the provided social network APIs
- Scalable crawling system
- Specialized content extraction
- Content recognition
- Probabilistic classifier
- Clustering
- Summerize
- Complete user profile

Research question & Aim

Is it possible to automate the user profiling from publicly available data in social network sites ?

Goals

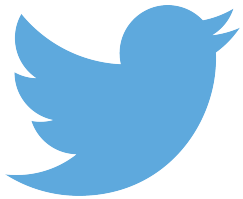
- Interface to enter user details
- Scripts to automating the queries for each social network API
- Web crawlers to use when API is not provided or sufficient
- Content extraction algorithm
- Content classification/clustering algorithm
- Summarization algorithm
- Interface to present processed user details

Final evaluation - Questionnaire

- Usability
- Accuracy
- Relavantness
- Completeness
- Uniqueness

Constraints

- Social Networks to be used :
  - Facebook
  - Twitter
  - LinkedIn
  - Google+
- Analized content will be limited to English.

Related work

- Abusing Social Networks for Automated User Profiling by Balduzzi, Marco Platzer, Christian Holz, Thorsten Kirda, Engin Balzarotti, Davide Kruegel, Christopher in 2010 [1]
- Automation of constructing an eProfile from Web Contents by Hashan Silva and Dr. Ajantha Athukorale in 2013, unpublished work [7]
- Combining Data Mining and Machine Learning for Effective User Profiling by Tom Fawcett and Foster Provost in 1996 [6]

References

[1] M. Balduzzi, C. Platzer, T. Holz, E. Kirda, D. Balzarotti, and C. Kruegel. Abusing social networks for automated user profiling. In Recent Advances in Intrusion Detection , 2010
[2] Acquisti, A., and Gross, R. Information Revelation and Privacy in Online Social Networks. In Proc. WPES 2005,ACM Press(2005).
[3] https://www.facebook.com/robots.txt
[4] http://www.facebook.com/apps/site_scraping_tos_terms.php
[5] https://www.linkedin.com/robots.txt
[6] Fawcett, Tom, and Foster J. Provost. "Combining Data Mining and Machine Learning for Effective User Profiling." KDD. 1996.
[7] Hashan Silva, Dr. Ajantha Athukorale (2013), Automation of Constructing an eProfile from Web Contents, Unpublished manuscript.