A Novel Methodology for Highly Accurate Hardware Trojan Detection

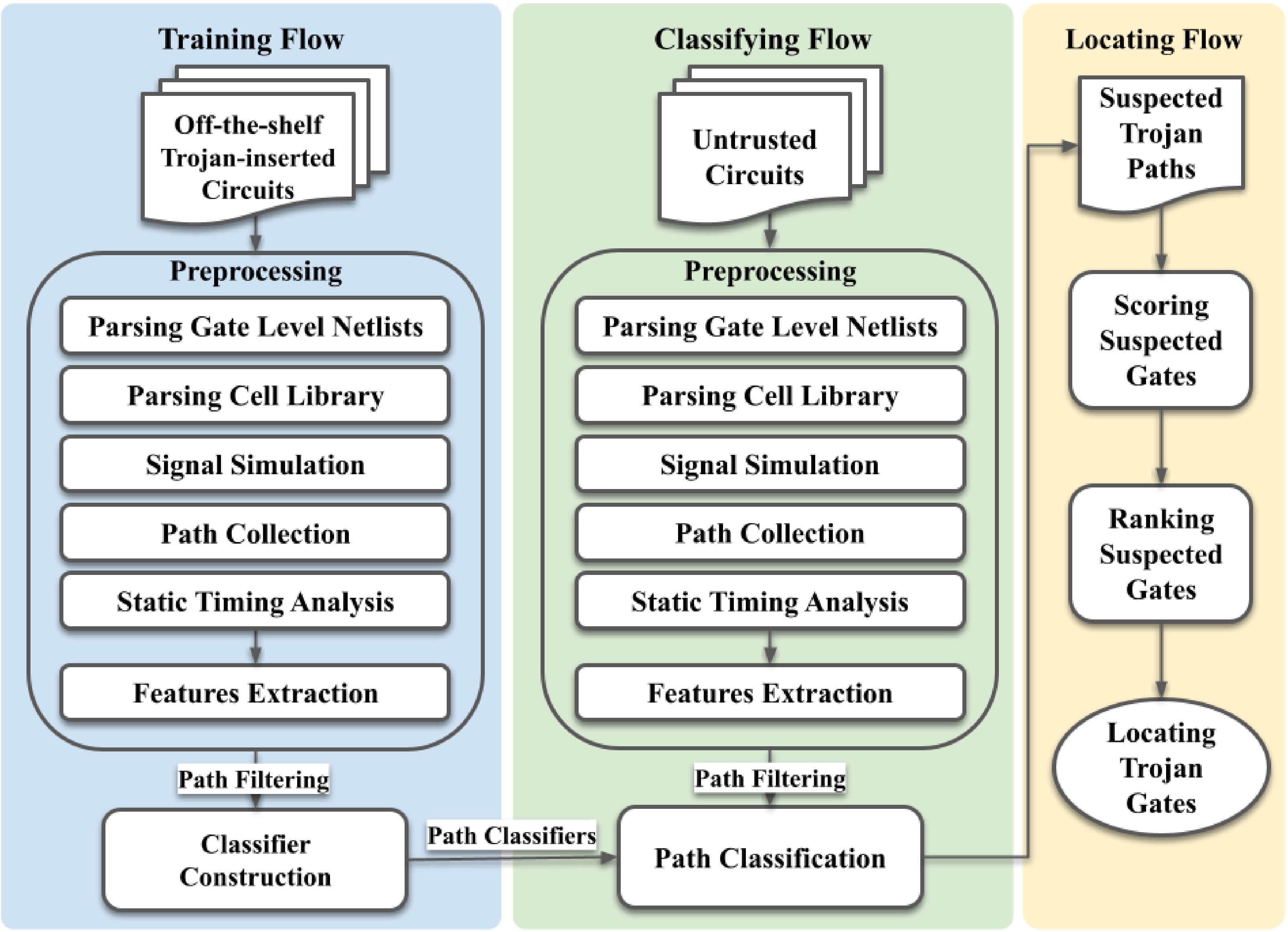
Student: Jung-Che Tsai

Advisor: Kai-Chiang Wu

Graduate Degree Program of Cyber Security, National Yang Ming Chiao Tung University

