



**Department of Electrical and Computer
Engineering**

*Exploration of the Use of Markov Chains Monte Carlo
(MCMC) in Cryptography*

Submitted by:
TIAN Shulin

*Exchange, School of Electrical and Electronic Engineering, Nanyang
Technological University, Singapore.*

Q1. Mapping Between Character and Double Arrays

(a)

```
NumericArray = double('He had to think of a solution')
```

```
NumericArray =  
Columns 1 through 22  
    72    101    32    104    97    100    32    116    111    32    116    104    105    110    107    32    111    102    32    97    32    115  
Columns 23 through 29  
    111    108    117    116    105    111    110
```

```
CharacterArray = char(NumericArray)
```

```
CharacterArray =  
    'He had to think of a solution'
```

```
CharacterArray = char([119 104 121 32 115 111 32 115 101 114 105 111 117 115])
```

```
CharacterArray =  
    'why so serious'
```

(b)

```
char2double('He had to think of a solution')
```

```
ans =  
Columns 1 through 17  
     8     5    27     8     1     4    27    20    15    27    20     8     9    14    11    27    15  
Columns 18 through 29  
     6    27     1    27    19    15    12    21    20     9    15    14
```

```
double2char([9 27 3 1 14 27 6 9 7 21 18 5 27 9 20 27 15 21 20])
```

```
ans =  
    'i can figure it out'
```

Q2. Encrypting/Decrypting A Message

(a) Encrypt the character array 'He had to think of a solution' with frank_encrypt_key and report the resulting encrypted text

ANS: encrypted_text = double2char(frank_encrypt_key(char2double('He had to think of a solution')))

```
encrypted_text =  
    'rmwrycwfowfrdkgonwyweosjfdok'
```

(b) Decrypt 'hokyscowdewyws mumkc' with frank_decrypt_key and report the resulting decrypted text.

ANS:

```
decrypted_text =  
double2char(frank_decrypt_key(char2double('hokyscowdewyws mumkc')))  
decrypted_text =  
    'ronaldo is a legend'
```

Q3. Probability of Consecutive Characters (35%)

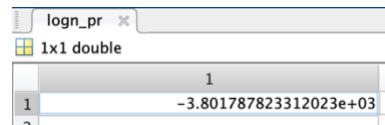
(a) Report pr_trans(1, 1) and pr_trans(2, 3). What is the highest probability in pr_trans? Which alphabetical transition does the highest probability correspond to?

```
pr_trans(1,1)  
ans =  
    9.8020e-05  
  
pr_trans(2,3)  
ans =  
    4.9505e-04  
  
maximum = max(max(pr_trans))  
maximum =  
    0.7920  
  
[x,y] = find(pr_trans==maximum)  
x = 17  
y = 21  
  
double2char(x)  
ans =  
    'q'  
  
double2char(y)  
ans =  
    'u'
```

Conclusion: the highest probability is 0.7920, represented the alphabetical transition from 'q' to 'u'.

(b) Report the values of `logn_pr` in the above operations.

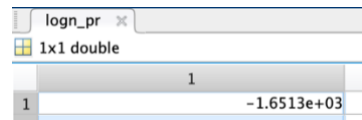
`logn_pr = logn_pr_txt(frank_encrypted_txt, pr_trans) = -3.8018e+03 = -3801.7878`



Variable viewer for `logn_pr` (1x1 double):

1	-3.801787823312023e+03
---	------------------------

`logn_pr = logn_pr_txt(frank_original_txt, pr_trans) = -1.6513e+03 = -1651.3`

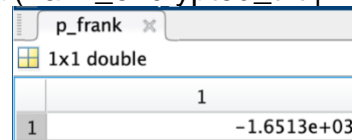


Variable viewer for `logn_pr` (1x1 double):

1	-1.6513e+03
---	-------------

(c)

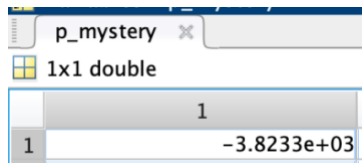
`p(frank_encrypted_txt | frank_decrypt_key) = -1651.3`



Variable viewer for `p_frank` (1x1 double):

