Family Matters: on the Investigation of [Malicious] Mobile Apps Clustering

Thalita Scharr Rodrigues Pimenta[1][0000−0003−0017−4644], Rafael Duarte Coelho dos Santos[3][0000−0002−8313−6688], and André Grégio[2][0000−0003−1766−5757]

[1] Instituto Federal do Parana, Irati, Brazil
thalita.pimenta@ifpr.edu.br
[2] INPE, São José dos Campos,SP, Brazil
rafael.santos@inpe.br
[3] Universidade Federal do Paraná, Curitiba, Brazil
gregio@inf.ufpr.br

Abstract. As in the classification of biological entities, malicious software may be grouped into families according to their features and similarity levels. Lineage identification techniques can speed up the mitigation of malware attacks and the development of antimalware solutions by aiding in the discovery of previously unknown samples. The goal of this work is to investigate how the use of hierarchical clustering on malware statically extracted features can help on explaining the distribution of applications into specific groups. To do so, we collected 76 samples of several versions from popular, legitimate mobile applications and 111 malicious malicious applications from 11 well-known scareware families, produced their dendograms, and discussed the outcomes. Our results show that the proposed apporach is promising for the verification of relationships found between samples and their attributes.

Keywords: Clustering · Mobile Malware · Lineage

1 Appendix

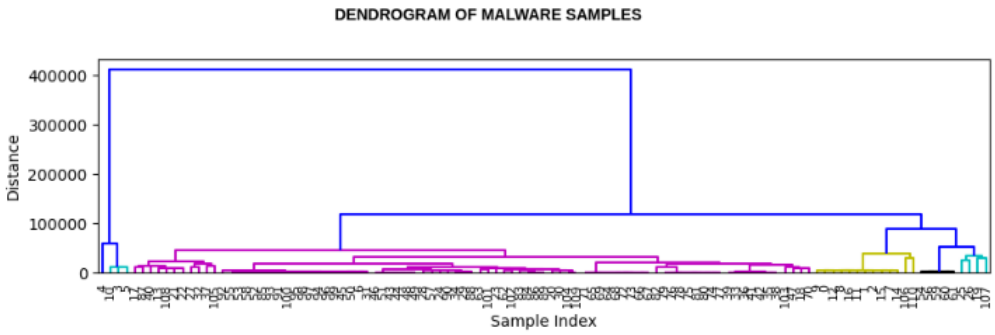

Fig. 1. Dendrogram of malware samples.

Table 1. Features of Instagram, Youtube, Netflix and Tik Tok samples

Id	Name	Permissions	Activities	Receivers	Classes	Services	Size
0	Instagram.v170.0	32	122	35	330	71	43.47 Mb
1	Instagram.v170. 2	32	122	35	813	71	43.46 Mb
2	Instagram.v171.0	32	122	35	818	71	43.67 Mb
3	Instagram.v172.0	32	122	35	822	71	43.80 Mb
4	Instagram.v173.0	32	122	43	810	74	44 Mb
5	Instagram.v174.0	33	123	43	810	74	33.56 Mb
6	YouTube.v16.01.34	33	59	25	256	18	96.94 Mb
7	YouTube.v16.02.35	33	59	25	256	18	96.68 Mb
8	YouTube.v16.04.34	33	49	25	260	18	94.72 Mb
9	YouTube.v16.04.36	33	49	25	260	18	94.72 Mb
10	Netflix v7.82.1	16	89	0	157	0	57.60 Mb
11	Netflix v7.82.2	16	88	0	157	0	57.64 Mb
12	Netflix v7.83.0	16	90	0	156	0	57.83 Mb
13	Netflix v7.84.1	16	91	0	156	0	58.22 Mb
14	Netflix v7.86.1	16	90	0	157	0	58.38 Mb
15	Netflix v7.87.2	16	90	0	157	0	59.57 Mb
16	Netflixv7.89.0	16	90	0	162	0	60.56 Mb

