Authenticating the Remote Learner: Linguistic Forensics

Paul Thompson
12 November 2008
Outline

• Early Research
• Literary Computing
• Intelligence and Security Informatics
• NSF OSTP Workshop on Detecting Deception in Language
• Conclusions
Early Research

• Authorship Attribution
  – Mosteller and Wallace – Federalist papers
  – Stylometrics

• Automatic genre detection
  – Schuetze and Nunberg
Literary Computing

• Who wrote Shakespeare?
• More elaborate mathematical models, e.g., principal component analysis
• Decades or research
Intelligence and Security Informatics

- Jonathan Lebed case – multiple user accounts, single person
- Admissibility as evidence, UK vs. US
- Linguistic fingerprint?
- NSF/CIA KDD program – Madigan et al. at DIMACS
Intelligence and Security Informatics

- International Association of Forensic Linguists
Madigan et al. KDD program

- Design of a collection or corpus of texts of known authorship for use as a research / development test bed
- Definition of standard attribution tasks for comparing techniques
- Formal evaluation methods
- 2006 DIMACS workshop
Madigan et al. (cont.)

• “Author Identification on the Large Scale”
• Attribution of
  – Which of these K authors wrote this particular document?
  – Did any of these K authors write this particular document?
Madigan et al. conclusions

- Bayesian logistic regression with high-dimensional document representations shows considerable promise for authorship attribution
- Challenges remain: different representations can give different attribution results
- No clear method exists to account for this uncertainty
Detecting Deception in Language

• First of Seven NSF OSTP Workshops on Security Evaluations – summer 2005
  – Background investigations, interviews, and interrogations
  – Most participants psychologists
Detecting Deception in Language (cont.)

• Research needs
  – urgent to develop a large, shared corpus of empirically validated truthful and untruthful messages drawn from a variety of situations, including both laboratory and real-world settings

• For authorship attribution similar need for ground truth data sets
Conclusions

• Problem not yet directly addressed
• Technology is available to address the problem
• Suggestive, but not conclusive, evidence for efficacy