硕士学位论文答辩

基于多源知识的开源软件漏洞的补丁识别方法

Finding Patches for Open Source Software Vulnerabilities from Multi-Source Knowledge

答辩人：许聪颖　　导师：陈碧欢

復旦大學
FUDAN UNIVERSITY

目录

01 背景介绍

问题

开源软件安全漏洞越来越多
+
开源软件被广泛使用

→ 大量安全漏洞被引入软件系统

98%的应用使用开源软件，84%的应用含有开源软件漏洞。[1]

举措

CVE、NVD、CVE Details…

漏洞知识库
（漏洞软件名、漏洞补丁…）

→ 安全维护工作
（漏洞检测、漏洞影响分析、漏洞修复…）

基于漏洞补丁知识，进行深度安全维护工作。

[1] Synopsys 公司，《开源安全和风险分析报告》.

举措

基于漏洞补丁知识，进行深度安全维护工作。

本文研究内容

目录

- 补丁知识的**质量和特征研究**
- 自动化补丁查找的**方法设计**

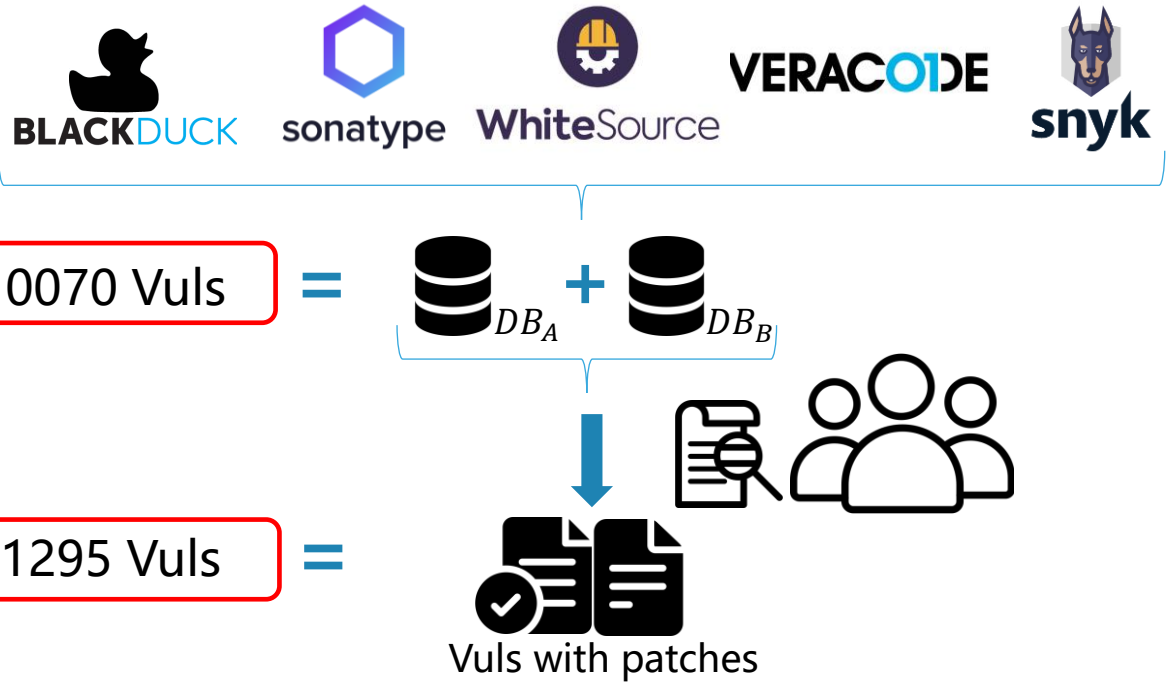
开源软件漏洞补丁的经验研究

研究目的　　探究商业漏洞库中漏洞补丁的质量和特征

研究问题

- RQ1 补丁覆盖率分析
- RQ2 补丁一致性分析
- RQ3 补丁类型分析
- RQ4 补丁映射分析
- RQ5 补丁准确性分析

数据准备



02 开源软件漏洞补丁的经验研究 > 2.2 研究结果

RQ1: 补丁覆盖率分析

补丁覆盖率为45.7% (4602/10070)

(a) 开源软件漏洞　　(b) 含补丁知识的开源软件漏洞

RQ2: 补丁一致性分析

DB_A 与DB_B的补丁一致率为19.7% (907/4602)

不全?　　不准?

表 3-1　DB_A 与 DB_B 补丁一致性分析结果

补丁一致	存在性不一致			内容不一致		
	总数	某一数据库中无漏洞	某一数据库中无补丁	总数	补丁为包含关系	补丁非包含关系
907（19.7%）	3,185（69.2%）	1,392（30.2%）	1,793（39.0%）	510（11.1%）	176（3.8%）	334（7.3%）

一致

RQ3: 补丁类型分析

90+%漏洞补丁都是GitHub代码提交类型

表 3-2　补丁类型分析结果

补丁总数	GitHub 代码提交	SVN 代码提交	其他 Git 平台代码提交
3,043	2,852（93.7%）	136（4.5%）	55（1.8%）
漏洞总数	仅 GitHub 代码提交	仅 SVN 代码提交	仅其他 Git 平台代码提交
1,295	1,202（92.8%）	4（0.3%）	30（2.3%）