Introduction

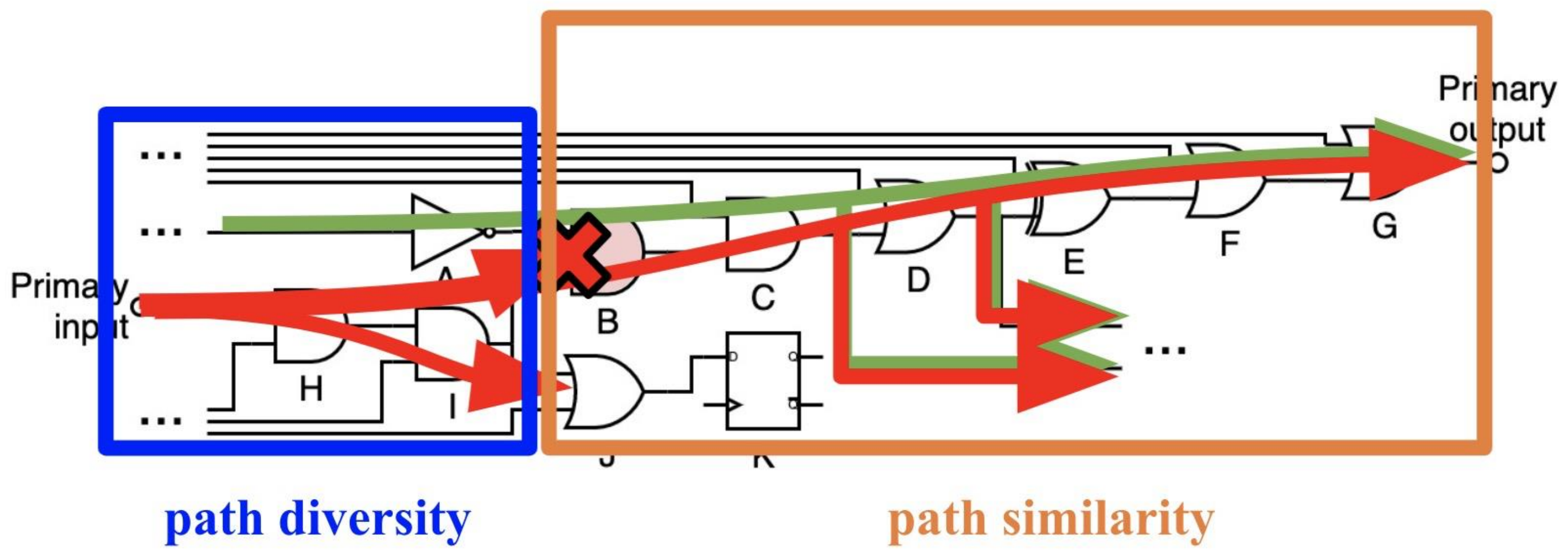
With the popularization of advanced embedded systems such as self-driving cars, the internet of things, intelligent mobile devices, and so on over the past decades, designing and manufacturing IC/SoC has become more sophisticated. The reduction of cost and the specialization in each production stage have brought up the trend of outsourcing. However, outsourcing also requires the involvement of untrusted third parties, which leads to the unreliability of IC-related products. As the IC/SoC applications spread widely, hardware security has turned into a non-negligible issue.



