Handycipher: a Low-tech, Randomized, Symmetric-key Cryptosystem

Bruce Kallick Curmudgeon Associates Winnetka, IL 60093 curmudgeon@rudegnu.com

Handycipher is a low-tech, randomized, symmetric-key, stream cipher, simple enough to permit penand-paper encrypting and decrypting of messages, while providing a significantly high level of security by using a nondeterministic encryption procedure, multiple encryption, and randomly generated session keys.

1.	Introduction	1
2.	The core cipher	2
3.	Example encryption with the core cipher	6
4.	Handycipher	10
5.	Example encryption with Handycipher	11
6.	Extended Handycipher	12
7.	Example encryption with Extended Handycipher	13
8.	Cryptanalytic vulnerability	16
9.	Challenge cryptograms	18
10.	Implementation notes	18
Ref	erences	19
App	pendix 1	20
App	pendix 2	21
App	pendix 3	22

1. Introduction

For several thousand years cryptography was concerned largely with developing various kinds of substitution and transposition ciphers which, through sharing a manageably sized secret key, permitted easy encryption and decryption of messages using nothing more than pen and paper. This has all changed, of course, within our lifetime and now with public key cryptosystems, employing massively powerful computers, so-called hand ciphers are for the most part interesting only to historians and hobbyists.

Yet one can conceive of circumstances in which a highly secure pen-and-paper cipher would be invaluable; for example, someone needing to send or receive a secret message might not have access to a secure computer, or might need to refrain from using one to avoid arousing suspicion that messages are being exchanged secretly. Indeed, Bruce Schneier, a cryptographer and fellow at Harvard's Berkman Center, designed the Solitaire cipher [7] used in the novel *Cryptonomicon* for such a scenario.

Moreover, apart from any consideration of potential real-world applications, it is an interesting challenge to explore how much security against a large-scale computer-based cryptanalytic attack can be achieved using nothing more than a few hours of effort with pen and paper. The problem of designing such a cipher has received little attention in the recent cryptographic literature, and Schneier's Solitaire is widely regarded as the best serious attempt to deal effectively with this problem yet to have been devised. In

this paper we describe a cipher which compares favorably in that it is somewhat easier to implement by hand, is less subject to error propagation, and needs no additional equipment besides pen and paper (unlike Solitaire which requires an ordered deck of cards).

In his seminal 1949 paper which heralded the emergence of modern cryptography, Shannon [8] observed:

...we can frame a test of ciphers which might be called the acid test. It applies only to ciphers with a small key (less than, say, 50 decimal digits), applied to natural languages, and not using the ideal method of gaining secrecy. The acid test is this: How difficult is it to determine the key or a part of the key knowing a small sample of message and corresponding cryptogram? [...] Note that the requirement of difficult solution under these conditions is not, by itself, contradictory to the requirements that enciphering and deciphering be simple processes.

In this spirit, then, the cipher described in this paper is proposed as a candidate for a modern formulation of Shannon's acid test. Using a 165-bit key (just small enough to fit Shannon's definition although larger than the Advanced Encryption Standard 128-bit minimum key size), Handycipher incorporates a nondeterministic encryption procedure along the lines described by Rivest and Sherman [6], and employs multiple encryption as suggested by Merkle and Hellman [5], as well as a randomly generated session key for each message. Combining a simple 31-character substitution cipher with a 3,045-token nondeterministic homophonic substitution cipher results in a novel system which, while quite easy to implement by hand, confers enough complexity to the relationship between ciphertext and plaintext and that between ciphertext and key to achieve a significant level of computational security against both statistical analysis and known-plaintext, chosen-plaintext, and chosen-ciphertext attack models.

The basic approach of the cipher is to take each plaintext character, convert it to a keydefined pattern of length five and, using this pattern as a template with one to five holes, select certain ciphertext characters from a 5 x 5 key-defined grid.

2. The core cipher

Handycipher is based on a core cipher which operates on plaintext strings over the ordered 31-character alphabet A

```
A = \{A B C D E F G H I J K L M N O P Q R S T U V W X Y Z , . - ? ^ \}
```

and generates ciphertext strings over A*, the same alphabet together with the ten decimal digits 0-9.¹ Some permutation of the 41 characters of A* is chosen as the secret shared key K, say for example,

Z D B 9 H A ? G V 8 1 J M T O U K - Y 5 0 Q 4 L ^ W F E R 6 I N . C , 7 2 X S 3 P

The 40 non-space characters of K are displayed as a 5×8 table, T_{K}

¹ It's important, of course, to be able to distinguish the digits 0 and 1 from the letters 0 and 1.

Z	D	В	9	Н	Α	?	G
v	8	1	J	М	Т	0	U
к	-	Y	5	0	Q	4	L
w	F	E	R	6	I	Ν	•
С	,	7	2	Х	S	3	Р

A 31-plaintext-character subkey P is derived from K by omitting the decimal digits Z D B H A ? G V J M T O U K - Y Q L ^ W F E R I N . C , X S P

and is displayed as a substitution table, ξ_P

m: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z , . - ? ^ $\xi_{P}(m)$:5 3 27 2 22 21 7 4 24 9 14 18 10 25 12 31 17 23 30 11 13 8 20 29 16 1 28 26 15 6 19

Then, by referring to T_K and ξ_P , plaintext characters are encrypted into k-tuples of ciphertext characters by means of the following scheme:

Regarding the first five columns of T_K as a 5 x 5 matrix comprising five rows, five columns, and ten diagonals, each plaintext character m is encrypted by first expressing $\xi_P(m)$ as a five digit binary number $b_1b_2b_3b_4b_5$ and by using the position of the 1's in this number as a pattern, associating the plaintext character m with a subset of the ciphertext characters comprising a randomly chosen row, column, or diagonal. Then a randomly chosen permutation of that subset is taken as the corresponding k-tuple of ciphertext characters.

For example, the plaintext character \cup occupying position 21 = 10101 is encrypted into one of the six permutations of one of the twenty 3-tuples

{ZKC D-, BY7 952 HØX ZBH V1M KYØ WE6 C7X ZYX D5C BØ, 9K7 H-2 Z5, DØ7 BK2 9-X HYC}

whereas the plaintext character Moccupying position 10 = 01010 is encrypted into one of the two permutations of one of the twenty 2-tuples

{WW 8F 1E JR M6 D9 8J -5 FR ,2 8R 16 JW MF VE ME VR 86 1W JF}

This roughly sketched scheme is now defined more precisely as follows.

A plaintext message M is encrypted into a ciphertext cryptogram C using a 41-character key K by means of the encryption algorithm E defined as follows:

Core cipher encryption algorithm: $C \leftarrow E(K,M)$

First, omitting ^ the remaining 40 characters of K are displayed as a 5 x 8 table T_K by writing successive groups of eight characters into the five rows of the table.

The first five columns of T_K comprise a 5 x 5 square array (or matrix) M_K and the rows, columns, and diagonals of M_K are designated R_1-R_5 , C_1-C_5 , and D_1-D_{10} , respectively. We refer to them collectively as *lines*, and call two characters colinear if

they lie in the same line. The 15 characters comprising columns C_6-C_8 are said to be *null characters*.

Also, a 31-character *plaintext-subkey* P is derived from K by omitting the ten decimal digits, and a simple (numerical coding) substitution ξ_P is applied, transforming each character m of M into the number $\xi_P(m)$ representing its position in P (i.e., if P = $p_1p_2...p_{31}$ then $\xi_P(m) = i$ where $m = p_i$).

Then the following three steps are applied in turn to each character m of M.

