

Reversing Undocumented File Formats

Using a Hex Editor and your Brain

2009.07.04

CodeEngn
3rd CodeEngn ReverseEngineering Seminar

Reverse What?

000000000	0F13	0DDA	32A2	AB23	342E	D49E	EE85	0AFA	B2BC	1404	9108	ADB4	B947	1C59	0509	A400	...2..#4.....G.Y....
000000020	92AB	62AC	2D6D	6C2E	90C4	D13B	2B07	7CA4	396B	49DA	6293	DD18	AB82	9A62	5701	AD96	..b.-m!...;+.. _9kI.b.....bW...
000000040	FB47	E985	4648	0353	6206	4041	02D8	D85E	B3F4	6CD8	323C	6368	89EE	0FD2	AB9E	957F	.G..FH.Sb.@A...^..I.2<ch.....
000000060	B92F	4F65	85B3	2747	CAB0	237E	177D	5F9C	28C6	7C9A	2601	D45E	B03D	A6E9	EAB7	E738	./0e..'G..#~.}_.(. .&..^.=.....8
000000080	73C1	A955	EA03	776D	0F49	A9E4	D8E7	A7BB	05F1	6DB9	E6A2	DAF6	88C9	5538	8329	91D4	s..U..wm.m.....U8.)..
0000000A0	4020	33B6	739E	DBFE	AA93	830F	7C14	4D56	1855	E77F	D156	40A3	8138	7192	98DE	3D1E	@ 3.s..... ..MV.U..V@..8q...=.
0000000C0	98CC	D5BD	FB3F	0A7C	D406	BE58	C2AA	9005	E51A	BE18	FA58	058D	B1ED	CA8D	AE28	E3F3? ..X.....X.....(...
0000000E0	65CC	485B	06F1	57EF	F1ED	AD97	910F	F507	AE29	F0A9	96B4	ACB3	443B	44FC	67BD	BA8C	e.H[..W.....).....D;D.g...
000000100	ABD6	E1CC	4204	7551	0322	E878	1F4B	7C9A	EE11	3197	C45A	1BB1	24D8	2462	220C	30A4	...B.uQ."..x.K ...1..Z..\$.b".O.
000000120	58BE	46FD	A46D	3C07	98A1	C980	9DCD	6437	A356	2771	DED6	2EF7	94EA	9A8A	8F7D	F664	X.F..m<.....d7.V'q.....}..d
000000140	4972	6FB7	D1BC	F8EE	C0F1	1C86	43E5	3BBE	E72C	66BE	E5E7	07D4	7A2D	A139	6C95	89EC	Iro.....C.;..f....z-91...
000000160	9965	722E	20D8	86D0	8D61	8052	1C10	7FB8	E9E4	C9A7	B240	E015	C1C7	0462	C9D7	645C	.er.....a.R.....@.....b..dt#
000000180	91E8	50CA	7EF1	3457	F17B	D141	7438	C400	1DOC	6422	3C02	AAEE	01ED	711E	F224	FB9C	.P..~.4W.{.At8...d"<.....q..\$..
0000001A0	2D73	1965	DAF0	663D	0584	BAD7	735D	6171	066A	77A0	27C6	1700	8293	62BE	2A65	5927	-s.e..f=...s]aq.jw.'.....b.*eY'
0000001C0	FF72	A340	D392	34C6	6C25	EB64	F75C	2424	7C64	C0F5	1E96	9F3A	675F	E07A	ABBf	84E5	.r@..4..1%.d.₩\$\$!d.....g_z....
0000001E0	753B	A846	74E6	AC4F	9436	EC10	AAA7	0F1F	E9CC	3A49	6BB1	F757	4B9F	219E	5F0A	A54C	u..Ft..0.6.....:Ik..WK.!_..L
000000200	71D0	EF98	534D	A5BE	8B69	9280	3E5A	9C60	A40E	969C	72Cf	BB84	E4E7	28A5	A1A0	90EB	q...SM....>Z`.....r.....(....
000000220	2599	1875	A675	9233	C33A	E7CA	9152	3657	185C	D585	24BB	42FC	5CFF	F8A5	2667	B045	%..u..u.3....R6W.W..\$.B.W...&g.E
000000240	B7A5	FAA6	5B77	6B66	026E	01FD	0E3F	ADF5	B86A	569E	93D0	A901	A931	BFE3	878D	ECBD	...[wkf.n...?....jV.....1.....
000000260	5AC2	BFD5	89B6	10B5	6B3C	9C1F	3CBD	BB30	94A1	C2BE	262F	EEA4	7A80	3CE4	2A71	20C0	Z.....K<..<..0....&/..z.<.*q ..
000000280	827C	0903	5643	09DF	3582	394D	8A6D	6952	B8A7	D32A	FEDD	3C0D	74B5	2CF8	1C8E	C0C8	. ..VC..5.9M.miR...*..<.t.....
0000002A0	65AC	83CF	BFDC	861D	93D6	EC9B	7CB9	4DBE	0391	0432	2821	ECB9	30A9	BC76	B49F	9657	e..... ..M....2(!..0..v..W
0000002C0	FE77	1445	CA17	060E	EDCE	3112	8CE3	3738	B459	D54D	F63E	D0D8	2CD7	9EE9	B3F2	F508	.w.E.....1...78.Y.M.>.....
0000002E0	1023	1926	512D	3479	70E7	5FC2	C560	7A38	9E6C	1975	E284	2C42	309A	E02C	1DF6	1189	.#.8Q-4yp._..`z8.I.u...B0.....
000000300	2379	CCA6	544D	976F	DED3	2AA8	0F5A	B14F	AF19	EF79	6D19	FE10	C294	75F0	A842	AF23	#y..TM.o..*..Z.O...ym.....u..B.#
000000320	62C1	8D2E	8031	D8EC	1877	0AEB	2BD4	C158	62BF	0661	13AF	AFF8	4DCD	DFF0	A283	87BD	b.....1....w.+..Xb..a....M.....
000000340	2A60	0F6C	90B3	A356	71A1	1602	1B8D	48AB	6F45	CF7B	B3B3	2C5B	B3E0	OB95	145C	BA26	*..I...Vq.....H.oE.{...,[....W.&
000000360	7025	A565	D9B1	2B09	1060	B414	598B	DF3F	BDAC	814C	D5C5	9B91	1301	2331	06FF	B411	p%.e..+...Y...?..L.....#1....

Why use a Hex Editor and not a Disassembler?

- } When the program that reads the File is not accessible(i.e. Firmware)
- } When the program is in your possession, but is too hard to reverse(packed, protected etc.)
- } When you don't know how to reverse executable files, but still want to reverse a file format
- } Sometimes, reversing a file itself is faster than reversing the executable file(Game Archives)



Methodology

- } Search Wotsit.org
- } Study the general structures of File Formats
- } Using a Hex Editor, try to match the structure of the data with the structures you already know
- } “Look” at the Hex Bytes and try to find patterns
- } Make Assumptions and verify
- } If the theory proves to be correct, make more assumptions based on the new facts until the whole(or part) of the file format is revealed

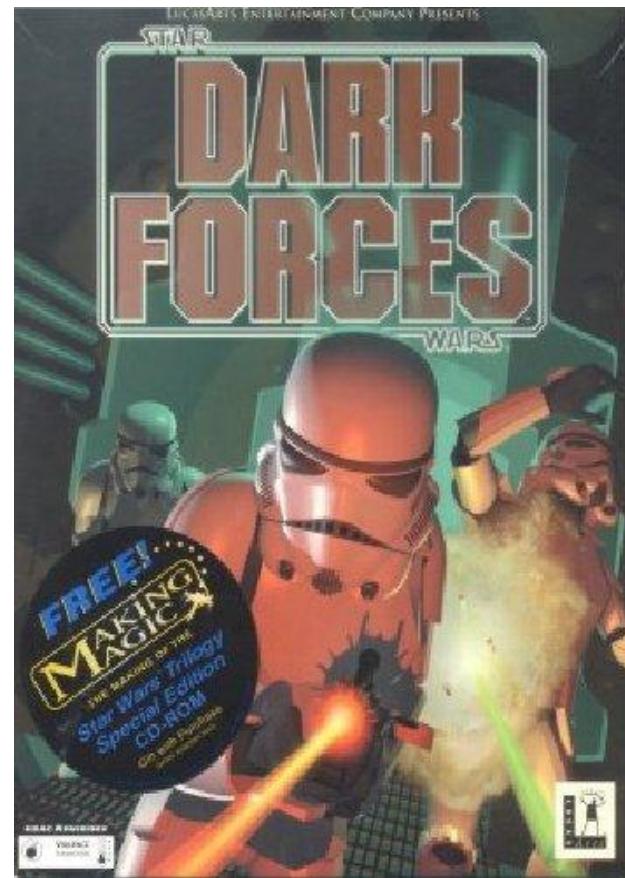
Real World Reversing

- } Case I : Cannot write but readable
- } Case II : readable/writable
- } Case III : Only 1 sample, cannot read/write

("can read/write" means the program that reads/writes to the file is accessible)



Case I : Reversing a Game Archive File Format



Things to look out for

- } GRAIS(Game Resource Archive Identity Strings)
- } Offsets
- } Size Fields
- } Number of Files
- } Filenames



Basic Facts

- } The file that contains the resources are bigger than the other files
- } Most values are stored in a 4 byte field
- } There is usually at least 1 meaningful field in the beginning of the file
- } Strings can be null-terminated, or fixed-length in which case the string is padded with nulls



*.GOB File

The screenshot shows the Hex Workshop application interface. The title bar reads "Hex Workshop - [SOUNDS.GOB]". The menu bar includes File, Edit, Disk, Options, Tools, Window, and Help. The toolbar contains various icons for file operations like Open, Save, Copy, Paste, and Find. The main window displays a grid of hex bytes and their corresponding ASCII representation. The ASCII column shows a mix of characters, numbers, and symbols, including some non-printable characters. A vertical scroll bar is visible on the right side of the main window.

Address	Hex	ASCII
00000000	474F 420A F7C9 1300 4372 6561 7469 7665 2056 6F69 6365 2046 696C 651A 1A00 0A01	GOB.....Creative Voice File.....
00000020	2911 0182 3B00 A500 8080 8080 8080 807F 7F80 807F 7E7E 7F7F 7F7F 8180 8080).....
00000040	8082 837E 7E7E 7D7B 7A79 7879 7C7E 7D7C 7E80 7E7C 7D7D 7B7A 7B79 7676 7879 7A7F	...~~~}{zxyx ~} ~ }}{z{vvxyz.
00000060	8384 8985 7E80 8584 837E 7D81 8180 8080 7F7B 7677 7B7A 7979 7C81 7D7D 7E81 807A~...~}.....{vw{zyv .}}~..z
00000080	7575 797F 827F 7C81 827D 8181 7B7E 807A 7977 7D89 8182 807F 8886 7F7E 7472 7683	uuy.... ..){~.zyw}.....~zrv.
000000A0	8670 787F 8C99 9CA0 A195 8D8F 8C80 7C75 7179 7973 747A 7875 7775 726E 737C 8282	.lx.....}{luqystzxuwurns ..
000000C0	8484 F780 8183 817D 7D7C 7674 7A89 8A78 767B 787C 8081 8382 837D 7676 6E6F 7475}{} vtz..xv{ })....}vvnotu
000000E0	8288 888B 918D 8784 7E81 8586 8283 8D8A 8581 7E7D 8688 8584 878A 918E 8282 8F93~.....~}.....
00000100	867F 7D7A 766A 6A71 6D6C 686B 7986 8F91 8D84 8783 7C7D 848B 8688 888C 867A 7973	.}zvjjqmlhky..... }.....zys
00000120	6869 717B 7880 888C 8988 7F7E 827B 7C7D 7E76 7485 7572 7D92 8879 747C 7677 7D7D	hiq{x.....~.{{ }~vz.ur}..yt vw}}
00000140	6D6E 8181 6573 8075 6F88 8280 767F 727B 8881 8186 7E74 8074 7981 727E 787E	mn..es.uo..v.r{.....tyt.py.r~x^
00000160	8972 838B 7B72 8574 8380 7176 6F7F 7563 758F 8278 8888 5D6A 7A88 7C82 7276 7E8C	.r..{r.t..avo.ucu..x..]jz. .rv^.
00000180	967B 537B 8357 607A 6A67 707F 8988 6A76 9695 776D 748F 8F82 777C 8684 6D6E 8373	.{S{.W`zigp..jv..wmt...wl..mn.s
000001A0	686E 8B71 7177 7582 9799 887D 7D97 967C 707E 8285 876B 6367 6D78 5E5F 6276 756F	hn.qawu..}}}.lp...kcgmx^_bvuo
000001C0	7B7C 6977 8889 7A75 787B 7E89 8678 6F75 736F 7570 646B 787D 816E 6F8E 938F 808E	{ iw..zux{..xousoupdkx}.no.....
000001E0	8D84 8884 7A80 8284 8371 6578 826F 6666 6865 5864 6061 6576 7986 8D96 908F A187z...aex.offheldlaevv.....
00000200	837E 6B65 8288 7A7C 899F 9995 907C 7972 5E55 545D 7680 7B95 9E9F ACA2 9D48 AA8D	.~ke..zlyr^UT]v.i.....
00000220	6669 7465 5A62 6E7D 7F75 6F70 7874 7B70 7A91 9883 8193 A0A8 9288 909E 9074 7166	fiteZbn}.uopxt{pz.....taf
00000240	6C6E 6C72 797B 7F9D 9891 9980 8987 7881 9E9A 9B9A 8D93 8686 7566 6C74 8C9D 9980	Inlry{.....x.....ufit....
00000260	6F67 5133 2732 3C43 425E 6A64 6D7B 7067 7D90 7E83 8872 8494 8577 717C 8781 7886	o@Q3'2<CB`jdml{pg}..~.. ..wq ..x.
00000280	A39A 9DAE B1A6 AEB3 B3AD 9B92 8566 575F 5B77 735A 584F 391C 170C 161B 3359 759BfW_[wsZX09.....3Yu.
000002A0	BFC6 CDC9 AF96 907C 6F72 7E7B 8487 827C 6E6F 6458 5C6C 7780 8D45 9D8D 9596 8896lor~{... nodX#Iw.....
000002C0	8D79 6E6F 5F57 6584 8686 9098 A69F 96A4 ADA4 B4BF CB8C B29B 7869 4B21 2243 625B	.yno_We.....xik!"Cb[
000002E0	5E6D 8481 908D 9899 9796 939D A095 9086 7672 7173 6162 5B4B 4664 6965 708A 968A	^m.....vrqsab[KFdie}...
00000300	BEB2 A1A2 A997 8972 6669 6C5E 595C 545C 5F67 7078 7994 9688 90A9 AB8C 908E 7843	..rfi ^YWTW_gpxy.....xC
00000320	4667 5846 5058 6050 6063 6673 8E91 9B8B 898E 7677 7F58 595E 7867 5649 6868 5B6F	FgXFPX`P`cfs.....vw.XY`xgVlhh[o
00000340	756F 818B 8E9A 827D 8492 8293 716E 8089 7666 5171 726B 6573 747A 7B81 A38F 7779	uo.....}....qn.ivfUqrkestzi...wy
00000360	7A71 6977 7268 5271 A180 7897 9984 8F92 7F67 7180 7B99 8854 5664 7970 5A52 7895	zqjwrhRq..x.....gq.{..TvdyZRx.
00000380	8989 736E AAA9 8C9A 999A 8583 867D 8680 616C 615D 788D 7473 7E9B 7B81 8F95 8EAF	..sn.....}..ala]x.ts~.i.....
000003A0	BDB6 BDA4 AA93 8285 8E87 7469 7588 776F 6771 758A 8A81 6C6A 797F 7E74 7A89 9082tiu.wogqu...ijy.^tz...uv}ha{..su}yw}....z{vy{...
000003C0	8E8E 8783 7576 7D68 617B 767D 7F73 757D 7D8E 8380 8E99 7A7B 7679 7B89 848C	..wzs]ZPPNOMOM_ifk.wu{ssqy....
000003E0	9189 777A 735D 5A50 504E 4F4D 4F4D 5F69 666B 8077 757B 7E73 7371 7681 858E 96A5	

OK... So I'm looking at a bunch of Hex Bytes... Now What???

