

doi:10.1631/FITEE.1500373

题目：改进的软件模块化二元相似度测量

目的：各种各样的二元相似度测量在聚类方法中被用来确定数据中的相似实体的同类组。这些相似度测量大多数仅基于特征的存在或缺失。二元相似度测量在软件模块化中亦能与不同的聚类方法一起用于提高软件系统的可理解性与可管理性。每种相似度测量都有其优势与不足，分别能使聚类结果优化或恶化。

创新点：本文强调了软件模块化中一些已有的著名的二元相似度测量的优势。此外，基于这些已有的相似度测量，新提出了几种改进的相似度测量。

方法：首先，介绍了一些软件模块化中已有的著名的二元相似度测量的优势。接着，提出了几种改进的新的相似度测量。结合具体例子，说明这些新方法整合了 JC、JNM 和 RR 这几种已有的二元相似度测量的优势。最后，通过实验比较新方法与已有方法的结果，验证所提新方法的有效性。

结论：实验结果表明相较于已有的相似度测量，本文所提出的新的二元相似度测量结果的可信度更高。这些新方法能减少任意决策的数量，增加聚类过程中聚类的数量。尽管这些新方法仅基于数据的二元特征向量表达，它们能被用来测试任何编程语言编写的软件系统。

关键词：二元相似度测量；二元特征；测量组合；软件模块化