Automatic Author Identification Using Personality Insights using Internet Relay Chat
Sicong Shao, Cihan Tunc, Salim Hariri
NSF Center for Cloud and Autonomic Computing

Motivation
- The user with illegitimate purposes can exploit the powerful dissemination feature of social network platform to spread malicious information and influence people opinion.
- With the anonymous nature, Internet Relay Chat (IRC) is widely used for cybercrime including hacking, cracking, and carding.
- Performing author identification on IRC is more challenging task, due to difficulties of real-time data collection, synchronous computer-mediated communication, many sophisticated suspects that are needed to be considered.

Objectives
- Creating autonomic IRC monitoring bot for continuous monitoring and performing pre-processing for unstructured data in real-time.
- Developing an effective personality feature extraction unit.
- Designing a robust learning unit for creating author identification model.
- Performing author identification to identify anonymity.

Approach
- Personality Feature Extraction Unit

Support Vector Machine Learning Unit
Given a personality feature training dataset \( (x, y) \),
The dual form is:
\[
L(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j K(x_i, x_j)
\]
s.t. \( \sum_{i=1}^{n} a_i y_i = 0 \), \( 0 \leq a_i \leq C, i = 1, ..., n \)
The classification function is:
\[
y = \text{sign} \sum_{i=1}^{n} a_i y_i K(x_i, x) + b
\]
The binary classification SVM can be extended to a multi-classification SVM by one-against-one method.

System Architecture
-Watson Assistant powers bot’s conversation module
-Transforming unstructured data to structured data in real-time.
-Parent bot can automatically generate children bot that inherits all the capabilities
-Automatic analyzing user’s personality feature based on three models: Big Five, Needs, and Values.

Experiment and evaluation
- Monitoring IRC channels involving the topics of hacking, cracking, carding, and political warfare.
- Evaluating author identification model using leave-one-out cross-validation.