References

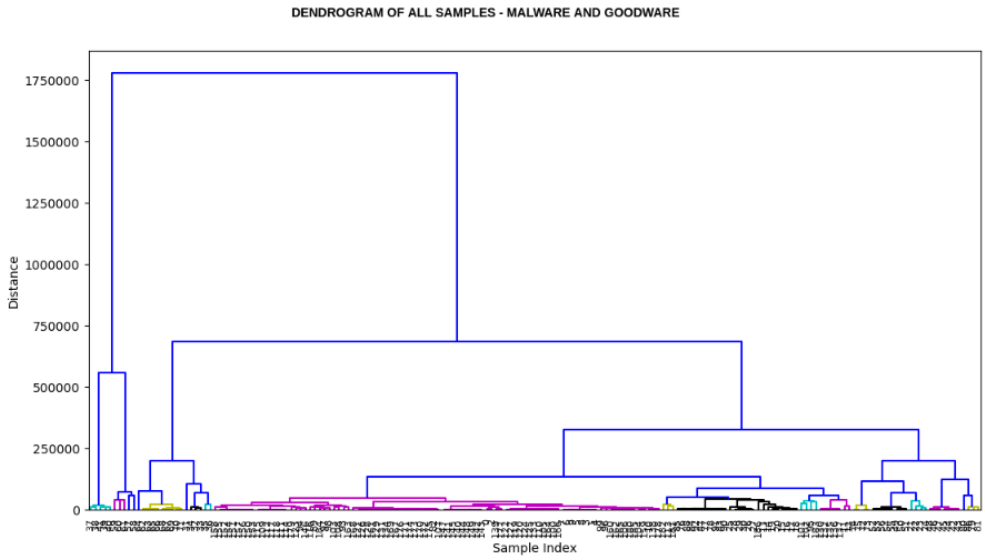

Fig. 2. Dendrogram of all samples.

Table 2. Features of Google Meet, Messenger Text and Video, Telegram and Whats App samples.

Id	Name	Permissions	Activities	Receivers	Classes	Services	Size
17	Tik Tok v18.3.6	61	342	31	1241	68	89.39 Mb
18	Tik Tok v18.4.41	61	345	31	1238	68	90.14 Mb
19	Tik Tok v18.5.3	61	360	30	1242	66	91.54 Mb
20	Tik Tok v18.5.5	61	360	30	1242	68	91.55 Mb
21	Tik Tok v18.5.6	61	360	30	1242	68	91.55 Mb
22	GoogleMeetv2020.06.42.0.3	16	20	15	84	11	13.50Mb
23	GoogleMeetv2020.07.42.5.3	16	20	15	84	11	13.56Mb
24	GoogleMeetv2020.07.43.5.3	16	20	15	81	11	13.26Mb
25	GoogleMeetv2020.07.44.0.3	16	20	16	81	12	13.25Mb
26	GoogleMeetv2020.07.44.5.3	16	20	16	79	12	13.50Mb
27	GoogleMeetv2020.10.04	16	34	17	94	14	13.25Mb
28	GoogleMeetv2020.11.01	17	35	18	113	14	17.85Mb
29	GoogleMeetv2020.11.29	17	36	19	117	12	18.22Mb
30	GoogleMeetv2020.11.45.0.3	16	20	16	79	12	13.70Mb
31	GoogleMeetv2021.01.11	17	38	19	117	12	18.51Mb
32	Telegram v7.1.3	56	13	20	138	22	40.106.000
33	Telegram v7.2.0	56	13	20	139	22	40.500.900
34	Telegram v7.2.1	56	13	20	139	22	40.542.500
35	Telegram v7.3.0	56	13	21	142	22	41.839.000
36	Telegram v7.3.1	56	13	21	142	22	41.844.200
37	Telegram v7.4.0	56	13	21	145	23	42.316.400
38	Telegram v7.4.1	56	13	21	145	23	42.316.800
39	Telegram v7.4.2	56	13	21	145	23	42.332.400

Table 3. Among Us, Garena Free Fire, Whats App, and Messenger samples.