- 1. A random choice is made (with equal probability of each of the 20 possible rows, columns or diagonals) between:
 - 1.1. *Column-encryption*: One of the five columns in M_K, say C_j, is randomly chosen (with equal probability), or
 - 1.2. *Row-encryption*: One of the five rows in M_K, say R_j, is randomly chosen (with equal probability) subject to the following three restrictions, where \hat{m} denotes the character immediately following m in M, $\xi_P(m) \neq 1, 2, 4, 8, \text{ or } 16$ $\xi_P(\hat{m}) \neq 2^{5 \cdot j}$, if the position of the character \hat{m} in M is an odd number $\xi_P(\hat{m}) \neq 2^{j-1}$, if the position of the character \hat{m} in M is an even number or
 - 1.3. *Diagonal-encryption*: One of the ten diagonals in M_K , say D_j , is randomly chosen (with equal probability) subject to the restriction that $\xi_P(m) \neq 1, 2, 4, 8$, or 16.
- 2. $\xi_P(m)$ is expressed as a five digit binary number, $b_1b_2b_3b_4b_5$, and if the position of the character m in M is an odd number, then
 - 2.1. If 1.1 was chosen in step 1, then for each i such that $b_i = 1$, the i-th element of C_j is chosen, yielding a subset of the five characters comprising C_j , or
 - 2.2. If 1.2 was chosen in step 1, then for each i such that $b_i = 1$, the i-th element of R_j is chosen, yielding a subset of the five characters comprising R_j , or
 - 2.3. If 1.3 was chosen in step 1, then for each i such that $b_i = 1$, the i-th element of D_j is chosen, yielding a subset of the five characters comprising D_j .

but if the position of the character m in M is an even number, then

- 2.4. If 1.1 was chosen in step 1, then for each i such that $b_i = 1$, the (6-i)-th element of C_j is chosen, yielding a subset of the five characters comprising C_j , or
- 2.5. If 1.2 was chosen in step 1, then for each i such that $b_i = 1$, the (6-i)-th element of R_j is chosen, yielding a subset of the five characters comprising R_j , or
- 2.6. If 1.3 was chosen in step 1, then for each i such that $b_i = 1$, the (6-i)-th element of D_j is chosen, yielding a subset of the five characters comprising D_j .²

² Thus for each successive plaintext character the process alternates between reading rows left-to-right or right-to-left and between reading columns and diagonals top-down or bottom-up.

- 3. The elements of the subset specified in Step 2 are concatenated in a randomly chosen order. If this string, composed of 1 to 5 ciphertext characters, satisfies both of the following two restrictions, where \overline{m} denotes the character immediately preceding m in M, then it is taken as $\sigma(m)$. Otherwise, Step 1 is restarted.³
 - 3.1. The first character of $\sigma(m)$ must never lie in the line used to encrypt \overline{m} (although it may be either colinear or non-colinear with the last character of $\sigma(\overline{m})$).
 - 3.2. If $\xi_{P}(\overline{m}) = 1, 2, 4, 8$, or 16 then the first character of $\sigma(m)$ must be non-colinear with the single character of $\sigma(\overline{m})$ (which is a stronger requirement than 3.1).

Finally, the strings produced in Step 3 for each character of M are concatenated forming C.

As a result of the restrictions contained in Steps 1 and 3, the resulting ciphertext cryptogram C, consisting of the string $\sigma(m_1)\sigma(m_2)\sigma(m_3)$... can be unambiguously decrypted into the plaintext message $M = m_1m_2m_3$... by means of the decryption algorithm D defined as follows:

Core cipher decryption algorithm: $M \leftarrow D(K,C)$

C is divided into contiguous groups of characters, proceeding from left to right, at each stage grouping as large an initial segment of the remaining ciphertext as possible composed of colinear characters of M_K , then inverting the association between binary numbers and subsets of column, row, or diagonal elements invoked in step 2 of the encryption algorithm, and finally decoding that number by inverting the substitution ξ_P .

Thus each plaintext character m is encrypted by randomly choosing a line of the key matrix M_K and representing that character's numerical code $\xi_P(m)$ by an n-tuple $\sigma(m)$ of characters lying in the chosen line. So that in decryption it will be possible to tell where one encrypted character ends and the next begins, $\sigma(m)$ is not allowed to begin with any character lying in the line chosen for $\sigma(\bar{m})$.

With any key, of the 31 characters comprising the plaintext alphabet A: five are mapped by step 3 into one of 5 length-1 ciphertext unigrams, ten are mapped by step 3 into one of 20 x 2! = 40 length-2 ciphertext bigrams, ten are mapped by step 3 into one of 20 x 3! = 120 length-3 ciphertext trigrams, five are mapped by step 3 into one of 20 x 4! = 480 length-4 ciphertext 4-grams, and one is mapped by step 3 into one of 20 x 5! = 2400 length-5 ciphertext 5-grams, resulting in a total of 3,045 possible cipher tokens.

³ It's fairly straightforward to show that some combination of choices made in Steps 1 and 3 satisfying all the restrictions must exist unless $\xi_P(m) \ge \xi_P(\bar{m}) = 16$ for two consecutive plaintext characters, which would require the two consecutive ciphertext characters to lie in the same row. Accordingly, for each key there are five bigrams which cannot be encrypted by the algorithm. (See Appendix 1.)

3. Example encryption with the core cipher

Although any permutation of the entire ciphertext alphabet can be chosen as K, the problem of remembering and secretly sharing it can be made easier by formalizing a way of generating the key from a more memorable key passphrase. The following method is designed to work well with Handycipher. The passphrase is processed from left to right, first replacing all spaces by the number of characters in the preceding word, and then (again proceeding from left to right) omitting all repetitions; then ^ and all other characters missing from the resulting string are appended in reverse order, i.e., in the order:

{9-0 ^ ? - . , ZYXWVUTSRQPONMLKJIHGFEDCBA}

A passphrase can be more easily communicated secretly than the key, for example by using, on the nth day of the year, the first fifty characters on the nth page of a previously agreed upon book. As a more fanciful example, the passphrase could be the first verse of a folk song, as in:

ON TOP OF OLD SMOKY, ALL COVERED WITH SNOW, I LOST MY TRUE LOVER FOR COURTING TOO SLOW.

which generates the key K

ON 2 T P 3 F L D S M K Y , 5 A C V E R 7 W I H 4 1 U G 8 . 9 6 0 ^ ? - Z X Q J B

and plaintext subkey P

ONTPFLDSMKY, ACVERWIHUG. ^?-ZXQJB

and associated table $T_{\ensuremath{K}}$

0	N	2	Т	Р	3	F	L
D	S	М	К	Y	,	5	Α
С	v	E	R	7	W	Ι	Н
4	1	U	G	8	•	9	6
0	?	-	Z	Х	Q	J	В

The substitution ξ_P can be written

m: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z , . - ? ^ $\xi_{P}(m)$:13 31 14 7 16 5 22 20 19 30 10 6 9 2 1 4 29 17 8 3 21 15 18 28 11 27 12 23 26 25 24 and the encryption process can be summarized as

A odd	13	01101	DCØ N2P	SV? SMY	ME- VE7	KRZ 1U8	Y7X ?-X	SEX YR?	MRØ D7-	K7? SCZ	YC- MVX	DVZ KEØ
A even	13	01101	4C0 T20	1VN KMD	UE2 REC	GRT GU4	87P Z-0	GEO URO	8RN G7N	472 8C2	1CT 4VT	UVP 1EP
В	31	11111	ODC40 ON2TP	NSV1? DSMKY				OSEGX OYRU?			TYC1- TMV4X	PDVUZ PKE10
C	14	01110	DC4 N2T	SV1 SMK	MEU VER	KRG 1UG	Y78 ?-Z	SEG YRU	MR8 D7G	K74 SC8	YC1 MV4	DVU KE1
D odd	7	00111	C40 2TP	V1? MKY	EU- ER7	RGZ UG8	78X -ZX	EGX RU?	R80 7G-	74? C8Z	C1- V4X	VUZ E10
D even	7	00111	ODC ON2	NSV DSM	2ME CVE	TKR 41U	PY7 0?-	OSE OYR	nmr ND7	2K7 2SC	TYC TMV	PDV PKE
E odd	16	10000	0	N	2	Т	Р					
E even	16	10000	0	?	-	Z	X					
F odd	5	00101	CØ 2P	V? MY	E- E7	RZ U8	7X -X	EX R?	RØ 7-	7? CZ	C- VX	VZ EØ
F even	5	00101	0C 02	NV DM	2E CE	TR 4U	P7 0-	OE OR	NR N7	27 2C	TC TV	PV PE

etc., where, in each row, the groups of characters comprising the rightmost ten columns are the subsets referred to in Step 2 of the encryption algorithm. In other words, A is randomly transformed into one of the six permutations of one of the twenty triples in either row 1 or row 2, depending on whether its location in M is odd or even; B is randomly transformed into one of the 120 permutations of one of the twenty quintuples in row 3; C is randomly transformed into one of the six permutations of one of the twenty triples in row 4; D is randomly transformed into one of the six permutations of one of the twenty triples in either row 5 or row 6, depending on whether its location in M is odd or even; E is randomly transformed into one of the five characters in either row 7 or row 8, depending on whether its location in M is odd or even; F is randomly transformed into one of the twenty doubles in either row 9 or row 10, depending on whether its location in M is odd or even; etc., subject to the restrictions specified in steps 1 and 3.