RQ4: 补丁映射分析

43.7%(567/1295)漏洞与补丁为一对一映射关系

41.1%(533/1295)漏洞与补丁为一对多映射关系

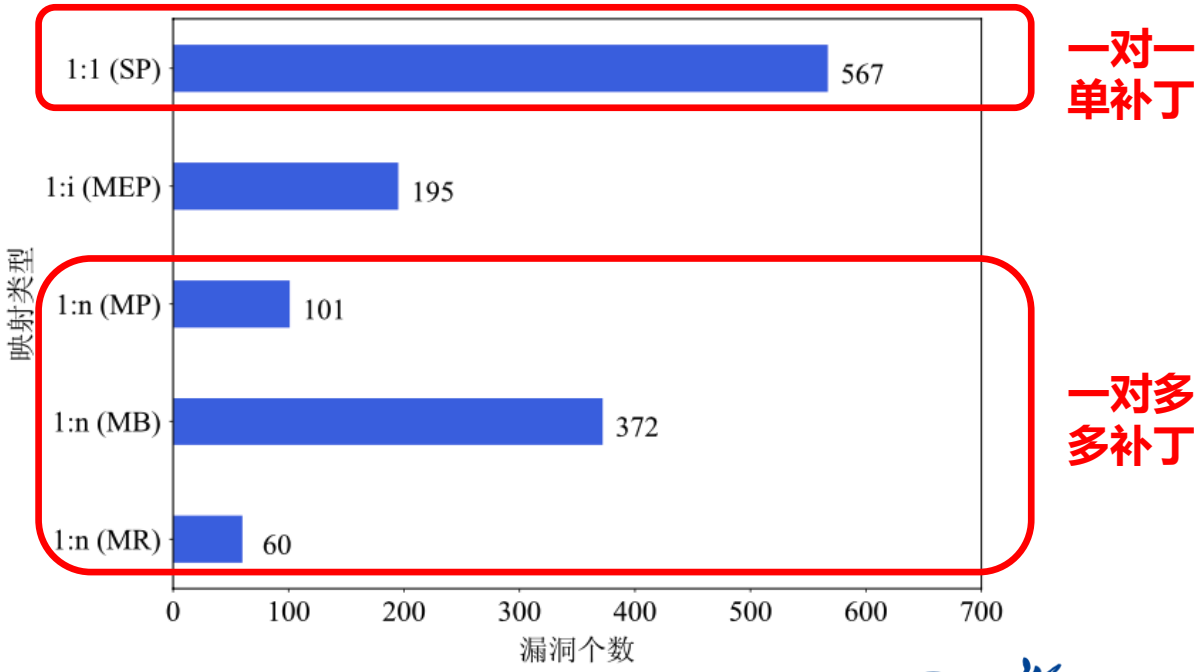


图 3-4　漏洞及其补丁映射类型统计

RQ5: 补丁准确性分析

表 3-3　DB_A 和 DB_B 补丁准确性评估结果

	映射类型	数量	DB_A			DB_B		
			Pre.	Rec.	F1	Pre.	Rec.	F1
单补丁漏洞	1:1 (SP)	567	0.908	0.915	0.910	0.900	0.921	0.906
	1:*i* (MEP)	195	0.935	0.898	0.902	0.924	0.909	0.906
多补丁漏洞	1:*n* (MP)	101	0.923	0.483	0.616	0.911	0.520	0.638
	1:*n* (MB)	372	0.941	0.510	0.620	0.932	0.436	0.555
	1:*n* (MR)	60	0.913	0.610	0.695	0.964	0.526	0.636
	总计	1,295	0.923	0.748	0.793	0.917	0.730	0.771

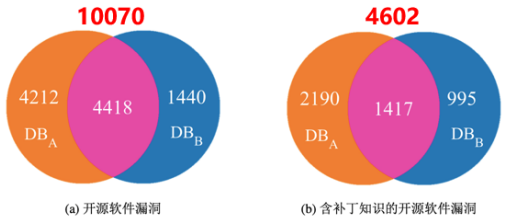
商业漏洞库具有较高的精确率，但近20%的漏洞补丁不全。

02 开源软件漏洞补丁的经验研究 > 2.3 研究发现

商业漏洞数据库中漏洞补丁质量并不理想

RQ1: 补丁覆盖率分析

补丁覆盖率为45.7% (4602/10070)



(a) 开源软件漏洞　(b) 含补丁知识的开源软件漏洞

RQ2: 补丁一致性分析

DB_A 与DB_B的补丁一致率为19.7% (907/4602)

表 3-1　DB_A 与 DB_B 补丁一致性分析结果

不全?　不准?

补丁一致	存在性不一致			内容不一致		
	总数	某一数据库中无漏洞	某一数据库中无补丁	总数	补丁为包含关系	补丁非包含关系
907（19.7%）	3,185（69.2%）	1,392（30.2%）	1,793（39.0%）	510（11.1%）	176（3.8%）	334（7.3%）

一致

RQ5: 补丁准确性分析

表 3-3　DB_A 和 DB_B 补丁准确性评估结果

	映射类型	数量	DB_A			DB_B		
			Pre.	Rec.	F1	Pre.	Rec.	F1
单补丁漏洞	1:1 (SP)	567	0.908	0.915	0.910	0.900	0.921	0.906
	1:*i* (MEP)	195	0.935	0.898	0.902	0.924	0.909	0.906
多补丁漏洞	1:*n* (MP)	101	0.923	0.483	0.616	0.911	0.520	0.638
	1:*n* (MB)	372	0.941	0.510	0.620	0.932	0.436	0.555
	1:*n* (MR)	60	0.913	0.610	0.695	0.964	0.526	0.636
	总计	1,295	0.923	0.748	0.793	0.917	0.730	0.771

商业漏洞库具有较高的精确率，但近20%的漏洞补丁不全。

开源软件漏洞补丁在类型、映射关系方面有一定的特殊性

RQ3: 补丁类型分析

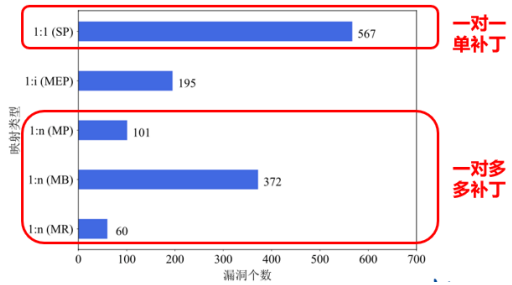
90+%漏洞补丁都是GitHub代码提交类型

表 3-2　补丁类型分析结果

补丁总数	GitHub 代码提交	SVN 代码提交	其他 Git 平台代码提交
3,043	2,852（93.7%）	136（4.5%）	55（1.8%）
漏洞总数	仅 GitHub 代码提交	仅 SVN 代码提交	仅其他 Git 平台代码提交
1,295	1,202（92.8%）	4（0.3%）	30（2.3%）