Id	Name	Permissions	Activities	Receivers	Classes	Services	Size
40	Among Us v2020.10.22	5	8	0	37	0	75Mb
41	Among Us v2020.2.17	5	4	0	25	0	69.82Mb
42	Among Us v2020.6.9	6	4	2	28	2	72.61Mb
43	Among Us v2020.3.29	5	4	0	25	0	69.83Mb
44	Among Us v2020.4.2	5	4	0	25	0	69.85Mb
45	Among Us v2020.5.9	5	4	0	25	0	70.33Mb
46	Among Us v2020.6.9	5	4	0	25	0	70.36Mb
47	Among Us v2020.8.12	5	4	0	25	0	70.34Mb
48	Among Us v2020.8.31	5	4	0	25	0	69.57Mb
49	Among Us v2020.9.9	5	4	0	25	0	69.63Mb
50	Among Us v2021.3.9	5	4	0	25	0	69.63Mb
51	GarenaFreeFireTheCobra v1.51.2	26	21	9	123	14	46.48Mb
52	GarenaFreeFireTheCobra v1.52	26	21	9	123	14	46.46Mb
53	GarenaFreeFireTheCobra v1.53.2	26	16	9	100	12	45.10Mb
54	GarenaFreeFireTheCobra v1.54.1	23	15	7	106	11	45.98Mb
55	GarenaFreeFireTheCobra v1.56.1	23	15	7	102	11	46.34Mb
56	GarenaFreeFireTheCobra v1.57	23	15	7	102	11	46.33Mb
57	GarenaFreeFireTheCobra v1.58	23	15	7	101	11	43.89Mb
58	GarenaFreeFireTheCobra v1.58.3	23	15	7	101	11	47.23Mb
59	GarenaFreeFireTheCobra v1.59.1	23	16	7	103	11	46.19Mb
60	GarenaFreeFireTheCobra v1.59.5	23	16	7	103	11	42.93Mb
61	WhatsAppMessenger.v2.20.205.13	56	233	29	391	40	32.560.800
62	WhatsAppMessenger.v2.20.205.17	56	233	29	389	40	42.314.300
63	WhatsAppMessenger.v2.20.206.16	56	232	29	394	40	42.314.300
64	WhatsAppMessenger.v2.20.206.19	56	232	29	392	40	40.987.200
65	WhatsAppMessenger.v2.20.206.24	56	232	29	390	40	31.188.600
66	WhatsAppMessenger.v2.21.1.10	57	227	29	386	40	41.472.500
67	WhatsAppMessenger.v2.21.1.13	57	227	29	387	40	31.608.400
68	WhatsAppMessenger.v2.21.1.16	57	227	29	389	40	31.604.300
69	WhatsAppMessenger.v2.21.2.15	57	228	29	391	40	41.709.200
70	WhatsAppMessenger.v2.21.2.18	57	228	29	390	40	31.919.000
71	Messenger v295	58	334	65	810	103	48.996.700
72	Messenger v296	58	335	65	796	103	48.832.400
73	Messenger v297	58	336	65	803	104	47.869.600
74	Messenger v298	58	337	65	815	104	48.635.400
75	Messenger v299	58	336	65	796	104	48.683.300

Table 4. Features of Android spy 277, AVForAndroid, and Android Defender samples.

