

Raw Linguist Analysis and Word Count Output for Big Five and Dark Triad traits

Talk Title : Can Twitter help expose psychopath killer's traits

Presenters: Chris Sumner & Randall Wald

Abstract:

Recent research has identified links between Psychopaths and the language they use (Hancock et al 2011), with media reports suggesting that such knowledge could be applied to social networks in order help Law Enforcement Agencies expose "Psychopath killers' traits". This is the first public study to research Psychopathy in the context of social media.

This study explored the extent to which it is possible to determine Psychopathy, and other personality traits based on Twitter usage. This was performed by comparing self-assessment 'Dark Triad' (Psychopathy, Machiavellianism, Narcissism) and 'Big Five' (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) personality traits with the Twitter information, usage and language of 2927 participants.

Results show that there are a number of statistically significant correlations between an individual's darker personality traits and their Twitter activity. We also identified links between users' attitudes to privacy, their personality traits and their twitter use. We will present the improvement gains possible through the use of machine learning for personality prediction and share the models and techniques employed.

In addition to presenting our results, this talk will provide an introduction into identifying psychopathic traits using the Hare Psychopathy Checklist (PCL-R), present the technical approaches to collecting, storing and analyzing Twitter data using Open Source technologies and discuss the current ethical, privacy and human rights concerns surrounding social media analysis, vetting and labeling.

We will conclude with two proof of concept works, the first using the visualization tool Maltego to explore how visual analysis could be used to identify potential troublemakers at events such a far right demonstrations; the second to look at how personality traits influence response and interaction with a benign Twitter Bot.

The results highlight that in certain contexts, personality prediction through social media can perform with a reasonably high degree of accuracy.

Biographies:

Chris Sumner is a contributor in the emerging discipline of Social Media Behavioral Residue research where he combines his interests in Psychology, Social Networks, Data Mining and Visual Analytics. He has previously spoken about these topics at BlackHat and DEF CON and is scheduled to speak at the European Conference on Personality in July 2012 with a team of academic personality researchers.

Chris has been directly involved in Corporate Information Security at Hewlett-Packard since 1999 and is currently focused on Security in the Development Lifecycle. Outside of work and together with a small group of likeminded individuals, he co-founded the not-for-profit Online Privacy Foundation to conduct topical research and raise security awareness at a community level.

Twitter: [@TheSuggmeister](#)

<https://www.facebook.com/onlineprivacyfoundation>

<http://www.onlineprivacyfoundation.org/>

Randall Wald is a researcher studying data mining and machine learning at Florida Atlantic University. Following his BS in Biology from the California Institute of Technology, Randall chose to shift his focus to computer science, applying his domain knowledge towards bioinformatics and building models to predict disease. He also studies machine learning for other domains, including machine condition monitoring, software engineering, and social networking.

<http://www.ceecs.fau.edu/directory/randallwald>

Raw Linguist Analysis and Word Count Output for Big Five and Dark Triad traits