So, for example, the plaintext CATS AND DOGS can be encrypted as follows⁴:

⁴ In the middle column $\xi_P(m)$ is expressed in binary; in the fourth column the row, column, or diagonal chosen in Step 1 is indicated.

<u>m</u>	<u>ξ_P(m)</u>		<u>C/R/D</u>	<u>σ(m)</u>	
С	14	01110	R1	2NT	
Α	13	01101	С3	EU2	
Т	3	00011	D1	GX	
S	8	01000	C2	1	
۸	24	11000	R2	DS	
Α	13	01101	D1	OGE	(O chosen to be colinear with preceding S)
Ν	2	00010	C1	4	
D	7	00111	С3	E2M	
۸	24	11000	C5	PY	
D	7	00111	D10	KPE	(K chosen to be colinear with preceding Y)
0	1	00001	C2	?	
G	22	10110	С3	EM-	
S	8	01000	C4	К	

yielding the ciphertext

2NTEU2GX1DSOGE4E2MPYKPE?EM-K

Note that 0 could not have been chosen instead of ? for $\sigma(0)$ according to restriction 3.1. but - could have been, if a colinear character was called for. Similarly, neither -EM nor -ME could have been chosen instead of ME- for $\sigma(G)$ according to restriction 3.2. Also note that R₂ could not have been used to encrypt G for then it would have been impossible to encrypt the following S. Except for the second A and the second D, non-colinearity was chosen instead of colinearity.

The ciphertext would be decrypted by dividing it, according to the table T_K , into its constituent k-tuples and then finding each group's associated binary number, converting to decimal, and decoding by inverting the substitution ξ_P

n: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 {P-1(n):0 N T P F L D S M K Y , A C V E R W I H U G . ^ ? - Z X Q J B

For a slightly larger example consider the 230-character plaintext⁵

It haunts me, the passage of time. I think time is a merciless thing. I think life is a process of burning oneself out and time is the fire that burn-s you. But I think the spirit of man is a good adversary. --Tennessee Williams

which can be encrypted by the core cipher in $(63 \times 5) \times (104 \times 2 \times 20) \times (55 \times 6 \times 20) \times (5 \times 24 \times 20) \times (3 \times 120 \times 20) \approx 1.5 \times 10^{17}$ ways including, for example, this 471-character ciphertext:

ZØXSN DPR?E M-OXE 8DOM1 ?PNZ7 YZ8-G ØENUZ 7TO2D 1ZSCR KZPG8 -VP?Ø S21-T DKNK? 72DO1 480NØ ?MDØN MGY1M 2DP1Ø PCRNK YN8ØU S078P MN24N XOUYR 814E1 XD8DN KTØ-S YD8?X -84UG 7RXØZ GX1?M Y1?NK UMXGR GOØD8 UØTM2 K?MZX CSZ1Ø

⁵ A dash is included in the plaintext word "burn-s" because this choice of key does not allow the bigram NS to be encrypted (see footnote 3).

70P?D 7GPGM ?107P 0?EKP 1020P ?28TZ K8VDZ NMTUX K-RGP VOSP? VNTYD S40CG Y7PS2 Y0ZUP G4PG- ØZXUT YMTEC X14-0 U2-0? T8XTP MY?RY 2S?ØK P1ØE? VNYD1 R7VYZ G81GP RN02M -E410 40DCX 7PM2D NY-TD PCV27 OSX1M 4SNYT MDXDN 4E?SN XOZDP 4?8GØ TCNO8 2XY7S Ø?2SZ YØ1ZD VGXRE CNZND 7S01P EYKMN 8MR2X ØST1C P08S2 ØR8NU Z41ZD UN8MZ XR7Z1 D21D? P?4RC M2-8C KENG- TV4ZK 8 parsed as: I t ^ h a u nt s ^ m e , ^ t h e ^ p a s s a g e Z0X SN DP R? EM- OXE 8 DO M 1? PN Z 7Y Z8 -G 0E N UZ 7 TO2 D 1 ZSC RKZ P ^ of ^ t i m e . ^ I ^ t h i n k ^ t i m e ^ G8 - VP ?0 S2 1-T DK N K?72 DO 148 ON 0? MD 0NM G Y1 M2 DP 10P CR N KY i s ^ a ^ m er c i l es s ^ t h i ng . ^ I N80 U SO 78P MN 24 N XO UYR 814 E1 X D 8 DN KT 0- SYD 8 ?X- 84UG 7R X0Z ^ t h i n k ^ l i f e ^ i s ^ a ^ p r o c e s s ^ GX 1? MY 1?N K UM XG RG 00D 8U 0 TM 2K? M ZX CSZ 10 7 OP ? D7G P G M ?1 of ^ b rning ^ oneself ^ ou u t ^ 0 7P 0? EKP10 20P ?2 8 TZK 8 VDZ NM T U X K - RG PV OS P ?VN TY DS 40C G me^is^the^fire^tha d ^ t i Y7P S2 YO ZUP G4 P G- OZX U TY MT EC X 14 -0 U2- 0? T 8X TP MY ?RY 2S ?0 b u r n - s ^ y o u . ^ B u t ^ I ^ t h KP10E ?VN YD 1 R7V Y ZG 81G P RN0 2M-E 41 040DC X7P M2 DN Y-T DP CV 27 i n k ^ t h e ^ s p i r i t ^ o f ^ m a n ^ i s ^ OSX 1 M4 SN YT MD X DN 4 E ?SN XO ZDP 4? 8G 0 TC NO 82 XY7 S 0? 2SZ Y 01 ^ a ood d V er sa ZDV GX ŘEC N Z ND7 SO 1PE YKM N8MR 2 XØ S T1C PO 852 ØR8N UZ 41 ZDU N8M ^ T ennessee ^ W i 1 l i а m s ZX R7 Z 1 D 2 1 D ? P ?4 RC M2- 8C KE NG- TV4 ZK 8

Although the average bandwidth expansion factor averaged over all possible keys and all possible messages uniformly distributed, is

 $(5 \times 1 + 10 \times 2 + 10 \times 3 + 5 \times 4 + 1 \times 5) / 31 \approx 2.58$

for the example key above, noting the distribution of length-n expansions among the characters of A, namely

length-1: E N O P S
length-2: F H K L M R T W , ^
length-3: A C D G I U X Y - ?
length-4: J Q V Z .
length-5: B

and using the usual frequency distribution of these 31 characters in English, an average bandwidth expansion factor can be computed as:

 $1 \times 0.28 + 2 \times 0.45 + 3 \times 0.23 + 4 \times 0.02 + 5 \times 0.01 \approx 2.2$

while that of this particular encryption is $470 / 230 \approx 2.0$.

4. Handycipher

Although the core cipher affords a reasonable level of security when used to encrypt relatively short plaintexts, with increasing message length it becomes more vulnerable to statistically based hill-climbing attacks along the lines described by Dhavare, et al [3]. Indeed, an earlier version of Handycipher was broken by just such an attack [1][2]. However, the cipher can be made significantly resistant to such attacks by the simple expedient of randomly dispersing so-called *null characters*, the fifteen characters comprising the last three columns of T_K , as decoys throughout the ciphertext. This is accomplished according to the following encryption algorithm E^+ defined as follows:

Handycipher encryption algorithm: $C \leftarrow E^{\dagger}(K,M)$

This algorithm is identical to the core cipher encryption algorithm except that the final sentence

Finally, the strings produced in Step 3 for each character of M are concatenated forming C.

is replaced by the following text:

Finally, the strings produced in Step 3 for each character of M are concatenated forming C*, and then null characters are inserted throughout C* in a statistically-balanced manner producing the cryptogram C by the following process:

To create C, start with the stream of characters C*.

- With probability 5/8 insert the current character from C* into C and repeat from (1) considering the next character in C*. If there is no next character, still repeat from (1) and stop only when there is a demand for a non-null (i.e. be prepared to insert more nulls).
- (2) Instead choose to insert a null into C. This null N, should be randomly chosen from the set of 15, but potentially rejected in favor of another null by considering the current last six characters of C. If N last appears at a position n characters back from the end of C, that N should be rejected with probability (6-n)/5. This leads to 100% rejection at n=1, i.e. consecutive identical characters are not allowed. Once a null is inserted, repeat (1) with the same current character in C* as before, i.e. all characters in C* end up in C.

