

# XÁC ĐỊNH TẬP HUẤN LUYỆN PHÙ HỢP TRONG ƯỚC LƯỢNG DỰ ÁN PHẦN MỀM

*Nguyễn V. Vũ*

Khoa Công nghệ Thông tin,  
Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM  
nvu@fit.hcmus.edu.vn

## **Tóm tắt**

Mô hình ước lượng phần mềm thường được xây dựng dựa vào dữ liệu dự án lịch sử. Tuy nhiên, khi phương pháp phát triển phần mềm thay đổi theo thời gian, mô hình được xây dựng dựa vào dữ liệu quá khứ có thể không còn chính xác nữa khi ước lượng những dự án mới. Điều này dẫn đến vấn đề là dữ liệu quá khứ nào phù hợp để huấn luyện mô hình ước lượng. Chúng tôi nghiên cứu và áp dụng cửa sổ di chuyển (moving windows) để xác định bộ dữ liệu phù hợp cho hiệu chỉnh mô hình COCOMO. Chúng tôi đánh giá độ chính xác của các mô hình ước lượng được hiệu chỉnh dựa trên cửa sổ di chuyển và dựa trên toàn bộ dữ liệu lịch sử. Kết quả thu được cho thấy việc sử dụng số ít những dự án hoàn thành gần nhất (nằm trong cửa sổ nhỏ) cho kết quả tốt hơn việc sử dụng toàn bộ dữ liệu lịch sử. Khi cửa sổ lớn, độ chính xác giảm dần. Nghiên cứu này đưa ra bằng chứng khoa học cho việc sử dụng số ít các dự án gần nhất thay vì sử dụng nhiều dự án quá khứ. Hơn nữa, khi phương pháp phát triển phần mềm thay đổi nhanh chóng, việc sử dụng cửa sổ nhỏ mang lại lợi ích lớn hơn là khi phương pháp phát triển phần mềm thay đổi ít.

Từ khóa: ước lượng phần mềm, COCOMO, cửa sổ di chuyển.

# DETERMINING RELEVANT TRAINING DATA FOR EFFORT ESTIMATION OF SOFTWARE PROJECTS

*Vu Nguyen*

Faculty of Information Technology, University of Science, VNU-HCM  
nvu@fit.hcmus.edu.vn

## **Abstract**

A software estimation model is often built using historical project data. As software development practices change over time, however, a model based on past data may not make accurate predictions for a new project. We investigate the use of moving windows to determine relevant training data for COCOMO calibration. We present a windowing calibration approach to calibrating COCOMO and assess the performance of effort estimation models calibrated using windows and all data. Our results show that calibrating COCOMO using small windows of the most recently completed projects generates superior estimates than using all available historical projects. Large windows tend to produce worse estimates. This study provides empirical evidence to support the use of small windows of projects completed so far to calibrate models when COCOMO-like data is available. Additionally, when the change in software development over time is rapid, the use of windows is more justifiable for improving estimation accuracy.

Key words: software estimation, COCOMO, moving windows