Methodology

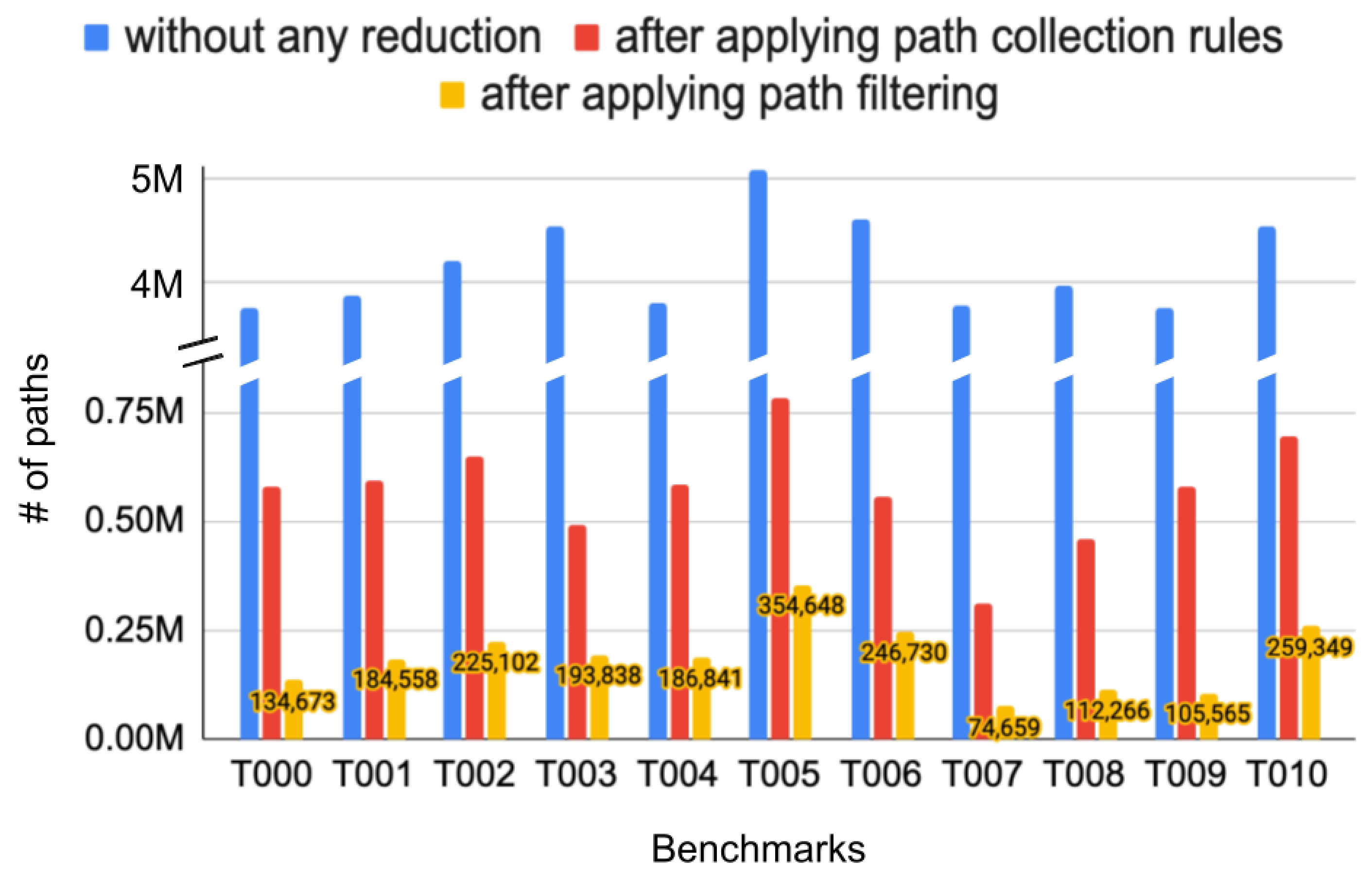
This work consists of three parts: the training flow, the classifying flow, and the locating flow. The preprocessing steps are conducted in both training flow and classifying flow. Paths in the gate-level netlist of circuit designs are collected, and path features are extracted after preprocessing. Next, in the training flow, the path classifiers are constructed with machine learning algorithms using the extracted path features. While in the classifying flow, paths collected from the untrusted circuits are classified into suspected Trojan paths and Trojan-free paths with the trained classifiers. Finally, gates on the suspected Trojan paths are located according to the results of path classification in the locating flow.

