Development Cycle Estimation Modeling with the Statistical Agent-based Model of Development and Evaluation

Introduction

Communication of development risk can be problematic. For software security risk, the Open Web Application Security Project (OWASP) professional standard is commonly used. It is based on a conceptual equation (Risk = Likelihood * Impact) presented with a color-coded 10-point scale [1]. Unfortunately, this presentation is neither compelling nor adequately informative for business and some technical decision-makers. This concern is not limited to software development. Risk assessment professionals have reached similar conclusions in more general cases [2]. A framework or model that allows consistent and quantitative risk calculation is needed.

Motivated by years of application security testing experience, this work presents results from a domain-independent development model. The concern is not limited to software development. Risk assessment professionals have reached similar conclusions in more general cases [2]. A framework or model that allows consistent and quantitative risk calculation is needed.

SAbMDE uses Wang’s process algebra ideas [3, 4] to represent each development cycle phase so that analytical techniques can be uniformly applied across the entire cycle. Wang develops a desired end product (DEP) by sequentially composing intermediary products from sets of fundamental elements: vocabulary items and relations. An agent decides which vocabulary item and relation to compose, and a correct decision set produces the DEP. SAbMDE recognizes that each decision is one of a set of alternatives, (2) that the hierarchical super-set of alternatives forms a development space (dSpace), and (3) that the correct DEP decision set is a path (dPath) through dSpace.

DPath

The DPath Probability graph below illustrates these calculations and shows that prediction confidence degrades quickly. So, predictions must be repeated periodically.

\[
\begin{align*}
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(1.1)} \\
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(1.2, 1.4)} \\
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(1.3, 1.5)} \\
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(2.0)} \\
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(3.1, 3.2)} \\
\text{DPath Probability} & = \frac{1}{n} \sum_{i=1}^{n} P(D_i) \\
\text{(4.0)}
\end{align*}
\]

Pricing and Resource Utilization

Estimate the resource utilization associated with a dPath by assigning a price to each vocabulary item and relation, and price their composition with some appropriate function. Similarly, assign a price to their decomposition when an agent makes an incorrect decision. The equations below describe a simple pricing system. There, the number of attacks, n, is calculated with a hypergeometric distribution to ensure that a correct decision is eventually made and that the estimate is an upper limit.

\[
\begin{align*}
P_f &= \text{function price} \\
P_r &= \text{relation price} \\
P_{1, p} &= \text{composition pricing function} \\
P_{0, p} &= \text{decomposition pricing function} \\
P_{1, p} &= \frac{n}{N} \\
P_{0, p} &= \frac{n}{N} \\
(6.1) \\
P_{1, p} &= \frac{n}{N} \\
P_{0, p} &= \frac{n}{N} \\
(6.2)
\end{align*}
\]

Calculate a dPath’s resource utilization by summing the composition prices for each product in the dPath. Such a summation for as-yet untraversed portions of the dPath is an estimate. The graph below shows such estimates made from each composition level and for each agent skill level. Note that, at composition level 0, the ratio of higher to lower estimates is approximately 16:1. This is similar to Boehm’s Cone of Uncertainty (COU).

Conclusions

- The reasonably close correlation of SAbMDE estimates to COCOMO estimates implies that SAbMDE has potential.
- Estimates are built entirely on the current project characteristics: vocabulary items, relations, prices, pricing function, and composition order.
- The characteristics are measurable and they can be recorded.
- Each characteristic is well-known. Together, they are a common basis for communication and collaboration among all parties: developers, decision-makers, business managers, etc.
- Development risk is reduced by getting more vulnerabilities fixed sooner.

References

[6] Texas Tech University

Dr. Atila Ertas
Dr. Susan Mengel
Sam Denard (sam.denard@ttu.edu)
Dr. Stephen Ekwo-Arisi

Samuel Denard (sam.denard@ttu.edu)