RQ4: 补丁映射分析

43.7%(567/1295)漏洞与补丁为一对一映射关系
41.1%(533/1295)漏洞与补丁为一对多映射关系



一对一单补丁

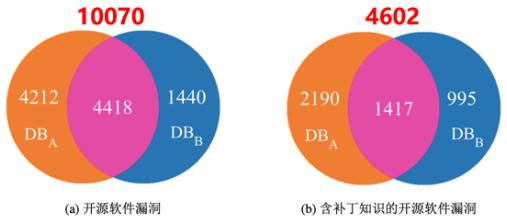
一对多多补丁

02 开源软件漏洞补丁的经验研究 > 2.3 研究发现

商业漏洞数据库中漏洞补丁质量并不理想

RQ1: 补丁覆盖率分析

补丁覆盖率为45.7% (4602/10070)



(a) 开源软件漏洞　(b) 含补丁知识的开源软件漏洞

RQ2: 补丁一致性分析

DB_A 与DB_B的补丁一致率为19.7% (907/4602)

表 3-1　DB_A 与 DB_B 补丁一致性分析结果

不全?　不准?

补丁一致	存在性不一致			内容不一致		
	总数	某一数据库中无漏洞	某一数据库中无补丁	总数	补丁为包含关系	补丁非包含关系
907 (19.7%)	3,185 (69.2%)	1,392 (30.2%)	1,793 (39.0%)	510 (11.1%)	176 (3.8%)	334 (7.3%)

一致

RQ5: 补丁准确性分析

表 3-3　DB_A 和 DB_B 补丁准确性评估结果

	映射类型	数量	DB_A			DB_B		
			Pre.	Rec.	F1	Pre.	Rec.	F1
单补丁漏洞	1:1 (SP)	567	0.908	0.915	0.910	0.900	0.921	0.906
	1:*i* (MEP)	195	0.935	0.898	0.902	0.924	0.909	0.906
多补丁漏洞	1:*n* (MP)	101	0.923	0.483	0.616	0.911	0.520	0.638
	1:*n* (MB)	372	0.941	0.510	0.620	0.932	0.436	0.555
	1:*n* (MR)	60	0.913	0.610	0.695	0.964	0.526	0.636
	总计	1,295	0.923	0.748	0.793	0.917	0.730	0.771

商业漏洞库具有较高的精确率，但近20%的漏洞补丁不全。

开源软件漏洞补丁在类型、映射关系方面有一定的特殊性

RQ3: 补丁类型分析

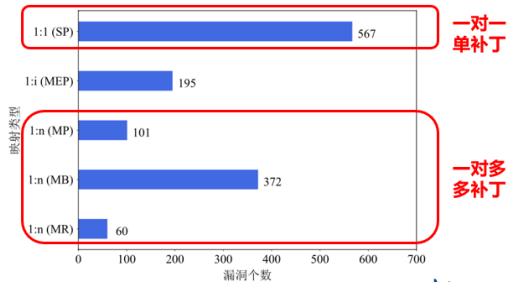
90+%漏洞补丁都是GitHub代码提交类型

表 3-2　补丁类型分析结果

补丁总数	GitHub 代码提交	SVN 代码提交	其他 Git 平台代码提交
3,043	2,852（93.7%）	136（4.5%）	55（1.8%）
漏洞总数	仅 GitHub 代码提交	仅 SVN 代码提交	仅其他 Git 平台代码提交
1,295	1,202（92.8%）	4（0.3%）	30（2.3%）

RQ4: 补丁映射分析

43.7%(567/1295)漏洞与补丁为一对一映射关系
41.1%(533/1295)漏洞与补丁为一对多映射关系



需求：使用自动化工具完善漏洞库

02 开源软件漏洞补丁的经验研究 > 2.3 研究发现

商业漏洞数据库中漏洞补丁质量并不理想

RQ1: 补丁覆盖率分析

补丁覆盖率为45.7% (4602/10070)

(a) 开源软件漏洞　(b) 含补丁知识的开源软件漏洞

RQ2: 补丁一致性分析

DB_A 与 DB_B 的补丁一致率为19.7% (907/4602)

表 3-1　DB_A 与 DB_B 补丁一致性分析结果

不全?　不准?

补丁一致	存在性不一致			内容不一致		
	总数	某一数据库中无漏洞	某一数据库中无补丁	总数	补丁为包含关系	补丁非包含关系
907 (19.7%)	3,185 (69.2%)	1,392 (30.2%)	1,793 (39.0%)	510 (11.1%)	176 (3.8%)	334 (7.3%)

一致

RQ5: 补丁准确性分析

表 3-3　DB_A 和 DB_B 补丁准确性评估结果

	映射类型	数量	DB_A			DB_B		
			Pre.	Rec.	F1	Pre.	Rec.	F1
单补丁漏洞	1:1 (SP)	567	0.908	0.915	0.910	0.900	0.921	0.906
	1:*i* (MEP)	195	0.935	0.898	0.902	0.924	0.909	0.906
多补丁漏洞	1:*n* (MP)	101	0.923	0.483	0.616	0.911	0.520	0.638
	1:*n* (MB)	372	0.941	0.510	0.620	0.932	0.436	0.555
	1:*n* (MR)	60	0.913	0.610	0.695	0.964	0.526	0.636
	总计	1,295	0.923	0.748	0.793	0.917	0.730	0.771

