Early Lifecycle Capability-Based Software Cost Estimation for Software-Intensive Programs

A Discussion with the Agile Collaboration Group
July 2023

Anandi Hira

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

DM23-0732

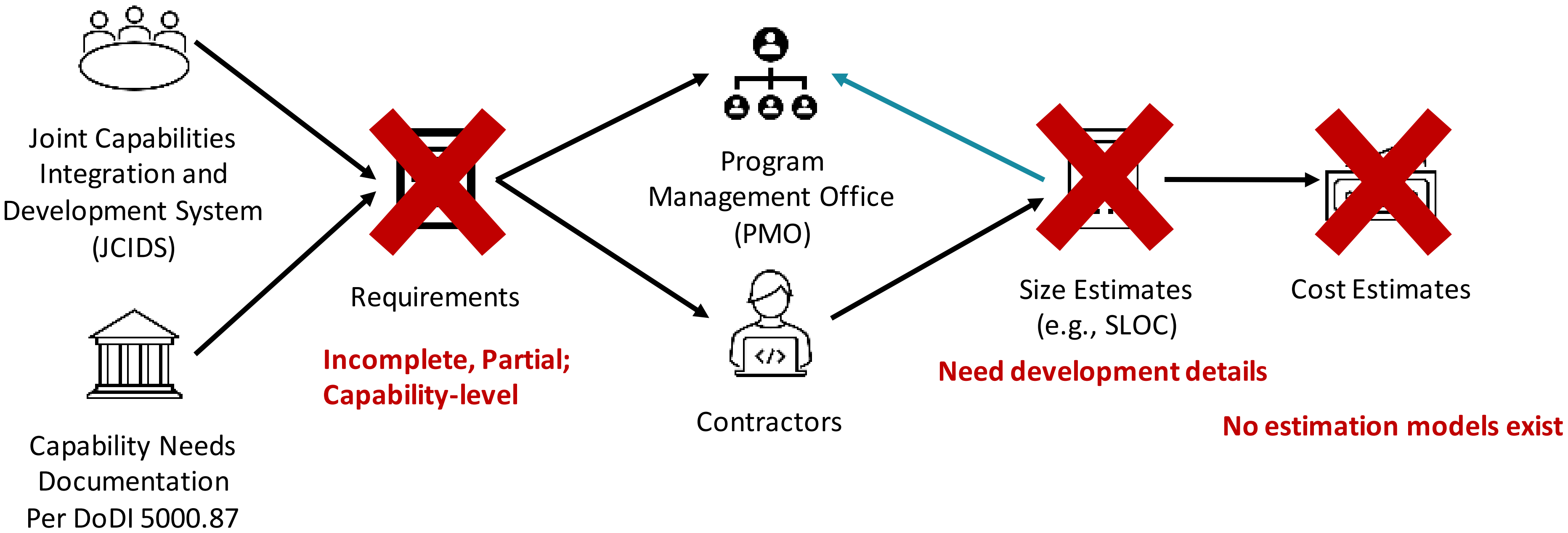
Military Application - Why are Cost Estimates Important?

- Give DoD and Service Cost Agencies and Congress basis of budget and estimated budget of program
- Answer questions such as: will capabilities be completed by deadline, when will capabilities be completed, what are effects of budget cut?
- Inform margins necessary to mitigate cost and schedule risks.

Unsound estimates have consequences

- Program portfolio cannot balance costs with benefits
- Underfunded programs in the POM (Program Objective Memo) will be perceived as failing by stakeholders, possibly triggering Nunn-McCurdy breaches
- Delayed capabilities increases risk to warfighters and mission accomplishment

DoD Challenge for Software Cost Estimation



Need size and estimation model based on capabilities

Examples of Problems with Existing Methods

Examples of struggles to build reliable, evidence-based cost estimates:

- A software-intensive program does not have faith in SLOC-based parametric model derived estimates due to significant changes in SLOC estimates over time – lacking credible estimate of effort needed to deliver critical capabilities or assess current progress.
- A DevSecOps program budget developed with Level of Effort (LOE) cut due to inability to justify budget request with demonstration of capability completion dates.

SEI is looking into potential research to address and overcome these challenges.

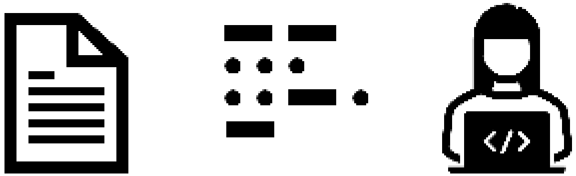
Potential Research Project Objectives

- New software sizing method – capability clusters/groups based on similar software functions
  - Based on information known at program initiation (capabilities)
  - Accounts for uncertainties (e.g., implementation variations)
  - Note! The size is not a number, it's a capability type!

- Cost model based on capability clusters/groups sizing
  - Effort ranges of similar capabilities
  - Evidence-based but also flexible

What are clusters of capabilities with similar functions?

GPS/Navigation vs Compilers/Parsers



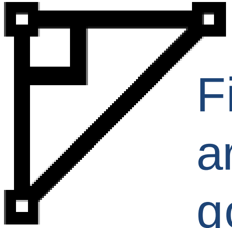


Figure out where we are, where we want to go. Some math/trig on how to get there.

Parse text/code for prespecified patterns.

