SW/CIS Development Project Estimation:
An Overview

1. SW/CIS D&D Project planning involves estimating how much time, effort, money, and resources will be required to build a specific software system.

2. After the project scope is determined and the problem is decomposed into smaller problems, SW/CIS Project Managers use historical project data (as well as personal experience and intuition) to determine estimates for each part of the project.

3. The final estimates are typically adjusted by taking project complexity and multiple associated risks into account.

4. Managers will not know that they have done a good job estimating until the project is over.

5. It is essential to track resources and revise estimates as project progresses.
Software D&D Project Estimation

- Project scope must be understood
- Elaboration (decomposition) is necessary
- Historical metrics are very helpful
- At least several (2…4) different techniques should be used
- Uncertainty is inherent in the process of estimation

Estimation Techniques

1. Historic Data-based (past, similar) project experience

2. Decomposition techniques:
   - task breakdown and effort estimates
   - functional decomposition
   - size-based (e.g., FP) estimates

3. Empirical models

4. Automated tools
1. Historic Data-Based Estimate

2. Decomposition Techniques

- **Software sizing** (LOC, function points, standard components, change)

- **Problem-based estimation** (using LOC decomposition focuses on software functions, using FP decomposition focuses on information domain characteristics)

- **Process-based estimation** (decomposition based on tasks required to complete the software process framework)

- **Use-case estimation** (promising, but controversial due to lack of standardization of use cases)
Conventional Estimation Techniques: Functional Decomposition (based on functions to be developed)

Statement of Scope

Perform a Grammatical “parse”

Example: Lines of Code (LOC)-Based Approach (CAD Systems)

<table>
<thead>
<tr>
<th>Function (in Computer-Aided Design, or CAD System)</th>
<th>Estimated LOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Interface (UI)</td>
<td>2,300</td>
</tr>
<tr>
<td>Two-dimensional geometric analysis (2DGA)</td>
<td>5,300</td>
</tr>
<tr>
<td>Three-dimensional geometric analysis (3DGA)</td>
<td>6,800</td>
</tr>
<tr>
<td>Database Management (DBM)</td>
<td>3,300</td>
</tr>
<tr>
<td>Computer Graphics Display Functions</td>
<td>4,900</td>
</tr>
<tr>
<td>Control of Peripherals</td>
<td>2,100</td>
</tr>
<tr>
<td>Design Analysis and Dynamic Simulation Modules</td>
<td>8,400</td>
</tr>
<tr>
<td><strong>Total (estimated lines of code)</strong></td>
<td><strong>33,200</strong></td>
</tr>
</tbody>
</table>

1. Average productivity for systems of this type = 620 LOC/pm.
2. Burdened labor rate = $8000 per month, the cost per line of code is approximately $13.
3. Based on the LOC estimate and the historical productivity data, the total estimated project cost is $431,000 and the estimated effort is 54 person-months.
Example: Function Points (FP)-Based Approach

<table>
<thead>
<tr>
<th>Information Domain Value</th>
<th>Optimal forgiveness</th>
<th>Number of FP estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of inputs</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>Number of outputs</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>Number of inquiries</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>Number of data internal structures</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Number of external interfaces</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>320</td>
</tr>
</tbody>
</table>

The estimated number of FP is derived:

\[ \text{FP}_{\text{estimated}} = \text{count-total} \times 0.65 + 0.01 \times \left( \sum F_i \right) \]

\[ \text{FP}_{\text{estimated}} = 375 \]

organizational average productivity = 6.5 FP/pm.

burdened (above and beyond) labor rate = $8000 per month, approximately $1230/FP.

Based on the FP estimate and the historical productivity data, total estimated project cost is $461,000 and estimated effort is 58 person-months.

