TABLE I: Effectiveness of PatchDiscovery

Program	TP	TN	FP	FN	Precision	Recall	F1	Accuracy
binutils	343	643	61	19	0.849	0.948	0.896	0.925
expat	4	4	0	0	1.000	1.000	1.000	1.000
ffmpeg	2386	2267	187	336	0.927	0.877	0.901	0.899
freetype	282	470	2	22	0.993	0.928	0.959	0.969
libexif	19	56	0	2	1.000	0.905	0.950	0.974
libpng	22	140	0	0	1.000	1.000	1.000	1.000
libxml2	373	343	14	38	0.964	0.908	0.935	0.932
libxslt	7	8	0	0	1.000	1.000	1.000	1.000
openssl	941	1388	102	34	0.902	0.965	0.933	0.945
openvpn	27	19	0	0	1.000	1.000	1.000	1.000
tcpdump	84	168	0	0	1.000	1.000	1.000	1.000
sqlite3	55	32	2	8	0.965	0.873	0.917	0.897
Total	4543	5538	368	459	0.925	0.908	0.917	0.924

TABLE II: PatchDiscovery's Resilience to Version Gap

Version Gap	TP	TN	FP	FN	Precision	Recall	F1	Accuracy
0	544	549	3	8	0.995	0.986	0.990	0.990
1	402	535	9	17	0.978	0.959	0.969	0.973
2	373	434	10	20	0.974	0.949	0.961	0.964
3	280	425	15	23	0.949	0.924	0.936	0.949
4	253	379	28	22	0.900	0.920	0.910	0.927
5	224	339	33	20	0.872	0.918	0.894	0.914
6	207	309	31	21	0.870	0.908	0.888	0.908
7	189	262	23	19	0.892	0.909	0.900	0.915
8	156	240	19	17	0.891	0.902	0.897	0.917
9	142	187	18	17	0.887	0.893	0.890	0.904
$\geq$10	1773	1879	179	275	0.908	0.866	0.886	0.889
Total	4543	5538	368	459	0.925	0.908	0.917	0.924

TABLE III: PatchDiscovery's Resilience to Function Size

Function Size	TP	TN	FP	FN	Precision	Recall	F1	Accuracy
0-100	3438	4280	201	348	0.945	0.908	0.926	0.934
100-200	948	1038	149	87	0.864	0.916	0.889	0.894
200-300	101	101	17	0	0.856	1.000	0.922	0.922
300-400	14	21	0	7	1.000	0.667	0.800	0.833
400-500	24	64	1	8	0.960	0.750	0.842	0.907
$\geq$500	18	34	0	9	1.000	0.667	0.800	0.852
Total	4543	5538	368	459	0.925	0.908	0.917	0.924

TABLE IV: PatchDiscovery's Resilience to Patch Size

Patch Size	TP	TN	FP	FN	Precision	Recall	F1	Accuracy
0-10	2910	3864	273	361	0.914	0.890	0.902	0.914
10-20	822	790	53	46	0.939	0.947	0.943	0.942
20-30	48	118	3	3	0.941	0.941	0.941	0.965
30-40	103	132	0	0	1.000	1.000	1.000	1.000
40-50	194	306	18	31	0.915	0.862	0.888	0.911
50-60	185	99	1	0	0.995	1.000	0.997	0.996
60-70	17	13	0	2	1.000	0.895	0.944	0.938
70-80	4	9	0	0	1.000	1.000	1.000	1.000
$\geq$80	260	207	20	16	0.929	0.942	0.935	0.928
Total	4543	5538	368	459	0.925	0.908	0.917	0.924

TABLE V: Effectiveness Comparison between PatchDiscovery and BinXray

Program	TP		TN		FP		FN		Precision		Recall		F1		Accuracy	
	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX
binutils	236	237	363	347	20	36	10	9	**0.922**	0.868	0.959	0.963	**0.940**	0.913	**0.952**	0.928
expat	3	3	4	4	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
ffmpeg	1851	1580	1715	1357	122	385	110	476	**0.938**	0.804	**0.944**	0.768	**0.941**	0.786	**0.939**	0.773
freetype	175	169	324	324	0	0	0	6	**1.000**	1.000	**1.000**	0.966	**1.000**	0.983	**1.000**	0.988
libexif	13	14	7	7	0	0	1	0	**1.000**	1.000	0.929	1.000	0.963	1.000	0.952	1.000
libpng	22	22	140	133	0	7	0	0	**1.000**	0.759	**1.000**	1.000	**1.000**	0.863	**1.000**	0.957
libxml2	373	389	334	348	14	0	23	7	0.964	1.000	0.942	0.982	0.953	0.991	0.950	0.991
libxslt	7	7	8	8	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
openssl	922	871	1307	1370	86	24	34	84	0.915	0.973	**0.964**	0.912	0.939	0.942	0.949	0.954
openvpn	27	27	19	19	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
tcpdump	81	81	162	162	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
sqlite3	23	23	17	17	2	2	0	0	**0.920**	0.920	**1.000**	1.000	**0.958**	0.958	**0.952**	0.952
Total	3733	3423	4400	4096	244	454	178	582	**0.939**	0.883	**0.954**	0.855	**0.947**	0.869	**0.951**	0.879

TABLE VI: Resilience to Version Gap Comparison between PatchDiscovery and BinXray

