Preliminary Causal Analysis Results with Software Cost Estimation Data

Anandi Hira, Bob Stoddard, Mike Konrad, Barry Boehm

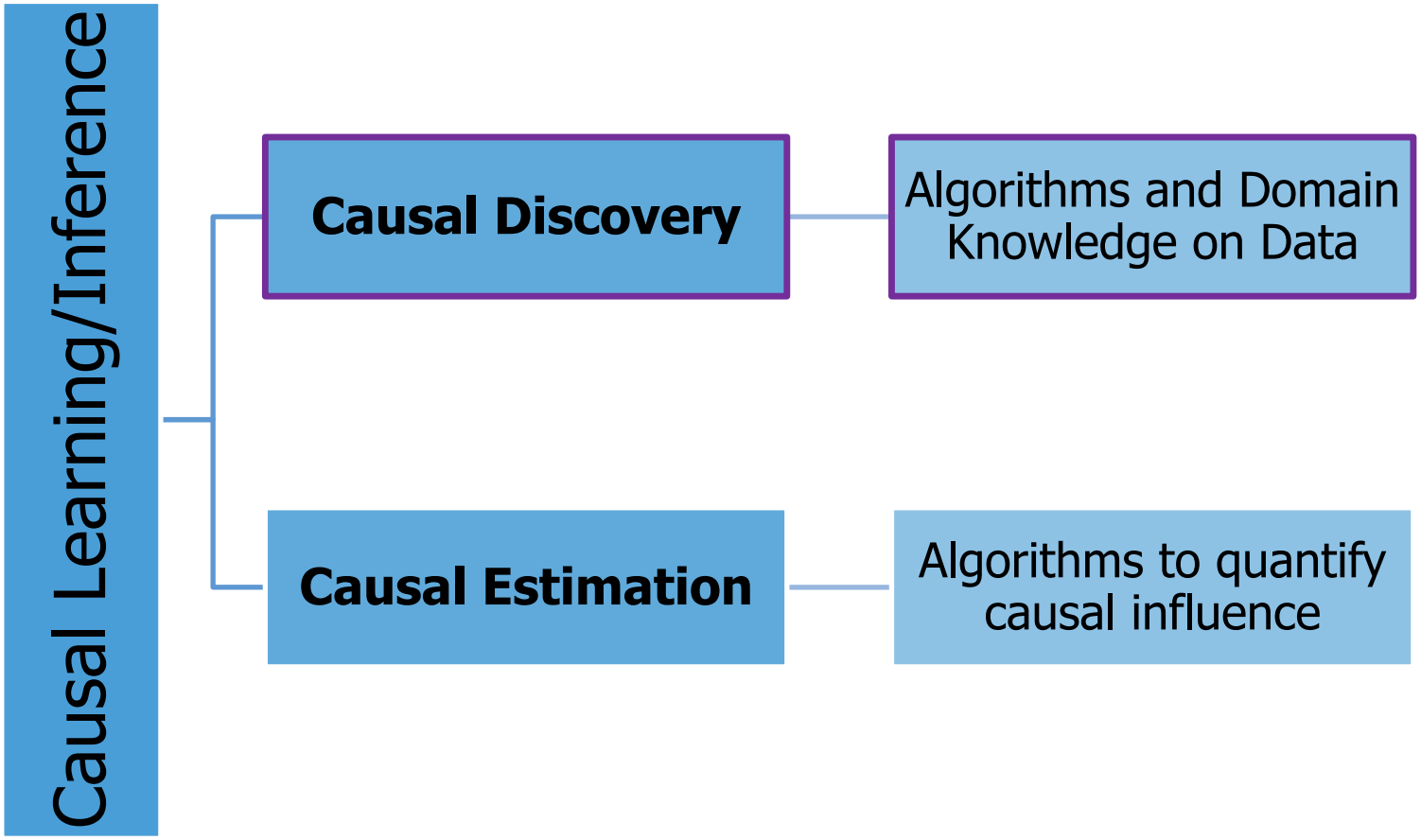
Parametric Cost Models

COCOMO® II

$$Effort = 2.94 \times Size^{E} \times \prod_{i=1}^{17} EM_i$$

- Input: size, product and personnel attributes
- Effort in Person-Months (PM)
- Size in KSLOC (1000 SLOC)
- Domain Experts
- Data calibration
- No causal analysis

Causal Inference



Past Causal-Type Analyses

Dr. Boehm COCOMO® 81

- In-depth behavioral analyses for effort factors

Evidence-Based SE

- Experiments
- Cause precede effect
- Cause covaries with effect
- Alternative explanations are implausible

Cuoto et al

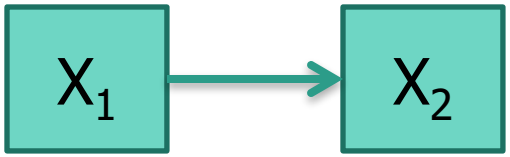
- Granger's causality test for software defect predictability
  - Doesn't get to heart of causality

Hu et al

- Bayesian networks with causality constraints for software risk factors

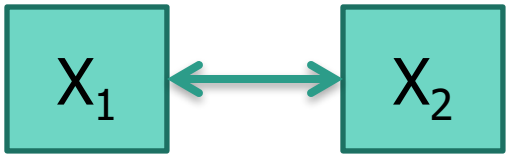
PC Search

- Named after Peter Spirtes and Clark Glymour
- First scalable discovery algorithm



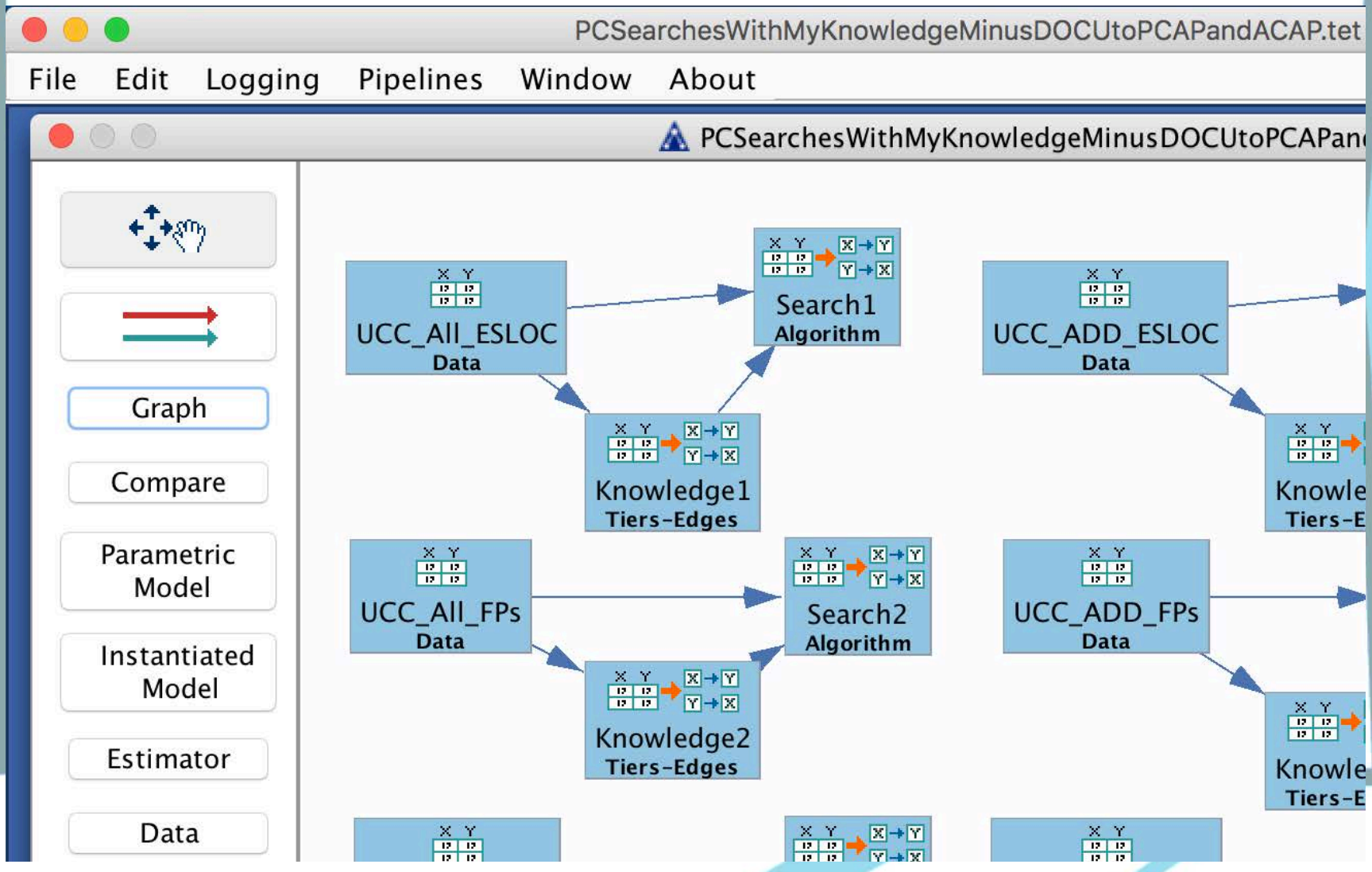
Change in X_1 causes change in X_2

Insufficient data to select orientation



May be common confounder of both variables, missing from dataset

Tetrad


PCSearchesWithMyKnowledgeMinusDOCUtoPCAPandACAP.tet
File
Edit
Logging
Pipelines
Window
About
PCSearchesWithMyKnowledgeMinusDOCUtoPCAPan
Graph
Compare
Parametric Model
Instantiated Model
Estimator
Data
UCC_All_ESLOC
Data
Search1
Algorithm
UCC_ADD_ESLOC
Data
Knowledge1
Tiers-Edges
UCC_All_FPs
Data
Search2
Algorithm
UCC_ADD_FPs
Data
Knowledge2
Tiers-Edges

Dataset: Unified Code Count (UCC)

