The NASA Software Cost Model: Formalizing analogy based cost estimation

Jairus Hihn
Jet Propulsion Laboratory, California Institute of Technology

Tim Menzies
North Carolina State University

Naveen Lekkalapudi
West Virginia University

James Johnson
National Aeronautics and Space Administration

October 21-23, 2014
29th International Forum on COCOMO and Systems Software Cost Modeling

© 2014. All rights reserved.
Introduction

- Purpose of this talk is to describe a new NASA Software Cost Model that is under development
- It is built around a spectral clustering algorithm that can be used to estimate software size and effort that is effective for
  - small sample sizes
  - noisy data
  - and uses high level systems information
Background

- The NASA Software CER Development Task is funded by the Cost Analysis Division to develop a software cost model that
  - Can be used in the early lifecycle
  - Can be used effectively by non-software specialists
  - Uses data from NASA in-house built and funded software “projects”
    - CADRe but also other Center level data sources
  - Supplement to current modeling and bottom up methods not a replacement
  - Can be documented as a paper model
  - Acceptable for use with both the cost and software communities
- Year 1 building a prototype model for robotic flight software
Theory
Why explore alternative modeling methods?

- For most of our history the cost community has relied upon regression type modeling methods
  - Regression method have the underlying assumption of
    - clean and complete data with large sample sizes
  - Cost data suffers from sparseness, noise, and small sample sizes
  - There are alternative methods that handle these conditions better than regression
Anscombe’s Quartet

Models especially regression models built on small samples with noisy data can be very misleading
Anscombe’s Quartet

All four of the displayed plots have virtually identical statistics
- Means, Medians, Variances
- Regression line, $R^2$, F and T tests
- But visual inspection clearly shows they are very different

MRE can distinguish between the models.

Plotting the absolute values of the relative error, it is easily seen that Model 3 fits its data best just as intuition would indicate.

MRE = Magnitude of Relative Error, \( \text{abs}(\text{Predicted} - \text{Actual})/\text{Actual} \)
Data mining techniques provided us with the rigorous tool set we needed to explore the many dimension of the problem we were addressing in a repeatable manner

- Analyze standard and non-standard models
  - Is there a best functional form
- Perform exhaustive searches over all parameters and records in order to guide data pruning
  - Rows (Stratification)
  - Columns (variable reduction)
- Measure model performance by multiple measures
  - $R^2$, MRE, Pred, F-test, etc.
- Is there a ‘best’ way to tune or calibrate a model
Effort Estimation with Data Mining Methods

References


Spectral Clustering

- PCA finds eigenvectors in numerical data
- Spectral Clustering
  - Spectral Clustering is like PCA on steroids but uses an eigenvector approximation method
  - Recursively splits the data on synthesized dimension of greatest variance/spread

- Why use it
  - Can handle numerical and symbolic data
  - Can work on small, sparse and somewhat noisy data sets but also works well on large consistent data sets
  - Can use as estimator with partial information
1. Select measure of distance
2. Pick point A at random (near middle works better)
3. Find furthest point from A (B)
4. Find Furthest point for B (C)
5. Draw line B-C
6. Project all points onto the line and find the median. This is first eigenvector.
7. Split data set by median point
8. Repeat and stop when subsample hits √N
9. Can use other stopping rules
Pure clustering

- Median measures always win
  - Has implications for our commonly used regression based models which are regression to the mean
- Interpolation beats centroid
  - Produces lower over all MRE
- **Median distance between two clusters is best**
  - Produces lower over all MRE
NASA
Analogy Software Estimation Tool
Estimation Experiments

Mission Descriptors

Size Distribution
- SLOC Range Estimate
- COCOMO Multiplier Range

Clusters

Ranges

COCOMO Monte Carlo Estimate

Spectral Clustering Effort Estimate

Model developed for this task

Total Effort CDF (Requirements through SW I&T)

Recommended Budget (70th Percentile) = 489.1 WM
Recommended Minimum (50th Percentile) = 402.7 WM

Likelihood of Occurrence
- Effort (Work Months)
- Size Distribution

Mission Descriptors

Model developed for this task
Data Items

- Total development effort in work months
- Delivered and equivalent logical lines
- COCOMO model inputs
  - Translated from CADRE which has SEER model inputs
- System parameters
  - Mission Type (deep-space, earth-moon, rover-lander, observatory)
  - Multiple element (probe, etc.)
  - Number of instruments (Simple, Medium&Complex)
  - Number of deployables (Simple, Medium&Complex)
  - Flight Computer Redundancy
  - Heritage
Data Sources