Version Gap	TP		TN		FP		FN		Precision		Recall		F1		Accuracy	
	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX
0	532	538	535	538	3	0	6	0	0.994	1.000	0.989	1.000	0.992	1.000	0.992	1.000
1	350	353	443	447	5	1	10	7	0.986	0.997	0.972	0.981	0.979	0.989	0.981	0.990
2	307	308	352	351	9	8	11	12	0.972	0.975	**0.965**	0.963	0.968	0.969	**0.971**	0.971
3	236	227	331	327	12	14	8	19	**0.952**	0.942	**0.967**	0.923	**0.959**	0.932	**0.966**	0.944
4	215	202	287	280	21	24	9	26	**0.911**	0.894	**0.960**	0.886	**0.935**	0.890	**0.944**	0.906
5	195	180	260	235	20	36	11	35	**0.907**	0.833	**0.947**	0.837	**0.926**	0.835	**0.936**	0.854
6	168	141	237	206	18	45	9	40	**0.903**	0.758	**0.949**	0.779	**0.926**	0.768	**0.938**	0.803
7	154	120	202	170	15	38	10	53	**0.911**	0.759	**0.939**	0.694	**0.925**	0.725	**0.934**	0.761
8	127	89	190	156	10	40	8	50	**0.927**	0.690	**0.941**	0.640	**0.934**	0.664	**0.946**	0.731
9	121	95	142	120	12	28	8	40	**0.910**	0.772	**0.938**	0.704	**0.924**	0.736	**0.929**	0.760
$\geq$10	1328	1170	1421	1266	119	220	88	300	**0.918**	0.842	**0.938**	0.796	**0.928**	0.818	**0.930**	0.824
Total	3733	3423	4400	4096	244	454	178	582	**0.939**	0.883	**0.954**	0.855	**0.947**	0.869	**0.951**	0.879

TABLE VII: Resilience to Function Size Comparison between PatchDiscovery and BinXray

Function Size	TP		TN		FP		FN		Precision		Recall		F1		Accuracy	
	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX
0-100	3001	2752	3746	3416	128	382	163	488	**0.959**	0.878	**0.948**	0.849	**0.954**	0.864	**0.959**	0.876
100-200	632	582	567	582	105	72	14	82	0.858	0.890	**0.978**	0.877	**0.914**	0.883	**0.910**	0.883
200-300	76	64	58	69	11	0	0	12	0.874	1.000	**1.000**	0.842	**0.933**	0.914	**0.924**	0.917
300-400	7	7	11	11	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
400-500	13	13	13	13	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
$\geq$500	4	5	5	5	0	0	1	0	**1.000**	1.000	0.800	1.000	0.889	1.000	0.900	1.000
Total	3733	3423	4400	4096	244	454	178	582	**0.939**	0.883	**0.954**	0.855	**0.947**	0.869	**0.951**	0.879

TABLE VIII: Resilience to Patch Size Comparison between PatchDiscovery and BinXray

Patch Size	TP		TN		FP		FN		Precision		Recall		F1		Accuracy	
	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX	PD	BX
0-10	2651	2449	3318	3132	201	304	145	430	**0.930**	0.890	**0.948**	0.851	**0.939**	0.870	**0.945**	0.884
10-20	722	659	673	611	43	104	32	96	**0.944**	0.864	**0.958**	0.873	**0.951**	0.868	**0.949**	0.864
20-30	30	30	44	44	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
30-40	37	37	37	37	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
40-50	110	90	153	126	0	20	0	27	**1.000**	0.818	**1.000**	0.769	**1.000**	0.793	**1.000**	0.821
50-60	17	17	29	28	0	0	0	1	**1.000**	1.000	**1.000**	0.944	**1.000**	0.971	**1.000**	0.978
60-70	6	7	3	3	0	0	1	0	**1.000**	1.000	0.857	1.000	0.923	1.000	0.900	1.000
70-80	2	2	3	3	0	0	0	0	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000	**1.000**	1.000
$\geq$80	158	132	140	112	0	26	0	28	**1.000**	0.835	**1.000**	0.825	**1.000**	0.830	**1.000**	0.819
Total	3733	3423	4400	4096	244	454	178	582	**0.939**	0.883	**0.954**	0.855	**0.947**	0.869	**0.951**	0.879

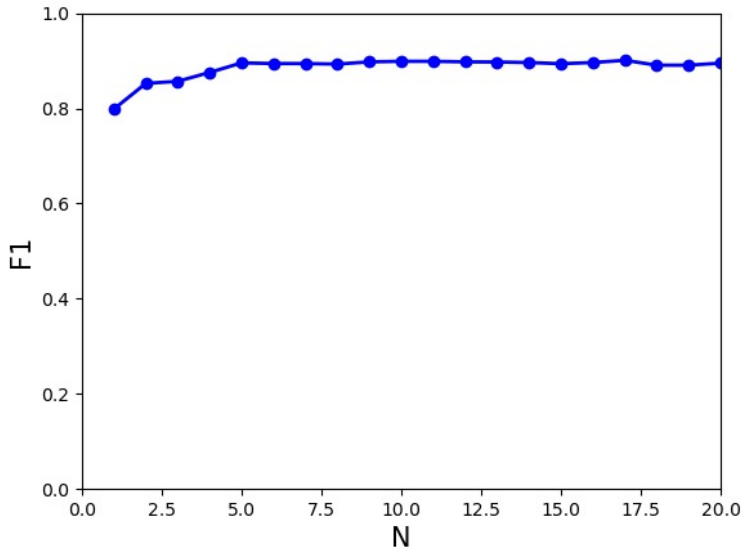

Fig. 1: **F1 of Different N**