This process should ensure that each individual character in C (null or non-null) is roughly equally common and that nulls are not betrayed by repeating too often within a few characters. Non-null characters are suppressed in their ability to repeat by the algorithm given the presence of the colinear groups, which can be as long as five characters. The likelihood of a null being the first, last, or any other character is constant.

The corresponding decryption is simply accomplished as:

Handycipher decryption algorithm: $M \leftarrow D^{\dagger}(K,C)$

This algorithm is identical to the core cipher decryption algorithm except that the phrase

proceeding from left to right,

is amended to read:

proceeding from left to right and omitting null characters,

5. Example encryption with Handycipher

Continuing with the example in Section 3, encrypting the Tennessee Williams quotation with Handycipher instead of the core cipher might yield this 753-character ciphertext:

Z0XBS .IN26 S-7.M R60QW TZIR4 NB6OM 1W5?P NZLFY RXWZP T,FH8 UN5BZ XCN1H VYGQY CJ-?B K7T?Q 1X2EQ DISTM 6DQKY 32UNX .6WTV MOQY5 2W?KN B0149 RNX0F ?8DUT M06SW 8G4LN GP-6G C73R1 OASU. 2EW41 0I4P4 98EK1 2SFZ7 G9LFX K8BVQ CJOHS I34HU WKTJZ 1679Y X2TPO 89XQ- Q2UGA 8L?UX WQ-FR 5CIW. 37K5J VRXZQ 4V.2L M-P25 Z0JST KZMGJ 8EAL. FV71? EDY07 4K01M Z8BAX N,VFO QDEIU M2-89 U4,P0 UW6IT Z3AKU S,ZCP AD,3, 01BF5 ,XK9X C68VN A412R TBZIM 2DPH5 QRU,O B16WJ 2JM65 QES2Y 30Z6X 5I024 L.PGZ .1-T8 ?,0LA HS07J 2H0IY 9BLOV NZ05X 2?QP- 1X4P2 8L16Q UND2S 9LTWY Q13CJ .,-27 I?TXU 5,7VR 9QK5H BXI8Q K6?L3 9HI4N ,ALFC EQ7D- 7NVCG KRQZA TE7CF ,PJKC V0XSB ITKND .TRJS 05FXU HS86K QT,JS 0720. JLTY4 RJ2ZS 0PSFX 0KYQ- 1XI20 1W4PI 86EVQ 7SPB. QY419 8DQ8A G?F6. Y,LIR 36X4P 87A50 9ZEJ3 FCV3M N7BNQ LW,HG CA91B -04DC P?H2Y QFZ0A -T-HK 432WG ATRIZ 3L?12 MZDU9 WJ0?6 Z7,RX 4F-4A Y21DX .N5U? T59IL 45-3N ID5FR EY71T JFA6- 8PW.W AF.QJ 6W7K6 Q0G

parsed as:

It ^ ha u nt s^ me, ^ th e ZØX BS.IN 26S -7 .MR60 QWTZIR 4 NB60 M 1W5? PN Z LFYR XWZ PT ,FH8U N ٨ ٨ passa ge^of^time. 5BZX C N1HV Y G QYCJ- ?BK7 T ?Q1 X 2E QDIS TM 6DQKY 32U N X.6WTVM 0QY ^ th i nk ^ t i m e ^ i s ^ a ^ Ι 52W?K NBO 14 9RN X0F? 8 DU TM 06S W8G4 LNG P -6G C73R 1 0AS U.2E W41 erciless[^] thing.[^] OI4 P 498 EK1 2SFZ 7G 9LFX K 8 BVQC JOHS I34HU WKTJZ 1 679YX 2TPO 89X Ι ^thi nk ^li fe^is^ Q-Q2U GA8 L?U XWQ- FR5CIW.37 K 5JVR XZ Q4V .2LM- P2 5Z OJS TKZ M GJ8 a ^process^of ^ b u r ni EAL.FV7 1? E DY Ø 74K O 1 M Z8 BAX N,V FOQD EIUM2- 89U4 ,PØ U W6ITZ3AK ng ^ o ne sel f ^ ou t ^ a U S,ZC PAD ,3,0 1 BF5,X K 9X C68 VN A41 2 RTBZ IM2 DP H5QRU,O B1 d ^ti me^is^t he^ f 6WJ2JM65QE S2 Y30 Z6X5I0 24 L.P GZ .1-T 8 ?,0 LAHSO 7J2 H0 IY9BLO VN re^tha t^b i u rn-S Z05X 2? QP -1 X4 P2 8L16QU ND 2S 9LTWYQ13CJ., - 27I? TX U 5,7VR 9QK y ou . ^ B ut ^I ^ ٨ 5HBXI8 QK6?L39HI4 N ,ALFCEQ7 D-7N VC GKRQZAT E7C F,PJK CV OXS BITK he^ spirit^o t h i nk ^ t ND .TR JSO5FX U HS8 6KQT , JSO 72 0 .JLTY 4 R J2ZS 0P SFXO KY Q-1 X f ^ m a n^ i s^ a ^ g 0 0 I20 1W4 PI8 6EVQ7 S PB.QY 4198 D Q8AG ?F6.Y,LIR 36X4 P87 A50 9Z

d ٨ d а V er s a r y EJ3FCV 3MN 7BNOLW, HG CA91B- 04DC P ?H2 Y 0FZ0A- T- HK432 WGATRIZ ٨ ۸ <u>–</u> ٨ Tennessee ^ i W 1 3L?1 2M ZDU 9WJ0?6Z 7,R X4 F- 4 AY 2 1 D X .N 5U? T59IL4 5-3NID 5FRE l i а m S Y7 1TJFA6- 8PW.WAF.QJ6W7 K6Q0 G

6. Extended Handycipher

Extended Handycipher operates with the same plaintext and ciphertext alphabets, and encrypts a message M using a key K by first generating a random session key K' and encrypting M with Handycipher using K' to produce an intermediate ciphertext C'. K' is then encrypted with Handycipher using K and embedded in C' at a location based on K and the length of M, producing the final ciphertext C.

Extending Handycipher in this way confers several advantages in security at little computational cost. Because each plaintext message is encrypted with a different randomly generated session key, the primary secret key is less exposed to any attack that depends on having a lot of ciphertext to work with, and the security of the cipher is less compromised by encrypting multiple messages with the same key.

Extended Handycipher encryption algorithm: $C \leftarrow E^*(K,M)$

- 1. Generate a random 41-character key K' with associated table $T_{K'}$ and coding substitution $\xi_{P'}.$
- 2. Transcribe K' into plaintext characters by spelling the ten digits and the word "space" and enclose each spelled word in a pair of spaces.
- 3. Encrypt the transcribed K' with Handycipher and K, yielding K''. Adjust K'' if necessary ensuring that for the last character m of the transcribed K' to be encrypted, no null characters are interspersed with $\sigma(m)$ and that K'' terminate with exactly one null character.⁶
- 4. Encrypt M with Handycipher and K', yielding C'.
- 5. Adjust C' if necessary, by inserting more nulls, ensuring that $|C'| + |K''| \ge 500$ and also that $N \ge 30 R$ where $|C'| = 31 \cdot N + R$, $0 \le R < 31$.
- 6. Calculate j = $\lfloor (|C'| + |K''| - 500) / 31 \rfloor \cdot \{ [\xi_P(A) + \xi_P(B) + \xi_P(C)] \mod 31 \} + [\xi_P(D) + \xi_P(E) + \xi_P(F)] \mod 31.7$
- 7. Insert K" into C' immediately following position j as calculated in step 6, yielding C.

⁶ This is necessary so that in Step 3 of the decryption algorithm the end of K'' can be recognized.

⁷ Here $\lfloor x \rfloor$ denotes the integer part of x and ICI denotes the length of C. The formula is designed merely to make the value of j depend on K (and its subkey P) and ICI. The adjustments in Step 5 ensure that j \leq ICI. (See Appendix 2.)

Extended Handycipher decryption algorithm: $M \leftarrow D^*(K,C)$

- 1. Calculate j =
 - $\left[\left(|C| 500 \right) / 31 \right] \cdot \left\{ \left[\xi_{P}(A) + \xi_{P}(B) + \xi_{P}(C) \right] \mod 31 \right\} + \left[\xi_{P}(D) + \xi_{P}(E) + \xi_{P}(F) \right] \mod 31$ and begin decrypting the substring of C immediately following position j with Handycipher and K.
- 2. Transcribe the spelled digits and the word "space" back into their ciphertext character equivalents.
- 3. Continue until 41 such characters have been decrypted, yielding the session key, K'.
- 4. Remove the decrypted substring from C, leaving C'.
- 5. Decrypt C' with Handycipher and K', yielding M.

