1 Related Works

Email is a pervasive medium in modern society - particularly in the workplace. [Cite stat from [1]]. [Talk about email in the workplace] [2].

Email has been a common research topic over the past decade. Most notably, the Enron email database was released in 2004, following the company's collapse [3]. This dataset has been extensively researched on topics ranging from spam classification [4] to statistical analysis [5] to [Ex3]. However, there are known issues with even the most recent versions of this dataset, as summarized by [6].

The existing literature on email research is mainly divided into two categories: feature-based and social-based [7]. However, the goal of this paper is to unite the two factions by using traditional statistical email features as well as features derived from representing the dataset as a social network.

References

[1] "Email-Statistics-Report-2015-2019-Executive-Summary.pdf." [Online]. Available: http://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf

[2] D. Fisher, A. J. Brush, E. Gleave, and M. A. Smith, "Revisiting Whittaker & Sidner's email overload ten years later," in *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM, 2006, pp. 309–312. [Online]. Available: http://dl.acm.org/citation.cfm?id=1180922

[3] J. Shetty and J. Adibi, "The Enron email dataset database schema and brief statistical report," *Information sciences institute technical report, University of Southern California*, vol. 4, 2004. [Online]. Available: http://foreverdata.com/1009/Enron_Dataset_Report.pdf

[4] S. Martin, B. Nelson, A. Sewani, K. Chen, and A. D. Joseph, "Analyzing Behavioral Features for Email Classification." in *CEAS*, 2005. [Online]. Available: http://blaine-nelson.com/research/pubs/Martin-Sewani-CEAS-2005.pdf

[5] J. Shetty and J. Adibi, "Discovering important nodes through graph entropy the case of enron email database," in *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005, pp. 74–81. [Online]. Available: http://dl.acm.org/citation.cfm?id=1134282

[6] A. Nordb, "Data Visualization for Discovery of Digital Evidence in Email," 2014. [Online]. Available: http://brage.bibsys.no/xmlui/handle/11250/198551

[7] G. Tang, J. Pei, and W.-S. Luk, "Email mining: tasks, common techniques, and tools," *Knowledge and Information Systems*, vol. 41, no. 1, pp. 1–31, Jun. 2013. [Online]. Available: http://link.springer.com/10.1007/s10115-013-0658-2