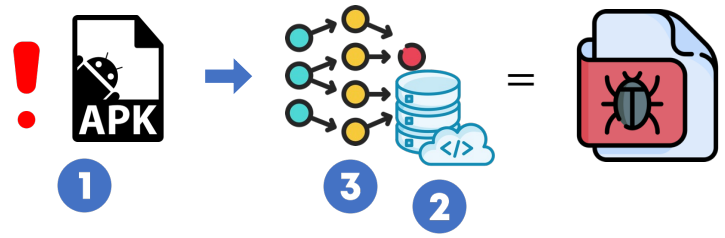
S2Lab Research

LAMD: Context-driven Android Malware Detection and Classification with LLMs

Xingzhi Qian*, Xinran Zheng*, Yiling He, Shuo Yang, Lorenzo Cavallaro

S2Lab Research: Background

Problem: Android malware detection



1. Evolving attacks
2. Dataset biases
3. Limited explainability

→ **LLM**

- Pretrained knowledge
- Zero-shot inference
- Generative capability

Intuition: How analysts examine Android malware

Suspicious APIs → Contextual relationships → Maliciousness detection

Predefined suspicious APIs → Sliced CFG → Explainable code reasoning

Challenges

Excessive support codes in Android applications

Complex program structures

LLM context limit

LLM sequential modeling

→

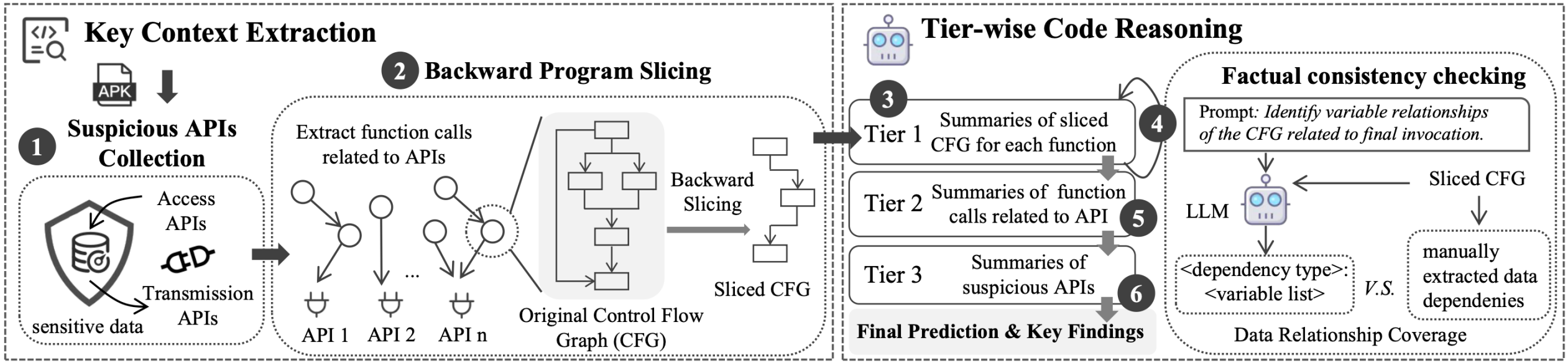
Key context extraction.

Tier-wise Code Reasoning.

Can we extract crucial semantic and structural information from complete application to guide LLMs in detecting Android malware?

S2Lab Research: Workflow

LAMD: LLM-powered practical Android malware detection framework



APK

- *Sensitive data access*
- *Sensitive data transmission*

✓ *Identify suspiciousness*

CFG

1. *Variable retrieval*
2. *Slices extraction*

✓ *Structured representation*

Sliced CFG

1. *Function Behavior Summarization*
2. *API Intent Awareness*
3. *APK Maliciousness Judgement*

✓ *Malware detection*

✓ *Mitigate context limit*

Sliced CFG

1. *Data dependencies*
2. *DRC Metric*

✓ *Factual summary*

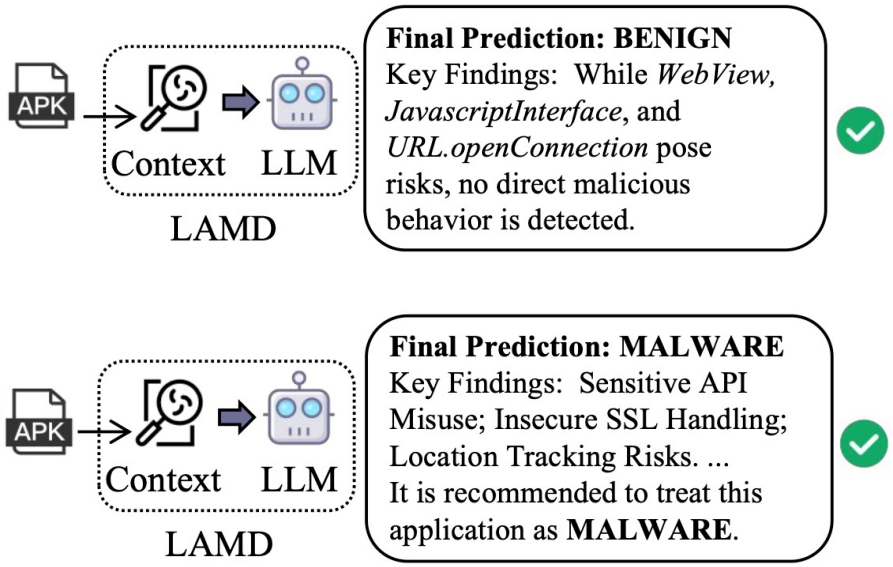
S2Lab Research: Results and Contribution

Evaluation results

LAMD improves F1-scores by 23.12% and reduces FNR by 71.59% on average, enhancing detection reliability. For effectiveness, 81 out of 100 samples are correctly classified into their respective categories.

TABLE 1: Overall detection performance (%) across three test sets with increasing distribution drift (Test 1 < Test 2 < Test 3).

Model	Test 1			Test 2			Test 3		
	F1	FPR	FNR	F1	FPR	FNR	F1	FPR	FNR
Drebin [11]	81.33	**0.40**	24.21	73.60	0.99	31.14	61.59	4.36	37.00
Deep Drebin [12]	71.92	0.62	34.12	69.22	0.85	36.13	66.11	0.95	38.58
Malscan [52]	70.49	0.75	35.12	66.07	**0.73**	33.06	63.91	**0.94**	33.05
LAMD-R	75.34	5.39	15.24	75.71	4.79	12.99	75.83	4.84	11.19
LAMD-F	87.63	2.00	10.37	87.28	1.85	9.74	87.21	1.77	9.83
LAMD	**90.24**	1.26	**8.44**	**90.16**	1.38	**7.79**	**89.85**	1.30	**8.47**



Contribution

- LAMD: LLM-powered practical Android malware detection framework with explainability.
- Key context extraction and tier-wise code reasoning to capture semantics and structural dependencies with factual consistency verification to ensure accuracy.
- Effectiveness in detecting and explaining Android malware, outperforming conventional detectors.