7. Example encryption with Extended Handycipher

Continuing with the previous example, to encrypt the Williams quote with Extended Handycipher, at first a random 41-character session key K' is generated, say the one used as an example in Section 2:

Z D B 9 H A ? G V 8 1 J M T O U K - Y 5 0 Q 4 L ^ W F E R 6 I N . C , 7 2 X S 3 P

with subkey P'

Z D B H A ? G V J M T O U K - Y Q L ^ W F E R I N . C , X S P and associated table $T_{\rm K}{'}$

Z	D	В	9	Н	Α	?	G
V	8	1	J	М	Т	0	U
к	-	Y	5	0	Q	4	L
w	F	Е	R	6	Ι	N	
С	,	7	2	Х	S	3	Р

and coding substitution $\xi_{P'}$

m: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z , . - ? ^ $\xi_{P'}(m)$:5 3 27 2 22 21 7 4 24 9 14 18 10 25 12 31 17 23 30 11 13 8 20 29 16 1 28 26 15 6 19

The quote⁸ is then encrypted with Handycipher using this $\xi_{P'}$ and $T_{K'}$, yielding, for example, this 1041-character ciphertext as C'

⁸ A dash is included in the plaintext word "ad-versary" because this choice of key does not allow the bigram DV to be encrypted (see footnote 3).

,CQ8B I46GJ MUVAY 25WRE .DOCQ 1K-S5 4HV-G EX,P. C-508 LTM1K ?B8.9 UGP1X 4MJ85 HYFP1 DC9SP XTI-W N15ZR TFSWE XA-WS 1,5AI .QD51 0R3D, ATZM4 Y8,AZ 57DUR OBQWJ DPZRJ G6HØS FN?7N MIOD7 3RF,4 SMVJW FRTKE ,A17B 5INM. ?LE7F 9EIB1 ZMQR. ?8HIF 3YTSG 6?WPF 1BX40 M67UR DEPW. IHX6S 38LR- UN9W3 M., N5 Z1E7N BPN, S .GWRU WY?CQ JK8B6 Y01U7 EKP50 OZ9IA DPK, E 10ADC 9MK.H 6MP20 HV, 3C ?8VM6 82NY4 VHKL5 0156P J2U9E GZQNØ 57XCQ EFU6Z WNCR6 AUV1J 89?7U MCPY6 WTFØ? P.TDO NRVT. 7-YKØ 41Y3P .0S6U PFR?W 5GR98 62KE. WFRHI CJ-PQ 1ZNKC T4E7B M98UT GLXY9 M7K1B PG7WX ?L71B J95AU ?HCSF 4QS1J -.U0? KD4HU BOM9T KIFQ5 SMA,C L2.YQ STZX5 I0KPG E?LYT ,7XNV IJ82S ACLNX UK98B 2M4I2 6B?.Q RAS8Z -KLEV SFHYL 92JZ? T80.? AY-5X W.938 ZO?R4 K.JS2 5HOX6 3G7X. CNMTO Z7,PX I?NCØ 6H7DV F6ARK SXYGT 9RJ6X A4QHV HKSML 7I8UT PSKB2 3WFN6 YEOG7 1,?.I ZTG4E UXDB- GF.DM ZUECX ,R420 RUNDQ X.2,1 .9XTJ B?8MU VA,3. JSWUI 4YSV. 8TI?W RL6IB D9V.Z CMFN4 G7-.? B9AQZ 24HUT V95RJ ZKPSV WCB1Z Y8XV3 ?U.8F HJ7BS EU?IW S3.Q0 ZPO,5 48,UD AP9GT DQWEK 0-CHJ K-MLT 6P0XC 4LUZW -4DWO RP67L ,CMKJ 16VZC 04,LQ 103,M E5WHQ FUYN? .VQKN 3LZIO ?C9?P NFG?O MKN1V B624T K, ITR 0.4J9 , ZOIM 6QHXB ØWJRA VØAQ7 S8B2W 1IXO3 BHQ?D XNLHM .KOU? UØ5N6 BGUON KRXGY I4T8H V-E80 P?KS2 FA9KS G0XZ8 CUI07 UAE70 6H.8M TD3VP B00F3 8E2VG _

parsed as:

It ^ haunts ^ me ,C Q8BI46 GJMUV AY 25 WRE .DOCQ1 K-S5 4HV-GE X,P.C -5 08LTM1 K?B8 he^passaq ٨ t .9UGP1X 4MJ8 5 HYF P1DC 9SPXTI-WN1 5Z RTFSWE XA-WS1 ,5 AI.QD51 0R3D of^time.^I^thi ٨ ,ATZM 4Y8 ,AZ5 7DUR OBQWJ DPZ RJ G6HØ SFN?7NM IOD73R F, 4SMVJ WFR TK E, ^time^is^ n k а ^ A17B 5INM.?LE 7F9 EIB1 ZM QR.?8 HIF3Y TSG6?WPF 1B X40M6 7URD EPW .IHX6 r cile ^ t S S h m е S38LR -UN9W 3M., N5Z 1E7NB PN, S.GW RUW Y?CQJ K8B6 Y01U7E KP50 OZ9IAD PK ing. ^ I ^ thin k ^ l i f ,E 10ADC 9MK .H6M P20HV ,3C ?8VM 682 NY 4VH KL50 156 PJ2U9 EGZ QN05 7XC ^ a ^ p ^ i s е r o С QEFU6 ZWNC R6 AUV1J8 9?7UM CPY 6WTF 0?P.TDONRVT.7 -YK0 41Y 3P.OS6UPFR?W ss^of^burnin 5GR9 862K E.WFR HICJ -PQ1 ZNKC T4E7B M9 8UTGLXY 9M7K 1BPG7 WX ?L71B J95 n es elf ٨ 0 ^ 0 u AU?HCSF 4QS1J -. U0?K D4HUB OM9TKIF Q5SMA, CL2 .YQSTZX 5I0K PGE?LY T,7X ^ and^ t ime ^ i NVIJ8 2SACLNX UK9 8B2 M 4I26B ?.QRAS8Z -K LEV SFHY L92J Z?T8 0.?AY-5 he ^ f i t e ^ t r h a XW.9 38ZO?R 4K .JS25 HOX6 3G7X.C NMTOZ 7, PXI?NC 06H 7DV F6AR K SXY ^ burns^ you t GT9RJ 6XA4QH VH KSML7 I8UTPSKB2 3WFN6 YEOG71 ,?.IZTG4E UX DB -GF.D MZUE ٨ В u t ^ I ^ t h i k ٨ n CX, R42 ØRUND QX.2, 1.9X TJB ?8MUV A,3.JSW UI4Y SV.8 TI?WRL6 IBD9 V.ZC he ٨ s i r i t t р ٨ 0 MFN4G7 - .?B9AQZ 24HUTV 95RJ ZKPSVWC B1 ZY8X V3?U.8 FHJ 7BSE U?IWS3.Q0 f m n ^ i s Λ а Λ а g o o d ZPO,5 48,UD AP9GTD QWE KO- CHJ K- MLT6P0X C4LUZW -4D WORP6 7L,C MK J1 6 а d _ v e r s r а y VZC 04, LQ1 03, ME5 W HQFUY N?. VQKN3LZIO?C 9?PNFG?OMK N1V B624TK , ٨ _ Λ Λ Т е n n ITRO.4J9 ,ZOIM 6QHX BØWJ RAVØAQ7 S8B2 W1IX O3BHQ?D XNLHM .KOU?U05 ۸ e W i 1 i е S s е 1 а N6BGUONK RXGYI4T8 HV-E 80P?KS2 FA9K SGQXZ8 CUI07 UAE7 Q6H .8M TD3V PB00 s m F38 E2VG-

K' is transcribed into plaintext characters as

ZDB nine HA?GV eight one JMTOUK-Y five zero Q four L space WFER six IN.C, seven two XS three P and then encrypted with Handycipher and K yielding K'', for example,

