

Analytical Hierarchy Process (AHP) Approach to Size Estimation

Bruce E. Fad
PRICE Systems L.L.C.
17000 Commerce Parkway, Suite A
Mount Laurel, New Jersey, 08003, USA

Abstract

Thomas Saaty developed the Analytical Hierarchy Process (AHP) to be an effective means of dealing with complex decision-making. AHP helps capture both subjective and objective evaluation measures, providing a useful mechanism for checking their consistency relative to considered alternatives, thus reducing bias in decision making.

AHP concepts can be applied to problems of size estimating in support of cost modeling. Since many cost models use some measure of size as a cost driver and since accurate size measurement is often elusive, AHP type sizing can be invaluable in overcoming challenges of credible estimating.

This presentation will explain some of the basic concepts and uses of AHP. Application of AHP to size estimating will be illustrated through a software sizing methodology that approaches the sizing problem with four distinct measures:

- Pairwise comparison of the software product components
- PERT (Program Evaluation and Review Technique) bounding of the size of the components
- Sorting of components into defined size intervals
- Ranking of components by relative size

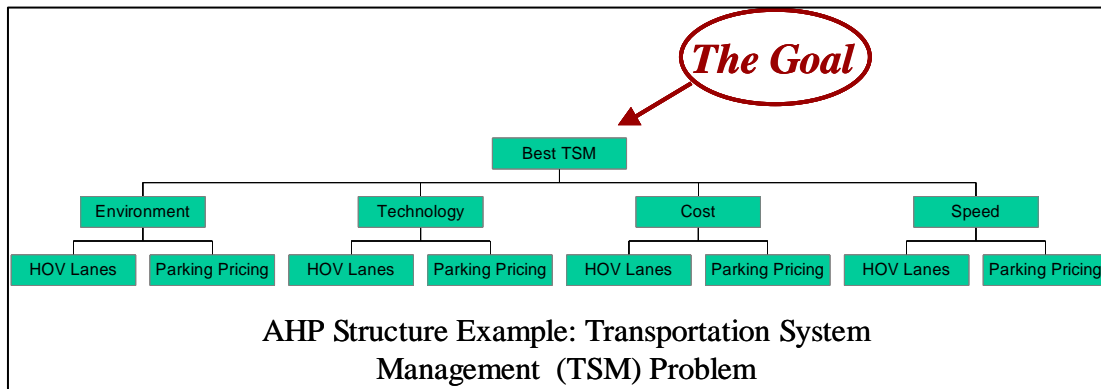
When subjected to independent assessment, these four measures correlate well to produce credible and unbiased estimates of software module and system size. A simple interactive exercise with workshop attendees will wrap up the session and help to establish the point that this AHP approach can actually be applied to a wide range of quantitative assessment problems – not just software size.

AHP: Analytical Hierarchy Process

Background. Analytical Hierarchy Process, or AHP, is a decision-making theory associated with Thomas Saaty of the University of Pennsylvania. Saaty traces his experience with AHP to contingency planning work he performed for the U.S. Department of Defense (DOD). Some of his earliest applications dealt with electricity rationing, transportation planning, and analysis of terrorism.

The precepts of AHP are reflected in observations of workings of the human mind. When confronted with a complex problem, humans tend to group elements of the problem by certain

properties that we believe we can compare. Through a series of analytical groupings at successively higher levels of problem abstraction, we eventually arrive at single element resulting from multiple level pairwise comparisons. That single element is the answer to our



problem - the goal of the decision making exercise. When an interrelated diagram of comparisons represents the process, we have a decision tree or hierarchy of comparisons that helps name the process (see figure above).

The Big Question.

Since factors of a decision are usually interrelated, it is necessary to establish a measuring scheme that allows each factor to influence the goal in proportion to its importance relative to all other factors. This poses the question for each comparison factor: *How strongly do the factors at the lowest level of the hierarchy influence the top factor (goal)?* In most cases, the answer to this is that each has a non-uniform influence, which necessitates use of an intensity measure – one that not only defines the most influential factors, but also yields relative measures of influence differentials.

AHP uses simple pairwise comparison of components of a decision to produce intensity measures. As a level of the decision hierarchy is addressed through a sequence of pairwise comparisons, an overall intensity emerges for each level.

Intuitive Sense for Numbers.

In the 1930s, Tobias Dantzig published observations about the most fundamental of mathematical concepts in his book, *Number, the Language of Science*. Among his theories is one in which Dantzig claims that the human mind has a sense for numbers that predates counting. He claims that from the earliest of humans, we have been able to recognize that a small collection of objects has increased or decreased when objects are added or removed, and that we understand this before we understand how to establish the size of a collection. This latent ability of the human mind is often referred to as qualitative assessment or fuzzy differentiation.