First Things First

- } File Size = 1299033
- } File size fits in 3 bytes(0~16777216)
- } Most fields are stored as 4 bytes
- } All offset and size fields will end with a NULL byte



Identifying the Header

00000000	474F	420A	F7C9	1300	4372	6561	7469	7665	2056	6F69	6365	2046	696C	651A	1A00	0A01	GOB.Creative Voice File.....
00000020	2911	0182	3B00	A500	8080	8080	8080	807F	7F80	807F	7E7E	7F7F	7F7F	7F7F	8180	8080)~}{{zyxy ~} ~,~ }}{z{ywwwz.
00000040	8082	837E	7E7E	7D7B	7A79	7879	7C7E	7D7C	7E80	7E7C	7D7D	7B7A	7B79	7676	7879	7A7F~~~}{{zyxy ~} ~,~ }}{z{ywwwz.	
00000060	8384	8985	7E80	8584	837E	7D81	8180	8080	7F7B	7677	7B7A	7979	7C81	7D7D	7E81	807A~ ..~ }.....{vw{zyyl.}}~..z	
00000080	7575	797F	827F	7C81	827D	8181	7B7E	807A	7977	7D89	8182	807F	8886	7F7E	7A72	7683	uuy... ..}..{~,zyw}.....~zrv.	
000000A0	867C	787F	8C99	9CA0	A195	8D8F	8C80	7C75	7179	7973	747A	7875	7775	726E	737C	8282	. x..... uqvystzxuwurns ..	
000000C0	8484	7F80	8183	817D	7D7C	7674	7A89	8A78	767B	7B7C	7D81	8382	837D	7676	6E6F	7475})}{vtz..xvt{} }....}vvnotu	
000000E0	8288	888B	918D	8784	7E81	8586	8283	8D8A	8581	7E7D	8688	8584	878A	918E	8282	8F93~~ }.....~}	
00000100	867F	7D7A	766A	6A71	6D6C	686B	7986	8F91	8D84	8783	7C7D	848B	8688	888C	867A	7973	.}zvjjqmlhky,..... }.....zys	
00000120	6869	717B	7880	888C	8988	7F7E	827B	7C7D	7E76	7A85	7572	7D92	8879	747C	7677	7D7D	hiqfx.....~.{ }~vz.ur}..yt vw}}	
00000140	6D6E	8181	6573	8075	6F88	8280	767F	727B	8B81	8186	8E74	7674	9070	7981	727E	787E	mn..es.uo...v.r{....tvt.py.r~^~	

Verifying

0013C9E0	BF07	1756	BF07	1656	BF07	1556	BF07	1401	8F28	5900	FF2F	0066	0000	0008	0000	00A1	..V...V...V....(Y.../f.....
0013CA00	3B00	0044	4F4F	5232	2D31	2E56	4F43	0000	A93B	0000	8411	0000	444F	4F52	322D	322E	;..DOOR2-1.VOC.....,..DOOR2-2.
0013CA20	564F	4300	002D	4D00	008E	0C00	0044	4F4F	5232	2D33	2E56	4F43	0000	BB59	0000	A128	VOC..-M.....,DOOR2-3.VOC...Y...()
0013CA40	0000	454C	4556	322D	312E	564F	4300	005C	8200	0007	4300	0045	4C45	5632	2D32	2E56	..ELEV2-1.VOC..W...C..ELEV2-2.V
0013CA60	4F43	0000	63C5	0000	C126	0000	454C	4556	322D	332E	564F	4300	0024	EC00	00C1	5100	OC..c...&..ELEV2-3.VOC..\$...Q..
0013CA80	0044	4F4F	522E	564F	4300	0000	0000	E53D	0100	011B	0000	4C4F	434B	4544	2D31	2E56	,DOOR.VOC.....=.....LOCKED-1.V
0013CAA0	4F43	00E6	5801	0061	1800	0047	414D	4F52	2D33	2E56	4F43	0000	4771	0100	112B	0000	OC..X..a...GAMOR-3.VOC..Gq...+..
0013CAC0	5245	4559	4545	2D31	2E56	4F43	0058	9C01	0001	1900	0042	4F53	534B	2D31	2E56	4F43	REEYEE-1.VOC.X.....,BOSSK-1.VOC
0013CAE0	0000	59B5	0100	01F9	0000	4352	4541	5455	5231	2E56	4F43	005A	AE02	0061	3500	0050	..Y.....CREATUR1.VOC.Z...a5..P
0013CB00	524F	4245	2D31	2E56	4F43	0000	BBE3	0200	212A	0000	494E	5441	4C45	5254	2E56	4F43	ROBE-1.VOC.....!*..INTALERT.VOC
0013CB20	00DC	0D03	00A1	2900	0052	414E	4F46	4330	322E	564F	4300	7D87	0300	2152	0000	5241),RANFC02.VOC,}7..!R..RA
0013CB40	4E4F	4643	3034	2E56	4F43	009E	8903	00AE	4A00	0052	414E	4F46	4330	352E	564F	4300	NOF04.VOC.....,J..RANFC05.VOC.
0013CB60	4CD4	0300	2119	0000	5241	4E4F	4643	3036	2E56	4F43	006D	ED03	0021	3400	0052	414E	L...!...RANFC06.VOC.m...!4..RAN
0013CB80	5354	4F30	312E	564F	4300	8E21	0400	213F	0000	5241	4E53	544F	3032	2E56	4F43	00AF	ST001.VOC..!...!?.RANST002.VOC..
0013CBA0	6004	0021	3F00	0052	414E	5354	4F30	332E	564F	4300	D09F	0400	2152	0000	5241	4E53	`..!?.RANST003.VOC..!..!R..RANS
0013CBC0	544F	3034	2E56	4F43	00F1	F104	00E1	3700	0052	414E	5354	4F30	352E	564F	4300	D229	T004.VOC.....,7..RANST005.VOC..)
0013CBE0	0500	2122	0000	5241	4E53	544F	3036	2E56	4F43	00F3	4B05	0021	4300	0052	414E	5354	..!"..RANST006.VOC..K..!C..RANST
0013CC00	4F30	372E	564F	4300	148F	0500	212A	0000	5241	4E53	544F	3038	2E56	4F43	0035	B905	007.VOC.....!*..RANST008.VOC..5..
0013CC20	00A1	2500	0041	5845	2D31	2E56	4F43	0000	0000	D6DE	0500	211D	0000	494E	5453	5455	..%..AXE-1.VOC.....,!..INTSTU
0013CC40	4E2E	564F	4300	00F7	FB05	0021	0F00	0050	524F	4246	4952	312E	564F	4300	180B	0600	N.VOC.....!..PROBFIR1.VOC.....

Is indeed an offset J
Looks like a directory...

Guessing the Directory entry length

The screenshot shows a debugger interface with two main sections:

Memory Dump: Displays a list of memory addresses and their corresponding byte values and ASCII strings. The strings include file names like EX-TINY1.VOC, HEALTH1.VOC, YLEDIE1.VOC, SHIELD1.VOC, USH.VOC, LAND-1.VOC, M-IN.VOC, and GOGGLES1.VOC.

Address	Length
0000	EX-TINY1.VOC...KEY.VOC...
004B	...A#..HEALTH1.VOC.....V...K
0000	YLEDIE1.VOC.j^..!~..FALL.VOC...
4352SHIELD1.VOC..x...I{..CR
0000	USH.VOC.....h.....JUMP-1.VOC...
5749	.y.....LAND-1.VOC...3...AF..SWI
000B	M-IN.VOC..t.....QUARTER.VOC...
4747	..!4..GOGGLES1.VOC.,,!..!?.GOGG
150A	...L...VOC...W...O...WAV1.VOC

Find Results: A table showing the results of a search operation.

Address	Length
0013CE08	Sel: 0x15 bytes

At the bottom, there is a menu bar with items: Compare, Checksum, Find, Bookmarks, Output. The "Find" item is highlighted.

The excess data

7 2356 BF07 2256 BF07 2156 BF07 2056	..V...&U..%V..\$V..#V.."V..!V.. V
7 1B56 BF07 1A56 BF07 1956 BF07 1856	..V...V...U...V...V...V...V...V...
3 5900 FF2F 0066 0000 0008 0000 00A1	..V...V...V....(Y.../.f...
3 0000 8411 0000 444F 4F52 322D 322E	;..DOOR2-1.VOC.....DOOR2-2.
2 2D83 2E56 4F43 0000 BB59 0000 A128	VOC..-M.....DOOR2-3.VOC..Y...()
3 0007 4300 0045 4C45 5632 2D32 2E56	..ELEV2-1.VOC..#...C..ELEV2-2.V
3 332E 564F 4300 0024 EC00 00C1 5100	OC..c...&..ELEV2-3.VOC..\$.Q.
3 011B 0000 4C4F 434B 4544 2D31 2E56	.DOOR.VOC.....=.....LOCKED-1.V
3 2E56 4F43 0000 4771 0100 112B 0000	OC..X..a...GAMOR-3.VOC..Gq...+..
1 1900 0042 4F53 534B 2D31 2E56 4F43	REYEYEE-1.VOC.X.....BOSSK-1.VOC
1 2E56 4F43 005A AE02 0061 3500 0050	..Y.....CREATUR1.VOC.Z...a5..P
4 0000 494E 5441 4C45 5254 2E56 4F43	ROBE-1.VOC.....!*..INTALERT.VOC
3 564F 4300 7D37 0300 2152 0000 5241).RANOF02.VOC.)7..!R..RA
3 0052 414E 4F46 4330 352E 564F 4300	NOFC04.VOC.....J..RANOF05.VOC.
3 4F43 006D ED03 0021 3400 0052 414E	L...!.RANOF06.VOC.m...!4..RAN
3 5241 4E53 544F 3082 2E56 4F43 00AF	ST001.VOC..!..!?.RANST002.VOC..
3 4300 D09F 0400 2152 0000 5241 4E53	`..!?.RANST003.VOC.....!R..RANS
2 414E 5354 4F30 352E 564F 4300 D229	T004.VOC.....7..RANST005.VOC..)

Possibly the number of Directory entries in the
Directory?

Verifying

The screenshot shows a debugger interface with several panes:

- Assembly Pane:** Shows assembly code with addresses 0, B, C, D, E, F, and 5. The assembly code includes instructions like ADD, SUB, and CMP.
- Memory Dump Pane:** Shows memory dump data with addresses 0, B, C, D, E, F, and 5. The data includes ASCII strings such as ".VOC.....", ".ELEV3-1", ".VOC.....", ".ELEV3-2", ".VOC.....", ".ELEV3-3", ".VOC.....", ".STAR-THM.GMD.....", ".FIGHT-01.GMD.b...%a..", ".STALK-01.", ".GMD..L..p}..", ".EXECMUS.GMD..", and ".".
- Bookmarks Pane:** Shows a table with one bookmark entry:

Address	Length	Description
0013C9F7	00000000	Directory He
- Bottom Navigation Bar:** Includes tabs for Compare, Checksum, Find, Bookmarks (which is selected), Output, and a status bar showing Offset: 0013C9FB, Sel: 0x85e bytes, and a value of 1.

$0x85E / 0x15 = 0x66$
Verified! ↴

Guessing the fields in the Directory Entry

00000000	0800	0000	A13B	0000	444F	4F52	322D	312E	564F	4300	00A9	3B00	0084	1100	0044	4F4F	5232	2D32	2E56	4F43	0000;..DOOR2-1.VOC..
0000002A	2D4D	0000	8E0C	0000	444F	4F52	322D	332E	564F	4300	00BB	5900	00A1	2800	0045	4C45	5632	2D31	2E56	4F43	0000	-M.....DOOR2-3.VOC..
00000054	5C82	0000	0743	0000	454C	4556	322D	322E	564F	4300	0063	C500	00C1	2600	0045	4C45	5632	2D33	2E56	4F43	0000	W....C..ELEV2-2.VOC..
0000007E	24EC	0000	C151	0000	444F	4F52	2E56	4F43	0000	0000	00E5	3D01	0001	1B00	004C	4F43	4B45	442D	312E	564F	4300	\$....Q..DOOR.VOC.....
000000A8	E658	0100	6118	0000	4741	4D4F	522D	332E	564F	4300	0047	7101	0011	2B00	0052	4545	5945	452D	312E	564F	4300	.X...a...GAMOR-3.VOC..
000000D2	589C	0100	0119	0000	424F	5353	4B2D	312E	564F	4300	0059	B501	0001	F900	0043	5245	4154	5552	312E	564F	4300	X.....BOSSK-1.VOC..
000000FC	5AAE	0200	6135	0000	5052	4F42	452D	312E	564F	4300	00BB	E302	0021	2A00	0049	4E54	414C	4552	542E	564F	4300	Z...a5..PROBE-1.VOC..
00000126	DC0D	0300	A129	0000	5241	4E4F	4643	3032	2E56	4F43	007D	3708	0021	5200	0052	414E	4F46	4330	342E	564F	4300)..RANOFC02.VOC..
00000150	9E89	0300	AE4A	0000	5241	4E4F	4643	3035	2E56	4F43	004C	D408	0021	1900	0052	414E	4F46	4330	362E	564F	4300J..RANOFC05.VOC..
0000017A	6DED	0300	2134	0000	5241	4E53	544F	3081	2E56	4F43	008E	2104	0021	3F00	0052	414E	5354	4F30	322E	564F	4300	m...!4..RANST001.VOC..
000001A4	AF60	0400	213F	0000	5241	4E53	544F	3083	2E56	4F43	00D0	9F04	0021	5200	0052	414E	5354	4F30	342E	564F	4300	...`!?.RANST008.VOC..
000001CE	F1F1	0400	E137	0000	5241	4E53	544F	3035	2E56	4F43	00D2	2905	0021	2200	0052	414E	5354	4F30	362E	564F	43007..RANST005.VOC..
000001F8	F34B	0500	2143	0000	5241	4E53	544F	3087	2E56	4F43	0014	8F05	0021	2A00	0052	414E	5354	4F30	382E	564F	4300	.K...!C..RANST007.VOC..
00000222	35B9	0500	A125	0000	4158	452D	312E	564F	4300	0000	00D6	DE05	0021	1D00	0049	4E54	5354	554E	2E56	4F43	0000	5....%..AXE-1.VOC..
0000024C	F7FB	0500	210F	0000	5052	4F42	4649	5231	2E56	4F43	0018	OB06	00E0	8700	0043	5245	4154	5552	322E	564F	4300!..PROBFIR1.VOC..
00000276	F892	0600	2118	0000	5354	2D44	4945	2D31	2E56	4F43	0019	AB06	0061	1D00	0047	414D	4F52	2D32	2E56	4F43	0000!..ST-DIE-1.VOC..
000002A0	7AC8	0600	4121	0000	5245	4559	4545	2D32	2E56	4F43	00BB	E906	00C1	1200	0042	4F53	534B	2D33	2E56	4F43	0000	z...A!..REYEE-2.VOC..
000002A1	70FC	0600	271F	0000	A352	A5A1	544A	525A	2F5A	4FA3	00A3	1A07	00A1	5000	00A2	4F53	534R	41A9	452F	564F	4300	I`!`!..PREFATHRT.VOC..

- } First field is increasing in each successive entries
→ Possibly a **File Offset**?
- } If first theory is correct, second field has a high chance to be the **File Size**

Verifying

00003BA0	8080 8080 8080 8080 0043 7265 6174 6976 6520 566F 6963 6520 4669 6C65 1A1A 000A	Creative Voice File....
00003BC0	0129 1106 0200 00FF FF01 5B11 00A5 007F 808E 8F96 959D 9D9B 9D9F A19E A4A4 ACAA	(.....[.....
00003BE0	B0AF B1AD A9A4 9E94 8C7F 7472 6565 6267 6C74 7873 7B7A 6F67 6056 5757 545C 6064treebgltxs{zog`VWWTW`d
00003C00	707A 828B 9194 989F 9A9B 9596 8D8B 8482 898B 8A87 8586 8787 8996 9BA3 AFB9 C1C8	pz.....
00003C20	C2C2 C2B4 A695 868A 827E 7C78 8087 848A 8C91 A2A6 ABB1 B0AB 9C90 7669 6861 6465~ x.....vihae
00003C40	6B72 8786 9098 9DA5 A6AA A6A5 AAA7 9D9F 9696 8E8C 857F 7473 6764 5C5F 5959 5B50	kr.....tsgdW_YY[]
00003C60	5A5D 5757 554D 4B46 3E47 454F 6057 7374 8487 8B91 958D 9785 8881 7E76 7E79 7C75	7]WWWWIMKE?GEN^Ww+ ~v~vlu

0000EC20	8080 8000 4372 6561 7469 7665 2056 6F69 6365 2046 696C 651A 1A00 0A01 2911 01A2	...Creative Voice File.....)
0000EC40	5100 A500 7F7F 7F7F 7E7F 7F7E 8080 8080 8080 8080 8181 8282 8281 8181 8180	Q.....~...~.....
0000EC60	8080 8080 8282 8080 8080 7F80 8080 8082 8180 8080 8080 8081 8080 8080 8080 8080
0000EC80	8080 8081 8180 8181 8080 8080 7F7F 8080 807F 7F7F 7E7E 7D7E 8080 8080 8080 7F7F~~}~.....
0000ECA0	7F7E 7E7E 7E7F 7F7F 7F7E 7F7F 7F7F 7F7F 7F7F 7E7E 7E7F 7F80 807F 7F80 8080	~~~~.....~.....
0000ECC0	7F7F 807F 7E7E 7F80 8180 8080 8080 8080 8080 8080 7F7F 807F 807F 8080 8080~.....
0000ECF0	7F84 7C7E 7F7D 8181 7D81 807A 827F 807D 7C86 7C7B 7F83 7881 7A82 7A83 7786 8178	~~} } z ~ } {~ x z z w x

00000000	474F 420A F7C9 1300 4372 6561 7469 7665 2056 6F69 6365 2046 696C 651A 1A00 0A01	G0B.....Creative Voice File.....
00000020	2911 0182 3B00 A500 8080 8080 8080 807F 7F80 807F 7E7E 7F7F 7F7F 7F7F 8180 8080).....~...~.....
00000040	8082 837E 7E7E 7D7B 7A79 7879 7C7E 7D7C 7E80 7E7C 7D7D 7B7A 7B79 7676 7879 7A7F	...~~}{zyxy ~} ~ ~ }}{z{yvvxyz.
00000060	8384 8985 7E80 8584 837E 7D81 8180 8080 7F7B 7677 7B7A 7979 7C81 7D7D 7E81 807A~...~}.....{vw[zyy .}}~..z
00000080	7575 797F 827F 7C81 827D 8181 7B7E 807A 7977 7D89 8182 807F 8886 7F7E 7A72 7683	uuy..... ..}{~,zyw}.....~zrv.
000000A0	867C 787F 8C99 9CA0 A195 8D8F 8C80 7C75 7179 7973 747A 7875 7775 726E 737C 8282	. x.....luavvstzxuwurns ..

