TEKNIK PENYEMBUNYIAN MESEJ DALAM STEGANOGRAFI TEKS MENGGUNAKAN PENDEKATAN WARNA RGB DAN PENEMPATAN RAWAK

BAHARUDIN BIN OSMAN

DOKTOR FALSAFAH
UNIVERSITI UTARA MALAYSIA
2020

Awang Had Salleh
Graduate School
of Arts And Sciences

Universiti Utara Malaysia

PERAKUAN KERJA TESIS / DISERTASI
(Certification of thesis / dissertation)

Kami, yang bertandatangan, memperakukan bahawa
(We, the undersigned, certify that)

BAHARUDIN OSMAN

calon untuk Ijazah
(candidate for the degree of) **PhD**

telah mengemukakan tesis / disertasi yang bertajuk:
(has presented his/her thesis / dissertation of the following title):

"TEKNIK PENYEMBUNYIAN MESEJ DALAM STEGANOGRAFI TEKS MENGGUNAKAN PENDEKATAN WARNA RGB DAN PENEMPATAN RAWAK"

seperti yang tercatat di muka surat tajuk dan kulit tesis / disertasi.
(as it appears on the title page and front cover of the thesis / dissertation).

Bahawa tesis/disertasi tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan, sebagaimana yang ditunjukkan oleh calon dalam ujian lisan yang diadakan pada : ***26 Februari 2020.***
That the said thesis/dissertation is acceptable in form and content and displays a satisfactory knowledge of the field of study as demonstrated by the candidate through an oral examination held on: ***February 26, 2020.***

Pengerusi Viva: *(Chairman for VIVA)*	**Assoc. Prof. Dr. Mohd Hasbullah Omar**	Tandatangan *(Signature)*
Pemeriksa Luar: *(External Examiner)*	**Assoc. Prof. Dr. Ainuddin Wahid Abdul Wahab**	Tandatangan *(Signature)*
Pemeriksa Luar: *(External Examiner)*	**Assoc. Prof. Ts Dr. Nur Izura Udzir**	Tandatangan *(Signature)*
Nama Penyelia/Penyelia-penyelia: *(Name of Supervisor/Supervisors)*	**Assoc. Prof. Dr. Azman Yasin**	Tandatangan *(Signature)*
Nama Penyelia/Penyelia-penyelia: *(Name of Supervisor/Supervisors)*	**Dr. Mohd Nizam Omar**	Tandatangan *(Signature)*

Tarikh:
(Date) ***February 26, 2020***

Kebenaran Mengguna

Tesis ini dikemukakan sebagai memenuhi sebahagian daripada keperluan pengijazahan Doktor Falsafah di Universiti Utara Malaysia. Saya bersetuju membenarkan pihak perpustakaan universiti mempamerkan sebagai bahan rujukan umum. Saya juga bersetuju bahawa sebarang bentuk salinan sama ada secara keseluruhan atau sebahagian daripada tesis ini untuk tujuan akademik adalah dibenarkan dengan kebenaran penyelia tesis atau Dekan Awang Had Salleh Graduate School of Arts and Sciences. Sebarang bentuk salinan dan cetakan bagi tujuan komersial adalah dilarang sama sekali tanpa kebenaran bertulis daripada penulis. Pernyataan rujukan kepada penulis dan Universiti Utara Malaysia perlulah dinyatakan jika terdapat sebarang rujukan ke atas tesis ini. Kebenaran untuk menyalin dan menggunakan tesis sarjana ini sama ada secara keseluruhan ataupun sebahagian daripadanya hendaklah dipohon melalui:

Dekan Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 Sintok Kedah

Abstrak

Steganografi merupakan teknik untuk melindungi kerahsiaan dan integriti data di dalam medium pelindung agar data yang disembunyikan tidak dicurigai. Penyembunyian mesej di dalam medium teks boleh dilakukan terhadap pelbagai atribut teks seperti jenis, gaya, saiz, warna tulisan dan sebagainya untuk menjana teks stego. Kajian ini mengenalpasti dua masalah utama yang mendorong kepada kecurigaan teks stego iaitu perubahan warna yang ketara terhadap teks stego yang dijana dan perwakilan aksara mesej rahsia secara statik menggunakan pemilihan lokasi penyembunyian secara berjujukan. Oleh itu, objektif utama kajian ialah mencadangkan penggunaan nilai tertentu bagi setiap kombinasi warna Merah, Hijau, Biru (RGB) untuk mengurangkan perubahan warna ketara terhadap teks stego yang dijana. Kajian ini turut mencadangkan kaedah perwakilan mesej rahsia berdinamik berdasarkan lokasi aksara terpilih secara rawak. Jadual *Homophonic* sifer diadaptasi sebagai kaedah untuk menjana aksara mesej rahsia agar bersifat dinamik. Disamping itu, Teorem Baki Hasil Bahagi Peringkat Kedua dicadangkan untuk menukarkan aksara mesej rahsia ke bentuk perwakilan 3D dengan memetakan nilai (x,y,z) kepada warna RGB. Model kiub warna RGB bernilai di antara RGB(0,0,0) hingga RGB(15,15,15) digunakan untuk memformatkan aksara teks pelindung terpilih menggunakan Penjana Nombor Pseudorawak. Prestasi teks stego yang dihasilkan melalui kajian ini dinilai menggunakan tiga ukuran utama iaitu kapasiti, ketakbolehkelihatan dan keteguhan. Hasil kajian mendapati kaedah yang dicadangkan menghasilkan prestasi yang lebih baik dengan kapasiti penyembunyian mesej rahsia meningkat sebanyak 41.31% serta skala ketakbolehkelihatan Jaro Winkler‟s bersamaan 1. Prestasi keteguhan bagi teks stego terbukti apabila tiada perbezaan di antara teks stego dengan teks pelindung sebelum dan selepas proses pemampatan. Kesimpulannya, kaedah yang dicadangkan berjaya mengurangkan keketaraan terhadap perubahan warna teks stego terjana yang menjurus kepada kecurigaan wujudnya mesej tersembunyi. Malahan, kaedah ini berupaya menghasilkan perwakilan mesej rahsia berdinamik dengan menggunakan teks pelindung tunggal.

Kata Kunci: Jadual *Homophonic* sifer, Kecurigaan, Teorem Baki Hasil Bahagi Peringkat Kedua, Steganografi teks, Model kiub warna RGB

Abstract

Steganography is a technique that protects the confidentiality and integrity of data in a protective medium from suspicion of hidden data. The hiding of a message in a text medium can be performed on various text attributes such as type, style, size, and font color to generate a stego text. This study have identified two main problems that lead to the suspicion towards the stego text which is the obvious change of colors of the generated stego and the static representation of the secret message characters using sequential selection of hiding location. Therefore, the main objective of this study is to propose the use of specific value for each combination of Red, Green, Blue (RGB) color to reduce the generated stego text obvious color changes. This study also recommends a dynamic secret message representation method based on a randomly selected character location. A Homophonic Cipher Table was adapted as a method to generate the dynamic secret message characters. Besides, the Second Quotient Remainder Theorem was proposed to convert the secret message characters into a 3D representation by mapping (x,y,z) values to RGB color. The RGB color cube model values of RGB(0,0,0) to RGB(15,15,15) were used to format a selected cover text characters using the Pseudorandom Number Generator. The performance of stego text produced in this study was evaluated using three main measures namely capacity, imperceptibility, and robustness. The results revealed that the proposed method produces a better performance of secret message hiding by 41.31% increase in capacity and the Jaro Winkler's scale imperceptibility score of 1. The performance of stego text is proven to be robust as there is no difference compared to the cover text before and after the compression process. In conclusion, the proposed method has successfully reduced the generated stego text obviousness in the change of colors that lead to suspicion of existence of hidden message. Beside, this method also capable of producing dynamic secret messages using a single cover text.

Keywords: Homorphonic Cipher Table, Suspicious, Second Quotient Remainder Theorem, Text steganography, RGB color cube model

Penghargaan

Alhamdullillah, syukur saya kehadrat Allah S.W.T kerana dengan rahmatNya dapat saya menyiapkan tesis ini. Selawat dan salam ke atas junjungan besar Nabi Muhammad S.A.W, ahli keluarga dan sahabat baginda.

Pertama sekali saya ingin mengucapkan jutaan terima kasih yang tidak terhingga kepada kedua-dua penyelia saya iaitu Prof. Madya Dr. Azman Bin Yasin dan Dr. Mohd Nizam Bin Omar di atas pengorbanan mereka membimbing, mengajar, memberi teguran dan menaikkan semangat saya sepanjang menjalankan penyelidikan ini. Tanpa sokongan dan bantuan anda, tesis ini mungkin tidak dapat disiapkan. Tidak dilupakan juga rakan-rakan pensyarah dan rakan-rakan seperjuangan yang terus memberi sokongan berterusan dengan memberi idea, komen dan sokongan untuk memantapkan lagi kajian ini.

Saya juga mengucapkan terima kasih kepada Dekan Pusat Pengajian Pengkomputeran UUM, Prof. Dr. Huda Hj Ibrahim yang sentiasa memantau perkembangan dan memberi sokongan dan dorongan kepada saya untuk menyiapkan tesis ini. Jutaan terima kasih kepada Kementerian Pendidikan Malaysia selaku penaja serta majikan Universiti Utara Malaysia yang meluluskan cuti belajar untuk membolehkan saya melanjutkan pengajian di peringkat PhD ini.

Terima kasih kepada ibu tersayang, Bashah Hj Daud yang sentiasa mendoakan kejayaan ini. Buat isteri dan keluarga tersayang; Hamiza Hj Mohd Munir, Badrul Hazmi, Nur Badrina, Badrul Hakimi dan Nur Batrisyia, terima kasih di atas sokongan moral, kesabaran serta pengorbanan tanpa henti yang diberikan sepanjang perjalanan menyiapkan tesis ini.

Akhir kalam, terima kasih kepada semua yang telah membantu saya untuk menyiapkan tesis ini sama ada secara langsung atau tidak langsung. Semoga Allah S.W.T merahmati dan memberkati segala ilmu dan pengorbanan yang diberikan.

Isi Kandungan

Senarai Jadual

Senarai Rajah

Senarai Lampiran

Glosari

Dicampuradukkan	*Scrambled*
Fungsi Nombor Pseudorawak	*Pseudorandom Number Function*
Kaedah Berasaskan Format	*Format Base Method*
Kaedah Linguistik	*Linguistic Method*
Kaedah Statistik dan Rawak	*Random and Statistical Method*
Kapasiti	*Capacity*
Ketakbolehkelihatan	*Imperceptibility*
Ketakbolehkesanan	*Undetectibility*
Keteguhan	*Robustness*
Kod Sifer	*Cipher Code*
Medium Pelindung	C*over Medium*
Mesej Rahsia	*Secret Message*
Nilai Kecergasan	*Fitness Value*
Padanan Corak Piksel	*Matching Pattern Pixel*
Pemilihan	*Selection*
Pembiakan	*Propagation*
Penjana Nombor Pseudorawak	*Pseudorandom Number Generators*
Penjana Nombor Rawak Sebenar	*True Random Number Generators*
Penjanaan Rawak dan Statistik	*Statistical and Random Generation*
Penggantian	*Substitution*
Ruang Lewah	*Redundant Space*
Skor Jaro	*Jaro Score*
Steganografi Berasaskan Format	*Format Based Steganography*
Steganografi Linguistik	*Linguistic Steganografi*
Suntikan	*Injection*
Teorem Baki Hasil Bahagi Peringkat Kedua	*Second Quotient Remainder Theorem*
Teks Pelindung	*Cover Text*
Teks Stego	*Stego Teks*
Tera Air	*Watermark*

Senarai Ringkasan

BER	*Bit Error Rate*
CALP	*Changing in Alphabet Letter Patterns*
CASE	*Capital Shape Alphabet Encoding*
CEBTS	*Crossover Encryption Based Text Steganography*
EDDS	*Evolution Detection Steganalysis System*
VERT	*Vertical Straight Line*
FNP	*Fungsi Nombor Pseudorawak*
GATS	*Genetic Algorithm Based Text Steganography*
MMSN	*Mathematical Model System Number*
PNP	Penjana Nombor Pseudorawak
PNRS	Penjana Nombor Rawak Sebenar
POS	*Part-of-Speech*
QRT	*Quotient Remainder Theorem*
QUAD	*Quadruple Categorization*
SnR	*Symmetric and Reflection*
SQRT	*Second Quotient Remainder Theorem*

BAB SATU
PENGENALAN

Bab ini menghuraikan tentang kepentingan bidang penyembunyian maklumat yang merangkumi steganografi dan kriptografi serta diikuti dengan penerangan mengenai isu-isu yang ditimbulkan yang menjurus kepada persoalan keperluan penyembunyian mesej dalam steganografi teks. Objektif dan skop penyelidikan diutarakan bagi menggarap signifikan kajian ini dalam bidang ilmu dan juga pihak-pihak tertentu. Bab ini diakhiri dengan rumusan organisasi tesis ini.

1.1 Latar Belakang Kajian

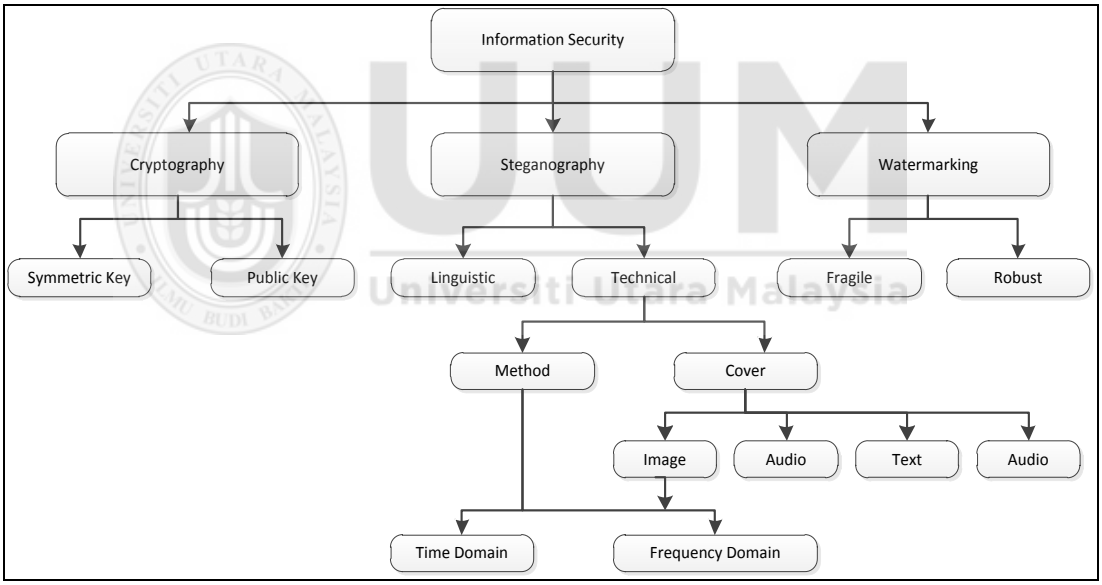
Internet merupakan platform utama digunakan oleh individu atau organisasi untuk pertukaran pelbagai bentuk maklumat seperti teks, imej, audio, video dan sebagainya. Selari dengan itu, peningkatan jumlah komunikasi melalui Internet turut meningkat dengan ketara sama ada untuk tujuan penghantaran maklumat sulit atau sebaliknya dan ini memerlukan satu sistem keselamatan yang efisien untuk tujuan pemeliharaan data. Menurut Mahato, Khan dan Yadav (2017) komunikasi data melalui Internet telah meningkat dengan ketara dan menyebabkan trafik data yang tinggi serta menimbulkan kebimbangan terhadap keselamatan data yang dihantar. Perkembangan kemajuan teknologi Internet menyebabkan maklumat boleh dikongsi di mana sahaja di seluruh dunia. Maniriho dan Ahmad (2017) menegaskan bahawa maklumat yang dihantar melalui talian Internet sentiasa mendapat perhatian penceroboh yang seterusnya menyebabkan keselamatan data sensitif menjadi satu masalah yang serius dan perlu dipelihara agar tidak berlaku kebocoran (Bhat, Prabhu, & Renuka, 2017; Joseph & Vishnukumar, 2015). Pertumbuhan Internet yang pantas,

mendorong keperluan untuk melindungi data sensitif daripada dicapai oleh pihak yang tidak bertanggungjawab (Pujari & Shinde, 2016). Justeru itu, penghantaran maklumat terutamanya melalui talian komunikasi perlulah dikawal bagi menghindarinya daripada dikesan atau dipantau oleh pihak ketiga semasa berlakunya proses pemindahan data. Pihak ketiga yang sentiasa memantau penghantaran data akan cuba untuk mendapatkan maklumat penting yang dihantar melalui talian komunikasi (Krishnan, Thandra, & Baba, 2017). Maklumat merupakan aset yang bernilai dan kegagalan menjaganya dengan baik boleh menyebabkan ia di ceroboh oleh pihak yang tidak bertanggungjawab. Kerahsiaan dan integriti data diperlukan untuk melindungi kerahsiaan data (Babu, 2010) serta memastikan data yang diterima tidak diubah oleh pihak ketiga bagi memastikan data tersebut adalah sahih atau asli.

Seiring dengan peningkatan penggunaan Internet untuk tujuan pemindahan data maka satu sistem keselamatan yang dapat melindungi data seperti menggunakan kaedah penyulitan sangat diperlukan. Penyulitan (*encryption*) merupakan teknik yang digunakan untuk menukarkan sesuatu teks ke bentuk yang tidak boleh dibaca (*unreadable*). Namun begitu, menurut Ray, Sanyal, Das dan Nath (2012), penggunaan kaedah penyulitan sahaja adalah tidak memadai untuk melindungi data sulit seperti maklumat perbankan atau sebarang maklumat sulit yang lain. Penyulitan adalah berkait rapat dengan kriptografi yang merupakan sebahagian daripada sistem keselamatan, namun ia tidak menyediakan ciri-ciri kerahsiaan (Krishnan et al., 2017) serta tidak dapat melindungi data secara berkesan (Pujari & Shinde, 2016). Pihak ketiga akan sentiasa memerhati data yang dihantar dan dengan mudah mengenal pasti data yang dihantar. Oleh itu, kaedah penyembunyian maklumat *(information hiding)* merupakan kaedah alternatif yang boleh digunakan untuk melindungi data kerana

mesej yang dihantar bukan sahaja disulitkan, malah kewujudan mesej tersebut tidak disedari oleh penjenayah siber semasa berlakunya proses pemindahan data (Baawi, Mokhtar, & Sulaiman, 2018; Zhang, Huang, Wang, Lin, & Gao, 2017). Malah Krishnan et al., (2017) mencadangkan steganografi merupakan salah teknik yang boleh digunakan untuk mengatasi masalah ini.

Steganografi, kriptografi dan tera air (*watermarking)* merupakan kaedah yang digunakan untuk melindungi maklumat agar ia tidak dapat dikesan oleh pihak ketiga. Ketiga-tiga kaedah ini diklasifikasikan di bawah sistem keselamatan maklumat seperti yang ditunjukkan di dalam Rajah 1.1 (Amirthrajan & Rayappan, 2013).



Rajah 1.1 Pengelasan Sistem Keselamatan Maklumat (Amirthrajan & Rayappan, 2013)

Kriptografi dan steganografi merupakan dua pendekatan di dalam sistem keselamatan maklumat yang digunakan dengan meluas untuk melawan ancaman kepada keselamatan maklumat (Almuhammadi & Al-shaaby, 2017). Kriptografi merupakan seni penulisan kod rahsia dan matlamatnya adalah untuk memastikan data tidak boleh

dibaca oleh pihak ketiga dan ianya tidak selalu menyediakan komunikasi yang selamat (Din, Samsudin, & Lertkrai, 2012). Tujuan utama kriptografi adalah untuk melindungi mesej rahsia yang dihantar di mana mesej tersebut akan dicampuradukkan (*scrambled*) atau ditukar dalam bentuk yang berbeza daripada bentuk asal (Antony, Sobin, & Sherly, 2012). Menurut Reddy, Subramanyam dan Reddy (2012), teknik kriptografi akan menukarkan mesej rahsia ke bentuk yang tidak boleh dibaca, namun begitu mesej yang tidak boleh dibaca ini mudah menarik perhatian pihak yang berkeinginan untuk menganalisis mesej tersebut.

Ini berbeza dengan steganografi, di mana steganografi merupakan kaedah yang digunakan untuk menyembunyikan maklumat rahsia di dalam medium-medium tertentu (Agath, Sidpara, & Upadhyay, 2018) dan teks yang dihantar masih boleh dibaca oleh sesiapa sahaja tanpa menyedari kewujudan mesej rahsia yang disembunyikan (Joseph & Vishnukumar, 2015) atau dicurigai oleh pihak ketiga. Selain itu, steganografi lebih sukar dan kompleks untuk diserang oleh pihak ketiga berbanding kriptografi (Zielińska, Mazurczyk, & Szczypiorski, 2014) dan kecurigaan yang wujud terhadap mesej yang dihantar dapat dihapuskan (Bhat et al., 2017).

Tera air *(Watermark*) merupakan proses menyembunyikan penanda seperti label, tandatangan atau hak cipta ke dalam media digital seperti teks, audio, video dan imej (Alotaibi & Elrefaei, 2018) dan kebiasaannya digunakan untuk perlindungan hak cipta atau tanda dagangan (*trademark*) (Kumar, Malik, Singh, & Chand, 2016). Penanda yang dimasukkan ke dalam media digital boleh dilihat dengan mata kasar atau sebaliknya (Douglas, Bailey, Leeney, & Curran, 2018) dan tidak boleh diubah dengan mudah (Atoum, 2018) dan ini berbeza dengan steganografi di mana ketakbolehkelihatan merupakan satu keperluan utama.

Steganografi digunakan di dalam pelbagai bidang seperti ketenteraan, diagnosis perubatan, maklumat berkaitan perniagaan, kewangan dan sebagainya (Malik, Sikka, & Verma, 2017). Walaupun steganografi telah bermula sejak zaman dahulu lagi, namun menurut Singh dan Singh (2013), minat penyelidik terhadap steganografi makin meningkat dan menarik perhatian para penyelidik serta berkembang semula atas dua faktor. Pertama, perkembangan industri penerbitan dan penyiaran yang pesat telah menarik minat pihak-pihak tertentu untuk menggunakan teknik penyembunyian data seperti menyembunyikan tanda hak cipta dan nombor siri di dalam filem-filem digital, rakaman audio, buku-buku dan produk multimedia yang bertujuan untuk mengenal pasti ketulenan sesuatu produk. Kedua, terdapat juga sektor-sektor kerajaan dan swasta yang tidak membenarkan perkhidmatan penyulitan digunakan di dalam organisasi dan menjadikannya sebagai satu polisi di dalam organisasi. Selain itu, terdapat beberapa negara barat seperti di Amerika Syarikat dan British melarang penggunaan kaedah penyulitan kerana pihak ketiga sentiasa peka terhadap komunikasi yang berlaku dan ia tidak begitu selamat (Kataria, Singh, Kumar, & Shekhawat, 2013). Justeru itu, kedua-dua faktor di atas telah mendorong penyelidik menggunakan kaedah steganografi untuk mengkaji dan mengenal pasti teknik yang sesuai bagi menyembunyikan maklumat rahsia di dalam medium pelindung (*cover medium*).

Menurut Sumathi, Santanam dan Umamaheswari (2013), steganografi merupakan topik panas di dalam domain penyembunyian maklumat dan kebanyakan kajian memfokuskan kepada imej, audio dan video (Osman, Yasin, & Omar, 2016) sebagai medium pelindung dan kurang tumpuan diberikan kepada medium teks. Namun, teks merupakan medium utama digunakan secara meluas pada masa kini untuk penghantaran data seharian (Bhaya, Rahma, & Al-nasrawi, 2013) dan masih

mempunyai ruang untuk kajian dijalankan (Iyer, 2017). Menurut Ahvanooey, Li, Shim, dan Huang (2018), teks merupakan salah satu sumber data utama dan merupakan media digital yang paling banyak digunakan di Internet, bahagian penting laman web, buku, artikel, kertas harian dan sebagainya. Selain daripada itu, kebanyakan dokumen-dokumen penting wujud dalam bentuk teks seperti surat pelantikan, sijil, laporan, dokumen sulit dan sebagainya (Din & Utama, 2018). Tambahan pula, dari aspek teknikal fail teks memerlukan kurang memori dan transmisi penghantaran data lebih pantas dan mudah berbanding dengan medium lain (Kouser, Khan, & Qamar, 2017; Kumar, Pabboju, & Desai, 2014). Penggunaan teknik penyembunyian yang sesuai berkemungkinan boleh meningkatkan kapasiti penyembunyian walaupun ruang penyembunyian adalah terhad.

Kriptografi dan steganografi merupakan dua pendekatan yang berbeza, di mana kriptografi boleh di implementasi di dalam steganografi tetapi tidak sebaliknya seperti yang dijelaskan di dalam Jadual 2.1. Walau bagaimanapun, terdapat pelbagai teknik kriptografi digunakan di dalam proses penyulitan sesuatu mesej seperti AES,DES, RSA dan Blowfish sebagainya. Namun, hasil akhir daripada proses penyulitan merupakan satu sifer teks yang masih menunjukkan kewujudan pelbagai aksara berulang seperti yang ditunjukkan di dalam Jadual 1.1. Selain itu, saiz teks sifer yang dihasilkan selepas proses penyulitan melebihi saiz asal mesej rahsia yang mana ini akan menjejaskan kapasiti teks stego yang dijana. Kajian yang dijalankan oleh Malalla dan Shareef, (2016) menunjukkan saiz mesej rahsia meningkat sehingga 10-12 kali ganda selepas melalui proses penyulitan. Selain itu, kajian yang dijalankan oleh Patel dan Patro, (2017) menunjukkan bahawa saiz optimum bit yang diperlukan untuk penyulitan satu aksara meningkat dan berbeza di antara satu teknik dengan

teknik yang lain (DES-27bit, RSA-44 bit, Blowfish-128 bit dan AES-256 bit). Ini menunjukkan saiz mesej rahsia menggunakan teknik kriptografi akan menyebabkan saiz yang diwakilkan semakin bertambah dan seterusnya menyebabkan kapasiti penyembunyian akan semakin berkurang.

```
Mesej Rahsia    :meetyouatten                   (12 aksara)
Kekunci         :2b7e151628aed2a6abf71589       (24 aksara)
```

Jadual 1.1

Teknik-Teknik Kriptografi

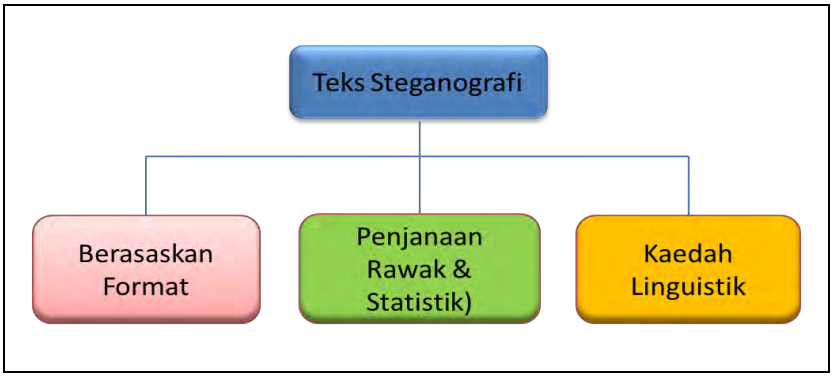
Teknik Kriptografi	Mesej Selepas Penyulitan	Saiz Mesej Selepas Penyulitan
AES	Edjd3Cwxn/uyPzCliT1tyA==	24
DES	s22qpZY9hxPPeMq6HxfCcg==	24
BLOWFISH	Ucwv7DCV2ZVOXGQgKycgRA==	24
SEED	xlvkvla3qT2dvUx/xa/M8g==	24
RSA	J2D53KfZ7s5f4zUTqSwGUZdQ77fXht RJSCAeP5Pab+dB95wrgnpiD99ym1Va NY+DRKX/XQ0c/iThvqOXsXElVg==	88

Walaupun kriptografi boleh digunakan di dalam steganografi, namun kajian ini tidak memfokuskan kepada teknik kriptografi kerana ia merupakan dua pendekatan yang berbeza selain tumpuan utama ialah perwakilan aksara yang terdapat di dalam mesej rahsia terutama aksara yang berulang. Selain itu, kajian yang dijalankan oleh Patil menunjukkan bahawa saiz optimum bit yang diperlukan untuk penyulitan satu aksara meningkat dan berbeza di antara satu teknik dengan teknik yang lain (DES-27bit, RSA-44 bit, Blowfish-128 bit dan AES-256 bit). Ini menunjukkan bahawa saiz mesej rahsia menggunakan teknik kriptografi akan menyebabkan saiz yang dihasilkan semakin bertambah dan seterusnya menyebabkan kapasiti penyembunyian akan semakin berkurang jika digunakan bersama steganografi. Oleh itu, walaupun

kriptografi boleh digunakan di dalam steganografi, namun kajian ini tidak memfokuskan kepada teknik kriptografi kerana ia merupakan dua pendekatan yang berbeza selain tumpuan utama ialah perwakilan aksara yang terdapat di dalam mesej tersembunyi terutama aksara yang berulang.

