

TABLE 1: Effectiveness Comparison between *PatchDiscovery* and *BinXray*

Program	PatchDiscovery				BinXray				Test Rate
	P	R	F <sub>1</sub>	A	P	R	F <sub>1</sub>	A	
binutils	0.866	0.940	0.902	0.929	0.678	0.967	0.797	0.837	0.871
Expat	1.000	1.000	1.000	1.000	1.000	0.750	0.857	0.875	1.000
FFmpeg	0.931	0.866	0.897	0.897	0.773	0.715	0.743	0.736	0.959
Freetype	0.983	0.954	0.968	0.976	0.940	0.866	0.901	0.921	0.903
Libexif	1.000	1.000	1.000	1.000	0.583	1.000	0.737	0.805	1.000
Libpng	1.000	1.000	1.000	1.000	0.759	1.000	0.863	0.957	1.000
Libxml2	0.983	1.000	0.991	0.991	1.000	0.982	0.991	0.991	1.000
Libxslt	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Openssl	0.972	0.930	0.951	0.962	0.973	0.905	0.938	0.952	1.000
Openvpn	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Sqlite3	0.932	0.873	0.902	0.876	0.833	0.789	0.811	0.785	0.670
Tcpdump	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Total	0.943	0.903	0.922	0.930	0.831	0.803	0.817	0.831	0.958

TABLE 2: Resilience to Version Gap Comparison between *PatchDiscovery* and *BinXray*

Version Gap	PatchDiscovery				BinXray			
	P	R	F <sub>1</sub>	A	P	R	F <sub>1</sub>	A
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	1.000	0.978	0.989	0.991	1.000	0.961	0.980	0.984
2	0.982	0.969	0.976	0.979	0.935	0.942	0.938	0.948
3	0.979	0.934	0.956	0.965	0.899	0.873	0.886	0.907
4	0.966	0.941	0.953	0.965	0.858	0.833	0.845	0.877
5	0.930	0.938	0.934	0.950	0.772	0.795	0.783	0.824
6	0.882	0.907	0.894	0.922	0.741	0.714	0.727	0.779
7	0.883	0.901	0.892	0.918	0.679	0.658	0.668	0.734
8	0.902	0.876	0.889	0.917	0.660	0.601	0.629	0.715
9	0.924	0.882	0.902	0.918	0.659	0.663	0.661	0.718
≥10	0.918	0.840	0.877	0.889	0.780	0.734	0.756	0.776
Average	0.942	0.924	0.933	0.947	0.816	0.798	0.807	0.842

TABLE 3: Resilience to Function Size Comparison between *PatchDiscovery* and *BinXray*

Function Size	PatchDiscovery				BinXray			
	P	R	F <sub>1</sub>	A	P	R	F <sub>1</sub>	A
0-40	0.955	0.927	0.941	0.944	0.832	0.780	0.805	0.818
40-80	0.932	0.881	0.906	0.928	0.820	0.806	0.813	0.846
80-120	0.930	0.901	0.915	0.919	0.751	0.780	0.765	0.767
120-160	0.961	1.000	0.980	0.983	0.986	0.969	0.977	0.979
160-200	0.881	0.744	0.807	0.828	0.892	0.847	0.869	0.870
≥200	0.937	0.867	0.901	0.915	0.833	0.909	0.870	0.880
Average	0.933	0.887	0.908	0.919	0.853	0.848	0.850	0.860

TABLE 4: Resilience to Patch Size Comparison between *PatchDiscovery* and *BinXray*

Patch Size	PatchDiscovery				BinXray			
	P	R	F <sub>1</sub>	A	P	R	F <sub>1</sub>	A
0-10	0.942	0.880	0.910	0.924	0.821	0.784	0.802	0.825
10-20	0.940	0.941	0.940	0.940	0.871	0.861	0.866	0.862
20-30	0.930	0.957	0.943	0.958	0.912	1.000	0.954	0.959
30-40	1.000	1.000	1.000	1.000	0.722	1.000	0.838	0.842
40-50	0.928	0.883	0.905	0.919	0.834	0.736	0.782	0.796
50-60	0.974	1.000	0.987	0.982	1.000	0.688	0.815	0.878
60-70	0.933	1.000	0.966	0.948	1.000	1.000	1.000	1.000
70-80	1.000	1.000	1.000	1.000	0.467	1.000	0.636	0.636
≥80	0.932	0.942	0.937	0.930	0.858	0.818	0.837	0.821
Average	0.953	0.956	0.954	0.956	0.832	0.876	0.837	0.847