Confirmed! J

Extracting the resources using a custom script

```
import struct
import os
import sys
import string

if len(sys.argv) < 2:
    print "usage : swdf_fileextract filename"
    exit(1)

data = open(sys.argv[1],"rb").read()
gaif = struct.unpack("4s",data[0:4])[0]
directory_offset = struct.unpack("L",data[4:8])[0]
filenum = struct.unpack("L",data[directory_offset:directory_offset+4])[0]
newdirname = "./"+sys.argv[1][:sys.argv[1].index('.')]
try:
    os.makedirs(newdirname)
except OSError:
    pass
print "Game Archive Identification String :",gaif
print "Directory Offset :",hex(directory_offset)
print "Number of Files :",filenum

for i in range(filenum):
    directory_index = directory_offset+4+(i*0x15)
    file_offset = struct.unpack("L",data[directory_index:directory_index+4])[0]
    file_size = struct.unpack("L",data[directory_index+4:directory_index+8])[0]
    file_name = string.join(struct.unpack("13s",data[directory_index+8:directory_index+8+13]))
    file_data = struct.unpack("%ds"%file_size,data[file_offset:file_offset+file_size])[0]
    file_name = file_name.strip('\0')
    outfile = open("./%s/%s"%(newdirname,file_name),"wb")
    outfile.write(file_data)
    outfile.close()
```

The extracted Resources

Name	Size	Type
EXECMUS.GMD	31,3 KB	GMD
STAR-THM.GMD	28,1 KB	GMD
STALK-01.GMD	24,2 KB	GMD
FIGHT-01.GMD	25,8 KB	GMD
WEAPON1.VOC	7,81 KB	VOC
SCRSHOT.VOC	8,03 KB	VOC
PLASMA4.VOC	8,28 KB	VOC
MISSILE1.VOC	7,40 KB	VOC
M01KYL01.VOC	28,6 KB	VOC
ELEV3-3.VOC	9,78 KB	VOC
ELEV3-2.VOC	13 KB	VOC
ELEV3-1.VOC	9,50 KB	VOC
EEEK-2.VOC	6,87 KB	VOC
EEEK-1.VOC	4,95 KB	VOC
DOOR1-3.VOC	2 KB	VOC
DOOR1-2.VOC	16,7 KB	VOC
DOOR1-1.VOC	12,8 KB	VOC
CONCUSS6.VOC	2,96 KB	VOC
CONCUSS5.VOC	10,2 KB	VOC

Files extracted! ↴ Moving on...

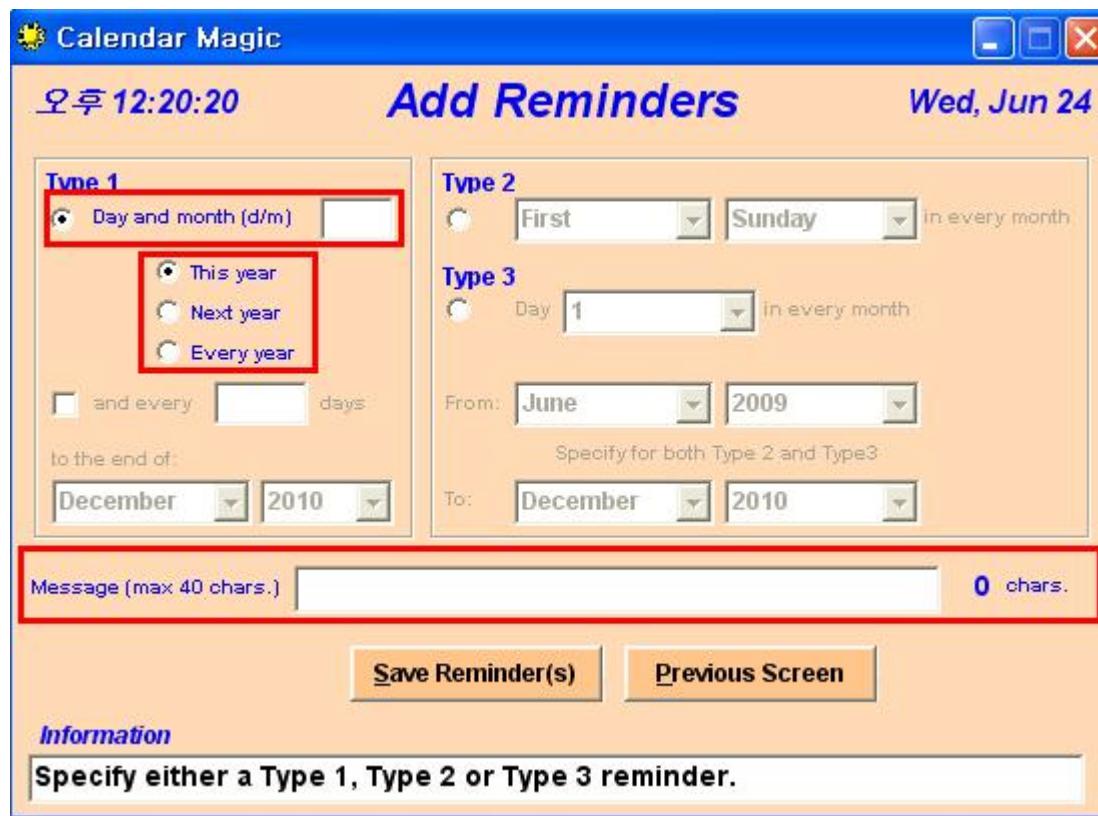
Other things to look out for when reversing Game Archive Files

- } First File Offset
- } Archive Name
- } Filename Offset
- } Filename Directory Offset
- } Total File Data Size
- } Total Directory Size
- } Archive Size
- } Number Of Directories
- } Directory Offset
- } File Extension / Type
- } File ID
- } Archive Version
- } Filename Length
- } Decompressed File Size
- } Checksum
- } Timestamp

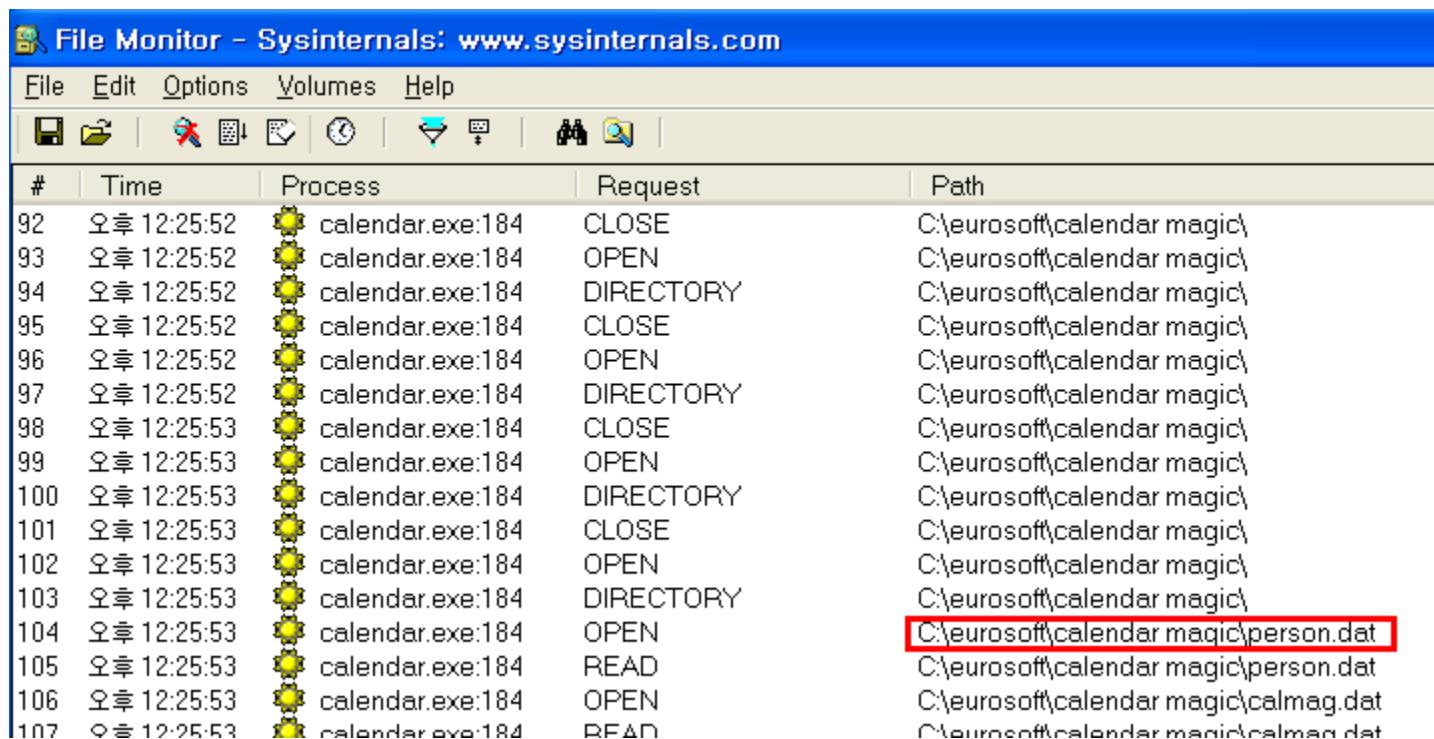
Case II : Reversing a Calendar Data File



Guess what's going to be stored in the Data File



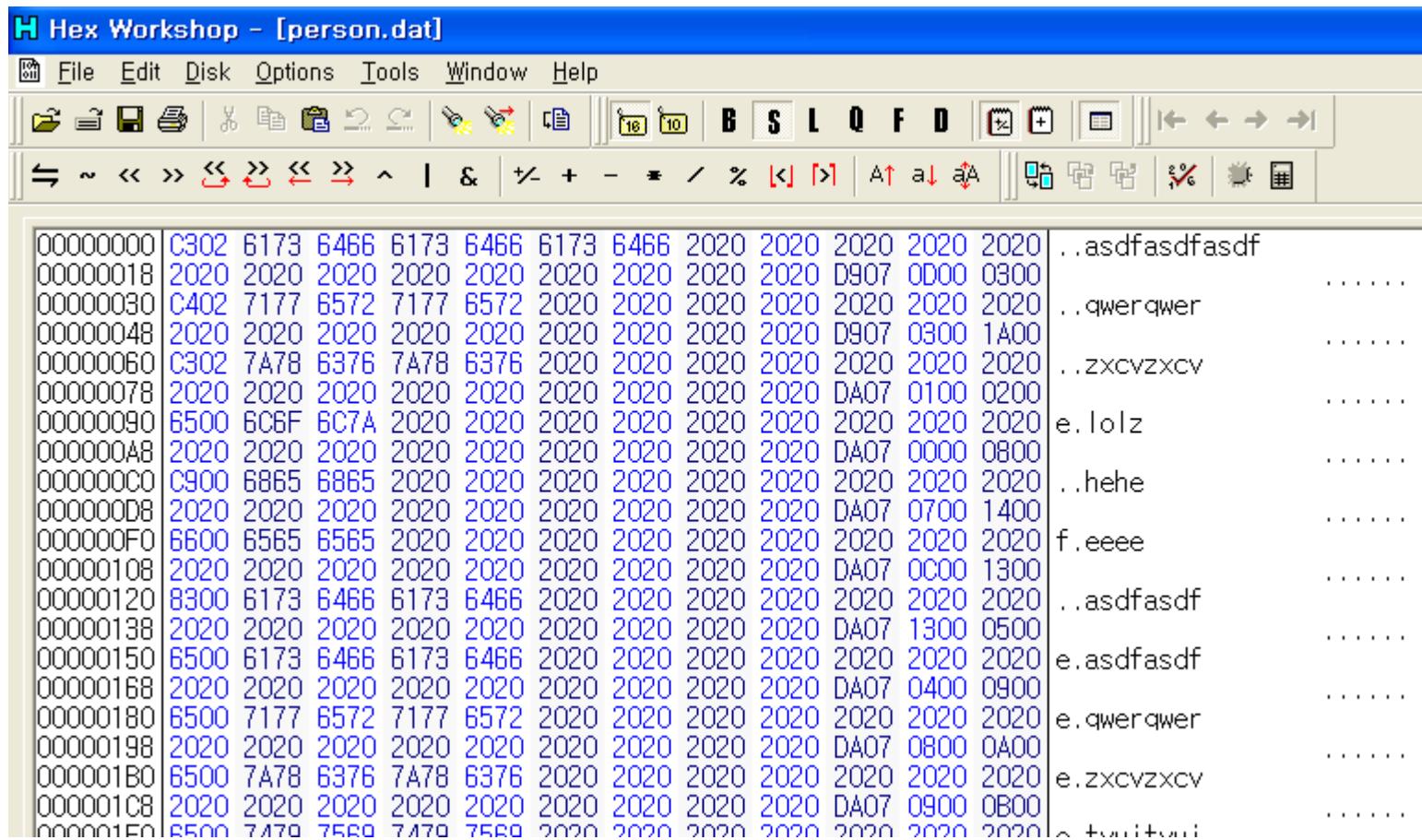
Find the Data File



The screenshot shows the Sysinternals File Monitor application interface. The title bar reads "File Monitor - Sysinternals: www.sysinternals.com". The menu bar includes File, Edit, Options, Volumes, and Help. Below the menu is a toolbar with icons for file operations like Open, Save, Find, and Filter. The main window is a table with columns: #, Time, Process, Request, and Path. The table lists 117 entries of file activity for the process "calendar.exe:184". The "Path" column shows the full file path for each request. A red box highlights the entry at row 104, which is "OPEN" for the file "C:\eurosoft\calendar magic\person.dat".