商业漏洞库具有较高的精确率，但近20%的漏洞补丁不全。

开源软件漏洞补丁在类型、映射关系方面有一定的特殊性

RQ3: 补丁类型分析

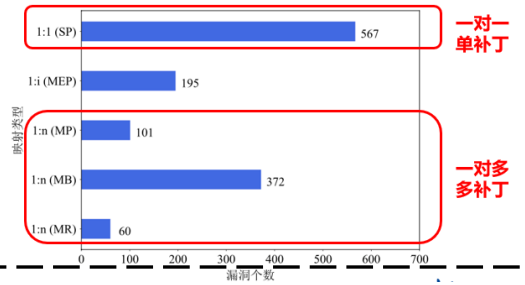
90+%漏洞补丁都是GitHub代码提交类型

表 3-2　补丁类型分析结果

补丁总数	GitHub 代码提交	SVN 代码提交	其他 Git 平台代码提交
3,043	2,852（93.7%）	136（4.5%）	55（1.8%）
漏洞总数	仅 GitHub 代码提交	仅 SVN 代码提交	仅其他 Git 平台代码提交
1,295	1,202（92.8%）	4（0.3%）	30（2.3%）

RQ4: 补丁映射分析

43.7%(567/1295)漏洞与补丁为一对一映射关系
41.1%(533/1295)漏洞与补丁为一对多映射关系



一对一 单补丁

一对多 多补丁

需求： 使用自动化工具完善漏洞库

启发： 设计自动化补丁查找方法

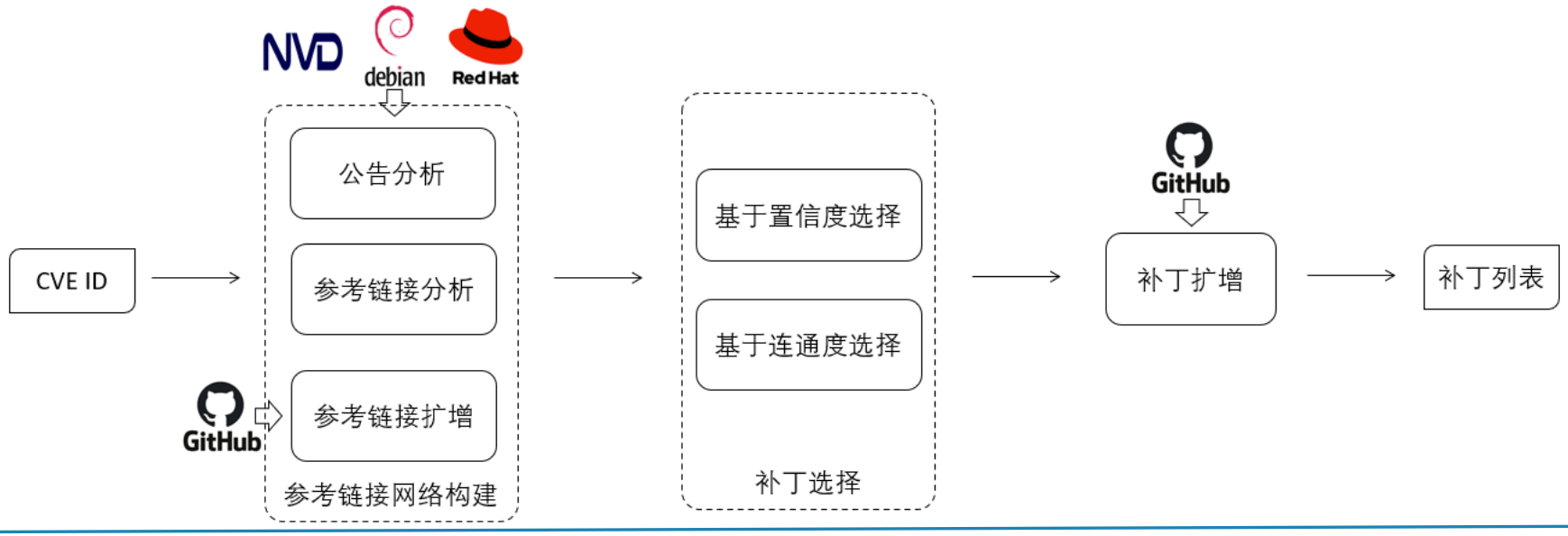
FUDAN UNIVERSITY

开源软件漏洞的补丁识别方法

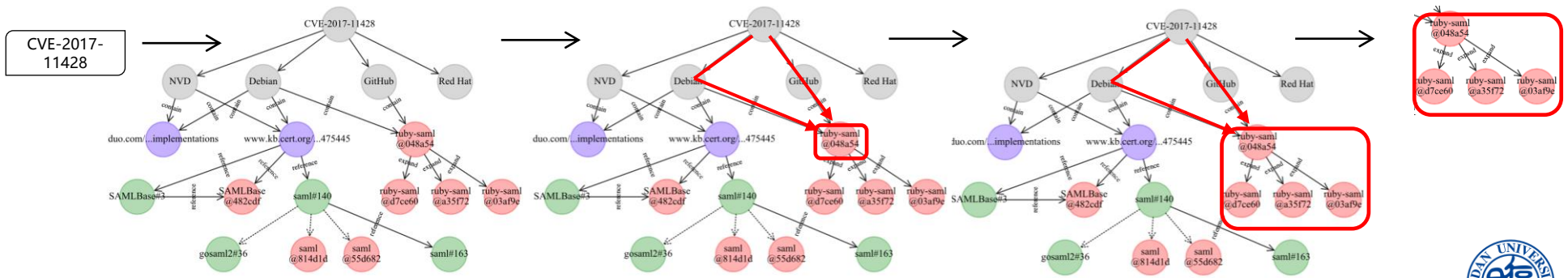
03 开源软件漏洞的补丁识别方法--TRACER

核心思想: 漏洞补丁会在讨论和解决漏洞的、多种来源的漏洞公告、分析报告等参考链接中被频繁地提及和引用。

方法概览



样例展示



实验评估

RQ6: 准确性评估

表 5-1 TRACER VS. 基于启发式规则的方法和商业数据库

映射类型	数量	TRACER				检索 NVD			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.951	0.881	282 (49.7%)	0.973	0.986	0.977
1:*i* (MEP)	195	189 (96.9%)	0.886	0.918	0.888	70 (35.9%)	0.932	0.925	0.921
1:*n* (MP)	101	81 (80.2%)	0.872	0.741	0.761	33 (32.7%)	0.980	0.552	0.683
1:*n* (MB)	372	349 (93.8%)	0.861	0.788	0.795	148 (39.8%)	0.979	0.416	0.546
1:*n* (MR)	60	56 (93.3%)	0.831	0.620	0.659	14 (23.3%)	1.000	0.708	0.794
总计	1,295	1,140 (88.0%)	0.864	0.864	0.837	527 (40.7%)	0.970	0.805	0.842
映射类型	数量	检索 GitHub				检索 NVD 以及 GitHub			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	95 (16.8%)	0.416	0.642	0.471	345 (60.8%)	0.839	0.930	0.864
1:*i* (MEP)	195	33 (16.9%)	0.472	0.490	0.452	91 (46.7%)	0.821	0.867	0.820
1:*n* (MP)	101	28 (27.8%)	0.536	0.445	0.461	49 (48.5%)	0.779	0.605	0.647
1:*n* (MB)	372	126 (33.9%)	0.445	0.236	0.284	201 (54.0%)	0.704	0.393	0.465
1:*n* (MR)	60	23 (38.3%)	0.627	0.345	0.413	33 (55.0%)	0.801	0.539	0.604
总计	1,295	305 (23.6%)	0.461	0.417	0.386	719 (55.5%)	0.793	0.732	0.720
映射类型	数量	DB_A				DB_B			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	100.0%	0.908	0.915	0.910	100.0%	0.900	0.921	0.906
1:*i* (MEP)	195	100.0%	0.935	0.898	0.902	100.0%	0.924	0.909	0.906
1:*n* (MP)	101	100.0%	0.923	0.483	0.616	100.0%	0.911	0.520	0.638