Path Collection Rules

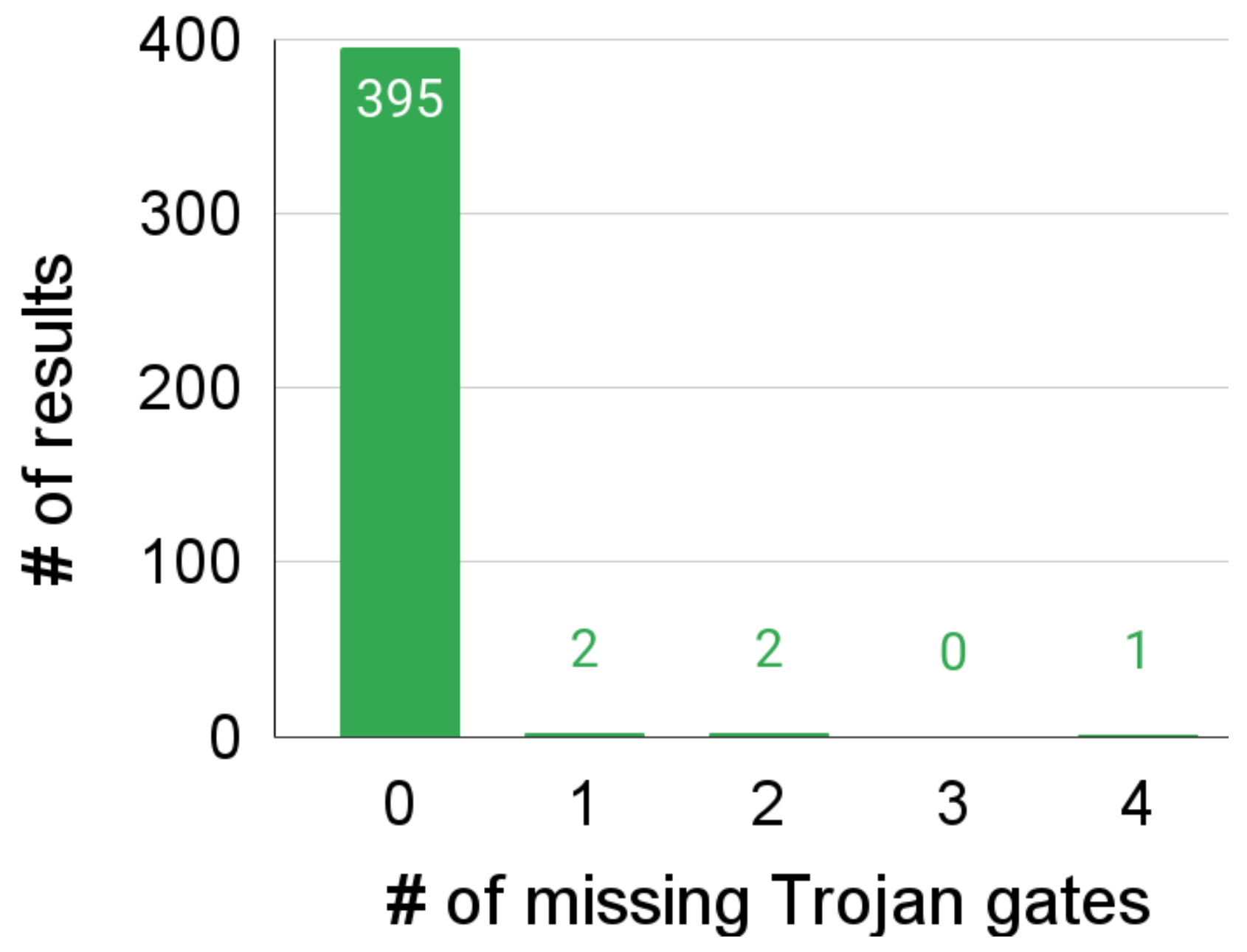


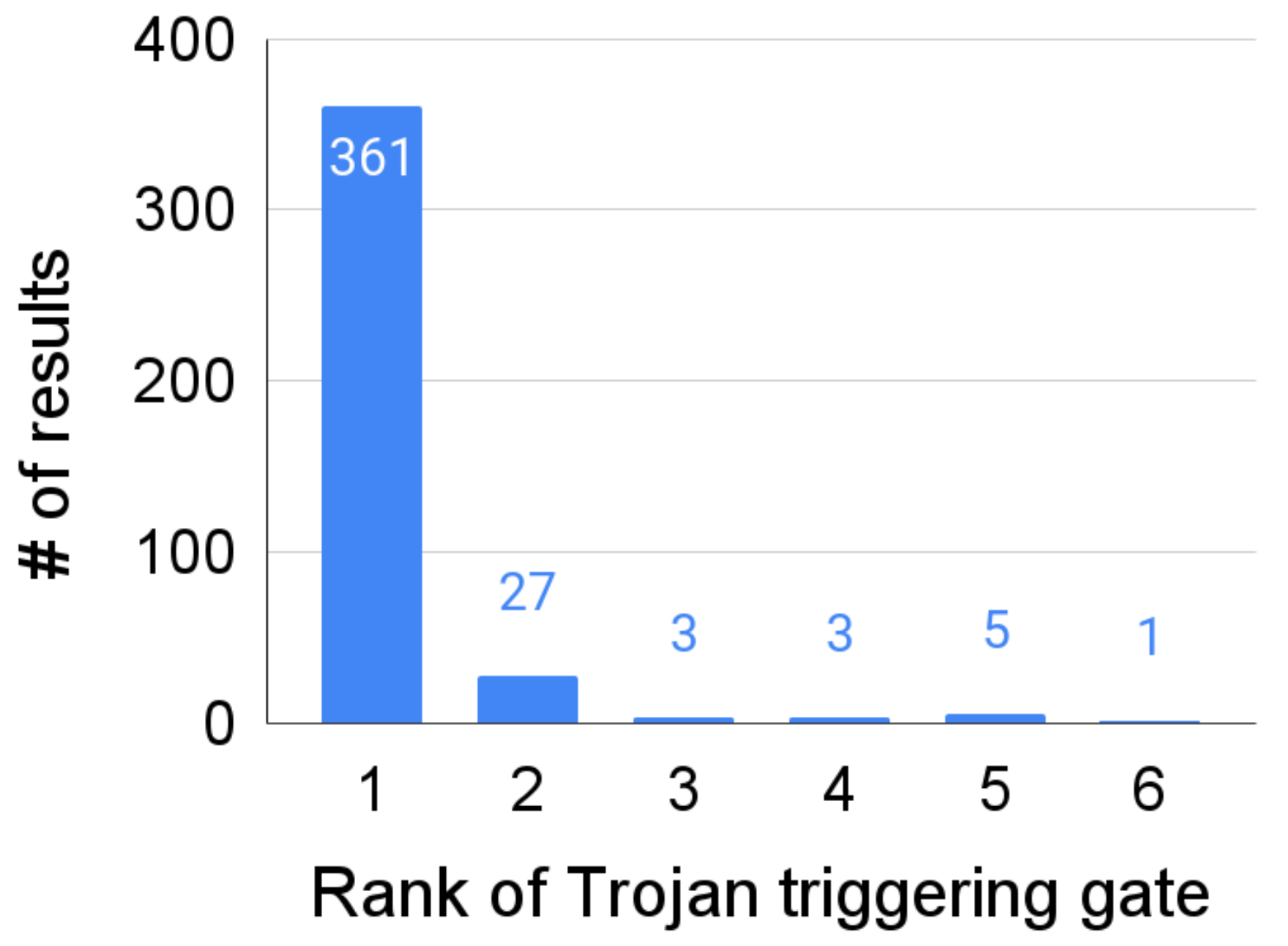
Results

The number of paths in a circuit design may be explosive. We propose path collection rules based on the path diversity and path similarity, and the path filter, to reduce the number of paths. Over 80% of the original paths are stripped with the path collection rules. While with the path filter, 60% more paths are filtered out.