Project Description

- Maintained at USC
- Code metrics tool (logical SLOC, cyclomatic complexity)
- Implemented in C++
- 45 to 1425 logical SLOC
- 2010 to 2014
- Modularized architecture
- 4-month time-boxed increments

Project Types

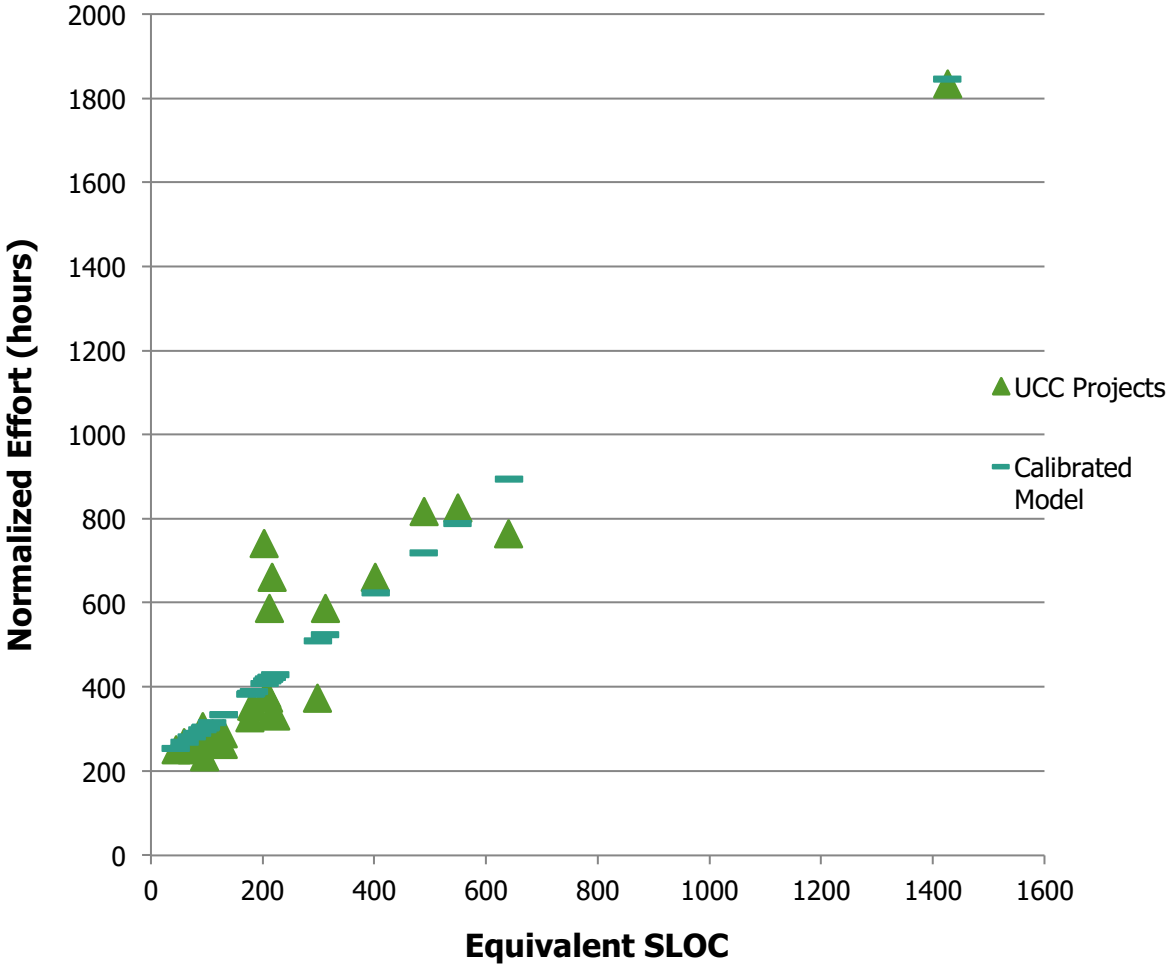
- Add Functions
  - New language parsers
  - New features, such as GUI
- Modify Functions
  - Cyclomatic complexity support (modify existing language parsers with mathematical operation and algorithms)

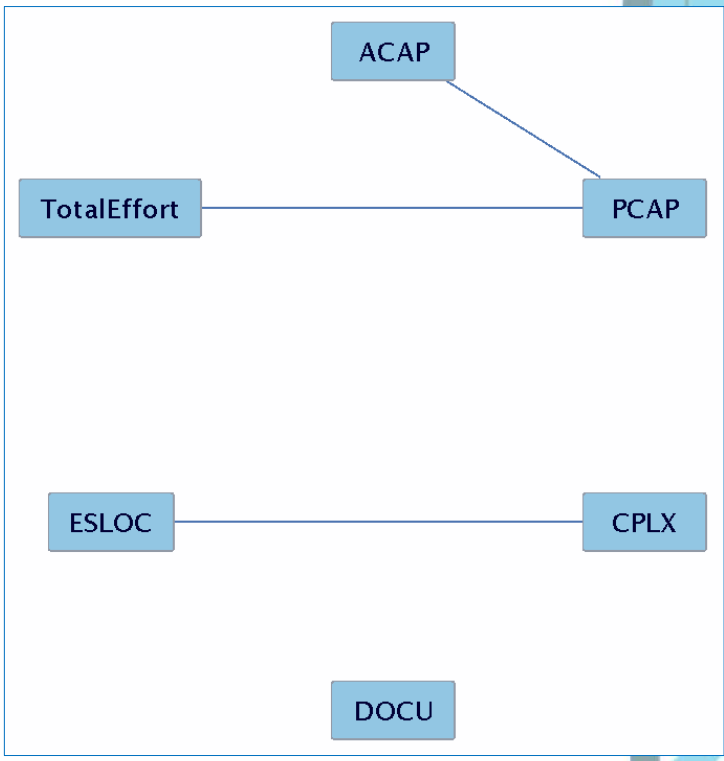
Dataset Attributes

1. Equivalent SLOC
2. IFPUG Function Points
3. IFPUG Software Non-functional Assessment Process
4. COSMIC Function Points
5. Total Effort
6. Applications Experience
7. Platform Experience
8. Use of Software Tools
9. Personnel Continuity
10. Documentation Match to Needs
11. Analyst Capability
12. Programmer Capability
13. Product Complexity

