

FCS-ScriptRR-ESM

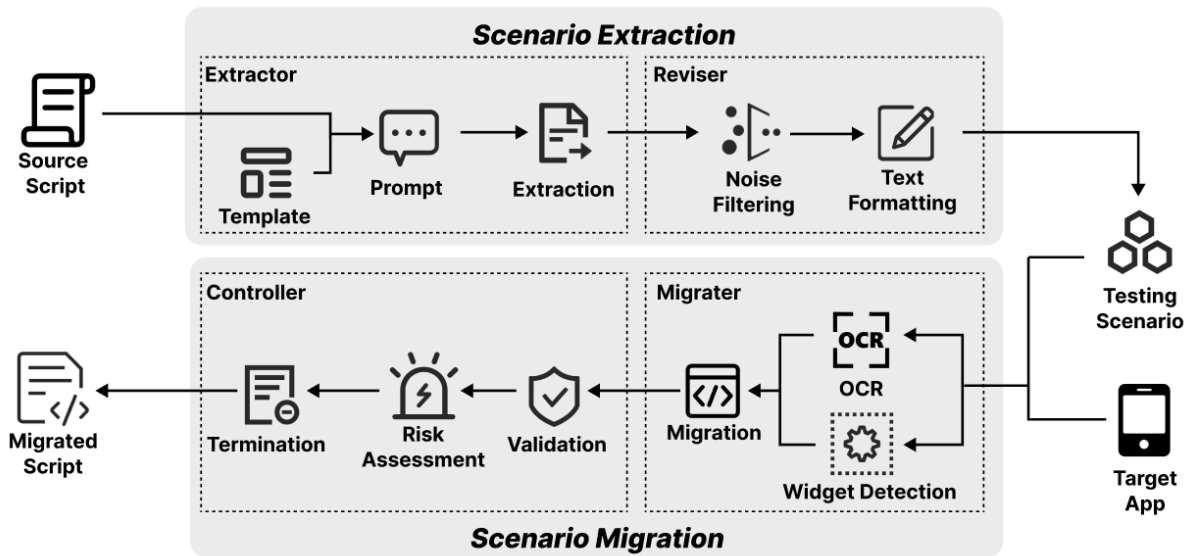
GUI Test Migration via LLM with Scenario-Granularity Understanding

Overview

Graphical User Interface (GUI) is a crucial component of mobile applications. The effectiveness and reusability of GUI testing directly impact the application quality and the testing cost. Given that apps within the same category often share similar functions and GUIs, researchers aim to reuse existing scripts through cross-app GUI test migration, thereby reducing cost and manual effort. Existing studies predominantly adopt operation-matching and widget-matching methods for cross-app migration. However, these matching methods often fail when source and target apps implement similar functions or GUIs using divergent operations or widgets. Moreover, these methods tend to focus on local matching results, lacking a comprehensive understanding of the general testing intention and the GUI states of apps. As a consequence, they struggle to explore modern mobile apps and tackle challenging migration tasks.

To address the above challenges and the limitations of existing tools, this paper proposes a novel GUI test migration via LLM with scenario-granularity understanding — **ScriptRR**. In this work, the testing scenario is defined as complete user interactions containing generalized business logic across apps, including critical contextual information such as the general testing intention and the expected state of apps. **ScriptRR** leverages the computer vision technology and vision language models to comprehensively extract the testing scenarios from scripts and GUIs, and organizes scenario-driven migration process based on the multi-agent mechanism. It addresses the lack of contextual understanding inherent in existing matching methods, enabling effective and efficient cross-app migration. The results show that **ScriptRR** outperforms the representative baselines, demonstrating its superior effectiveness and efficiency in GUI test migration.

Framework



Experiment

Dataset

Category	Applications
browser	chrome, Firefox, Foss Browser, Microsoft Edge: AI Browser
communication	TikTok, REDnote
mail	Gmail, Mail.ru, Outlook, Yahoo Mail
music	spotify, youtube music, KKBox, Music Worx
news	Sina News, Google News, CNN, Netease News
shopping	Amazon, Temu, eBay, AliExpress
to-do	todoist, To-Do List, google task, Any.do, Microsoft To Do

We construct a [dataset](#) including 18 typical testing scenarios, 27 popular mobile apps from 7 categories.

Reproduction Package

We provide code for reproduction [here](#).

Result Presentation

We provide detailed experimental results [here](#).

RQ1: How effective is ScriptRR in cross-app test migration compared to existing approaches?

Category	SR			GSR			OSR		
	Craft	S-GPT	S-Qwen	Craft	S-GPT	S-Qwen	Craft	S-GPT	S-Qwen
browser	0.167	0.792	0.771	0.125	0.667	0.708	0.125	0.646	0.708
comm.	0.125	0.750	0.750	0.000	0.375	1.000	0.000	0.375	0.750
mail	0.188	0.479	0.521	0.188	0.479	0.500	0.229	0.479	0.500
music	0.167	0.521	0.563	0.063	0.542	0.542	0.063	0.458	0.542
news	0.188	0.771	0.750	0.188	0.833	0.729	0.188	0.750	0.708
shopping	0.000	0.531	0.500	0.000	0.563	0.500	0.000	0.531	0.500
to-do	0.140	0.500	0.560	0.140	0.520	0.480	0.140	0.420	0.440
Total	0.148	0.610	0.609	0.120	0.599	0.578	0.127	0.546	0.561

RQ2: How does ScriptRR's critical module affect the test migration?

Category	SR			GSR			OSR		
	DLLM	S-GPT	S-Qwen	DLLM	S-GPT	S-Qwen	DLLM	S-GPT	S-Qwen
to-do	0.400	0.500	0.560	0.460	0.520	0.480	0.400	0.420	0.440
browser	0.104	0.792	0.771	0.083	0.667	0.708	0.042	0.646	0.708
shopping	0.375	0.531	0.500	0.375	0.563	0.500	0.375	0.531	0.500
mail	0.500	0.479	0.521	0.500	0.479	0.500	0.500	0.479	0.500
news	0.104	0.771	0.750	0.188	0.833	0.729	0.083	0.750	0.708
music	0.208	0.521	0.563	0.104	0.542	0.542	0.021	0.458	0.542
comm.	0.375	0.750	0.750	0.125	0.375	1.000	0.125	0.375	0.750
Total	0.287	0.610	0.609	0.277	0.599	0.578	0.227	0.546	0.561

RQ3: How does ScriptRR compare to existing approaches in terms of efficiency?

Approach	Time	Cost
Craftdroid	206s	/
ScriptRR-GPT	246s	0.69\$
ScriptRR-Qwen	214s	0.10\$