Id	Hash	Permissions	Activities	Receivers	Classes	Services	Family	Size
0	2c5f158e2be5b0a67fe7378d6cff0d2d	23	16	6	58	7	Android.spy.277	45.05 MB
1	2e34a05a8dbd10579ec745872ebdecad	13	10	6	48	5	Android.spy.277	16.08 MB
2	7fcfb3327f7593c3b579c1d5e45d44dd	30	56	15	108	12	Android.spy.277	12.60 MB
3	a73f6b0e4b6497757208e5d302554fee	11	5	5	14	3	Android.spy.277	3.30 MB
4	aa5fe233d1e080f3ebe4b1968fa53fda	18	11	7	68	5	Android.spy.277	4.43 MB
5	b92b038965bdcd9ae3b91679f7cd3ab6	11	16	4	55	2	Android.spy.277	26.20 MB
6	00cf4e209bb1719646f5309e0b5c1583	40	24	9	175	13	Android Defender	6.59 MB
7	00fbcc473c451ac5baa52246a7aed0ce	40	24	9	175	13	Android Defender	6.77 MB
8	03d4af908c194c4492e57b986a52d9b9	40	24	9	175	13	Android Defender	6.61 MB
9	08026e2b63ec51cb36bc6cff00c28909	79	98	31	380	38	Android Defender	31.72 MB
10	090f717dec14c0198e6c235accee7cd0	87	95	31	285	38	Android Defender	9.56 MB
11	094f67a3a682a0cd4305d720cc786e00	79	98	31	378	38	Android Defender	31.74MB
12	0c029418165cc8a7fbb25de0f37b720d	6	1	1	2	1	Android Defender	67.06 KB
13	0dc29a73d74be55df4c84e26ee31b716	40	24	9	175	13	Android Defender	7.07 MB
14	11517a3faa093728a24a3b7f044bf9c2	40	24	9	175	13	Android Defender	6.34 MB
15	11521e42f4ed27605dfab6aab6d7f06e	40	24	9	175	13	Android Defender	6.51 MB
16	14292932679d6930f521a21de4e8bffd	79	100	31	378	38	Android Defender	31.33MB
17	161a51d8e66dfd5b0d7a290e0eb29036	40	24	9	175	13	Android Defender	6.70 MB
18	168aa2734efba288e4025275ab1dec3c	40	24	8	175	13	Android Defender	6.57 MB
19	199cf86600c9e7a518ac9f1795b307c7	16	14	8	78	0	Android Defender	30.25 MB
20	19ae2cd63c8ea0b6e5cb68b4af96ca9e	40	24	9	175	13	Android Defender	6.74 MB
22	1f0f79475e428c84aa26e51fef472f3b	40	24	9	175	13	Android Defender	6.43 MB
22	2299b0b039c1bc23ed0dc9abe0227435	40	24	9	175	13	Android Defender	69.93 MB
23	028733be662b4c4b8b7f2676c2987411	12	5	1	42	1	AvForAndroid	5.79 MB
24	06dddec91d1efd55363aaab40b3ff2aa	11	5	0	11	1	AvForAndroid	2.76 MB
25	070194b27034d592f62e98887489e00b	27	91	10	216	11	AvForAndroid	25.80MB
26	0e06bbfdeb4a7c4bcf7f0a6fcfc6fa38	29	97	7	174	9	AvForAndroid	19.26 MB
27	1021f51f46879b802d3d40803dcd2fc5	12	38	3	81	4	AvForAndroid	3.67 MB
28	17c290bc50f76292ebb10586babb8ac0	16	5	0	20	0	AvForAndroid	11.26 MB
29	18f0a4edf574a68346244b0e02cf8b88	12	5	1	11	1	AvForAndroid	3.29 MB
30	1d2bd3908ab93ac56ba69019791e1a63	12	6	3	13	4	AvForAndroid	1.56 MB
31	309fecc4cb7e80023ef83fd9ba75cfbb	3	1	0	6	0	AvForAndroid	44.43 MB
32	3455aff554ef42c1fc41e4bcf6acebb7	12	43	9	51	4	AvForAndroid	7 MB

Table 5. Features of AVPass, FakeApp, FakeApp AL, and FakeAV samples.