Pairwise Assessment Integration.

Fuzzy differentiation is the foundation for pairwise assessment of AHP. Almost anyone can compare two objects against some number related criterion, but to effectively compare many objects in a pairwise fashion against multiple criteria, it is important that priority and its measurement be understood.

Priority is reflected in the weight assigned to factors of a decision or to alternatives of a course of action. Both the priorities and the weights depend upon goals. Measurement of priorities is the relative difference between alternative or factor weights. The absolute value of

any weight is not critical even though the level/hierarchy/goal may be an effective single weight (measure).

Saaty proposes an eigenvalue approach as the best way for calibrating a numerical scale resulting from pairwise comparison. The eigenvector provides the priority ordering and the eigenvalue provides a measure of judgment consistency in pairwise comparisons.

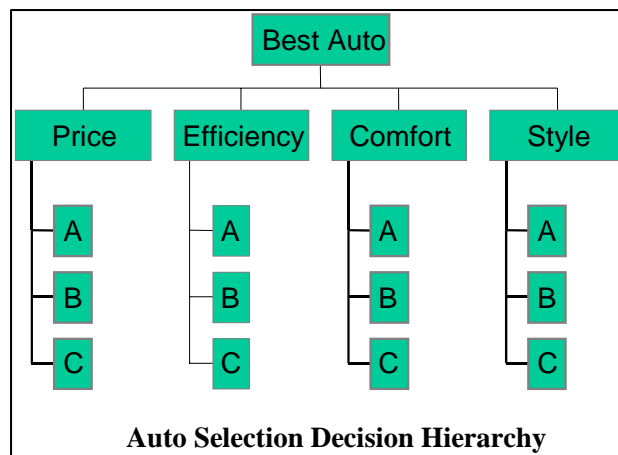
An AHP Example

Select the Best Auto. Suppose you are in the market for a new automobile and have narrowed your choice to three options. The table left summarizes the salient features of your options. You need a method for selecting the one auto that best meets those capabilities most important to you. This example will illustrate how AHP can be used to arrive at your choice.

	A	B	C
<i>Price</i>	45K	28K	17.5K
<i>Efficiency</i>	18 MPG	24 MPG	32 MPG
<i>Interior</i>	LE	SE	ST
<i>Body</i>	4-Door, Full-size	2-Door, Sport	2-Door, Compact
<i>Sound System</i>	6-CD Bose SS	AM/FM, 1-CD	AM/FM
<i>Engine</i>	VORTEC-6	Turbo-4	4-cylinder

Features of Three Auto Purchase Options

Decision Hierarchy. There are four important features to you in selecting you auto: Price, Fuel Efficiency, Comfort, and Style. Of these, price is two to three times as important to you as any other feature. Style is the next most important feature, followed by comfort. Fuel efficiency is the least important feature to you. It should be noted that the most important and the least important features are somewhat at odds with one another since fuel efficiency has a significant impact on automobile ownership cost. For purpose of this simple example, the price measure is the purchase price only of the auto. The hierarchy of decisions for this example is illustrated below, with each option evaluated under each feature and the best overall choice emerging at the top.



Priority Matrix. This decision requires creation of five priority matrices – four that represent the three options relative to each of the four important features and one that represents

	A	B	C
A	1	.33333	.25
B	3	1	.5
C	4	2	1
Price Feature Priority Matrix			

the relative importance of each of the four features relative to the goal (selecting the best auto). An example of one of the four options to feature priority matrices (for price) is shown at the left. The values of the matrix convey preference information about one option versus another with respect to the feature in question. In this case, the feature is price and the value 3 of the second row, first column indicates that auto option B is moderately preferred to auto option A as far as price is concerned. The higher the value, the more the option of a row is preferred to its optional pair