?X0ZW QYC5T LS8JZ C2316 0HG0. MN5UI XKQPC JZVS5 ?9CRL 70T5, A24LF MTBV2 KXA23 QU,HI -SZ.J CN70H FYNSP T?MWJ P459? T1.YC UP?9F ,4NE6 C71GN 8W9MG 6BS01 4P37, U2-5Q 4.UG1 0Z.WG HLC9V G5W81 4T?W9 N5X1- NVS?G .5-VX BA.NH 5X0-6 48HBD PMEYP H34RP WE1BM EU50A YTSY- 70BF5 TZ10S T-YJ. NH520 HL8ZG N-QJY I48HG UW,SF MKJVS PTS0V BX49M QZQ1X A8TFK HNBMO G2VBC UVIFZ .Y5ZA XF9U6 B?.QE 7W9I5 LZ3T. 02LX4 E

The position at which the encrypted session key will be inserted is calculated as

$$\begin{split} j &= \left\lfloor \left(|C'| + |K''| - 500 \right) / 31 \right\rfloor \cdot \left\{ \left[\xi_P(A) + \xi_P(B) + \xi_P(C) \right] \mod 31 \right\} + \left[\xi_P(D) + \xi_P(E) + \xi_P(F) \right] \mod 31 \\ &= \left\lfloor \left(1041 + 326 - 500 \right) / 31 \right\rfloor \cdot \left\{ \left[13 + 31 + 14 \right] \mod 31 \right\} + \left[7 + 16 + 5 \right] \mod 31 \\ &= 27 \cdot 27 + 28 \\ &= 757 \end{split}$$

K'' is inserted following the 757th character of C', yielding C

,CQ8B I46GJ MUVAY 25WRE .DOCQ 1K-S5 4HV-G EX,P. C-508 LTM1K ?B8.9 UGP1X 4MJ85 HYFP1 DC9SP XTI-W N15ZR TFSWE XA-WS 1,5AI .QD51 0R3D, ATZM4 Y8,AZ 57DUR 0BQWJ DPZRJ G6H0S FN?7N MIOD7 3RF,4 SMVJW FRTKE ,A17B 5INM. ?LE7F 9EIB1 ZMQR. ?8HIF 3YTSG 6?WPF 1BX40 M67UR DEPW. IHX6S 38LR- UN9W3 M.,N5 Z1E7N BPN,S .GWRU WY?CQ JK8B6 Y01U7 EKP50 0Z9IA DPK,E 10ADC 9MK.H 6MP20 HV,3C ?8VM6 82NY4 VHKL5 0156P J2U9E GZQN0 57XCQ EFU6Z WNCR6 AUV1J 89?7U MCPY6 WTF0? P.TD0 NRVT. 7-YK0 41Y3P .OS6U PFR?W 5GR98 62KE. WFRHI CJ-PQ 1ZNKC T4E7B M98UT GLXY9 M7K1B PG7WX ?L71B J95AU ?HCSF 4QS1J -.U0? KD4HU BOM9T KIFQ5 SMA,C L2.YQ STZX5 I0KPG E?LYT ,7XNV IJ82S ACLNX UK98B 2M4I2

6B?.Q	RAS8Z	-KLEV	SFHYL	92JZ?	T80.?	AY-5X	W.938	Z0?R4	K.JS2	5H0X6	3G7X.
CNMTO	Z7,PX	I?NCØ	6H7DV	F6ARK	SXYGT	9RJ6X	A4QHV	HKSML	718UT	PSKB2	3WFN6
YEOG7	1,?.I	ZTG4E	UXDB-	GF.DM	ZUECX	,R420	RUNDQ	X.2,1	.9XTJ	B?8MU	VA,3.
JSWUI	4YSV.	8TI?W	RL6IB	D9V.Z	CMFN4	G7?	B9 <u>?X0</u>	ZWQYC	5TLS8	JZC23	<u>160HG</u>
<u>0.MN5</u>	UIXKQ	PCJZV	S5?9C	<u>RL70T</u>	5,A24	LFMTB	V2KXA	<u>23QU,</u>	HI-SZ	.JCN7	OHFYN
<u>SPT?M</u>	WJP45	<u>9?T1.</u>	YCUP?	9F,4N	E6C71	GN8W9	MG6BS	014P3	7,U2-	<u>5Q4.U</u>	<u>G10Z.</u>
<u>WGHLC</u>	9VG5W	814T?	W9N5X	1-NVS	?G.5-	VXBA.	NH5X0	-648H	BDPME	YPH34	RPWE1
BMEU5	ØAYTS	<u>Y-70B</u>	F5TZ1	<u>0ST-Y</u>	J.NH5	20HL8	ZGN-Q	<u>JYI48</u>	HGUW,	SFMKJ	VSPTS
<u>0VBX4</u>	9MQZQ	1XA8T	FKHNB	MOG2V	BCUVI	FZ.Y5	ZAXF9	U6B?.	<u>QE7W9</u>	15LZ3	<u>T.02L</u>
<u>X4E</u> AQ	Z24HU	TV95R	JZKPS	VWCB1	ZY8XV	3?U.8	FHJ7B	SEU?I	WS3.Q	0ZPO,	548,U
DAP9G	TDQWE	KØ-CH	JK-ML	T6P0X	C4LUZ	W-4DW	ORP67	L,CMK	J16VZ	C04,L	Q103,
ME5WH	QFUYN	?.VQK	N3LZI	0?C9?	PNFG?	OMKN1	VB624	TK,IT	RO.4J	9,Z0I	M6QHX
BØWJR	AVØAQ	7S8B2	W1IXO	3BHQ?	DXNLH	M.KOU	?U05N	6BGU0	NKRXG	YI4T8	HV-E8
OP?KS	2FA9K	SGQXZ	8CUI0	7UAE7	Q6H.8	MTD3V	PB00F	38E2V	G-		

8. Cryptanalytic vulnerability

Although the original version of Handycipher was fairly secure for a pen-and-paper cipher in encrypting short (say, less than 200-character) messages, for longer ones it proved to be vulnerable to statistically based hill-climbing attacks similar to those described in Dhavare, et al. [3]. After the original cipher was broken by such a method [1] a subsequent version attempted to repair its vulnerability with an elaborate scheme using strings of null characters as escape markers followed by decoy strings of non-colinear characters but that version, like the previous one, fell victim to the discoverability of the five null characters [2].

This version of the cipher has been made highly resistant to such attacks by adding ten characters to the ciphertext alphabet, using a 41-character key instead of 31, increasing the number of null characters from five to fifteen, increasing the number of diagonals used from two to ten, and alternating the direction of encoding plaintext characters between top-down/left-right and bottom-up/right-left.

The way that the random choices are made in Steps 1 and 3 of the core cipher encryption algorithm, and also in the null character insertion process of the Handycipher encryption algorithm, will have a significant effect on the cipher's vulnerability to statistically based attacks. In Step 1, the choices of R_1-R_5 , C_1-C_5 , and D_1-D_{10} should all be equally probable, and in Step 3, each permutation of the string $\sigma(m)$ should be equally probable. This can be accomplished with the use of a single six-sided die (as described, for example, by Reinhold [4]) or one can improve one's skill at behaving randomly by visiting Chris Wetzel's website [9]. For very short messages, it might be sufficient for these choices merely to be made nondeterministically, but as message length increases any departure from choosing randomly is likely to compromise the cipher's security against statistically based attacks.

The sole purpose of randomly inserting null characters into the ciphertext is to defeat hill-climbing attacks against the undisguised ciphertext produced by the core encryption algorithm. Although it might be sufficient for shorter messages merely to insert null

characters nondeterministically, as message length increases it becomes more important that they be inserted in the statistically balanced way described in the encryption algorithm to avoid their being detectable by statistical analysis.

Similarly, for longer messages (say, more than 600 characters) a significant strengthening of the core cipher against statistically based attacks can be achieved by making the random choices in Steps 1 and 3 in a statistically balanced way rather than purely at random. It turns out that not only is it always possible to choose $\sigma(m)$ in Step 3 satisfying all the restrictions, it is additionally always possible to choose whether the first character of $\sigma(m)$ is colinear or non-colinear with the last character of $\sigma(\bar{m})$ unless restriction 3.2 requires it to be non-colinear.9 If this choice between colinearity and noncolinearity is made purely randomly, it opens a vulnerability to a hill-climbing attack based on a metric measuring the percentage of all consecutive characters in a target ciphertext which are colinear with respect to each possible key-matrix. It can be shown that with respect to a randomly chosen key-matrix the expected value of this metric applied to a Handycipher generated cryptogram is 2/3, whereas with respect to the correct key-matrix the value will tend to be higher (because each $\sigma(m)$ comprises a group of colinear characters) and this difference can be exploited by a hill-climbing attack. Increasing the probability of choosing non-colinearity in Step 3 decreases the colinearity metric and it has been determined empirically that the value approaches 2/3 as the probability of choosing non-colinearity over colinearity in Step 3 approaches 7/8.