Kajian teks steganografi yang dijalankan adalah berasaskan kepada data digital di mana teks steganografi menggunakan kurang memori serta boleh berkomunikasi dengan banyak maklumat serta memerlukan kos percetakan yang rendah berbanding dengan medium lain (Kingslin & Kavitha, 2015; T. Kumar, Abhinav, Jyoti, & Nehra, 2014; Memon, Khowaja, & Kazi, 2008; Shirali-shahreza, 2006). Oleh itu, tiada sebarang kos percetakan terlibat di dalam kajian ini kerana teknik penyelidikan adalah berdasarkan data digital.

Steganografi boleh dibahagikan kepada dua kategori iaitu; Steganografi Bahasa Semula Jadi (*Natural Language Steganography*) dan Steganografi Teknikal. Menurut Din dan Samsudin, (2009) dan Vennice et al., (2012) steganografi teknikal melibatkan penyembunyian maklumat di dalam medium seperti imej, audio dan video manakala steganografi Bahasa Semula Jadi melibatkan penyembunyian maklumat di dalam teks. Steganografi berasaskan teks boleh diklasifikasikan kepada tiga kategori iaitu Berasaskan Format (*Format based*), Penjanaan Rawak dan Statistik (*Statistical and Random Generation)* serta Steganografi Linguistik (*Linguistic Steganografi*) seperti yang ditunjukkan di dalam Rajah 1.2 (Bennett, 2004). Ketiga-tiga kaedah ini dibincangkan secara terperinci pada sub-topik 2.3.1, 2.3.2 dan 2.3.3.



Rajah 1.1 Kategori Steganografi Teks (Bennett, 2004)

Steganografi teks berasaskan format memfokuskan kepada penyembunyian sesuatu mesej rahsia berasaskan kepada ciri-ciri teks pelindung seperti gaya tulisan, ruang kosong, ruang antara perenggan, bentuk abjad dan sebagainya serta merupakan kaedah yang popular digunakan oleh kebanyakan penyelidik.

Manakala steganografi penjanaan rawak dan statistik menyembunyikan mesej rahsia di dalam teks stego yang dijana berdasarkan kepada ciri-ciri seperti panjang perkataan, frekuensi huruf, jujukan aksara, jujukan perkataan, dan sebagainya (Baawi et al., 2018; Bhattacharyya, Banerjee, & Sanyal, 2011). Kaedah statistik dan rawak menjana teks pelindung menggunakan dua teknik iaitu teknik rawak dan teknik statistik. Teknik rawak menyembunyikan aksara menggunakan teknik penggantian secara jujukan rawak manakala teknik statistik menentukan nilai statistik seperti min, varian, frekuensi huruf atau perkataan, panjang perkataan dan sebagainya untuk menentukan jumlah maklumat yang boleh disembunyikan. Kaedah ini akan menjana teks stego berdasarkan kepada ciri-ciri statistik sesuatu dokumen (Saraswathi dan Kingskin, 2014).

Akhir sekali, steganografi linguistik melibatkan penyembunyian mesej rahsia dengan melakukan penggantian perkataan atau tatabahasa terhadap teks pelindung dengan

menggunakan teknik penggantian sintaks atau semantik sesuatu bahasa. Bagi teknik sintaks, tanda bacaan seperti tata tanda koma (,) , seruan (!) , titik noktah (.), tanda tanya (?) dan sebagainya digunakan untuk menyembunyikan mesej rahsia manakala bagi teknik semantik, kebiasaannya penggantian perkataan sinonim digunakan untuk menyembunyikan mesej rahsia. Kaedah ini memerlukan tahap pengetahuan yang tinggi terhadap nahu sesuatu bahasa agar penjanaan teks stego tidak mengganggu semantik ayat yang dijana. Kajian yang dijalankan oleh Sunariya, Din, & Mahmudin (2016) mendapati hanya 25% penyelidik memfokuskan kajian di dalam linguistik steganografi berbanding steganografi teks.

1.2 Motivasi Kajian

Pada masa kini, memastikan maklumat digital disimpan dalam keadaan selamat daripada penyalahgunaan merupakan kriteria amat penting di dalam sistem keselamatan maklumat. Ini kerana penggodam atau pihak yang tidak bertanggungjawab akan sentiasa mencuba memecahkan kaedah-kaedah atau protokol sedia ada untuk mendapat maklumat sensitif (Samer, Zaid, Sami, & Zainab, 2015). Walaupun penjanaan versi teks stego sedia ada mempunyai pengubahsuaian yang minimum dilakukan terhadap teks pelindung bagi menghindarkannya daripada dikesan oleh visual manusia, namun terdapat steganalisis moden yang dapat mengesan perubahan minimum tersebut (Dasgupta, Mondal, & Dutta, 2013). Oleh itu, sebarang pengubahsuaian yang dilakukan terhadap teks pelindung perlulah tidak merosakkan integriti teks pelindung.

Ahvanooey, Li, Hou, Rajput dan Yini (2019) dalam kajiannya merumuskan bahawa tidak dapat dilihat (*invisibility*) atau dikenali sebagai ketakbolehkelihatan

(*imperceptibility*) atau ketakbolehkesanan (*undetectibility*) atau ketelusan (*transperancy*) boleh diukur menggunakan skala Jaro-Winkler. Menurut Malalla & Shareef (2017) secara amnya, ketakbolehkesanan atau ketakbolehkelihatan boleh diimplementasikan dengan melakukan pengubahsuaian yang tidak kelihatan terhadap teks stego. Justeru itu, diperhatikan bahawa ketakbolehkelihatan merupakan satu keperluan penting di dalam steganografi. Kesinambungannya, ketakbolehkelihatan telah mendorong motivasi kepada penyelidik untuk mereka bentuk teknik steganografi yang boleh melindungi teks stego yang dijana daripada dikesan oleh pihak ketiga di samping dapat meningkatkan kapasiti mesej rahsia di dalam teks stego yang dijana.

Sebagai kesimpulannya, steganografi memainkan peranan penting di dalam memelihara keselamatan dokumen seperti urus niaga data rahsia, pemeliharaan data e-dagang, pemeliharaan hak cipta digital dan pengenalan maklumat terutamanya semasa proses pemindahan maklumat melalui Internet (Agath et al., 2018; Din, Ani, & Samsudin, 2012). Selain itu, kebocoran maklumat rahsia seperti dokumen-dokumen rahsia kerajaan, kata kunci dan sebagainya sering kali didengari di media-media penyiaran. Ini disebabkan kebocoran maklumat mudah tersebar melalui komunikasi Internet seperti e-mel, media sosial dan sebagainya. Antara faktornya ialah maklumat-maklumat rahsia tersebut tidak dipelihara sebaik yang mungkin semasa berlakunya proses pemindahan data. Oleh itu, sebagai alternatif teknik steganografi boleh digunakan untuk menyediakan satu sistem keselamatan yang lebih cekap agar dokumen rahsia tidak mudah diterjemahkan apabila tersebar (Agath et al., 2018) serta tidak menimbulkan kecurigaan oleh sistem visual manusia (Ahvanooey, Li, Hou, Rajput, & Yini, 2019).

1.3 Pernyataan Masalah

Steganografi teks menggunakan medium digital masih mendapat sambutan penyelidik disebabkan beberapa faktor, antaranya ialah kadar pemindahan data yang pantas, penggunaan ruang ingatan yang kecil (Arya & Soni, 2018), serta boleh dicapai dengan jalur lebar yang rendah berbanding dengan medium-medium lain (Malik et al., 2017; Kingslin & Saraswathi, 2015; Mahajan & Singh, 2012). Selain itu, teks merupakan medium utama digunakan di dalam komunikasi harian di seluruh dunia sama ada dalam bentuk digital atau bercetak (Ahvanooey, Li, Shim, et al., 2018; Tutuncu & Hassan, 2015) dan digunakan dengan meluas di dalam kebanyakan organisasi (Krishnan et al., 2017).

Kapasiti penyembunyian, keteguhan dan ketakbolehkelihatan merupakan beberapa isu utama yang mempengaruhi prestasi teks stego di dalam domain steganografi teks (Al-Azzawi, 2018). Menurut Ahvanooey et al., (2019), penyembunyian mesej di dalam teks pelindung mestilah tidak boleh dilihat (*invisible*) dan perlu mengelakkan kecurigaan terhadap sistem visual manusia. Selain itu, menurut Archana, Judice dan Kaliyamurthie (2013), dua cabaran utama steganografi ialah isu kapasiti (*capacity*) penyembunyian dan ketakbolehkelihatan mesej rahsia di dalam teks stego yang dihasilkan. Penyembunyian mesej rahsia di dalam teks pelindung boleh menyebabkan perubahan terhadap teks pelindung. Oleh itu, sensitiviti terhadap perubahan teks pelindung merupakan isu penting di dalam steganografi (Shivani, Yadav, & Batham, 2015) seperti perubahan saiz fail, gaya, jenis, warna, bentuk, format teks, struktur ayat dan sebagainya. Menurut Rasmi dan Mohanapriya (2016), kaedah steganografi yang baik mestilah tidak mengubah atribut (*attribute*) medium pelindung (warna, jenis

tulisan, gaya tulisan, saiz tulisan, latar belakang, dan sebagainya) dengan ketara selepas berlakunya proses penyembunyian yang dikenali sebagai ketakbolehkesanan.

Teknik penyembunyian menggunakan ruang kosong seperti ruang kosong di antara perkataan, ruang kosong di antara perenggan dan ruang kosong di hujung ayat dan sebagainya mempunyai kelemahan dari segi kapasiti penyembunyian yang rendah disebabkan jumlah ruang kosong yang terhad yang terdapat di dalam sesuatu teks pelindung. Menurut Kouser (2016) kelemahan utama teknik ruang kosong ialah saiz teks pelindung yang besar diperlukan untuk menyembunyikan sebilangan kecil bit sahaja. Selain itu, teknik ini juga mempunyai risiko kehilangan mesej rahsia terutama semasa melakukan proses pemampatan dan penyahmampatan. Kehilangan satu atau lebih ruang kosong semasa melakukan proses pemampatan dan penyahmampatan akan menyebabkan berlakunya ralat terhadap mesej yang diekstrakkan. Kajian yang dijalankan oleh Rauf, Rose, Jamal, dan Nur Hafizah (2014) mendapati bahawa steganalisis telah berjaya mengesan steganografi teks mesej rahsia dengan efektif di mana prestasi pengesanan mencecah sehingga 96.67% menggunakan teknik *visualization*.

Kajian yang dijalankan oleh Wang, Huang, Chen, Yang dan Miao (2013) terhadap kaedah linguistik telah mengenal pasti bahawa, kesalahan penggantian perkataan menyebabkan teks stego yang dijana dicurigai dan berpotensi tinggi untuk dikesan melalui steganalisis kerana kualiti teks stego yang dihasilkan tidak mengikut nahu bahasa yang digunakan (Bhattacharyya, Indu, Dutta, Biswas, & Sanyal, 2011). Kaedah ini tidak efektif digunakan terutama bagi mesej yang panjang (Mulunda & Wagacha, 2013) kerana ianya dapat dikesan hampir 100% oleh teknik steganalisis

seperti di dalam kajian yang dilakukan oleh Yang dan Cao (2010) dan Xiang, Sun, Luo dan Xia (2014).

Justeru itu, teknik penyembunyian berasaskan warna RGB telah diperkenalkan oleh beberapa penyelidik seperti (Al-Azzawi, 2018; Joshi, 2018; Malik, Sikka, & Verma, 2016; Malik et al., 2017; Singh & Diwakar, 2014; Singh, Diwakar, & Upadhyaya, 2014; Tang & Chen, 2013; Wang & Li, 2014). Teknik warna RGB menyembunyikan mesej rahsia dengan cara memformatkan aksara terpilih di dalam teks pelindung dengan warna RGB tertentu. Pelbagai julat warna RGB antara 0 hingga 255 digunakan untuk menyembunyikan bit mesej rahsia. Namun begitu, terdapat beberapa kelemahan utama telah dikenal pasti terhadap kajian berasaskan warna RGB, antaranya ialah perubahan ketara terhadap warna RGB aksara teks stego yang dihasilkan dapat dikenal pasti dengan mudah oleh sistem visual manusia sepertimana yang dinyatakan oleh Malik et al. (2016) dan Malik et al. (2017) serta menimbulkan kecurigaan.

Selain itu, beberapa kajian lepas mewakilkan aksara mesej rahsia yang berulang dan tunggal dengan satu nilai yang sama (Bhaya et al., 2013; Dulera, Jinwala, & Dasgupta, 2011; Fuad, 2014; Koley & Mandal, 2016, 2017; Kouser, 2016) sebelum dilakukan proses penyembunyian sama ada dalam bentuk ASCII atau binari. Perwakilan aksara mesej rahsia yang berulang ini, mengakibatkan pembentukan corak perwakilan yang sama dihasilkan dan seterusnya mendorong kepada steganalisis untuk mengekstrak mesej yang disembunyikan. Perwakilan terhadap aksara mesej rahsia berulang dengan pemetaan yang sama menyebabkan berlakunya kecurigaan terhadap teks stego yang dijana seperti yang dilakukan di dalam kajian perwakilan *emoticons* (Iranmanesh,

Wei, Dao-ming, & Arigbabu, 2015) dan aksara (Lee, Iranmanesh, & Quiroz, 2016) dalam Sistem Pesanan Ringkas (*SMS*). Oleh itu isu perwakilan mesej berulang adalah penting di dalam perwakilan aksara mesej rahsia (Agarwal, 2017) bagi memastikan aksara berulang diwakilkan dengan pelbagai nilai berbeza bagi mengelakkan berlakunya kecurigaan.

Penyelidikan steganografi berasaskan atribut warna RGB telah dijalankan oleh beberapa penyelidik dengan menggunakan teknik seperti perubahan warna aksara (Al-Asadi & Bhaya, 2016; Singh & Diwakar, 2014), warna aksara dan warna garis bawah (Wang & Li, 2014), warna ruang kosong, warna margin muka surat dan perenggan (Stojanov et al., 2014), gabungan aksara dan garis bawah (Tang & Chen, 2013) dan sebagainya dengan menggunakan pelbagai nilai warna RGB dengan julat antara (0,0,0) hingga (255,255,255). Setiap aksara mesej rahsia yang berulang diwakilkan dengan corak warna yang sama. Ini mewujudkan isu perubahan warna RGB yang sama terhadap teks stego yang dijana seperti kajian yang dilakukan oleh Al-Asadi & Bhaya (2016); Malik et al. (2016) dan Malik et al. (2017) selain daripada perwakilan nilai RGB yang maksimum (255,255,255).

Menurut Khairullah (2019) teknik warna yang digunakan pada e-mel untuk proses penyembunyian di dalam kajian yang dijalankan oleh Malik et al., (2017) dikenal pasti mempunyai kelemahan utama dari segi warna yang digunakan. Selain itu, isu lain yang dikenal pasti di dalam semua kajian di atas ialah pemilihan lokasi penyembunyian yang dilakukan secara jujukan di mana ia dapat memudahkan teknik steganalisis untuk memanipulasikan kedudukan aksara mesej rahsia. Kesemua faktor-faktor ini telah memberi ruang kepada penyelidik untuk membuat kajian terhadap

teknik warna RGB dengan memberi tumpuan kepada kapasiti penyembunyian, perwakilan aksara yang dinamik, dan lokasi penyembunyian rawak bagi meningkatkan prestasi teks stego yang dihasilkan.

Kesimpulannya, kajian ini memfokuskan kepada teknik penyembunyian berdasarkan warna RGB dengan memfokuskan kepada kapasiti penyembunyian, perwakilan aksara mesej rahsia yang dinamik dan lokasi penyembunyian rawak. Penyembunyian aksara mesej rahsia yang mempunyai corak yang sama merupakan salah satu cabaran yang ditimbulkan oleh penyelidik Satir dan Isik (2012b) serta Mandal, Koley, dan Dhar (2016) di dalam kajian mereka. Satir dan Isik (2012b) dan Hamdan dan Hamarsheh, (2017) mencadangkan penggunaan ciri rawak boleh digunakan untuk mengatasi masalah tersebut dan digunakan di dalam beberapa penyelidikan lepas (Elmahi, Wahbi, & Sayed, 2017). Seterusnya, kajian lepas juga menunjukkan kapasiti penyembunyian menggunakan teknik warna RGB meningkat sehingga 20% berbanding teknik-teknik lain.

Menurut Chaudhary & Dave (2016), penyembunyian menggunakan atribut warna RGB berpotensi tinggi untuk meningkatkan prestasi kapasiti penyembunyian (Kaur, Gupta, Sandhu, & Kaur, 2010) selain daripada atribut saiz teks, gaya tulisan dan sebagainya. Selain itu, teknik ini juga lebih selamat dan sukar dikesan oleh pihak ketiga (Kaur, Gupta, Sandhu, & Kaur, 2010) yang cuba untuk mengesan kehadiran mesej rahsia. Teks stego yang dijana akan dinilai prestasinya berdasarkan kepada kapasiti, keteguhan dan ketakbolehkelihatan yang merupakan topik panas di dalam domain steganografi teks seperti yang disarankan di dalam kajian yang dijalankan oleh Baawi et al. (2018).

1.4 Persoalan Kajian

i. Bagaimana cara untuk menentukan perwakilan aksara mesej rahsia yang berulang dan tunggal agar mempunyai nilai rawak yang pelbagai?

ii. Bagaimana aksara mesej rahsia diwakilkan dengan perwakilan warna RGB untuk melaksanakan proses penyembunyian?

iii. Bagaimana prestasi teks stego yang dihasilkan dinilai berbanding dengan teknik lain?

1.5 Objektif Kajian

Objektif utama kajian ini ialah untuk menyembunyikan mesej dalam teks menggunakan teknik warna RGB dan penempatan rawak. Objektif khusus ialah:

1. Untuk menentukan perwakilan aksara mesej rahsia yang berulang dan tunggal agar mempunyai nilai rawak yang pelbagai dengan menjana Jadual *Homophonic* sifer serta penggunaan nombor rawak.

2. Untuk menentukan aksara mesej rahsia diwakilkan dengan perwakilan tiga dimensi untuk dipetakan kepada warna RGB.

3. Untuk menilai kapasiti, keteguhan dan ketakbolehkelihatan teks stego yang dihasilkan berbanding dengan teknik-teknik yang menggunakan warna RGB.

1.6 Kepentingan Kajian

Kajian ini dapat memberi idea di dalam bidang penyembunyian maklumat khususnya di dalam bidang steganografi teks. Beberapa signifikan penting dikenal pasti di dalam kajian yang dicadangkan. Pertama, kebanyakan teknik steganografi pada masa kini memfokuskan kepada imej (57%) dan audio (33%) tetapi kurang kajian di dalam teks (6%) dan audio (4%) steganografi seperti yang dinyatakan di dalam kajian yang dilakukan oleh Osman et al., (2016). Tambahan pula, komunikasi melalui Internet yang paling popular ialah menggunakan medium teks kerana ia menggunakan kurang memori dan tidak kompleks (Pawar & Kakde, 2014) berbanding dengan medium lain. Oleh itu medium teks amat sesuai digunakan sebagai teks pelindung di dalam kajian ini di mana kadar pemindahan data dan saiz fail adalah lebih kecil berbanding dengan medium lain.

Kedua, kajian ini adalah untuk menghasilkan teks stego yang disembunyikan dengan mesej rahsia bertujuan untuk memastikan keteguhan dan ketakbolehkelihatan dapat dikekalkan di samping memastikan kapasiti penyembunyian mesej rahsia dapat dipertingkatkan (El Rahman, 2019). Ketiga-tiga aspek ini merupakan ukuran utama yang digunakan untuk mengukur prestasi teks stego yang dijana. Menurut Khadim, Khan, Ahmad, dan Khan, (2015), keteguhan dan ketakbolehkelihatan merupakan dua aspek penting dalam teks steganografi bagi memastikan data boleh bertahan dengan selamat selepas berlakunya proses penyembunyian di mana data asal tidak berubah. Steganografi berbeza dengan kriptografi dan tera air dari aspek ketakbolehkelihatan di mana ia memastikan bagaimana untuk menyembunyikan maklumat tanpa dapat dilihat dengan nyata (*unnoticeably*) bagi mengelakkan berlakunya kecurigaan terhadap teks stego yang dihasilkan (Ahvanooey, Li, Shim, et al., 2018). Menurut

Ramakrishnan, Thandra, dan Srinivasula (2017), perubahan yang berlaku terhadap teks stego perlu memastikan agar ia mencapai tahap kerahsiaan yang tinggi dan tidak menarik perhatian terhadap penampilan dari aspek visualnya (ketakbolehkelihatan). Ini disebabkan kekerapan aksara teks stego boleh dianalisis berdasarkan kepada teks pelindung dan bukan berdasarkan frekuensi mesej rahsia.

Ketiga, kajian lepas menunjukkan bahawa teknik steganalisis sedia ada mampu mengesan hampir 100% mesej rahsia terutama menggunakan teknik ruang kosong selain daripada nilai ketakbolehkelihatan kurang dari 1 yang menunjukkan terdapat kecurigaan terhadap teks stego. Oleh itu teknik penyembunyian berdasarkan lokasi abjad mampu untuk meningkatkan prestasi teks stego yang dijana dan mengelakkan berlakunya kecurigaan.

Keempat, penyembunyian mesej rahsia secara rawak seperti yang digunakan di dalam kajian terhadap steganografi imej oleh (Sumathi et al., 2013) boleh diadaptasikan di dalam kajian ini kerana sifat rawak (*randomize*) dan penyembunyian berdasarkan kepada kedudukan abjad (Bhattacharyya, Indu, et al., 2011) boleh meningkatkan tahap kerahsiaan dan prestasi serta kapasiti penyembunyian mesej. Kebolehan menyembunyi mesej rahsia secara rawak di dalam pelbagai lokasi di dalam teks pelindung boleh meningkatkan tahap kesukaran steganalisis untuk mengekstrak mesej rahsia.

Akhir sekali, di dalam kajian ini penjanaan jadual *homophonic* berdasarkan kekerapan frekuensi abjad dapat mewakilkan aksara mesej rahsia dalam bentuk yang dinamik terutama bagi aksara mesej rahsia yang berulang.

1.7 Skop Kajian

Imej, teks, audio, video dan protokol merupakan medium yang digunakan di dalam steganografi. Walau bagaimanapun, kajian ini hanya memberi tumpuan kepada medium teks serta kaedah berasaskan format untuk menyembunyikan mesej rahsia. Kajian lepas lebih menumpukan kepada aksara (A-Z) untuk proses penyembunyian berbanding aksara lain seperti nombor, simbol dan sebagainya. Antara penyelidik yang menjalankan kajian berasaskan aksara (A-Z) ialah Majumder dan Changder, (2013); Odeh, Elleithy, Faezipour, dan Abdelfattah, (2015); Samer et al., (2015); Malik et al., (2017); Ramakrishnan, Thandra, dan Srinivasula, (2017); Naqvi, Abbasi, Hussain, Khan, dan Ahmad, (2018); Win dan Oo, (2018); Xiao, Zhang, dan Zheng, (2018); Mandal, Chatterjee, dan Chakraborty, (2019). Selain itu, kajian yang dijalankan oleh Grigas dan Juškevičienė, (2018) menunjukkan peratusan taburan aksara di dalam sesuatu dokumen kebanyakannya adalah aksara ruang kosong dan A-Z. Oleh itu, kajian ini hanya mengehadkan kepada mesej rahsia yang terdiri daripada aksara A hingga Z sahaja tanpa melibatkan simbol-simbol lain tetapi sebaliknya ia tidak terhad untuk teks pelindung.

Data kajian yang digunakan adalah berdasarkan kepada data piawaian dalam bidang penyembunyian maklumat yang diperoleh daripada koleksi Reuters-21578 (Aghdam, Ghasem-Aghaee, & Basiri, 2009; Maiti & Samanta, 2010). Set data ini merupakan koleksi ujian pengkategorian teks yang kumpulkan oleh Carnegie Group, Inc. and Reuters, Ltd. dan digunakan oleh para penyelidik sehingga kini.

1.8 Organisasi Tesis

Kajian yang dijalankan ini telah melalui beberapa proses yang disusun lengkap di dalam penulisan ini yang merangkumi tujuh bab. Bab 1 menghuraikan latar belakang kajian yang meliputi perkembangan keselamatan data, sejarah penyembunyian maklumat, motivasi kajian, pernyataan masalah, persoalan kajian, objektif kajian, skop kajian serta kepentingan kajian di dalam bidang penyembunyian maklumat.

Bab 2 menghuraikan pengenalan steganografi, proses-proses steganografi, ukuran-ukuran prestasi yang digunakan serta menganalisis ulasan karya berkaitan dengan penyelidikan lepas yang telah dijalankan. Selain itu penjanaan Jadual *Homophonic*, pembentukan formula SQRT, teknik warna RGB dan penjanaan nombor rawak yang digunakan turut dibincangkan di dalam bab ini.

Bab 3 menghuraikan metodologi kajian yang dijalankan yang meliputi empat fasa iaitu Fasa Kajian Awalan, Reka bentuk dan Pembangunan, Implementasi serta Penilaian Prestasi. Selain itu, pemilihan sampel data teks tersembunyi dan teks pelindung turut dibincangkan di dalam bab ini.

Bab 4 menghuraikan proses reka bentuk teknik penyembunyian yang meliputi perwakilan aksara dengan nilai yang dinamik menggunakan teorem SQRT. Di samping itu proses pemetaan warna RGB dan mengenal pasti lokasi rawak untuk tujuan penyembunyian diterangkan di dalam bab ini. Akhir sekali, algoritma untuk proses penyembunyian dan pengekstrakan turut disertakan di penghujung bab.

Bab 5 Bab ini membincangkan hasil dan analisis terhadap teks stego yang telah dijana. Selain itu, analisis terhadap fail mesej rahsia dan fail teks pelindung turut diterangkan

di dalam bab ini. Bab ini turut membincangkan prestasi teks stego yang dijana dinilai berdasarkan kepada ukuran kapasiti, keteguhan dan ketakbolehkelihatan yang telah ditetapkan. Hasil dapatan kajian dibandingkan dengan kajian lepas berdasarkan kepada ukuran-ukuran tersebut.

Akhir sekali, Bab 6 membincangkan pencapaian terhadap objektif kajian yang telah ditetapkan. Selain itu sumbangan kajian, limitasi serta kajian masa hadapan di dalam bidang steganografi turut dibincangkan di akhir bab ini.

1.9 Ringkasan

Steganografi merupakan sub disiplin dalam bidang penyembunyian maklumat yang bertujuan untuk melindungi mesej rahsia daripada dikesan oleh pihak ketiga. Teks, imej, audio dan video merupakan medium yang digunakan dalam steganografi untuk menyembunyikan teks. Dalam kajian ini, mesej rahsia diwakilkan dalam bentuk dinamik, di mana kedua-dua aksara mesej rahsia tunggal dan berulang diwakilkan dengan pelbagai nilai menggunakan teknik yang diperkenalkan. Selain itu, teknik yang dicadangkan juga dapat menyembunyikan mesej rahsia pada lokasi rawak dengan memformatkan aksara pada lokasi terpilih dengan warna RGB yang telah ditentukan. Teknik ini dapat meningkatkan kapasiti penyembunyian mesej di samping dapat menghindari sistem visual manusia daripada mengesan kewujudan mesej rahsia di dalam teks stego yang dijana. Selain itu, tahap keteguhan teks stego yang dijana diuji untuk memastikan prestasi teks stego yang dihasilkan adalah teguh.