#	Time	Process	Request	Path
92	오후 12:25:52	calendar.exe:184	CLOSE	C:\eurosoft\calendar magic\
93	오후 12:25:52	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\
94	오후 12:25:52	calendar.exe:184	DIRECTORY	C:\eurosoft\calendar magic\
95	오후 12:25:52	calendar.exe:184	CLOSE	C:\eurosoft\calendar magic\
96	오후 12:25:52	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\
97	오후 12:25:52	calendar.exe:184	DIRECTORY	C:\eurosoft\calendar magic\
98	오후 12:25:53	calendar.exe:184	CLOSE	C:\eurosoft\calendar magic\
99	오후 12:25:53	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\
100	오후 12:25:53	calendar.exe:184	DIRECTORY	C:\eurosoft\calendar magic\
101	오후 12:25:53	calendar.exe:184	CLOSE	C:\eurosoft\calendar magic\
102	오후 12:25:53	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\
103	오후 12:25:53	calendar.exe:184	DIRECTORY	C:\eurosoft\calendar magic\
104	오후 12:25:53	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\person.dat
105	오후 12:25:53	calendar.exe:184	READ	C:\eurosoft\calendar magic\person.dat
106	오후 12:25:53	calendar.exe:184	OPEN	C:\eurosoft\calendar magic\calmag.dat
107	오후 12:25:53	calendar.exe:184	RFDN	C:\eurosoft\calendar magic\calmag.dat

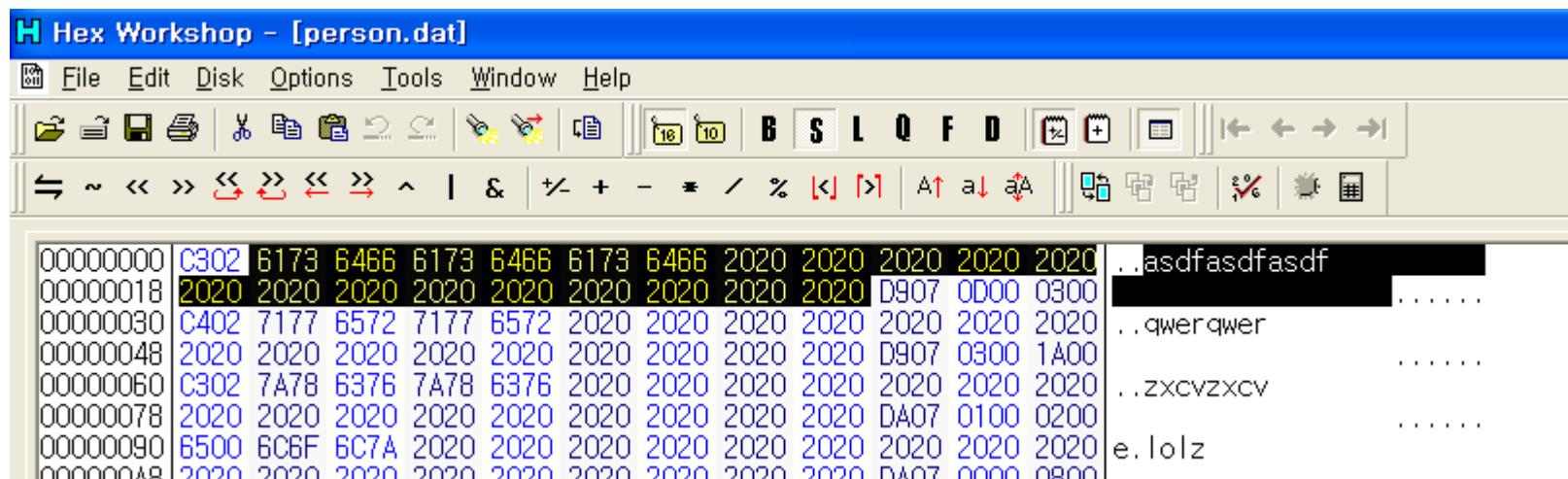
Start Digging! 😊



The screenshot shows the Hex Workshop application interface. The title bar reads "H Hex Workshop - [person.dat]". The menu bar includes File, Edit, Disk, Options, Tools, Window, and Help. The toolbar contains various icons for file operations like Open, Save, Copy, Paste, Find, and Replace. Below the toolbar is a set of buttons labeled B, S, L, Q, F, D. The main window displays a hex dump of a file. The left column shows memory addresses from 00000000 to 000001F0. The right column shows the corresponding ASCII representation. Some data is redacted with ellipses (...). The file contains several strings of text, such as ".asdfasdf", ".qwerqwer", ".zxcvzxcv", ".lolz", ".hehe", ".eeee", ".asdfasdf", ".qwerqwer", ".zxcvzxcv", and ".fonfon".

Address	Hex Value	ASCII
00000000	C302 6173 6466 6173 6466 6173 6466 2020 2020 2020 2020 2020 2020	..asdfasdf
00000018	2020 2020 2020 2020 2020 2020 2020 2020 D907 0D00 0300
00000030	C402 7177 6572 7177 6572 2020 2020 2020 2020 2020 2020 2020 2020	..qwerqwer
00000048	2020 2020 2020 2020 2020 2020 2020 2020 D907 0300 1A00	..qwerqwer
00000060	C302 7A78 6376 7A78 6376 2020 2020 2020 2020 2020 2020 2020 2020	..zxcvzxcv
00000078	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0100 0200
00000090	6500 6C6F 6C7A 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020	e.lolz
000000A8	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0000 0800
000000C0	C900 6865 6865 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020	..hehe
000000D8	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0700 1400
000000F0	6600 6565 6565 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020	f.eeee
00000108	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0000 1300
00000120	8300 6173 6466 6173 6466 2020 2020 2020 2020 2020 2020 2020 2020	..asdfasdf
00000138	2020 2020 2020 2020 2020 2020 2020 2020 DA07 1300 0500
00000150	6500 6173 6466 6173 6466 2020 2020 2020 2020 2020 2020 2020 2020	e.asdfasdf
00000168	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0400 0900
00000180	6500 7177 6572 7177 6572 2020 2020 2020 2020 2020 2020 2020 2020	e.qwerqwer
00000198	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0800 0A00
000001B0	6500 7A78 6376 7A78 6376 2020 2020 2020 2020 2020 2020 2020 2020	e.zxcvzxcv
000001C8	2020 2020 2020 2020 2020 2020 2020 2020 DA07 0900 0B00
000001E0	6500 7A70 75E0 7A70 75E0 2020 2020 2020 2020 2020 2020 2020 2020	e.fonfon

Message Part



Hex Workshop - [person.dat]

File Edit Disk Options Tools Window Help

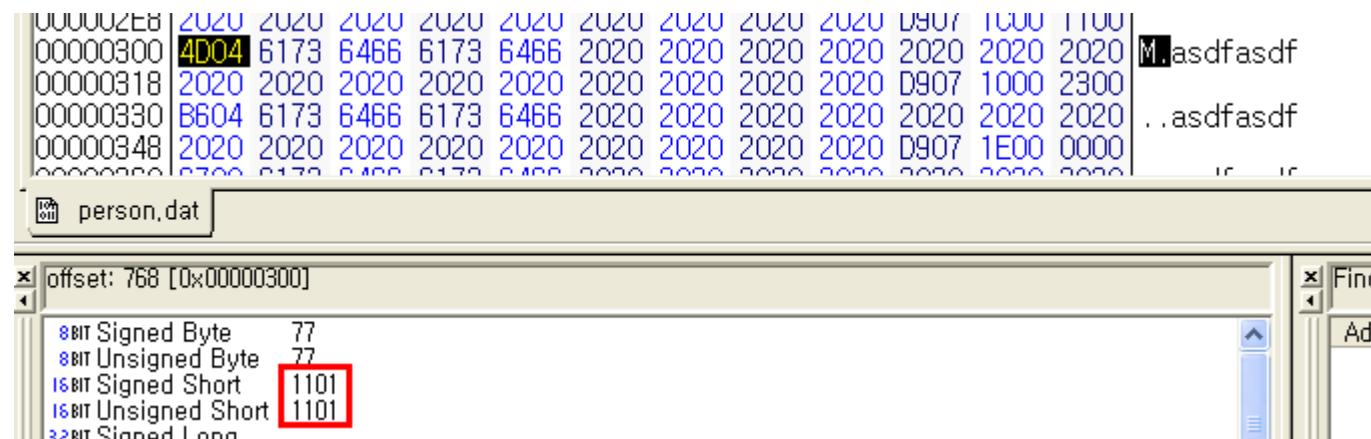
B S L Q F D

00000000 C302 6173 6466 6173 6466 6173 6466 2020 2020 2020 2020 2020 2020 2020 2020 asdfasdfasdf
00000018 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020
00000030 C402 7177 6572 7177 6572 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 ..qwerqwer
00000048 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020
00000060 C302 7A78 6376 7A78 6376 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 ..zxcvzxcv
00000078 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020
00000090 6500 6C6F 6C7A 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 e.lolz
000000A8 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020

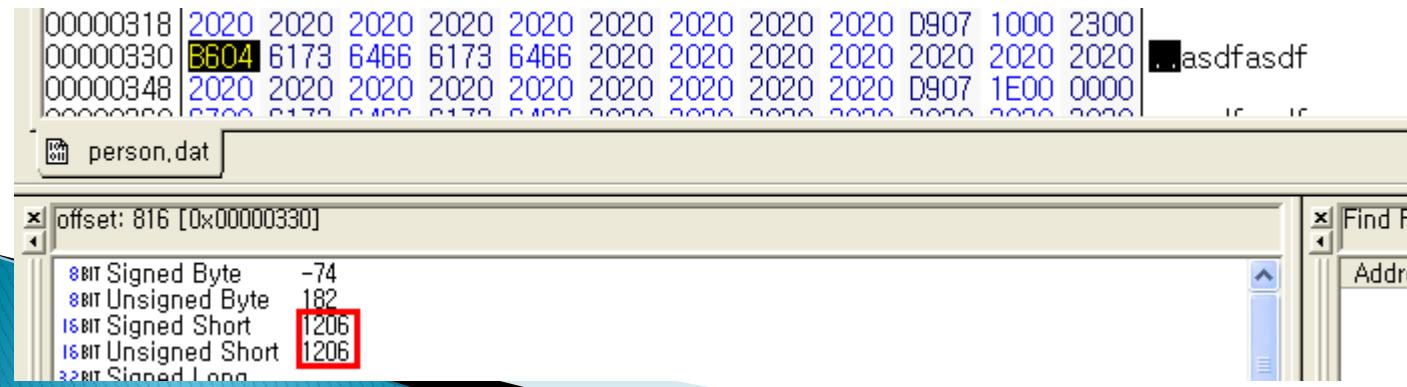
} Max 40 Chars

Date Field

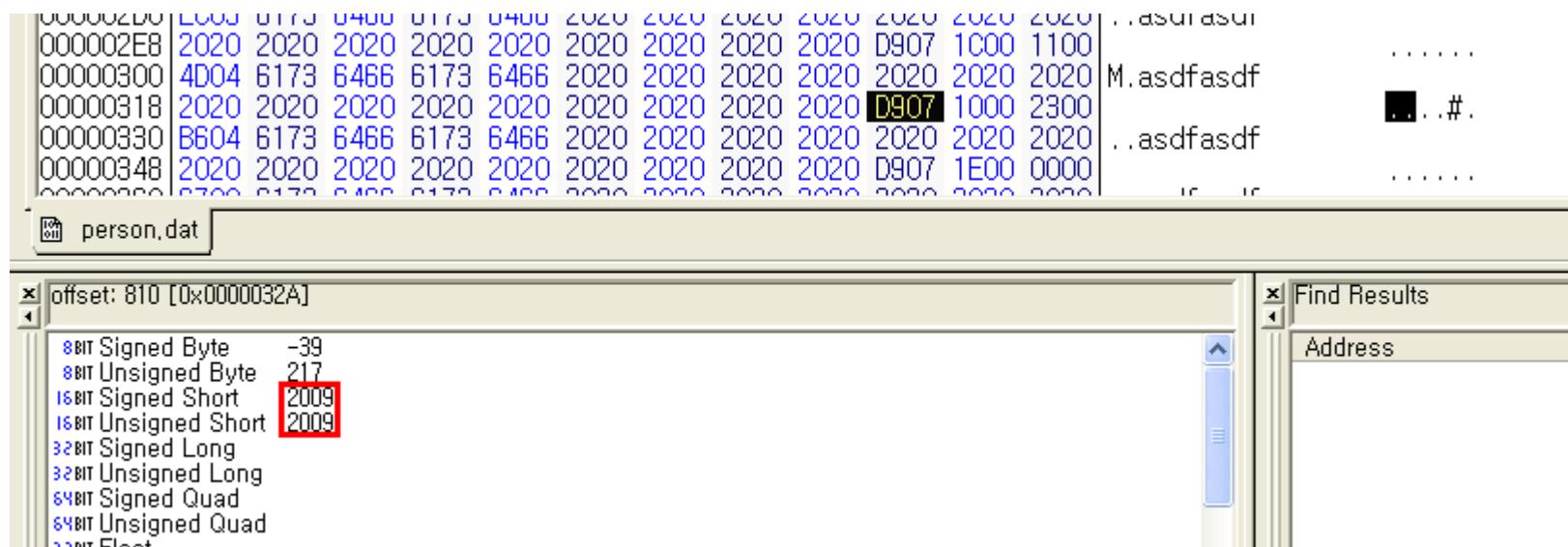
} 11/1



} 12/6



Year Field



The Excess Data

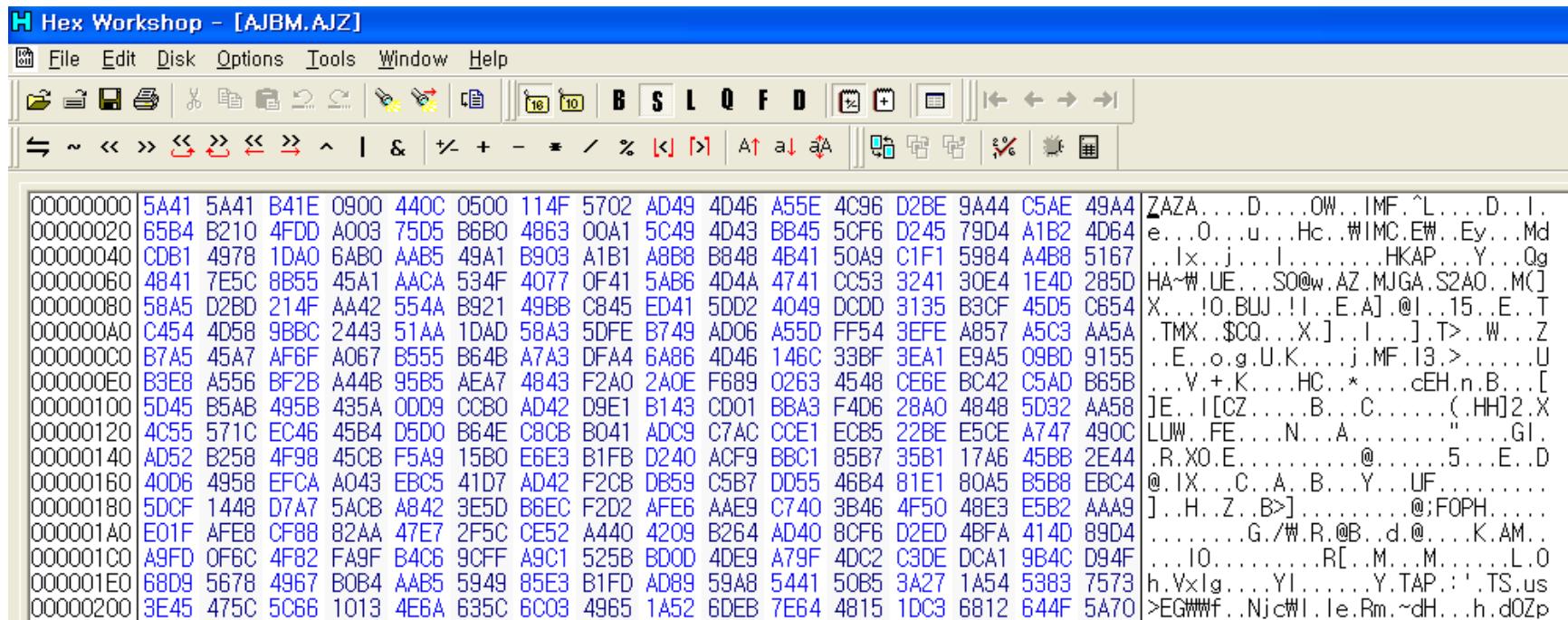
} Possibly some kind of unique identifier

Person.dat File Format

Chunk Structure

Size	Meaning
2	Date
40	Message
2	Year
4	Identifier

Case III : Reversing Firmware



The screenshot shows a hex editor window titled "Hex Workshop - [AJBM.AIZ]". The menu bar includes File, Edit, Disk, Options, Tools, Window, and Help. The toolbar contains various icons for file operations like Open, Save, Copy, Paste, and Find. Below the toolbar is a set of navigation and search buttons. The main pane displays a hex dump of memory starting at address 00000000. The data consists of pairs of hex digits followed by ASCII characters and symbols. A vertical scrollbar is on the right side of the main pane.