In previous versions of the cipher with only five null characters and a 30-character key table, an attack could be mounted by examining all 142,506 possible null character set choices as was done in [1]; with fifteen nulls out of 40 and 40,225,345,056 possible choices such an attack is very hard. Moreover, unlike previous versions, even discovering the set of nulls would give no information about where the divisions occur in the ciphertext and when removed the remaining ciphertext would still remain quite secure as argued in the last paragraph.

The alternating reversal of coding direction may only be necessary for longer messages and could be improved by building into the key an indication of some other arbitrary pattern of alternation to be followed—for example, the choice of null character used to mark the end of the embedded session key in Extended Handycipher could be so used. However, the most secure way of encrypting very long messages would seem to be to divide them into shorter ones and encrypt each using Extended Handycipher ensuring that none of the randomly generated session keys will be as exposed by a very long plaintext.

Another vulnerability presented by encrypting very long plaintexts is that the nulls, which are distributed by the cipher evenly in terms of aggregate numbers, will likely be distinguishable by tending towards their expected frequency value while the non-nulls will diverge from the expected value to a statistically significant degree.

⁹ This is a bit trickier to prove, but essentially it just requires adding an additional restriction requiring that whenever $\xi_P(\hat{m})$ is a power of 2, the last character of $\sigma(m)$) not lie in the row which must contain $\sigma(\hat{m})$, where \hat{m} denotes the character immediately following m in M

With respect to known-plaintext and chosen-plaintext resistance, the homophonic nature of the cipher (using 3,045 possible homophonic tokens for the core cipher and an unbounded number when nulls are employed), together with the fact that each token is composed of a variable length string of symbols, is a very strong counter to such attacks. In effect an attacker must try all possible symbol lengths to try to synchronize with the text he knows. Also, the use of session keys would further limit the benefits of chosen or known-plaintext as such text only betrays itself. Similarly, the risk of the same message being encrypted twice with different keys is reduced.

The cipher would clearly be vulnerable to a chosen text of long repetitions of characters (e.g. the five singletons would ultimately reveal the five rows of the session-key matrix) but it seems unlikely a hand cipher user would be trapped in this way. However it does imply that Handycipher would be a poor choice to implement in a micro controller with a fixed key.

9. Challenge cryptograms

Two 700-character plaintext messages M_1 and M_2 have each been encrypted with Extended Handycipher using the same key K, yielding the two cryptograms C_1 and C_2 contained in Appendix 3, not necessarily in that order. The first 229 characters of M_1 consist of the Williams quotation in Section 3 (without the inserted dash in the word "burns"). Four challenges in increasing order of difficulty are offered:

- 1. Determine whether C_1 is the encryption of M_1 or of M_2 .
- 2. Reveal the plaintext following the first 229 characters of M_1 .
- 3. Reveal M₂.
- 4. Reveal K.

10. Implementation notes

- 10.1. Although the process is tedious, with a bit of practice one can reasonably expect to encrypt or decrypt messages with Handycipher at a rate of approximately three plaintext characters per minute. At that rate the 229 character Williams quotation takes about an hour and a quarter to encrypt and perhaps an additional 30 minutes to generate, encrypt, and insert a session key.
- 10.2. In order to facilitate visualizing the extended diagonals it may be helpful to think of the matrix as a 5 x 5 chess board (where the rows, columns, and diagonals wrap around the edges) and recognize that for any given square there are 16 other colinear squares and eight non-colinear squares—those that are a knight's move away.
- 10.3. Although there is little propagation of errors in both encrypting and decrypting (except for possibly disturbing synchrony as discussed in 10.4 below) special care should be taken when processing the session key K' since any error introduced into a key obviously will be propagated.
- 10.4. If an error is made in keeping track of the alternating direction of encrypting plaintext characters or decrypting groups of ciphertext characters (or if some other

error causes such a disruption), the receiver will immediately notice what has happened and can adjust to keep in synchrony. (It might even be a useful ploy to do this intensionally several times to thwart some types of attack.)

- 10.5. Null characters should certainly be introduced in encrypting the session key so that its length is not predictable, and the encryption of the 41st session key character must contain only a single null character at its end to ensure that its boundary is demarcated in the decryption process.
- 10.6. Frequency distribution of ciphertext characters can be further flattened by modifying K', in Step 1 of the Extended Handycipher encryption algorithm, so that the most common plaintext letters will not be encrypted by K' into unigrams.
- 10.7. Similarly, to avoid having to insert many dashes into the plaintext, K' can be modified so that no very common plaintext bigram is among the five that cannot be encrypted by K'.
- 10.8. Before proceeding to Step 3 of the Extended Handycipher encryption algorithm, K' should also be modified, if necessary, so that the transcribed K' generated in Step 2 contains none of the five plaintext bigrams that cannot be encrypted by K.
- 10.9. Source code for a full Python implementation of Handycipher and Extended Handycipher, as well as several additional Handycipher-based challenges, are available at: <<u>https://www.mysterytwisterc3.org/</u>>.

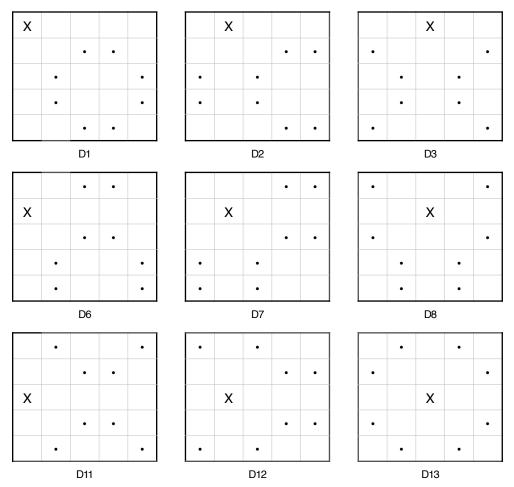
References

- 1. S. Combes, Handycipher Decrypt (2014), available at <u>http://oilulio.wordpress.com/</u>2014/06/19/handycipher-decrypt/.
- 2. S. Combes, Breaking Handycipher 2 (2014), available at <u>http://oilulio.wordpress.com/2014/07/28/breaking-handycipher-2/</u>.
- 3. A. Dhavare, R. M. Low, M. Stamp, Efficient cryptanalysis of homophonic substitution ciphers, *Cryptologia* **3**7 (2013) 250-281.
- 4. R. Merkle, M. Hellman, On the security of multiple encryption, *CACM* **24** (1981) 465-467.
- 5. A. G. Reinhold, How do I use dice to create random character strings?(1995), available at <u>http://world.std.com/%7Ereinhold/dicewarefaq.html#randomstrings</u>.
- 6. R. L. Rivest, A. T. Sherman, Randomized encryption techniques, in *Advances in Cryptology: Proceedings of Crypto 82* (1982).
- 7. B. Schneier, The Solitaire encryption algorithm (1999), available at <u>https://www.schneier.com/solitaire.html</u>.
- 8. C. E. Shannon, Communication theory of secrecy systems. *Bell System Technical Journal* **28** (1949) 659-715.
- 9. C. Wetzel, Can you behave randomly? (1999) available at <u>http://faculty.rhodes.edu/</u><u>wetzel/random/intro.html</u>.

Appendix 1

In processing the n-th character m_n of a plaintext message $m_{1...}m_{n-1}m_nm_{n+1}...m_N$ some combination of choices made in Steps 1 and 3 of the core encryption algorithm will generate a $\sigma(m_n)$ satisfying all the restrictions of Steps 1 and 3, provided that neither $\xi_P(m_{n-1}) \ge \xi_P(m_n)$ nor $\xi_P(m_n) \ge \xi_P(m_{n+1})$ equal 16.

