Bot-or-Not: social bot detection

Onur Varol, Clayton Davis, Prashant Shiralkar,
Emilio Ferrara, Filippo Menczer, and Alessandro Flammini

CNetS

INDIANA UNIVERSITY
BLOOMINGTON

I. Detecting social bots

Can we verify the authenticity of users who participate in discussions and produce content in social media?

Our system detects with **95%** accuracy (measured by AUC) *whether a user is a bot or not*, by exploiting over 1K features that capture **user meta-data**, **social contacts**, **diffusion networks**, **content**, **sentiment**, and **temporal patterns**.

II. Dataset

14.8K Legitimate users

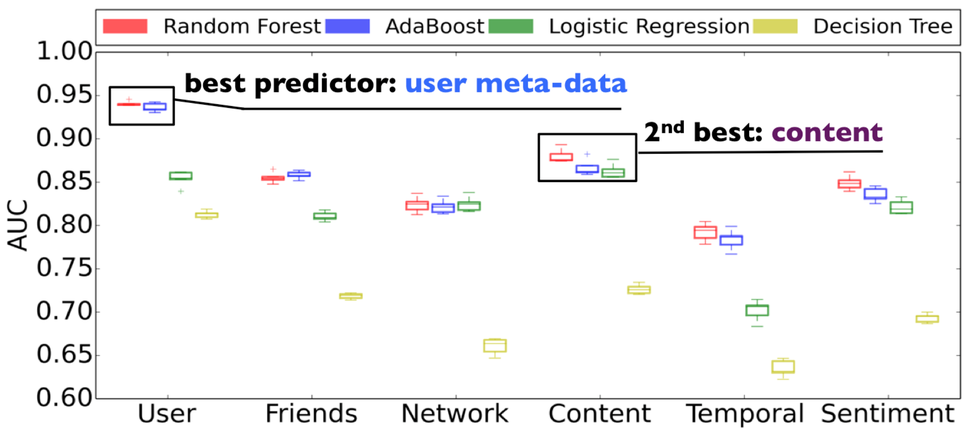
15.8K Social bot users

Lee, Kyumin, Brian David Eoff, and James Caverlee. "*Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter.*" ICWSM. 2011.