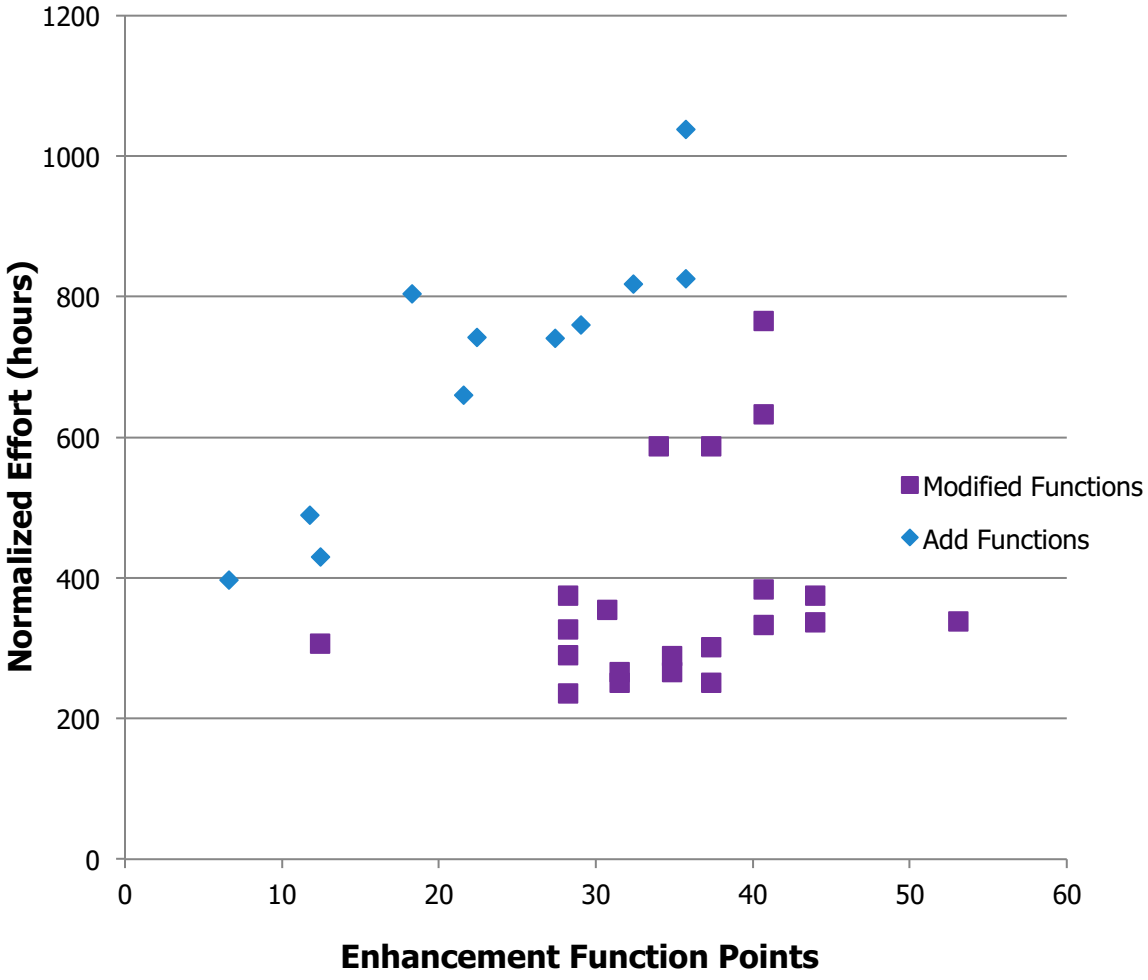
All Data Points

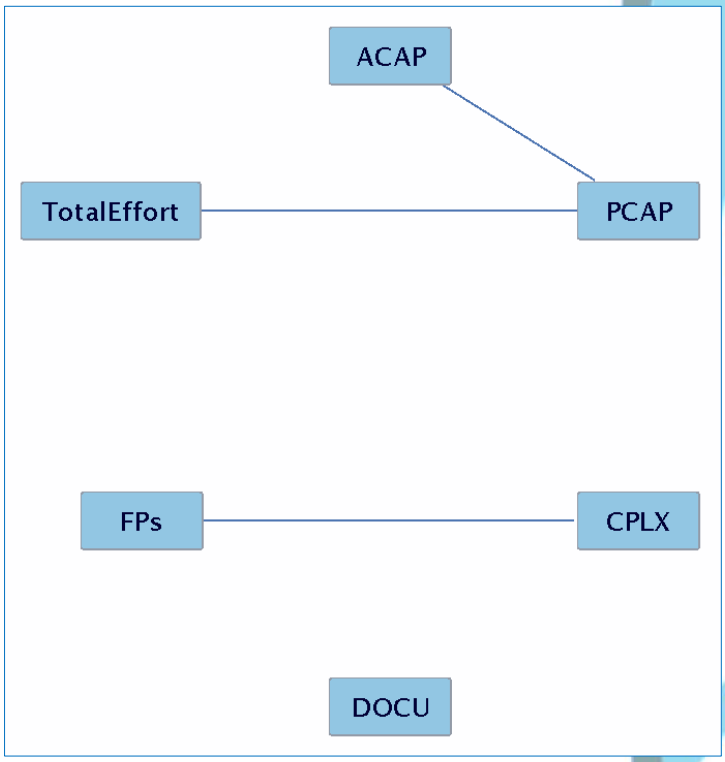
ESLOC


2000
1800
1600
1400
1200
1000
800
600
400
200
0
Normalized Effort (hours)
0
200
400
600
800
1000
1200
1400
1600
Equivalent SLOC
UCC Projects
Calibrated Model

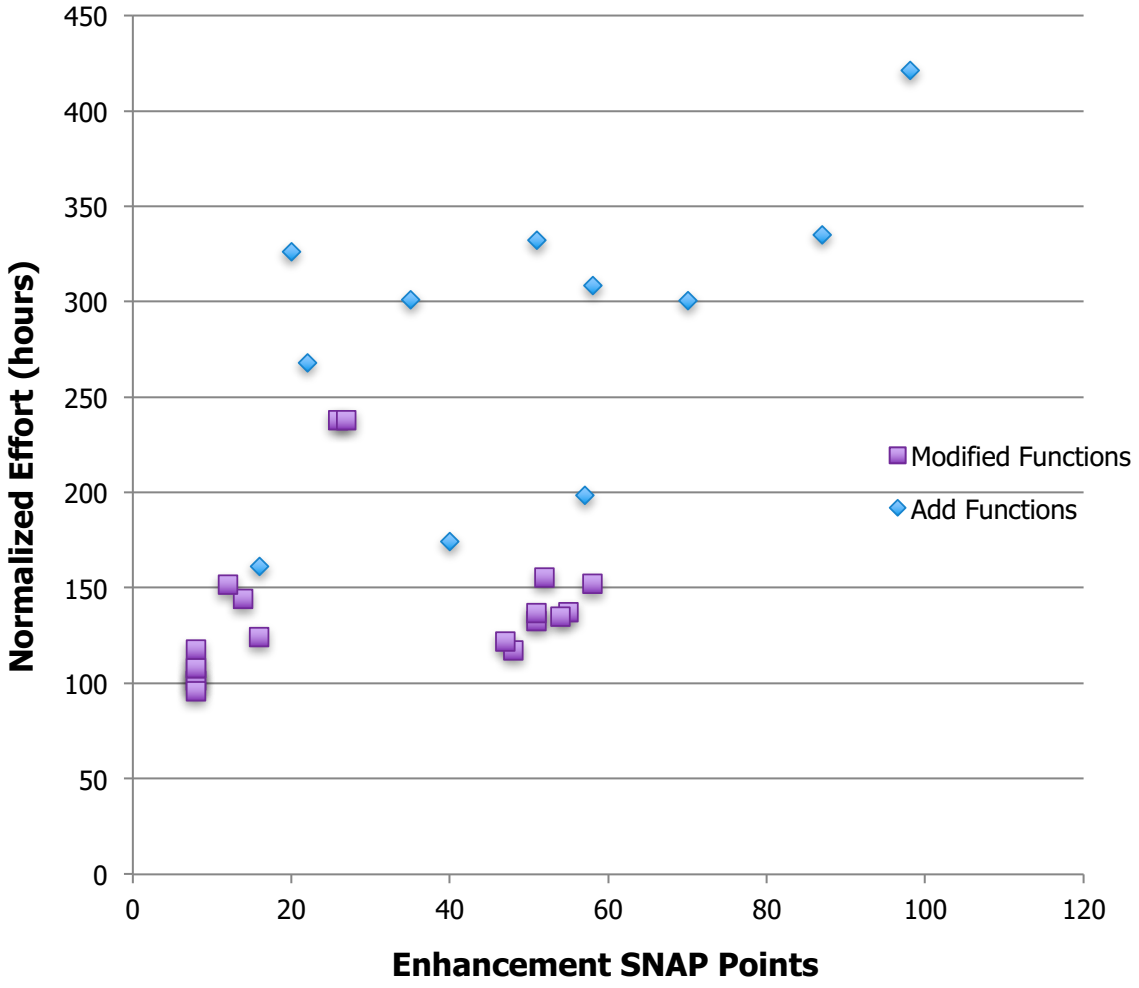

ACAP
TotalEffort
PCAP
ESLOC
CPLX
DOCU

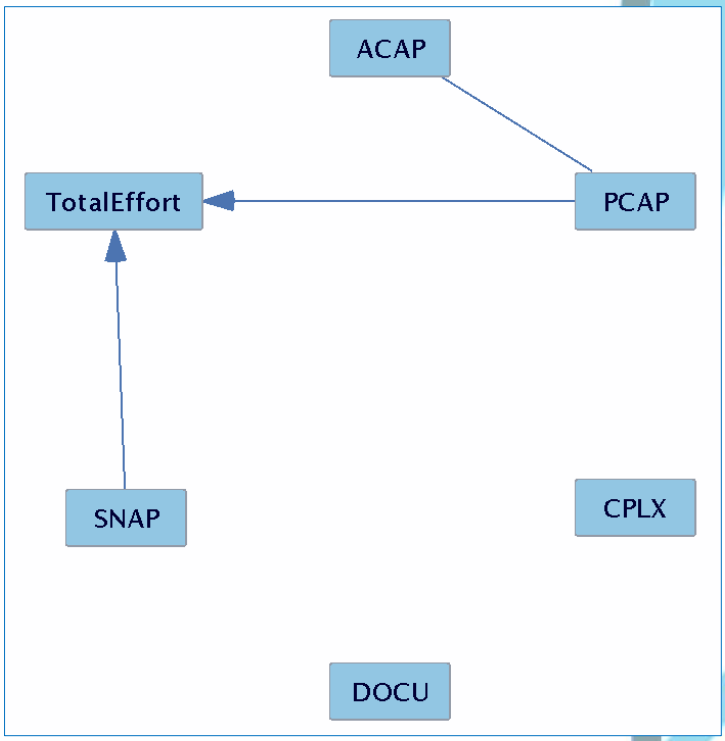
IFPUG FPs


1200
1000
800
600
400
200
0
0
10
20
30
40
50
60
Normalized Effort (hours)
Enhancement Function Points
Modified Functions
Add Functions