BAB DUA
ULASAN KARYA

Bab ini menerangkan kajian terdahulu yang berkaitan dengan bidang penyembunyian maklumat yang dilakukan oleh penyelidik-penyelidik untuk menyokong kajian yang dicadangkan. Sejarah, kaedah-kaedah dan teknik-teknik steganografi yang digunakan dipetik dan dikritik bagi menyokong kajian yang dijalankan. Akhir sekali, bab ini menjelaskan pendekatan yang sesuai digunakan di dalam kajian yang telah dijalankan.

2.1 Steganografi Teks

Steganografi merupakan seni menyembunyikan maklumat dengan matlamat untuk menyembunyikan data daripada pihak yang tidak bertanggungjawab supaya tiada yang mengesyaki kewujudan mesej rahsia di dalam sesuatu mesej yang dihantar. Perkataan steganografi berasal daripada bahasa Greek (*steganos* + *graphy*) yang mana *steganos* bermaksud "dilindungi atau tersembunyi" dan *graphy* bermaksud "penulisan". Penggunaan pertama steganografi telah mula diperkenalkan pada tahun 440 sebelum Masihi dalam sejarah Herotodus di mana pembawa mesej akan dicukur rambutnya dan kemudian mesej rahsia akan ditulis pada kulit kepala seperti tatu (Herodotus, 1992). Pertumbuhan rambut akan menutupi mesej tersebut dan pembawa mesej akan dikirimkan ke penerima. Seterusnya, semasa perang dunia kedua, steganografi digunakan untuk penghantaran mesej rahsia dengan pelbagai kaedah seperti menebuk lubang menggunakan pin di atas keratan akhbar atau artikel-artikel, menggunakan jus, lilin atau dakwat yang tidak kelihatan apabila ditulis dan dipanaskan untuk pengekstrakan semula mesej tersebut. Selain itu, antara kaedah lain yang digunakan ialah menggunakan papan sebagai medium pelindung, di mana mesej

rahsia ditulis di atas papan tersebut dan kemudian dilitupi dengan lapisan penggilap supaya papan pelindung kelihatan kosong tanpa sebarang kecurigaan (Ingemar, Matthew, Jeffrey, Fridrich, & Kalker, 2008). Steganografi pernah digunakan oleh Bangsa Rumawi dengan menulis mesej menggunakan tinta tak nampak (*invisible ink*) yang dibuat daripada campuran jus buah-buahan, susu dan cuka. Teknik tinta tak nampak ini terbukti telah digunakan semasa perang dunia kedua untuk penyembunyian mesej (Johnson & Jajodia, 1998). Mesej yang ditulis menggunakan teknik ini tidak kelihatan pada kertas kecuali perlu dipanaskan untuk mengekstrak semula mesej yang disembunyikan. Rajah 2.1 menunjukkan sebahagian daripada kaedah steganografi yang digunakan pada masa lalu.

Rajah 2.1 Kaedah Steganografi Pada Masa Lalu

Rajah 2.1 menunjukkan sebahagian daripada teknik tradisional yang digunakan untuk melakukan proses penghantaran data secara tersembunyi, antaranya ialah dengan mencukur rambut, menebuk lubang pada kertas dan penggunaan tinta tak nampak.

Perkembangan teknologi menyebabkan kaedah-kaedah moden telah digunakan di dalam steganografi seperti penggunaan teknik Microdot dan kod Morse. Pada tahun 2004, pengilang-pengilang mesin cetak seperti Brother, Dell, Canon dan HP telah menggunakan kaedah steganografi untuk menyembunyikan maklumat penting seperti nombor siri pencetak, kod pengilang dan tarikh cetakan. Maklumat tersebut disembunyikan di dalam dokumen bercetak berkod corak yang bertujuan untuk mengelakkan daripada dipalsukan atau tuntutan hak cipta (Bloom et al., 1999) oleh pihak yang tidak bertanggungjawab. Penyelidikan terkini yang dilakukan oleh Xao et al., (2018), menyembunyikan maklumat seperti tarikh, masa, lokasi GPS disembunyikan di dalam fotograf digital.

Di era moden ini, kajian steganografi telah berkembang setiap tahun dengan pelbagai teknik telah diperkenalkan dengan menyembunyikan mesej di dalam dokumen berbentuk teks, imej audio video dan sebagainya. Sebagai contoh, keratan teks di bawah menunjukkan mesej rahsia "Come to our place at midnight" telah disembunyikan di dalam satu dokumen yang dihantar kepada penerima tanpa disedari oleh pembaca. Teknik ini menyembunyikan mesej pada perkataan pertama setiap baris agar ia tidak menjejaskan dokumen asal.

> Come see how the swallows fly
> to the south island helped by
> our fine weather. They locate the
> place in the immense ocean by navigating
> at night, following the clouds which form around
> midnight above the shore.

Sumber : Salomon, (2003)

Kajian saintifik mengenai steganografi bermula pada tahun 1983 apabila Simmons, (1984) memperkenalkan satu analogi komunikasi di penjara di antara *Alice*, *Bob* dan *Warden*. *Alice* dan *Bob* tinggal di dalam sel yang berbeza dan cuba membuat

perancangan untuk melarikan diri dari penjara. Satu-satu cara untuk mereka berkomunikasi hanyalah melalui *Warden* yang memantau secara terperinci aktiviti mereka. Maka di sinilah bermulanya idea steganografi, bagaimana *Alice* dan *Bob* membuat perancangan untuk melepaskan diri melalui berkomunikasi secara biasa dengan menyembunyikan mesej rahsia di dalam komunikasi mereka bagi mengelakkan daripada dikesan oleh *Warden*.

Steganografi merupakan salah satu cabang sistem keselamatan dan ia berbeza dengan kriptografi di mana kriptografi menukarkan mesej rahsia ke bentuk yang tidak boleh dibaca. Namun begitu, mesej yang tidak boleh dibaca ini lebih mudah menarik perhatian pihak yang berkeinginan berbanding teknik steganografi yang bersifat menyembunyikan maklumat di dalam dokumen. Jadual 2.1 menunjukkan perbandingan antara kriptografi dan steganografi yang dipetik dari Zaidan, Zaidan, Al-Frajat dan Jalab, (2010).

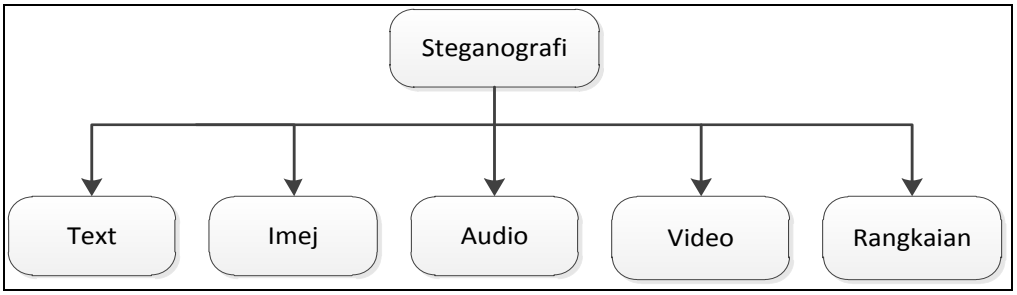
Jadual 2.1

Perbandingan di antara Kriptografi dan Steganografi (Zaidan, Zaidan, Al-Frajat dan Jalab, 2010)

Kriptografi	**Steganografi**
Huruf yang disulitkan boleh dilihat oleh sesiapa sahaja dan kriptografi membuatkan mesej yang dijana tidak difahami.	Steganografi menyembunyikan mesej di dalam medium lain supaya tiada yang menyedari kewujudan mesej tersebut.
Output bagi kriptografi ialah *cipher text*.	Output bagi steganografi ialah stego medium.
Matlamat utama kriptografi ialah mencegah penceroboh daripada memperoleh sebarang maklumat mengenai teks sifer.	Matlamat steganografi ialah untuk mencegah pemerhati daripada mengenal pasti kewujudan mesej rahsia.

Sesiapa sahaja mempunyai kebolehan untuk mengesan dan mengubah mesej yang disulitkan.	Mesej rahsia tidak kelihatan oleh sesiapa.
Steganografi tidak boleh diguna untuk mengadaptasikan sistem kriptografi.	Steganografi boleh diguna bersama-sama dengan kriptografi dengan menyembunyikan mesej yang disulitkan bagi tujuan keteguhan

Menurut Singh et al. (2014), kriptografi menjana kod sifer yang mudah diterjemahkan oleh pihak ketiga berbanding dengan steganografi yang menyembunyikan mesej. Antara salah satu kelebihan steganografi berbanding kriptografi ialah mesej yang dihantar adalah lebih selamat dan tidak menarik perhatian pihak ketiga serta sukar untuk penceroboh mengesan kehadiran mesej rahsia (Kant, Nath, & Chaudhary, 2008;Conklin & William, 2011). Seterusnya, Kumar dan Pooja (2010) menyatakan bahawa steganografi dan kriptografi boleh digabungkan bagi menghasilkan pencegahan yang lebih baik terhadap mesej rahsia. Oleh itu, dengan menggabungkan pendekatan steganografi dan kriptografi, mesej rahsia kadangkala boleh disulitkan terlebih dahulu sebelum melalui proses steganografi bagi memperteguh teks stego agar kekal selamat walaupun selepas melalui proses pemampatan dan penyahmampatan (Singh & Singh, 2013).



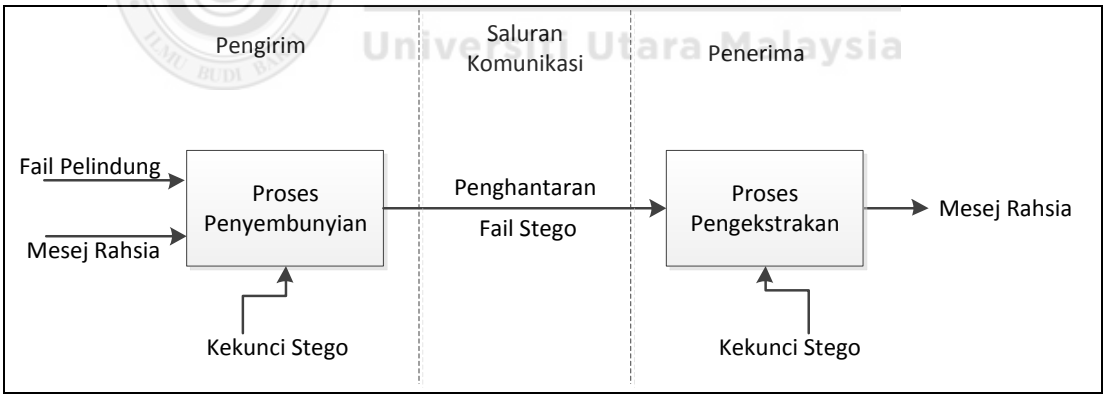
Rajah 2.2 Klasifikasi Medium Teks Pelindung (Alanazi, Zaidan, Zaidan, Jalab, & AL-Ani, 2010)

Medium pelindung yang digunakan untuk menyembunyikan mesej rahsia boleh diklasifikasikan kepada teks, imej, audio, video dan rangkaian atau protokol (Alanazi et al., 2010) seperti yang ditunjukkan di dalam Rajah 2.2. Walaupun medium-medium tersebut amat popular (Krishnan et al., 2017; Hmood, Jalab, Kasirun, Zaidan, & Zaidan, 2010) digunakan oleh penyelidik, namun medium teks kurang diberi tumpuan oleh penyelidik disebabkan oleh ruang lewah (*redundant space*) untuk menyembunyikan mesej rahsia adalah terhad (Krishnan et al., 2017; Sloan & Hernandez-castro, 2015; Bhattacharyya, Banerjee, & Sanyal, 2011). Walau bagaimanapun, medium teks tetap menjadi pilihan penyelidik kerana ia menggunakan kurang memori (Al-Azzawi, 2018; Lwin & Phyo, 2014; Mulunda & Wagacha, 2013; Shirali-shahreza, 2006; Shivani et al., 2015), penghantaran pantas (K. A. Kumar et al., 2014) , serta sukar dikesan oleh steganalisis (Bhattacharyya, Banerjee, et al., 2011) berbanding medium-medium lain. Kenyataan ini disokong oleh artikel yang ditulis oleh Zielińska, Mazurczyk dan Szczypiorski (2014) yang menyatakan bahawa dua ciri medium pelindung yang terbaik ialah; pertama, medium pelindung mestilah popular dan; kedua, sebarang perubahan terhadap objek stego mestilah tidak boleh dikesan atau dicurigai oleh pihak ketiga.

2.2 Proses Steganografi

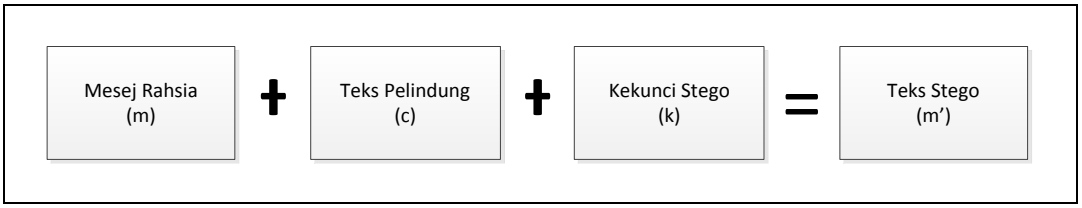
Steganografi merupakan proses yang melibatkan empat komponen utama iaitu mesej rahsia, medium pelindung, kekunci stego dan medium stego (output yang dihasilkan yang bergantung kepada jenis medium yang digunakan). Secara umumnya, terdapat pelbagai terma yang digunakan oleh penyelidik untuk mewakili mesej rahsia, teks pelindung, kekunci stego dan medium stego. Antaranya seperti mesej asli atau objek rahsia digunakan untuk mewakili mesej rahsia; teks pelindung atau medium pelindung

digunakan untuk mewakili teks pelindung; kunci stego atau fungsi stego digunakan untuk mewakili kekunci stego dan akhir sekali teks stego atau objek stego yang digunakan untuk mewakili mesej rahsia yang disembunyikan. Justeru itu, kajian ini menggunakan terma mesej rahsia (*secret message*), teks pelindung (*cover text*), kekunci stego (*stego key*) dan teks stego (*stego text*) di dalam perbincangan selanjutnya dengan memfokuskan kepada medium teks sebagai teks pelindung. Rajah 2.3 menunjukkan mekanisme steganografi teks yang melibatkan proses penyembunyian, penghantaran mesej dan proses pengekstrakan (Por & Delina, 2008). Proses penyembunyian melibatkan input fail pelindung, mesej rahsia dan kekunci stego bagi menjana teks stego yang akan dihantar kepada penerima melalui saluran komunikasi. Teks stego yang diterima akan dilakukan proses pengekstrakan menggunakan kekunci stego bagi mendapatkan semula mesej rahsia yang disembunyikan.



Rajah 2.3 Mekanisme Steganografi Teks (Por & Delina, 2008)

Proses penyembunyian steganografi teks secara am (Kumar dan Pooja, 2010) dapat ditunjukkan di dalam Rajah 2.4.



Rajah 2.4 Proses Penyembunyian Steganografi Teks Secara Am (Kumar & 3.2, 2010)

Berdasarkan kepada Rajah 2.4, hubungan bagi proses steganografi secara am boleh ditulis sebagai:

(2.1)

$$m'' = \{m,c,k\}$$

di mana,

m : mesej rahsia – mesej yang hendak disembunyikan

c : teks pelindung – medium yang digunakan untuk menyembunyikan mesej rahsia.

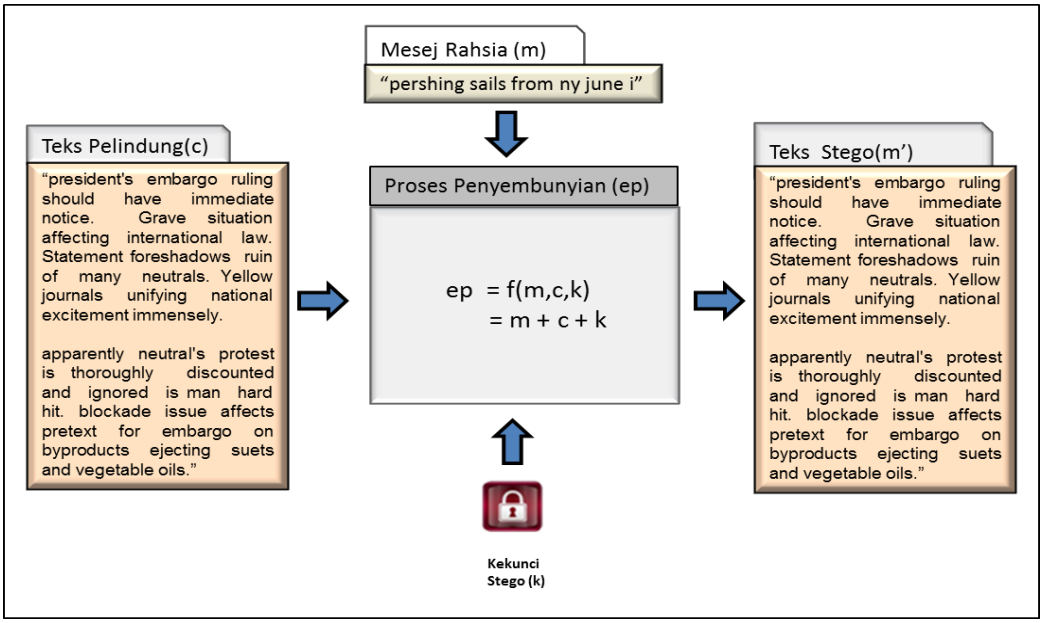
k : kekunci stego - digunakan untuk menyembunyikan dan mengekstrak mesej rahsia.

m'' : teks stego – medium yang mengandungi mesej rahsia.

medium : {teks, imej, audio, video, protocol}

Teks stego, *m''* yang dihasilkan akan diekstrakkan menggunakan kekunci stego untuk mendapatkan semula mesej rahsia, *m* yang disembunyikan seperti yang ditunjukkan di dalam Rajah 2.4.

Empat jenis medium pelindung gunakan oleh penyelidik di dalam steganografi iaitu teks, imej, audio, video dan protokol (Zielinska, Mazurczyk, & Szczypiorski, 2012). Rajah 2.5 di bawah menunjukkan contoh proses steganografi menggunakan teks sebagai medium pelindung yang dipetik daripada Osman, Din dan Idrus (2015).



Rajah 2.5 Proses Menyembunyikan Mesej

Rajah 2.5 menunjukkan mesej rahsia (*m*) disembunyikan ke dalam teks pelindung (*c*) menggunakan fungsi kekunci stego (*k*). Berdasarkan kepada rajah di atas, mesej rahsia "*pershing sails from ny june i*" disembunyikan ke dalam teks pelindung, *c* pada huruf pertama setiap perkataan bagi menghasilkan teks stego.

2.3 Kriteria Penilaian Prestasi di dalam Steganografi

Satir & Isik, (2012b) menyatakan bahawa tiga kriteria asas digunakan untuk mengukur prestasi teks stego ialah kapasiti (*capacity*), keteguhan (*robustness*) dan ketakbolehkelihatan (*imperceptibility*). Antara penyelidik lain yang menggunakan kriteria tersebut di dalam kajian mereka ialah Amirthrajan dan Rayappan, (2013), Agarwal, (2013), Gupta, Gupta, dan Singhal, (2014), Kumar et al, (2014) serta Joseph dan Vishnukumar, (2015). Menurut Amirthrajan dan Rayappan (2013), tiga faktor utama yang perlu dipertimbangkan di dalam steganografi ialah pertama, mesej rahsia mestilah tidak boleh dilihat oleh sesiapa sahaja termasuk penerima; kedua, ia perlulah dapat menghalang sesiapa sahaja yang cuba untuk mendapatkan semula mesej yang

disembunyikan dan akhir sekali jumlah aksara mesej rahsia yang boleh disembunyikan di dalam teks pelindung. Kebanyakan penyelidik terkini masih menggunakan ukuran kapasiti, keteguhan dan ketakbolehkesanan atau ketakbolehkelihatan untuk menilai prestasi steganografi teks, antaranya Akotoye, (2017); Aman, Khan, Ahmad, dan Kouser, (2017); Ishtiaq, Khan, dan Samim, (2017); Kothari, Thakkar dan Khara (2017); Al-Azzawi,(2018); Ahvanooey, Li, Hou, Mazraeh, dan Zhang, (2018); Naqvi et al., (2018) serta Ahvanooey et al., (2019).

Kapasiti merupakan ukuran utama untuk mengukur prestasi teks stego yang digunakan oleh kebanyakan penyelidik. Ia merujuk kepada jumlah bit data yang boleh disembunyikan di dalam teks pelindung (Ahvanooey, Li, Shim, et al., 2018; Htet & Phyo, 2016; Kingslin & Kavitha, 2015) yang mana ia merupakan hubungan di antara saiz mesej rahsia dengan saiz teks pelindung. Formula untuk pengiraan kapasiti ditunjukkan dalam persamaan 2.2.

$$Kapasiti\ Penyembunyian\ (c) = \frac{jumlah\ bit\ mesej\ tersembunyi}{jumlah\ bit\ teks\ pelindung} \quad (2.2)$$

Kajian yang dilakukan oleh Saniei dan Faez (2013) menggunakan pendekatan pemampatan teks aritmetik (*arithmetic text compression*) dan penskalaan menegaskan bahawa teknik yang digunakan oleh mereka mampu menyimpan kapasiti mesej rahsia dengan kadar 5.95%. Walau bagaimanapun, teknik ini menghasilkan ralat bagi mesej rahsia yang panjang. Sementara itu kajian yang dilakukan oleh Satir dan Isik (2012a), menggunakan teknik LZW berjaya menyembunyikan sebanyak 5.31% dan 7.042% mesej rahsia di dalam teks pelindung masing-masing bersaiz 100 dan 300 aksara. Jadual 2.2 menunjukkan kapasiti penyembunyian mesej rahsia menggunakan beberapa

teknik yang dipetik dari kajian yang dijalankan oleh Saniei dan Faez, (2013) dan beberapa kajian terkini.

Jadual 2.2

Kapasiti Penyembunyian Mesej Rahsia Berdasarkan Teknik Yang Digunakan

Teknik	Kapasiti (%)	Mesej Rahsia	Teks Pelindung
(Satir & Isik, 2012a)	6.92	200	850
(Satir & Isik, 2012b)	7.017	20,200	450,850
(Por et al.,2012)	20.0	1024	5000
(Saniei & Faez, 2013)	5.95	100	Tidak dinyatakan
(Agarwal, 2013)	8.92	508	Tidak dinyatakan
(Mahato, Yadav, & Khan, 2014)	9.1	Tidak dinyatakan	Tidak dinyatakan
(Iyer & Lakhtaria, 2016)	10.50	Tidak dinyatakan	Tidak dinyatakan
(Kumar, Malik, Singh, & Chand, 2016)	7.21	166	723
(Malik et al., 2016)	13.43	166	723
(Malik et al., 2017)	18.34	166	723
(Kouser et al., 2017)	23.25	800	2640
(Ramakrishnan et al., 2017)	28.11	500	1779
(Fateh & Rezvani, 2018)	10.6	200	847
(Al-Azzawi A. F., 2018)	25.5	34	202
(Naharuddin, Wibawa, & Sumpeno, 2018)	14.2	320	2241
(Khairullah, 2019) – Teks Bengali	7.88	315	4000
(El Rahman & Nourah, 2019)	12.27	104	847
(Baawi, Mokhtar, & Sulaiman, 2019)	12.02	198	847

Berdasarkan kepada Jadual 2.2, walaupun teknik UniSpace yang diperkenalkan oleh Por et al.,(2012) dapat menyimpan kapasiti mesej rahsia sebanyak 20%, namun teknik steganalisis menggunakan analisis statistik boleh mengesan kewujudan mesej rahsia seperti yang dinyatakan di dalam kajian tersebut.

Menyembunyikan mesej rahsia menggunakan abjad adalah lebih baik berbanding menggunakan nilai binari kerana ia boleh meningkatkan kapasiti penyembunyian seperti kajian yang dilakukan oleh Agarwal (2013b). Kajian beliau mendapati bahawa menyembunyikan mesej menggunakan satu aksara menghasilkan kapasiti penyembunyian yang tinggi berbanding menggunakan nilai binari. Kajian beliau

berjaya menyembunyikan mesej rahsia dengan purata kapasiti penyembunyian ialah 8.162%. Sementara itu, kajian yang dilakukan oleh Al-Azzawi, (2018) menggunakan teknik penyembunyian penandaan perkataan (*word tagging*) dan pengekodan warna RGB telah berjaya meningkatkan kapasiti penyembunyian sehingga 25.5%.

Ukuran kedua yang digunakan untuk mengukur prestasi steganografi ialah keteguhan. Keteguhan merupakan kesukaran untuk mengekstrak mesej rahsia daripada stego objek (Böhme, 2010) atau keupayaan untuk menahan serangan (Ahvanooey, Li, Shim, et al., 2018; Joseph & Vishnukumar, 2015), manakala menurut Mihaela (2011), keteguhan adalah merujuk kepada jumlah perubahan teks pelindung boleh bertahan walaupun mesej rahsia dimusnahkan oleh alatan-alatan tertentu. Selain daripada itu, menurut Saniei dan Faez (2013) keteguhan adalah merujuk kepada kemungkinan menahan daripada perubahan atau kemusnahan mesej rahsia di dalam saluran komunikasi. Secara umumnya keteguhan di dalam steganografi teks bermaksud memastikan agar mesej yang disembunyikan di dalam teks pelindung tidak terjejas walaupun berlaku perubahan ke atas teks stego seperti penghapusan ruang kosong, proses pemampatan, penyahmampatan , OCR (*Optical Character Recognition*) dan sebagainya. Teknik OCR pernah digunakan di dalam beberapa kajian seperti (Agarwal, 2013; Ali & Saad, 2013; Malalla & Shareef, 2017) untuk mengukur keteguhan teks stego yang dihasilkan.

Menurut Kumar et al. (2014), keteguhan adalah merujuk kepada kebolehan data yang disembunyikan di dalam teks pelindung untuk bertahan daripada pelbagai perubahan yang berlaku hasil daripada proses pemampatan dan penyahmampatan. Pemampatan (*compression*) merupakan salah satu kaedah yang digunakan untuk mengukur prestasi keteguhan seperti yang dinyatakan oleh Tiwari dan Sahoo (2011) dengan

membandingkan mesej rahsia sebelum dan selepas pemampatan (Kingslin & Kavitha, 2015). Pemampatan boleh dibahagikan kepada dua teknik iaitu pemampatan *lossless* dan pemampatan *lossy* di mana kedua-duanya digunakan untuk mengurangkan saiz fail manakala penyahmampatan bertujuan untuk mendapatkan semula isi kandungan asal fail yang dimampatkan.

Pemampatan *lossless* akan memastikan setiap bit yang dimampatkan tetap kekal tanpa kehilangan bit selepas proses penyahmampatan. Teknik *lossless* biasanya digunakan untuk pemampatan teks, aturcara (*program*) (Kavitha, 2016) atau data kewangan di mana kehilangan data merupakan sesuatu yang amat sensitif. Sebaliknya, pemampatan *lossy* akan mengurangkan saiz fail dengan menghapuskan beberapa maklumat tertentu (Kodituwakku dan Amarasinghe, 2014) terutamanya maklumat berulang selain maklumat yang dinyahmampat masih boleh diterima walaupun berlaku kehilangan data. Data yang dihapuskan menggunakan teknik ini tidak dapat diperoleh kembali semasa proses penyahmampatan dan pemprosesan teks data sensitif tidak sesuai menggunakan teknik ini (Sidhu & Garg, 2014) disebabkan terdapat perbezaan dengan fail yang asal. Kebiasaannya teknik *lossy* digunakan untuk pemampatan video dan audio di mana pengguna masih boleh bertolak ansur terhadap kehilangan bit selepas berlakunya proses penyahmampatan. Tan, Tan, dan Guo, (2013) menggunakan teknik pemampatan *lossless* untuk memampatkan mesej rahsia sebelum mesej tersebut disembunyikan di dalam imej. Namun begitu penyelidik tidak menggalakkan teknik pemampatan *lossy* digunakan untuk teks kerana ia boleh menyebabkan berlakunya kehilangan data semasa proses penyembunyian.

