Table 1: BFA Results - Fingerprints with printable ASCII characters

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	189,732	204,795	190,055	177,887	204,728	193,352	195,608	195,608	195,656	195,653
pdf	27.9	52.3	0.0	20.3	48.1	0.2	35.3	40.7	46.5	44.1
zip	20.2	26.6	0.0	13.3	28.0	0.1	24.9	29.2	24.7	28.2
text	21.3	4.9	98.0	50.4	4.4	95.5	14.1	6.0	7.1	7.2
doc	14.4	4.2	0.5	7.1	5.2	0.2	9.7	7.9	8.7	5.8
mp4	1.7	0.6	0.0	0.2	0.8	0.0	0.4	0.5	0.4	0.5
xls	1.2	0.0	1.4	0.8	0.1	3.9	1.0	0.2	0.0	0.1
ppt	3.2	2.2	0.0	1.8	2.7	0.0	3.3	3.3	2.7	2.9
jpg	0.5	0.1	0.0	0.1	0.0	0.0	0.1	0.1	0.0	0.1
ogg	2.8	2.2	0.0	1.4	3.0	0.0	2.8	3.0	2.7	2.7
png	6.8	6.9	0.0	4.6	7.7	0.0	8.3	9.1	7.2	8.5
Unclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Total = 34%
Perc = 33

Table 2: BFA Results - All bytes Ashims

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
pdf	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
zip	33.6	86.0	1.9	17.9	22.0	0.0	48.1	33.5	6.7	62.8
text	15.7	0.1	96.2	47.7	4.7	43	5.5	1.1	10.4	2.3
doc	2.1	0	0	0.5	0.6	0	0.4	0.1	8.2	0.3
mp4	10.1	4.5	0.4	4.1	27.2	0	12.3	25.2	18.2	11.4
xls	11.4	0.3	0.3	17.9	0.2	56.8	10.9	4.4	6.4	1.8
ppt	0	0	0	0	0	0	0	0	0	0
jpg	2.6	1.3	0.2	2	0.2	0	4.6	9.7	3.4	1.9
ogg	20.6	3	0.2	6.5	39.7	0	10.9	16.3	40.2	6.4
png	4.1	4.5	0.4	2.8	5	0	6.8	9.4	6.2	12.8
Unclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Total = 34%
Perc = 33

Table 3: BFA Results - Dominant Fingerprints

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	189,732	204,795	190,055	177,887	204,728	193,352	195,289	195,608	195,656	195,653
pdf	5.0	3.9	0	2.9	4.9	0	5.3	5.4	4.8	5.1
zip	20.4	26.8	0	13.4	28.2	0.1	25.1	29.5	24.9	28.4
text	27.9	6.8	98.4	51.9	6.4	81.7	17.3	8.6	10.6	9.0
doc	31.4	51.8	0.1	22.1	47.4	0.2	37.5	42.0	47.8	44.6
mp4	3.0	1.9	0	0.9	2.8	0	1.6	1.7	1.4	1.9
xls	1.8	0.3	1.5	2.6	0.4	17.8	1.8	0.4	0.4	0.3
ppt	6.7	6.5	0	4.7	7.2	0	8.5	9.2	7.5	8.1
jpg	1	0.3	0	0.3	0.3	0	0.5	0.6	0.3	0.4
ogg	2.2	1.5	0	1	2.1	0	1.9	2.1	1.9	1.9
png	0.7	0.3	0	0.2	0.4	0	0.5	0.5	0.4	0.4
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 4: BFA Results with Full Fingerprints to 0-75% Printable ASCII Threshold

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	165,840	204,795	1,491	157,196	204,728	192,044	192,236	194,582	195,656	195,651
pdf	31.5	52.3	3.5	22.9	48.1	0.2	35.9	40.9	46.5	44.1
zip	21.6	26.6	2.7	15.0	28.0	0.1	25.2	29.4	24.7	28.2
text	15.2	4.9	26.4	44.1	4.4	95.5	13.1	5.5	7.1	7.2
doc	16.0	4.2	59.6	7.9	5.2	0.2	9.7	7.9	8.7	5.8
mp4	0.6	0.6	0.1	0.3	0.8	0	0.4	0.5	0.4	0.5
xls	1.2	0	5.0	0.8	0.1	3.9	0.8	0.2	0	0.1
ppt	3.5	2.2	1.1	2.1	2.7	0	3.4	3.3	2.7	2.9
jpg	0.1	0.1	0.1	0.1	0	0	0.1	0.1	0	0.1
ogg	2.8	2.2	0.7	1.6	3.0	0	2.8	3.0	2.7	2.7
png	7.5	6.9	0.8	5.2	7.7	0	8.5	9.2	7.2	8.5
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 5: Training Set Analysis

ratio	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
0 - 25%	9,327	347	235	528,661	3,130	1,054,503	114,968	7,842	785	11,875
25 - 50%	1,332,849	1,680,052	436	768,686	1,585,760	576,755	1,547,585	1,685,320	1,684,877	1,674,301
50 - 75%	86,583	370	181	8,834	18	31,595	10,106	1,305	287	1,787
75 - 100%	265,275	2	1,621,682	161,133	0	21,521	10,785	4,410	5	4,850
Total:	1,694,034	1,680,771	1,622,534	1,467,314	1,588,908	1,684,374	1,683,444	1,698,877	1,685,954	1,692,813
0 - 25%	0.55	0.02	0.01	36.03	0.20	62.61	6.83	0.46	0.05	0.70
25 - 50%	78.68	99.96	0.03	52.39	99.80	34.24	91.93	99.20	99.94	98.91
50 - 75%	5.11	0.02	0.01	0.60	0	1.88	0.60	0.08	0.02	0.11
75 - 100%	15.66	0	99.95	10.98	0	1.28	0.64	0.26	0	0.29

