

Title: “Sociometrics in Software Engineering: Measuring Technical Skills of Software Engineer through GitHub”

Author: Rohan Bagwe

TCDID: [REDACTED]

Supervisor: Prof. Stephen Barrett

Date: September 7, 2020

Abstract:

Any software project relies gravely on the skill set of its engineers. Technical skill measurement and analysis are the key factors to observe the growth and performance of an engineer in a specific project. While the prevailing system does understand the utmost importance of skill measurement, the qualitative nature of it poses a challenge that has not met an optimized solution yet. So, how are the companies dealing with it, noticeably analyzing the self-authored resumes or LinkedIn profiles?

Many studies in the field to record the severity of the challenge have revealed that the primary issue faced by most recruiters when hiring engineers was to find sufficiently qualified candidates. Indeed, the majority of miss-hires in tech recruiting come down to an inaccurate measurement of the technical skills of the candidate. And even after the company has hired the engineers, they do not have quantified measurable parameters to assess their skill level. There is a need for more reliable metrics to evaluate engineers' technical skills and certify their knowledge on a particular subject. This could help management to assign the best resources to the project as per the budget and technical requirements of the task.

Traditionally, for the hiring process, the metrics used to evaluate the technical skills of a software engineer included education level, professional experience, online coding challenges, onsite interviews, and after the actual hiring, metrics change to the number of bugs fixed, number of new features developed or even the number of hours worked. These proxy variables do not offer any real or actionable insights to quantify the qualitative skills of individuals.

This research seeks to build a data model by capturing the qualitative signals from GitHub. This data model represents an opinionated view for measuring the technical skills of a software engineer. Using this conceptualized data model, this research also seeks to build an expert system that could rank order software engineers using a relative grading system. To evaluate the efficiency of the proposed solution, a sanity test is performed using manual evaluation by the human expert (the author of this research). The results indicate that the data model provides a reasonably accurate technical skill assessment of software engineers in Java technology by rank-ordering the contributors of the OSS Spring Boot project. We do see some loss in the conceptual fidelity of the proposed system. Thus, the limitations have also been discussed in the latter part of the thesis.