



VirginiaTech
Institute for Critical Technology and Applied Science

Ted and Karyn
Hume Center for National Security and Technology

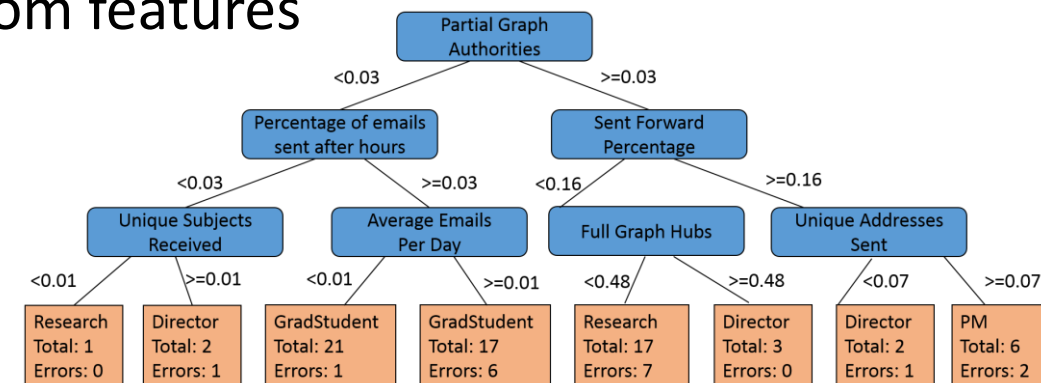
Data Mining Academic Emails to Model Employee Behaviors and Analyze Organizational Structure