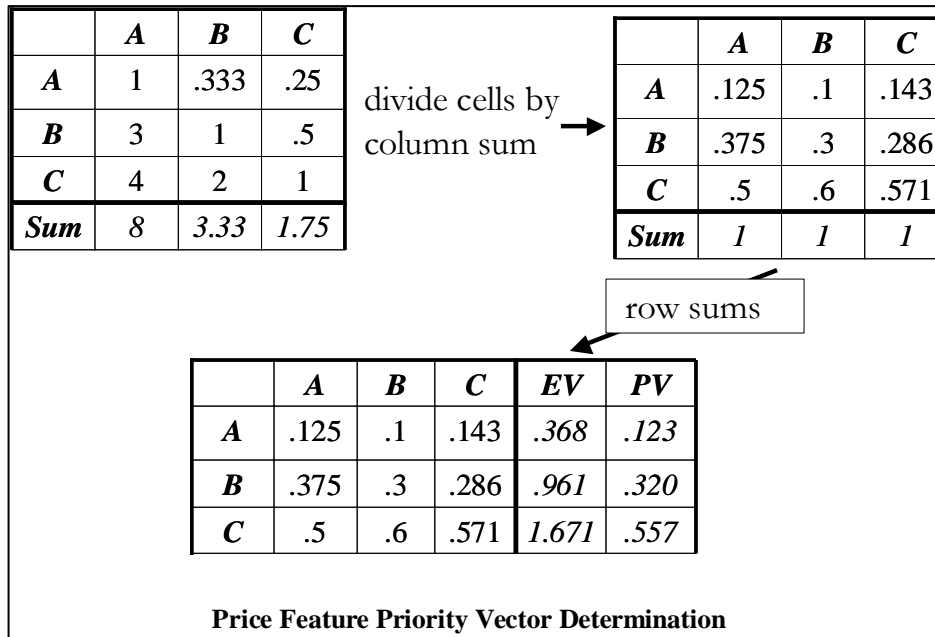
of the column. Diagonal values are always 1 for a priority matrix and matrix reflections are always reciprocals.

Comparison Scale. In the matrix above, auto B was described as being moderately preferred to auto A. Why moderately and not heavily preferred or just simply preferred? Because of the qualitative nature of preference expressions, a standard comparison scale is frequently

used with AHP. The scale, shown at the right serves as a fairly consistent mechanism for translating the language of comparison into the quantitative measures required for AHP. Even numbered rating scores are used to straddle the descriptions of preference described in this table.

Verbal Statement of Preference	# Rating
Extremely Preferred	9
Very Strongly Preferred	7
Strongly Preferred	5
Moderately Preferred	3
Equally Preferred	1
Pairwise Comparison Scale	

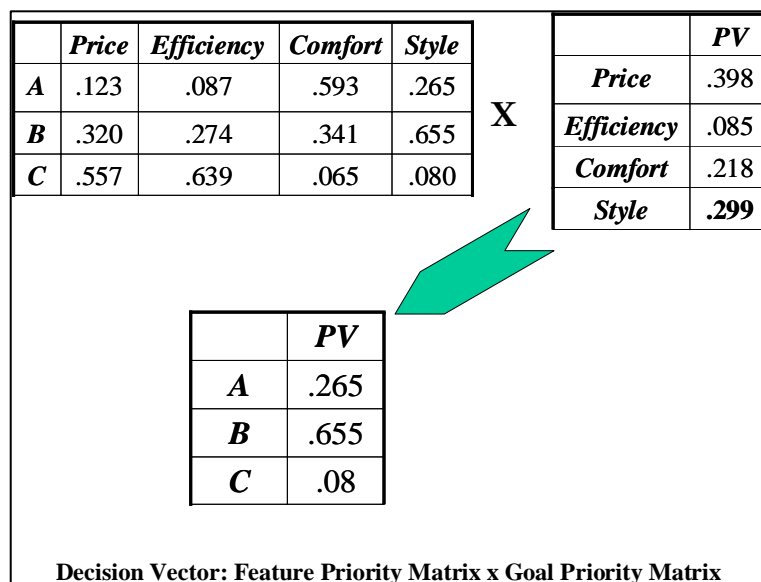
Priority Vectors. Simple matrix mathematics are used to compute a priority vector for each option to feature priority matrix. The process begins by computing the sum of each column. Each cell is then divided by its column sum and the rows added. The vector resulting from summing the rows is called the eigenvector (EV below). The priority vector (PV below) is the normalized eigenvector; normalized by dividing each eigenvector element by the elements sum. The process and calculations for determining the price priority vector are shown in the following diagram.



The same process applied to the other three important selection features yields the feature priority matrix below.

	<i>Price</i>	<i>Efficiency</i>	<i>Comfort</i>	<i>Style</i>
<i>A</i>	.123	.087	.593	.265
<i>B</i>	.320	.274	.341	.655
<i>C</i>	.557	.639	.065	.080
Overall Feature Priority Matrix				

Simple Computation. Following the same calculation process through pairwise comparison of each feature relative to the overall goal yields a goal priority vector. Matrix multiplication of the 3 x 4 feature priority and 4 x 1 goal priority matrices produces the decision vector that results in auto B as the best choice (see right).

