Software Productivity Measurement
Using Multiple Size Measures

Barbara Kitchenham and Emila Mendes
IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 2004

2006-04-19
Kim, Su-hyun
Productivity

- Simple ratio of product size to project effort
- Broadly used to benchmarking
- Difficult to measure
  - No standard model for aggregating these measures

\[
\text{Productivity} = \frac{\text{Amount of output}}{\text{Unit of input used}}
\]
Research trend

- Common assumptions
  - Size must be related to effort
  - Reused components should not be included
- Productivity measure method
  - Used single size measures
    - FP, SLOC, System Meter, Magnitude, Use Case Point
    - Standard size/effort ratio
  - Approach on multidimensional size measure
Reifer’s work [IEEE SOFTWARE, 2000]
- Estimating quick-to-market software
- Size of web application
  - Using Halstead’s equation
  
  \[ V^* = N \log_2(n) = (N_1^* + N_2^*) \log_2 (n_1^* + n_2^*) \]

- \( N \) = number of total occurrences of operands and operators
- \( n \) = number of distinct operands and operators
- \( N_1^* \) = total occurrences of operand estimator
- \( N_2^* \) = total occurrences of operator estimator
- \( n_1^* \) = number of unique operands estimator
- \( n_2^* \) = number of unique operators estimator
- \( V^* \) = volume of work involved represented as Web Objects
Reifer’s work (cont’d)
- Web based predictors of size

<table>
<thead>
<tr>
<th>Web Object predictors</th>
<th>Example operands</th>
<th>Example operators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of building blocks</td>
<td>Fine grained components (ActiveX, DCOM, OLE, etc.), widgets, …</td>
<td>Create, apply, call, dispatch, interface, terminate, …</td>
</tr>
<tr>
<td>Number of COTS components (includes any wrapper code)</td>
<td>Commercial packages, library routines, objects like shopping carts, …</td>
<td>Initiate, terminate, apply, bind, customize, export, wrap, …</td>
</tr>
<tr>
<td>Number of multimedia files</td>
<td>Text, video, sound, 3D objects, plug-ins, metatags (no graphics files), …</td>
<td>Create, cut, paste, clear, edit, animate, broadcast, …</td>
</tr>
<tr>
<td>Number of object or application points° (or others proposed)</td>
<td># server data tables, # states, # client data tables, percent reuse, …</td>
<td>Transform (inputs to outputs), access, generate, modify,…</td>
</tr>
<tr>
<td>Number of xml, sgml, html and query language lines</td>
<td># lines including links to data attributes</td>
<td>Create, call, browse, link, find, search, retrieve, optimize,…</td>
</tr>
<tr>
<td>Number of Web components</td>
<td>Applets, agents, guards, …</td>
<td>Create, schedule, dispatch, …</td>
</tr>
<tr>
<td>Number of graphics files</td>
<td>Templates, pictures, images, …</td>
<td>Apply, align, import, export, insert, …</td>
</tr>
<tr>
<td>Number of scripts (visual language, audio, motion, and so forth)</td>
<td>Macros, containers, …</td>
<td>Create, store, edit, distribute, serializa, generalize,…</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stensrud and Myrtveit’s work [TSE’03]
- Identifying high performance ERP projects
  - Productivity is an indicator
  - The smallest project is always the most productive
- Relative measure
  - Variable returns to scale (VRS)

\[
P = \frac{y}{x} \quad \iff \quad x = \frac{1}{P} \cdot y^B
\]

- \(P\): productivity
- \(x\): effort
- \(y\): FP or SLOC
- \(B > 1\) (COCOMO)

\[
Effort = a \times \text{Size}^b
\]
Stensrud and Myrtveit’s work (cont’d)

- Increasing return to scale (small project)
- Constant return to scale (medium project)
- Decreasing return to scale (large project)
Motivation

- Mendes’s work [METRICS’03]
  - Early web size measures and effort prediction for web costimation (133 companies)
    - From Tukutuku database
      - Company-specific dataset
      - Multi-company dataset
      - Case-based reasoning
      - Stepwise regression

Best?
Considerations

- Finding elements related to efforts

- Combinations of size measures
  - Multiplicative parameters (non-linear)

- Obtaining the total effort
  - Person-hours
    - No information about the proportion of total effort spent constructing new elements and the proportion of total effort spent on reused elements
Assumptions

- Effort is related to several different size measures
- Size-based effort estimation model
  - The expected value of productivity is one
  - A value greater than one is indicative of good productivity

\[
Productivity = \frac{AdjustedSize}{Effort}. \quad (2)
\]