Keteguhan fail teks stego boleh diterima apabila mesej yang disembunyikan dapat diekstrakkan semula tanpa kehilangan data sebelum dan selepas pemampatan.

Algoritma pemampatan dan penyahmampatan seperti *Lempel-Ziv-Welch* (LZW), *FLATE, CCITT*, *WinRAR*, *WinZip* dan *RAR* boleh digunakan untuk tujuan menguji keteguhan. Keteguhan algoritma teks stego boleh diterima apabila mesej rahsia boleh diekstrakkan sepenuhnya tanpa kehilangan data selepas melalui proses pemampatan dan penyahmampatan. Teknik pemampatan dan penyahmampatan telah digunakan di dalam beberapa kajian lepas antaranya oleh Lee dan Tsai (2010) yang menggunakan fail *pdf* dan menghasilkan saiz fail stego yang sama sebelum dan selepas proses penyahmampatan serta dapat mengekstrak semula 100% mesej rahsia selepas proses penyahmampatan. ZIP format merupakan aplikasi pemampatan *lossless* yang popular (Rani dan Singh, 2016; Sidhu dan Garg, 2014) digunakan untuk proses pemampatan data berbentuk teks. Selain itu, pemampatan lossless digunakan apabila kesamaan antara data asal dan data selepas penyahmampatan merupakan satu keperluan (Sidhu & Garg, 2014). Oleh itu, aplikasi WinZip digunakan sebagai alat pemampatan di dalam kajian ini untuk proses pemampatan fail teks stego yang dijana bagi tujuan mengukur prestasi keteguhan.

Ukuran ketiga untuk mengukur prestasi teks stego ialah tidak kelihatan (*invisible*) atau dikenali dengan nama lain seperti ketakbolehkelihatan (*imperceptibility*), kebolehkesanan (*detectability*) atau ketelusan (*transperancy*) (Ahvanooey et al., 2019). Menurut kamus oxford, ketakbolehkelihatan bermaksud tidak dapat diperhatikan atau dirasai disebabkan oleh perubahan yang kecil. Proses penyembunyian mesej rahsia boleh menyebabkan isi kandungan teks pelindung berubah. Kajian yang dijalankan oleh Joseph dan Vishnukumar (2015) menyarankan agar perubahan teks pelindung mestilah tidak dapat dikesan dan dapat mengelakkan kecurigaan sistem visual manusia (Ahvanooey, Li, Hou, et al., 2018; Ahvanooey et

al., 2019) ketika melakukan proses penyembunyian. Selain itu, menurut Kumar, Pabboju, Megha dan Desai (2014), kaedah steganografi yang baik seharusnya tidak mengubah ciri-ciri aksara teks pelindung secara nyata agar tidak berlaku sebarang kecurigaan. Oleh itu ketakbolehkelihatan bertujuan untuk memastikan agar data yang disembunyikan tidak mengubah kandungan teks pelindung secara jelas untuk mengelakkannya daripada dicurigai.

Menurut Khan (2015), isu ketakbolehkelihatan berlaku apabila terdapat perubahan terhadap corak abjad yang boleh menyebabkan teknik pengesanan statistik dapat membezakan antara teks pelindung dan teks stego. Seterusnya, menurut Baawi, Mokhtar, dan Sulaiman (2017) ciri-ciri teks pelindung yang telah diubahsuai, di manipulasi, atau dimodifikasi semasa proses penyembunyian mestilah kekal tidak kelihatan agar tidak berlaku kecurigaan terhadap pengguna yang tidak dibenarkan (*unauthorized user*). Kapasiti penyembunyian yang tinggi kurang berkesan jika ia menyebabkan gangguan terhadap teks pelindung. Menurut Rahman, Khalil, Yi, dan Dong, (2017) ketakbolehkelihatan merupakan salah satu ukuran di dalam steganografi yang mengukur tahap persepsi kewujudan mesej tersembunyi di dalam sesuatu teks pelindung.

Jaro-Wrinkler merupakan skala yang digunakan untuk mengukur ketakbolehkelihatan sesuatu teks stego dengan membandingkan kesamaan di antara dua renteten. Skala Jaro Wrinkler sesuai digunakan untuk mengukur kesamaan antara teks pelindung dan teks stego kerana ia melibatkan perbandingan rentetan (Agarwal, 2013) selain juga digunakan di dalam bidang pengesanan kebocoran maklumat untuk mengesan pertindihan. Semakin tinggi nilai skor (menghampiri 1), maka semakin hampir

kesamaan dua rentetan yang dibandingkan. Skala Jaro Winkler dikira berdasarkan kepada persamaan 2.3 di bawah.

$$d_j = \begin{cases} 0 & ; jika\ m = 0 \\ \frac{1}{3}\left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right) & ; sebaliknya \end{cases} \quad (2.3)$$

di mana

d_j	:	Skor Jaro
m	:	Bilangan aksara yang sepadan
s_1	:	Panjang rentetan pertama
s_2	:	Panjang rentetan kedua
t	:	Bilangan peralihan (*transposition*)

Skala Jaro-Winkler telah digunakan di dalam kajian steganografi untuk membandingkan kesamaan di antara teks pelindung dan teks stego seperti di dalam kajian Bhattacharyya, Indu, et al., (2011); Kingslin & Kavitha, (2015); Htet & Phyo, (2016); Iyer & Lakhtaria, (2016), Elmahi, Wahbi, & Sayed, (2017) dan Naqvi, Abbasi, Hussain, Khan, & Ahmad, (2018). Jadual 2.3 menunjukkan nilai skor Jaro Winkler bagi beberapa kajian lepas.

Jadual 2.3

Nilai Skor Jaro Winkler bagi Kajian Lepas

Penyelidik	**Skor Jaro-Winkler**
(Agarwal, 2013)	1
(Stojanov, Mileva, & Stojanovi, 2014)	Tidak teguh seperti yang dinyatakan oleh penyelidik
(Iyer & Lakhtaria, 2016a)	0.98
(Elmahi et al., 2017)	0.61
(Ramakrishnan et al., 2017)	1
(Banik & Bandyopadhyay, 2018)	0.97
(Naharuddin et al., 2018)	1
(Naqvi et al., 2018)	1
(Baawi et al., 2019)	0.98
(El Rahman & Nourah, 2019)	1

Mesej yang disembunyikan ke dalam teks pelindung berkemungkinan boleh menyebabkan perubahan terhadap teks stego. Mihaela (2011) di dalam kajiannya menyarankan agar memastikan bit yang disembunyikan tidak mengubah teks pelindung bagi mengelakkan isi kandungannya (teks stego) tidak dicurigai. Menurut Roslan, Mahmod, Udzir, & Zurkarnain (2014), sensitiviti data terhadap perubahan merupakan salah satu tajuk atau masalah penting yang mendorong kepada kecurigaan yang melibatkan ketakbolehkelihatan atau ketakbolehkesanan sesuatu teks stego dan perlu dibincangkan atau diperdebatkan. Menurut Ramakrishnan, Thandra, dan Srinivasula (2017) teknik steganografi yang baik mestilah dapat menyembunyikan mesej tanpa sebarang kecurigaan terhadap sistem visual manusia.

Kesimpulannya, kapasiti, keteguhan dan ketakbolehkelihatan merupakan tiga ukuran utama yang digunakan untuk mengukur prestasi steganografi yang merupakan tiga ciri-ciri utama di dalam bidang penyembunyian maklumat (Kumar et al., 2014). Kajian yang dijalankan oleh Sabri et al. (2018) menunjukkan kapasiti, keteguhan dan ketakbolehkelihatan merupakan ukuran utama yang digunakan oleh penyelidik untuk mengukur prestasi teks stego dengan peratusan penggunaan ukuran tersebut masing-masing ialah 32%, 14% dan 10% selain kelajuan, kecekapan dan andaian (*assumption*). Menurut Rahman et al., (2017), ketakbolehkelihatan adalah ukuran keselamatan dalam steganografi yang menentukan tahap persepsi kewujudan mesej rahsia dalam sesuatu teks pelindung.

Oleh itu, kajian ini memilih kapasiti, keteguhan dan ketakbolehkelihatan sebagai tiga ukuran utama untuk menilai prestasi teks stego yang dijana berdasarkan kepada

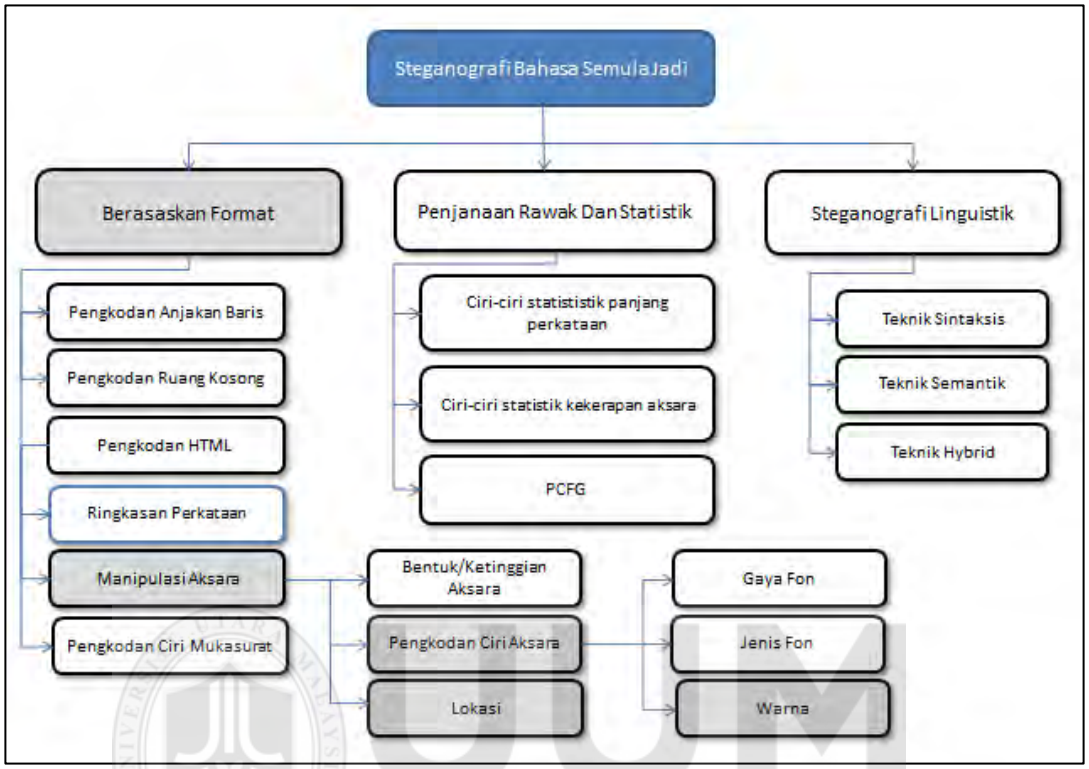
ukuran-ukuran yang telah digunakan oleh penyelidik terkini seperti Ahvanooey, Li, Shim, et al., (2018), Al-Azzawi, (2018) serta Alotaibi dan Elrefaei, (2018).

2.4 Penyembunyian Maklumat dan Steganografi

Steganografi merupakan sub disiplin di dalam bidang penyembunyian maklumat di mana tujuan utamanya adalah untuk melindungi mesej rahsia daripada dikesan dan mengelakkan penceroboh daripada mengasingkan mesej rahsia yang terdapat pada teks stego. Pelbagai medium pelindung digunakan untuk menyembunyikan mesej rahsia seperti teks, imej, audio dan video. Medium imej menjadi pilihan penyelidik untuk menyembunyikan mesej rahsia berbanding medium lain disebabkan medium ini mempunyai kapasiti ruang lewah yang besar berbanding dengan teks (Gongshen, Xiaoyun, Bo, & Meng, 2013). Walau bagaimanapun sejajar dengan perkembangan rangkaian komputer masa ini, pertukaran maklumat melalui Internet terutama menggunakan teks menjadi pilihan penyelidik kerana komunikasi melalui teks menggunakan sedikit ruang ingatan, kos percetakan yang rendah serta dapat berkomunikasi dengan lebih banyak maklumat (Kingslin & Kavitha, 2015). Justeru itu, steganografi teks masih mendapat perhatian penyelidik walaupun mempunyai cabaran yang tinggi dari segi ruang lewah yang kecil berbanding dengan medium lain.

Steganografi bagi bahasa semula jadi diklasifikasikan kepada tiga jenis iaitu; Berasaskan Format, Penjanaan Rawak dan Statistik serta Linguistik. Pelbagai teknik penyembunyian digunakan penyelidik seperti pengekodan perkataan, anjakan garis, pengekodan HTML, ringkasan perkataan, pengekodan ciri dan sebagainya. Dobriyal, Yadav dan Jain, (2015) telah mengklasifikasikan teknik-teknik tersebut berdasarkan kepada tiga kaedah di atas. Beberapa teknik-teknik lain telah dikenal pasti dan

dimasukkan ke dalam kategori yang berkaitan dan secara keseluruhan teknik penyembunyian tersebut dapat ditunjukkan dalam Rajah 2.6.



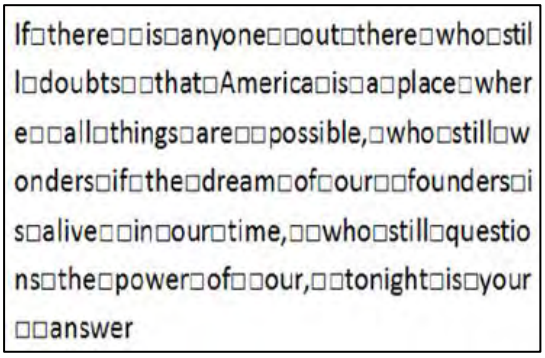
Rajah 2.6 Teknik-Teknik Penyembunyian

Kaedah linguistik serta kaedah rawak dan statistik mempunyai banyak limitasi seperti keperluan penggunaan kamus dan pengkomputan yang kompleks (Ahvanooey et al., 2019) dan hanya beberapa penyelidik memfokuskan kajian terhadap teknik tersebut berbanding kaedah berasaskan format. Oleh itu, teknik pengekodan ciri berasaskan format seperti yang ditunjukkan di dalam Rajah 2.6 digunakan di dalam kajian ini kerana kaedah ini menyediakan kadar ketakbolehkelihatan yang tinggi, kadar penyembunyian yang tinggi serta teguh terhadap serangan (Ahvanooey et al., 2019).

2.4.1 Kaedah Berasaskan Format

Kaedah berasaskan format memfokuskan kepada atribut atau ciri fizikal teks pelindung untuk menyembunyikan mesej rahsia. Menurut, Lwin dan Phyo (2014), kaedah ini tidak mengubah perkataan atau ayat serta tidak menjejaskan *„nilai"* sesuatu teks pelindung. Antara teknik yang menggunakan kaedah ini ialah seperti anjakan baris, anjakan perkataan, pengekodan HTML, ringkasan perkataan dan manipulasi aksara. Teknik penyembunyian pada ruang kosong antara perkataan, ruang kosong antara ayat atau ruang kosong di hujung ayat menyembunyikan bit "0" mesej rahsia pada satu ruang kosong manakala dua ruang kosong yang berturut-turut digunakan untuk menyembunyikan bit "1" seperti kajian yang dilakukan oleh Bhattacharyya, Banerjee dan Sanyal, (2010).

Teknik suntikan merupakan teknik yang dilakukan dengan membuat penambahan terhadap teks pelindung seperti penambahan ruang kosong antara perkataan, ruang kosong antara ayat, ruang kosong di hujung baris (Dobriyal et al., 2015), penambahan aksara tak kelihatan (*invisible character*) dan sebagainya seperti yang ditunjukkan di dalam Rajah 2.7. Kelemahan teknik ruang kosong ialah, proses pemampatan dan penyahmampatan boleh menyebabkan berlakunya kehilangan ruang kosong (tidak teguh) dan seterusnya menghasilkan ralat terhadap mesej rahsia yang diekstrakkan. Oleh itu, teknik ini berpotensi tinggi mengekstrak mesej rahsia yang salah sekiranya berlaku kehilangan ruang kosong.



Rajah 2.7 Teknik Suntikan Ruang Kosong

Selain daripada itu terdapat juga teknik-teknik lain yang digunakan untuk menyembunyikan mesej rahsia seperti sengaja membuat kesalahan pada sesuatu perkataan (Baawi et al., 2018), menambahkan ruang kosong, memformatkan gaya teks seperti *bold, italic* dan *underline* yang mana kaedah ini mungkin boleh menipu mata manusia tetapi sebaliknya ia boleh dikesan dengan mudah oleh komputer (Bennett, 2004).

Teknik memanipulasi aksara menyembunyikan mesej rahsia berdasarkan kepada bentuk, ciri-ciri sesuatu huruf seperti gaya tulisan, saiz, warna (Chaudhary & Dave, 2016) serta lokasi aksara. Teknik refleksi simetri yang dilakukan oleh (Majumder & Changder, 2013), membahagikan setiap abjad A-Z kepada empat kumpulan mengikut pantulan simetri abjad sama ada melintang, menegak, pantulan pada kedua-dua simetri atau tiada pantulan pada kedua-dua paksi seperti yang ditunjukkan di dalam Jadual 2.4 di bawah.

Jadual 2.4

Teknik Refleksi Simetri (Majumder & Changder, 2013)

Kumpulan	Ciri-Ciri	Aksara	Bit yang dipadankan
1	Tiada pantulan pada mana-mana paksi	C, F, G, J, L, N, P, Q, R, Z	00
2	Pantulan pada paksi melintang	B, D, E, K, S	01
3	Pantulan pada paksi menegak	A,M,T,U,V,W,Y	10
4	Pantulan pada kedua-dua paksi	H,I,O,X	11

Setiap kumpulan dipadankan dengan dua bit bagi setiap aksara untuk tujuan penyembunyian. Setiap dua bit mesej rahsia dipadankan dengan aksara pertama setiap ayat di dalam teks pelindung secara berjujukan. Teknik ini memilih sesuatu ayat sekiranya padanan berjaya dilakukan dan sebaliknya ayat tersebut akan diabaikan. Beberapa kelemahan dikenal pasti menggunakan teknik ini, antaranya pemilihan padanan sering kali menyebabkan ayat yang dijana tidak berkesinambungan di antara satu sama lain dan boleh mendorong kepada kecurigaan teks stego yang dijana terutama dari segi semantik ayat. Selain daripada itu, mesej yang disembunyikan adalah terhad kerana ayat yang tidak dapat dipadankan terpaksa diabaikan dan secara tidak langsung menyebabkan kapasiti penyembunyian berkurang. Namun begitu, bagi bit yang sama, terdapat kelebihan untuk mewakilkannya dengan aksara yang berbeza. Sebagai contoh, bit „01"boleh dipadankan dengan beberapa pilihan ayat yang bermula dengan salah satu abjad B, D, E, K atau S di dalam teks pelindung, namun ia dihadkan kepada lima pilihan sahaja. Seterusnya, adalah sukar untuk memilih ayat yang bermula dengan abjad-abjad tersebut di dalam teks pelindung kerana kemunculan abjad-abjad tersebut pada permulaan ayat kurang berlaku dan ini merupakan salah satu kelemahan yang terdapat di dalam teknik ini.

Kajian yang dijalankan oleh Kingslin dan Kavitha (2015), mendapati teknik penyembunyian berdasarkan jenis fon (*font type*) di dalam Microsoft Word dan Putaran Teks (*Text Rotation*)

di dalam Microsoft Excel telah menghasilkan nilai skor Jaro Winkler bersamaan 1. Walaupun perbezaan ketara terhadap teks stego tidak dapat dilihat oleh visual manusia, namun sistem komputer dapat mengenal pasti perbezaan jenis tulisan yang digunakan. Selain daripada itu, kedua-dua teknik di atas menghasilkan nilai kapasiti yang rendah iaitu 3.38% dan 2.90% masing-masing. Sementara itu, Kouser et al., (2017) menggunakan teknik pemetaan tiga bit mesej rahsia pada satu aksara teks pelindung dengan membahagikannya kepada 8 kumpulan seperti yang ditunjukkan di dalam Jadual 2.5.

Jadual 2.5

Teknik Zero-Text (Kouser et al., 2017)

Kumpulan	Ciri-Ciri	Aksara	Bit yang dipadankan
1	Aksara yang boleh ditulis dalam satu aliran, mempunyai garisan melintang atau menegak atau kedua-duanya, dan mempunyai lengkung separuh.	P, R, U	111
2	Aksara yang boleh ditulis dalam satu aliran, mempunyai garisan melintang atau menegak atau kedua-duanya, dan tiada lengkung	I,J,L M,N, Y, Z	110
3	Aksara yang boleh ditulis dalam satu aliran, tidak mempunyai garisan melintang atau menegak dan mempunyai lengkung separuh atau penuh.	C, G, O, S	101
4	Aksara yang boleh ditulis dalam satu aliran, tidak mempunyai garisan melintang atau menegak dan tiada lengkung.	V, W	100
5	Aksara yang tidak boleh ditulis dalam satu aliran, tidak mempunyai garisan melintang atau menegak dan mempunyai separuh lengkung.	A, B, D	011
6	Aksara yang tidak boleh ditulis dalam satu aliran, mempunyai garisan melintang atau menegak dan tidak mempunyai lengkung.	E, F, H, K, T	010
7	Aksara yang tidak boleh ditulis dalam satu aliran, tidak mempunyai garisan melintang atau menegak dan mempunyai lengkung separuh atau penuh.	Q	001
8	Aksara yang tidak boleh ditulis dalam satu aliran, tidak mempunyai garisan melintang atau menegak dan tidak mempunyai lengkung.	X	000

Jadual 2.5 menunjukkan setiap aksara teks pelindung dipetakan dengan tiga bit mesej rahsia. Teknik tersebut berjaya meningkatkan kapasiti penyembunyian kepada 23.25%

berbanding dengan teknik Refleksi Simetri. Namun, teknik ini tidak sesuai untuk mesej yang panjang kerana ia memerlukan saiz kekunci hampir tiga kali ganda bilangan aksara mesej rahsia. Antara kelemahan lain ialah, kekunci yang dijana mewakili lokasi aksara yang disembunyikan di dalam teks pelindung.

Kajian yang dilakukan oleh Sunariya, Din, dan Mahmudin (2016) mendapati 50% penyelidik menjalankan kajian steganografi menggunakan teknik berasaskan format berbanding dengan teknik linguistik (25%) dan teknik berasaskan *word-rule* (25%). Ini menunjukkan teknik berasaskan format masih mendapat sambutan berbanding dengan teknik-teknik lain.

2.4.2 Kaedah Penjanaan Rawak dan Statistik

Kaedah penjanaan statistik dan rawak digunakan untuk menjana teks pelindung secara automatik dan menyembunyikan mesej di dalam teks pelindung yang dijana mengikut ciri-ciri statistik sesuatu bahasa. Teks stego dijana berdasarkan kepada ciri-ciri statistik dokumen yang dipilih dengan menganggarkan statistik taburan abjad, perkataan, panjang perkataan dan sebagainya. Sebagai contoh, ciri-ciri statistik bagi panjang perkataan dan frekuensi aksara digunakan untuk membina perkataan baru yang mempunyai ciri-ciri yang hampir sama seperti teks asal. Mesej rahsia dipadankan dengan perkataan atau ayat di dalam teks pelindung berdasarkan kepada ciri-ciri di atas bagi menghasilkan teks stego di mana kualitinya bergantung sepenuhnya kepada nahu bahasa yang digunakan.

Markov Chains (Moraldo, 2012) dan *PCFG* merupakan teknik yang digunakan di dalam kaedah ini untuk menjana teks stego. Namun teks stego yang dijana sangat

cenderung kepada kesalahan tata bahasa dan semantik serta mudah dikesan oleh visual manusia (Hana‟a, 2008). Menurut Adesina, Nyongesa, dan Agbele (2010) teks stego yang dihasilkan dalam teknik ini kadangkala tidak bermakna dan tidak mengikut semantik yang betul di mana kelemahan ini boleh membawa kepada kecurigaan di dalam komunikasi rahsia. Secara amnya, kelemahan teknik ini ialah struktur, sintaks dan semantik teks stego yang dijana terdedah kepada kesilapan dan kecurigaan yang serius. Di samping itu, kualiti teks stego yang dijana juga bergantung sepenuhnya pada kualiti tatabahasa yang digunakan (Bhattacharyya, et al., 2011) selain mempunyai kapasiti penyembunyian mesej rahsia yang terhad. Fateh dan Rezvani (2018) menggunakan teknik penyembunyian pada alamat e-mel yang dijana berdasarkan kepada kandungan e-mel dengan kapasiti penyembunyian ialah ialah 10.6%. Teknik ini memerlukan penjanaan alamat e-mel yang banyak sekiranya saiz mesej rahsia semakin bertambah seperti yang dinyatakan di dalam kajian tersebut. Selain itu, alamat e-mel yang dijana menggunakan teknik ini merupakan alamat e-mel yang direka cipta serta tidak sah yang boleh mendorong kepada kecurigaan terhadap pihak ketiga. Menurut Ahvanooey et al. (2019) kaedah statistik dan rawak adalah lebih kompleks dan mengambil masa dan ruang untuk menjana teks stego.

2.4.3 Kaedah Linguistik

Kaedah linguistik menekankan kepada pengubahsuaian terhadap sifat linguistik sesuatu teks. Kaedah ini boleh dibahagi kepada dua iaitu kaedah sintaksis dan kaedah semantik. Kaedah sintaksis menekankan kepada penggunaan tanda baca, koma dan noktah dan sebagainya diletakkan di tempat yang betul di dalam sesuatu dokumen, manakala kaedah semantik pula menggantikan perkataan sinonim di dalam sesuatu teks. Penggunaan tata tanda adalah berbeza bagi beberapa bahasa, antaranya Bahasa

Hindi yang mempunyai tata tanda bahasa seperti …., -- , | , || dan sebagainya (Nagarhalli, Bakal, & Jain, 2016) berbanding bahasa yang menggunakan aksara A-Z yang kebiasaannya menggunakan tata tanda seperti noktah (.) dan koma (,) dan sebagainya di dalam penulisan.

Kaedah sintaksis akan menyembunyikan bit 0 pada tanda noktah (.) dan bit 1 pada tanda koma (,). Penggunaan tata tanda yang tidak konsisten akan menyebabkan teks stego yang dijana dapat dikesan dengan mudah oleh pembaca (Singh, Singh, & Saroha, 2009; Stojanov et al., 2014). Kaedah sintaksis adalah baik, namun penekanan yang terperinci perlu diberikan semasa menentukan kedudukan tata tanda di dalam sesuatu dokumen. Kajian yang dilakukan oleh Chaudhary, Dave, dan Sanghi (2016) mendapati bahawa, kaedah ini mempunyai dua kelemahan. Pertama penceroboh yang mempunyai pengetahuan yang tinggi di dalam sesuatu bahasa akan dapat mengetahui kedudukan sebenar tata tanda di dalam sesuatu dokumen; kedua, kaedah ini mempunyai kadar kapasiti penyembunyian yang rendah disebabkan jumlah tata tanda terhad yang terdapat di dalam sesuatu dokumen. *Spammimic* merupakan contoh teknik yang menggunakan kaedah ini di mana teks stego yang dihasilkan kelihatan seperti *spam* yang mungkin tidak dicurigai oleh penceroboh. Namun penggunaan mesej berbentuk *spam* adalah tidak dibenarkan dan akan ditapis atau dihapuskan di dalam kebanyakan organisasi.

Kaedah semantik akan menggantikan perkataan sinonim di dalam teks pelindung. Justeru itu, penggantian perkataan sinonim ini menyebabkan berlaku perbezaan yang nyata di antara teks pelindung dan teks stego. Contoh penggantian perkataan "*bagai*" kepada perkataan "*seperti*" akan menghasilkan perbezaan yang ketara di dalam teks stego yang dihasilkan seperti yang ditunjukkan di dalam Rajah 2.8 (a) dan Rajah 2.8

(b). Berdasarkan kepada Rajah 2.8(b), didapati terdapat perbezaan yang jelas telah berlaku terhadap teks stego yang dihasilkan. Selain berlakunya perbezaan terhadap teks stego yang dihasilkan, teknik ini juga menghasilkan kapasiti penyembunyian yang terhad (Mulunda & Wagacha, 2013) disebabkan limitasi terhadap penggantian perkataan yang sinonim dan tidak sesuai untuk mesej yang panjang.