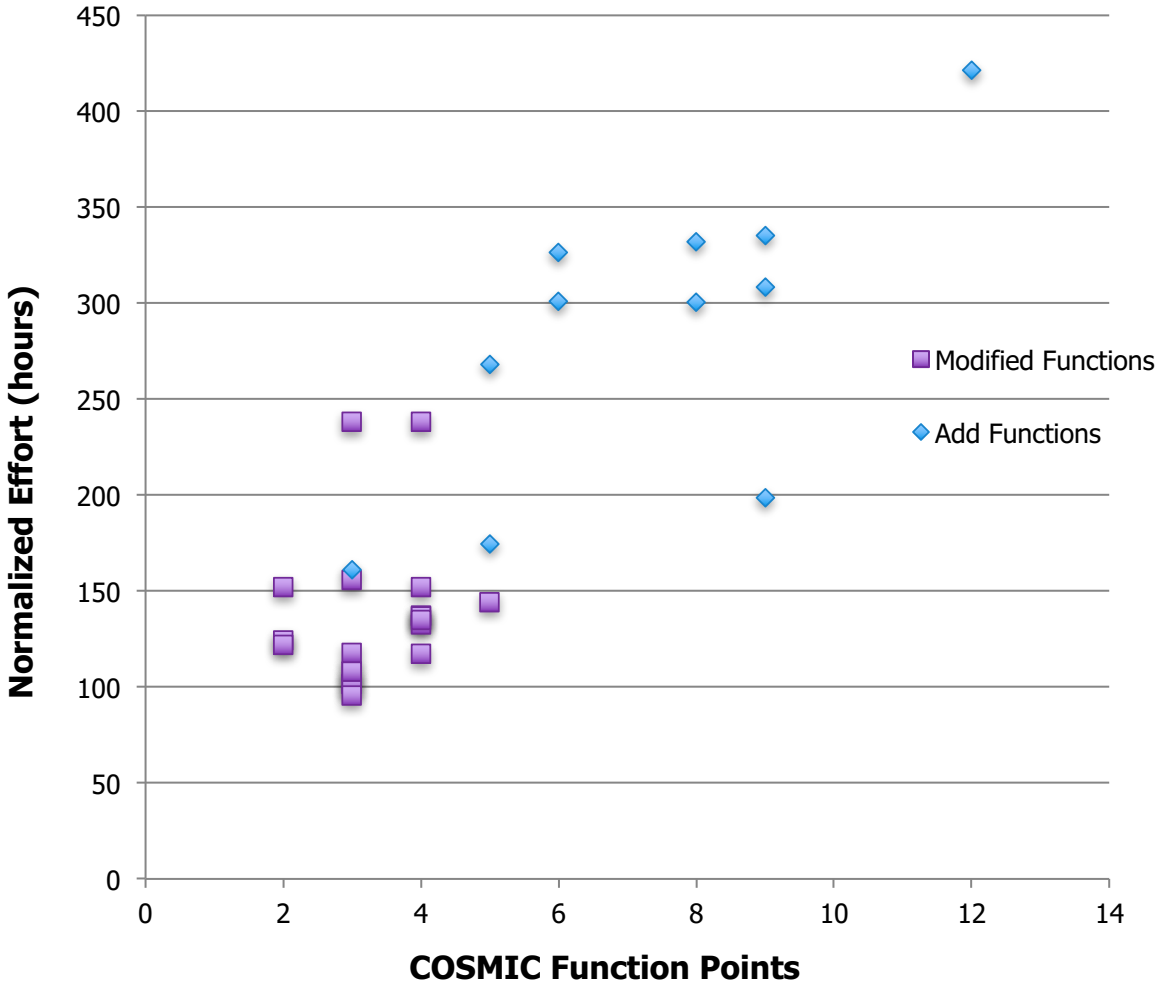

ACAP
TotalEffort
PCAP
FPs
CPLX
DOCU

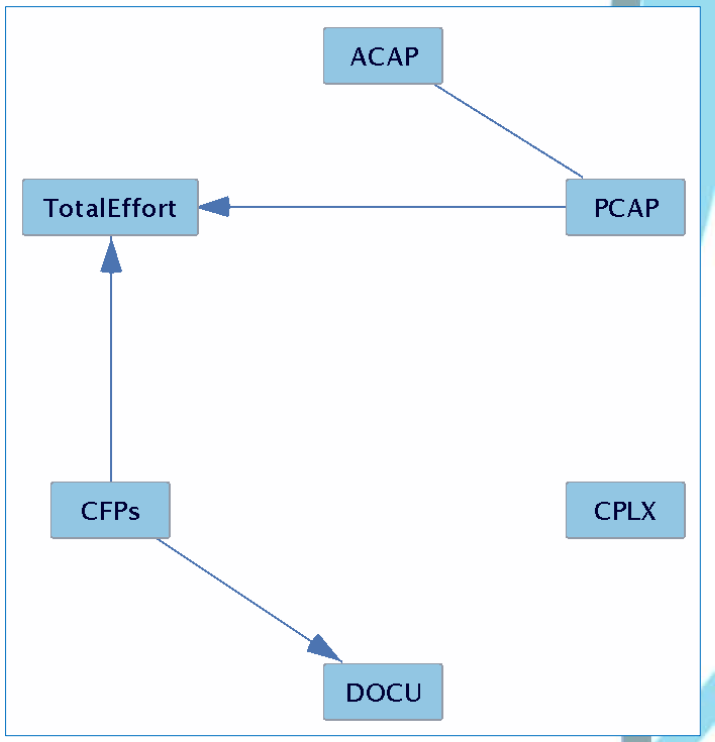
IFPUG SNAP

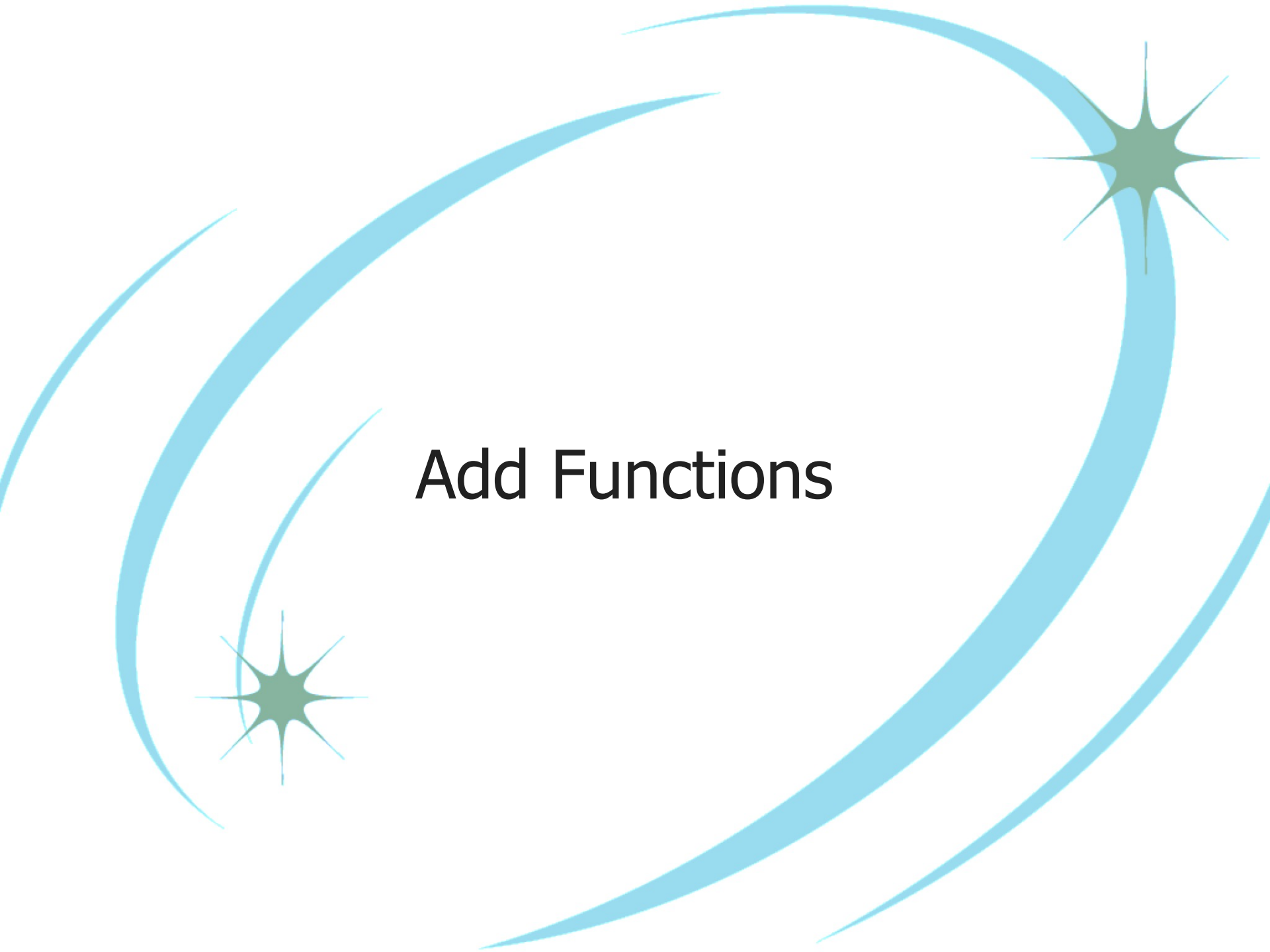

450
400
350
300
250
200
150
100
50
0
0
20
40
60
80
100
120
Normalized Effort (hours)
Enhancement SNAP Points
Modified Functions
Add Functions


