

# Second International Workshop on Detection of Software Clones

Arun Lakhotia and Andrew Walenstein  
Center for Advanced Computer Science  
University of Louisiana at Lafayette

walenste@ieee.org, arun@louisiana.edu

Rainer Koschke  
Institute for Software Technology  
University of Stuttgart

koschke@informatik.uni-stuttgart.de

## Abstract

*The aim of this half-day workshop is to bring together researchers within the field of clone detection to critically assess the current state of research, and to establish new directions and partnerships for research. Position papers and an index card-based brainstorming technique will be employed to organize the discussion and help form a foundation for a workshop report.*

## 1. Introduction

A software clone is generally defined to be a copy or near-copy of a portion of code appearing elsewhere in a system. The problem of detecting clones in systems is an established software engineering problem known to occur in many contexts, including during pattern detection, software refactoring and perfective maintenance, system quality evaluation, and class library reengineering.

This workshop expands upon the *First International Workshop on Detection of Software Clones*<sup>1</sup>, held in conjunction with ICSM'2002 and SCAM'2002 in Montreal in October of 2002. The goals of this workshop are to (1) bring together researchers within the field, (2) clarify and assess the current state of research, (3) establish a list of new directions and open and critical research questions, and (4) generate new research collaboration partnerships.

## 2. Scope, topics

The workshop scope is the general field of clone detection techniques, theories, and applications. Relevant topics include but are not limited to:

- software similarity models
- taxonomies of clone, redundancy, or duplication types
- clone detection techniques and methods

<sup>1</sup><http://www.bauhaus-stuttgart.de/clones/workshop.html>

- tool and technique evaluation or comparison
- empirical studies of clones in systems or of clone detection
- benchmarking and benchmarking issues, including data format issues, benchmark types, subject system selection, etc.

## 3. Session format

The half day workshop will be split into three sections:

- |             |                                 |
|-------------|---------------------------------|
| 45-90 min.  | Position papers and discussions |
| 90-120 min. | Brainstorming session           |
| 30-45 min.  | Summary and wrap-up session     |

The presentation and discussion of the position papers will be in WCRE style, i.e., with short, strict time limits and ample discussion time.

The goal of the brainstorming session is to bring focus to the discussions in order to produce the desired outcomes: consensus on the current state of research and a list of research directions. The brainstorming technique asks participants to write down onto index cards their research questions and evaluations of the state-of-research. The cards are then collected and collaboratively clustered, revealing points of consensus, disagreements, and directions for future research.

The summarization and wrap-up session is intended to raise outstanding questions and collect together any new insights that emerge from the prior two sessions. Participants will be individually asked to state their most pressing research concern and what they will be interested in working on in the near future. The aim is to provide the workshop participants with a summary overview of the important emergent issues in the field, and to identify potential future research collaborators.

The results of the brainstorming session are expected to help generate a report that is of broad interest. The identified topic clusters and the summary discussion will form the basis for a workshop report.