00000000	5A41 5A41 B41E 0900 440C 0500 114F 5702 AD49 4D46 A55E 4C96 D2BE 9A44 C5AE 49A4	ZAZA....D....OW..IMF.^L....D...I.
00000020	65B4 B210 4FDD A003 75D5 B6B0 48E3 00A1 5C49 4D43 BB45 5CF6 D245 79D4 A1B2 4D64	e...O...u...Hc..#IMC.EW..Ey...Md
00000040	CDB1 4978 1DAO 6AB0 AAB5 49A1 B903 A1B1 A8B8 B848 4B41 50A9 C1F1 5984 A4B8 5167	.Ix..j....I.....HKAP...Y...Qg
00000060	4841 7E5C 8B55 45A1 AAC4 534F 4077 0F41 5AB6 4D4A 4741 CC53 3241 30E4 1E4D 285D	HA~W.UE...SO@w.AZ.MJGA.S2AO..M()
00000080	58A5 D2BD 214F AA42 554A B921 49BB C845 ED41 5DD2 4049 DCDD 3135 B3CF 45D5 C654	X...!O.BUJ.!I..E.A).@I...15..E..T
000000A0	C454 4D58 9BBC 2443 51AA 1DAD 58A3 5DFE B749 AD06 A55D FF54 3EFE A857 A5C3 AA5A	.TMX..\$CQ...X.]..I...].T>..W...Z
000000C0	B7A5 45A7 AF6F A067 B555 B64B A7A3 DFA4 6A86 4D46 146C 33BF 3EA1 E9A5 09BD 9155	..E..o.g.U.K...j.MF.13,>....U
000000E0	B3E8 A556 BF2B A44B 95B5 AEA7 4843 F2A0 2A0E F689 0263 4548 CE6E BC42 C5AD B65B	.V.+K....HC..*...cEH.n.B...[
00000100	5D45 B5AB 495B 435A 0DD9 CCB0 AD42 D9E1 B143 CD01 BBA3 F4D6 28A0 4848 5D32 AA58	JE..I[CZ....B...C.....(HH]2.X
00000120	4C55 571C EC46 45B4 D5D0 B64E C8CB B041 ADC9 C7AC CCE1 ECB5 22BE E5CE A747 490C	LUM..FE....N...A....."....GI.
00000140	AD52 B258 4F98 45CB F5A9 15B0 E6E3 B1FB D240 ACF9 BBC1 85B7 35B1 17A6 45BB 2E44	.R.X0.E.....@.....5...E..D
00000160	40D6 4958 EFCA A043 EBC5 41D7 AD42 F2CB DB59 C5B7 DD55 46B4 81E1 80A5 B5B8 EBC4	@.IX...C..A..B...Y...UF.....
00000180	5DCF 1448 D7A7 5ACB A842 3E5D B6EC F2D2 AFE6 AAE9 C740 3B46 4F50 48E3 E5B2 AAA9	J..H..Z..B>].....@;FOPH.....
000001A0	E01F AFE8 CF88 82AA 47E7 2F5C CE52 A440 4209 B264 AD40 8CF6 D2ED 4BFA 414D 89D4G./W.R.@B..d.0...K.AM..
000001C0	A9FD 0F6C 4F82 FA9F B4C6 90FF A9C1 525B BD0D 4DE9 A79F 4DC2 C3DE DCA1 9B4C D94F	...10.....R[..M...M.....L.0
000001E0	68D9 5678 4967 B0B4 AAB5 5949 85E3 B1FD AD89 59A8 5441 50B5 3A27 1A54 5383 7573	h.Vxlg....YI.....Y.TAP.:'.TS.us
00000200	3E45 475C 5C66 1013 4E6A 635C 6C08 4965 1A52 6DEB 7E64 4815 1DC3 6812 644F 5A70	>EGWWf..Njc#I.Ie.Rm.~dH...h.d0Zp

You know what to do next! J

Try to identify the Header

	FileType	Unknown	Data Size	Checksum	
00000000	5A41	5A41	B41E	0900	440C 0500 114F 5702 AD49 4D46 A55E 4C96 D2BE 9A44 C5AE 49A4 ZAZA
00000020	65B4	B210	4FDD	A003	75D5 B6B0 4863 00A1 5C49 4D43 BB45 5CF6 D245 79D4 A1B2 4D64 e...0.
00000040	CDB1	4978	1DA0	6AB0	AAB5 49A1 B903 A1B1 A8B8 4B41 50A9 C1F1 5984 A4B8 5167 ..Ix..
00000060	4841	7E5C	8B55	45A1	AAAC 534F 4077 0F41 5AB6 4D4A 4741 CC53 3241 30E4 1E4D 285D HA~W.UB
00000080	58A5	D2BD	214F	AA42	554A B921 49BB C845 ED41 5DD2 4049 DCDD 3135 B3CF 45D5 C654 X...!O
000000A0	C454	4D58	9BBC	2443	51AA 1DAD 58A3 5DFE B749 AD06 A55D FF54 3EFE A857 A5C3 AA5A .TMX..S
000000C0	B7A5	45A7	AF6F	A067	B555 B64B A7A3 DFA4 6A86 4D46 148C 33BF 3EA1 E9A5 09BD 9155 ..E..o.
000000E0	B3E8	A556	BF2B	A44B	95B5 AEA7 4843 F2A0 2AOE F689 0263 4548 CE6E BC42 C5AD B65B ...V.+
00000100	5D45	B5AB	495B	435A	0DD9 CCB0 AD42 D9E1 B143 CD01 BBA3 F4D6 28A0 4848 5D32 AA58]E..!([O
00000120	4C55	571C	EC46	45B4	D5D0 B64E CBCB B041 ADC9 C7AC CCE1 ECB5 22BE E5CE A747 490C LUW..FB
00000140	AD52	B258	4F98	45CB	F5A9 15B0 E6E3 B1FB D240 ACF9 BBC1 85B7 35B1 17A6 45BB 2E44 .R.XO.B
00000160	40D6	4958	EFCA	A043	EBC5 41D7 AD42 F2CB DB59 C5B7 DD55 46B4 81E1 80A5 B5B8 EBC4 @.IX..
00000180	5DCF	1448	D7A7	5ACB	A842 3E5D B6EC F2D2 AFE6 AAE9 C740 3B46 4F50 48E3 E5B2 AAA9]..H..Z

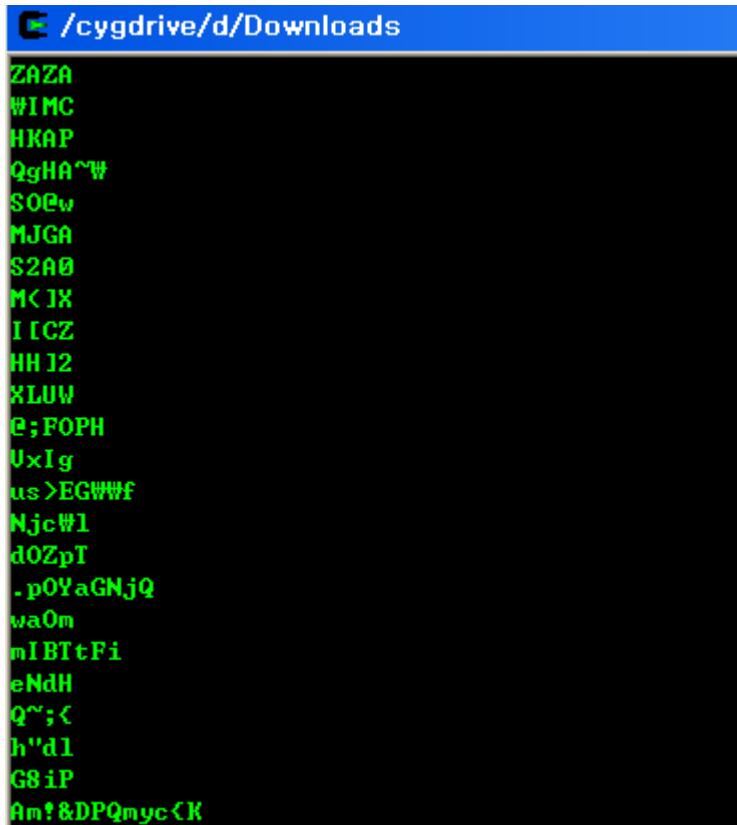
Algorithm	Checksum/Digest
Checksum-8	11
Checksum-16	4E11
Checksum-32	02574F11
Checksum-64	0000000002574F11
CRC-16	CCC2
CRC-16 CCITT	1003
CRC-32	96D73885
Custom CRC (32 bit)	96D73885
MD2	2E198F82949A24FB3

Compare Checksum Find Bookmarks Output
Offset: 00000010 Sel: 0x50c44 bytes 330836 b

Copying the data excluding the Header to a separate File

000000000	AD49	4D46	A55E	4C96	D2BE	9A44	C5AE	49A4	65B4	B210	4FDD	A003	75D5	B6B0	4863	00A1	t
000000020	5C49	4D43	BB45	5CF6	D245	79D4	A1B2	4D64	CDB1	4978	1DAO	6AB0	AAB5	49A1	B903	A1B1	z
000000040	A8B8	B848	4B41	50A9	C1F1	5984	A4B8	5167	4841	7E5C	8B55	45A1	AACA	534F	4077	0F41	z
000000060	5AB6	4D4A	4741	CC53	3241	30E4	1E4D	285D	58A5	D2BD	214F	AA42	554A	B921	49BB	C845	z
000000080	ED41	5DD2	4049	DCDD	3135	B3CF	45D5	C654	C454	4D58	9BBC	2443	51AA	1DAD	58A3	5DFE	z
0000000A0	B749	AD06	A55D	FF54	3EFE	A857	A5C3	AA5A	B7A5	45A7	AF6F	A067	B555	B64B	A7A3	DFA4	z
0000000C0	6A86	4D46	146C	33BF	3EA1	E9A5	09BD	9155	B3E8	A556	BF2B	A44B	95B5	AEA7	4843	F2A0	z
0000000E0	2AOE	F689	0263	4548	CE6E	BC42	C5AD	B65B	5D45	B5AB	495B	435A	0DD9	CCB0	AD42	D9E1	z
00000100	B143	CD01	BBA3	F4D6	28A0	4848	5D32	AA58	4C55	571C	EC46	45B4	D5D0	B64E	C8CB	B041	z
00000120	ADC9	C7AC	CCE1	ECB5	22BE	E5CE	A747	490C	AD52	B258	4F98	45CB	F5A9	15B0	E6E3	B1FB	z
00000140	D240	ACF9	BBC1	85B7	35B1	17A6	45BB	2E44	40D6	4958	EFCA	A043	EBC5	41D7	AD42	F2CB	z
00000160	DB59	C5B7	DD55	46B4	81E1	80A5	B5B8	EBC4	5DCF	1448	D7A7	5ACB	A842	3E5D	B6EC	F2D2	z
00000180	AFe6	AAE9	C740	3B46	4F50	48E3	E5B2	AAA9	E01F	AFe8	CF88	82AA	47E7	2F5C	CE52	A440	z
000001A0	4209	B264	AD40	8CF6	D2ED	4BFA	414D	89D4	A9FD	0F6C	4F82	FA9F	B4C6	9CFF	A9C1	525B	E
000001C0	BD0D	4DE9	A79F	4DC2	C3DE	DCA1	9B4C	D84F	68D9	5678	4967	B0B4	AAB5	5949	85E3	B1FD	z
000001E0	AD89	59A8	5441	50B5	3A27	1A54	5383	7573	3E45	475C	5C66	1013	4E6A	635C	6C03	4965	z
00000200	1A52	6DEB	7E64	4815	1DC3	6812	644F	5A70	54BF	2E70	4F59	6147	4E6A	5183	E85C	B1C3	z
00000220	9205	AF77	614F	6DEC	1094	EBFB	6D49	4254	7446	699C	BE50	5968	81A1	73F5	6163	956E	z
00000240	537B	9360	61C9	A772	ADB2	654E	6448	8278	4674	F386	4F53	31BA	CA5F	E966	4A13	517E	z
00000260	3B7B	9368	2264	6CBD	157D	BA68	D5CF	689E	61D7	6CDA	76A6	2A4B	8B4B	57C9	BB6F	FE86	z
00000280	5E6A	8360	A670	9277	088B	5A60	ED86	7D98	2355	1362	455E	151A	4738	6950	0815	7300	z
nnnnnn?An	A1BD	212B	4450	516D	7963	7R4R	09C0	FFRR	2917	9979	9A55	5B6F	6581	671F	27A3	A201	z

Search for readable strings

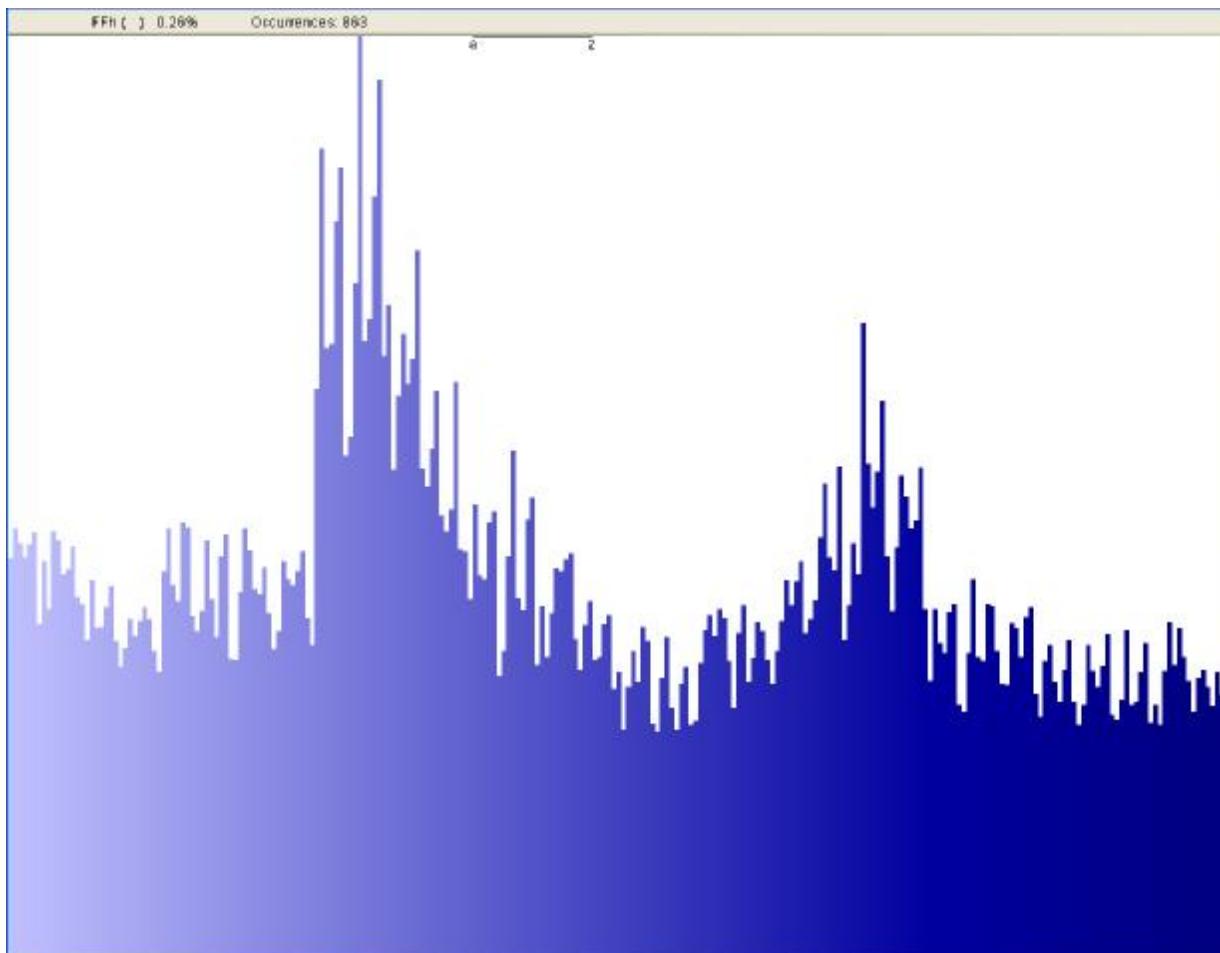


The screenshot shows a terminal window with the path `/cygdrive/d/Downloads` in the title bar. The content of the window is a list of strings, many of which are completely non-readable or compressed. The strings include:

```
ZAZA
WIMC
HKAP
QgHA~W
SOBw
MJGA
S2A0
MKJX
IICZ
HHJ2
XLUW
e;FOPH
UxIg
us>EGWWf
NjcW1
d0ZpT
.p0YaNjQ
waOm
mIBTtFi
eNdH
Q~;c
h"dl
G8iP
Am!&DPQmycKK
```

- } All strings are absolutely Incomprehensible
- } → Possibly Compressed, Encrypted
- } What to do next?
→ Frequency Analysis! ↗

