# Replication and impact analysis of "Scale-free topology of e-mail networks"

Jack Margeson*

University of Cincinnati, College of Engineering and Applied Science

**Abstract**

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 1    Introduction

This report serves as both a replication study and impact analysis on the text "Scale-free topology of e-mail networks" [EMB02] by Holger Ebel, Lutz-Ingo Mielsch, and Stefan Bornholdt.

Following this introduction is a summary outlining the goals and results of the original paper. The summary leads into the replication of major findings from the paper, completed with a cleaned copy of the original dataset using NetworkX, a Python package with the purpose of creating, manipulating, and analyzing complex networks. The full source code for all replications performed can be found in a public GitHub repository [Mar23] from the author of this report. After the presentation of replications, an analysis is conducted on two identified examples of impactful work that have been produced citing the "Scale-free topology of e-mail networks" paper. Finally, a section theorizing two interesting directions for future research in the related field is included.

## 2    Summary

In the original text, Ebel et al. study a network consisting of sent and recieved e-mails. The data set used for the study was obtained by observing the log files of an e-mail server located at Kiel University. The authors of this paper were able to contruct a network based upon this data in which several mathematical analyses were conducted in order to deduce information that could help classify the network.

The general conclusion found from these analyses was that their network " exhibits a scale-free link distribution and pronounced small-world behavior" [EMB02] that aligns with the findings of similar studies on various other social-based networks. Implications for these findings, including the importance of the scale-free property of the network on the spread of e-mail viruses, are touched upon briefly towards the end of the paper.

### 2.1    Strengths

The greatest strength of the original paper is undoubtedly that it established a line of reasoning to determine the degree distribution of e-mail networks to exhibit a power law, providing evidence towards e-mail networks being scale-free. Additionally, not only did the authors find evidence that e-mail networks could be considered scale-free, they also found that the e-mail network that they had created from their sample data exhibited small world behavior, which indicates that node neighbors of a node are likely to be connected themselves. These two findings have been instrumental in future e-mail related studies, two of which will be talked about in the impact analysis section of this report.

---

*Additional author information: https://marg.es/on

## 2.2 Weaknesses

There is a small albeit important weakness to mention present in the original paper. Justifications for this weakness are provided via opinion of the author of this report.

One issue that arises in the fact that the sampling process is restricted to exclusively one e-mail server. However, it would generally be unreasonable for the authors of the paper to aquire data from other universities or similar group-based entities with an emphasis on e-mail communication. This is due to the fact that a data set of e-mails as large as the one presented in the original paper would have to be extensively cleaned as to not expose sensitive data, which would take time and energy on the part of the group providing the dataset. Given this information, it is unlikely that another dataset to enforce preliminary findings would not be feasable to obtain for the original paper authors, leaving the task of replication with alternate datasets to other publications.

## 3 Reproduction

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 4 Impact Analysis

Hello, here is some text without a meaning. This text should show what a printed text will look

like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 5 New Directions

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## 6 Conclusion

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# References

[EMB02]   Holger Ebel, Lutz-Ingo Mielsch, and Stefan Bornholdt. "Scale-free topology of e-mail networks". In: *Phys. Rev. E* 66 (3 Sept. 2002), p. 035103. DOI: `10.1103/PhysRevE.66.035103`. URL: `https://link.aps.org/doi/10.1103/PhysRevE.66.035103`.

[Mar23]   Jack Margeson. *Email Topology Analysis*. Nov. 2023. URL: `https://github.com/jack-margeson/email-topology`.