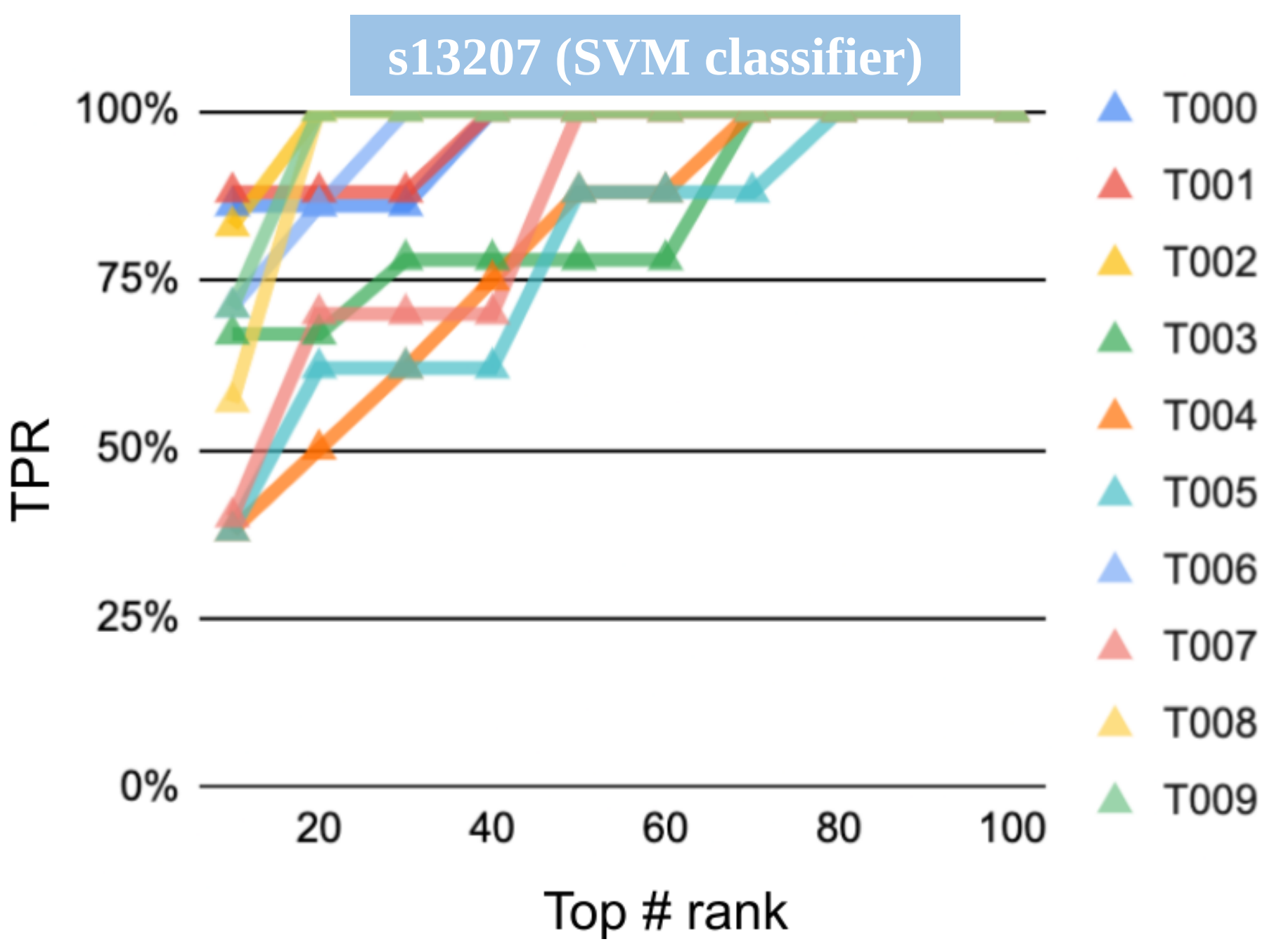


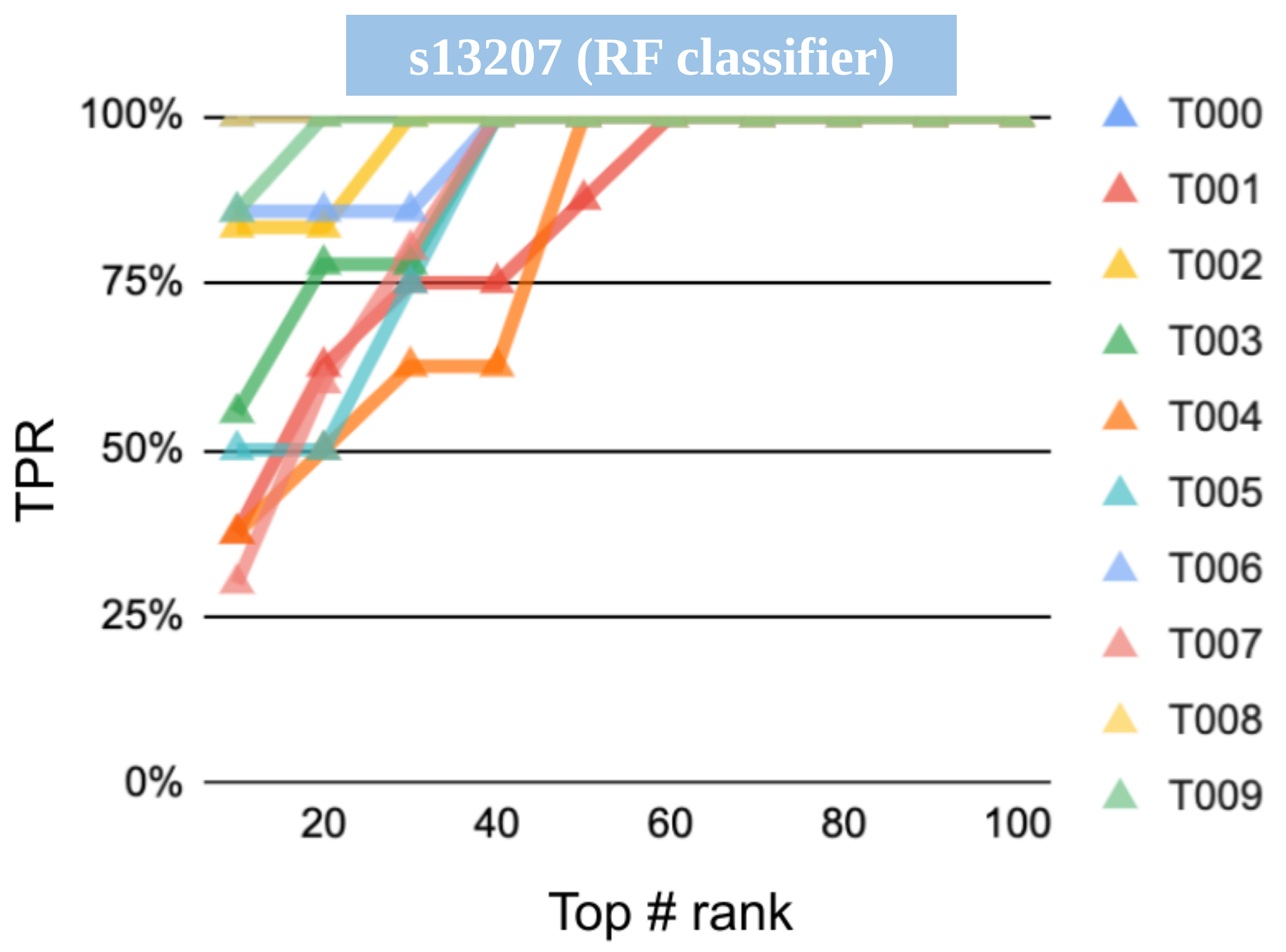
For the results of locating Trojan gates, there's no missing Trojan gates in most of the results. Furthermore, with the locating method we proposed, we are able to locate Trojan triggering gates at top ranks.

