

MEMORANDUM

To: Peter

From: Yuecai

Subject: Idea of Truck Factor Small Paper

Date: Monday, December 11, 2014

1 Introduction

Since we have done a lot of work on the Truck Factor, and have found some interesting observation related to Truck Factor, I think it is enough for a small paper.

2 Research Questions

For the small paper, I propose the following research questions:

2.1 Can we speed up the Truck Factor algorithm?—Optimized Truck Factor with a stopping condition

In these research question, we can present and explain the Optimized Truck Factor algorithm, and discuss its limitation on time complexity.

2.2 can we speed up the Truck Factor algorithm with a trade off of losing accuracy?—Simplified Randomized Truck Factor

In these research question, we can present the Simplified Randomized Truck Factor. explain how to make the upper bound estimation, and show the result from Linux and Chrome.

2.3 What it really means by the Truck Factor?—relation between Modularization and Vulnerability

In these research question, we can use a simplified model to explain that when a software team is high modularized, it will have a low Truck factor, otherwise, It depends on how the ownership distributes on the team, flat or skewed. a well modularized team will have a higher standard error on the sample file loss data, a skewed ownership distributing team will have a high std too. a flat ownership distributing team will have a low std and average file loss. Thus the randomized algorithm will reveal this fact. The data of Chrome and Linux confirms this.

3 Conclusion

In the small paper, we contribute a stopping condition for the Truck Factor algorithm to speed it up. However it is still too slow for large project. Then we create a new algorithm to estimate the Truck Factor based on random sampling. We also discuss about the relation between project developer team modularization and project vulnerability based on the Truck Factor findings.