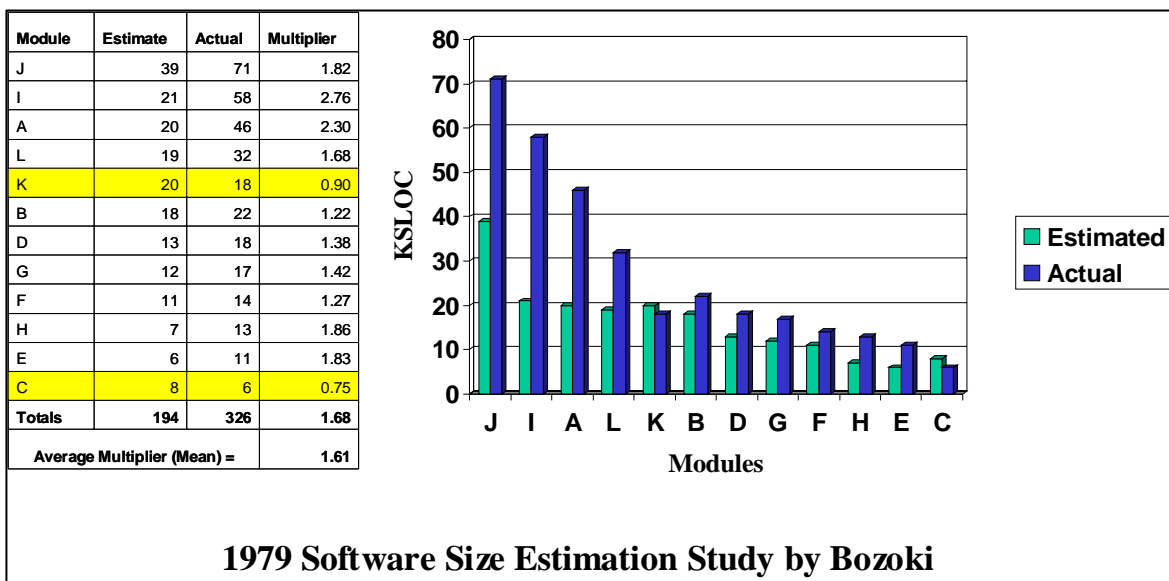


AHP to Size Software

SSM. Dr. George Bozoki has developed a software sizing model (SSM) based upon the principles of AHP. Though Bozoki's model uses geometric means as the mathematical basis for sizing, he applies much of the Saaty thinking about pairwise comparison of components and relative measures to arrive at size estimates. This approach can be viewed as AHP with software modules serving as the alternatives and size being the feature of measure for each alternative as well as size of the combined modules being the goal.