ACAP
TotalEffort
PCAP
SNAP
CPLX
DOCU

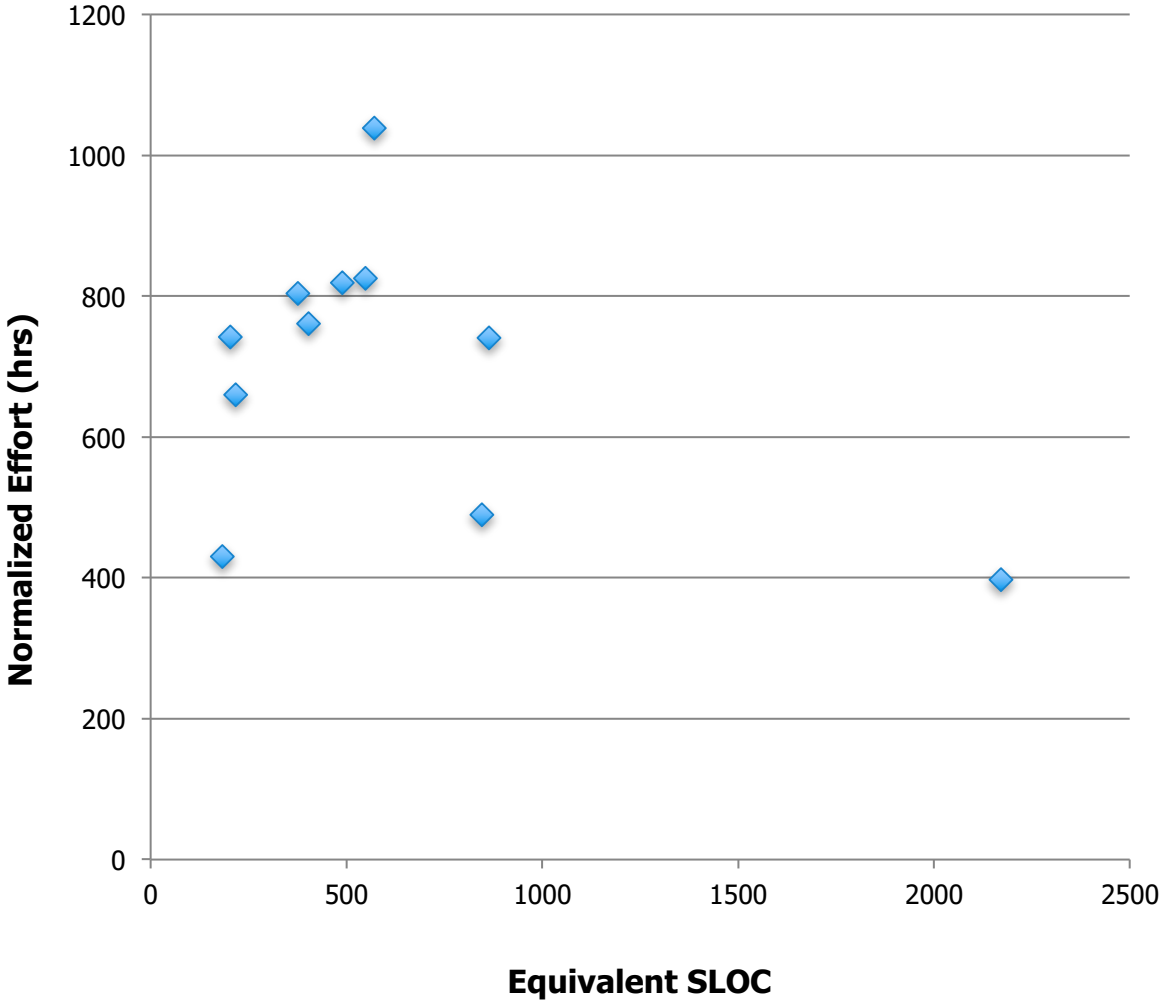
COSMIC FPs

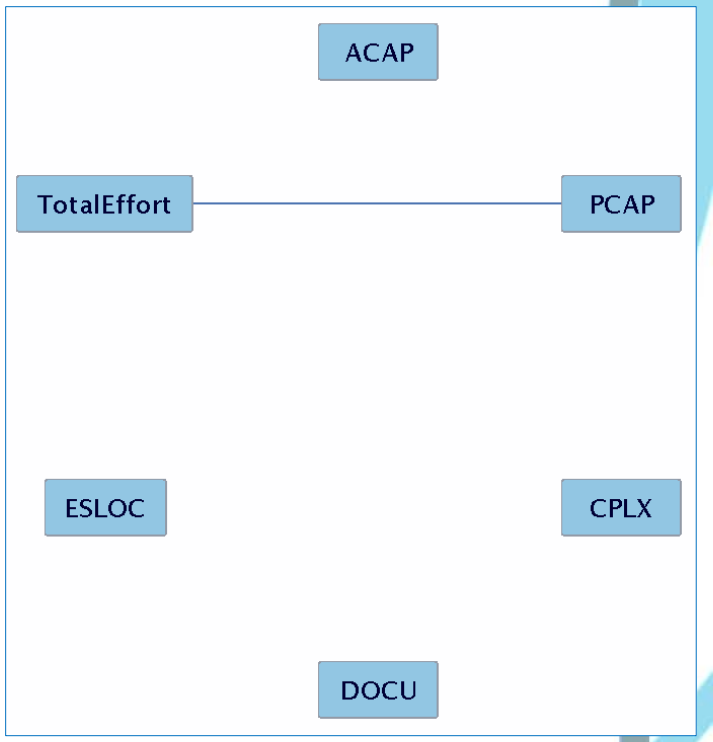

450
400
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
14
Normalized Effort (hours)
COSMIC Function Points
Modified Functions
Add Functions


ACAP
PCAP
TotalEffort
CFPs
CPLX
DOCU

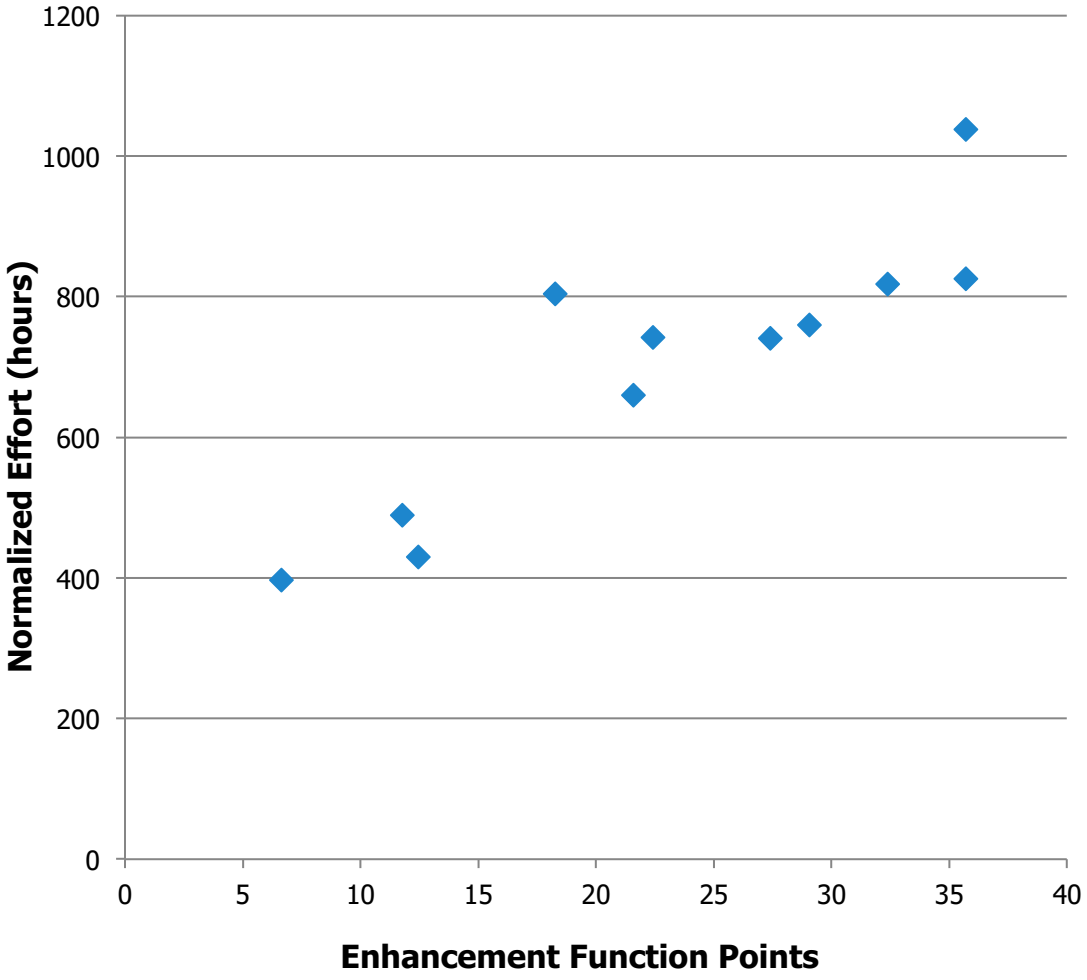

Add Functions

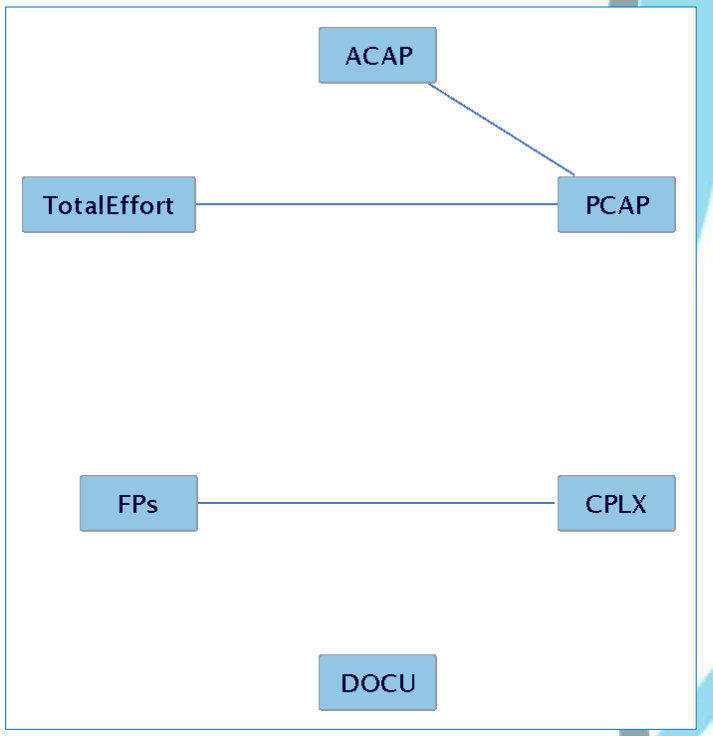
ESLOC


1200
1000
800
600
400
200
0
0
500
1000
1500
2000
2500
Normalized Effort (hrs)
Equivalent SLOC

