

# Automated detection on the security of the linked-list operations

Hongyu Kuang, Jian Wang, Ruilin Li, Chao Feng,  
YunFei Su, Xing Zhang

Frontiers of Computer Science, [10.1007/s11704-020-0312-6](https://doi.org/10.1007/s11704-020-0312-6)

# Problems & Ideas

- Problems: Currently existed method could not automatically check the security on linked-list operation.
  - How to identify linked-list operations?
  - How to classify the operations on linked lists with fine granularity?
  - How to examine the specifications of linked-list operations?
- Ideas: Parse the source code into token sequences in a fuzzy way through static analysis, and then extract the relationship between these tokens through neural network model to identify whether the operation of linked list meets the secure specifications.

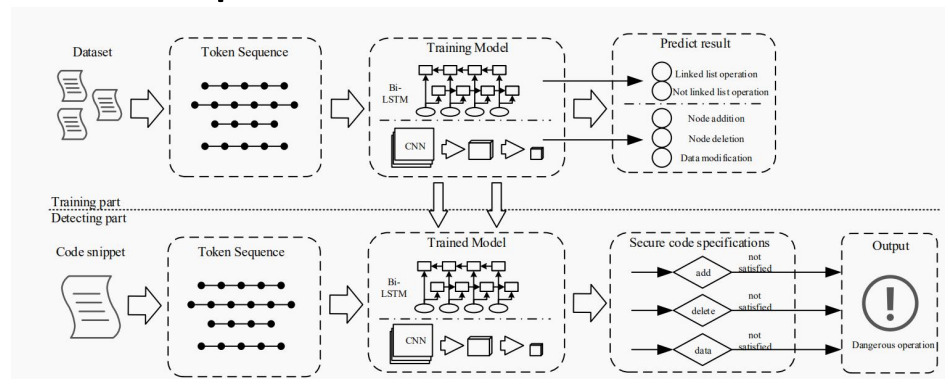


Fig.1 The overview of the framework. The neural network model receives a token sequence as input data. The BLSTM model is used to construct a binary-classifier and the CNN model is used to construct a three-classifier. Finally, output the result by three kinds of secure code specification we defined.

# Main Contributions

**Table 1** Evaluation results on DATA-BIN and DATA-MULTI

Dataset	Model	F1_M	Pr_M	Re_M	F1_W	Pr_W	Re_W
BIN	<b>BLSTM</b>	<b>0.979</b>	<b>0.999</b>	<b>0.960</b>	<b>0.979</b>	<b>0.999</b>	<b>0.960</b>
	BGRU	0.953	0.961	0.945	0.953	0.961	0.945
	CNN	0.970	0.987	0.955	0.970	0.987	0.955
MULTI	BLSTM	0.740	0.733	0.797	0.765	0.815	0.759
	BGRU	0.562	0.681	0.689	0.472	0.738	0.543
	<b>CNN</b>	<b>0.819</b>	<b>0.813</b>	<b>0.842</b>	<b>0.820</b>	<b>0.841</b>	<b>0.818</b>

- We performed experiment with BLSTM, BGRU and CNN respectively. On the DATA-BIN, the value of \*\_M is equal to the value of \*\_W and BLSTM reached best. F1\_\*, Pr\_\* and Re\_\* is 0.97, 0.99, 0.96 respectively. On the DATA-MULTI, CNN reached best. F1\_M, Pr\_M, Re\_M is 0.819, 0.813, 0.842 respectively. F1\_W, Pr\_W, Re\_W is 0.820, 0.841, 0.818 respectively.