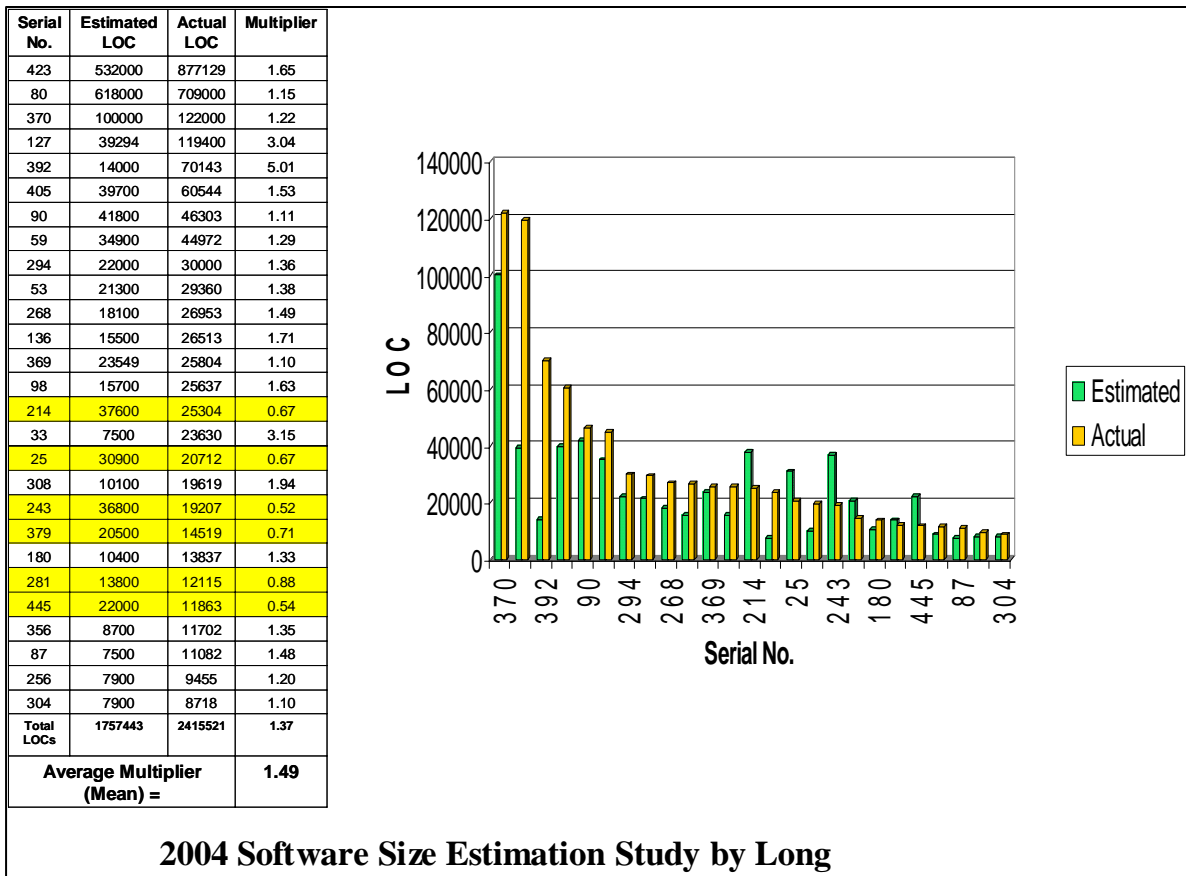
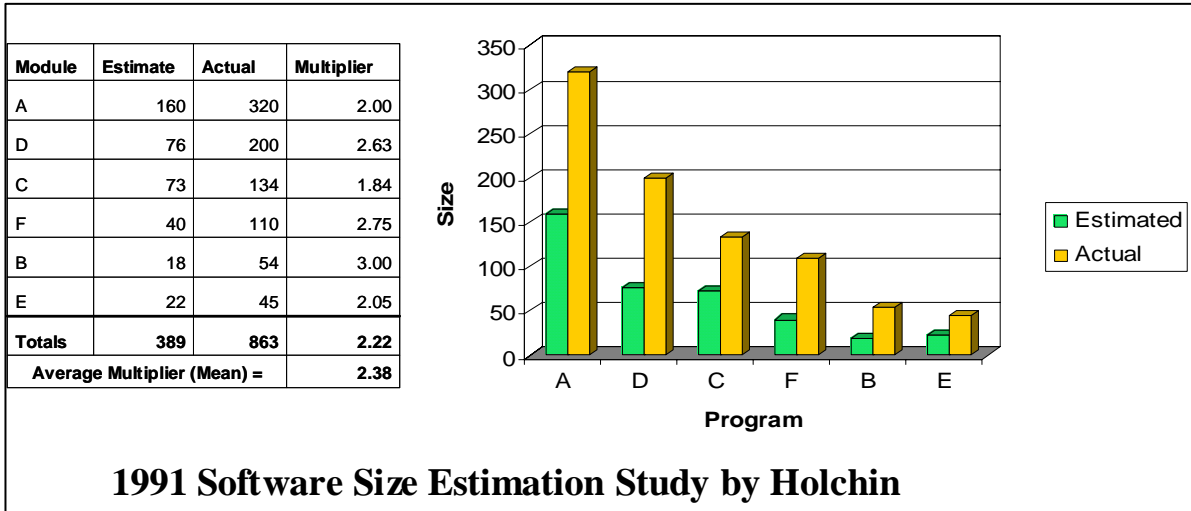
Bozoki's SSM is built around three premises: 1. In the early stages of a project, qualifying software size is a more accurate exercise than is quantifying size; 2. We estimate relative sizes of modules more reliably than we do absolute sizes; and 3. There is evidence of strong correlation between the estimated and actual relative size of software modules.

Correlation Evidence. Published studies of estimated versus actual size of software strongly indicate chronic underestimates throughout the industry. The Standish Group *Chaos* report is frequently cited for its shock value. However, there is at least one other phenomenon of size estimates versus actuals that goes almost unnoticed; that is the strong correlation between size estimates and actual size. So, at the same time one is likely to underestimate size, one is likely to accurately estimate the variation in size among modules of a software system. As an example, consider Bozoki's 1979 analysis of 12 software projects, as illustrated below.



10 of the 12 software modules of this study were underestimated. On average, a multiplier of 1.61 to the original estimate would have improved the accuracy of the results by more than 2x.

Later studies by Holchin in 1991 and Long in 2004 show similar results. Holchin's study is less optimistic with 100% of the data pointing to underestimates of over 100% on average. Long's study has a larger population and gives hope that software sizing is becoming more accurate. The data and graphs pertinent to each of these studies follows.



SSM Methodology. SSM employs four dimensions of relative size assessment together with a reference to estimate software size. Each dimension of assessment equates to a different way of viewing the relationship among software components in terms of size. The reference is a module of known size that can be related to each of the other modules. There is no requirement that the reference be part of the system being sized. In fact, it is preferable that it be a known entity from a different, but comparable system since the reference point is not included in the

system size estimated by SSM. An example of the SSM User Interface showing descriptive input and the resulting systems size estimates is shown below.