Rajah 2.8 Ketaksamaan antara Teks Pelindung dan Teks Stego Menggunakan Kaedah Penggantian.

Sengaja membuat kesalahan pada sesuatu perkataan merupakan salah satu teknik yang digunakan penyelidik untuk menyembunyikan mesej. Namun, menurut Potdar, Han dan Chang (2005) teknik tersebut mudah dikesan serta menimbulkan kecurigaan oleh pembaca mahir walaupun huruf-huruf di campur aduk (Mulunda & Wagacha, 2013) seperti “How ***is*** you" ditukar kepada “How ***iz*** you". Ini kerana perkataan yang sengaja disalah eja boleh diperoleh dengan bantuan penyemak ejaan atau kamus bahasa. Kecurigaan akan menjadi semakin ketara apabila saiz mesej rahsia semakin bertambah disebabkan pertambahan kesalahan ejaan yang perlu dilakukan. Oleh itu, bagi memastikan ketakbolehkelihatan adalah sentiasa terjamin, pengubahsuaian gaya

penyampaian dan struktur teks mestilah tidak menjejaskan secara ketara makna atau nada sesuatu teks stego yang dihasilkan.

Secara keseluruhannya, perbandingan ketiga-tiga kaedah di atas boleh diringkaskan berdasarkan kepada kriteria ketakbolehkelihatan, kapasiti penyembunyian dan keteguhan (Ahvanooey et al., 2019) seperti yang ditunjukkan di dalam Jadual 2.6.

Jadual 2.6

Perbandingan Kaedah-Kaedah Steganografi Berdasarkan Kriteria (Ahvanooey et al., 2019)

Kaedah Steganografi	Ketakboleh kelihatan	Kapasiti Penyembunyian	Keteguhan	Kelebihan/Kekurangan
Berasaskan Format	Ya/Tidak	Tinggi	Bergantung kepada teknik yang digunakan	Tiada atau kurang perubahan terhadap CT. Bergantung kepada atribut CT. Sesuai untuk pelbagai bahasa
Linguistik	Ya	Rendah	Sederhana	Kaedah yang kompleks kerana memerlukan kamus data untuk proses penggantian aksara/perkataan. Kapasiti penyembunyian rendah Keteguhan terhadap serangan visual agak sederhana.
Statistik & Rawak	Tidak	Sederhana	Tinggi	Kaedah yang kompleks kerana memerlukan algoritma pemampatan untuk SM. Keteguhan terhadap serangan visual adalah tinggi. Bergantung kepada bahasa yang digunakan

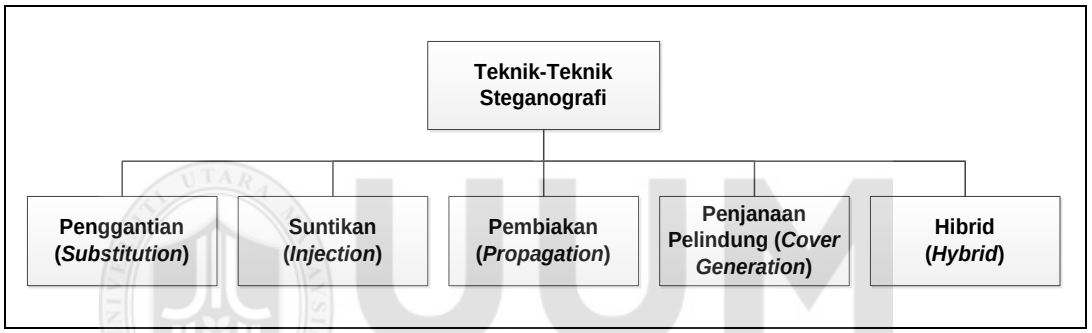
Nota : CT = Teks Pelindung SM = Mesej Rahsia

Jadual 2.6 menunjukkan ringkasan perbandingan di antara kaedah-kaedah steganografi yang digunakan dan didapati bahawa kaedah berasaskan format mempunyai potensi tinggi di dalam pelbagai aspek untuk meningkatkan prestasi

steganografi teks. Selain itu, kaedah berasaskan format juga sesuai digunakan untuk pelbagai bahasa berbanding dengan kaedah-kaedah lain (Ahvanooey et al., 2019).

2.5 Teknik-Teknik Penyembunyian

Teknik menyembunyikan maklumat di dalam steganografi teks boleh dikategorikan kepada teknik Penggantian, Suntikan/Masukan, Pembiakan, Penjanaan Pelindung, Hibrid dan Pelbagai (Rafat & Hussain (2017); Baawi et al., (2017); Odeh & Khaled, 2013) seperti yang ditunjukkan di dalam Rajah 2.9.



Rajah 2.9 Teknik-Teknik Steganografi

2.5.1 Teknik Penggantian

Teknik penggantian akan menggantikan atau membuat sedikit perubahan terhadap objek pelindung. Proses penggantian perlu dilakukan secara berhati-hati untuk mengelakkan objek stego yang dihasilkan tidak dicurigai. Teknik ini merupakan teknik yang popular digunakan di dalam medium imej, audio dan video. *Least Significant Bits* (LSB) dan *Most Significant Bit* (MSB) merupakan dua teknik yang paling popular digunakan di dalam medium imej, video dan audio di mana penggantian bit kurang menjejaskan kualiti objek stego (Jose, Chatterjee, Patodia,

Kabra, & Nath, 2016) dan dapat mengelakkan sebarang kecurigaan terhadap objek stego yang dihasilkan.

Teknik penggantian di dalam medium teks biasanya digunakan di dalam Linguistik Steganografi. Penggantian perkataan sinonim kadangkala menghasilkan teks stego yang berbeza dengan teks pelindung dan penggantian ini berpotensi untuk menyebabkan kesalahan semantik sesuatu ayat (El Rahman & Nourah, 2019) dan seterusnya boleh mendorong kepada kecurigaan terhadap teks stego yang dijana. Teknik linguistik memerlukan perhatian yang tinggi terhadap sintaks dan nahu sesuatu bahasa serta mempunyai kapasiti penyembunyian yang rendah (Ahvanooey et al., 2019) serta memerlukan saiz teks pelindung yang besar untuk melakukan proses penyembunyian (Chaudhary et al., 2016).

2.5.2 Teknik Suntikan

Teknik suntikan merupakan teknik menambahkan data ke dalam objek pelindung yang mana secara tidak langsung meningkatkan saiz objek stego dan berkemungkinan tinggi untuk terdedah kepada penceroboh apabila penambahan data berlaku dengan ketara. Teknik ini biasanya digunakan untuk medium imej, audio atau video manakala di dalam medium teks ialah dengan penambahan ruang kosong.

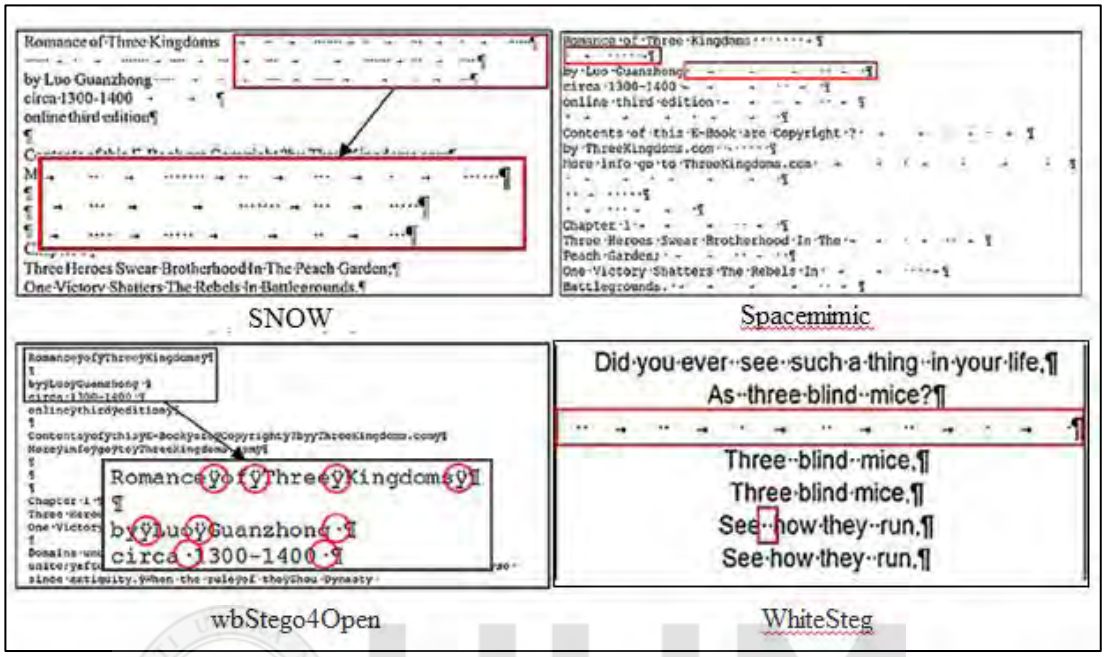
Teknik penyembunyian menggunakan ruang kosong mempunyai kelemahan dari segi ruang penyembunyian yang terhad serta tidak teguh terhadap pemampatan (*compression*) dan penyahmampatan (*decompression*). Kajian steganalisis terhadap ruang kosong yang dijalankan oleh Rauf, Rose, Jamal dan Nur Hafizah (2014) berkebolehan untuk mengesan steganografi teks secara efektif dengan prestasi

pengesanan mencapai sehingga 96.67% menggunakan teknik *visualization*. Bagi teknik memanipulasikan aksara seperti penambahan aksara tambahan, perubahan tanda baca, perubahan gaya tulisan dalam sesuatu dokumen akan mudah disedari dan dicurigai oleh pihak ketiga. Ini disebabkan teknik-teknik penyembunyian tersebut menyebabkan perubahan terhadap teks stego yang dihasilkan.

Teknik memanipulasikan ruang kosong, WhiteSteg (Por & Delina, 2008) hanya boleh menyimpan tiga belas bit mesej rahsia sahaja di dalam ruang antara perenggan dan tidak sesuai untuk menyembunyikan mesej rahsia yang panjang (Por, Wong & Chee, 2012). Selain itu, kajian yang dijalankan oleh Mahato et al. (2014) menyembunyikan mesej rahsia dengan mengubah fon saiz ruang kosong dengan kapasiti penyembunyian 2.01% sahaja kerana penyembunyian bergantung sepenuhnya kepada bilangan ruang kosong.

Kaedah memanipulasikan ruang kosong di dalam penyembunyian maklumat boleh menyebabkan kecurigaan terhadap sesuatu dokumen kerana teknik yang digunakan dapat dikesan oleh mata kasar atau peralatan steganalisis. Sebagai contohnya kajian yang dilakukan oleh Por et al. (2012) mendapati teknik *Dot and Arrow Show/Hide* (DASH) dapat mengesan dengan mudah kehadiran mesej rahsia di dalam dokumen Microsoft Word dengan menggunakan fungsi *show* dan *hide*. Menurut penyelidik, dokumen yang dimanipulasikan menggunakan teknik SNOW, *Spacemimic*, *wbStego4Open* dan *WhiteSteg* menunjukkan dengan jelas kewujudan ruang kosong dan tab yang dimanipulasikan yang diwakili oleh simbol "." dan "→" seperti yang ditunjukkan di dalam Rajah 2.10. Justeru itu, bagi mengatasi kelemahan tersebut, maka teknik yang dikenali sebagai UniSpaCh (Por et al., 2012) telah digunakan

dengan memanipulasikan penggunaan ruang kosong Unicode seperti dalam Rajah 2.11.



Rajah 2.10 Pendedahan Ruang Kosong "." dan Tab"" Menggunakan Fungsi *show/hide* dalam Microsoft Word 2007.

Space Character	Windows XP		Windows Vista		Windows 7	
	Hide	Show	Hide	Show	Hide	Show
Space	abc def	abc·def	abc def	abc·def	abc def	abc·def
En Quad	abc def	abc def	abc def	abc def	abc def	abc def
Em Quad	abc def	abc def	abc def	abc def	abc def	abc def
Three-Per-Em	abc def	abc def	abc def	abc def	abc def	abc def
Six-Per-Em	abc def	abc def	abc def	abc def	abc def	abc def
Figure	abc def	abc def	abc def	abc def	abc def	abc def
Punctuation	abc def	abc def	abc def	abc def	abc def	abc def
Thin	abc def	abc def	abc def	abc def	abc def	abc def
Hair	abc def	abc def	abc def	abc def	abc def	abc def

Rajah 2.11 Aksara Unicode Yang Digunakan Dalam Teknik UniSpaCh

Walaupun teknik UniSpaCh dapat mengatasi masalah yang dihadapi oleh keempat-empat teknik di atas, namun kehadiran mesej rahsia masih boleh dikesan

menggunakan analisis statistik seperti yang dinyatakan di dalam kajian tersebut. Ini disebabkan bit-bit yang disembunyikan adalah berdasarkan kepada saiz aksara Unicode yang dipadankan. Sebagai contoh, setiap pasangan mesej rahsia akan dipadankan dengan aksara Unicode seperti yang ditunjukkan dalam Jadual 2.7 di bawah.

Jadual 2.7

Perwakilan Aksara Unicode

Aksara	Kombinasi	Perwakilan
	Normal	00
	Thin + Normal	01
	Six-Per-Em + Normal	10
	Hair + Normal	11

Teknik UniSpaCh hanya teguh terhadap pemproses perkataan namun sebaliknya ia tidak teguh terhadap proses pemampatan dan penggunaan peralatan *Optical Character Recognition* (OCR). Kajian tersebut juga mendapati teknik *WhiteSteg*, *SNOW*, *Spacemimic* dan *wbStego4Open* masing-masing menghasilkan saiz fail teks stego yang lebih besar berbanding teknik UniSpaCh. Peningkatan saiz fail teks stego amat ketara apabila saiz mesej rahsia melebihi 64Kb terutama sekali bagi teknik *wbSteg4open* di mana ia mencecah hampir sembilan kali ganda lebih besar daripada teknik UniSpaCh berbanding dengan teknik-teknik lain. Walaupun teknik UniSpCh teguh terhadap pemproses perkataan, namun teknik padanan ruang kosong yang berbeza saiz menyebabkan berlakunya kecurigaan yang jelas terhadap visual manusia. Kelemahan ini jelas diterangkan di dalam kajian masa depan penyelidik tersebut yang akan memfokuskan terhadap ketakbolehkesanan terhadap teks stego yang dijana.

Baawi, Mokhtar, & Sulaiman (2019) menggunakan empat aksara tidak bercetak (*nonprintabale character)* iaitu U+200B,U+200D,U+200C dan U+200E) untuk menyembunyikan bit mesej rahsia pada lokasi permulaan, pertengahan, berasingan atau di hujung blok perkataan yang dijana. Secara amnya, kapasiti penyembunyian bagi teknik ini ialah 12.02% dengan satu lokasi dapat menyimpan empat bit mesej rahsia. Menurut Roy dan Manasmita (2011) penggunaan teknik ruang kosong adalah tidak teguh kerana mesej rahsia akan musnah apabila ruang kosong dihapuskan oleh perisian pemproses perkataan.

2.5.3 Teknik Pembiakan

Teknik pembiakan tidak bergantung pada medium pelindung, tetapi ia bergantung kepada enjin penjanaan. Mesej rahsia akan dimasukkan ke dalam enjin penjanaan yang akan mencipta satu fail tiruan (*mimic file*) sama ada dalam bentuk grafik, audio atau teks. Limitasi teknik ini ialah, fail yang dijana ini kadangkala di halang dalam komunikasi antara dua pihak kerana ia berunsurkan *spam*. Kegagalan berkomunikasi menyebabkan teks stego yang dihantar tidak diterima oleh penerima. Selain itu, penghantaran mesej yang bersifat *spam* secara berulang boleh mendatangkan kecurigaan terhadap mesej yang dihantar. *Spammimic* merupakan contoh peralatan yang menggunakan teknik ini dengan menjana teks stego tanpa berasaskan dokumen pelindung.

2.5.4 Teknik-Teknik Lain

Teknik memanipulasikan abjad berdasarkan kepada ciri-ciri fizikal abjad seperti CALP (Bhattacharyya, Indu, Dutta, Biswas, & Sanyal, 2011a), CURVE, VERT,

QUAD (Dulera, Jinwala, & Dasgupta, 2011), CASE (Sunita, Peeyush, Tarun, & Sharma, 2013), Symmetric Reflection (Majumder & Changder, 2013) adalah teguh terhadap pemampatan dan penyahmampatan. Walau bagaimanapun kecurigaan masih wujud pada teknik tersebut, kerana berlaku perubahan terhadap teks pelindung walaupun teknik-teknik tersebut menghasilkan kapasiti penyembunyian mesej hampir 8%. Selain itu, kajian steganalisis yang dijalankan oleh Din et al. (2013), mendapati sistem EDDS yang dibangunkan telah berjaya mengesan 100% mesej rahsia di dalam kajian yang dijalankan dengan nilai kecergasan melebihi dua puluh menggunakan teknik algoritma genetik.

Teknik pemampatan Huffman (Satir & Isik, 2014) berjaya meningkatkan kapasiti penyembunyian sehingga 7.96% bagi mesej rahsia bersaiz 300 aksara. Namun teknik tersebut menyebabkan berlaku kecurigaan disebabkan penggunaan alamat e-mel yang banyak apabila mesej rahsia yang digunakan melebihi 450 aksara. Justeru itu, kajian tersebut mencadangkan penggunaan Pengekodan Aritmetik bagi meningkatkan kapasiti dan penggunaan teknik rawakan bagi mengatasi masalah corak penyembunyian yang sama. Pengekodan Aritmetik telah berjaya meningkatkan kapasiti sehingga 5.95%, tetapi masih gagal menyembunyikan mesej rahsia yang panjang kerana berlakunya limpahan (Saniei & Faez, 2013a;Saniei & Faez, 2013b).

Kajian penjanaan ringkasan teks menggunakan teknik simetri refleksi telah dijalankan oleh Majumder dan Changder (2013) dengan membahagikan abjad kepada empat kumpulan yang mewakili bit 00,01,10 dan 11 berdasarkan kepada atribut simetri dan refleksi aksara seperti yang ditunjukkan di dalam Jadual 2.8.

Jadual 2.8

Pemetaan Aksara Mengikut Simetri Melintang dan Menegak

No	Ciri-Ciri Kumpulan	Aksara	Bit-Bit yang disembunyikan
1	Refleksi huruf bukan pada kedua-dua paksi.	C, F, G, J, L, N, P, Q, R, Z	00
2	Refleksi huruf pada paksi melintang.	B, D, E, K, S	01
3	Refleksi huruf pada paksi menegak.	A, M, T, U, V, W, Y	10
4	Refleksi huruf pada kedua-dua paksi.	H, I, O, X	11

Kajian tersebut berjaya menjana teks stego dengan memadankan aksara pertama setiap ayat dalam teks pelindung berdasarkan kepada Jadual 2.8. Ayat yang sepadan dengan bit yang disembunyikan akan dijana sebagai teks stego. Walau bagaimanapun, terdapat dua kelemahan yang dikenal pasti iaitu pertama; kemungkinan tinggi teks stego yang dijana menghasilkan ayat-ayat yang tidak berkesinambungan (Chaundhary, Dave, & Sanghi, 2016) antara satu sama lain yang berkemungkinan menyebabkan kecurigaan terhadap pihak penceroboh, kedua; kapasiti teks stego agak rendah (0.41%) kerana setiap ayat hanya dapat menyimpan dua bit aksara mesej rahsia dan tidak sesuai untuk mesej yang panjang.

Sementara itu, Roy dan Venkateswaran (2013) menggunakan taburan frekuensi abjad untuk menyembunyikan mesej dengan mengumpukkan nilai 0-15 kepada setiap abjad berdasarkan kepada peratusan frekuensi abjad (Lampiran A). Mesej yang disembunyikan ditukarkan kepada kod ASCII dan seterusnya ditukarkan ke bentuk nombor binari bersaiz lapan bit dan dibahagikan kepada empat bit setiap kumpulan. Setiap kumpulan bit dipadankan dengan jadual taburan frekuensi untuk proses penyembunyian seperti berikut

```
Mesej Rahsia          : text
Binari                : 0111 0100 0110 0101 0111 1000
Nilai Perpuluhan      :   7    4    6    5    7    8
Pemetaan ke Jadual    :   P    Y    G    B    M    D
```

Teks stego dijana berdasarkan aksara yang telah dipadankan seperti "***Promod Yadav has gone to Bangalore for the marriage of his daughter to Pormash Yadav***" Kelemahan teknik ini ialah, kecurigaan yang tinggi berlaku terhadap ayat yang dijana sekiranya ia tidak mengikut nahu tatabahasa. Sementara itu, kapasiti penyembunyian agak rendah iaitu 5.7% disebabkan hanya aksara terpilih sahaja boleh dilakukan proses penyembunyian.

Null Chiper merupakan teknik penyembunyian berdasarkan kedudukan abjad di dalam sesuatu perkataan. Teknik ini menyebabkan kesukaran untuk menjana teks stego terutama bagi mesej rahsia yang panjang di samping mempunyai kapasiti penyembunyian yang rendah. Contoh di bawah menunjukkan teknik penyembunyian menggunakan kedudukan huruf kedua setiap perkataan untuk menyembunyikan mesej rahsia.


```
"Apparently neutral's protest is thoroughly discounted and
 ignored. Isman hard hit. Blockade issue affects pretext for
 embargo on by-products, ejecting suets and vegetable oils".
 Mesej Rahsia : "Pershing sails from NY June 1".
```

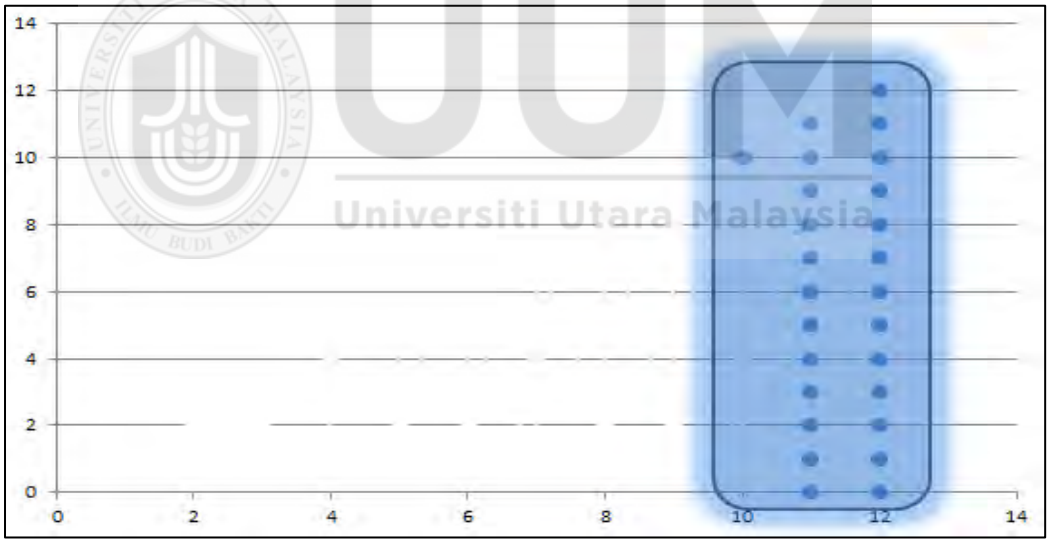
Sumber : David, (1967)

Teknik Simetri dan Refleksi, GATS dan CEBTS merupakan tiga teknik lepas yang menyebabkan ketaksamaan di antara teks pelindung dan teks stego (Lampiran B) yang mendorong kecurigaan terhadap teks stego. Sementara itu, Model Matematik Sistem Nombor (MMSN) merupakan skema yang digunakan untuk menukar dan mewakilkan sesuatu aksara dalam bentuk 2D atau koordinat (x,y) yang diperkenalkan oleh Mandal, Jana dan Agarwal (2014). MMSN menukarkan setiap aksara mesej

rahsia kepada nilai ASCII dan seterusnya menukarkannya ke bentuk koordinat *x* dan *y* berdasarkan kepada persamaan 2.4.

$$Nilai\ ASCII = (x * (x + 1)/\ 2) + y \qquad (2.4)$$

Kelemahan teknik ini ialah nilai (*x,y*) yang dihasilkan adalah sama dan statik bagi setiap aksara berulang mesej rahsia. Perwakilan nilai yang sama ini akan memberi laluan mudah kepada steganalisis untuk mengekstrak mesej yang disembunyikan. Ini kerana steganalisis akan cuba mengekstrak objek stego yang dihasilkan berdasarkan teknik pengesanan corak kerana perwakilan aksara hanya diwakili oleh 26 koordinat yang meliputi julat yang terhad di antara (10,10), (11,0) hingga (11,11) dan (12,0) hingga (12,12) seperti yang ditunjukkan di dalam Rajah 2.12.



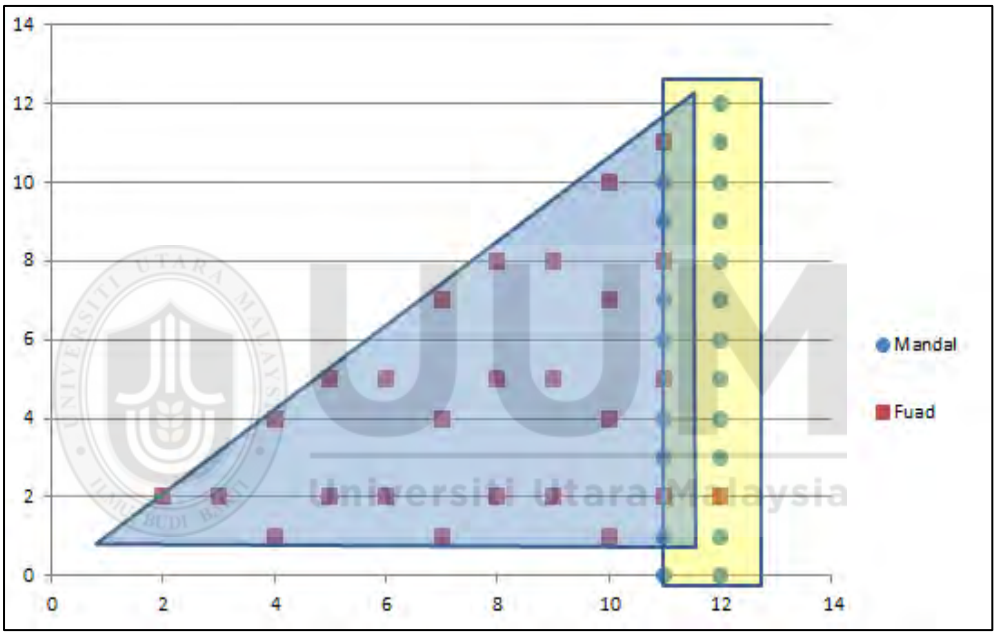
Rajah 2.12 Julat yang diwakilkan oleh Model Matematik Sistem Nombor

Kekangan terhadap kelemahan tersebut telah memberi kesan terhadap kapasiti penyembunyian mesej rahsia dan penjanaan stego objek. Fuad (2014) telah menambah baik formula yang diperkenalkan oleh Mandal et al. (2014) dengan menukarkan nilai

ASCII kepada perwakilan baru menggunakan persamaan 2.5 dan seterusnya menukarkan nilai tersebut kepada perwakilan (*x,y*) menggunakan persamaan 2.4.

$$z = ASCII - 60 + (n * 2), n = 0...25 ; \quad (2.5)$$

Penambahbaikan formula tersebut telah berjaya mengubah julat perwakilan koordinat tetapi masih meliputi kawasan yang terhad seperti yang ditunjukkan di dalam Rajah 2.13. Selain itu, aksara berulang mesej rahsia masih diwakilkan dengan nilai yang sama.



Rajah 2.13 Julat Perwakilan oleh Mandal et al., (2014) dan Fuad (2014).

