Enron Case Study: Analysis of Email Behaviour Using EmailTime
Minoo Erfani Joorabchi, Ji-Dong Yim, Mona Erfani Joorabchi and Christopher D. Shaw
DVH Research Group
School of Interactive Art and Technology, Simon Fraser University, Surrey, BC V3T 0A3 CANADA

1. Introduction

Research Question:
Are there specific patterns in the way people with different executive (organizational) positions use email?

More specifically,
Metrics are utilized to discover the impact of the organizational positions on the email behaviour of people by providing detailed quantitative information about them.

Email behavior: email activity of people regarding a series of measured metrics by Email Time, e.g., number of Sent and Received emails as To or Co, number of the email addresses, number of the created folders (by the owner of the email dataset) and size of the Sent emails.

EMAIITIME [1]:
- A visual analysis of email correspondence patterns over the course of time.
- Interactively portrays personal and interpersonal networks using the correspondence in the email dataset.

2. Benchmark: Enron Email Dataset

- A unique large dataset [2].
- Contains more than 200,000 emails.
- Time span for this case study is between January 2000 and December 2001.
- There are missing actors and emails in the dataset.
- Seven categories were identified [3] including 35 employees, 4 CEOs, 4 presidents, 21 vice presidents, 12 directors, 14 managers and 11 traders.

3. Analysis and Result

3.1 Analysis on the Activity Level

- Displays some groups behaved similarly regarding the number of Sent, Received, and exchanged emails (sum of Sent and Received).
- We divide them into three categories: Modest, Moderate and Active (dashed lines).
- Managers and employees were Active, Traders and Directors were Modest, and the rest were Moderate.
- Employees have the highest average number of Sent emails.
- Managers have the highest average number of Received emails.

3.2 Analysis on the Types (Sender/Receiver)

- We divided the Sent emails into four categories based on the number of recipients: Single nodes (one recipient), Small nodes (2-8), Medium nodes (10-29), and Large nodes (30 and up).
- Also, divide difference between CEOs and other groups in sending large size emails.
- Traders and managers sent medium size emails more than any other groups.

3.3 Analysis on the Size of Sent Emails

- Contribution index [4] is 
  \[ \text{email size} = \frac{\text{email size (To-CC)}}{\text{total size of email sent and received (To-CC)}} \]
  - For the receivers and for the senders.
- We expand this formula to receive as To (To-CC) and received as Co (Co-CC).
- Average Contribution Index of administrative people tends to be lower than that of lower positional staff.
- CI and To-CC follow a trend and Co-CC has a different trend.
- Co of Employees, Traders and Managers are near zero (they had same amount of send and received emails in average).
- Then there is a jump to directors, president and vice president when the three groups behaved as weak receivers.

Contact Information
DVH Research Group
School of Interactive Art and Technology
Simon Fraser University
Surrey, BC V3T 0A3 CANADA

References