The screenshot displays the SSM User Interface with the following components:

- Module Size Summary Table:**

Module Name	-Std Dev	Expected Size	+Std Dev	Std. Dev.
Archive Data Processing	314	320	329	6
Instrument Processing	238	250	262	12
System Status and Schedule	161	170	179	9
- System Size Summary:**
 - Expected System Size: 740
 - Standard Deviation: 16
 - Sizes Expressed In: FUNCTION POINTS
 - Confidence Limits Table:

Confidence Limits	Probability (%)	Size (-)
50%		740
60%		740
70%		750
80%		750
90%		760
99%		790
- Reference Module List:**

Reference Module	Input Size	Implementation Size
Data Acquisition	25000	80
- Software Implementation Menu:**
 - Select a Language or Implementation type for the estimated Modules
 - 3rd Generation Languages
 - 4th Generation Languages
 - ABAP
 - Access
 - Ada
 - Ada 95
 - Ada 95 w/ components
 - Algol
 - APL
- Module Properties (Data Acquisition):**
 - Name: Data Acquisition
 - Size: 25000 (Reference Module limits: Minimum Size: 5 ; Maximum size: 999999)
 - Language: ASSEMBLY
 - Description: The data acquisition module receives the spacecraft Full Resolution data and Telemetry (FRD&T) at data rates up to 20 MBPS for periods up to 5 minutes. The module receives the Low Resolution Data and Telemetry (LRD&T) at a KBPS data rate. The module error checks all received data blocks, decommutates the data, and stores the resulting level 0 data on a working medium. The module separates the recorded Health and Instrument Telemetry (H&I TLM) from the FRD&T and provides low rate service for transmission.
- Module Information Table:**

Module Name	Language	Size
Data Acquisition	ASSEMBLY	25000
Archive Data Processing		
Instrument Processing		
System Status and Schedule		

SSM UI Example: Results at Top, Descriptive Input at Bottom

The assessments are best performed by those who are both knowledgeable about software and the specific system being assessed. Four different relative size assessments are required with SSM. Ideally, a different person is used for each of the four assessments and there should be no attempt to cross correlate. The data sets associated with each comparison are: Pairwise Data Sets that indirectly rank each component, PERT Sizing Data Sets that produce a ratio ranking of scale differences among the components, Sorting Data Sets that slot differences into ranking scales, and Ranking Data Sets that directly rank all modules. Examples of each follow.

Pairwise Data

Click on the larger Module

System Status and Schedule

The system status and schedule provides the following functions:

- allocation and scheduling of system resources in support of S/C operations, ground segment data processing facilities & equipment maintenance, and development efforts.
- outage reporting and system reconfig to maintain services in both automatic and manual modes of system control.

Instrument Processing

The instrument processing module retrieves data from the working medium and generates the calibrated data products.

During product generation, the module applies:

- instrument calibration corrections
- instrument transfer function corrections

Pair	1st Module (A)	2nd Module (B)	Larger
1	System Status and Schedule	Instrument Processing	B
2	System Status and Schedule	Data Acquisition (25000)	A
3	Data Acquisition (25000)	Archive Data Processing	B
4	Archive Data Processing	Instrument Processing	A
5	Archive Data Processing	System Status and Schedule	A
6	Instrument Processing	Data Acquisition (25000)	A

Ranking Data

Data Acquisition (25000)

Description: The data acquisition module receives the spacecraft Full Resolution data and Telemetry (FRD&T) at data rates up to 20 MBPS for periods up to 5 minutes. The module receives the Low Resolution Data and Telemetry (LRD&T) at a KBPS data rate. The module error checks all received data blocks, decommutates the data, and stores the resulting level 0 data on a working medium. The module separates the recorded Health and Instrument Telemetry (H&I TLM) from the FRD&T and provides low rate service for transmission.

Modules

Smallest Module

Data Acquisition [size=25000]

- System Status and Schedule
- Instrument Processing
- Archive Data Processing

Examples of Indirect (top) and Direct Ranking Assessments

PERT Sizing Data

Module Name: Archive Data Processing

Lowest: 90000, Most Likely: 110000, Highest: 130000

Module Name	Lowest	MostLikely	Highest
Archive Data Processing	90000	110000	130000
Instrument Processing	75000	95000	115000
System Status and Schedule	54000	64000	74000

50RT Sizing Data

Module Name: Archive Data Processing

Size Interval: 10900 -- 14800, 14801 -- 20000, 20001 -- 27000, 27001 -- 36600, 36601 -- 49500, 49501 -- 66900, 66901 -- 90600

Select this Interval

Name	Size Interval
Archive Data Processing	90601 - 122500
Instrument Processing	66901 - 90600
System Status and Schedule	49501 - 66900

Examples of Ratio (top) and Interval Scaling Assessments

SSM Exercise

Size the New Project. Garf Corp. develops training and Management Information Systems (MIS) software for industry and government. A software size estimate is needed so that a cost estimate can be prepared to support a bid/no-bid decision for development of warehouse operation and control software for a new Customs Department system. There are five software components needed for the new system: Storage Control (Store), Item Shipment (Ship) Control, Item Receipt (Receive) Control, Warehouse Operations (Ops.), and System Diagnostics (Diagnostics).

