The Use of Cluster Analysis Techniques in Space Flight Project Cost Risk Estimation

George Fox, Ph.D.
Donald H. Ebbeler, Ph.D.
Edward J. Jorgensen, Ph.D.
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109, USA

Cluster analysis is a data mining technique for identifying similar objects in a database of objects and their attributes. We classify and display, in a simple tree, a database of historical missions and a proposed mission with attributes such as spacecraft mass, development duration, operations period, etc. Missions are grouped by comparing the separation among groups of missions. The separation between two missions is defined by the Euclidean metric in the (log) space of attributes. This analysis is used to identify missions similar to the proposed mission. For the proposed mission, simple cost risk elicitation is used to identify any internal, systemic cost risk: potential cost changes that arise from non-external sources only. The historical mission(s) identified by cluster analysis as closest to the proposed mission are used to estimate the ratio of total cost risk to internal cost risk and this factor is applied to the internal cost risk simulation of the proposed mission. A case study demonstrates the utility of this approach in evaluating proposed space flight projects.