Frequency Analysis of Individual Bytes

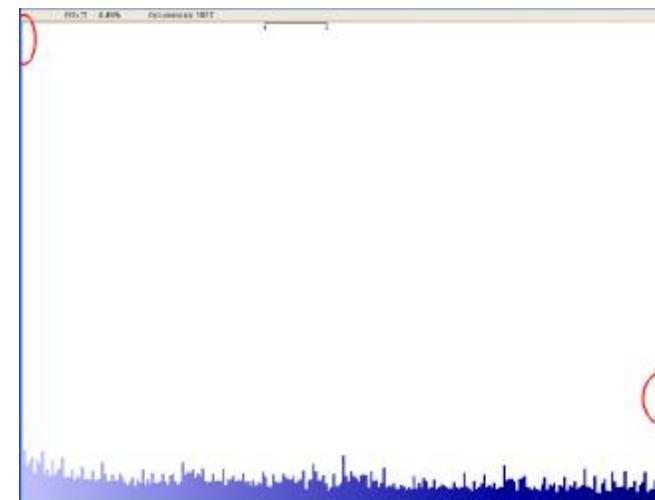
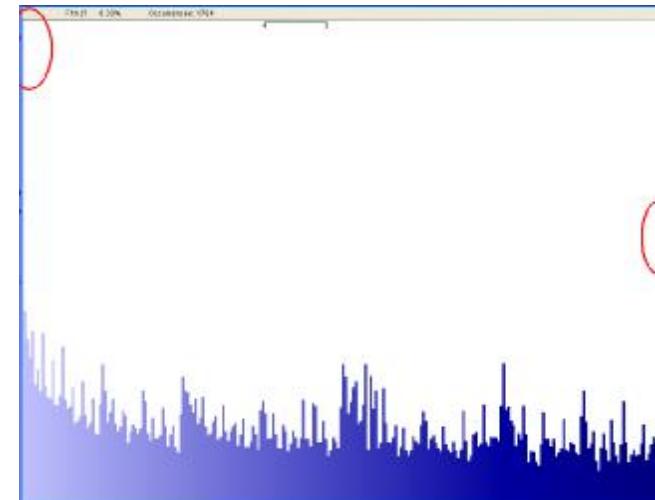
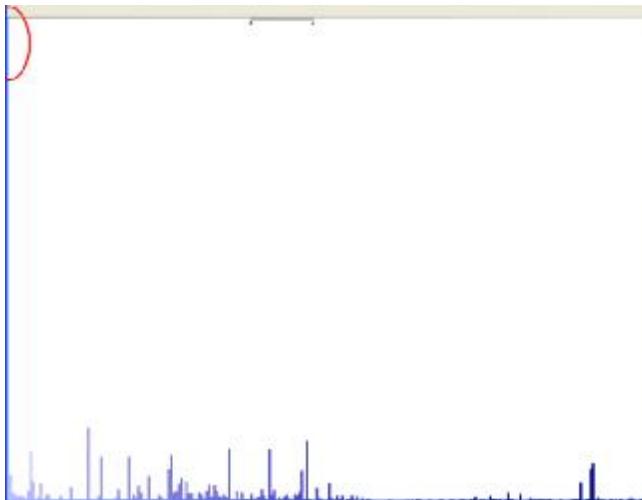


What can we understand from this?

CodeEngn

3rd CodeEngn ReverseEngineering Seminar

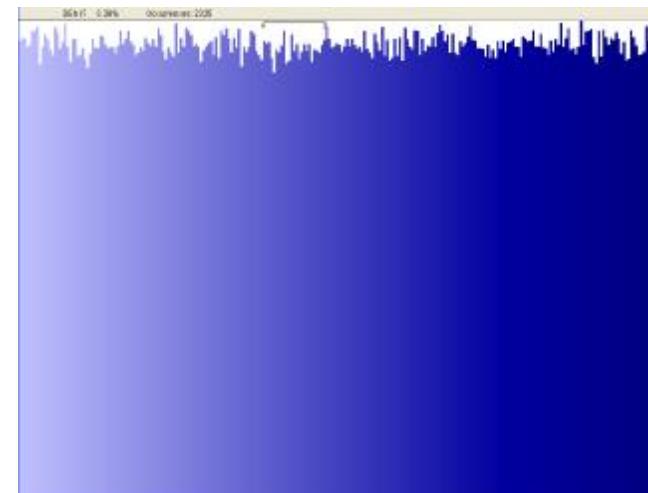
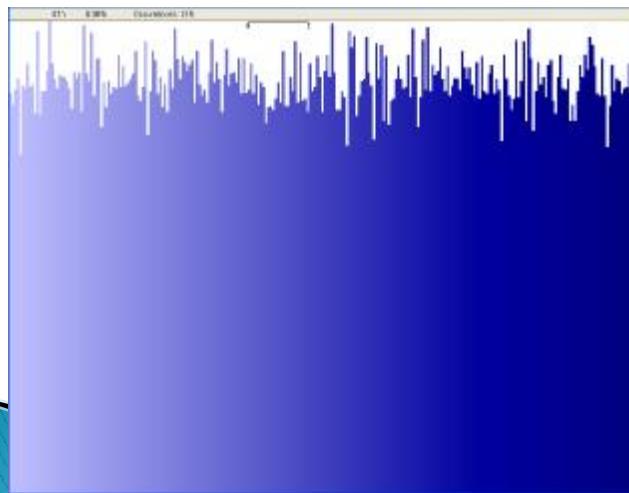
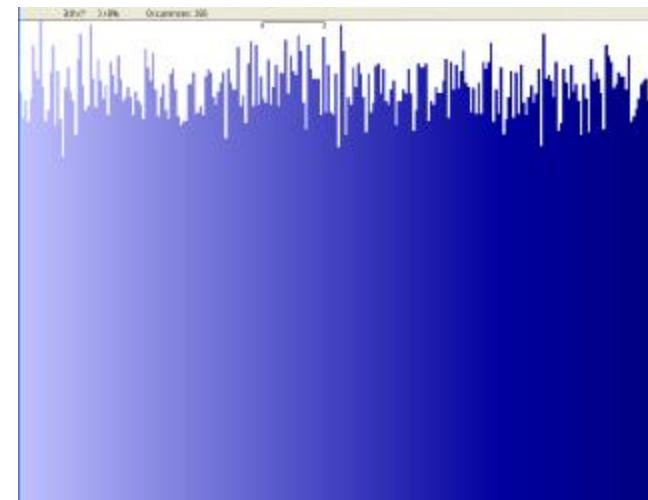
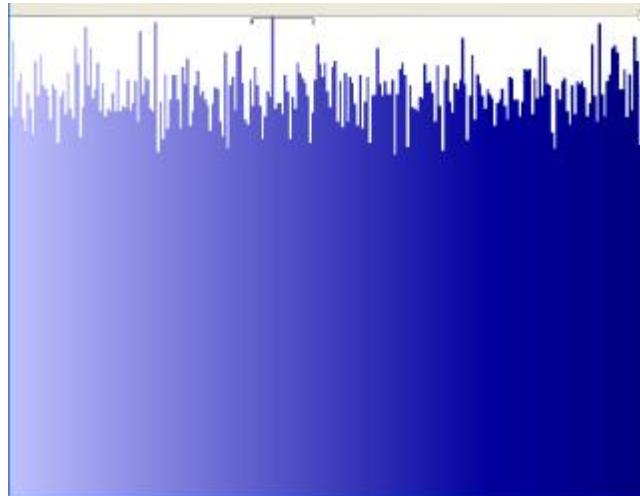
Frequency Analysis of Normal Executable Files



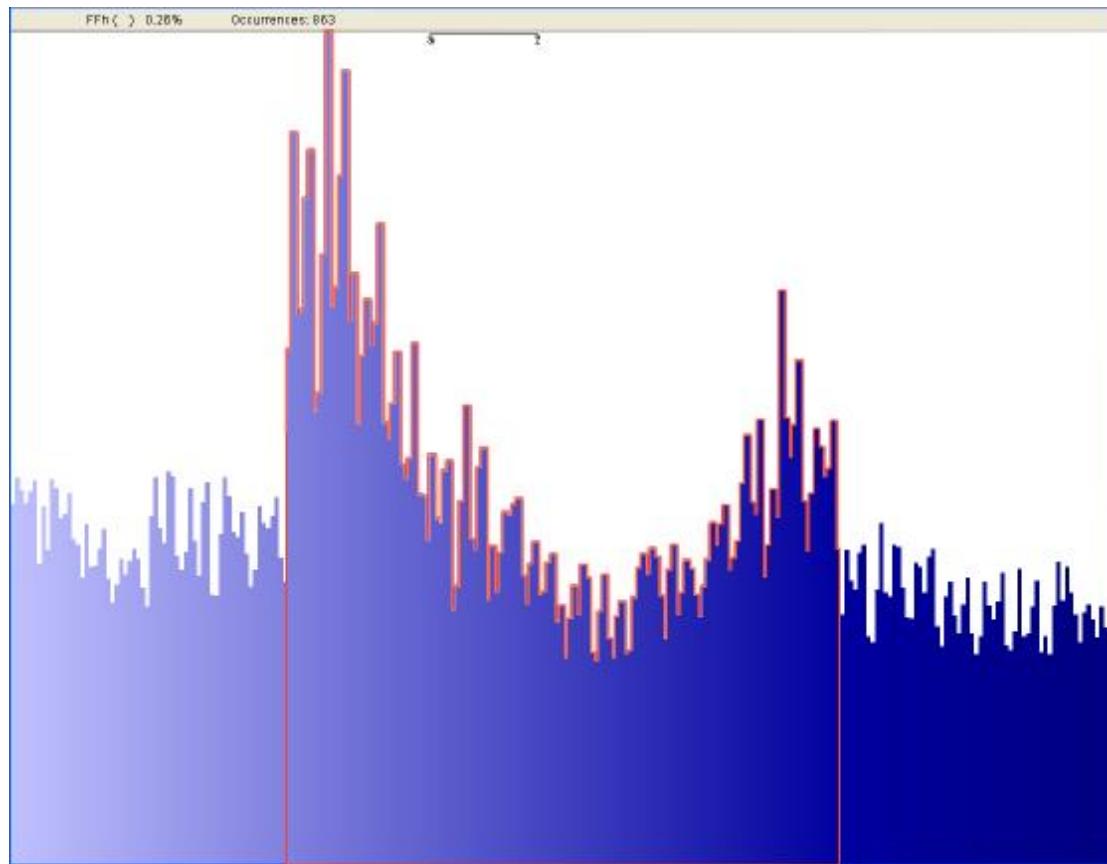
CodeEngn

3rd CodeEngn ReverseEngineering Seminar

Frequency Analysis of Encrypted Files(Modern Crypto)



The Symmetry of the Frequency Graph



Highly Symmetric → Pattern
→ **Sign of a Weak Cipher !!!**

CodeEngn

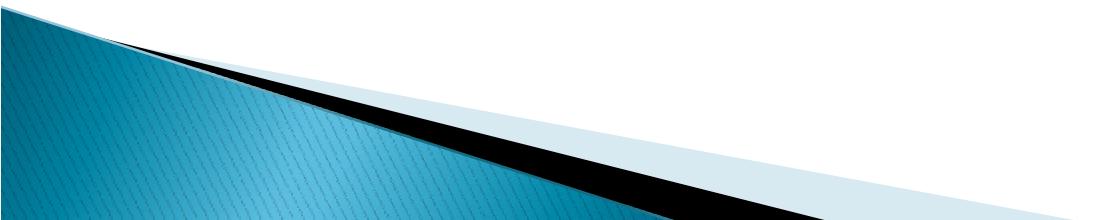
3rd CodeEngn ReverseEngineering Seminar

Guessing the Encryption Scheme

- } Monoalphabetic
- } Polyalphabetic
- } Transposition

~~}{~~ Monoalphabetic
(Frequency Doesn't Change)

~~}{~~ Polyalphabetic
~~}{~~ Transposition
(Frequency Doesn't Change)



Possible Source code of the Encryption Scheme

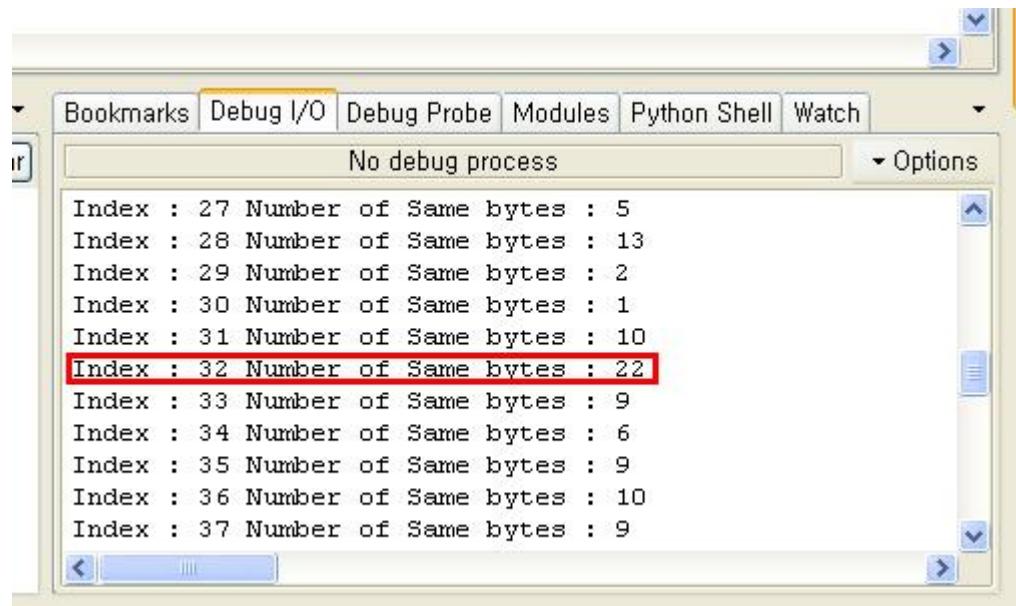
```
char[] key = "1234567890"

for(int i=0; i<filesize; i++){
    encrypted_data[i] = data[i]^key[i % strlen(key)]
}
```

- } Why? Because it is the simplest encryption algorithm with a key to implement! ↗
- } It also matches our theory of Polyalphabetic Cipher, or more like, a **Vigenère cipher**

Attacking Vigenère Ciphers

- Coincidence Counting -



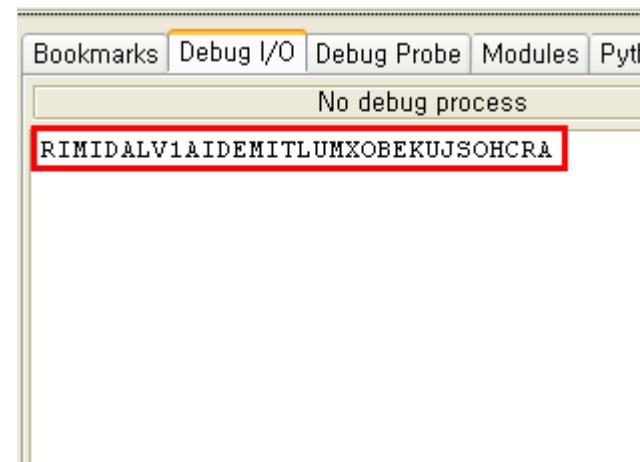
The screenshot shows a debugger interface with a tab bar at the top containing Bookmarks, Debug I/O (which is selected), Debug Probe, Modules, Python Shell, and Watch. Below the tab bar, a message box displays "No debug process". The main window contains a list of entries, each consisting of an index and the number of same bytes found at that index. The entry at index 32, which has 22 same bytes, is highlighted with a red box.

Index	Number of Same bytes
27	5
28	13
29	2
30	1
31	10
32	22
33	9
34	6
35	9
36	10
37	9

Key Length : 32 bytes

Making Assumptions

- } Key is XORed
- } Key is Added
- } Key is Subtracted
- }- Key is RORed
- }- Key is ROLed
- }{ ...
- }{ ...



