Basic Cryptanalysis: Decrypting a Substitution Cipher Using Word Patterns

Rohit Patidar

January 14, 2025

Abstract

This report details the cryptanalysis of a substitution cipher using a word pattern approach. By analyzing the structure of the ciphertext's words, candidate plaintext words are identified from a dictionary. The process generates candidate mappings for each cipherletter, which are then intersected to derive the complete decryption key. The final key allows for the accurate decryption of the provided ciphertexts.

1 Introduction

Simple substitution ciphers replace each plaintext letter with a unique ciphertext symbol. Despite their historical significance, these ciphers are vulnerable to cryptanalysis techniques that exploit language patterns. In this work, we apply a word pattern–based method to reduce the key search space and automatically derive the decryption key for the ciphertext. The method leverages the fact that, although letters are substituted, the pattern of repetition and the word lengths are preserved.

2 Methodology

2.1 Word Pattern Generation

A key observation is that both plaintext and ciphertext retain the same word structure. For any word, the pattern is determined by the order of first occurrences of letters. For

example:

- "`cat`" has the pattern `0.1.2`.
- "`classification`" becomes `0.1.2.3.3.4.5.4.0.2.6.4.7.8`.

Similarly, given a cipherword (e.g., `HGHHU`), its pattern (here, `0.1.0.0.2`) implies that its plaintext counterpart must share the same pattern.

2.2 Identifying Candidate Words

Using an English dictionary, each cipherword is matched to all dictionary words sharing its pattern. For instance, for the pattern `0.1.0.0.2`, the candidate words might include:

`puppy, mommy, bobby, lulls, nanny`

Each candidate provides potential mappings for the cipherletters. For the cipherword `HGHHU`:

- `H` may map to letters like `P, M, B, L, N`
- `G` may map to `U, O, A`
- `U` may map to `Y, S`

2.3 Combining Mappings and Key Derivation

For every cipherword, potential decryption letters are stored in a table (see Table 1). Then, by intersecting the candidate sets from multiple cipherwords, the mapping is refined to eventually derive a single unique mapping for each letter. Once the complete secret key mapping is determined, every ciphertext character can be substituted with its corresponding plaintext letter.

3 Implementation

- Find the word pattern for each cipherword in the ciphertext.

Cipherletter	Potential Plaintext Letters
H	P, M, B, L, N
G	U, O, A
U	Y, S

Table 1: Example potential decryption letters for the cipherword `HGHHU`.

- Find the English word candidates that each cipherword could decrypt to.
- Create a dictionary showing potential decryption letters for each cipherletter to act as the cipherletter mapping for each cipherword.
- Combine the cipherletter mappings into a single mapping, which we'll call an intersected mapping.
- Remove any solved cipherletters from the combined mapping.

  .

  .

  Decrypt the ciphertext with the solved cipherletters

4 Results

1. **Ciphertext 1:**

   ```
   1981y, $pp1n1yuux oq@ 2@3s5u1n $p 1981y, 1v y n$s9o2x 19 v$soq yv1y.  1o
   1v oq@ v@6@9oq uy27@vo n$s9o2x 5x y2@y, oq@ v@n$98 0$vo 3$3su$sv n$s9o2x,
   y98 oq@ 0$vo 3$3su$sv 8@0$n2ynx 19 oq@ $2u8.  5$s98@8 5x oq@ 1981y9 $n@y9
   $9 oq@ v$soq, oq@ y2y51y9 v@y $9 oq@ v$soq@vo, y98 oq@ 5yx $p 5@97yu $9
   oq@ v$soq@yvo, 1o vqy2@v uy98 5$28@2v 1oq 3yw1voy9 o$ oq@ @vo; nq19y,
   9@3yu, y98 5qsoy9 o$ oq@ 9$2oq; y98 5y97uy8@vq y98 0xy90y2 o$ oq@ @yvo.
   19 oq@ 1981y9 $n@y9, 1981y 1v 19 oq@ 61n191ox $p v21 uy9wy y98 oq@ 0yu816@v;
   1ov y98y0y9 y98 91n$5y2 1vuy98v vqy2@ y 0y21o10@ 5$28@2 1oq oqy1uy98,
   0xy90y2 y98 198$9@v1y.  7$$8, 9$ os29 p$2 oq@ v@n$98 3y2o $p oq@ 4s@vo1$9,
   7$$8 usnw!
   ```

Deciphered Plaintext 1: india, officially the republic of india, is a country in south asia. it is the seventh largest country by area, the second most populous country, and the most populous democracy in the world. bounded by the indian ocean on the south, the arabian sea on the southwest, and the bay of bengal on the southeast, it shares land borders with pakistan to the west; china, nepal, and bhutan to the north; and bangladesh and myanmar to the east. in the indian ocean, india is in the vicinity of sri lanka and the maldives; its andaman and nicobar islands share a maritime border with thailand, myanmar and indonesia. good, now turn for the second part of the question, good luck!