1:*n* (MB)	372	100.0%	0.941	0.510	0.620	100.0%	0.932	0.436	0.555
1:*n* (MR)	60	100.0%	0.913	0.610	0.695	100.0%	0.964	0.526	0.636
总计	1,295	100.0%	0.923	0.748	0.793	100.0%	0.917	0.730	0.771

- VS. 启发式规则，TRACER 显著提高覆盖率和 F1 值
- VS. 商业库DB_A 与DB_B ，TRACER 有更为显着的召回率，略低的精确率和覆盖率。

TRACER 具有较高准确性，
可用于补充现有漏洞数据库缺失的漏洞补丁数据。

RQ7: 削弱性分析

知识源、网络构建、补丁选择和补丁扩增等步骤都有一定的贡献度和必要性。

表 5-2 TRACER 削弱性分析结果（1）

映射类型	数量	TRACER				v_1^1: TRACER w/o NVD			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.951	0.881	281 (49.6%)	0.820	0.936	0.846
1:*i* (MEP)	195	189 (96.9%)	0.886	0.918	0.888	116 (59.5%)	0.882	0.935	0.886
1:*n* (MP)	101	81 (80.2%)	0.872	0.741	0.761	60 (59.4%)	0.881	0.728	0.766
1:*n* (MB)	372	349 (93.8%)	0.861	0.788	0.795	288 (77.4%)	0.876	0.780	0.800
1:*n* (MR)	60	56 (93.3%)	0.831	0.620	0.659	52 (86.7%)	0.848	0.551	0.624
总计	1,295	1,140 (88.0%)	0.864	0.864	0.837	797 (61.5%)	0.856	0.839	0.815
映射类型	数量	v_1^2: TRACER w/o Debian				v_1^3: TRACER w/o Red Hat			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	457 (80.6%)	0.847	0.943	0.869	454 (80.1%)	0.853	0.943	0.874
1:*i* (MEP)	195	187 (95.6%)	0.880	0.912	0.882	188 (96.4%)	0.883	0.918	0.886
1:*n* (MP)	101	79 (78.2%)	0.851	0.716	0.739	80 (79.2%)	0.880	0.736	0.760
1:*n* (MB)	372	344 (92.5%)	0.838	0.760	0.771	337 (90.6%)	0.844	0.761	0.767
1:*n* (MR)	60	55 (91.7%)	0.819	0.613	0.651	56 (93.3%)	0.738	0.640	0.618
总计	1,295	1,122 (86.6%)	0.848	0.849	0.821	1,115 (86.1%)	0.851	0.853	0.823
映射类型	数量	v_1^4: TRACER w/o GitHub				v_1^5: TRACER w/o Network			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	418 (73.7%)	0.898	0.943	0.908	390 (68.8%)	0.910	0.972	0.925
1:*i* (MEP)	195	176 (90.3%)	0.887	0.921	0.892	117 (60.0%)	0.956	0.959	0.941
1:*n* (MP)	101	73 (72.3%)	0.873	0.690	0.726	61 (60.4%)	0.943	0.669	0.743
1:*n* (MB)	372	333 (89.5%)	0.874	0.752	0.773	263 (70.7%)	0.908	0.575	0.659
1:*n* (MR)	60	53 (88.3%)	0.816	0.545	0.604	50 (83.3%)	0.920	0.641	0.712
总计	1,295	1,053 (81.3%)	0.883	0.841	0.835	881 (68.0%)	0.918	0.812	0.823