Pendekatan penyembunyian menggunakan formula MMSN telah dilakukan oleh Koley dan Mandal, (2016) dan Koley et al., (2017) dengan mewakilkan aksara mesej rahsia dalam format (*x,y*) menggunakan persamaan 2.4. Rajah 2.14 menunjukkan contoh teks stego menggunakan teknik tersebut.

Mandal dan Koley (2016) **Mesej Tersembunyi : "SOLARIS"**	Koley dan Mandal, (2017) **Mesej Rahsia : "Ami 1 Ghanta "**
"Narayan Gopal Guruwacharya was a prominent popular singer of nepali music, who was died on 12/05/1990. Swami Vivekananda, a famous Indian hindu monk was born on 12/01/1893. The famous Indian Bollywood actor, Amitabh Bachchan was born on 11/10/1942. A famous writer R. K. Narayan was born on 10/10/1942. A famous Indian scientist Srinivasa Ramanujan was born on 12/04/1920. Indian railway minister, Suresh Prabhu was born on 11/07/1953. International Nurse‟s Day will be held on every year 12th May." (a)	"Dear customer blue print of your bunglow is ready. The specified bedroom size is (13,10) lobby (17,02) second bedroom (16,15) chairs are (040) (001) (040) storeroom (14,02) bedroom (16,14) two bathrooms (16,07) and (17,03) bedroom (17,11) (16,07) fourth chair (040) ……" (b)

Rajah 2.14 Teks Stego Menggunakan Teknik Mandal et al. (2016) dan Koley et al. (2017).

Kebarangkalian perwakilan format yang sama adalah tinggi apabila saiz mesej rahsia semakin panjang yang mana ia boleh menimbulkan kecurigaan teks stego yang dijana. Rajah 2.14(b) menunjukkan mesej rahsia "Ami 1 Ghanta" diwakilkan dengan 12 nilai *(x,y)* dan didapati aksara ruang kosong dan „a" diwakili dengan nilai *(x,y)* yang sama iaitu (040) dan (16,07). Semakin panjang mesej rahsia, maka semakin tinggi kesukaran untuk menjana teks stego serta mempunyai tahap kecurigaan yang tinggi disebabkan mengandungi format (*x,y*) berulang yang banyak serta kekangan untuk menjana ayat yang sesuai merupakan cabaran utama yang dihadapi oleh penyelidik di dalam teknik ini.

Satir dan Isik (2012b) di dalam kajian beliau menyatakan bahawa antara kelemahan yang ditimbulkan oleh penyelidik ialah berkaitan dengan penyembunyian aksara yang mempunyai corak yang sama. Oleh itu penyelidik mencadangkan penggunaan ciri perawakan yang boleh digunakan untuk mewakili aksara berulang bagi mengelakkan berlakunya pembentukan corak yang sama. Justeru itu, kaedah berbentuk dinamik perlu diwujudkan bagi membolehkan setiap aksara mesej tersembunyi diwakilkan

dengan nilai yang berbeza kerana kebanyakan mesej rahsia mempunyai aksara yang berulang.

Teknik penyembunyian berdasarkan lokasi aksara telah dilakukan oleh Htet dan Phyo (2016) dengan menyembunyikan bit mesej rahsia pada lokasi pertama, kedua, kedua terakhir dan terakhir setiap perkataan. Setiap perkataan dipilih secara jujukan berdasarkan kepada beberapa peraturan yang telah ditetapkan. Teknik ini membandingkan dua aksara pada setiap perkataan. Sekiranya terdapat perbezaan, maka aksara tersebut akan dipilih sebagai kekunci stego untuk proses penyembunyian. Antara kelemahan yang dikenal pasti dalam teknik ini ialah dari aspek kapasiti penyembunyian yang agak rendah kerana hanya satu bit dapat disembunyikan pada setiap aksara. Selain itu, kekunci stego yang dijana bertambah sebanyak lapan kali ganda kerana ia bergantung sepenuhnya kepada bilangan aksara yang diperlukan dan tidak sesuai untuk mesej yang panjang.

Naharuddin et al. (2018) menggunakan teknik penyembunyian dengan memetakan setiap bit aksara mesej rahsia dengan aksara yang terdapat di dalam teks pelindung. Lokasi bit yang dipadankan (0-7) dijadikan sebagai kekunci stego untuk tujuan pengekstrakan. Purata kapasiti penyembunyian ialah 14.2% dengan satu aksara mesej rahsia memerlukan 7 aksara teks pelindung dengan skor Jaro Wrinkler ialah 1. Kelemahan utama teknik ini ialah penjanaan kekunci stego yang panjang kerana ia berkadar terus dengan saiz mesej rahsia seperti yang ditunjukkan di dalam Lampiran C. Sementara itu, Lampiran D menunjukkan ringkasan analisis kajian lepas berkaitan dengan teknik-teknik yang digunakan sejak tahun 2012 hingga 2019.

Secara ringkasnya, kapasiti penyembunyian bagi kesemua teknik agak rendah dengan kapasiti tertinggi ialah 24.4%. Jumlah bit yang dapat disembunyikan pada setiap aksara mempengaruhi kapasiti penyembunyian mesej rahsia. Selain kapasiti, aksara mesej rahsia yang berulang diwakilkan atau dipetakan dengan nilai yang sama boleh menyebabkan berlakunya corak perwakilan yang sama dan berpotensi tinggi untuk pihak ketiga mengekstrak mesej rahsia. Di samping itu, teknik penyembunyian mesej rahsia pada lokasi berjujukan merupakan isu penting yang perlu dielakkan semasa melakukan proses penyembunyian seperti yang ditimbulkan oleh Satir dan Isik (2012b) dan teknik rawakan boleh digunakan seperti yang dicadangkan oleh penyelidik. Sebarang perubahan terhadap teks pelindung boleh menyebabkan berlakunya kecurigaan terhadap teks stego yang dijana dan perlu diminimumkan sebaik mungkin agar visual manusia tidak dapat mengesan kehadiran mesej yang disembunyikan di samping memastikan kapasiti penyembunyian dapat dipertingkatkan.

2.6 Penyembunyian Berasaskan Aksara

Kajian yang dilakukan oleh Ramakrishnan, Thandra, dan Srinivasula (2017) menyimpulkan bahawa peratus taburan aksara di dalam dokumen Bahasa Inggeris boleh diklasifikasikan kepada tiga kategori iaitu Tinggi, Sederhana dan Rendah seperti yang ditunjukkan di dalam Jadual 2.9.

Jadual 2.9

Peratus Kekerapan Aksara di dalam Dokumen

Kebarangkalian Kejadian	Aksara	Kekerapan	Kebarangkalian Kejadian	Aksara	Kekerapan
Tinggi	SPACE	20.29%	Rendah	G	1.69%
	E	9.63%		Y	1.55%
	T	7.56%		F	1.50%
	A	6.84%		Dot	1.39%
	O	6.31%		B	1.36%
	I	5.45%		C	1.21%
	S	5.06%		K	1.14%
Sederhana	H	4.96%		P	0.86%
	N	4.84%		V	0.78%
	D	3.93%		J	0.30%
	R	3.57%		Z	0.08%
	L	3.01%		X	0.06%
	U	2.27%		Q	0.05%
	W	2.21%			
	M	2.08%			

Berdasarkan Jadual 2.9, didapati aksara E,T,A,O,I,S mewakili peratus kekerapan kategori tertinggi dengan peratus kekerapan antara 5.06% hingga 9.63% (tidak termasuk ruang kosong) dan diikuti dengan aksara kategori sederhana yang terdiri dari aksara H,N,D,R,L,U,W,M dengan peratus kekerapan antara 2.08% hingga 4.96%. Aksara lain hanya muncul kurang dari 2% dengan aksara J,Z,X dan Q merupakan aksara paling kurang muncul dengan peratus kekerapan kurang dari 0.5%. Kajian yang dijalankan oleh Grigas dan Juškevičienė, (2018) menghasilkan peratus taburan kekerapan yang sama seperti yang diperoleh oleh Ramakrishnan et al., (2017).

Bhattacharyya, Indu, et al. (2011) di dalam kajiannya memilih 12 aksara yang telah dikenal pasti dengan memadankan setiap kumpulan aksara tersebut dengan satu bit atau dua bit mesej rahsia. Setiap aksara yang disembunyikan akan diformatkan berdasarkan kepada bentuk aksara yang telah diubahsuai seperti yang ditunjukkan di dalam Rajah 2.15.

Aksara Asal	Perubahan aksara untuk 1 bit	Bit yang disembunyikan	Perubahan aksara untuk 2 bit	Bit yang disembunyikan
A	Ạ	1	Ạ	10
a	ạ	1	ạ	01
c	ḉ	1	ḉ	11
i	ị	0	ị	00
j	ȷ	0	ȷ	10
h			ḥ	01
x			ẋ	01
w			ẉ	10
n			ṉ	11
f			ḟ	11
p			ṗ	00
q			q̣	00

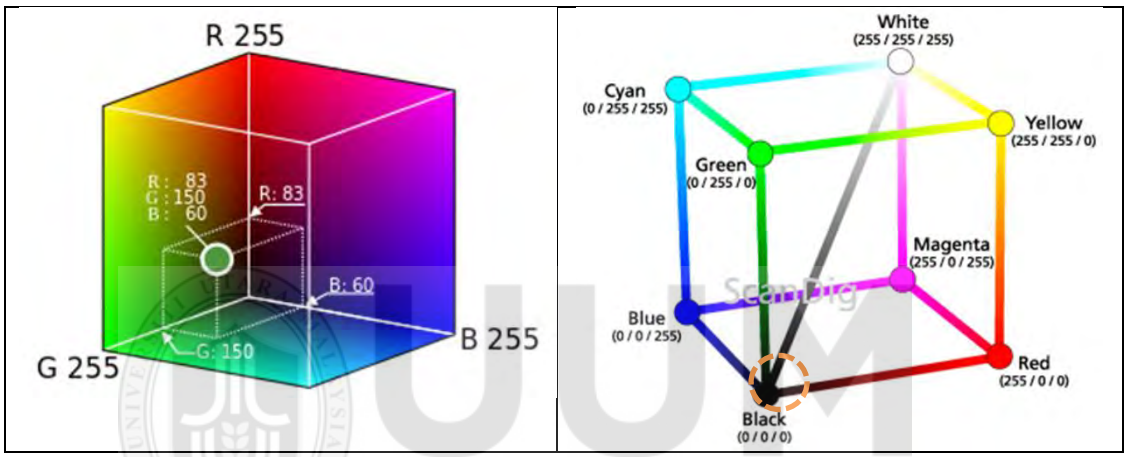
Rajah 2.15 Teknik Memformatkan Aksara (Bhattacharyya, Indu, et al., 2011)

Rajah 2.15 menunjukkan teknik penyembunyian dengan memetakan 1 bit atau 2 bit aksara mesej rahsia dengan aksara berikut; i,p,q (00), a,h,x (01), A,j,w (10) dan c,n,f (11). Kelemahan yang dikenal pasti ialah, kapasiti penyembunyian agak rendah kerana hanya dua belas aksara sahaja yang terlibat untuk proses penyembunyian. Selain itu, setiap aksara hanya boleh menyembunyikan maksimum dua bit sahaja di samping lokasi penyembunyian adalah secara berjujukan.

Namun begitu, Krishnan, Thandra, dan Baba (2017) yang melakukan kajian dengan membandingkan teknik penyembunyian peringkat aksara, peringkat bit serta teknik gabungan mendapati bahawa kapasiti penyembunyian bagi teknik penyembunyian peringkat aksara adalah yang terbaik kerana ia memerlukan saiz teks pelindung yang kecil untuk menyembunyikan mesej rahsia. Oleh itu penyembunyian mesej menggunakan abjad dijangka mampu untuk meningkatkan kapasiti mesej yang disembunyikan.

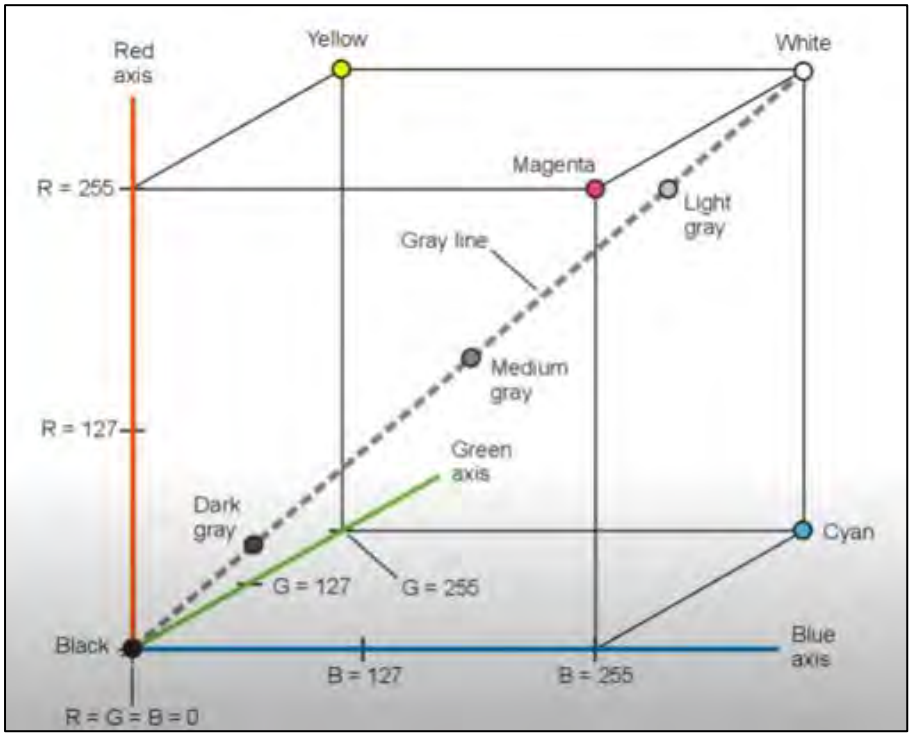
2.7 Model Kiub Warna RGB

Model warna RGB merupakan model gabungan warna merah, hijau dan biru dicampurkan bersama dengan pelbagai nilai RGB untuk menghasilkan warna baru. Satu bait (8 bit) setiap warna RGB boleh diwakili dengan nilai antara 0-255 dengan kombinasi setiap warna RGB akan menghasilkan sebanyak 255^3 atau 16777216 warna unik seperti yang ditunjukkan dalam Rajah 2.16.



Rajah 2.16 Model Kiub Warna RGB

Rajah 2.16 menunjukkan titik di sepanjang garisan diagonal bermula dengan warna gelap (0,0,0) dan bertukar kepada semakin cerah dengan perubahan nilai RGB dari kelabu gelap (*dark grey*), kelabu pertengahan (*medium* grey), kelabu cerah (*light* grey) dan putih (*white*) seperti yang ditunjukkan dalam Rajah 2.17. Nilai RGB (128,128,128) merupakan warna kelabu pertengahan dan RGB (255,255,255) merupakan warna putih manakala warna RGB yang paling hampir dengan warna hitam ialah RGB (15,15,15) iaitu 1/16 daripada nilai 256 warna. RGB (15,15,15) secara amnya mengandungi 5.8%. peratus warna RGB yang seimbang bagi setiap warna.



Rajah 2.17 Kiub Warna RGB Gelap

Perubahan julat RGB antara nilai 0-15 akan menghasilkan nilai kelabu gelap berbanding nilai lain seperti yang ditunjukkan di dalam Rajah 2.18.

15, 15, 15	15, 15, 15	15, 15, 15
15, 13, 13	15, 16, 16	36, 36, 36
15, 12, 12	15, 18, 18	58, 58, 58
15, 10, 10	15, 19, 19	81, 81, 81
15, 9, 9	15, 21, 21	105, 105, 105
15, 7, 7	15, 22, 22	130, 130, 130
15, 6, 6	15, 24, 24	156, 156, 156
15, 5, 5	15, 25, 25	183, 183, 183
15, 3, 3	15, 27, 27	210, 210, 210
15, 1, 1	15, 28, 28	239, 239, 239

Rajah 2.18 Warna Kelabu RGB

Pertambahan nilai RGB melebihi nilai 15 akan menyebabkan warna palet menjauhi warna gelap dengan ketara. Skala 0-15 merupakan skala warna RGB yang hampir dengan warna gelap. Menurut Alanazi, Zaidan, Zaidan, Jalab, dan AL-Ani, (2010) yang menjalankan kajian terhadap steganografi imej menyatakan bahawa model warna RGB merupakan teknik asas yang digunakan dalam pengkomputeran untuk mengekodkan warna dengan mewakilkannya dengan pelbagai nilai binari. Kajian yang dilakukan oleh Singh, Diwakar, dan Upadhyaya, (2014) menyatakan bahawa julat warna RGB (0,0,0) hingga (15,15,15) boleh digunakan di dalam teks steganografi kerana ia tidak merosakkan integriti visual dokumen yang dijana. Selain itu, kajian yang dijalankan oleh Bhadra, Bojamma, Prasad, dan Nachappa, (2014) mendapati bahawa palet warna skala kelabu yang digunakan dalam steganografi imej (.GIF) sukar dikesan oleh sistem visual manusia selepas melalui proses penyembunyian mesej rahsia.

2.8 Penyembunyian Berasaskan Warna RGB

Penyembunyian maklumat berasaskan atribut teks seperti fon (*font*), warna dan garis bawah (*underline*) merupakan beberapa teknik yang digunakan penyelidik di dalam steganografi teks. Menurut Chaudhary & Dave (2016), penyembunyian mesej rahsia berdasarkan atribut teks dapat meningkatkan kapasiti penyembunyian mesej rahsia. Atribut seperti jenis tulisan, saiz tulisan, gaya tulisan, warna tulisan boleh digunakan untuk proses penyembunyian. Justeru itu, atribut warna RGB telah digunakan di dalam kajian steganografi oleh beberapa penyelidik seperti perubahan warna aksara (Al-Asadi & Bhaya, 2016; Singh & Diwakar, 2014), warna aksara dan warna garis bawah (Wang & Li, 2014), warna ruang kosong, warna margin muka surat dan perenggan (Stojanov et al., 2014), gabungan warna aksara dan warna garis bawah

(Tang & Chen, 2013) dengan menggunakan pelbagai nilai warna RGB dengan julat antara (0,0,0) hingga (255,255,255).

Kapasiti penyembunyian menggunakan teknik warna RGB bergantung kepada jumlah bit yang boleh disembunyikan menggunakan warna tersebut. Kajian yang dilakukan oleh Tang dan Chen (2013) menyembunyikan lapan bit mesej rahsia pada setiap dua aksara teks pelindung menggunakan enam belas kombinasi warna RGB seperti yang ditunjukkan di dalam Jadual 2.10. Setiap aksara mesej rahsia ditukarkan kepada perwakilan perenambelasan (*hexadecimal)* dan dipadankan dengan jadual tersebut. Atribut warna dan garis bawah digunakan oleh penyelidik ini untuk dipadankan dengan warna RGB bagi menyembunyikan mesej rahsia di mana pemilihan lokasi penyembunyian adalah secara berjujukan.

Jadual 2.10

Perwakilan warna RGB oleh Tang & Chen, (2013)

Nilai Hexadecimal	**Warna RGB**	**Nilai Hexadecimal**	**Warna RGB**
0	0,0,0	A atau a	1,0,1
1	0,0,1	B atau b	1,0,2
2	0,0,2	C atau c	1,1,0
3	0,1,0	D atau d	1,1,1
4	0,1,1	E atau e	1,1,2
5	0,1,2	F atau f	1,2,0
6	0,2,0		
7	0,2,1		
8	0,2,2		
9	1,0,0		

Wang dan Li (2014) menggunakan gabungan atribut fon dan garis bawah aksara untuk menyembunyikan mesej rahsia. Setiap garis bawah dapat menyembunyikan 24 bit dengan setiap satu komponen warna RGB dipadankan dengan 8 bit mesej rahsia.

Sebagai tambahan bagi meningkatkan kapasiti penyembunyian, sebanyak 8 bit lagi disembunyikan pada aksara tersebut dengan setiap komponen RGB disembunyikan dengan 2, 2 dan 4 bit masing-masing. Secara keseluruhan sebanyak 32 bit dapat disembunyikan bagi setiap satu aksara teks pelindung. Walaupun aksara tersembunyi tidak dapat dilihat dengan visual manusia dan mempunyai kapasiti penyembunyian empat kali ganda saiz teks pelindung, namun kelemahan teknik ini ialah setiap aksara yang terlibat akan mempunyai garis bawah disebabkan proses perubahan warna RGB terhadap garis bawah aksara yang digunakan. Selain itu setiap aksara tersembunyi berulang diwakilkan dengan nilai ASCII yang sama selain penyembunyian dilakukan secara berjujukan seperti di bawah.

Mesej Rahsia	:	M	E	E	T	Y	O	U	A	T	T	E	N
Kod ASCII	:	77	69	69	84	89	79	85	65	84	84	69	78

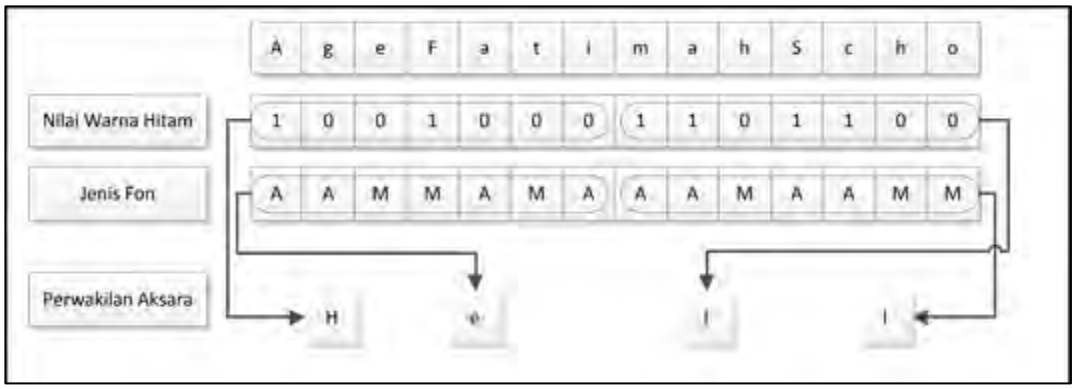
Sementara itu, Singh et al. (2014) menjalankan kajian dengan menggunakan dua nilai warna RGB iaitu 0,0,0 dan 1,1,1 untuk menyembunyikan bit 0 dan 1 masing-masing. Oleh itu, lapan aksara diperlukan untuk menyembunyikan satu aksara mesej rahsia yang disembunyikan secara berjujukan. Untuk meningkatkan kapasiti penyembunyian, kajian tersebut telah ditambah baik dengan menambahkan nilai variasi warna kepada 16 warna RGB iaitu (0,0,0), (1,1,1), (2,2,2) hingga (15,15,15). Namun kapasiti penyembunyian masih pada tahap yang rendah dengan purata 11.57% bagi teks pelindung bersaiz antara 220 hingga 4000 aksara.

Seterusnya, Stojanov et al. (2014) melakukan empat kaedah penyembunyian yang dikenali sebagai pengekodan ciri (*Property Coding*) dengan melakukan penyembunyian menggunakan empat teknik yang berbeza iaitu skala aksara, garis

bawah aksara, sempadan perenggan dan sempadan ayat. Hanya dua teknik menggunakan penyembunyian berasaskan warna RGB iaitu teknik garis bawah aksara dan teknik sempadan perenggan. Teknik garis bawah aksara dapat menyembunyikan 8 bit pada setiap aksara menggunakan 16 warna putih garis bawah yang berbeza dengan kapasiti penyembunyian ialah 95.6%. Walau bagaimanapun, aksara *g, j, p, q* dan *y* tidak digunakan di dalam teknik ini kerana kecurigaan pada aksara tersebut berlaku apabila format garis bawah dilakukan dan dapat dikenal pasti oleh sistem visual manusia. Manakala teknik sempadan perenggan yang dijalankan oleh penyelidik tersebut menggunakan 216 warna RGB antara (0,0,0) hingga (249,249,249) dapat menyembunyikan mesej dengan kapasiti penyembunyian yang agak rendah iaitu 2.1 peratus.

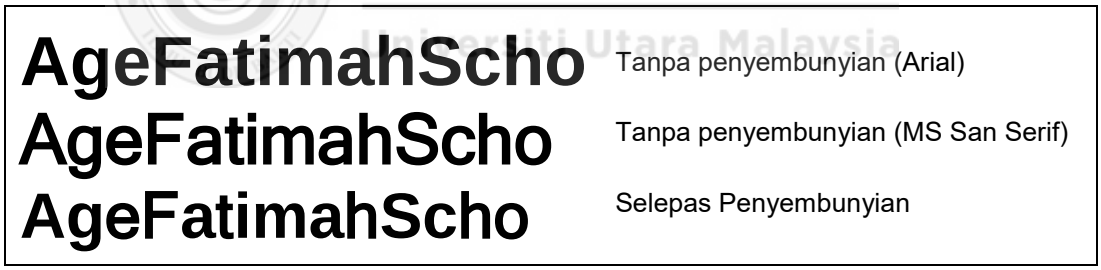
Microsoft Excel juga digunakan sebagai medium teks pelindung selain daripada Microsoft Word. Al-Asadi dan Bhaya (2016) menjalankan kajian dengan menyembunyikan bit mesej rahsia dengan menggabungkan warna hitam (0 dan 1) serta fon Arial dan Microsoft San Serif untuk mewakili bit 0 dan 1. Pemilihan sel teks pelindung berdasarkan kepada Jadual Index yang telah dijana. Penyembunyian dilakukan dengan mengambil 14 bit pertama mesej rahsia dan dibahagikan kepada dua kumpulan dengan tujuh bit setiap kumpulan. Tujuh bit pertama mesej rahsia dipadankan dengan tujuh aksara pertama teks pelindung yang diformatkan dengan warna RGB 0 atau 1, manakala tujuh bit seterusnya dipadankan dengan jenis fon Arial (0) atau Microsoft San Serif (1) untuk aksara yang sama. Proses penyembunyian berulang sehingga akhir mesej rahsia. Secara amnya, setiap aksara teks pelindung dapat menyembunyikan 2 bit mesej rahsia. Rajah 2.19 menunjukkan mesej rahsia "Hello" dengan nilai binari **1001000**1100101**1101100**1101100**1101111**

disembunyikan di dalam dokumen Microsoft Excel menggunakan teknik di atas dengan lokasi penyembunyian secara jujukan.



Rajah 2.19 Teknik Penyembunyian Menggunakan Warna Hitam

Kapasiti penyembunyian agak rendah kerana tujuh aksara teks pelindung diperlukan untuk menyembunyikan dua aksara mesej rahsia. Selain itu perubahan terhadap jenis tulis boleh menyebabkan kecurigaan terhadap sistem visual manusia dari segi bentuk tulisan dan panjang perkataan seperti yang ditunjukkan dalam Rajah 2.20.



Rajah 2.20 Perubahan Terhadap Jenis Tulisan dan Panjang Perkataan

Teknik penyembunyian berdasarkan pemampatan LZW dan pengekodan warna yang dilakukan oleh Malik, Sikka, dan Verma (2016) telah berjaya meningkatkan kapasiti mesej rahsia sehingga 13.43%. Teknik ini menggunakan 32 warna yang pelbagai dengan setiap satu aksara dapat menyembunyikan 10 bit mesej rahsia dengan memformatkan sempadan dan isi warna (*border and fill color*) aksara dengan warna yang dipadankan pada Rajah 2.21.

S. No.	Boundary line colors		Fill area Color		Binary code	S. No.	Boundary line Color		Fill area Color		Binary code
	Name	Color	Name	Color			Name	Color	Name	Color	
1	Rackley		Violet		00000	17	Green		Blue		10000
2	Rackley		Blue		00001	18	Green		Rackley		10001
3	Rackley		Rose		00010	19	Green		Rose		10010
4	Rackley		Black		00011	20	Green		Black		10011
5	Rackley		Scarlet		00100	21	Black		Violet		10100
6	Violet		Blue		00101	22	Black		Blue		10101
7	Violet		Rackley		00110	23	Black		Rackley		10110
8	Violet		Rose		00111	24	Black		Rose		10111
9	Violet		Black		01000	25	Black		Scarlet		11000
10	Violet		Scarlet		01001	26	Blue		Violet		11001
11	Rose		Violet		01010	27	Blue		Rackley		11010
12	Rose		Blue		01011	28	Blue		Rose		11011
13	Rose		Rackley		01100	29	Blue		Black		11100
14	Rose		Black		01101	30	Blue		Scarlet		11101
15	Rose		Scarlet		01110	31	Scarlet		Violet		11110
16	Green		Violet		01111	32	Scarlet		Blue		11111

Rajah 2.21 Perwakilan Warna Menggunakan Teknik Pemampatan LZW (Malik et al., 2016)

Walau bagaimanapun teknik ini menyebabkan perubahan ketara terhadap teks stego yang dihasilkan dan dapat dikesan dengan jelas oleh sistem visual manusia seperti yang ditunjukkan oleh Rajah 2.22. Selain itu, lokasi penyembunyian dilakukan adalah secara berjujukan.