Where the data came from

- NASA CADRe (When it exists and is usable)
  - Cost Analysis Data Requirements archived in ONCE database
- Contributed Center level data
- NASA software inventory
- Project websites and other sources for system level information if not available in CADRe

In tomorrow's talk we will refer to this dataset as NASA 2010
# System Descriptor Details (Example)

## System Descriptors

<table>
<thead>
<tr>
<th>Mission Type</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth/Lunar Orbiter</td>
<td>Robotic spacecraft that orbit the earth or moon conducting science measurements. These spacecraft are very similar if not identical to the many commercial satellites used for communication as well as many military satellites. They often can have have high heritage and even use production line buses from industry.</td>
<td>Aqua</td>
<td></td>
</tr>
<tr>
<td>Telecomm Sat</td>
<td>Earth orbiters that support very high bandwidth and designed for very long life.</td>
<td>TDRS</td>
<td></td>
</tr>
<tr>
<td>Observatory</td>
<td>Observatories are space based telescopes that support space based astronomy across a wide set of frequencies. They can be earth orbiters or earth trailing at the various lagrange points created by the gravity fields of the earth, sun and moon.</td>
<td>Hubble</td>
<td></td>
</tr>
<tr>
<td>Deep Space</td>
<td>Any robotic spacecraft that goes beyond the moons orbit. So this category includes any mission whose destination is a planet, planetoids, any planetary satellite, comet, asteroid or the sun. These mission can be orbiters or flybys or a mixture of both.</td>
<td>Deep Impact</td>
<td></td>
</tr>
<tr>
<td>Static Lander</td>
<td>A robotic spacecraft that does its science in-situ or from the surface of a solar system body. It does not move from its original location.</td>
<td>Phoenix</td>
<td></td>
</tr>
<tr>
<td>Rover</td>
<td>A robotic spacecraft that does its science in-situ or from the surface of a solar system body and has the ability to move on the surface. To date all rovers have wheels but in the future they may crawl, walk or hop.</td>
<td>Mars Exploration Rover (MER)</td>
<td></td>
</tr>
</tbody>
</table>

※ Complete list is in the backup slides
- 39 records with system descriptors mostly from GSFC and JPL
- 19 records have all data items
- 31 records have delivered LOC
- 21 records have effort
Results so far are promising

- Remember that software size growth of 50-100%+ is not uncommon

Half the time, estimates within 40% of actual, using early life cycle data

3 major outliers need to look into
Comparing Estimates: Model vs Clustering

Clustering on Systems Parameters does almost as well as COCOMO or a Regression!

- Clustering using just high level system descriptors/variables estimates almost as good as running the COCOMO model or a simple regression
- LSR - Effort/EM = aS^b
- Results biased
- There is no inherent reason to assume with similar inputs that other models would perform any better

67-73% of estimates within +/-50% of actual, using early life cycle data
### NASA SW Cluster Estimation Prototype

**Example Clusters**

#### COCOMO EM

<table>
<thead>
<tr>
<th>Language</th>
<th>Mission Type</th>
<th>Complexity</th>
<th>Secondary Complexity</th>
<th>Corresponding Effort</th>
<th>Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>DeepSpace</td>
<td>None</td>
<td>Medium</td>
<td>10 0 2 LowtoNone</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>C/++/Assembly</th>
<th>DeepSpace</th>
<th>SampleReturn</th>
<th>Simple</th>
<th>2 1 3 DualString</th>
<th>Coldbackup</th>
<th>Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>3 3 3 3 4 4 4 4 3 2 3 3 3 3 4 2 3 3 6 3</td>
<td>DeepSpace</td>
<td>SampleReturn</td>
<td>Simple</td>
<td>2 1 3 DualString</td>
<td>Coldbackup</td>
<td>5 0 0 Veryhigh</td>
</tr>
<tr>
<td>2003</td>
<td>3 2 3 3 4 5 3 3 3 5 4 3 4 3 3 3 2 6 3</td>
<td>DeepSpace</td>
<td>None</td>
<td>Complex</td>
<td>0 7</td>
<td>DualString</td>
<td>Coldbackup</td>
</tr>
<tr>
<td>1997</td>
<td>4 3 3 3 4 4 3 2 3 3 3 3 3 5 4 2 3 3 5 3</td>
<td>DeepSpace</td>
<td>None</td>
<td>Simple</td>
<td>3 1</td>
<td>DualString</td>
<td>Coldbackup</td>
</tr>
<tr>
<td>1998</td>
<td>4 3 3 3 4 4 3 3 3 3 3 3 3 3 5 4 2 3 3 5 3</td>
<td>DeepSpace</td>
<td>None</td>
<td>Medium</td>
<td>2 1</td>
<td>DualString</td>
<td>Coldbackup</td>
</tr>
</tbody>
</table>