Key Found! J

RIMIDALV1AIDEMITLUMXOBEKUJSOHCRA

Decrypted File

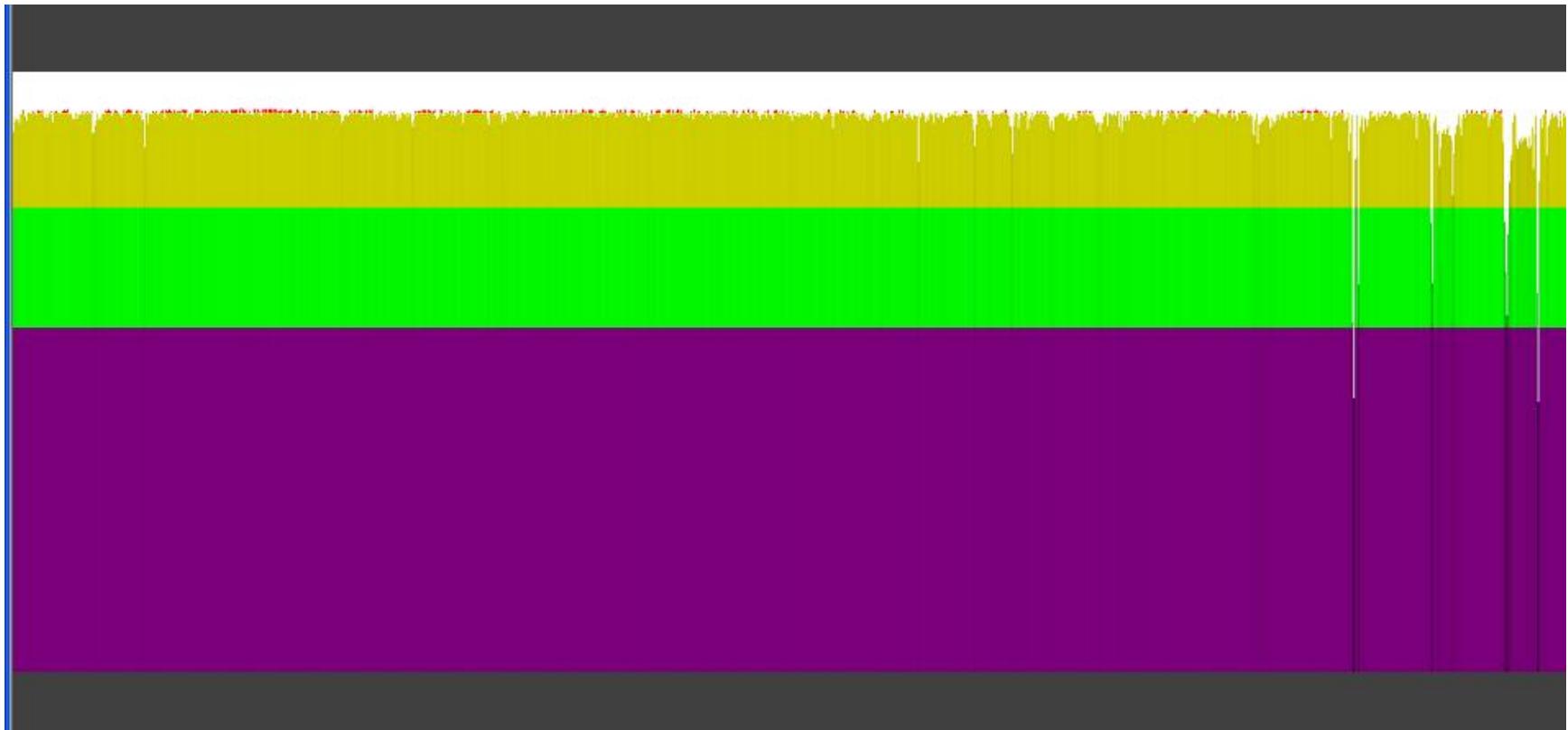
000000000	FF00 000F E11F 00C0 E3FF D300 80E3 00F0 29E1 FF48 009F E548 209F E5FF 0020 52E0) .H ..H . .R.
000000020	0E00 000A FF04 10A0 E304 3090 E4FF 0430 81E4 0420 52E2 2FFB FFFF 1AEE F140 F3F0	.0 ..0 ..R./ ..@ ..
000000040	FAF1 F501 0F00 1CFF F0B0 10C0 E1F5 1833 0414 3304 C417 00EA FF80 0000 0834 5D00	..3 ..3 ..4] ..
000000060	08FF 0003 0300 8005 0300 79A0 5B00 6109 14F0 9FE5 6E0D EF09 0000 EA6E 01F8 9A04	.y .[.a ..n ..n ..
000000080	BF08 109B 0408 908B 0074 FA8B 0098 8F00 8801 0000 D4FE 6108 04E0 4EE2 10E0 0FBF	.t ..a ..N ..
0000000A0	E500 E04F E11C B302 0FBF E113 E08E E30E FB00 08FF E02D E52C E01F E504 EFE0 8DE5	..0 ..- ..
0000000C0	38CF 000F 502D 7FE9 0FE0 A0E1 4CFO D801 FFBD E80E F069 E100 C0FF FDE8 0000 A0E1	8 ..P- ..L ..i ..
0000000E0	7847 BBC0 4622 091E FF2F F506 80E0 FF0F 1110 F8F3 0619 0611 5893 9FFF E501 8BA0	xG ..F" .. / ..X ..
000000100	E30A 8048 FFE2 B880 19E1 010C 187F E30C 0000 1A44 A304 00FF 809A E501 8088 E200	.H ..D ..
000000120	FF80 8AE5 88A0 A0E3 13FF AC8A E20A 0058 E107 FF00 00DA 0080 A0E3 46FF AEA0 E3BA	..X ..F ..
000000140	8009 E1B0 FF80 C9E1 04F0 5EE2 00F6 6710 0C83 0400 A088 E508 BE8F 1298 E501 A08A	^ ..g ..
000000160	8910 88FE 9914 0AE2 B0A0 C9E1 F0F5 A290 119A 5910 98E5 1F80 FD08 6D12 FEAF A093	.Y ..m ..
000000180	FDAF E7A0 8301 7710 7E11 01A7 A0FF E3FD AC4A E2B0 80CA C7E1 12AD 7C13 8611 F601	..w ..~ ..J ..I ..
0000001A0	1040 FF2D E901 C0A0 E3AC 02BE 0400 C080 E5AB 4234 0000 BFD4 E18C CF80 E182 001A	@ ..- ..B4 ..
0000001C0	EF44 00A0 E3DE 0194 F29F 95E5 DE01 901B 248C 1B20 0625 F5FF FFFF 0A06 CDAO E3BC	.D ..\$..% ..
0000001E0	FFC0 14E1 1000 1CE3 0B66 5310 16CE 3C27 7210 0A04 1324 5558 1B20 3013 2440 1B24	.fs ..<'r ..\$UX ..0.\$@.\$..
000000200	481B 20A2 3A25 0443 2C82 2156 2102 1324 18EA 6328 001B 240C 1B20 02CC A01F E382	H ..% ..C ..IVI ..\$..c(..\$..
000000220	C04C E23E 250E 21BA 21D5 A2BF 2804 0B00 3813 24C4 F112 1C23 D4EB 20BA 2920 C72F	L ..>% ..! ..(..8 ..\$..# ..) ..
000000240	0132 DE29 2588 EB24 90F3 200A 2105 CB2C 0A21 BEDE 0011 74F1 9F15 BA29 0250 033F	.2 ..)% ..\$..! ..! ..t ..) ..P ..?
000000260	6932 DE21 6625 20EB 243C F32C 9082 21CA 2D82 2182 39E4 6F00 DE01 0486 F32C ACC7	i2 ..!f% ..\$..! ..- ..! ..9 ..o ..
000000280	OC23 CE29 E231 DE21 39CA 1324 A8CB 340C 6F00 5E3A 0A1C 5051 1272 3A1F 4056 2141	# ..) ..1 ..! ..\$..4 ..o ..^ ..PQ ..r ..@VIA ..
0000002A0	1324 6C6F 0011 1D3B 4822 320F 4C80 A73F 6542 DE21 D517 1324 30CB 3450 6F00 1040	\$..lo ..;H"2 ..L ..?eB ..! ..\$0 ..4Po ..@ ..
0000002C0	FBBD E806 1100 0A08 00F0 FF5E 0000 C05C 0000 ECFA A340 E8A3 4002 0708 0090 FD44	..^ ..W ..@ ..@ ..D ..
0000002E0	A000 4704 0858 4304 FB08 E0DE 2054 1000 0098 FE46 2080 7304 0800 7404 9B08 D442	.G ..XC ..T ..F ..s ..t .._B ..

Searching for strings

```
/cygdrive/d/Downloads
uFw-'X2oDaR
: inhibyib
playlisWt!!~
CL?USTERSn
ing byt*!3
BC s_To'tal
HD ER?ROR in
[found dsir
RM lba zid"
AhqCI-D Xw
~qBcopied
p_Pa?thMax_
0id_eoEndJ
ayu0<.qper >q>
geStream?Header
v^r2nd 3rUdbu
@U$gI"h_d PC
:p0>tI>qX0
uld be <uZ
kP kP4kPU<kP,kP0kP
KPGET:Uwmax
k0~I~q81I<J
LINKw TO:P
zeichnis> d lesen
:
```

- } Strings are weirdly distorted
- } 1 byte trash in the Middle of Strings
- } What could this mean???
- } What to do next?
→ Calculate the Entropy

Entropy divided in 256 byte blocks



→ File is possibly Compressed

Refining the Header Structure

	FileType	Uncompressed Size	Compressed Size	Checksum	
00000000	5A41	5A41	B41E	0900	440C 0500 114F 5702 AD49 4D46 A55E 4C96 D2BE 9A44 C5AE 49A4 ZAZA ..
00000020	65B4	B210	4FDD	A008	75D5 B6B0 4863 00A1 5C49 4D43 BB45 5CF6 D245 79D4 A1B2 4D64 e...0..
00000040	CDB1	4978	1DAO	6AB0	AAB5 49A1 B903 A1B1 A8B8 B848 4B41 50A9 C1F1 5984 A4B8 5167 ..Ix..
00000060	4841	7E5C	8B55	45A1	AA0A 534F 4077 0F41 5AB6 4D4A 4741 CC53 3241 30E4 1E4D 285D HA~W.UE
00000080	58A5	D2BD	214F	AA42	554A B921 49BB C845 ED41 5DD2 4049 DCDD 3135 B3CF 45D5 C654 X...!O..
000000A0	C454	4D58	9BBC	2443	51AA 1DAD 58A3 5DFE B749 AD06 A55D FF54 3EFE A857 A5C3 AA5A .TMX..S
000000C0	B7A5	45A7	AF6F	A067	B555 B64B A7A3 DFA4 6AB6 4D46 146C 33BF 3EA1 E9A5 09BD 9155 ..E..o..
000000E0	B3E8	A556	BF2B	A44B	95B5 AEA7 4843 F2A0 2AOE F689 0263 4548 CE6E BC42 C5AD B65B ...V.+..
00000100	5D45	B5AB	495B	435A	0DD9 CCB0 AD42 D9E1 B143 CD01 BBA3 F4D6 28A0 4848 5D32 AA58]E..![(
00000120	4C55	571C	EC46	45B4	D5D0 B64E C8CB B041 ADC9 C7AC CCE1 ECB5 22BE E5CE A747 490C LUW..FB
00000140	AD52	B258	4F98	45CB	F5A9 15B0 E6E3 B1FB D240 ACF9 BBC1 85B7 35B1 17A6 45BB 2E44 .R.XO.B
00000160	40D6	4958	EFCA	A043	EBC5 41D7 AD42 F2CB DB59 C5B7 DD55 46B4 81E1 80A5 B5B8 EBC4 @.IX..
00000180	5DCF	1448	D7A7	5ACB	A842 3E5D B6EC F2D2 AFE6 AAE9 C740 3B46 4F50 48E3 E5B2 AAA9]..H..Z

} Compression Ratio : 1.8

Choosing from the candidates

- } Run Length Encoding
 - } Arithmetic coding
 - } Huffman
 - } LZ77
 - } LZ78
- ~~} Run Length Encoding
(Repeating bytes exist)~~
- ~~} Arithmetic coding
(No text remains)~~
- ~~} Huffman
(No text remains)~~
- ~~} LZ77~~
- ~~} LZ78~~



A little intro on LZ77 Compression

- } the_rain_<3,3>Sp<9,4>falls
- } <3,3> : go 3 bytes back, copy 3 bytes from there and paste at current location.

- } The_rain_in_Sp<9,4>falls
- } <9,4> : go 9 bytes back, copy 4 bytes from there and paste at current location.

- } The_rain_in_Spain_falls
 - Compressed bytes info data(CBID) contains <last occurred matching sequence position, length>

Recognizing Patterns

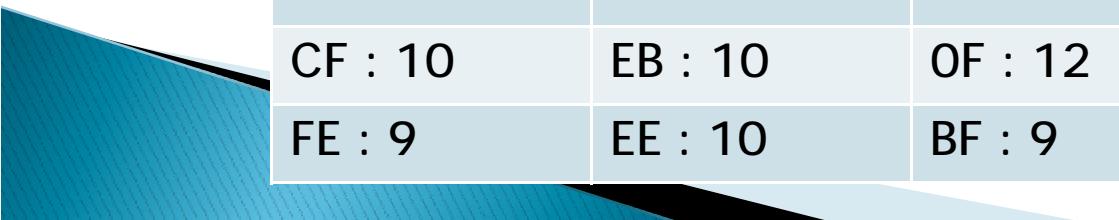
- } Every n bytes there exists a **Header byte**
 - } HeaderByte probably tells the number of bytes until the next HeaderByte, and the position of the CBID(s)

U2BD 9890 2450 U598 0/41 7000 E74b C38E 01A9 7200 B5B8 8BAU 1E73 38C6 99DU b41 | . . \$P . . Ap . . F . . r . . . s8 . . VA
E6F1 52FF 6573 756D 6520 696E FF20 706C 6179 6C69 7367 7421 217E 9151 D300 84F7 . . R . resume in . playlisWt!!~.Q . . .
905F C403 1408 8CF7 9044 8390 5590 F790 4883 9094 F790 4C83 90E5 98F7 9088 F790 | P . . H . .

- } HeaderByte 0xFF means the next 8 bytes are uncompressed

Trying to guess the meaning of the Header Bytes

- } We know that the Header byte contains the following information :
 1. The number of bytes until the next Header
 2. The position of the CBID(s)
- } Start attacking no.1 theory by collecting samples.



```
71DE 0558 52F1 0E21 E2A0 D905 58E7 E2A0 D505 583E 1141 5243 FD48 F210 4A75 6B65 q.,XR,,!....X.....X>,ARC,H,,Juke
626F FF78 00C0 4D75 6C74 69DF 6D65 6469 61D4 A04F 57BF 4552 4F46 465A B122 6220 bo,x, Multi media..OW,ER0FFZ,"b
FA27 63C5 C310 6966 0308 7BFE 3F60 A112 0308 B7AF 03FF 0885 0400 00A0 0503 BF00 .'c...if...{.?.....+....%
0000 0000 9071 FF00 0000 0000 90F5 9002 FR08 RA99 9074 CR13 0825 FFFF 0208 ..n + %
```

FB : 9	EF : 9	BE : 10	DF : 9
FF : 8	EA : 11	AB : 11	AF : 10
CF : 10	EB : 10	OF : 12	F9 : 10
FE : 9	EE : 10	BF : 9	FD : 9

Guessing

- } Compile Header Bytes with the same length(9)

FB : 9	EF : 9	DF : 9	FE : 9	BF : 9
--------	--------	--------	--------	--------

- } Use your brain and think. 'Why do all those different bytes end up in the same length?'
- } Make assumptions.
(Length info is contained in High nibble | Low nibble | bitfields | etc...)
- } Verify your theory by testing it on other samples.

Answer : Bitfields

- } Bit 0 means length = 2
- } Bit 1 means length = 1

ex) 0xBF = 10111111 = $7*1 + 1*2 = 9$

0xAC = 10101100 = $4*1 + 4*2 = 12$

- } Moreover, 0 is a **CBID**, while 1 is a **normal uncompressed byte**. Read from right to left.

2A0 D905 58E7 E2A0 D505 583E 1141 5243 F048 F210 4A/5 6B65 |q..XR...!....X.....X>.ARC.H.
3C74 69DF 6D65 6469 61D4 A04F 579F 4552 4F46 465A B122 6220 |bo,x..Multi media..OW,EROFF:
1A0A 7RFF 9F00 A112 030A R7AF 0AFF 0885 0400 00A0 050A RF00 |'c if f ?'

- } 0xDF = 11011111
- } = 5 uncompressed , CBID, 2 uncompressed
- } "media", ~~xD4~~~~xA0~~, "OW"

Confirm your theory by writing a parser/verifier



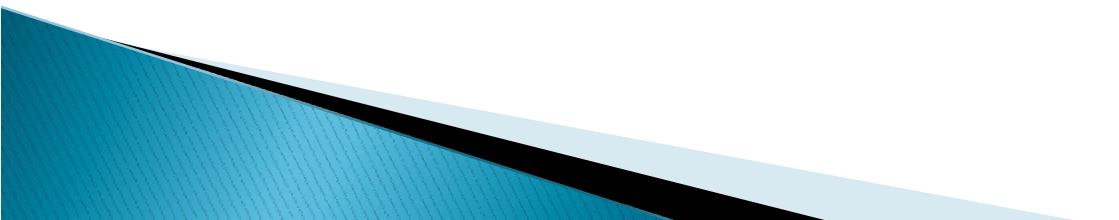
A screenshot of a debugger interface. At the top, there's assembly code with several bytes highlighted in red: EB78, 6D65, 579F, and 4552. Below the assembly is a memory dump table. The table has four columns: address, value, hex, and ASCII. The first few rows show memory starting at address 00011FC0. The row at address 6770 is highlighted with a light purple background, and the value column shows '0x11fc2' which is also highlighted with a red box.