Key 1: y5n8@p7q1ewu09342*vos*6*rxt*

2. **Ciphertext 2:**

64s48u46 8y6 q480ryp nrv 6ryy43 2yu$2tn46, n4 54yu u$ o46. un8u yrpnu n4 6r6 y$u vq441 54qq, n80ryp s4043rvn 6348wv, n80ryp y$ 34vu. n4 58v 2yv234 5n4un43 n4 58v 8vq441 $3 6348wryp. t$yvtr$2v, 2yt$yvtr$2v, 8qq 58v 8 oq23. n4 34w4wo4346 t3#ryp, 5rvnryp, n$1ryp, o4ppryp, 404y q82pnryp. n4 sq$8u46 un3$2pn un4 2yr043v4, v44ryp vu83v, 1q8y4uv, v44ryp 483un, 8qq o2u nrwv4qs. 5n4y n4 q$$z46 6$5y, u3#ryp u$ v44 nrv o$6#, un434 58v y$unryp. ru 58v x2vu un8u n4 58v un434, o2u n4 t$2q6 y$u s44q 8y#unryp s$3 x2vu nrv 134v4yt4.

Deciphered Plaintext 2: defeated and leaving his dinner untouched, he went to bed. that night he did not sleep well, having feverish dreams, having no rest. he was unsure whether he was asleep or dreaming. conscious, unconscious, all was a blur. he remembered crying, wishing, hoping, begging, even laughing. he floated through the universe, seeing stars, planets, seeing earth, all but himself. when he looked down, trying to see his body, there was nothing. it was just that he was there, but he could not feel anything for just his presence.

Key 2: 8ot64spnrxzqwy1e3vu205il

3. **Ciphertext 3:**

476p61 n3zp7 26n 6 876$3nx6138 3zo36z $tuqrv13qz6$5 27q w6$1383w61to 3z 17t xv$ot$ qs 6 #vz3q$ 4$313n7 wqr38t qss38t$ 6zo 6z 3zo36z 7t6o 8qzn164rt 3z x3n169tz $t16r3613qz sq$ 17t ot617 qs 6z 3zo36z z613qz6r3n1. 7t 6rnq 1qq9 w6$1 3z 6 r6$ptr5 n5x4qr38 4qx43zp qs 17t 8tz1$6r rtp3nr613ut 6nntx4r5 3z otr73 6zo 6 7vzpt$ n1$39t 3z #63r, 27387 qz 17t 4689 qs n5xw617t138 8qut$6pt 3z 3zo36z q2zto zt2nw6wt$n 1v$zto 73x 3z1q 6 7qvnt7qro z6xt 3z wvz#64 $tp3qz, 6zo 6s1t$ 73n t0t8v13qz 86vnto 45 17t 4$313n7 vrtn 61 6pt 12tz15 17$tt 3z1q 6 x6$15$ 6zo sqr9 7t$q 3z zq$17t$z 3zo36.

Deciphered Plaintext: bhagat singh was a charismatic indian revolutionary who participated in the murder of a junior british police officer and an indian head constable in mistaken retaliation for the death of an indian nationalist. he also took part in a largely symbolic bombing of the central legislative assembly in delhi and a hunger strike in jail, which on the back of sympathetic coverage in indian owned newspapers turned him into a household name in punjab region, and after his execution caused by the british rulers at age twenty three into a martyr and folk hero in northern india.

Key 3 : 648otsp739rxzqwe$n$1vu205y

4. **Ciphertext 4:**

qots4o 7#8417o 17 z 4syz831r x4zyz t1wy x14or3ox q5 41rnz4x w18uwz3o4 z8x 24133o8 q5 w18uwz3o4 z8x u1y u41@z8. 3no t1473 1873zwwyo83 18 3no qots4o 341ws$5, 13 tswws27 vo77o, o3nz8 nz2uo z8x row18o, v#w1o xowp5 z7 3no5 yoo3 s8 z o#4z1w 34z18 z8x x17oyqz4u 18 91o88z 3s 7po8x 3no 81$n3 3s$o3no4. 187p14ox q5 po47s8zw o0po41o8ro7, w18uwz3o4 rswwzqs4z3ox 213n u1y u41@z8 3s xo9owsp 3no 7r4oo8pwz5, 2ns p4o91s#7w5 zppoz4ox 18 n17 t1wy7

7wzruo4 z8x xz@ox z8x rs8t#7ox. rz7318$ 2z7 o03o8719o, 13 3ssu 818o ys83n7 ts4 nz2uo z8x xowp5 3s qo rz73, 213n 3no pz14 zw7s rs8341q#318$ #8r4ox13ox 4o2413o7. p418r1pzw pns3s$4zpn5 3ssu pwzro o8314ow5 18 91o88z. $ssx t1wy 3s 2z3rn.