Rajah 2.22 Teknik Pemampatan LZW dan Pengekodan Warna

Malik et al. (2017) seterusnya menggunakan teknik pemampatan Huffman dan Pengekodan warna RGB untuk menyembunyikan mesej pada setiap aksara. Sebanyak lapan warna digunakan dan dibahagikan kepada dua kumpulan untuk mewakili bit 0 dan 1 seperti yang ditunjukkan di dalam Rajah 2.23. Setiap aksara teks pelindung diformatkan dengan warna bit aksara tersembunyi yang sepadan dengan Rajah 2.23.

Kapasiti penyembunyian menggunakan teknik ini agak rendah kerana satu aksara hanya boleh menyembunyikan satu bit sahaja. Rajah 2.25 menunjukkan contoh teks stego yang dihasilkan menggunakan teknik ini berdasarkan kepada sebahagian bit mesej rahsia "01100100010001001010101001111111100111011100111011101111010 10 11011011110111110011010111010011000110 1….."

S.No.	Color	Color Name	Bit representation
1		Fluorescent Pink	0
2		Forest Green	0
3		French Blue	0
4		French Puce	0
5		[illegible]	1
6		Fuzzy Wuzzy	1
7		Dark Cyan	1
8		Deep Magenta	1

Rajah 2.23 Pemetaan Warna dan Bit Mesej Rahsia (Malik et al., 2017)

" in the research area of text steganography, algorithms based on font format have advantages of great capacity, good imperceptibility and wide application range. However, little work on steganalysis for such algorithms has been reported in the literature. based on the fact that the statistic features of font format will be changed after using font-format based steganographic algorithms "

Rajah 2.24 Teks Stego Menggunakan Teknik Pemampatan Huffman dan Pengekodan Warna (Malik et al., 2017)

Teks stego yang ditunjukkan dalam Rajah 2.22 dan Rajah 2.24 jelas menunjukkan perbezaan warna yang ketara antara teks stego dan teks pelindung. Menurut Fateh dan Rezvani (2018), teknik ini mengalami tahap keselamatan yang rendah kerana ia mengubah teks pelindung asal dengan perubahan yang ketara. Walaupun tiada

perubahan dari segi isi kandungan, namun visual manusia dapat melihat dengan jelas perubahan ini dan secara tidak langsung ketakbolehkelihatan dari segi visual wujud dengan jelas bagi kedua-dua teknik di atas.

Al-Azzawi (2018) menjalankan kajian dengan menggunakan teknik penyembunyian berdasarkan tag *part-of-speech* (POS) dan pengekodan warna RGB. POS merupakan tag yang dipadankan dengan setiap perkataan yang terdapat di dalam sesuatu dokumen. Beberapa perkataan dari teks pelindung dipilih dan disusun berdasarkan kepada bilangan lapisan yang dijana secara rawak. Bilangan lapisan digunakan sebagai kekunci untuk menentukan perkataan-perkataan yang terpilih. Setiap aksara pada perkataan terpilih digunakan untuk melakukan proses penyembunyian dengan memformatkan aksara tersebut dengan warna RGB. Setiap 12 bit mesej rahsia akan dipetakan dengan komponen warna RGB (R=4, G=4,B=4). Julat nilai RGB yang digunakan adalah di antara (0,0,0) hingga (15,15,15) yang mewakili nilai yang menghampiri warna hitam. Teknik ini dapat menyembunyikan 12 bit mesej rahsia pada setiap satu aksara pada perkataan yang terpilih dengan kapasiti penyembunyian adalah 100%. Walaupun kapasiti penyembunyian agak tinggi, namun jumlah kapasiti ini bergantung sepenuhnya kepada bilangan lapisan (kekunci) yang dipilih.

Secara keseluruhannya, teknik penyembunyian berasaskan warna RGB dapat diringkaskan berdasarkan kepada bilangan bit yang dapat sembunyikan dan julat warna yang digunakan seperti yang ditunjukkan di dalam Jadual 2.11.

Jadual 2.11

Ringkasan Teknik Penyembunyian Berasaskan Warna RGB

Bil	Penyelidik	Tahun	Jumlah Bit Penyembunyian	Julat/Warna yang Digunakan	Teknik	Kebolehkelihatan Visual Manusia
1.	Khairullah	2009	24 bit pada setiap ruang kosong.	Menggunakan RGB(0,0,0) hingga RGB(255,255,255).	• Menyembunyikan SM dalam ruang kosong, *tab* atau *Carriage Return*. • Menggunakan warna latar depan (*foreground*) untuk penyembunyian. • Penyembunyian secara jujukan.	Rendah
2.	Tang & Chen	2013	8 bit setiap 2 aksara	16 variasi warna - (0,0,0) hingga (1,2,0) Rujuk Jadual 2.10	• Menukarkan SM ke bentuk perenam belasan (4 bait) dan mewakilkan nilai tersebut dalam bentuk RGB (0,0,0) hingga (0,2,2). • RGB digunakan untuk memformatkan atribut warna aksara dan garis bawah • Penyembunyian secara jujukan.	Rendah
3.	Wang dan Li	2014	32 bit setiap aksara	Untuk garis bawah (0,0,0) hingga (255,255,255) Untuk aksara (0,0,0 hingga (15,15,15)	• Menyembunyi 4 bait pada satu aksara. ▪ 3 bait pada garis bawah aksara iaitu R(8),G(8),B(8) ▪ 1 bait pada atribut warna aksara - 2 bit pada R - 2 bit pada G - 4 bit pada B • Penyembunyian secara jujukan.	Rendah

4.	Singh et al.	2014	4 bit setiap aksara	16 variasi warna- (0,0,0), (1,1,1), (2,2,2), (3,3,3) hingga (15,15,15)	• Warna RGB mewakili nilai kekunci. Cth RGB(6,6,6) dipetakan dengan bit 0110 • Aksara pada lokasi terpilih akan diformatkan dengan warna RGB menggunakan RGB(0,0,0) hingga RGB(15,15,15) • Penyembunyian secara jujukan.	Rendah
5.	Stojanov et al.	2014	8 bit setiap aksara	Teknik garis bawah menggunakan 16 variasi warna putih.	• Memformatkan garis bawah aksara dengan warna RGB. • Menggunakan 16 variasi warna RGB (putih). • 1 aksara menyembunyikan 8 bit • Penyembunyian secara jujukan.	Rendah
6.	Stojanov et al .	2014	7 bit setiap sempadan perenggan	Menyembunyikan mesej menggunakan nilai RGB > 249.	• Menggunakan 4 sempadan perenggan untuk menyembunyikan mesej	Rendah
7.	Al-Asadi & Bhaya	2016	2 bit setiap aksara	Julat nilai hitam (0,0,0) dan (1,1,1)	• Menggunakan gaya tulisan dan warna aksara. • Memformatkan setiap aksara terpilih dengan warna aksara (0 atau 1) dan gaya tulisan. • Gaya tulisan Arial dan San Serif mewakili bit 1 dan 0 masing-masing.	Rendah
8.	Malik, Sikka, & Verma	2016	10 bits setiap aksara	Menggunakan 32 warna yang berbeza	• Pemampatan LZW & Pengekodan Warna • Memformatkan garis sempadan aksara dengan 5 bit dan warna aksara dengan 5 bit. • Penyembunyian secara jujukan.	Tinggi

9.	Malik et al.	2017	1 bit setiap aksara	8 warna yang berbeza Rujuk Rajah 2.20	• Pemampatan Huffman & Pengekodan Warna • Memformatkan aksara dengan 8 jenis warna yang berbeza untuk mewakili bit 0 dan 1 - 4 warna untuk bit 0 - 4 warna untuk bit 0 • Penyembunyian secara jujukan.	Tinggi
10.	Al-Azzawi	2018	12 bit setiap aksara	Julat warna (0,0,0) hingga (15,15,15)	• 12 bit disembunyikan pada satu aksara. • Perkataan dipilih berdasarkan kepada tag *Part-of-Speech* (POS) • Penyembunyian bergantung kepada bilangan lapisan (kekunci) yang digunakan. • Penyembunyian secara jujukan.	Rendah

2.9 Analisis Saiz Teks Pelindung, Mesej Rahsia dan Kapasiti Kajian Lepas

Jadual 2.12 menunjukkan saiz mesej rahsia, teks pelindung dan kapasiti penyembunyian kajian lepas menggunakan teknik RGB dan sebaliknya. Berdasarkan kepada Jadual 2.12, dapat disimpulkan bahawa saiz mesej rahsia yang digunakan dalam kajian lepas adalah bersaiz antara 26 hingga 2154 aksara, manakala saiz aksara teks pelindung ialah antara 202 dan 4000 aksara.

Jadual 2.12

Saiz Mesej Rahsia, Teks Pelindung dan Kapasiti Penyembunyian

Penyelidik	Saiz Mesej Rahsia (aksara)	Saiz Teks Pelindung (aksara)	Kapasiti %	Berasaskan Warna RGB
(Khairullah, 2009)	-	-	Tidak dinyatakan	Ya
(Tang & Chen, 2013)	43	336	50.00	Ya
(Stojanov et al., 2014)	2154	2252	95.60	Ya
(Singh & Diwakar, 2014)	26	220	11.81	Ya
(Wang & Li, 2014)	198	847	93.5	Ya
(Al-Asadi & Bhaya, 2016)	2 bit	7	3.57	Ya
(Malik et al., 2016)	198	847	13.43	Ya
(Kumar et al., 2016)	198	847	7.21	Tidak
(Malik et al., 2017)	198	847	18.34	Ya
(Ramakrishnan et al., 2017)	500	1779	28.11	Tidak
(Kouser et al., 2017)	800	2640	23.25	Tidak
(Fateh & Rezvani, 2018)	198	847	10.6	Tidak
(Al-Azzawi A. F., 2018)	34	202	25.5	Ya
(Khairullah, 2019)	315	4000	7.88	Tidak
(Baawi et al., 2019)	198	847	12.02	Tidak

Berdasarkan kepada Jadual 2.12, saiz fail mesej rahsia terendah yang digunakan ialah bersaiz 2 bit dengan teks pelindung bersaiz 7 aksara yang dijalankan oleh penyelidik Al-Asadi & Bhaya (2016). Namun begitu, saiz tersebut tidak digunakan dalam kajian ini memandangkan saiznya yang agak kecil. Kapasiti penyembunyian tertinggi berasaskan warna RGB diperoleh menggunakan teknik yang dicadangkan oleh Stojanov et al. (2014) dan diikuti dengan kajian yang dijalankan oleh (Tang & Chen, 2013) dengan kapasiti penyembunyian masing-masing ialah 95.6% dan 50.0%. Wang dan Li, (2014), dapat menyembunyikan mesej rahsia sehingga 4 kali ganda saiz teks pelindung disebabkan satu aksara teks pelindung dapat menyembunyikan 4 bait aksara mesej rahsia (3 bait untuk format garis bawah dan 1 bait untuk format warna aksara). Walau bagaimanapun teks stego yang dihasilkan jelas menunjukkan garis bawah pada setiap huruf yang diformatkan.

Kajian terkini yang dijalankan oleh Al-Azzawi (2018) menggunakan warna RGB hanya dapat menyembunyikan 25.5% sahaja mesej rahsia berbanding kajian yang dilakukan oleh Tang & Chen (2013) dan Stojanov et al. (2014). Berdasarkan kepada Jadual 2.12, secara keseluruhannya, pelbagai saiz mesej rahsia dan teks pelindung digunakan oleh penyelidik yang berbeza untuk menentukan kapasiti mesej tersembunyi.

Rajah 2.25 menunjukkan kapasiti penyembunyian menggunakan teknik berasaskan warna aksara RGB yang dilakukan oleh beberapa penyelidik.

Rajah 2.25 Kapasiti Penyembunyian Menggunakan Teknik Warna Aksara RGB

Rajah 2.25 menunjukkan kapasiti penyembunyian yang dihasilkan oleh Al-Asadi dan Bhaya (2016) paling rendah berbanding teknik lain dengan kapasiti penyembunyian hanya 3.57%. Manakala dua teknik penyembunyian mempunyai kapasiti penyembunyian melebihi 50% iaitu teknik Tang dan Chen (2013) dan Stojanov et al. (2014) dengan kapasiti penyembunyian masing-masing ialah 50% dan 95.6%. Walaupun teknik yang digunakan oleh Stojanov et al. (2014) boleh menyembunyikan 95.6% mesej rahsia dalam teks pelindung, namun teknik ini tidak teguh seperti yang dinyatakan oleh penyelidik dan Baawi et al. (2018) kerana garisan aksara akan dapat dilihat dengan menukar latar belakang dokumen.

2.10 Perwakilan Aksara Berulang Dengan Nilai Rawak

Aksara berulang dalam mesej rahsia akan meningkat apabila saiz mesej rahsia semakin bertambah. Kajian lepas mewakilkan aksara berulang mesej rahsia dengan satu nilai yang sama serta membentuk satu perwakilan statik. Perwakilan aksara berulang mesej rahsia dengan nilai yang berbeza boleh dilakukan menggunakan

Jadual sifer *Homophonic*. *Homophonic* sifer adalah berbeza dengan perwakilan sifer *Monoalphabetic* seperti yang dijelaskan pada sub-topik 2.10.1 dan 2.10.2.

2.10.1 Sifer *Monoalphabetic*

Teknik ini mewakilkan aksara dengan pemetaan satu-ke-satu (*one-to-one)* di mana setiap aksara diwakilkan dengan satu simbol sahaja. Menurut Patel dan Patro (2017), perwakilan sifer *Monoalphabetic* mudah di pecah oleh penceroboh kerana teknik penggantian yang digunakan adalah tetap (perwakilan *satu-ke-satu)* yang menggambarkan frekuensi data bagi aksara asal.

Contoh seperti di bawah.

Sumber	=	abcdefghijklmnopqrstuvwxyz
Kekunci	=	XNYAHPOGZQWBTSFLRCVMUEKJDI
Mesej Rahsia	=	thiscourserockstheblock
Perwakilan *Monoalphabethic*	=	MGZVYFUCVHCFYWVMGHNBFYW

Berdasarkan kepada contoh di atas, di dapati bahawa aksara mesej rahsia `c,e,h,k,o,r,s` dan `t` merupakan aksara berulang dan telah diwakilkan dengan perwakilan yang sama. Jika perwakilan *Monoalphabetic* ditukar ke nilai ASCII, maka kekerapan aksara berulang dalam mesej rahsia dapat ditunjukkan seperti di bawah.

Aksara Berulang	**c**	**e**	**h**	**k**	**o**	**r**	**s**	**t**
Kekerapan	3	2	2	2	3	2	3	2
Perwakilan *Monoalphabetic*	Y	H	G	W	F	C	V	M
Perwakilan ASCII	89	72	71	87	70	67	86	77

Berdasarkan kepada contoh di atas, kesemua aksara berulang diwakilkan dengan nilai yang sama. Sebagai contoh, aksara `c, o` dan `s` berulang sebanyak tiga kali dan diwakilkan dengan aksara `Y, F` dan `V` dengan nilai ASCII masing-masing ialah 89, 70 dan 86.

2.10.2 Sifer *Homophonic*

Sifer *Homophonic* berbeza dengan sifer *Monoalphabetic* di mana setiap aksara boleh diwakilkan dengan pelbagai perwakilan serta ia lebih sukar dan kompleks untuk di analisis oleh penceroboh (Patel & Patro, 2017). Teknik ini menggunakan konsep pemetaan *one-to-many* di mana kebaikannya ialah sesuatu aksara mesej rahsia boleh diwakilkan dengan pelbagai penggantian (Dhavare, Low, & Stamp, 2013). Sifer *Homophonic* biasanya digunakan untuk menyulitkan sesuatu mesej rahsia. *American Cryptogram Association* (*ACA)* memperkenalkan ACA *Homophonic* yang mewakilkan setiap aksara dengan pelbagai nilai tetapi dengan bilangan elemen yang sama banyak bagi setiap aksara seperti yang ditunjukkan di dalam Rajah 2.26.

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
08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	01	02	03	04	05	06	07
44	45	46	47	48	49	50	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
68	69	70	71	72	73	74	75	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67
84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	76	77	78	79	80	81	82	83

Rajah 2.26 Perwakilan ACA *Homophonic*

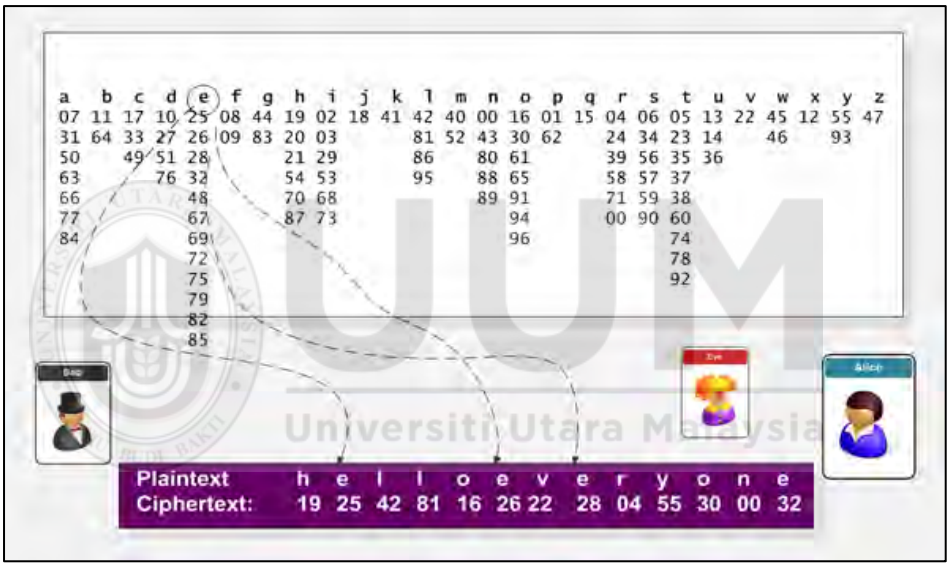
Berdasarkan kepada Rajah 2.26, mesej "MEETYOUATTEN" boleh diwakilkan dengan pelbagai nilai seperti yang ditunjukkan di dalam Jadual 2.13 di bawah.

Jadual 2.13

Kepelbagaian Perwakilan Mesej Rahsia Dengan Nilai Yang Dinamik

Perwakilan	**M E E T**	**Y O U**	**A T**	**T E N**
#1	19 48 12 61	82 32 39	44 01	37 12 96
#2	30 72 72 37	42 21 03	68 61	77 48 55
#3	54 12 72 01	66 21 78	44 37	61 72 31
#4	95 88 48 01	66 97 62	08 77	01 12 20

Berdasarkan Rajah 2.26, terdapat corak perwakilan berturutan yang mewakilkan sesuatu aksara antaranya 08 hingga 25 mewakili aksara A hingga S, 26 hingga 43 mewakili aksara H hingga Z dan seterusnya. Perwakilan seperti ini boleh mendorong steganalisis untuk mengekstrak mesej rahsia dengan mengenal pasti corak perwakilan aksara. Selain daripada perwakilan seperti di atas, perwakilan bagi setiap aksara boleh diwakilkan dengan nilai rawak seperti yang ditunjukkan dalam Rajah 2.27. Teknik ini membentuk satu perwakilan yang lebih bersifat rawak berbanding dengan perwakilan ACA *Homophonic*.



Rajah 2.27 Perwakilan Aksara Dengan Pelbagai Nilai

Rajah 2.27 menunjukkan perwakilan aksara yang boleh diwakilkan dengan nilai yang berbeza kecuali beberapa aksara yang diwakilkan dengan satu perwakilan sahaja. Sebagai contoh, aksara `e` diwakilkan dengan 12 nilai yang berbeza dan diikuti dengan aksara lain yang diwakili lebih dari dua perwakilan kecuali aksara `j,k,q,v,x` dan `z` yang diwakili dengan satu perwakilan sahaja. Rajah 2.30 menunjukkan aksara `e` boleh diwakilkan dengan pelbagai nilai berbeza, antaranya ialah 25, 26, 28, 32, 48, 67 dan seterusnya. Menurut Dhavare, Low, dan Stamp (2013) , teknik *monoalphabetic* memerlukan 4.9 x 10^{12} tahun untuk memecahkan kod berbanding teknik *Homophonic*

yang memerlukan 1.59 x 10^{74} tahun untuk memecahkan 62 aksara jika Jadual *Homophonic* mempunyai 100 perwakilan.

Kesimpulannya, berdasarkan kepada Rajah 2.26, satu Jadual *Homophonic* boleh dijana berdasarkan kepada lokasi aksara di dalam sesuatu dokumen. Penjanaan jadual ini dapat menjana pelbagai set mesej rahsia dengan pelbagai perwakilan.

2.11 Kaedah Pembahagian Aritmetik

Kaedah Pembahagian Matematik merupakan satu kaedah matematik yang dilaksanakan dengan membahagikan satu nombor dengan satu nilai di mana hasilnya dikenali sebagai *Quotient* dan *Remainder*. Kaedah ini telah digunakan oleh Jassim (2013) dan Chandini dan Ganesh Kumar (2018) untuk menukarkan mesej rahsia (ASCII) ke bentuk perwakilan *Divisor(D)*, *Quotient(Q)* dan *Remainder(R)* dan masing-masing menggunakan medium imej dan video sebagai teks pelindung. Perwakilan nilai ASCII mesej rahsia ditukar ke bentuk perwakilan *D* ,*Q* dan *R* menggunakan persamaan berikut:

$$Divisor * Quotient + Remainder = Message \quad (2.6)$$

Chandini dan Ganesh Kumar (2018), menyembunyikan nilai *D*, *Q* dan *R* yang diperoleh ke dalam bingkai (*frame*) video menggunakan teknik Padanan Corak Piksel (*Matching Pattern Riel*). Teknik tersebut menyembunyikan setiap 2 bit mesej tersembunyi ke dalam warna RGB pada bingkai lokasi yang dipilih secara rawak.

2.12 Penjanaan Nombor Rawak

Nombor rawak memainkan peranan penting di dalam penyulitan terutama di dalam pelbagai aplikasi keselamatan berasaskan rangkaian (Elmahi et al., 2017). Penjana Nombor Rawak (*Random Number Generator-RNG*) boleh dibahagikan kepada tiga jenis iaitu Penjanaan Nombor Rawak Sebenar (*True Random Number Generator-TRNG*), Penjana Nombor Pseudorawak (*Pseudorandom Number Generators-PRNG*) dan Fungsi Nombor Pseudorawak (*Pseudorandom Number Function-PNF)*. Penjana Nombor rawak jenis pertama, *TRNG* menghasilkan output yang tidak boleh dihasilkan semula. *TRNG* adalah berdasarkan kepada eksperimen fizikal seperti melambung duit syiling beberapa kali dan hasilannya direkodkan dalam bentuk bit binari. Nombor Pseudorawak telah digunakan di dalam steganografi imej untuk mendapatkan lokasi piksel secara rawak seperti yang dijalankan oleh Bailey dan Curran, (2006). Srikanth, Mehta, Yadav, Singh, dan Singhal (2017) menggunakan pendekatan nombor pseudorawak untuk menjana kekunci bagi proses penyulitan. Nombor rawak boleh digunakan untuk menyembunyikan aksara mesej rahsia pada lokasi rawak seperti di dalam kajian yang dilakukan oleh Elmahi et al. (2017) untuk mendapatkan nombor rawak bagi menyembunyikan mesej rahsia di dalam teks pelindung. Antara penyelidik lain yang menggunakan nilai rawak di dalam kajian mereka ialah Rahman et al. (2017) yang menggunakan nombor pseudorawak untuk mendapatkan lokasi ruang kosong yang terdapat di dalam teks pelindung. Walaupun dapat menyembunyikan mesej dengan sempurna, namun teknik tersebut memerlukan saiz teks pelindung yang besar kerana satu ruang kosong hanya dapat menyembunyikan satu bit mesej rahsia di samping ia bergantung pada jumlah ruang kosong terhad yang terdapat di dalam teks pelindung. Menurut Stephen, Reddy, Naidu, Sonali, dan Heymaraju (2012),

meletakkan aksara pada kedudukan rawak adalah idea yang lebih baik berbanding meletakkan aksara dalam kedudukan yang diinginkan untuk menjadikannya lebih sukar bagi pihak yang tidak bertanggungjawab untuk memecahkan mesej yang disembunyikan dan akan menjadikan mesej tersebut lebih selamat.

2.13 Ringkasan Teknik, Kelebihan dan Limitasi Steganografi Teks

Secara amnya, teknik, kelebihan dan limitasi daripada ketiga-tiga kategori steganografi telah diringkaskan seperti yang ditunjukkan dalam Jadual 2.14.

Jadual 2.14

Kaedah, Kelebihan dan Kelemahan Steganografi Mengikut Kategori

Kategori	Kaedah	Kelebihan	Kelemahan/Kekurangan
Berasaskan Format	• Ruang antara baris • Ruang antara perkataan (satu ruang kosong, ruang kosong berganda). • Ruang di hujung perenggan • Ruang tab. • Hibrid (antara perkataan+ antara perenggan) • Hibrid (ruang antara perkataan+ penjajaran kanan teks. • Perubahan saiz fon. • Kesalahan ejaan disengajakan.	• Kurang mencederakan teks pelindung. • Teknik yang ringkas dan mudah. • Sukar untuk dikesan oleh sistem visual manusia (jika menggunakan teknik lain)	• Mesej rahsia yang boleh disembunyikan ke dalam teks pelindung adalah terhad. • Mesej rahsia akan dihapuskan apabila berlaku proses pemampatan dan penyahmampatan. • Stego teks yang dijana boleh menipu mata manusia tetapi tidak boleh menipu sistem komputer.
Penjanaan Rawak & Statistik	• Berdasarkan jujukan aksara atau jujukan perkataan. • Menganggarkan beberapa taburan statistik yang terdapat dalam teks sebenar • Beberapa teknik yang digunakan adalah jujukan aksara, jujukan perkataan urutan, penjanaan statistik bagi peniruan jujukan-teks. • Kebarangkalian tatabahasa bebas-kontek (PCFG) • *Markov Chains*	• Menjana teks stego mengikut ciri-ciri statistik sesuatu teks. • Sukar untuk dikesan oleh mata manusia dan sistem komputer (jika secara rawak)	• Mencurigakan teks stego yang dijana jika menggunakan teknik penggantian kerana terdapat perubahan pada teks pelindung.
Kaedah Linguistik	• Melibatkan tanda baca seperti koma dan noktah diletakkan di tempat yang	• Sukar untuk dikesan oleh mata manusia dan sistem komputer (jika semantik	• Memerlukan banyak pengetahuan mengenai sintaks dan semantik

	betul dalam dokumen itu. • Kaedah semantik akan menggantikan perkataan sinonim.	adalah betul)	sesuatu bahasa. • Penggunaan kaedah sintaksis dan semantik umumnya akan mengubah teks pelindung. • Mencurigakan teks stego yang dijana kerana terdapat perubahan pada teks pelindung • Kapasiti untuk menyimpan mesej rahsia yang terhad

Menurut Mandal (2012) steganalisis merupakan proses mengenal pasti atau mengesan steganografi dengan memeriksa pelbagai parameter di dalam teks stego. Matlamat steganalisis adalah untuk mengenal pasti aliran maklumat yang disyaki, menentukan sama ada wujud atau tidak mesej rahsia yang dikodkan di dalam aliran maklumat tersebut, dan jika boleh ia akan cuba mengekstrak mesej rahsia yang terdapat di dalam aliran maklumat tersebut. Kajian yang dilakukan oleh Bhattacharyya, et al.,(2011) menjelaskan bahawa mesej yang disembunyikan di dalam medium teks adalah sukar untuk diekstrakkan oleh penceroboh berbanding dengan medium lain.

