Finding Spammy Names in Social Networks

David Freeman, Ph.D.

Head of Security Data Science
LinkedIn

Abstract: Many social networks are predicated on the assumption that a member's online information reflects his or her real identity. In such networks, members who fill their name fields with fictitious identities, company names, phone numbers, or just gibberish are violating the terms of service, polluting search results, and degrading the value of the site to real members. Finding and removing these accounts on the basis of their spammy names can both improve the site experience for real members and prevent further abusive activity. In this talk we describe how to use the Naive Bayes classification algorithm to find accounts whose names do not represent real people. The model can detect both automated and human abusers and can be used at registration time, before other signals such as social graph or clickstream history are present. We use member data from LinkedIn to train and validate our model and to choose parameters.

Bio: *David Freeman* is head of Security Data Science at LinkedIn, where he leads a team charged with detecting and preventing fraud and abuse across the LinkedIn site and ecosystem. He has a Ph.D. in mathematics from UC Berkeley and did postdoctoral research in cryptography and security at CWI and Stanford University.