表 5-3 TRACER 削弱性分析结果（2）

映射类型	数量	v_2^1: TRACER w/o Selection				v_2^2: TRACER w/o Connectivity			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.632	0.961	0.680	322 (56.8%)	0.892	0.978	0.913
1:*i* (MEP)	195	189 (96.9%)	0.622	0.976	0.682	84 (43.1%)	0.929	0.939	0.915
1:*n* (MP)	101	81 (80.2%)	0.615	0.933	0.656	45 (44.6%)	0.953	0.685	0.764
1:*n* (MB)	372	349 (93.8%)	0.616	0.903	0.657	181 (48.7%)	0.927	0.787	0.821
1:*n* (MR)	60	56 (93.3%)	0.368	0.891	0.394	33 (55.0%)	0.885	0.722	0.772
总计	1,295	1,140 (88.0%)	0.611	0.940	0.658	665 (51.4%)	0.910	0.889	0.871
映射类型	数量	v_2^3: TRACER w/o Confidence				v_2^4: TRACER with Path Length			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.942	0.879	465 (82.0%)	0.833	0.957	0.859
1:*i* (MEP)	195	189 (96.9%)	0.888	0.913	0.889	189 (96.9%)	0.848	0.945	0.867
1:*n* (MP)	101	81 (80.2%)	0.880	0.722	0.751	81 (80.2%)	0.849	0.760	0.742
1:*n* (MB)	372	349 (93.8%)	0.871	0.765	0.784	349 (93.8%)	0.830	0.798	0.770
1:*n* (MR)	60	56 (93.3%)	0.849	0.462	0.550	56 (93.3%)	0.652	0.747	0.590
总计	1,295	1,140 (88.0%)	0.869	0.844	0.826	1,140 (88.0%)	0.827	0.882	0.812
映射类型	数量	v_2^5: TRACER with Path Number				v_3: TRACER w/o Expansion			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.805	0.951	0.837	465 (82.0%)	0.871	0.948	0.889
1:*i* (MEP)	195	189 (96.9%)	0.849	0.920	0.858	189 (96.9%)	0.910	0.914	0.902
1:*n* (MP)	101	81 (80.2%)	0.801	0.756	0.726	81 (80.2%)	0.873	0.696	0.732
1:*n* (MB)	372	349 (93.8%)	0.833	0.811	0.791	349 (93.8%)	0.860	0.506	0.590
1:*n* (MR)	60	56 (93.3%)	0.789	0.630	0.644	56 (93.3%)	0.847	0.567	0.629
总计	1,295	1,140 (88.0%)	0.819	0.873	0.809	1,140 (88.0%)	0.873	0.771	0.776

知识源、网络构建、补丁选择和补丁扩增等步骤都有一定的贡献度和必要性。

RQ8: 敏感度分析

TRACER 准确率对参数变化不是非常敏感。
网络深度为 5 层 、提交时间跨度为 30 天时，
效果相对最优。

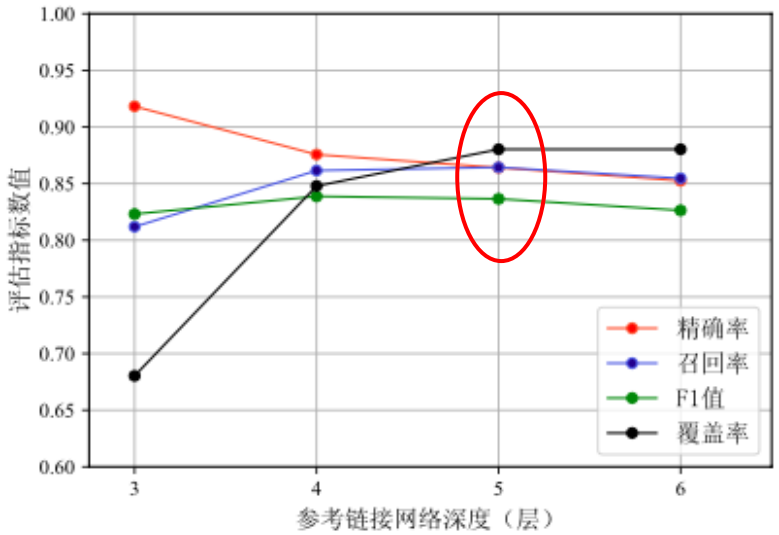


图 5-3　网络深度限制（层数）敏感性分析结果

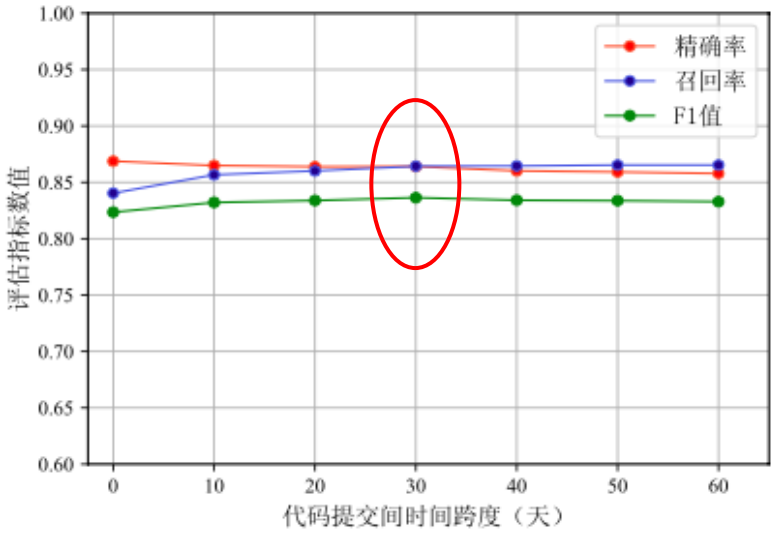


图 5-4　提交时间跨度的敏感性分析结果

- 随着网络层数的增加，网络中将包含更多补丁。
- 随着代码提交时间跨度的增加，TRACER 会搜索到更广的代码提交。

TRACER 准确率对参数变化不是非常敏感。
网络深度为 5 层 、提交时间跨度为 30 天时，效果相对最优。

RQ9: 通用性分析

- 数据集一 3,185个漏洞，TRACER 补丁覆盖率 67.7% (2,155/3,185) 。
- 数据集二 5,468个漏洞，TRACER 补丁覆盖率 51.5% (2,816/5,468) 。