#### System Descriptors

- Rovers
- DeepSpace
- Earth/Moon

**Enter Data**

- Cost Analysis Division
- Jet Propulsion Laboratory
- West Virginia University
Conclusions and Next steps

- Initial results very promising:
  - Reasonably accurate LOC estimators for early lifecycle data
  - Effort estimators for early lifecycle data

- Next Steps under consideration
  - Expand and improve SC flight software data set and add Instrument flight software
  - Test with SEER-SEM

- Tomorrow
  - The results described within reveal that we can estimate with system descriptors with acceptable uncertainty ranges
  - But COCOMO and even linear regression can beat the proposed model when there is a complete data set
  - This raised the question we will address in tomorrows talk
    - Just how good is COCOMO
Back Up Slides
# System Descriptors - 1

## System Descriptors

<table>
<thead>
<tr>
<th>Mission Type</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth/Lunar Orbiter</td>
<td>Robotic spacecraft that orbit the earth or moon conducting science measurements. These spacecraft are very similar if not identical to the many commercial satellites used for communication as well as many military satellites. They often can have high heritage and even use production line buses from industry.</td>
<td>Aqua</td>
<td></td>
</tr>
<tr>
<td>Telecomm Sat</td>
<td>Earth orbiters that support very high bandwidth and designed for very long life.</td>
<td>TDRS</td>
<td></td>
</tr>
<tr>
<td>Observatory</td>
<td>Observatories are space based telescopes that support space based astronomy across a wide set of frequencies. They can be earth orbiters or earth trailing at the various lagrange points created by the garvoty fields of the earth, sun and moon.</td>
<td>Hubble</td>
<td></td>
</tr>
<tr>
<td>Deep Space</td>
<td>Any robotic spacecraft that goes beyond the moons orbit. So this category includes any mission whose destination is a planet, planetoids, any planetary satellite, comet, asteroid or the sun. These missions can be orbiters or flybys or a mixture of both.</td>
<td>Deep Impact</td>
<td></td>
</tr>
<tr>
<td>Static Lander</td>
<td>A robotic spacecraft that does its science in-situ or from the surface of a solar system body. It does not move from its original location.</td>
<td>Phoenix</td>
<td></td>
</tr>
<tr>
<td>Rover</td>
<td>A robotic spacecraft that does its science in-situ or from the surface of a solar system body and has the ability to move on the surface. To date all rovers have wheels but in the future they may crawl, walk or hop.</td>
<td>Mars Exploration Rover (MER)</td>
<td></td>
</tr>
</tbody>
</table>

## Secondary Element

<table>
<thead>
<tr>
<th>Secondary Element</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>No secondary element</td>
<td>Mars Reconcns Obiter (MRO)</td>
<td></td>
</tr>
<tr>
<td>Probe or Impactor</td>
<td>A simple impactor with little or no guidance and navigation capability and once released it simply transmits data from its instruments. A moderate-complexity impactor which may receive commands after separation, may have some internal guidance control, and several moderately complex instruments.</td>
<td>Cassini-Huygens was a simple probe. Deep Impact had a medium complexity probe.</td>
<td></td>
</tr>
<tr>
<td>Entry Descent and La</td>
<td>EDL can be simple with a ballistic trajectory or complex with precision landing and hazard avoidance. All landers and Rovers will have an EDL element.</td>
<td>Mars Pathfinder is an example of a simple EDL. MSL is an example of a complex EDL</td>
<td></td>
</tr>
<tr>
<td>Sample return</td>
<td>A simple sample return is like a simple probe but returning to earth. A complex sample return would be a return from a planet surface and requires an ascent stage.</td>
<td>Stardust is an example of a simple sample return</td>
<td></td>
</tr>
</tbody>
</table>
## System Descriptors - 2