6769	0x11fb8	0xfd	0x9
6770	0x11fc2	0xff	0x8
6771	0x11fcbb	0xdf	0x9
6772	0x11fd5	0x9f	0xa
6773	0x11fe0	0xfa	0xa

- } Offsets found! **J**
- } Verify for 5 or more header bytes just to make sure.
- } The theory is correct.
- } Something new learned : CBID is 2 bytes

Trying to understand the meaning of the CBID

- } What we already know about the CBID :
 1. Contains the size of the compressed bytes
 2. Contains the position of the compressed bytes
- } Start attacking no.1 theory by collecting samples.



Figuring out which byte contains the length field

- } "Firmware Version" ≈ "Firmware Version"
 - } ∴ C8B0 → 3 bytes

A988 03FF 0842 7269 6768 746E 7B65 7355 884C 756D 69BE 51FA F7F9 487B 8069 676B BrightnesU.Lumi.Q...Higk

- } “H{.igkei” = “Helligkei”
} ∴ 7B80 → 3 bytes

```
5 ..Pi.tch.,~o.OonAh0@.o.o.n.1..p.  
6 @R.0mf Br.ight.KLum.inos..Hel.  
7 ligkei.Yr0i1loP<.ua...?..q./0.7  
8 C.oror.<Cou.leur{Farb..=n...>..  
9 $ Video3!_pA
```

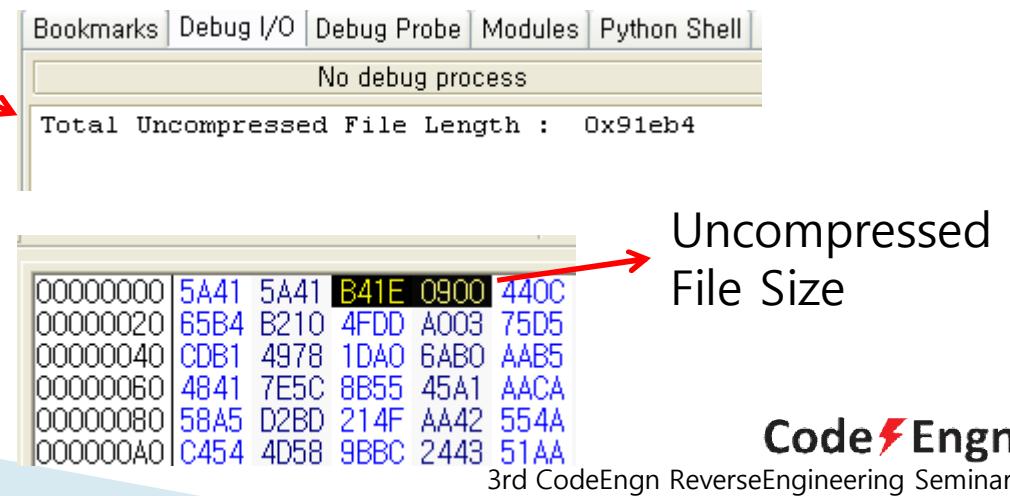
1D04 08E6 9143 7265 FD61 B3A0 466F 8064 6572 FEC2 C543 72E9 6572 | Cre.a..Folder ..

- } “Cre.a..Folder” = Create Folder
} ∴ B3A0 → 3 bytes

Guessing

- ∴ C8B0 → 3 bytes
- ∴ 7B80 → 3 bytes
- ∴ B3A0 → 3 bytes
- } The last nibble is all 0 for all CBIDs.
- } (Compressed bytes length = Last Nibble + 3) ?
- } Verify by parsing the file using that assumption.

- } Matches ! ↴
- } Theory is Correct.



**Uncompress the whole file using
the newly found length field**

0008F400	0044	E970	6C00	0000	0000	0065	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	D.pl.....e...
0008F420	0000	6965	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	.ie.....1.....
0008F440	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	311C	0400	0000	0000	0000	
0008F460	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	
0008F480	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	
0008F4A0	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	
0008F4C0	0000	0000	0000	0000	0000	6700	0000	0000	F8B3	1308	3C4D	5083	3E00	0000	0000	0000	
0008F4E0	004A	5047	0000	0000	0041	5649	3E00	C046	C500	0000	00B4	1308	5365	7475	7000	0000	JPG.....AVI>..F.....Setup.....
0008F500	0000	0000	0000	0000	4272	6F77	7300	0000	4C89	6272	6172	7900	4564	6900	0000	0000	Brows...Library.Edi...
0008F520	52E9	676C	0000	0000	0000	0000	0000	0000	4E61	7669	6700	0000	0000	0000	0000	0000	R.gI.....Navig.....
0008F540	0000	0000	0000	4D65	6EFC	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	Men.....
0008F560	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	
0008F580	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	
0008F5A0	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	F.....
0008F5C0	9800	0000	D800	0000	FC55	0808	4857	0808	9458	0808	E059	0808	2C5B	0808	785C	0808	U..HW..X..Y..[..x#..
0008F5E0	C45D	0808	105F	0808	5C60	0808	A861	0808	F462	0808	4064	0808	8C65	0808	D866	0808].....#`..a..b..@d..e..f..
0008F600	2468	0808	7069	0808	BC6A	0808	086C	0000	0000	546D	0000	0000	A06E	0808	\$h..pi..j..l..Tm..n..		
0008F620	EC6F	0808	3871	0808	8472	0808	D073	0808	1C75	0808	6876	0808	B477	0000	0079	0000	o..8q..r..s..u..hy..w..y..

- } The compressed bytes are still unknown and are filled with zeroes

Guessing the remaining 3 nibble field meaning

- } We know that it represents the position of the last occurred matching bytes
- } We know by the length field that the minimum length of the matching bytes is 3
- } Collect samples where you can guess the compressed bytes, in which the bytes identical to the resulting bytes are located somewhere before the compressed bytes, where you can see them.

Collecting Samples (1)

- } "a...save" = "and save"
 - } Compressed bytes = "nd "
 - } Last occurred at 0x3826A and 0x48 bytes before the compressed bytes
 - } CBID position field = 0x258

Collecting Samples (2)

00005340	00D1	9BE4	3E00	0000	0000	00D8	96E4	3F00	0000	0000	03D1	3E00	0000	0000	0FD0	0000	..>.....?....>.....
00005360	0000	0000	0000	0000	0000	08D1	1700	0000	0000	00D0	2E00	0000	0C00	0000	68E4	3000h.0.
00005380	0000	6430	3000	0000	78E4	C046	0DDA	0208	0000	0000	2D2D	2D3E	0000	0000	0000	0000	.d00..x.F...-->.....
000053A0	2045	5252	0000	0000	0000	0000	0000	0043	4F4E	5452	4F4C	2044	4953	0000	004E	ERR.....CONTROL.....DIS.....N	
000053C0	4543	5445	4400	0000	4C00	0000	BC00	0000	E800	0000	25BB	0308	49E3	0008	3000	0000	ECTED.....L.....%....I....0....
000053E0	5800	0000	51C5	0108	3C00	0000	2000	0000	2400	0000	E5F4	0008	15C3	0008	7D20	0408	X...Q..<....\$.....}
00005400	1222	2122	2522	2122	1552	2222	2222	2122	2122	2222	2122	2222	2122	2222	2122	2222	5

- } "DIS...NECTED" = "DISCONNECTED"
 - } Compressed bytes = "CON"
 - } Last occurred at 0x53B1 and 0xB bytes before the compressed bytes
 - } CBID position field = 0x39F

Collecting Samples (3)

```
10|.....0.....E.....  
18|already exists|R..W..1m...  
10|g..l..0...j..T..
```

2000 011D 011E 0200 01 03 0403 0001 0142 4000 0200 0000 0000 0000 |u.....w..d.....
061C 03D0 10C1 7E90 C1D5 9801 30D5 906E B11F C0DC 0BE3 450C D064 zb.u.....~....0..n....E..d
6057 7064 7920 65FF 7869 7374 7300 C052 FA12 F057 12F0 316D 0408 ... a!Wpdy e.xists..R..W..1m..
6703 087C DFFB 1308 308E 60E0 6A08 EF08 549C 1365 82AC A0B9 7E99g..|..0`j...T..e...~.
F742 11EA 67AA 8206 8EAD 4E01 1CE3 A0ED DDAB 180C 900C 980D 90FF ..C..~Dh..~E

- } "al...dy exists" = "already exists"
 - } Compressed bytes = "rea"
 - } Last occurred at 0x13769 and 0x6B9 bytes before the compressed bytes
 - } CBID position field = 0x757

Analyzing the samples

- } $0x258 \neq 0x3826A, 0x48$
- } But $0x258$ is **very close** to $0x26A$!
- } In fact, $0x757$ is also very close to $0x769$, and so is $0x39F$ and $0x3B1$
- } Carefully observing the difference between the 2 numbers, we realize that the difference is **constant** !
- } The difference between the two numbers is **$0x12$** , therefore, the format of the CBID is
- } <Offset of the last occurred identical byte sequence - $0x12$ (low 3 nibbles), Length – 3>

File Parsing Script

- Using the Above facts, uncompress the entire file and dump it into a new file.

```
uncompressed_filesize = 0
index = 0

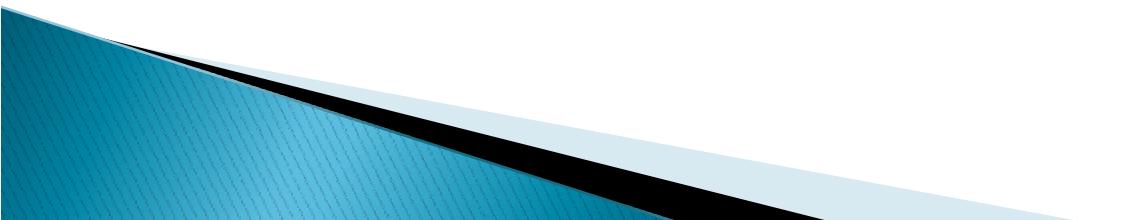
while index < filesize:
    headerbyte = struct.unpack("B",data[index])[0]
    savedheaderbyte = headerbyte
    chunksize = 0
    for i in range(8):
        if (headerbyte & 1) == 0:
            compress_code = struct.unpack("H",data[index+chunksize+1:index+chunksize+3])[0]
            compressed_bytes_size = ((compress_code & 0xFOO) >> 8) + 3
            compressed_bytes_offset = (compress_code & 0xFF) + ((compress_code & 0xF000) >> 4) + 0x12
            saved_offset = outfile.tell()
            compressed_bytes_offset += (saved_offset & 0xFFFFF000)
            if compressed_bytes_offset >= saved_offset:
                compressed_bytes_offset -= 0x1000
            outfile.seek(compressed_bytes_offset)
            uncompressed_bytes = outfile.read(compressed_bytes_size)
            outfile.seek(saved_offset)
            outfile.write(struct.pack("%ds"%compressed_bytes_size,uncompressed_bytes))
            chunksize += 2
        else:
            outfile.write(data[index+chunksize+1])
            chunksize += 1
    if index+chunksize+1 == filesize:
        break
    headerbyte = headerbyte >> 1
    index += chunksize + 1
print hex(uncompressed_filesize)
outfile.close()
```

Fully Uncompressed File

0008F160	702E	0043	6F70	7920	416C	6C20	746F	2048	4400	0000	0000	436F	7079	2041	6C6C	p..Copy All to HD.....Copy All to HD.....
0008F180	2074	6F20	4844	0000	0000	0000	0000	0000	1DC4	0308	D001	0000	CCB0	1308	0400	0000
0008F1A0	5374	6172	7420	506C	6179	6C69	7374	0000	0000	0000	0044	E96D	6172	7265	7220	6C61
0008F1C0	206C	6973	7465	0000	0000	506C	6179	6C69	7374	2073	7461	7274	656E	0000	0000	0053
0008F1E0	7461	7274	2050	6C61	796C	6973	7400	0000	0000	0000	5374	6172	7420	506C	6179	6C69
0008F200	7374	0000	0000	0000	0000	2DBB	0208	0000	0000	5361	7665	2050	6C61	796C	6973	
0008F220	7400	0000	0000	0000	0053	6175	7665	672E	206C	6120	6C69	7374	6500	0000	0000	506C
0008F240	6179	6C69	7374	2073	7065	6963	6865	726E	0000	0053	6176	6520	506C	6179	6C69	7374
0008F260	0000	0000	0000	5361	7665	2050	6C61	796C	6973	7400	0000	0000	0000	0000	0000	0000
0008F280	45BF	0208	0000	436C	6561	7220	506C	6179	6C69	7374	0000	0000	0000	0045	6666	
0008F2A0	6163	6572	206C	6120	6C69	7374	6500	0000	0000	506C	6179	6C69	7374	206C	F673	6368
0008F2C0	656E	0000	0000	0043	6C65	6172	2050	6C61	796C	6973	7400	0000	0000	0000	436C	6561
0008F2E0	7220	506C	6179	6C69	7374	0000	0000	0000	0000	75BB	0208	0000	0000	5368	7566	
0008F300	666C	6520	506C	6179	6C69	7374	0000	0000	004C	6973	7465	2061	6C69	6174	6F69	7265
0008F320	0000	0000	0000	5A75	6661	6C6C	732D	506C	6179	6C69	7374	0000	0000	0053	6875	6666
0008F340	6C65	2050	6C61	796C	6973	7400	0000	0000	5368	7566	666C	6520	506C	6179	6C69	7374
0008F360	0000	0000	0000	B7BB	0208	5C01	0000	9C82	1308	0300	0000	4164	6420	5472	6163	
0008F380	6B00	0000	0000	0000	0000	0041	6A6F	7574	6572	2075	6E20	7469	7472	6500	0000	
0008F3A0	0000	5469	7465	6C20	4869	6E7A	7566	FC67	656E	0000	0000	0041	6464	2054	7261	636B
0008F3C0	0000	0000	0000	0000	0000	4164	6420	5472	6163	6B00	0000	0000	0000	0000	0000	
0008F3E0	0000	0000	911B	0408	0200	0000	4D6F	7685	2054	7261	636B	0000	0000	0000	0000	0000
0008F400	0044	E970	6C61	6365	7220	6C65	2074	6974	7265	0000	0000	5469	7465	6C20	5665	7273
0008F420	6368	6965	6265	6E00	0000	004D	6F76	6520	5472	6163	6B00	0000	0000	0000	0000	
0008F440	4D6F	7665	2054	7261	636B	0000	0000	0000	0000	0000	0000	311C	0408	0000	0000	
0008F460	4465	6C65	7465	2054	7261	636B	0000	0000	0000	0000	0053	7570	7072	696D	6572	206C
0008F480	6520	7469	7472	6500	0000	5469	7465	6C20	4CF6	7363	6865	6E00	0000	0000	0000	0044
0008F4A0	656C	6574	6520	5472	6163	6B00	0000	0000	0000	0000	4465	6C65	7465	2054	7261	636B
0008F4C0	0000	0000	0000	0000	0000	671C	0408	1600	0000	F8B3	1308	3C4D	5033	3E00	0000	
0008F4E0	3C4A	5047	3E00	0000	3C41	5649	3E00	C046	C500	0000	1084	1308	5365	7475	7000	0000
0008F500	496E	666F	0000	0000	4272	6F77	7365	0000	4069	6272	6172	7900	4564	6974	0000	0000
0008F520	52E9	676C	2E00	0000	496E	666F	0000	0000	4E61	7669	672E	0000	4C69	6272	6172	7900
0008F540	4564	6974	0000	0000	4D65	6EFC	0000	0000	496E	666F	0000	0000	5469	7465	6C85	0000
0008F560	4C69	6272	6172	7900	4564	6974	0000	0000	5365	7475	7000	0000	496E	666F	0000	0000
0008F580	4272	6F77	7365	0000	4C69	6272	6172	7900	4564	6974	0000	0000	5365	7475	7000	0000
0008F5A0	496E	666F	0000	0000	4272	6F77	7365	0000	4069	6272	6172	7900	4564	6974	0000	4600

} All strings are neatly displayed! Which means the whole file is uncompressed. J

Questions



CodeEngn

3rd CodeEngn ReverseEngineering Seminar

Thank You

안기찬 ; Ahn Ki-Chan

winger@paran.com

www.Externalist.org