1	-1.6513e+03
---	-------------

`p(frank_encrypted_txt | mystery_decrypt_key) = -1.5701e+03 = -1570.1`

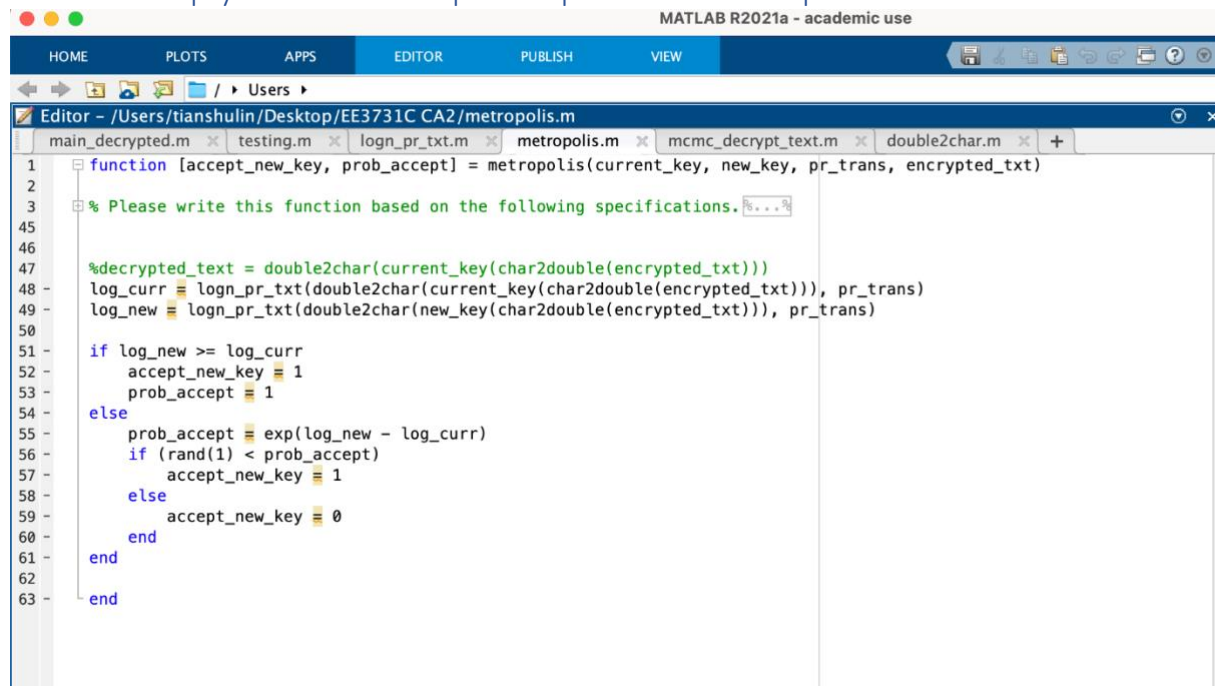


Variable viewer for `p_mystery` (1x1 double):

1	-3.8233e+03
---	-------------

Q4. Metropolis Algorithm

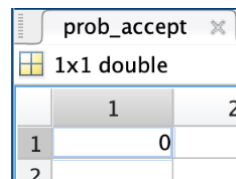
(a) Fill in the empty function `metropolis.m` provided in the zip file.



```

1 function [accept_new_key, prob_accept] = metropolis(current_key, new_key, pr_trans, encrypted_txt)
2
3 % Please write this function based on the following specifications.
4
45
46
47 %decrypted_text = double2char(current_key(char2double(encrypted_txt)))
48 log_curr = logn_pr_txt(double2char(current_key(char2double(encrypted_txt))), pr_trans)
49 log_new = logn_pr_txt(double2char(new_key(char2double(encrypted_txt))), pr_trans)
50
51 if log_new >= log_curr
52     accept_new_key = 1
53     prob_accept = 1
54 else
55     prob_accept = exp(log_new - log_curr)
56     if (rand(1) < prob_accept)
57         accept_new_key = 1
58     else
59         accept_new_key = 0
60     end
61 end
62
63 end
  
```

- (i) Apply metropolis.m using frank_decrypt_key as the current key, mystery_decrypt_key as the new key, pr_trans computed from Q3a and frank_encrypted_txt. Report the probability of accepting the new key as returned by metropolis.m.



A screenshot of the MATLAB variable viewer showing a variable named 'prob_accept' of type '1x1 double'. The value is displayed as a 2x2 matrix:

	1	2
1		0

When accept_new_key = 0, , prob_accept = 0

- (ii) Create a new key from frank_decrypt_key, by swapping the 12-th and 13-th elements of the array. Apply metropolis.m using frank_decrypt_key as the current key, the newly generated key, pr_trans computed from Q3a and frank_encrypted_txt. Report the probability of accepting the new key as returned by metropolis.m.

prob_accept1 =

9.0532e-303

When accept_new_key1 = 0, ,
prob_accept1 = 9.0532e-303

(b)

- (i) Report the final decrypted text (decrypt_txt) and its log probability. The final decrypted text is:

decrypted_txt_mcmc =

'sometimes i could cope with the sullen despair that overwhelmed me but sometimes the whirlwind passions of my soul drove me to seek by bodily exercise and by change of place some relief from my intolerable sensations it was during an access of this kind that i suddenly left my home and bending my steps towards the near alpine valleys sought in the magnificence the eternity of such scenes to forget myself and my ephemeral because human sorrows my wanderings were directed towards the valley of chamounix i had visited it frejuently during my boyhood six years had passed since then i was a wreck but nought had changed in those savage and enduring scenes '

And the log probability is: log_pr = -1651.29406848247

- (ii) How are the keys different? How does these differences show up in the final decrypted text? Explain why the algorithm does not give exactly the correct answer?