覆盖率评估

表 5-4 TRACER 通用性分析结果

评估对象	数据集一（91 个漏洞）				数据集二（89 个漏洞）			
	Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
TRACER	100.0%	0.823	0.845	0.784	100.0%	0.888	0.899	0.867
DB_A	62 (68.1%)	0.935	0.827	0.858	0.0%	–	–	–
DB_B	29 (31.8%)	0.885	0.664	0.725	0.0%	–	–	–

准确率评估

对于更大范围的漏洞，TRACER 依旧具有较好的准确率，通用性较好。

**对于更大范围的漏洞，
TRACER 依旧具有较好的准确率，通用性较好。**

RQ10: 实用性分析

- 从国内外多所高校和科技公司共招募了 10 名实验人员，包括：博士后、博士生、硕士生以及工程师。
- 随机选取10 个漏洞，对比分析在有无TRACER 的情况下，用户查找补丁的用时和准确性。

表 5-5　用户研究中任务的用时和准确率

任务	w/o TRACER				with TRACER			
	用时	Pre.	Rec.	F1	用时	Pre.	Rec.	F1
全部 10 个任务	5.66 mins	0.880	0.677	0.765	4.66 mins	0.983	0.920	0.951
5 单补丁任务	5.60 mins	0.960	0.960	0.960	3.84 mins	1.000	1.000	1.000
5 多补丁任务	5.72 mins	0.800	0.393	0.527	5.48 mins	0.967	0.840	0.899

TRACER 有助于用户更准确、更快速地查找到补丁。

04 实验评估 > 4.3 实验结论

本文从准确性、削弱性、敏感度、通用性及实用性对TRACER 进行了评估。结果表明:

RQ6: 准确性评估

表 5-1 TRACER VS. 基于启发式规则的方法和商业数据库

映射类型	数量	TRACER				检索 NVD			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.951	0.881	282 (49.7%)	0.973	0.986	0.977
1:1 (MEP)	195	189 (96.9%)	0.886	0.918	0.888	70 (35.9%)	0.932	0.925	0.921
1:n (MP)	101	81 (80.2%)	0.872	0.741	0.761	33 (32.7%)	0.980	0.552	0.683
1:n (MB)	372	349 (93.8%)	0.861	0.788	0.795	148 (39.8%)	0.979	0.416	0.546
1:n (MR)	60	56 (93.3%)	0.831	0.620	0.659	14 (23.3%)	1.000	0.708	0.794
总计	1,295	1,140 (88.0%)	0.864	0.864	0.837	527 (40.7%)	0.970	0.805	0.842
映射类型	数量	检索 GitHub				检索 NVD 以及 GitHub			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	95 (16.8%)	0.416	0.642	0.471	345 (60.8%)	0.839	0.930	0.864
1:1 (MEP)	195	100.0%	0.935	0.898	0.902	100.0%	0.924	0.909	0.906
1:n (MP)	101	100.0%	0.923	0.483	0.616	100.0%	0.911	0.520	0.638
1:n (MB)	372	100.0%	0.941	0.510	0.620	100.0%	0.932	0.436	0.555
1:n (MR)	60	100.0%	0.913	0.610	0.695	100.0%	0.964	0.526	0.636
总计	1,295	100.0%	0.923	0.748	0.793	100.0%	0.917	0.730	0.771

- VS. 启发式规则，TRACER 显著提高覆盖率和 F1 值
- VS. 商业库DB_A 与DB_B ，TRACER 有更为显着的召回率，略低的精确率和覆盖率。

TRACER 具有较高准确性，
可用于补充现有漏洞数据库缺失的漏洞补丁数据。

RQ7: 削弱性分析

表 5-2 TRACER 削弱性分析结果（1）

映射类型	数量	TRACER				v_1^1: TRACER w/o NVD			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.951	0.881	281 (49.6%)	0.820	0.936	0.846
1:1 (MEP)	195	189 (96.9%)	0.886	0.918	0.888	116 (59.5%)	0.882	0.935	0.886
1:n (MP)	101	81 (80.2%)	0.872	0.741	0.761	60 (59.4%)	0.881	0.728	0.766
1:n (MB)	372	349 (93.8%)	0.861	0.788	0.795	288 (77.4%)	0.876	0.780	0.800
1:n (MR)	60	56 (93.3%)	0.831	0.620	0.659	52 (86.7%)	0.848	0.551	0.624
总计	1,295	1,140 (88.0%)	0.864	0.864	0.837	797 (61.5%)	0.856	0.839	0.815
映射类型	数量	v_1^2: TRACER w/o Debian				v_1^3: TRACER w/o Red Hat			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	457 (80.6%)	0.847	0.943	0.869	454 (80.1%)	0.853	0.943	0.874
1:1 (MEP)	195	187 (95.6%)	0.880	0.912	0.882	188 (96.4%)	0.883	0.918	0.886
1:n (MP)	101	79 (78.2%)	0.851	0.716	0.739	80 (79.2%)	0.880	0.736	0.760
1:n (MP)	101	73 (72.3%)	0.873	0.690	0.726	61 (60.4%)	0.943	0.669	0.743
1:n (MB)	372	333 (89.5%)	0.874	0.752	0.773	263 (70.7%)	0.908	0.575	0.659
1:n (MR)	60	53 (88.3%)	0.816	0.545	0.604	50 (83.3%)	0.920	0.641	0.712
总计	1,295	1,053 (81.3%)	0.883	0.841	0.835	881 (68.0%)	0.918	0.812	0.823