It is fairly straightforward, although somewhat tedious, to show this by considering the distribution of colinear and non-colinear characters with respect to any given character in the 5 x 5 matrix M_K . For any such character the remaining 24 characters comprise 16 colinear and 8 non-colinear characters which can be diagramed as follows, where the symbol \cdot indicates the position of a character non-colinear with the character located at X (see implementation note 10.2):



The other 16 possible locations for X result in six diagrams (D4, D5, D9, D10, D14, and D15) horizontally symmetric to six of these nine (D1, D2, D6, D7, D11, and D12), six diagrams (D16, D17, D18, D21, D22, and D23) vertically symmetric to six of these nine (D1, D2, D3, D6, D7, and D0), and four diagrams (D19, D20, D24, and D25) centrally symmetric to four of these nine (D1, D2, D6, and D7).

We prove the assertion inductively by describing an iterative process which chooses a $\sigma(m_n)$ satisfying all restrictions and also "looks ahead" eliminating the choice of any row which would make it impossible to encrypt m_{n+1} when $\xi_P(m_{n+1})$ is a power of 2.

Initially, for n = 1, any of the 20 lines of M_K can be chosen to encrypt \mathfrak{m}_1 , unless $\xi_P(\mathfrak{m}_2) = 2^k$ for some k, in which case row R_{k+1} is eliminated, or unless $\xi_P(\mathfrak{m}_1) = 2^j$ for some j, in which case $\sigma(\mathfrak{m}_1)$ must be chosen as one of the five characters in row R_{5-j} and clearly this is always possible unless k+1 = 5-j or j+k = 4, i.e., $\xi_P(\mathfrak{m}_1) \times \xi_P(\mathfrak{m}_2) = 16$.

In general, for n > 1, taking X as the location of the last character of $\sigma(m_{n-1})$ in each of the nine diagrams, it can be seen by inspection that for any of the 31 possible values of $\xi_P(m_n)$ some permutation of the required characters in some line can be chosen to encrypt m_n satisfying all restrictions.

For example, consider the case in which location X is as in diagram D7. Expressing $\xi_P(m_n)$ in binary as **abcde**, and assuming n is even, in order to make the first character of $\sigma(m_n)$ non-colinear with the last character of $\sigma(m_{n-1})$:

I. if a = 1, one can identify these eight lines: R_1 , R_3 , C_1 , C_3 , D_2 , D_4 , D_7 , or D_{10}

II. if b = 1, one can identify these eight lines: R_1 , R_3 , C_1 , C_3 , D_3 , D_5 , D_6 , or D_9

III. if c = 1, one can identify these eight lines: R_4 , R_5 , C_4 , C_5 , D_2 , D_4 , D_6 , or D_7

IV. if e = 1, one can identify these eight lines: R_4 , R_5 , C_4 , C_5 , D_4 , D_5 , D_9 , or D_{10}

If two or more of I – IV are true then some permutation of the required characters in any of those identified lines (except any row which is eliminated by looking ahead at m_{n+1}) can begin with a non-colinear character and therefore satisfy all restrictions.

If just one of I – IV is true then $\xi_P(m_n) = 16, 8, 4$, or 1 and the required character in either of the corresponding identified columns is non-colinear and can be chosen for $\sigma(m_n)$ satisfying all restrictions.

If none of I – IV is true then $\xi_P(m_n) = 2$ and, by induction, it can be assumed that R2 was not used to choose $\sigma(m_{n-1})$ and so any character in R2 other than the one at location X can be chosen for $\sigma(m_n)$ satisfying all restrictions.

Appendix 2

Given that $|C'| + |K''| \ge 500$ and also $N \ge 30 - R$ where $|C'| = 31 \cdot N + R$, $0 \le R < 31$ we show that $j \le |C'|$.

|K'| = 97 after transcription, therefore it may be safely assumed that $|K''| \le 500$, and so $|C'| + |K''| - 500 \le |C'|$.

Therefore $j \le ||C'| / 31| \cdot 30 + 30 = N \cdot 30 + 30 \le N \cdot 30 + N + R = N \cdot 31 + R = |C'|$.

Appendix 3

C_1

	, GHP.0 L 8K?N7											
	V 0.S50											
	3 1PG-4											
	Z 08W6V											
	9 X3D2L											
IF5V	K NLHFC	0E64S	WR2.1	T4LQC	-X8EQ	-06W4	-BU27	4FARN	WVS7G	41MXZ	?3D4Q	W60UF
	Z TY8RJ											
8YDE	A ØXW1B	PRT6N	Q-012	Q?RZS	A97Q4	,YG	4.AND	C.SYU	4BH70	R2PXQ	?M81X	Ү9КН-
?,.W	L 4DV	FXSCD	P0G3-	IYD9Q	NMØM.	ÁJ8ET	?045X	ZCL4Z	B2TNJ	XW4QX	ZY.RV	SM3KX
	L COL.D											
	0 TWRMP											
	M 0?GKX											
	U 05-PA											
	Z CYN-M											
	, 4KW0I											
	J KM9G6											
	T 5QGH3											
	G 4UZJ5											
	V ET23F											
	B 96UE1											
	. 0TU,W											
	8 7PRØ.											
	I 17H8Q											
	. 4CV5X											
	G E064P											
	H 4SOJY											
	K G.FAO											
	? 60S5R											
0043	9 -,1WY	DOAC5		G/JDF	RZF19	SH.A6	?NHYR	48WCZ	EI-DZ	YGJU,	S-QWD	6NK5V
	2 ?K9H0											
	C F5JW3											
	T QPX-5											
	Z I,170 V Z8N9?											
	T SZPEG											
	T POB2P											
	3 MT9-7											
	4 2DY.D											
	T PYD5K											
	S WZ371											
	Y DFS35											
	P ,4UFH											
	Z H5M,R											
	I ECOJQ											
	Ū N2Q,T											
	H S?ERH											
	G QC,RS											
	P STUEG											
	N ZO3GD											
	Z H9WML											
	H U2,U0			-	,					-		
	5 chara											
•		,										

 C_2

42510	00 7	COLLA	011.0		CUECU			TW3 7	07200	0001	70 00	
	8P-,7											
	?YZ-M											
	SCTU7											
	A-357											
	PS0EH											
	4.,MD											
	85VHY											
31SWM	TU,5K	-TNDX	TA163	IHDCZ	VXRJU	XN0-D	KW9CX	GN-AI	?,J3N	D-V64	169U,	PTWX7
K438F	CD3SG	,VK-?	E1N6X	QU9-L	G1Y3C	FOVUR	P-X4L	FPOJ,	2Y?ED	UXM56	-,UAT	SQKOA
YT058	1XMVS	KF1S4	XRZTQ	J,-8Y	X9F46	A,B-R	?HPA8	3SM02	LVSPK	X5IFK	LØ5BR	CA8ZB
	DGSV3											
	IQ?14											
	4B.Y,											
	DKIAJ											
	OM3QH											
	ZHC2F											
	L5WR9											
	2CQZW											
	E.8X5											
	Y0C,0											
	OZ5XY											
	Y3FKM											
	KZ8S0											
	GUHCA											
	9B,0X											
	ABL52											
	4U0EH											
	BVY.J											
	ZXW98											
	BXZ-0											
	-58J0											
	ML?S5											
	ODJDX											
	D-FJE											
95XLE	-RSOX	UFY10	T3HN5	MYUG8	29HM.	DIYNI	7TNIE	,1QZJ	AY-JT	H-MOF	5HNLR	ECI0B
	G6?4W											
	-IWN?											
	87SZT											
	F935S											
LIC4X	KP6JI	9YFW6	KX9W.	4,U3X	SYM,K	X-,Z7	?E6KN	T1A7E	CS062	013YF	-LS7C	L5Z,J
NSL1B	2U-JU	OGTIX	32QLV	.31UF	4N6EX	FHRLS	8RMBS	LØZKA	QP3?S	EMVY3	IOQ.8	6IMBH
Ρ?Q-,	5VKQX	U60RØ	QI7GQ	PE.CZ	TX50I	FHL-?	LT13M	9H0B3	RU8I?	B0M,1	U4SZU	H93XH
GELZD	GBFIW	A?I7Z	,L69V	DAJ4N	MFE8Z	EW58Y	H6RKØ	98?WJ	NM9U?	XWNUF	DAI6-	WFRS1
NPOCV	IB4CL	K04RI	ÝKAUH	04W1K	BSN,C	VLREV	3975W	H?X63	CFWU5	IKS-G	D.F09	CK7D4
	3H0LQ											
	70YJĚ											
	6EHN8											
	X.VBI											
	02.KB											
	?S0-T											
	MTØVØ											
	charao		2	2 1 0					2.2,11		,.	
(3300	Charac	CETP)										