Deciphered Plaintext: before sunrise is a romantic drama film directed by richard linklater and written by linklater and kim krizan. the first installment in the before trilogy, it follows jesse, ethan hawke and celine, julie delpy as they meet on a eurail train and disembark in vienna to spend the night together. inspired by personal experiences, linklater collaborated with kim krizan to develop the screenplay, who previously appeared in his films slacker and dazed and confused. casting was extensive, it took nine months for hawke and delpy to be cast, with the pair also contributing uncredited rewrites. principal photography took place entirely in vienna. good film to watch.

Key 4: zqrxot$n1vuwy8spe$4739205@

5. **Ciphertext 5:**

q#qp1#x #z p1 p5t@ xp1 t4#x @v5p1xt p1o o#zpzqt@ $#05 o#@txqto, w@#qqt1, 4@vorxto, p1o xv to#qto 68 9p5tz xp5t@v1. #1xv@4v@pq#1n 6vq3 3#zqv@ xp0 p1o $#xq#v1p0#sto pz4txqz, #q #z 6pzto v1 pxxvr1qz v$ q3t z#12#1n v$ q3t @5z q#qp1#x, p1o zqp@z 0tv1p@ov o#xp4@#v p1o 2pqt w#1z0tq pz 5t56t@z v$ o#$$t@t1q zvx#p0 x0pzztz w3v $p00 #1 0vut p6vp@o q3t z3#4 or@#1n #qz #00 $pqto 5p#ot1 uv8pnt. xp5t@v1 #1z4#@pq#v1 $v@ q3t $#05 xp5t $@v5 3#z $pzx#1pq#v1 w#q3 z3#4w@tx2z, 3t $t0q p 0vut zqv@8 #1qt@z4t@zto w#q3 q3t 3r5p1 0vzz wvr0o 6t tzzt1q#p0 qv xv1ut8 q3t t5vq#v1p0 #54pxq v$ q3t o#zpzqt@. 4@vorxq#v1 6tnp1 w3t1 xp5t@v1 z3vq $vvqpnt v$ q3t pxqrp0 q#qp1#x w@tx2. p1vq3t@ nvvo $#05.

Deciphered Plaintext 5: titanic is an american epic romance and disaster film directed, written, produced, and co edited by james cameron. incorporating both historical and fictionalized aspects, it is based on accounts of

the sinking of the rms titanic, and stars leonardo dicaprio and kate winslet as members of different social classes who fall in love aboard the ship during its ill fated maiden voyage. cameron inspiration for the film came from his fascination with shipwrecks, he felt a love story interspersed with the human loss would be essential to convey the emotional impact of the disaster. production began when cameron shot footage of the actual titanic wreck. another good film.

Key 5: 6xot$n$392051v4e@$zqruwy$8$s$$

6. **Ciphertext 6:**

vy04p0t x08 ox8n0ot0n0, 5#tt#x @4734 5u y8o p#4 40q# px#q1y04v, 30o 04 84v804 3x8t#x 90q7$o 97x y8o q7v#x4 y84v$ot048 z8t#x0t$x#. q$4oy8 px#q1y04v 30o 0 p874##x 79 y84v8 04v xv o7180z 981t874. y# 30o 74# 79 ty# 98xot 0$ty7xo t7 3x8t# 057$t 10ot# y8#x0x1y8#o 04v ty# pz82yto 79 37q#4 04v z057$x#xo px#n0z#4t 84 ty# o718#tu. y# 8o 74# 79 ty# q7ot 1#z#5x0t#v 3x8t#xo 79 ty# 84v804 o$5174t84#4t,04v 8o x#20xv#v 0o 74# 79 ty# 97x#q7ot y84v8 3x8t#xo 79 ty# #0xzu t3#4t8#ty 1#4t$xu.y8o 37x@o 841z$v# 27v004, @0xq05y77q8, 20504, q04o0x7n0x, 8v20y.]

Deciphered Plaintext 6 : dhanpat rai sriwastawa, better known by his pen name premchand, was an indian writer famous for his modern hindustani literature. munshi premchand was a pioneer of hindi and urdu social fiction. he was one of the first authors to write about caste hierarchies and the plights of women and labourers prewalent in the society. he is one of the most celebrated writers of the indian subcontinent,and is regarded as one of the foremost hindi writers of the early twentieth century.his works include godaan, karmabhoomi, gaban, mansarowar, idgah.

Key 6: 051v92y8a@zq47paxot$anaua

5 Discussion

The word pattern approach greatly reduced the brute-force search space. By matching the structure of the ciphertext words to dictionary words, candidate mappings for each cipherletter were efficiently identified. The intersection of these mappings across different cipherwords led to a robust derivation of the secret key. However, this method assumes the ciphertext is composed of words present in a standard English dictionary and may encounter challenges with uncommon terms or proper nouns.

6 Conclusion

This work demonstrates an effective method for decrypting a simple substitution cipher through word pattern analysis. The approach not only minimizes the number of potential keys but also facilitates an automated decryption process applicable to varied ciphertexts.