Id	Hash	Permissions	Activities	Receivers	Classes	Services	Family	Size
33	02e38abd7ed8714ba37e30e21a334648	20	15	2	76	2	AvPass	2.30 MB
34	06deb2b73527f730b68b126519fa4c55	18	1	1	18	1	AvPass	129.36 KB
35	97d44ba157aece681c07a3d4f650d49	20	15	2	88	2	AvPass	2.23 MB
36	09c6697dbd0e5f0dc598e3d2bcd8c0b1	20	15	2	76	2	AvPass	2.10 MB
37	0d2d1fb3dc67717216f542bca0156f3d	21	47	5	50	5	AvPass	2.80 MB
38	0e3cc4510cbe53840b471bf140b83eef	20	15	2	88	2	AvPass	2.23 MB
39	0fb48dda4462519bb762930c82248b56	20	15	2	76	2	AvPass	2.38 MB
40	10261625aa8ca539ecd2a870c8bbc086	17	33	4	51	3	AvPass	2.47 MB
41	10875400d5a9201bf8a430f1fa247060	20	15	2	76	2	AvPass	2.10 MB
42	10c17fcef5c4d18a9e933d0c4d0f6d23	20	15	2	76	2	AvPass	2.10 MB
43	01cd469e80f79e7237a6d4fd8bc17c44	5	2	3	20	1	FakeApp	2.62 MB
44	0318653999a8096d2babea0dc665a17e	5	2	3	20	1	FakeApp	2.57 MB
45	05ab6089d7b8e474881c7e5aa5ed9b06	5	1	0	1	0	FakeApp	48.03 MB
46	05b9c3619fe68968d55bfcbad21f0b23	5	2	3	20	1	FakeApp	2.57 MB
47	0b23d608afc90c8acd0b7f8d807ede2b	39	5	0	16	3	FakeApp	1 MB
48	0fc9234b6bdb33aea39e3957fa582f62	5	2	3	20	1	FakeApp	11.26 MB
49	101ef95c05043fc95e69514d221f704e	5	2	3	20	1	FakeApp	3.29 MB
50	11d389146b923895f6867ef403d1af4b	6	1	0	1	0	FakeApp	43.83 KB
51	2303e8a4d1bee66bcdcff31770d1109c	5	2	3	20	1	FakeApp	2.58 MB
52	2375950e3580a3c4a1ba7b79a3448454	2	1	0	1	0	FakeApp	4.92 MB
53	0a573eb7a13964cbd94a70a90406fb1e	0	1	0	6	0	FakeApp_AL	1.27 MB
54	0a58fd1542cdd4d84550ec5deefed2b6	25	11	18	137	28	FakeApp_AL	2.40 MB
55	0c3d8620d1d813e7d3d015682736659f	0	1	0	6	0	FakeApp_AL	1.31 MB
56	16848ed83702e6fbecb0abf38b63771c	25	11	18	137	28	FakeApp_AL	2.38 MB
57	18ba00b2666aa67497425c4e8bc905f2	15	7	6	2	2	FakeApp_AL	477.76 KB
58	19960ed01a8d5769eedea76db6d14f40	0	1	0	6	0	FakeApp_AL	1.31 MB
59	1e071e3042dceca4a0fa92f3e89c998c	25	11	18	137	28	FakeApp_AL	2.41 MB
60	209d9a746c16e94e695891c1cdca8312	25	11	18	137	28	FakeApp_AL	2.39 MB
61	2186f65199f8c6be0d27c40a65637d70	25	11	18	137	28	FakeApp_AL	2.38 MB
62	21d12ef6cd4f247a91880c44dd7449a6	11	23	0	11	0	FakeApp_AL	878.62 KB
63	26676538cbdedc1285699d9af0dcf06b	30	1	1	31	1	FakeApp_AL	275.53 KB
64	002e34c2b615c13fe21498013c5daa16	29	5	4	3	7	FakeAV	546.22 KB
65	00eb3d75ca5624af970482bafc3d6a73	29	5	4	3	7	FakeAV	284 KB
66	0102f20a81eb1cf3c86abd49a2f4c06b	29	5	4	3	7	FakeAV	1.14 MB
67	010474247ebfd78a765cfc5b7db550e1	29	5	4	3	7	FakeAV	268.13 KB
68	0125b812c1a948e0cb39d22f4b5615bf	29	5	4	3	7	FakeAV	1 MB
69	020241a8933c0798d5537180c84085e79	29	5	4	3	7	FakeAV	547.30 KB
70	02b222e277c8a44e2fe67e4e734571c0	33	2	6	27	7	FakeAV	479.35 KB
71	02d23d111dbf8812c635d4f0eadc75b8	29	5	4	3	7	FakeAV	275.53 KB
72	0306fbe880a11761c775589ed9ff7381	29	5	4	3	7	FakeAV	283.15 KB
73	030a142b6c1913dabd1d2186ba2a0e48	29	5	4	3	7	FakeAV	276.43 KB

Table 6. Features of FakeJobOffer, Faketaobao, Penetho, and Virus Shield samples.