Redefining software size - demystifying the software sizing process while providing an objective, flexible, and defensible estimation method

Capability Clusters

SRDR Data

CSCI	Descr	Effort
1234	ABC 123	#
2234	XYZ acd	#
1345	mn ABC	#
2346	ABC yyy	#
3456	dco XYZ	#
4456	XYZ qrs	#

NLP/ML on Descriptions to identify clusters

Iterate for improvements

Capability Clusters

CSCI	Descr	Effort
1234	ABC 123	#
2234	XYZ acd	#
1345	mn ABC	#
2346	ABC yyy	#
3456	dco XYZ	#
4456	XYZ qrs	#

Orange and purple represent 2 capability clusters

Validation

- ≥ 0.5 Silhouette coefficient
- ≤ 10% calculation error

Cost Model

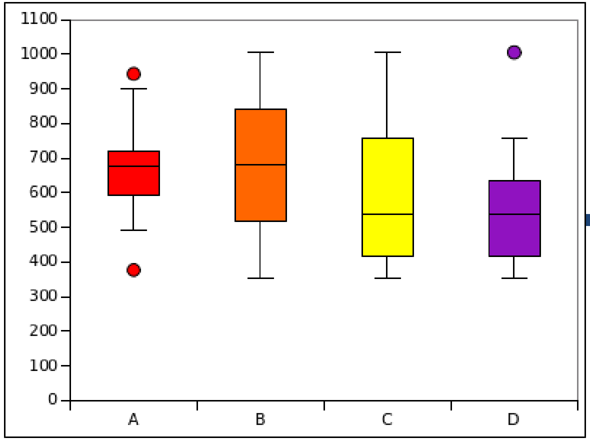
Capability Clusters

CSCI	Descr	Effort
1234	ABC 123	#
2234	XYZ acd	#
1345	mn ABC	#
2346	ABC yyy	#
3456	dco XYZ	#
4456	XYZ qrs	#

Get effort ranges per capability cluster

Min, Max, Mean, Median, Variance, Standard Deviation

Cost Model



Validation

- Within 90% confidence bound 90% of the time

Orange = Cluster B
Purple = Cluster D

Relevance and Impact of Potential Research

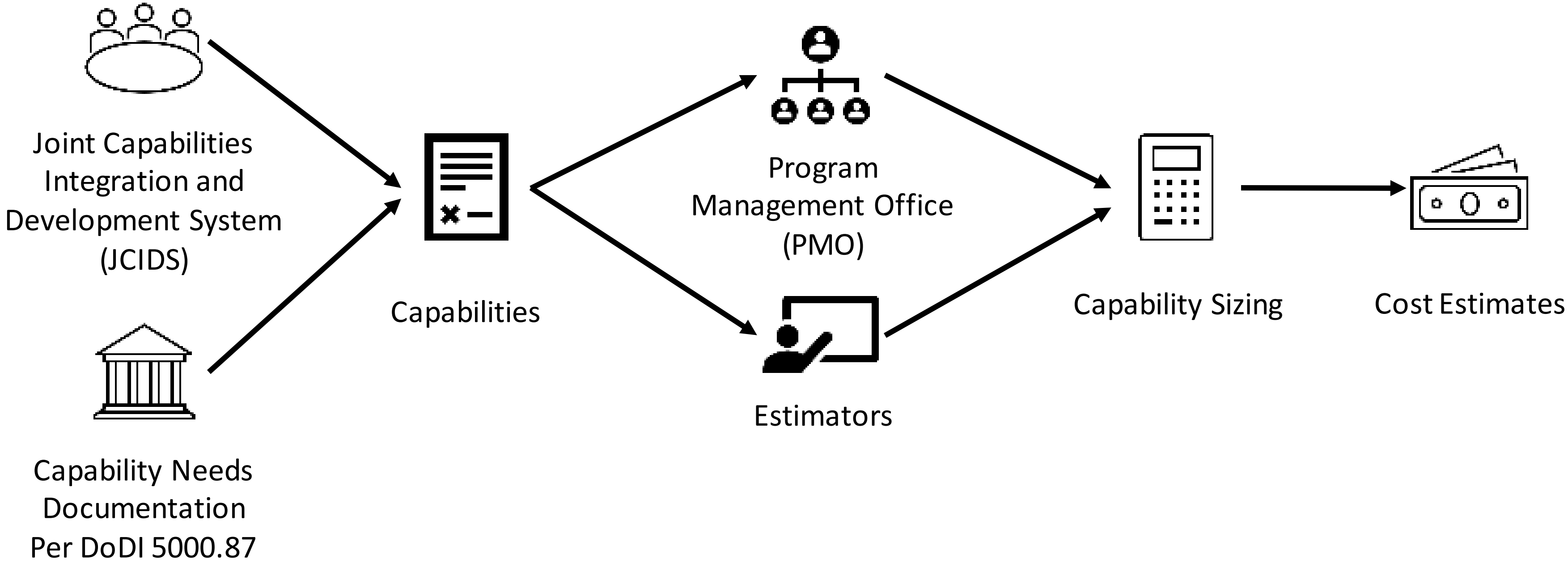
- Overcomes weaknesses of other size metrics
- 10% size error at early lifecycle, compared to
  - SLOC: 85% calculation error
  - Function Points: 18% calculation error
- Estimates within 90% confidence bound 90% of time, compared to
  - Historical average of 138% cost growth using traditional estimation
- 3-6 months learning time and 2-year transition time (50% more efficient), compared to
  - Function Points: ½ - 1 year learning time, 3-4 years transition time

Call for Participation

- Looking for capability and requirements documents (i.e., Capability Needs Statement (CNS) and Software Initial Capabilities Documents (SW-ICDs))
  - Ensure the capability groups/clusters we identify map to capability and functionality requirements
  - Evaluate how the capabilities evolve through the lifecycle

Backup Slides

Potential DoD Use Case – DoD Estimation Process



Potential DoD Use Case – Estimators' Steps