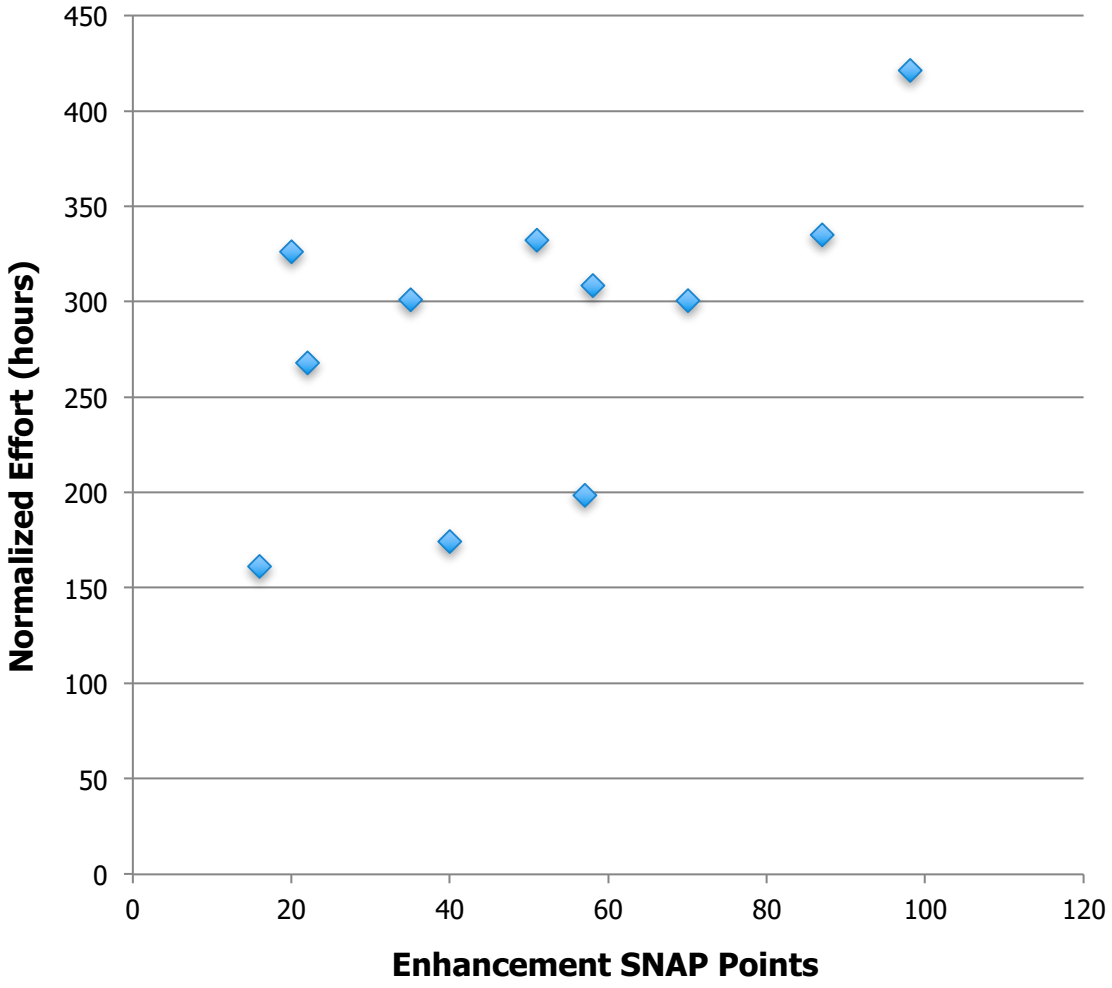

ACAP
TotalEffort
PCAP
ESLOC
CPLX
DOCU

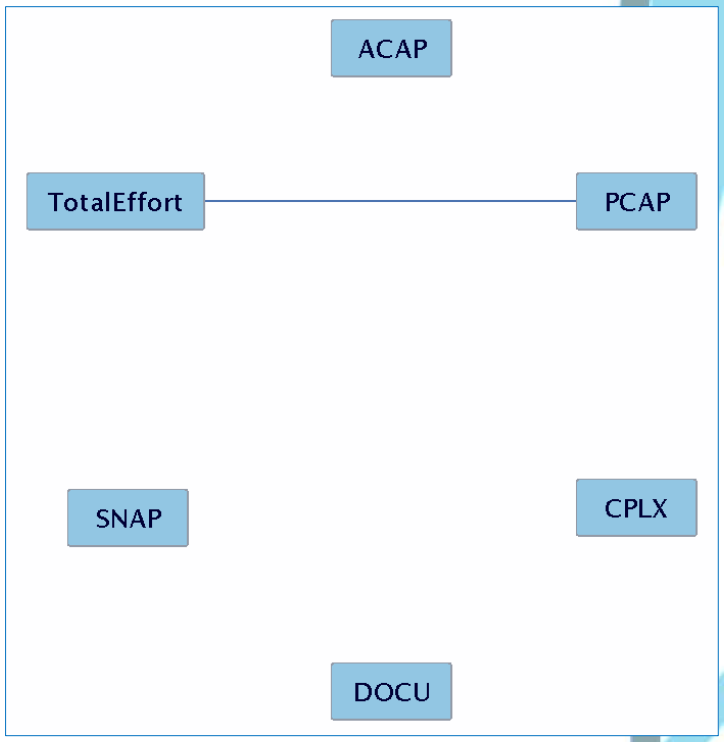
IFPUG FPs


1200
1000
800
600
400
200
0
Normalized Effort (hours)
0
5
10
15
20
25
30
35
40
Enhancement Function Points

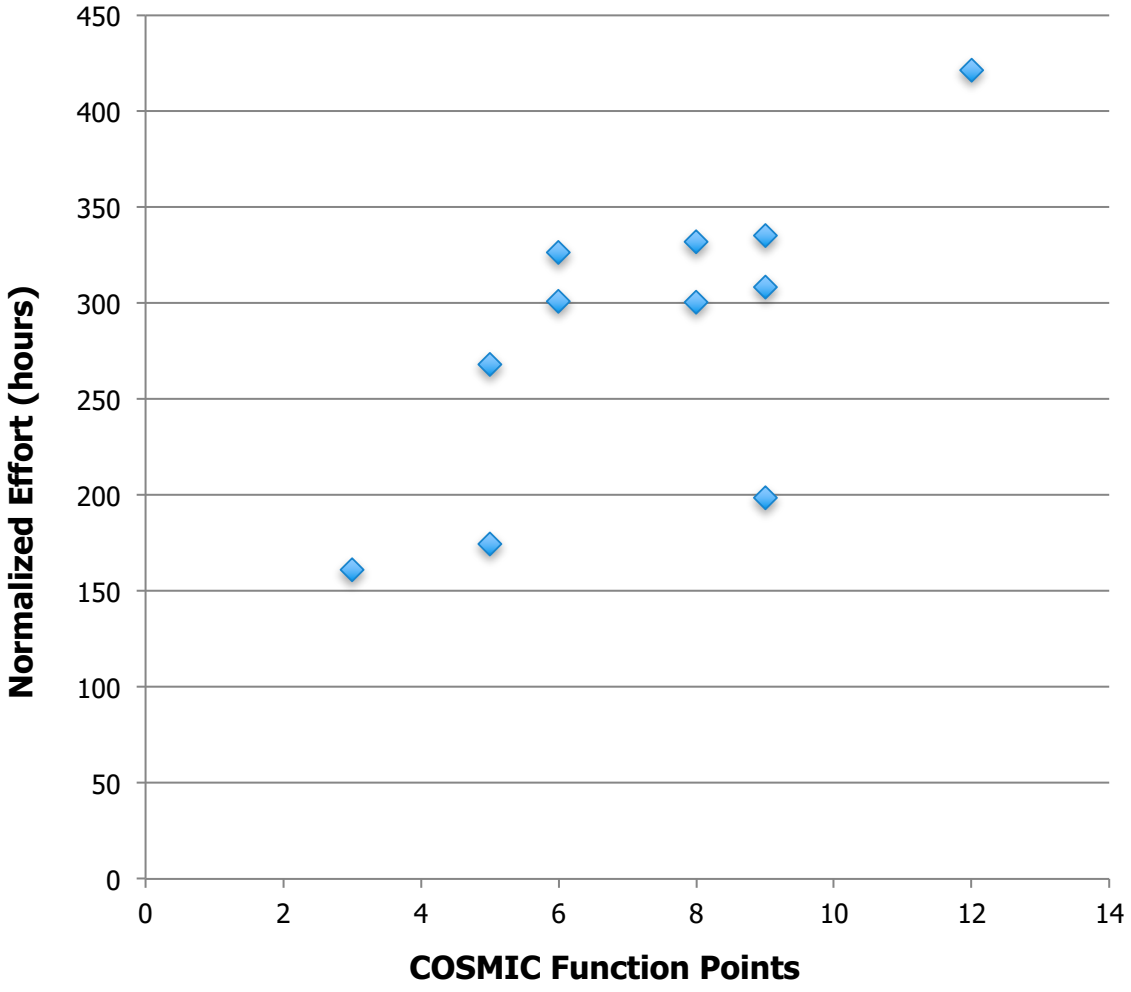

ACAP
TotalEffort
PCAP
FPs
CPLX
DOCU

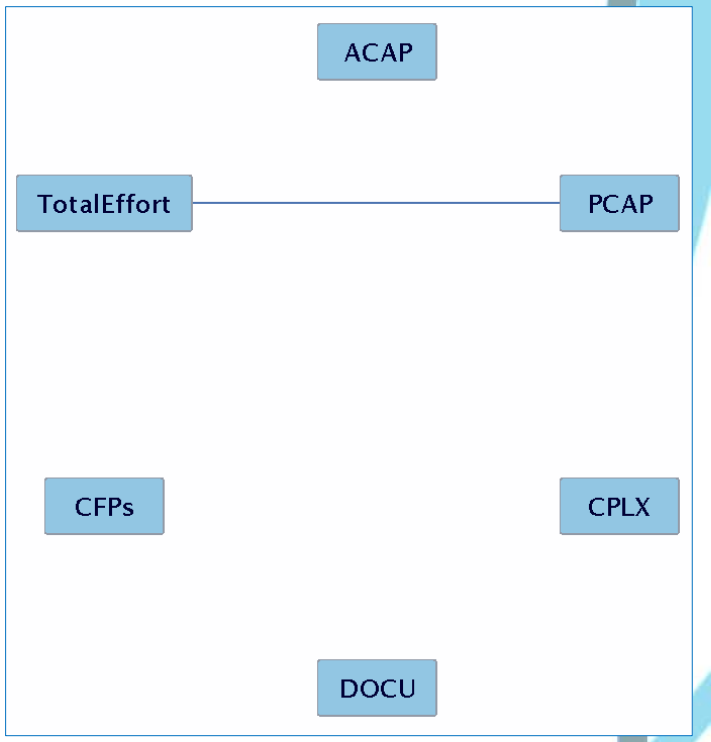
IFPUG SNAP


Normalized Effort (hours)
450
400
350
300
250
200
150
100
50
0
0
20
40
60
80
100
120
Enhancement SNAP Points