Decrypted_key_mcmc																										
16	13	4	9	19	20	11	18	17	21	14	22	5	6	15	3	2	8	12	25	7	10	27	24	1	26	23
Frank_decrypted_key																										
16	13	4	9	19	20	11	18	10	21	14	22	5	6	15	3	2	8	12	25	7	17	27	24	1	26	23

In decrypted_key_mcmc and frank_decrypted_key, there is only one difference in the position of “10” and “17” decryption key.

The decryption key “10” represented alphabetic “j”; the decryption key “17” represented alphabetic “q”.

Decrypted_txt_mcmc
sometimes i could cope with the sullen despair that overwhelmed me but sometimes the whirlwind passions of my soul drove me to seek by bodily exercise and by change of place some relief from my intolerable sensations it was during an access of this kind that i suddenly left my home and bending my steps towards the near alpine valleys sought in the magnificence the eternity of such scenes to forget myself and my ephemeral because human sorrows my wanderings were directed towards the valley of chamounix i had visited it frequently during my boyhood six years had passed since then i was a wreck but nought had changed in those savage and enduring scenes
Frank_decrypted_txt
sometimes i could cope with the sullen despair that overwhelmed me but sometimes the whirlwind passions of my soul drove me to seek by bodily exercise and by change of place some relief from my intolerable sensations it was during an access of this kind that i suddenly left my home and bending my steps towards the near alpine valleys sought in the magnificence the eternity of such scenes to forget myself and my ephemeral because human sorrows my wanderings were directed towards the valley of chamounix i had visited it frequently during my boyhood six years had passed since then i was a wreck but nought had changed in those savage and enduring scenes

After the comparison between decrypted_txt_mcmc and frank_decrypted_txt, I noticed that there is only 1 bit different, which is the part that I have marked as yellow, they are “frejuently” and “frequently”, which corresponds to the difference in the decryption keys.

Conclusion: The reason why MCMC didn’t get the exact answer, I guess it is because when the program reached its 15000th iteration, the Markov chain hasn’t converged.

Proposed solution: increase the number of iterations.

(c)

- (i) Report the final decrypted text (decrypt_txt) and its log probability.
The final decrypted text is:

decrypt_txt_mystery_mcmc =

'i went up the bank about fifty yards and then i doubled on my tracks and slipped back to where my canoe was a good piece below the house i qumped in and was off in a hurry i went up stream far enough to make the head of the island and then started across i took off the sun bonnet for i didn t want no blinders on then when i was about the middle i heard the clock begin to strike so i stops and listens the sound come faint over the water but clear eleven when i struck the head of the island i never waited to blow though i was most winded but i shoved right into the timber where my old camp used to be and started a good fire there on a high and dry spot '

Log probability = -1.651294068482473e+03

- (ii) How are the keys different? How does these differences show up in the final decrypted text? Why did the algorithm not give exactly the correct answer?

Decrypted_key_mystery_mcmc																										
24	16	27	2	7	18	26	5	1	9	6	8	10	22	17	3	14	12	11	20	19	4	13	15	25	21	23
Mystery_decrypted_key																										
24	16	27	2	7	18	26	5	1	9	6	8	17	22	10	3	14	12	11	20	19	4	13	15	25	21	23

In decrypted_key_mystery_mcmc and mystery_decrypted_key, there is only one difference in the position of “10” and “17” decryption key.

The decryption key “10” represented alphabetic “j”; the decryption key “17” represented alphabetic “q”.

This difference is the same as the difference between original frank key and mcmc decryption key for frank text!

I guess the reason should be similar to that “why algorithm didn’t give the exact answer for frank text”, that is: the number of iterations is not enough for the algorithm to converge. However, I’m not sure about why they will appear to have the same difference, and I’ll try to figure it out in the future.

decrypt_txt_mystery_mcmc
i went up the bank about fifty yards and then i doubled on my tracks and slipped back to where my canoe was a good piece below the house i qumped in and was off in a hurry i went up stream far enough to make the head of the island and then started across i took off the sun bonnet for i didn t want no blinders on then when i was about the middle i heard the clock begin to strike so i stops and listens the sound come faint over the water but clear eleven when i struck the head of the island i never waited to blow though i was most winded but i shoved

right into the timber where my old camp used to be and started a good fire there on a high and dry spot '

Mystery_decrypted_txt

i went up the bank about fifty yards and then i doubled on my tracks and slipped back to where my canoe was a good piece below the house i jumped in and was off in a hurry i went up stream far enough to make the head of the island and then started across i took off the sun bonnet for i didn t want no blinders on then when i was about the middle i heard the clock begin to strike so i stops and listens the sound come faint over the water but clear eleven when i struck the head of the island i never waited to blow though i was most winded but i shoved right into the timber where my old camp used to be and started a good fire there on a high and dry spot

The difference between the 2 pieces of text implies the difference in the decryption key.