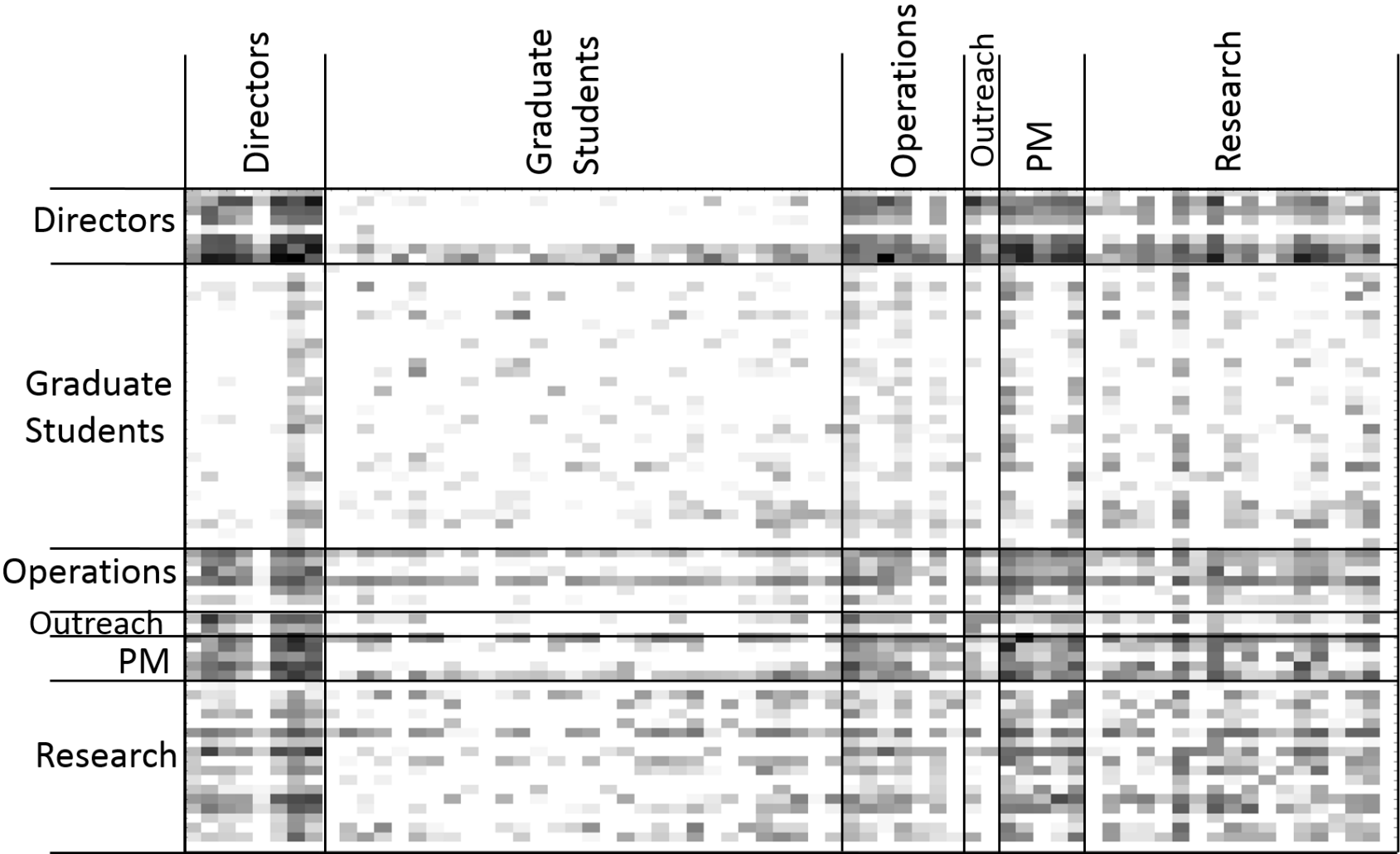
2/3/2016

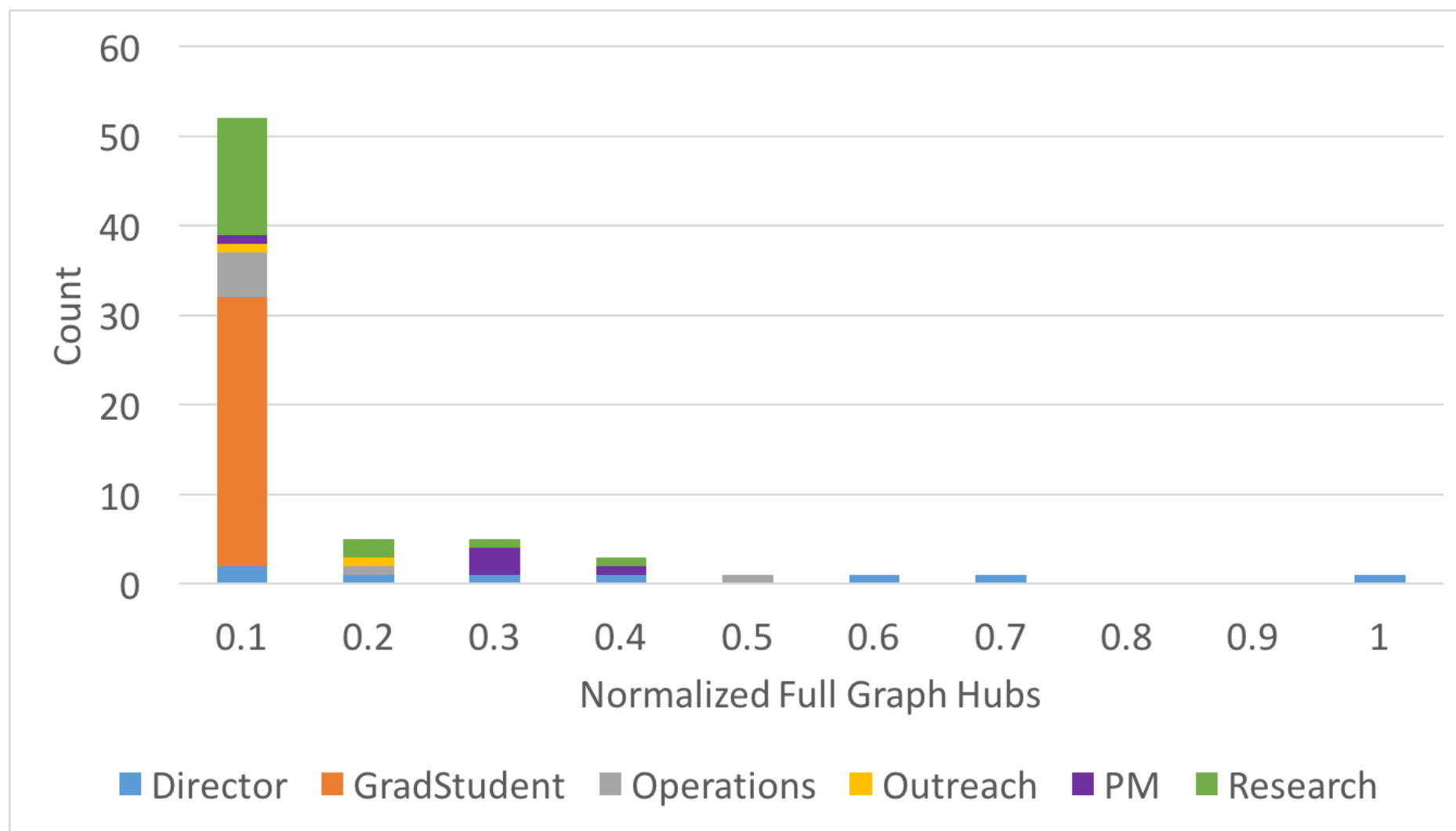
hume@vt.edu
www.hume.vt.edu

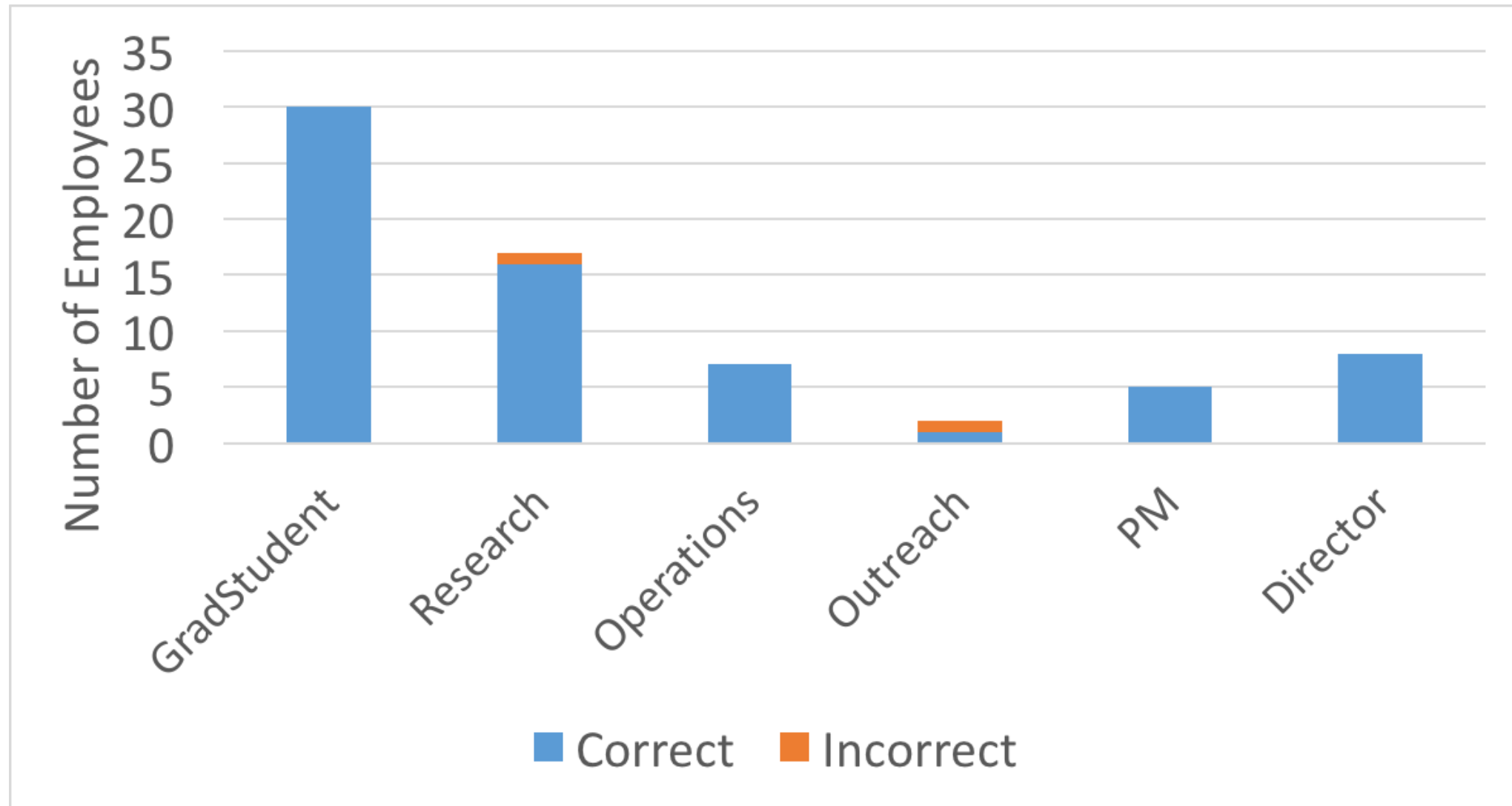
Hume Center Overview Presentation

- Problem
 - Email is everywhere!
 - Can a better understanding of an organization be gleaned from internal emails?
 - Application: business reorganization
- Approach
 - Enron dataset has issues, so collect anonymous email metadata from volunteers
 - Calculate 98 different features, including traffic-based and social-based
 - Unique subjects received, number of forwards received, hubs
 - Use random forest algorithm to predict job title from features
 - Compare trends in data with organizational chart











VirginiaTech

Institute for Critical Technology and Applied Science

Ted and Karyn

Hume Center for National Security and Technology

Questions?