“LOC per FP” Correspondence Table (by QSM)

http://www.qsm.com/resources/function-point-languages-table
CIS Engineering Process-Based (Framework Activities)-Based Estimation

Obtained from “process framework”

<table>
<thead>
<tr>
<th>application functions</th>
<th>framework activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort required to accomplish per each framework activity for each application function</td>
<td></td>
</tr>
</tbody>
</table>

### Process-Based Estimation Example

<table>
<thead>
<tr>
<th>Activity</th>
<th>CC</th>
<th>Planning</th>
<th>Risk Analysis</th>
<th>Engineering</th>
<th>Construction Release</th>
<th>CE</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td></td>
<td>analyses</td>
<td>design</td>
<td>code</td>
<td>test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UICF</td>
<td>0.50</td>
<td>2.50</td>
<td>0.40</td>
<td>5.00</td>
<td>n/a</td>
<td>8.40</td>
<td></td>
</tr>
<tr>
<td>2DGA</td>
<td>0.75</td>
<td>4.00</td>
<td>0.60</td>
<td>2.00</td>
<td>n/a</td>
<td>7.35</td>
<td></td>
</tr>
<tr>
<td>3DGA</td>
<td>0.50</td>
<td>4.00</td>
<td>1.00</td>
<td>3.00</td>
<td>n/a</td>
<td>8.50</td>
<td></td>
</tr>
<tr>
<td>CGDF</td>
<td>0.50</td>
<td>3.00</td>
<td>1.00</td>
<td>1.50</td>
<td>n/a</td>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>DSM</td>
<td>0.50</td>
<td>3.00</td>
<td>0.75</td>
<td>1.50</td>
<td>n/a</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>PCF</td>
<td>0.25</td>
<td>2.00</td>
<td>0.50</td>
<td>1.50</td>
<td>n/a</td>
<td>4.25</td>
<td></td>
</tr>
<tr>
<td>LRM</td>
<td>0.50</td>
<td>2.00</td>
<td>0.50</td>
<td>2.00</td>
<td>n/a</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>0.25</td>
<td>2.25</td>
<td>0.35</td>
<td>3.50</td>
<td>20.50</td>
<td>4.50</td>
<td>16.50</td>
</tr>
<tr>
<td>% effort</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>8%</td>
<td>45%</td>
<td>10%</td>
<td>36%</td>
</tr>
</tbody>
</table>

Based on an average burdened labor rate of $8,000 per month, the total estimated project cost is $368,000 and the estimated effort is 46 person-months.
Estimation with Use-Cases

<table>
<thead>
<tr>
<th>Use Case</th>
<th>Scenarios</th>
<th>Pages</th>
<th>LOC</th>
<th>LOC Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>User interface subsystem</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>3.36</td>
</tr>
<tr>
<td>Engineering subsystem group</td>
<td>10</td>
<td>20</td>
<td>8</td>
<td>31.23</td>
</tr>
<tr>
<td>Infrastructure subsystem group</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>7.97</td>
</tr>
<tr>
<td>Total LOC Estimate</td>
<td></td>
<td></td>
<td></td>
<td>42.56</td>
</tr>
</tbody>
</table>

Using 620 LOC/pm as the average productivity for systems of this type and a burdened labor rate of $8000 per month, the cost per line of code is approximately $13. Based on the use-case estimate and the historical productivity data, the total estimated project cost is $552,000 and the estimated effort is 68 person-months.

3. Empirical Estimation Models
(based on best cases=previous projects in industry or company)

General form:

\[ \text{effort} = \text{tuning coefficient} \times \text{size} \]

- Usually derived as person-months of effort required
- Usually LOC but may also be function point
- Either a constant or a number derived based on complexity of project
- Empirically derived

...
Empirical Estimation Models

General form:

\[ \text{effort} = \text{tuning coefficient} \times \text{size}^{\text{exponent}} \]

- Usually derived as person-months of effort required
- Either a constant or a number derived based on complexity of project


COCOMO-II Model

- COCOMO II is actually a hierarchy of estimation models that address the following areas:
  - **Application composition model.** Used during the early stages of software engineering, when prototyping of user interfaces, consideration of software and system interaction, assessment of performance, and evaluation of technology maturity are paramount.
  - **Early design stage model.** Used once requirements have been stabilized and basic software architecture has been established.
  - **Post-architecture-stage model.** Used during the construction of the software.
CS/CIS Development Empirical Model

A dynamic multivariable model

\[ E = \left[ \text{LOC} \times B^{0.333}/P \right]^3 \times (1/t^4) \]

where

- \( E \) = effort in person-months or person-years
- \( t \) = project duration in months or years
- \( B \) = “special skills factor”
- \( P \) = “productivity parameter”

4. Automated Tools

1. OHLOH system (LOC → rates → project cost)
2. K-LOC calculator
3. CLOC calculator
4. etc.

Different approach: use SourceForge to get OSS system

Free, secure and fast downloads from the largest Open Source applications and software directory - SourceForge.net
Automated tools for LOC-based approach
(example: http://www.openhub.net/)

- Go to Projects
- Specify Project Name
- Open the link
- Click on Code Analysis for lines of code...