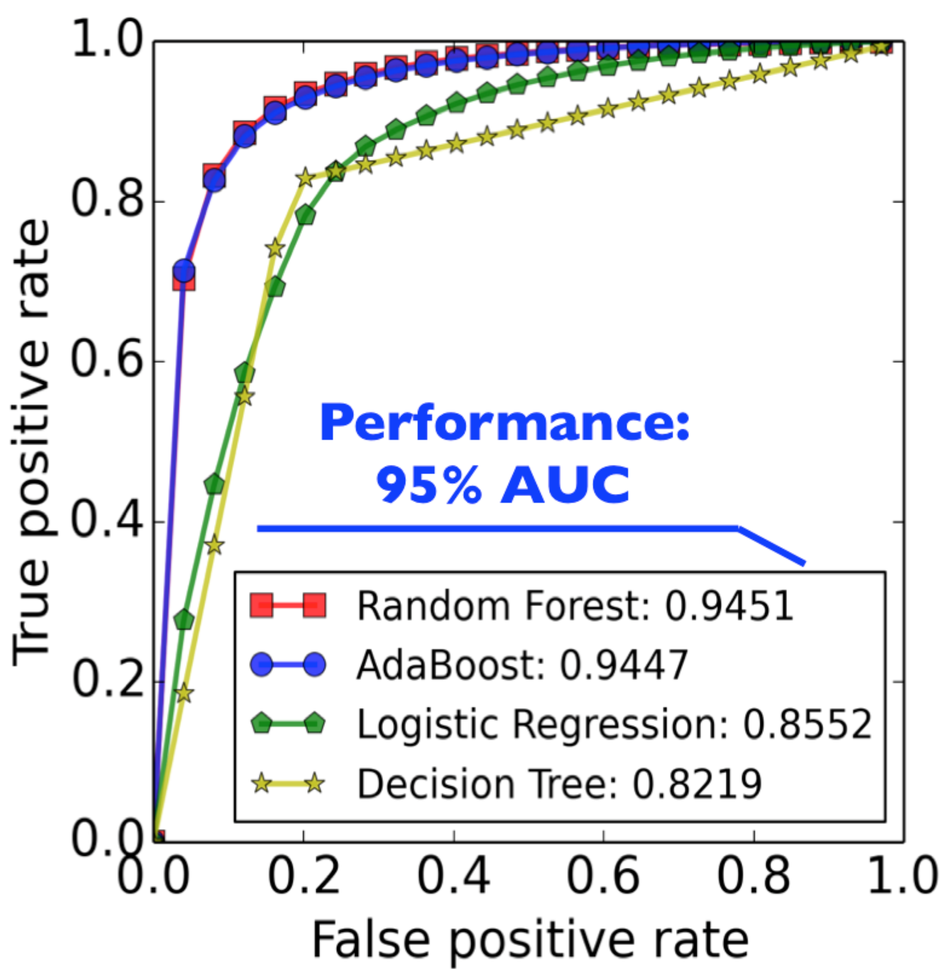
III. Features

Class (# of features)	Description
Network (**112**)	Retweet, Mention, HT co-occurrence networks
User (**56**)	User meta-data
Friend (**208**)	Followees statistics
Timing (**24**)	Temporal information
Content (**411**)	Part of Speech (PoS) tags and meme information
Sentiment (**339**)	Sentiment Analysis features

IV. Feature selection



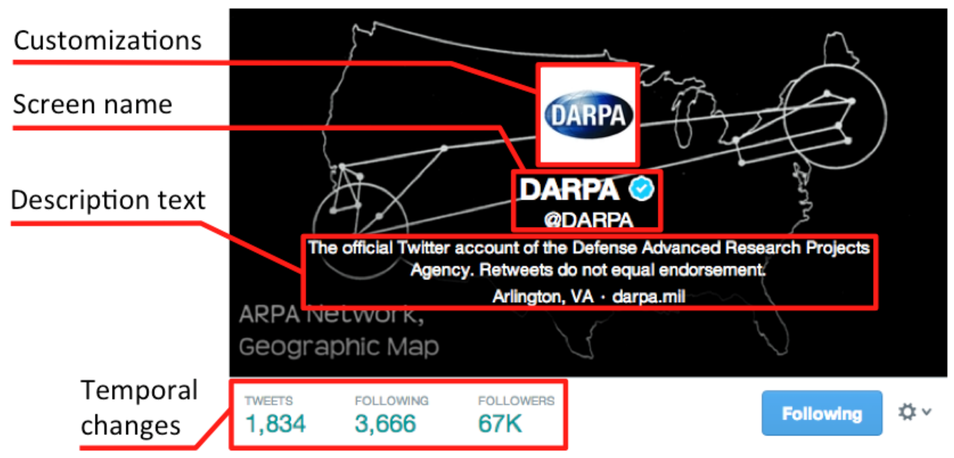
V. Performance evaluation: ROC



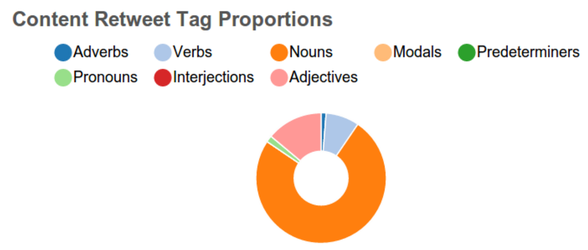
VI. Demo: `truthy.indiana.edu/botornot`

Front-end: **HTML5/Javascript** — Back-end: **Python/Flask** — Interface: **D3.js** — Authentication: **Twitter API**