表 5-3 TRACER 削弱性分析结果（2）

映射类型	数量	v_1^4: TRACER w/o Selection				v_1^5: TRACER w/o Connectivity			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.632	0.961	0.680	322 (56.8%)	0.892	0.978	0.913
1:1 (MEP)	195	189 (96.9%)	0.622	0.976	0.682	84 (43.1%)	0.929	0.939	0.915
1:n (MP)	101	81 (80.2%)	0.615	0.933	0.656	45 (44.6%)	0.953	0.685	0.764
1:n (MB)	372	349 (93.8%)	0.616	0.903	0.657	181 (48.7%)	0.927	0.787	0.821
1:n (MR)	60	56 (93.3%)	0.368	0.891	0.394	33 (55.0%)	0.885	0.722	0.772
总计	1,295	1,140 (88.0%)	0.611	0.940	0.658	665 (51.4%)	0.910	0.889	0.871
映射类型	数量	v_1^6: TRACER w/o Confidence				v_1^7: TRACER with Path Length			
		Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
1:1 (SP)	567	465 (82.0%)	0.860	0.942	0.879	465 (82.0%)	0.833	0.957	0.859
1:1 (MEP)	195	189 (96.9%)	0.888	0.913	0.889	189 (96.9%)	0.848	0.945	0.867
1:n (MP)	101	81 (80.2%)	0.880	0.722	0.751	81 (80.2%)	0.849	0.760	0.742
1:n (MP)	101	81 (80.2%)	0.801	0.756	0.726	81 (80.2%)	0.873	0.696	0.732
1:n (MB)	372	349 (93.8%)	0.833	0.811	0.791	349 (93.8%)	0.860	0.506	0.590
1:n (MR)	60	56 (93.3%)	0.789	0.630	0.644	56 (93.3%)	0.847	0.567	0.629
总计	1,295	1,140 (88.0%)	0.819	0.873	0.809	1,140 (88.0%)	0.873	0.771	0.776

知识源、网络构建、补丁选择
和补丁扩增等步骤都有一定的贡献度和必要性。

RQ8: 敏感度分析

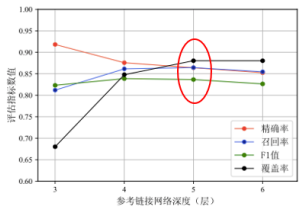


图 5-3 网络深度限制（层数）敏感性分析结果

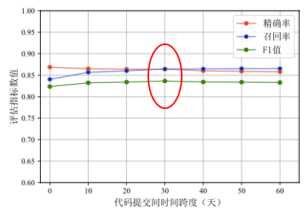


图 5-4 提交时间跨度的敏感性分析结果

TRACER 准确率对参数变化不是非常敏感。
网络深度为 5 层 、提交时间跨度为 30 天时，
效果相对最优。

RQ9: 通用性分析

- 数据集一 3,185个漏洞，TRACER 补丁覆盖率 67.7% (2,155/3,185) 。
- 数据集二 5,468个漏洞，TRACER 补丁覆盖率 51.5% (2,816/5,468) 。

覆盖率评估

表 5-4 TRACER 通用性分析结果

评估对象	数据集一（91 个漏洞）				数据集二（89 个漏洞）			
	Coverage	Pre.	Rec.	F1	Coverage	Pre.	Rec.	F1
TRACER	100.0%	0.823	0.845	0.784	100.0%	0.888	0.899	0.867
DB_A	62 (68.1%)	0.935	0.827	0.858	0.0%	–	–	–

准确率评估

对于更大范围的漏洞，
TRACER 依旧具有较好的准确率，通用性较好。

RQ10: 实用性分析

- 从国内外多所高校和科技公司共招募了 10 名实验人员，包括：博士后、博士生、硕士生以及工程师
- 随机选取10 个漏洞，对比分析在有无TRACER 的情况下，用户查找补丁的用时和准确性。

表 5-5 用户研究中任务的用时和准确率

任务	w/o TRACER				with TRACER			
	用时	Pre.	Rec.	F1	用时	Pre.	Rec.	F1
全部 10 个任务	5.66 mins	0.880	0.677	0.765	4.66 mins	0.983	0.920	0.951
5 单补丁任务	5.60 mins	0.960	0.960	0.960	3.84 mins	1.000	1.000	1.000
5 多补丁任务	5.72 mins	0.800	0.393	0.527	5.48 mins	0.967	0.840	0.899

TRACER 有助于用户更准确、更快速地查找到补丁。

总结与展望

05 总结与展望

本文总结

- 本文开展了一项针对开源软件漏洞补丁质量和特征的经验研究，涵盖补丁覆盖度、补丁一致性、补丁类型、补丁映射关系以及补丁准确性。
  - 发现: 商业漏洞库中补丁的质量并不理想，且漏洞补丁在类型、映射关系方面有特殊性。
- 基于经验研究的发现，本文提出了基于多源知识的开源软件漏洞的补丁识别方法(TRACER)。
- 本文进行大量实验，从准确性、削弱性、敏感度、通用性及实用性对TRACER进行评估。

未来展望

- 扩增输入类型：CVE ID + Advisory ID、Issue ID ……
- 扩增知识源：NRDG+ CNNVD、GitHub Advisory ……
- 升级补丁选择方法：基于置信度和连通度 -> 基于语义

硕士学位论文答辩

谢谢聆听！

基于多源知识的开源软件漏洞的补丁识别方法

Finding Patches for Open Source Software Vulnerabilities from Multi-Source Knowledge

答辩人：许聪颖　　导师：陈碧欢

复旦大學
FUDAN UNIVERSITY