Mozilla Firefox
https://www.openhub.net/p/firefox/analyses/latest/languages_summary
Mozilla Firefox

LOC by Language

<table>
<thead>
<tr>
<th>Language</th>
<th>Code Lines</th>
<th>Comment Lines</th>
<th>Comment Ratio</th>
<th>Blank Lines</th>
<th>Total Lines</th>
<th>Total Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>C++</td>
<td>4,035,234</td>
<td>1,275,220</td>
<td>17.2%</td>
<td>982,641</td>
<td>5,008,100</td>
<td>26.2%</td>
</tr>
<tr>
<td>ENGLISH</td>
<td>3,995,564</td>
<td>915,378</td>
<td>15.9%</td>
<td>110,682</td>
<td>4,112,612</td>
<td>21.4%</td>
</tr>
<tr>
<td>C</td>
<td>3,145,091</td>
<td>854,816</td>
<td>27.3%</td>
<td>369,850</td>
<td>3,452,935</td>
<td>16.9%</td>
</tr>
<tr>
<td>HTML</td>
<td>1,044,052</td>
<td>115,120</td>
<td>5.6%</td>
<td>181,285</td>
<td>1,235,417</td>
<td>11.2%</td>
</tr>
<tr>
<td>PHP</td>
<td>118,669</td>
<td>17,870</td>
<td>2.6%</td>
<td>80,859</td>
<td>217,827</td>
<td>5.0%</td>
</tr>
<tr>
<td>Java</td>
<td>11,040</td>
<td>146,235</td>
<td>22.0%</td>
<td>120,270</td>
<td>277,525</td>
<td>4.8%</td>
</tr>
<tr>
<td>CSS</td>
<td>278,154</td>
<td>1,103,260</td>
<td>39.9%</td>
<td>52,480</td>
<td>370,273</td>
<td>2.8%</td>
</tr>
<tr>
<td>Script</td>
<td>103,185</td>
<td>29,070</td>
<td>11.8%</td>
<td>26,914</td>
<td>139,169</td>
<td>1.0%</td>
</tr>
<tr>
<td>Others</td>
<td>158,649</td>
<td>65,964</td>
<td>57.0%</td>
<td>27,204</td>
<td>115,520</td>
<td>1.6%</td>
</tr>
<tr>
<td>Autoconf</td>
<td>136,039</td>
<td>2,127</td>
<td>1.6%</td>
<td>18,000</td>
<td>157,256</td>
<td>0.8%</td>
</tr>
<tr>
<td>Makefile</td>
<td>16,973</td>
<td>15,670</td>
<td>91.4%</td>
<td>1,236</td>
<td>19,929</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

K-LOC Calculator

http://www.analogx.com/contents/download/Programming/kloc/Freeware.htm
CLOC Calculator
http://cloc.sourceforge.net/

Libraries of Open Source Software
(like http://sourceforge.net/)
Source Code of Web apps

(for ex: Open Source Software
(like http://sourceforge.net/)

“LOC per FP” Correspondence Table (by QSM)

http://www.qsm.com/resources/function-point-languages-table
Topic 12

SW/CIS Project Estimates
(LOC, FP, efforts, cost, etc.)

Additional information.

The Make-Buy Decision

system X

build

reuse

buy

contract

simple (0.30)
difficult (0.70)
minor changes (0.40)
major changes (0.50)
simple (0.20)
complex (0.80)
minor changes (0.70)
major changes (0.30)
without changes (0.30)
with changes (0.40)

$380,000
$460,000
$275,000
$310,000
$490,000
$210,000
$400,000
$360,000
$500,000
Computing Project Expected Cost

\[
\text{expected cost} = \sum \text{(path probability)} \times \text{(estimated path cost)}
\]

For example, the expected cost to build is:

\[
\text{expected cost}_{\text{build}} = 0.30 \times $380,000 + 0.70 \times $450,000 = 0.30 \times $380,000 + 0.70 \times $450,000 = $429,000
\]

Similarly,

\[
\text{expected cost}_{\text{reuse}} = $382,000
\]
\[
\text{expected cost}_{\text{buy}} = $267,000
\]
\[
\text{expected cost}_{\text{contract}} = $410,000
\]