Table 6: Sample table

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments										
pdf										
zip										
text										
doc										
mp4										
xls										
ppt										
jpg										
ogg										
png										
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 7: BFA - Fingerprints Trained in 75-100% and tested in 75-100%

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	23,892	0	188,564	20,691	0	1,308	3,053	1,026	0	2
pdf	7.6	0	0.3	0.3	0	0	0.5	0	0	0
zip	0.7	0	0.4	0.5	0	3.7	5.9	1.2	0	0
text	11.8	0	1.4	3.4	0	6.2	2.3	1.8	0	0
doc	2	0	8.2	43.2	0	17.7	5.3	1.4	0	0
mp4	49.3	0	86.5	48.6	0	68.3	78.0	74.4	0	0
xls	7.9	0	0.6	1.2	0	0.9	2.9	0.1	0	0
ppt	0.8	0	0.7	0.1	0	0	0.3	0	0	100
jpg	4.2	0	1.4	1.7	0	1.3	0.8	20.6	0	0
ogg	4.3	0	0.4	0.9	0	1.8	3.7	0.7	0	0
png	11.4	0	0	0	0	0	0.4	0	0	0
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 8: BFA - Fingerprints Trained in 50-75% and tested in 50-75%

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	12,421	43	1,203	2,101	15	3,158	2,393	147	66	89
pdf	39.1	23.3	6.2	1.8	0	1.6	1.6	2	3	1.1
zip	4.8	16.3	6.7	10.4	0	0.4	3.1	5.4	1.5	14.6
text	0.6	2.3	1.6	5.9	0	0	2.3	2	0	9
doc	6.2	7	40.9	7.5	0	2.4	18.2	4.1	3	1.1
mp4	12.2	27.9	1.2	37.6	100	27.2	42.6	40.8	36.4	12.4
xls	13.5	0	1.4	19.6	0	65.5	18.7	35.4	15.2	1.1
ppt	16.0	0	17.5	1.4	0	1.5	0.6	0.7	1.5	0
jpg	5.3	0	15.1	1.2	0	1.2	7	1.4	3	3.4
ogg	0.6	0	8.8	3.7	0	0.2	4.9	0	36.4	0
png	1.5	23.3	0.5	10.9	0	0	0.9	8.2	0	57.3
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 9: BFA - Fingerprints Trained in 25-50% and tested in 25-50%

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	147,705	204,662	285	102,831	201,859	41,013	178,816	193,103	195,368	187,688
pdf	6.9	4	4.9	5.5	5.1	0.1	6	5.8	5.2	5.3
zip	25.2	26.7	14	23.7	28.4	0.6	27.6	30	25.3	29.5
text	32.6	40.1	14.7	30.9	38.8	0.8	33.4	34.7	36.5	37.3
doc	16.3	17.7	4.9	15.6	15.8	0.4	14.8	14.4	19.4	15.1
mp4	2	0.8	2.1	1.1	1.7	0	1.2	1.1	1.1	1.1
xls	3.9	1.7	49.1	11.5	0	97.9	4.2	0.8	1.3	0.4
ppt	9.2	6.7	7.7	8.8	7.4	0.2	9.5	9.8	8.1	8.6
jpg	0.7	0.3	0.7	0.5	0.2	0	0.6	0.6	0.4	0.4
ogg	2.6	1.5	1.8	2	2.2	0.1	2.2	2.2	2.2	2
png	0.6	0.3	0	0.5	0.4	0	0.6	0.5	0.5	0.5
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 10: BFA - Fingerprints Trained in 0-25% and tested in 0-25%

	pdf	zip	text	doc	mp4	xls	ppt	jpg	ogg	png
num.of fragments	5,714	90	3	52,264	2,854	147,873	11,027	1,332	222	7,874
pdf	0	0	0	0	0	0.3	0	0.1	0	0
zip	0	0	0	0	0	0	0	0	0.5	0
text	0	0	0	0.1	0	0.7	0	0	0	0
doc	0	0	0	0	0	0	0	0.1	0	0
mp4	0	0	0	0	0.1	0.1	0	0	0	0
xls	99.6	95.6	100	99.6	99.9	97.3	98.3	95.3	96.3	99.9
ppt	0	0	0	0	0	0	0	0	0	0
jpg	0.3	4.4	0	0.2	0	0.9	1.6	4.5	2.7	0.1
ogg	0	0	0	0	0	0.2	0	0.1	0	0
png	0	0	0	0	0	0.4	0		0.5	0
Unclassified	0	0	0	0	0	0	0	0	0	0

Table 11: Sample table

	Training Set		Testing Set	
num.of fragments				
pdf				
zip				
text				
doc				
mp4				
xls				
ppt				
jpg				
ogg				
png				