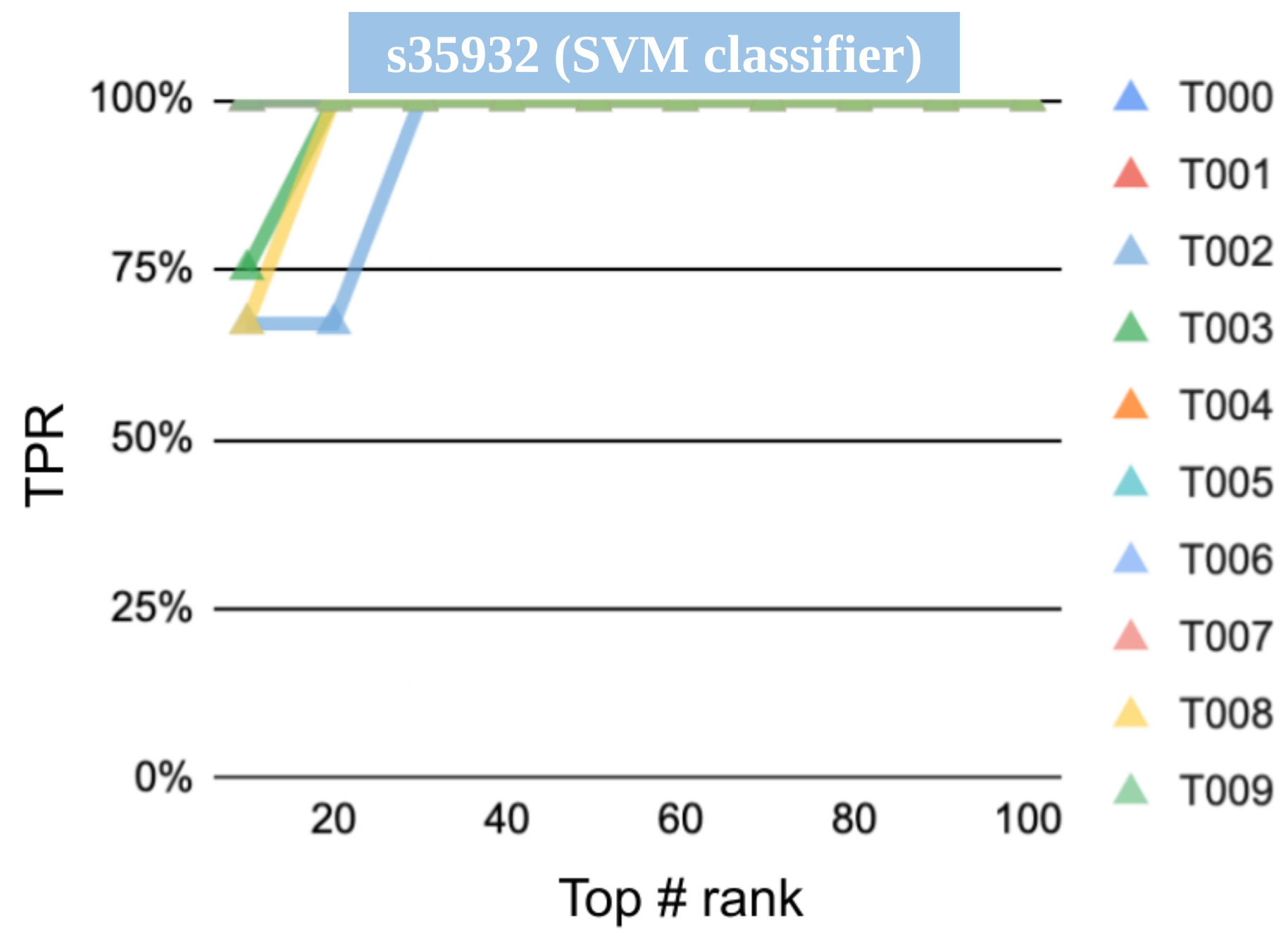


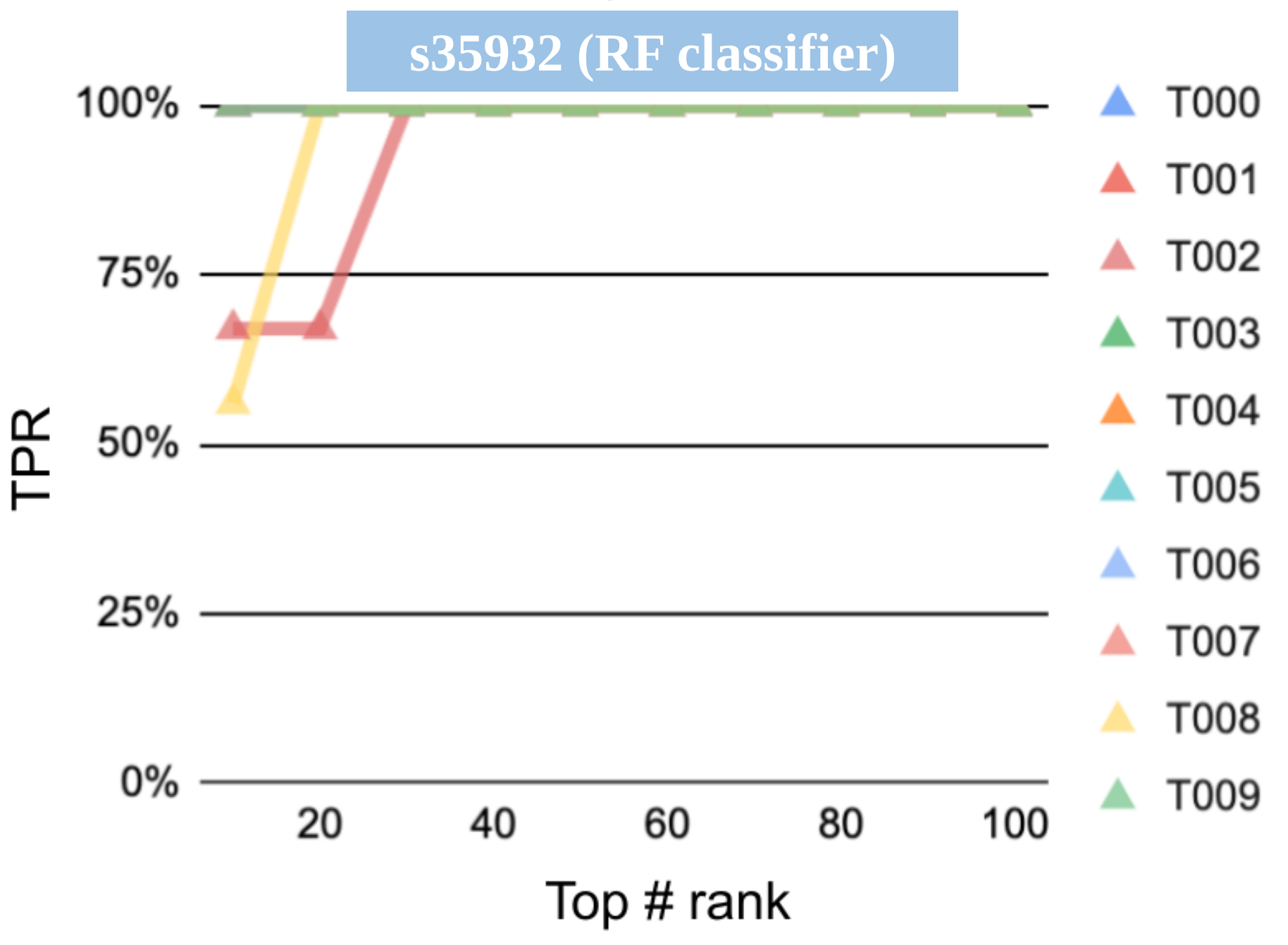


For the results of locating other Trojan gates, we are able to rise the TPRs as we consider lower ranks. For most of the designs below, the TPRs reach 100% before top 100. That is, all Trojan gates are found before rank 100.


s13207 (SVM classifier)
TPR
Top # rank
T000
T001
T002
T003
T004
T005
T006
T007
T008
T009


s13207 (RF classifier)
TPR
Top # rank
T000
T001
T002
T003
T004
T005
T006
T007
T008
T009


s35932 (SVM classifier)
TPR
Top # rank
T000
T001
T002
T003
T004
T005
T006
T007
T008
T009


s35932 (RF classifier)
TPR
Top # rank
T000
T001
T002
T003
T004
T005
T006
T007
T008
T009