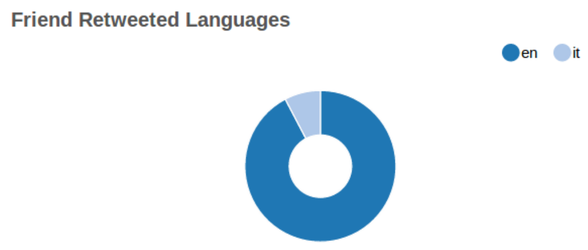
VII. User information



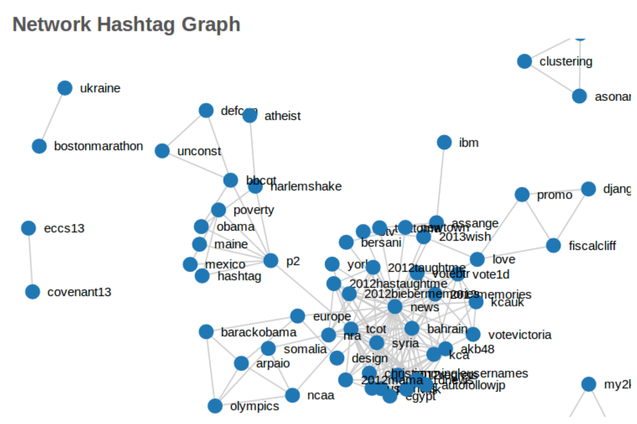
VIII. Content POS tags



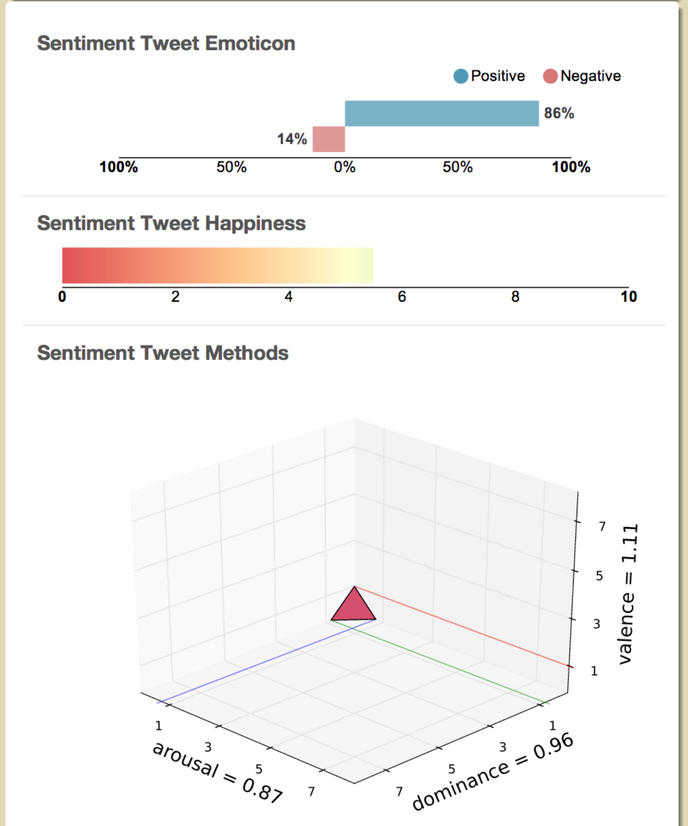
IX. Contact languages



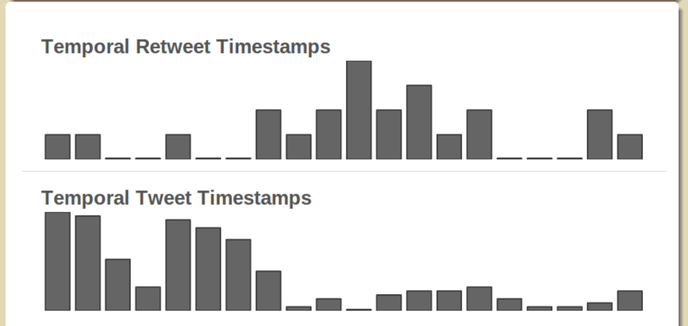
X. Network



XI. Sentiment Analysis



XII. Temporal signals



DARPA Grant Award W911NF-12-1-0037 - NSF Grant Award CCF-1101743