<table>
<thead>
<tr>
<th>Over All Complexity</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple, Medium, or Complex</td>
<td>These are based on the mission type and secondary element so are derived from the descriptions above</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Instruments</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Any science instrument for which the FSW need only pass through commands and receive and store telemetry.</td>
<td>Magneometer</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Any science instrument for which the FSW must provide control logic that is relatively simple and requires no or only loose real time control. E.g., MER instruments.</td>
<td>MER Instruments</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>Any science instrument for which the FSW must provide control logic that is complicated or requires tight real-time control.</td>
<td>Telescope</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flight Computer Redundancy</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single String</td>
<td>Spacecraft has no redundancy in the flight computer</td>
<td>Most Earth Orbiters</td>
</tr>
<tr>
<td></td>
<td>Dual String - Cold backup</td>
<td>Spacecraft has redundant flight computers. Backup is normally off, is powered up and boots when prime string goes down</td>
<td>Most Deep space missions</td>
</tr>
<tr>
<td></td>
<td>Dual String - Warm backup</td>
<td>Backup computer is powered on and monitoring state of prime computer, but does not need to maintain continuous operation (e.g., a sequence may be restarted, attitude control restarts with last known state, etc.)</td>
<td>MSL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Deployables</th>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Simple deployable(s) which activate one time and remain in the deployed position for the duration of the mission.</td>
<td>Magnetometer boom</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Moderately complex deployables which require some sequencing of deployment events, or may require deployment and retraction.</td>
<td>Deployable Solar arrays</td>
</tr>
<tr>
<td></td>
<td>Complex</td>
<td>Complex deployables with detailed deployment sequences, many deployments and retractions which may require additional control algorithms to compensate for changing system characteristics, or deployables which are critical to mission safety and/or success.</td>
<td>Parachute, bag inflation and retraction, rover standup, ramp extension, complex robotic arms.</td>
</tr>
</tbody>
</table>
## Number of Deployables

<table>
<thead>
<tr>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>Simple deployable(s) which activate one time and remain in the deployed</td>
<td>Magnetometer boom</td>
</tr>
<tr>
<td></td>
<td>position for the duration of the mission.</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>Moderately complex deployables which require some sequencing of deployment</td>
<td>Deployable Solar arrays</td>
</tr>
<tr>
<td></td>
<td>events, or may require deployment and retraction.</td>
<td></td>
</tr>
<tr>
<td>Complex</td>
<td>Complex deployables with detailed deployment sequences, many deployments</td>
<td>Parachute, bag inflation and retraction,</td>
</tr>
<tr>
<td></td>
<td>and retractions which may require additional control algorithms to</td>
<td>rover standup, ramp extension, complex</td>
</tr>
<tr>
<td></td>
<td>compensate for changing system characteristics, or deployables which are</td>
<td>robotic arms).</td>
</tr>
<tr>
<td></td>
<td>critical to mission safety and/or success.</td>
<td></td>
</tr>
</tbody>
</table>

## Inheritance

<table>
<thead>
<tr>
<th>Values</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low to None</td>
<td>Software to be inherited has never flown in space. Significant new design</td>
<td>Mar Pathfinder or MSL EDL software</td>
</tr>
<tr>
<td></td>
<td>and basically all new code.</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>Basic design has been used before but significant portion is new design and</td>
<td>MSL</td>
</tr>
<tr>
<td></td>
<td>a code is newly written.</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Software to be inherited has flown in space and performed satisfactorily.</td>
<td>Many Planetary orbiters</td>
</tr>
<tr>
<td></td>
<td>Inherited SW architecture but majority of code is newly developed.</td>
<td></td>
</tr>
<tr>
<td>Very High</td>
<td>Software to be inherited was developed as a product line, has flown</td>
<td>Many earth orbiters</td>
</tr>
<tr>
<td></td>
<td>successfully in space at least once, has been successfully re-used in at</td>
<td></td>
</tr>
<tr>
<td></td>
<td>least two missions, and has extensive documentation.</td>
<td></td>
</tr>
</tbody>
</table>