Id	Hash	Permissions	Activities	Receivers	Classes	Services	Family	Size
74	09714aac017c9acc96d270eb9bad9dc1	11	24	4	8	3	FakeJobOffer	874.91 KB
75	39d84c6ac4837d210690f65e60a27378	11	24	4	8	3	FakeJobOffer	883.41 KB
76	40aeb22aba6dca5eb926b0ba4834177d	11	24	4	8	3	FakeJobOffer	877.34 KB
77	57ebf71043609ba4e2d42f4c08f8a5c5	11	24	4	8	3	FakeJobOffer	873.5 KB
78	6009c8c4c66771c67357283b7d665dec	11	24	4	8	3	FakeJobOffer	1 MB
79	9e8fa23dfc817bdcad42b2f6ada6e658	11	24	4	8	3	FakeJobOffer	950.49 KB
80	da8f8d68f6ded154378b25d82234d8a7	11	24	4	8	3	FakeJobOffer	854.40 KB
81	fd4b9f97d3534c009ccd95858032eb51	11	24	4	8	3	FakeJobOffer	854.40 KB
82	fe7a44137739c8543d066ec0c2aa15b9	11	24	4	8	3	FakeJobOffer	860 KB
83	18565907b3a04006b6bcd3a7eb402dc3	21	2	3	6	1	FaketaoBao	33.47 KB
84	1c0c1837e99107f137c47e87570be00f	21	3	4	6	2	FaketaoBao	334 KB
85	45dae1ee4ca1980c140cb5c9da2a7ed5	3	3	0	2	0	FaketaoBao	454.7 KB
86	5376d63e7c498b52548e1c487d972a9c	21	3	4	6	2	FaketaoBao	344 KB
87	8be7ac1e01b3a5db141031872324d4f75	14	33	2	22	2	FaketaoBao	1.23 MB
88	c57194d05a30d53c764983c70e471791	10	1	4	5	2	FaketaoBao	143.37 KB
89	e97be7f2450dd363d2d53711dbf4c4ae	21	3	4	6	2	FaketaoBao	327.97 KB
90	eda9098498a6201383e311f9b3757b4f	6	1	4	0	2	FaketaoBao	187.96 KB
91	7f521bbef670ab04d27b30ac96832734	5	3	0	2	0	Penetho	155.84 KB
92	b92765d38cd1c6f49dfc2da5d857efc2	9	4	0	4	0	Penetho	510.73 KB
93	c1191c7efc4b1fb674694f9920076da5	5	3	0	2	0	Penetho	95.85 KB
94	c16a2685052edbb2bfcca7ea12f580cb	5	3	0	2	0	Penetho	118.33 KB
95	d9626d54c63f8ce4c32e9061b62009d0	5	3	0	2	0	Penetho	142.55 KB
96	dabf1f1c058ef3f95ba497fef4e9195f	5	3	0	2	0	Penetho	116.98 KB
97	dbe952dd2becf9725f84780947582fb2	5	3	0	2	0	Penetho	117.91 KB
98	eb7c0cf99ebc58e6a05391cc2319c89a	5	3	0	2	0	Penetho	142.70 KB
99	ed323f9a664c535afc7fb75aabecb9c6	5	3	0	2	0	Penetho	140.97 KB
100	f3d11149cf49e48c74bee08da507469a	5	3	0	2	0	Penetho	146.74 KB
101	002485c5c96f0681a4eccee5b69c5f50	22	3	6	1	1	Virus Shield	294.21 KB
102	00c921f069c02d88c96900c212db4b39	18	7	1	4	4	Virus Shield	584.46 KB
103	16fb420aa291325e8dfd72c55d8c832	16	33	5	19	2	Virus Shield	8.90 MB
104	42577326f0407873d900bb65e401891	13	6	2	11	5	Virus Shield	8 MB
105	04cda5773c1bfa9a5387c9e5b8bfefac	32	11	6	45	10	Virus Shield	7.29 MB
106	6ea4781b1ecba14454a234a3361daa0	31	11	10	51	14	Virus Shield	7.82 MB
107	0f050fb655e4624925d9d32582805601	30	27	18	35	43	Virus Shield	9.94 MB
108	119658211908d17d4f78f36168477af1	22	19	2	79	4	Virus Shield	3.47 MB
109	13bca905ccc119771040fd5fd30afff1	13	5	2	11	5	Virus Shield	7.74 MB
110	15ca5ad27034f5b9c516afc94511adba	33	7	14	78	18	Virus Shield	7.98 MB

Table 7. Calculated indices(%) for malware clustering results

Indexes	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Cluster 6**	**Cluster 7**	**Cluster 8**	**Cluster 9**	**Cluster 10**
Precision	100	100	0	88	100	100	100	100	100	100
Recall	100	100	0	88	100	100	100	100	100	61
F1-Score	100	100	0	88	100	100	100	100	100	75
Indexes	**Cluster 11**	**Cluster 12**	**Cluster 13**	**Cluster 14**	**Cluster 15**	**Cluster 16**	**Cluster 17**	**Cluster 18**	**Cluster 19**	**Cluster 20**
Precision	100	100	0	100	100	100	0	100	100	33
Recall	100	100	0	100	60	100	0	100	63	20
F1-Score	100	100	0	100	75	100	0	100	77	25
Indexes	**Cluster 21**	**Cluster 22**	**Cluster 23**	**Cluster 24**	**Cluster 25**	**Cluster 26**	**Cluster 27**	**Cluster 28**	**Cluster 29**	**Cluster 30**
Precision	0	75	6	100	0	0	0	0	0	0
Recall	0	30	100	50	0	0	0	0	0	0
F1-Score	0	42	12	66	0	0	0	0	0	0