Within the past year, Garf delivered a software system to the USDA for controlling the operation of a surplus produce distribution warehouse. That software was developed as three modules: Receive, Store and Ship. The Store module for the USDA system is 65,400 SLOC of Ada.

Do the Assessments. After the five Customs modules and the USDA Store reference are described, the SSM assessment begins with pairwise comparison of each of the six modules. Store and Ship are assessed as being of similar size, with ship being a bit larger due to a few more operations being required for shipping than for storing. Receiving often involves more operations than shipping and storing, but Operations always involves the most. The diagnostics for this system are to be very basic; this is judged to be the smallest of the modules. The figure at the left contains the pairwise assessment.

Pair	1st Module (A)	2nd Module (B)	Larger
1	Diagnostics	USDA Store (65400)	B
2	Ship	USDA Store (65400)	B
3	Receive	Diagnostics	A
4	USDA Store (65400)	Store	A
5	Diagnostics	Ops	B
6	Ops	Ship	A
7	Ops	Receive	A
8	Receive	Ship	A
9	Diagnostics	Store	B
10	USDA Store (65400)	Ops	B
11	Ship	Diagnostics	A
12	Ops	Store	A
13	Store	Ship	B
14	Receive	Store	A
15	USDA Store (65400)	Receive	A

Pairwise Comparison of Customs & USDA Reference Modules

The next assessment is PERT sizing to establish relative size scale differences among the modules. The reference module is not subjected to this assessment since its size is known. There are a variety of approaches one can take in performing

this assessment. In this example, the analyst believes that the Customs Store module is one-half the size of the USDA Store reference and that a 25% variance around that guess is a fair bracketing for PERT sizing of this and every other module. The most likely value for each of the modules other than Store is established from a factor based upon module function and this analysts experience with these type systems (e.g. Diagnostics are half the size of Store). The overall PERT assessment is

Module Name	Lowest	Most Likely	Highest
Store	24525	32700	40875
Diagnostics	12262	16350	20438
Ship	30657	40875	51094
Ops	46214	61619	77024
Receive	38320	51093	63867

PERT Data Sizing Notes

Rules of Thumb:
 Diagnostics is 50% of Store
 Ship is 25% greater than Store
 Receive is 25% greater than Ship
 Operations is 2/3 of Ship + Receive

PERT Assessment of Customs Modules with Notes of Explanation

summarized at the bottom of the previous page. The next assessment is to sort the modules into intervals created from the information supplied so far. The result of the analyst's sorting is shown below left. The final assessment is to directly rank the modules (all 6) from smallest to largest. The analyst's ranking assessment is shown below right.

Sort Interval Assessment of Customs Modules

Name	Size Interval
Store	29301 - 39700
Diagnostics	11801 - 16000
Ship	29301 - 39700
Ops	53701 - 72600
Receive	39701 - 53700

Ranking of All Modules

Smallest Module

- Diagnostics
- Store
- Ship
- Receive
- USDA Store (size=65400)
- Ops

Largest Module

Roll up the Assessments. The quad-chart below shows the expected module size including plus and minus 1 standard deviation ranges and the expected system size estimates (at 50% probability) along with confidence limits for 60, 70, 80, 90, and 99% probabilities.

SSM Customs Sizing Estimate

Module Name	Std. Dev.	Expected Size	+Std. Dev.	-Std. Dev.
Store	29540	32400	35260	2960
Diagnostics	17020	20100	23180	3080
Ship	39190	40900	42610	1710
Ops	60360	63900	67440	3540
Receive	50120	51200	52280	1080

System Size Summary

Expected System Size: 208500

Standard Deviation: 5860

Sizes Expressed In: Ada

Confidence Limits	Probability (%)	Size (+)
	50%	208500
	60%	210000
	70%	211600
	80%	213400
	90%	216000
	99%	226700

Reference Module List

Reference Module	Input Size	Implementation Size
USDA Store	65400	65400

Software Implementation Menu

Select a Language or Implementation type for the estimated Modules

- FUNCTION POINTS
- 3rd Generation Languages
- 4th Generation Languages
- ABAP
- Access
- Ada
- Ada 95
- Ada 95 w/ components
- Algol
- APL
- Assembly
- Atlas
- Basic
- Basic Assembler
- C
- C#
- C++
- Clipper

The table below shows the estimated versus actual size of the Customs modules and the total system. Overall, this application with SSM resulted in accuracy within 2%.