ACAP
TotalEffort
PCAP
SNAP
CPLX
DOCU

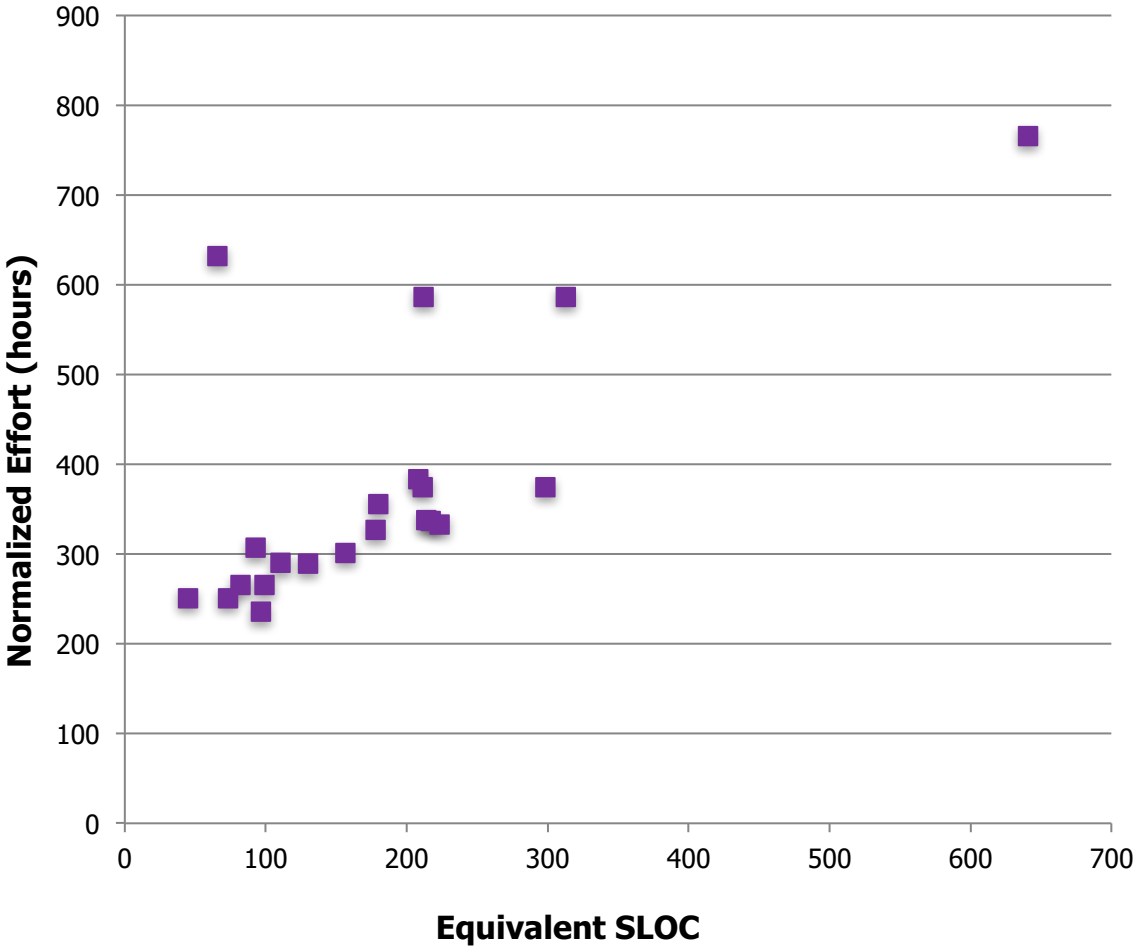
COSMIC FPs

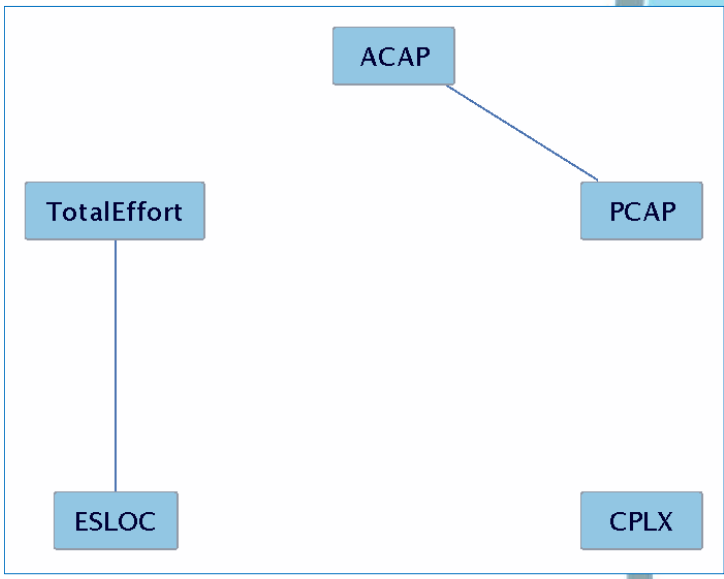

450
400
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
14
Normalized Effort (hours)
COSMIC Function Points


ACAP
TotalEffort
PCAP
CFPs
CPLX
DOCU

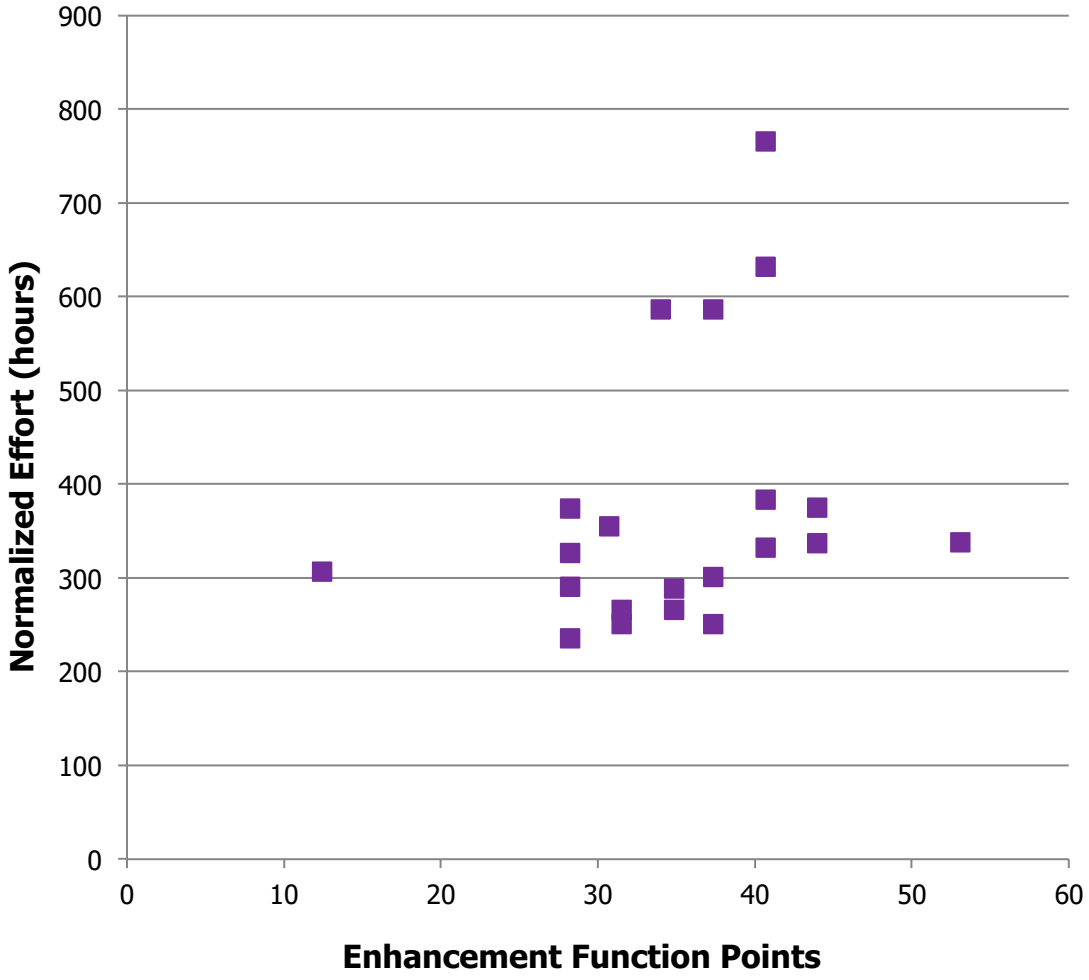
Modify Functions

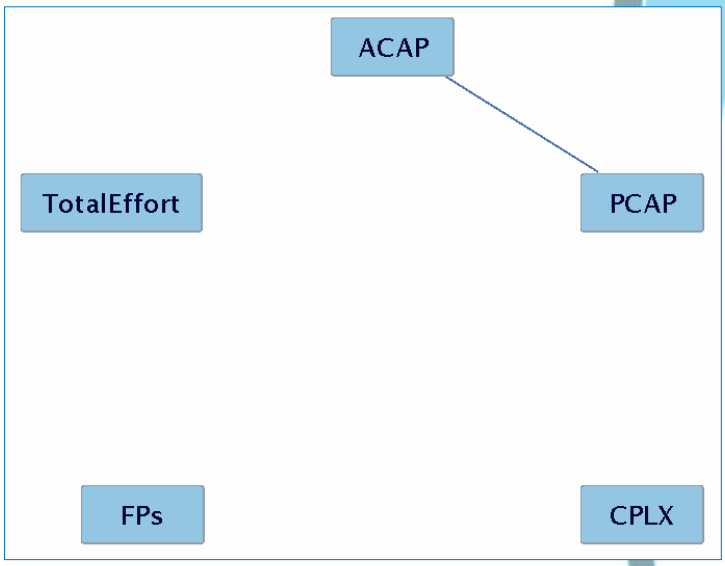
ESLOC


900
800
700
600
500
400
300
200
100
0
Normalized Effort (hours)
0
100
200
300
400
500
600
700
Equivalent SLOC