(Interestingly, the value of effort is not truly of interest for any sort of productivity cepartment, as it does not appear in the equation itself, with the exception of the denominator which is not generally of interest in practice.)
Assumptions (cont’d)

- Economies and diseconomies of scale should be accounted for

\[ \text{Effort} = a \times \text{Size}^b \]
\[ \text{Productivity} = \frac{\text{Size}}{a} \times \text{Size}^b = \text{Size}^{(1-b)} / a \]

<table>
<thead>
<tr>
<th>(b)</th>
<th>Economies and diseconomies of scale</th>
<th>size ↑ productivity ↓</th>
<th>Large size</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1</td>
<td>diseconomies of scale</td>
<td>size ↑ productivity ↓</td>
<td>Large size</td>
</tr>
<tr>
<td>&lt;1</td>
<td>economies of scale</td>
<td>size ↑ productivity ↑</td>
<td>Small size</td>
</tr>
<tr>
<td>=1</td>
<td>linear relationship</td>
<td>Productivity is constant</td>
<td>Medium size</td>
</tr>
</tbody>
</table>
Stepwise regression
  - Best fitting model (adjusted $R^2=0.6054$)
    - 54 web projects in 25 companies
    - Using STATA tool (http://www.stata.com/)
    - Reused web pages were not considered

$$\text{AdjustedSize} = 9.6 \times (\text{totWebPages})^{0.442}$$
$$\times (\text{highefffns} + 1)^{0.753}$$
$$\times (\text{newimages} + 1)^{0.13}. \quad (7)$$

<table>
<thead>
<tr>
<th>AdjustedSize</th>
<th>Total estimated effort (staff hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Effort includes all activities during development</td>
</tr>
<tr>
<td>totWebPages</td>
<td>Total number of web pages (new pages)</td>
</tr>
<tr>
<td></td>
<td>- excluding dynamically-generated pages</td>
</tr>
<tr>
<td>highefffns</td>
<td>Total number of high effort features/functions</td>
</tr>
<tr>
<td></td>
<td>- The function takes more than 12 hours</td>
</tr>
<tr>
<td>newimages</td>
<td>Total number of new images</td>
</tr>
</tbody>
</table>
Stepwise regression (cont’d)

- Model validation
  - Repeating the regression analysis excluding the high influence projects does not result in any major changes to the model coefficients
  - Asking data provider (13 projects)

<table>
<thead>
<tr>
<th>Project Id</th>
<th>Total effort</th>
<th>Estimated Effort</th>
<th>Total Web pages</th>
<th>New Web pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1786</td>
<td>283.42</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>363</td>
<td>2191.14</td>
<td>600</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>72.12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>625</td>
<td>69.49</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>32</td>
<td>3150</td>
<td>357.53</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

Mean : 1.58
Median : 0.98
Standard deviation : 1.945
Reused vs. no reused web pages
- Reuse is having a significant effect on productivity
Among different countries
- Excluding reuse factor
Conclusion

_constructing software productivity measure
- Combining several size measures
- Using regression-based effort prediction model
  - Built-in baseline (expected effort)
  - A value greater than one is indicative of good productivity
- Considering economies or diseconomies of scale

Current analysis is restricted to a specific data set
- Using random sample from a well-defined population
- Identifying robust conclusions
Critiques

- Applying only to the projects in the current database
- Absence of identifying project size
  - Small, medium, large project
- Undecided about reused project
- No consideration about input-side (actual effort)
My approach

- Productivity measurement on CS550 project
  - Comparing the result to well-known productivity measure methods
  - Finding factor that affects productivity

Measurement Tools

Effort report

Product size

Actual effort

Size report

Statistics Analysis

Factors

Experience, Rework, Meeting time

Survey
## Discussion (3/3)

### Milestone

<table>
<thead>
<tr>
<th>Milestone</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collecting effort data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making survey form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training statistics tool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting data set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data analysis (productivity)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applying hypothesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Software Engineering Lab, KAIST
Q & A

○ Thank you!!!
Product size metrics

- Source lines of code (SLOC)
  - The most practical (+)
  - The lack of international standards (-)
  - Produced by program generator (-)
  - Paradoxically decrease as the level of the language (-)
- Function points [Albrecht, 1970s]
  - Inputs, outputs, logical files, inquiries, and interfaces
  - High algorithmic complexity (-)
- Feature points [Software productivity Research, 1986]
  - Extended function point
  - Applied experimentally to embedded S/W
Tukutuku benchmarking project