Module	SSM Estimate	Actual	Variance
Store	32400	32150	0.8%
Diagnostics	20100	18850	6.6%
Ship	40900	39930	2.4%
Operations	63900	63320	0.9%
Receive	51200	50990	0.4%
Total	208500	205240	1.6%

Accuracy of SSM Assessment

Summary

Value of a Reference. The most widely used tools for estimating cost, effort, and time of a software development project all use size as a first order effect variable. Software engineers usually estimate software size. Historically, size estimates are fairly well correlated to the later known actual size, but the relationship is routinely one that shows the size estimated is too low. The byproduct has been project estimates that are also too low. So, while software size estimators have a good grasp of relative sizes of familiar packages, without a known point of reference, they are likely to significantly underestimate size. Tools like SSM, that use an AHP like approach to establish accurate relative size measures as well as incorporate a reference of known size can significantly improve the ability of the software size estimator to accurately predict size. This, in turn, leads to improved software project cost estimating accuracy, better planning and greater likelihood of project success.

References

- Reference List.** Air Force, 1987, *A Descriptive Evaluation of Software Sizing Models*, Prepared for HQ USAF/Air Force Cost Center (AFCCE), Washington, D.C., Prepared by IIT Research Institute, Lanham, MD, September 1987.
- Bozoki, G.J., 1982, *Software Sizing Methodology, A Study Prepared for NASA/GSFC*, S-98771-B.
- Bozoki, G.J., 1991, *Performance Simulation of SSM (Software Sizing Model)*, Proceedings, 13th Annual Conference, ISPA, 1991, Volume X, pp CM-14.
- Bozoki, G.J., 1992, *A Trade Study of PERT and SSM Software Sizing Models*, Proceedings, 14th Annual Conference, ISPA, 1992, Volume XI.
- Bozoki, G.J., 1993, *An Expert Judgment Based Software Sizing Model*, Journal of Parametrics, Volume XIII, No. 1, May 1993.
- Bozoki, G.J., 2004, *Treatment of Code Growth*, Proceedings, 26th International Conference, ISPA, 2004.
- Saaty, T.L., 1980, *The Analytical Hierarchy Process*, McGraw Hill, N.Y.
- Torgerson, W.S., 1958, *Theory and Methods of Scaling*, John Wiley and Sons, N.Y.
- Foster, S. T., Jr. and LaCava, G., *The Analytical Hierarchy Process: A Step-by-Step Approach*, Boise State University, Department of Operations Management.
- Boulter, J., *The Analytical Hierarchy Process*, University of Illinois at Urbana-Champaign, Department of Urban and Regional Planning

Atthirawong, W. and MacCarthy, B., 2004, *An Application of the Analytical Hierarchy Process to International Location Decision-Making*, University of Nottingham, Operations Management Group

Software Productivity Consortium, Inc., *Applied Software Measurement Course*, Version 01.02.04, March 1994, Herndon, VA, 22070.

Biography

Bruce E. Fad. Bruce E. Fad is Vice President, Product Division of PRICE Systems, LLC. Mr. Fad, along with Anthony A. DeMarco and Georges T. Teoglou, formed a Limited Liability Company and led the PRICE Systems buyout from Lockheed Martin in 1998. A recognized expert in international cost estimating, Mr. Fad has executive responsibility for all software product development, maintenance, enhancement, sales, training, and support worldwide. In his 24 years at PRICE Systems he has been focused at delivery of products and services to clients.

Mr. Fad has been a graduate teaching assistant in the Mathematics Department of the University of Delaware, a US Army Air Defense Artillery Officer, a Cryptanalytic Mathematician with NSA, a PRICE-guy with RCA, GE, Martin-Marietta, and Lockheed Martin, an Acquisition Management Manager for Lockheed Missiles & Space Company, and a college football player and Academic All-America baseball player at the University of Delaware.

Mr. Fad is an active member of the Society for Cost Estimating and Analysis (SCEA), the International Society of Parametric Analysts (ISPA) and was recognized by ISPA as Parametrician of the Year in 1995. He is a patent holder of a parametric costing technique, and has served as an expert witness and as an expert consultant in litigation involving software system valuation and project claims. You are welcome to contact Mr. Fad (especially if the subject is sports) at 856-608-7217 or by email at: bruce.fad@pricesystem.