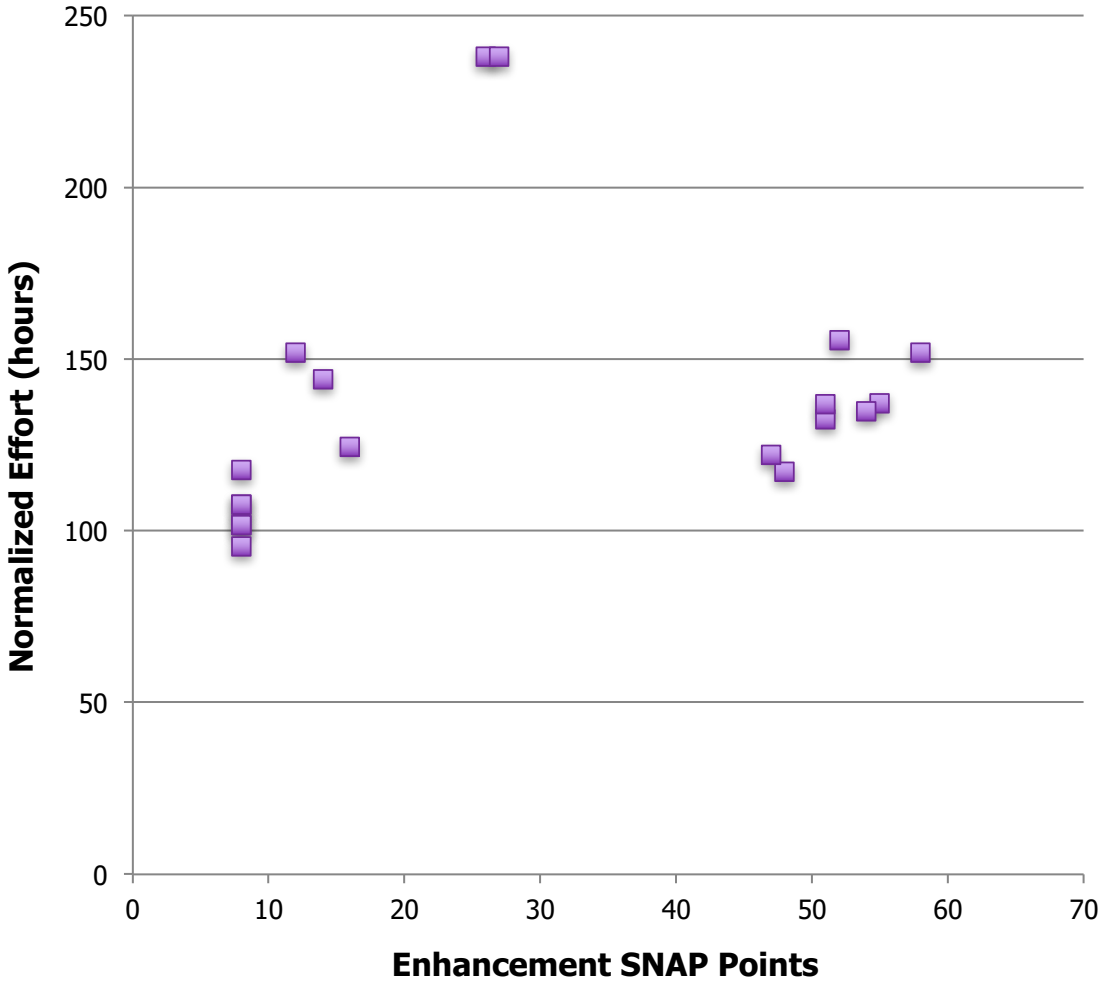

ACAP
TotalEffort
PCAP
ESLOC
CPLX

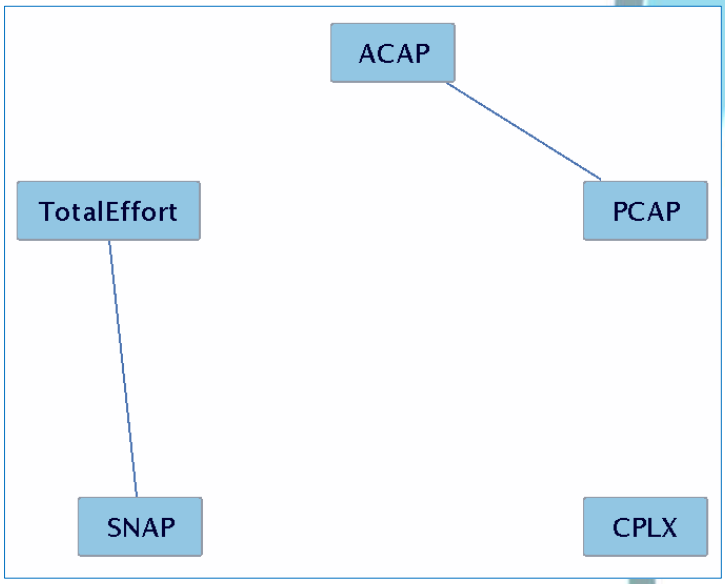
IFPUG FPs


900
800
700
600
500
400
300
200
100
0
0
10
20
30
40
50
60
Normalized Effort (hours)
Enhancement Function Points

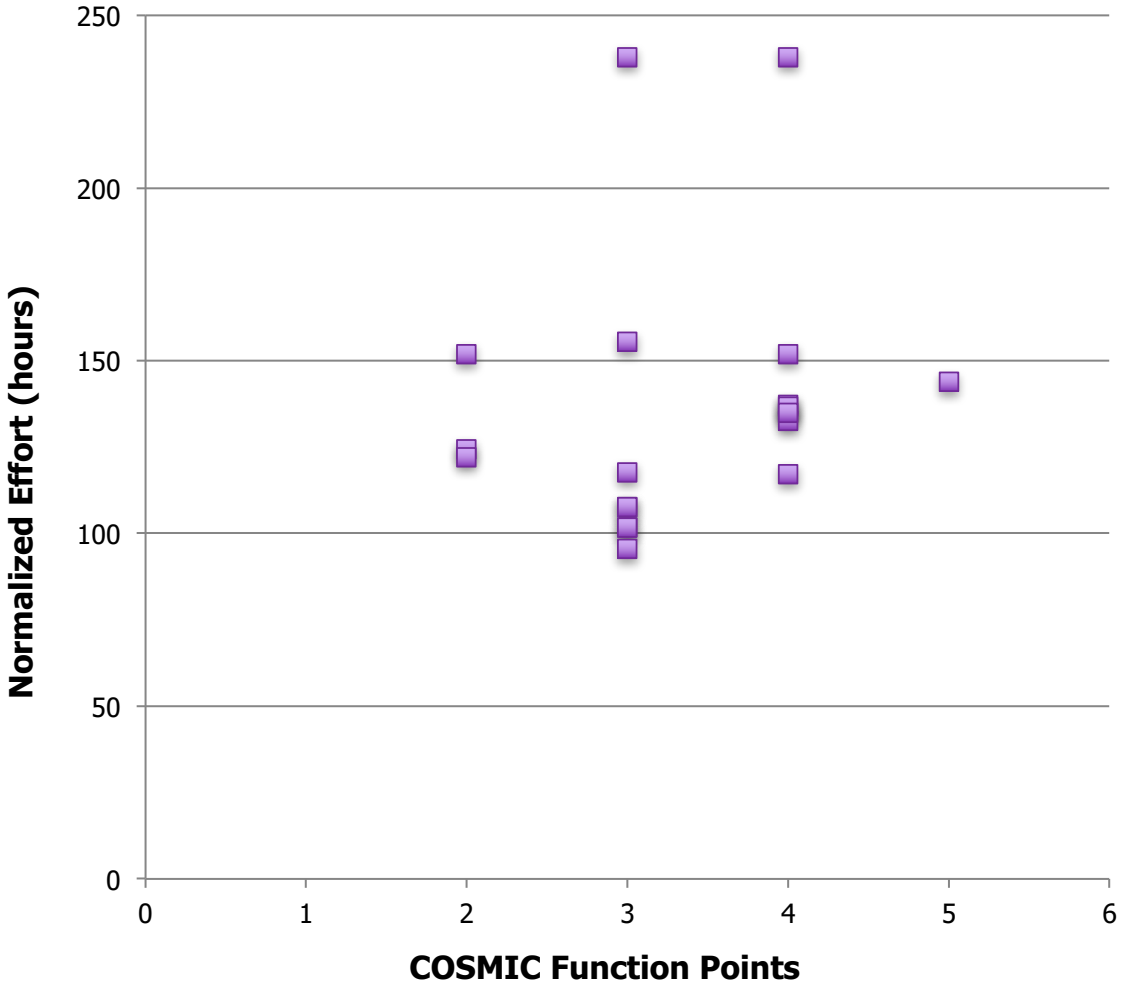

ACAP
TotalEffort
PCAP
FPs
CPLX

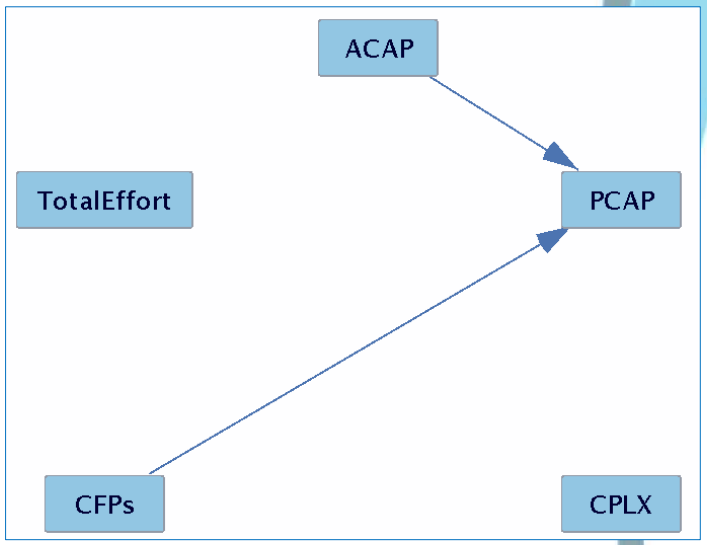
IFPUG SNAP


250
200
150
100
50
0
Normalized Effort (hours)
0
10
20
30
40
50
60
70
Enhancement SNAP Points


ACAP
TotalEffort
PCAP
SNAP
CPLX

COSMIC FPs


250
200
150
100
50
0
0
1
2
3
4
5
6
Normalized Effort (hours)
COSMIC Function Points


ACAP
TotalEffort
PCAP
CFPs
CPLX

Conclusion

Conclusions

General Conclusions

- All Data Points
  - SNAP -> Total Effort
  - CFPs -> Total Effort
  - PCAP – Total Effort
  - ACAP – PCAP
- Add Functions
  - PCAP – Total Effort
- Modify Functions
  - ESLOC – Total Effort
  - SNAP – Total Effort
  - ACAP – PCAP

"Interesting" Results

- All Data Points
  - CFPs -> DOCU
- Modify Functions
  - CFPs -> PCAP
  - ACAP -> PCAP