- Variables
  - Web hypermedia applications
  - Web software applications

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webpages</td>
<td>number of Web pages</td>
</tr>
<tr>
<td>NewWP</td>
<td>number of Web pages created from scratch</td>
</tr>
<tr>
<td>WPCustomer</td>
<td>Web pages given by the customer</td>
</tr>
<tr>
<td>WPOutsourced</td>
<td>Web pages developed by third party</td>
</tr>
<tr>
<td>TextTyped</td>
<td>number of text pages typed (600 words)</td>
</tr>
<tr>
<td>TextElectronic</td>
<td>number of text pages in electronic format</td>
</tr>
<tr>
<td>TextScan</td>
<td>number of text pages that had to be scanned</td>
</tr>
<tr>
<td>ImgNew</td>
<td>number of new images developed</td>
</tr>
<tr>
<td>Img3rdParty</td>
<td>number of images developed by third party</td>
</tr>
<tr>
<td>ImgScanned</td>
<td>number of images that had to be scanned</td>
</tr>
<tr>
<td>ImgLibrary</td>
<td>number of images reused from a library</td>
</tr>
<tr>
<td>Animnew</td>
<td>number of new animations</td>
</tr>
<tr>
<td>AnimLib</td>
<td>number of animations reused from a library</td>
</tr>
<tr>
<td>AVNew</td>
<td>number of new audio/video files</td>
</tr>
<tr>
<td>AVLlib</td>
<td>number of reused audio/video files</td>
</tr>
<tr>
<td>TotdiffProducts</td>
<td>number of different products offered</td>
</tr>
<tr>
<td>Hfots</td>
<td>high effort FOTS (features off the shelf)</td>
</tr>
<tr>
<td>Hfots-a</td>
<td>high effort FOTS-A (features off the shelf adapted to the application)</td>
</tr>
<tr>
<td>Hnew</td>
<td>high effort new features</td>
</tr>
<tr>
<td>Tot-high</td>
<td>number of high effort features</td>
</tr>
<tr>
<td>Fots</td>
<td>low effort FOTS</td>
</tr>
<tr>
<td>Fots-a</td>
<td>low effort FOTS-A</td>
</tr>
<tr>
<td>New</td>
<td>low effort new features</td>
</tr>
<tr>
<td>Tot-low</td>
<td>number of low effort features</td>
</tr>
<tr>
<td>Teffort</td>
<td>effort to develop the Web application</td>
</tr>
</tbody>
</table>

http://www.cs.auckland.ac.nz/tukutuku
회귀분석 및 상관계수 (단순회귀분석)

둘 또는 그 이상의 변수 사이의 관계 특히 변수 사이의 인과관계를 분석하는 추측통계로 회귀분석은 독립변수가 하나인 경우와 2개 이상인 경우로 구분되는데, 하나인 경우를 단순회귀분석, 2개 이상인 경우를 다중회귀분석이라고 한다.

http://blog.naver.com/slmask?Redirect=Log&logNo=10002862273
상자-수염 그림 (box-whisker plot)

자료의 최소값, 제1사분위수, 중앙값, 제3사분위수, 최대값을 나타냄으로써 중심위치, 산포도, 왜도, 꼬리 길이 및 이상치 등의 정보를 제공해 주는 그림

IEEE International Software Metrics Symposium

- Main stream (since 1993)
  - Estimation of effort, maintenance, quality
  - Metrics
  - Empirical study
  - Data analysis

- New trend
  - Web application
  - Process Improvement
  - Open source
  - Function Point Adoption

By Dongwon Kang
System architecture
- Measure and Jasmine

Client side application
- Monitor main (Java)
- Win32 API
- Eclips plug-in

Server Side application
- Repository

UDP packet
Every 5 min

<activity>
Time
Application
filename

Working hours
By event hooking

Number of defects:
By JUnit test result
### Comparison table

<table>
<thead>
<tr>
<th>Measure</th>
<th>KAIST</th>
<th>ICU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer</td>
<td>Event hooking</td>
<td>Monitoring files</td>
</tr>
<tr>
<td></td>
<td>(When mouse on window)</td>
<td>(When writing a file)</td>
</tr>
<tr>
<td>Checking boundary</td>
<td>Predefined applications</td>
<td>Predefined working directory</td>
</tr>
<tr>
<td>Unit testing data</td>
<td>X</td>
<td>0</td>
</tr>
</tbody>
</table>
Measurement tools (3/3)

- Collected data set
  - <timestamp><id><application><filename>

Measure

Jasmine