2.14 Ringkasan

Secara ringkas, kapasiti merujuk kepada jumlah bit data yang boleh disembunyikan di dalam teks pelindung. Manakala ketakbolehkelihatan adalah berkaitan dengan keupayaan untuk mendedahkan maklumat yang tersembunyi teks stego dan keteguhan pula menekankan kepada kemungkinan untuk mengubahsuai atau memusnahkan data tersembunyi. Kajian lepas menunjukkan bahawa terdapat beberapa limitasi dalam ketiga-tiga kategori steganografi seperti yang ditunjukkan di dalam Jadual 2.14. Kapasiti penyembunyian yang rendah, lokasi secara jujukan, kecurigaan teks stego

yang dijana merupakan limitasi utama yang telah dikenal pasti di dalam kajian-kajian lepas dan selari dengan objektif yang digariskan dalam kajian ini. Oleh itu kajian yang dicadangkan ini memperkenalkan perwakilan aksara dalam bentuk 3D bagi menghasilkan kepelbagaian nilai terutama bagi setiap aksara berulang mesej rahsia supaya ia lebih dinamik serta meningkatkan kapasiti mesej rahsia di samping memelihara ketakbolehkelihatan dan keteguhan teks stego yang dihasilkan.

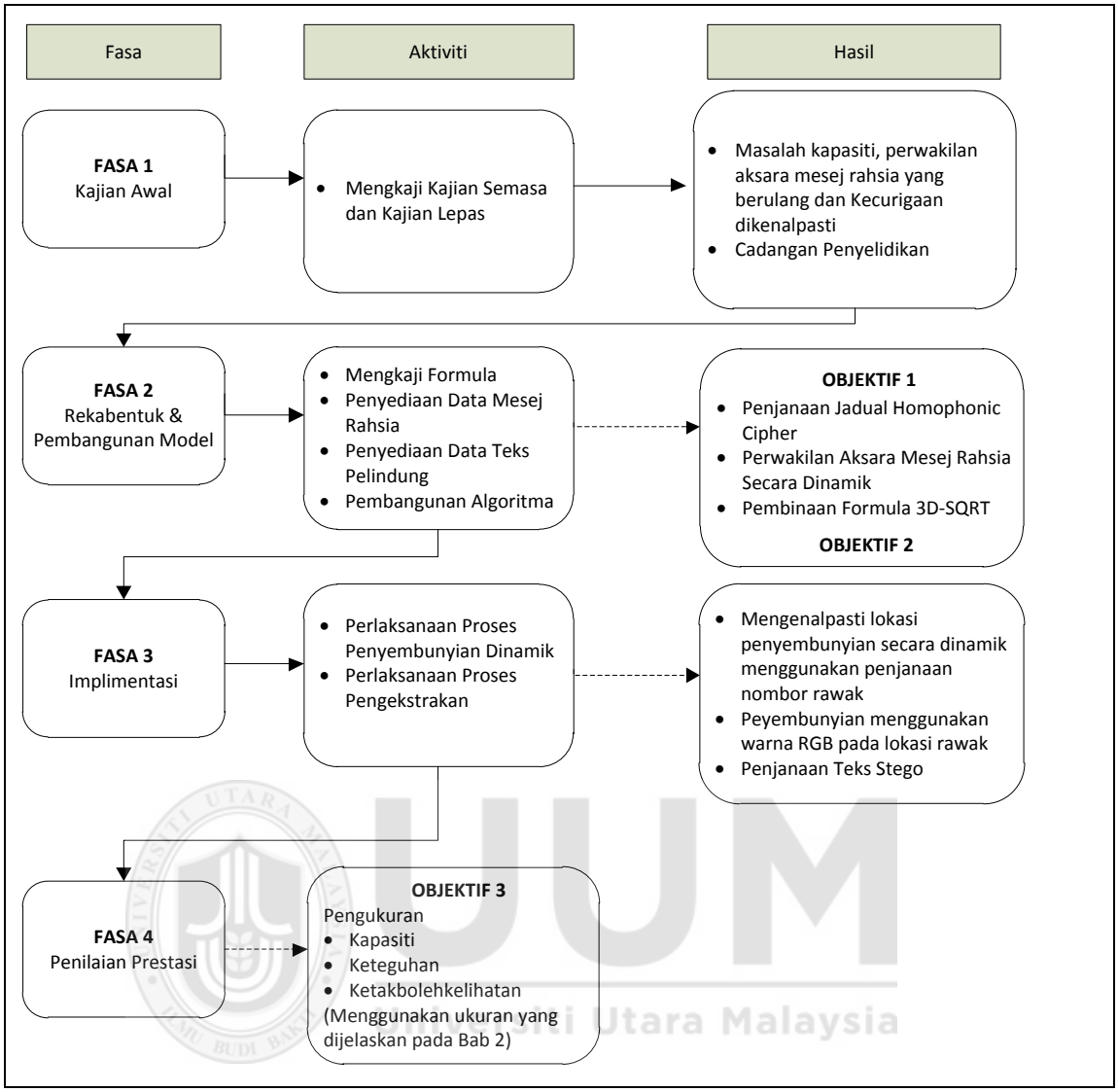
Menurut Ray et al. (2012), penyembunyian bit mesej rahsia pada lokasi berturutan akan memudahkan penceroboh untuk mengekstrak semula mesej yang disembunyikan di dalam teks pelindung. Oleh itu, memanipulasikan abjad berdasarkan kepada lokasi rawak digunakan di dalam kajian ini untuk menghasilkan kepelbagaian perwakilan di samping memastikan tiada perubahan ketara berlaku terhadap teks stego yang dijana bagi mengelakkan sebarang kecurigaan.

BAB TIGA
METODOLOGI KAJIAN

Bab ini menerangkan kaedah yang digunakan untuk menjalankan kajian yang dicadangkan. Pendekatan yang digunakan di dalam kajian ini adalah berdasarkan penyelidikan eksperimental yang merangkumi empat fasa iaitu Kajian Awalan, Reka bentuk dan Pembangunan Model, Implementasi dan Penilaian Prestasi. Rangka kerja penyelidikan yang merangkumi keempat-empat fasa tersebut dibincangkan secara terperinci dalam bab ini bagi menjelaskan aliran kajian untuk mencapai objektif yang telah ditetapkan. Di samping itu, penerangan berkaitan data yang digunakan turut dijelaskan di akhir bab ini.

3.1 Kerangka Metodologi Kajian

Kajian ini melibatkan empat fasa iaitu Fasa Kajian Awalan, Fasa Reka Bentuk dan Pembangunan Prototaip, Fasa Implementasi dan Fasa Penilaian Prestasi seperti yang ditunjukkan di dalam Rajah 3.1. Output yang diperolehi bagi setiap fasa merupakan objektif penyelidikan yang telah ditetapkan di dalam kajian ini dan diterangkan dengan terperinci di bahagian berikutnya.



Rajah 3.1 Rangka Kerja Kajian

3.1.1 Fasa 1 : Kajian Awalan

Fasa ini melibatkan kajian terhadap penyelidikan terdahulu dan semasa melalui penilaian kritikal terhadap artikel, prosiding, journal dan sumber-sumber akademik yang lain. Fasa ini merangkumi pemahaman mendalam terhadap proses steganografi dan isu yang dihadapi terhadap teks stego yang dijana atau dihasilkan. Tujuan fasa ini adalah untuk mengenal pasti ruang yang berpotensi untuk dibuat penyelidikan. Hasil

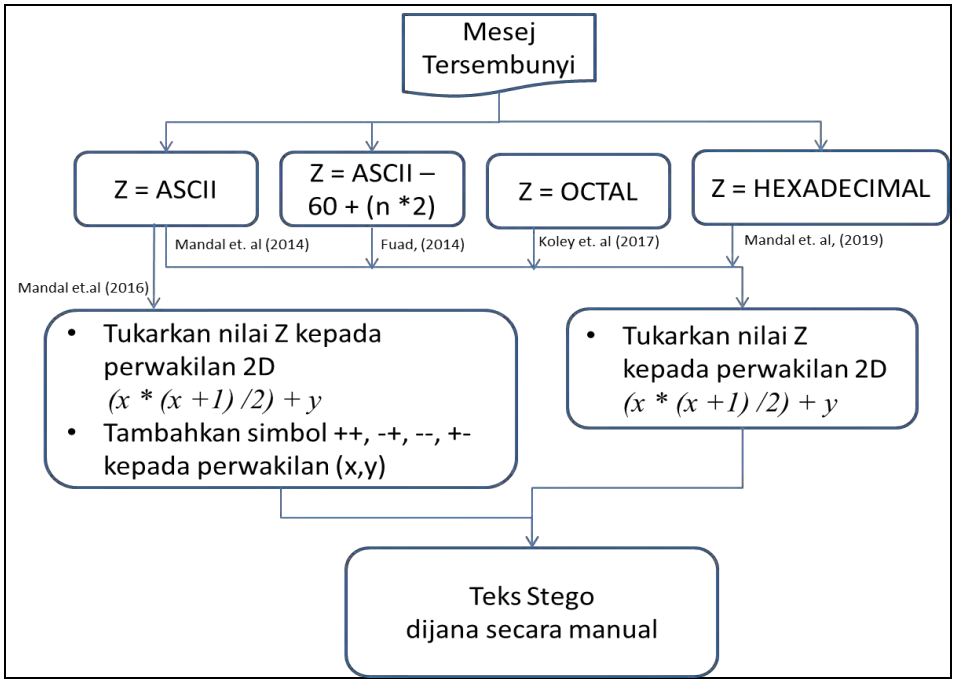
dapatan daripada fasa ini ialah pembentukan penyataan masalah, persoalan kajian, objektif kajian dan seterusnya pembentukan kepada cadangan penyelidikan ini.

3.1.2 Fasa 2 : Reka Bentuk dan Pembangunan Model

Fasa Reka bentuk dan Pembangunan Model merangkumi analisis terhadap Model Dua Dimensi (2D) dan penerangan terperinci mengenai proses reka bentuk pembangunan Model Tiga Dimensi (3D) yang menerbitkan formula perwakilan aksara dalam bentuk (*x,y,z*). Jadual *Homophonic* sifer dijana berdasarkan kepada data teks pelindung. Proses Pembangunan Model Tiga Dimensi diterangkan pada subtopik 3.1.2.2 Seterusnya algoritma penyembunyian dan pengekstrakan dibangunkan pada fasa ini sebelum proses implementasi dilakukan pada fasa seterusnya.

3.1.2.1 Model Penyembunyian Dua Dimensi (2D)

Model penyembunyian dua dimensi (2D) merupakan model yang menukarkan nilai aksara mesej rahsia kepada perwakilan koordinat (*x,y*) menggunakan formula berdasarkan kepada Model Matematik Sistem Nombor (MMSN). MMSN merupakan model yang menukarkan sesuatu nilai ke bentuk perwakilan koordinat 2D dan seterusnya digunakan untuk penjanaan teks stego. Kajian yang dilakukan oleh Mandal et al. (2014), Fuad (2014), Mandal et al. (2016), Koley et al. (2017) dan Mandal et al. (2019) mewakilkan aksara mesej rahsia dalam bentuk perwakilan 2D yang diwakilkan dengan nilai (x,y). Rajah 3.2 menunjukkan model penjanaan teks stego bagi teknik-teknik di atas.



Rajah 3.2 Model Penyembunyian Menggunakan Teknik Mandal et al. (2014), Fuad (2014), Mandal et al. (2016), Koley et al. (2017) dan Mandal et al. (2019)

Kajian menggunakan MMSN telah dijalankan oleh Mandal et al., (2014) dengan memperkenalkan persmaaan 3.1 dibawah untuk menukarkan nilai ASCII mesej rahsia (z) ke bentuk nilai 2D yang diwakilkan dengan nilai (*x,y*).

$$z = (x * (x+1)) / 2 + y \tag{3.1}$$

Perwakilan nilai (*x,y*) ditentukan berdasarkan kepada jadual Jumlah Isi Kandungan bagi sesuatu kumpulan yang diperkenalkan oleh Fuad, (2014) seperti yang ditunjukkan di dalam Jadual 3.1.

Jadual 3.1

Jadual Jumlah Isi Kandungan

Kumpulan (*n*)	Isi Kandungan	Jumlah Isi Kandungan (*j*)
1	(1)	1
2	(1,2)	3
3	(1,2,3)	6
4	(1,2,3,4)	10
5	(1,2,3,4,5)	15
6	(1,2,3,4,5,6)	21
7	(1,2,3,4,5,6,7)	28
8	(1,2,3,4,5,6,7,8)	36
9	(1,2,3,4,5,6,7,8,9)	45
10	(1,2,3,4,5,6,7,8,9,10)	55
11	(1,2,3,4,5,6,7,8,9,10,11)	66
12	(1,2,3,4,5,6,7,8,9,10,11,12)	78
13	(1,2,3,4,5,6,7,8,9,10,11,12,13)	91
14	(1,2,3,4,5,6,7,8,9,10,11,12,13,14)	105
15	(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15)	120
16	(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16)	136
17	(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17)	153
18	(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18)	171

Berdasarkan kepada, Jadual 3.1, satu persamaan diterbitkan untuk mendapatkan Jumlah Isi Kandungan (*j*) bagi sesuatu kumpulan iaitu :

$$Jumlah\ Isi\ Kandungan\ (j) = \sum_{i=1}^{n} \mathrm{i}$$

di mana

n = nombor kumpulan

Sesuatu nilai dapat ditukarkan ke bentuk (*x,y*) berdasarkan kepada Jadual 3.1 di atas. Contoh, perwakilan nilai ASCII, *z* boleh diperolehi berdasarkan kepada langkah berikut:

Langkah 1 : (Nilai *x*) Kenal pasti kumpulan (*n*) di mana nilai $z \geq j$ dengan nilai perbezaan yang paling minimum. Nilai *x = kumpulan (j)*

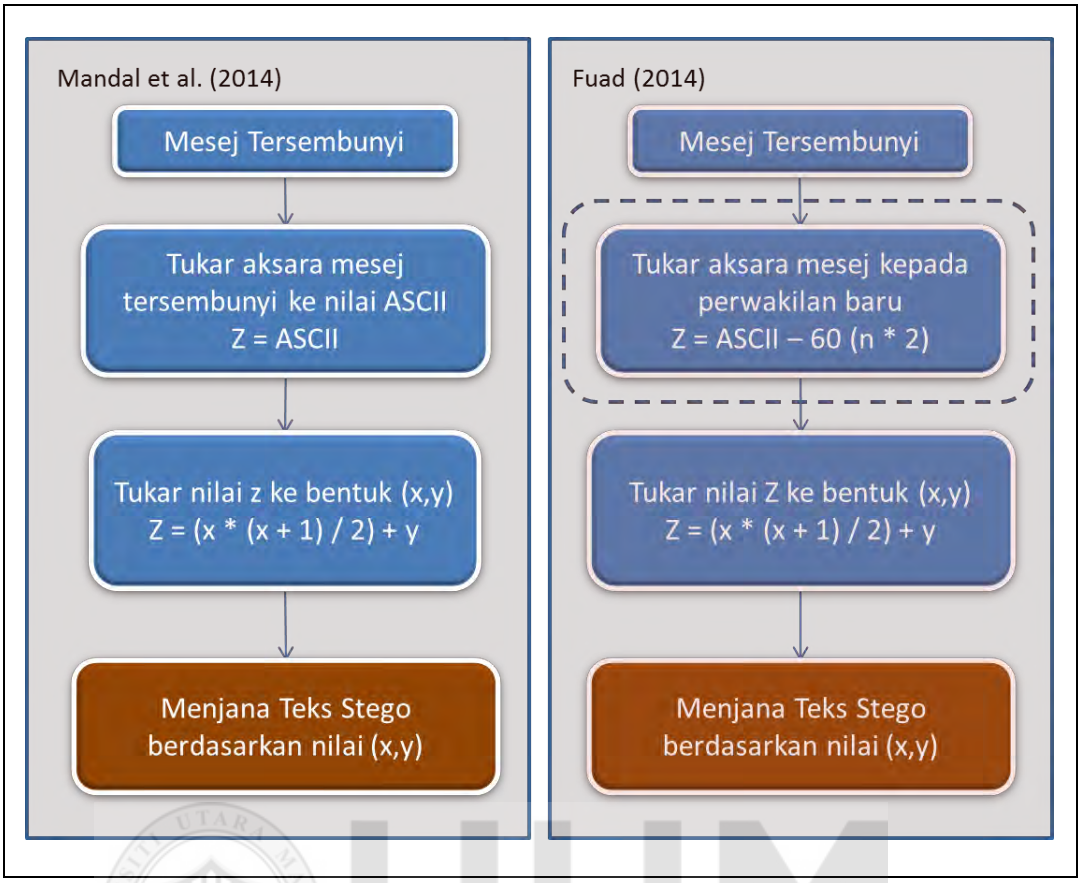
Langkah 2 : (Nilai *y*) Kira perbezaan nilai *z* dan jumlah isi kandungan *j, (y = z – j)*

Contoh : Perwakilan bagi nilai $z = 69$

Langkah 1 : Kumpulan 11 (69 > 66 dgn perbezaan yang paling minimum)
(Nilai x) $x = 11$

Langkah 2 : 3 ($y = 69 - 66$)
(Nilai y)

Oleh itu, nilai 69 dapat diwakilkan dengan nilai (*11,3*) dan boleh ditukar semula ke nilai asal berdasarkan kepada persamaan 3.1 iaitu ($x * (x+1)/2) + y$) = (11 * 12)/2 + 3) = 69. Fuad, (2014) telah menambah baik kajian yang dijalankan oleh Mandal et al. (2014) dengan melakukan perubahan terhadap nilai z untuk mewakilkan nilai baharu ASCII mesej rahsia sebelum ditukarkan ke bentuk koordinat 2D menggunakan persamaan 3.1. Kajian yang dilakukan oleh Fuad (2014) menukarkan nilai ASCII mesej rahsia kepada satu nilai baharu berdasarkan persamaan z = ASCII – 60 + (n * 2) sebelum menukarkan nilai tersebut kepada bentuk perwakilan (x,y) seperti yang ditunjukkan di dalam Jadual 3.2. Rajah 3.3 menunjukkan model pertukaran nilai ASCII yang digunakan oleh kedua-dua penyelidik tersebut.



Rajah 3.3 Perwakilan nilai (*x,y*) menggunakan teknik Mandal et al., (2014) dan Fuad, (2014)

Walaupun kedua-dua model tersebut berjaya menghasilkan koordinat 2D yang berbeza seperti yang ditunjukkan dalam Jadual 3.2, sebaliknya teknik ini menghasilkan nilai (*x,y*) yang sama (statik) bagi aksara mesej rahsia yang berulang seperti yang ditunjukkan dalam Jadual 3.4. Selain itu, kedua-dua kajian ini juga menghasilkan perwakilan julat 2D yang terhad kepada 26 koordinat sahaja seperti yang ditunjukkan dalam Jadual 3.2. Berdasarkan kepada Jadual 3.2, julat perwakilan (*x,y*) bagi teknik Mandal et. al., (2014) ialah (10,10), (10,11), (11,1) … (11,12), (12,1) …. (12,12) manakala julat perwakilan bagi Fuad (2014) pula ialah (02,02) hingga (12,02).

Jadual 3.2

Perwakilan Nilai (x,y) Menggunakan Teknik Mandal,(2014) dan Fuad,(2014)

	Teknik #1 Mandal et al., (2014)			Teknik #2 Fuad, (2014)	
Aksara	Nilai ASCII Z = ASCII	Perwakilan (*x,y*)	n	Perwakilan Baru Nilai ASCII Z=ASCII – 60 + (n *2)	Perwakilan (*x,y*)
A	65	(10,10)	0	65 - 60 + (0 * 2) = 5	(02,02)
B	66	(10,11)	1	66 – 60 + (1 * 2) = 8	(03,02)
C	67	(11,01)	2	67 – 60 + (2 * 2) = 11	(04,01)
D	68	(11,02)	3	68 – 60 + (3 * 2) = 14	(04,04)
E	69	(11,03)	4	69 – 60 + (4 * 2) = 17	(05,02)
F	70	(11,04)	5	70 – 60 + (5 * 2) = 20	(05,05)
G	71	(11,05)	6	71 – 60 + (6 * 2) = 23	(06,02)
H	72	(11,06)	7	72 – 60 + (7 * 2) = 26	(06,05)
I	73	(11,07)	8	73 – 60 + (8 * 2) = 29	(07,01)
J	74	(11,08)	9	74 – 60 + (9 * 2) = 32	(07,04)
K	75	(11,09)	10	75 – 60 + (10 * 2) = 35	(07,08)
L	76	(11,10)	11	76 – 60 + (11 * 2) = 38	(08,02)
M	77	(11,11)	12	77 – 60 + (12 * 2) = 41	(08,05)
N	78	(11,12)	13	78 – 60 + (13 * 2) = 44	(08,08)
O	79	(12,01)	14	79 – 60 + (14 * 2) = 47	(09,02)
P	80	(12,02)	15	80 – 60 + (15 * 2) = 50	(09,05)
Q	81	(12,03)	16	81 – 60 + (16 * 2) = 53	(09,08)
R	81	(12,04)	17	82 – 60 + (17 * 2) = 55	(10,00)
S	83	(12,05)	18	83 – 60 + (18 * 2) = 59	(10,04)
T	84	(12,06)	19	84 – 60 + (19 * 2) = 62	(10,07)
U	85	(12,07)	20	85 – 60 + (20 * 2) = 65	(10,10)
V	86	(12,08)	21	86 – 60 + (21 * 2) = 68	(11,02)
W	87	(12,09)	22	87 – 60 + (22 * 2) = 71	(11,05)
X	88	(12,10)	23	88 – 60 + (23 * 2) = 74	(11,08)
Y	89	(12,11)	24	89 – 60 + (24 * 2) = 77	(11,11)
Z	90	(12,12)	25	90 – 60 + (25 * 2) = 80	(12,02)

Koley et al., (2017), menambah baik kajian yang dijalankan sebelumnya dengan menukarkan nilai ASCII mesej rahsia (*z*) kepada nilai OCTAL dan seterusnya menukarkan kepada perwakilan (*x,y*) menggunakan persamaan 3.1. Jadual 3.3 menunjukkan perwakilan baru yang dihasilkan menggunakan nilai *OCTAL*.

Jadual 3.3

Perwakilan Nilai (x,y) Menggunakan Teknik Koley et al., (2017)

	Teknik #3 Koley et al., (2017)		
Aksara	Nilai ASCII	Nilai OCTAL (Z)	Perwakilan (*x,y*)
A	65	101	(13,10)
B	66	102	(13,11)
C	67	103	(13,12)
D	68	104	(13,13)
E	69	105	(14,00)
F	70	106	(14,01)
G	71	107	(14,02)
H	72	108	(14,03)
I	73	109	(14,04)
J	74	110	(14,05)
K	75	111	(14,06)
L	76	112	(14,07)
M	77	113	(14,08)
N	78	114	(14,09)
O	79	115	(14,10)
P	80	116	(14,11)
Q	81	117	(14,12)
R	81	118	(14,13)
S	83	119	(14,14)
T	84	120	(15,00)
U	85	121	(15,01)
V	86	122	(15,02)
W	87	123	(15,03)
X	88	124	(15,04)
Y	89	125	(15,05)
Z	90	126	(15,06)
a	97	141	(16,05)
b	98	142	(16,06)
i	105	149	(16,07)
m	109	153	(17,00)
t	116	160	(17,07)
z	122	172	(18,01)

Berdasarkan kepada Jadual 3.2 dan Jadual 3.3, mesej rahsia yang mempunyai aksara berulang diwakili oleh nilai (*x,y*) yang sama bagi setiap teknik. Situasi ini dapat dilihat berdasarkan contoh yang ditunjukkan di dalam Jadual 3.4.

Jadual 3.4

Perwakilan Nilai (x,y) Bagi Aksara Berulang

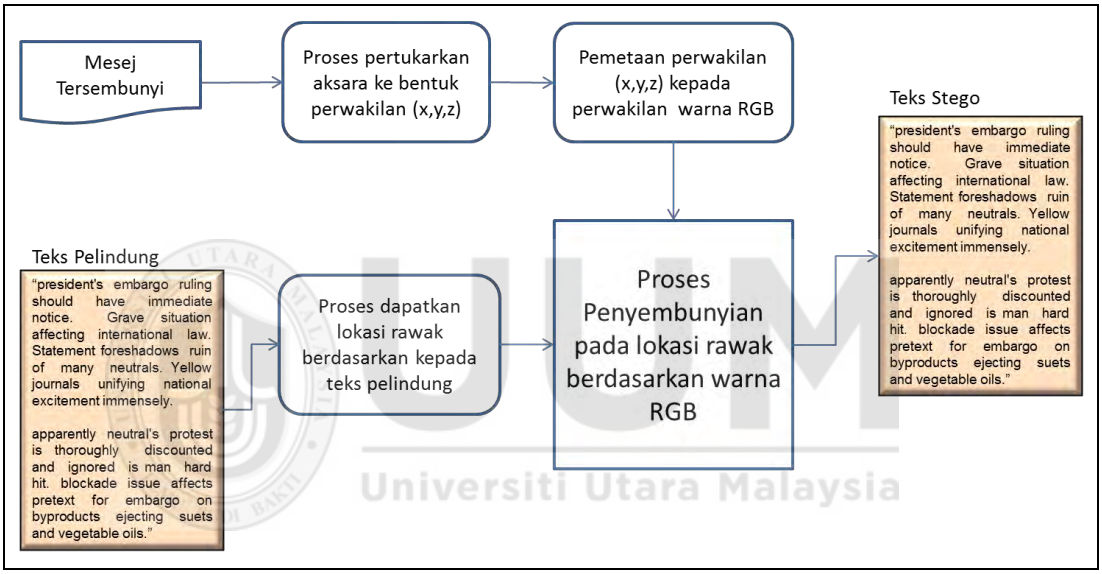
Mesej Rahsia : MEETYOUATTEN			
	Teknik #1 Mandal et al., (2014)	Teknik #2 Fuad, (2014)	Teknik #3 Koley & Mandal, (2017)
M	(11,11)	(08,05)	(14,08)
E	(11,03)	(05,02)	(14,00)
E	(11,03)	(05,02)	(14,00)
T	(12,06)	(10,07)	(15,00)
Y	(12,11)	(11,11)	(15,05)
O	(12,01)	(09,02)	(14,10)
U	(12,07)	(10,10)	(15,01)
A	(10,10)	(02,02)	(13,10)
T	(12,06)	(10,07)	(15,00)
T	(12,06)	(10,07)	(15,00)
E	(11,03)	(05,02)	(14,00)
N	(11,12)	(08,08)	(14,09)

Berdasarkan kepada Jadual 3.4, aksara “E” dan “T” merupakan aksara berulang di dalam mesej rahsia “MEETYOUATTEN”. Ketiga-tiga aksara E dalam mesej rahsia masing-masing diwakili oleh nilai yang sama iaitu (11,03), (05,02), (14,00) bagi Teknik #1, Teknik #2 dan Teknik #3. Hal yang sama juga berlaku bagi aksara berulang “T” yang mempunyai nilai yang sama iaitu (12,06), (10,07) dan (15,00). Perwakilan ini akan memberi kesan terhadap proses penjanaan teks stego yang mempunyai aksara mesej rahsia yang berulang.

3.1.2.2 Model Penyembunyian Tiga Dimensi (3D)

Model penyembunyian 2D dijadikan sebagai panduan untuk proses penghasilan formula dalam bentuk 3D. Penjanaan Jadual *Homophonic* dan penggunaan nombor

rawak digunakan untuk mewakilkan aksara berulang yang terdapat di dalam mesej rahsia agar ia bersifat lebih dinamik. Hasilnya, model ini menukarkan aksara mesej rahsia ke bentuk pelbagai perwakilan (*x,y,z*) bertujuan untuk meningkatkan kapasiti penyembunyian mesej rahsia menggunakan formula SQRT yang telah dicadangkan. Nilai (*x,y,z*) yang dihasilkan dipadankan dengan warna RGB dan disembunyikan pada lokasi rawak yang telah dikenal pasti. Model am bagi proses penyembunyian ditunjukkan di dalam Rajah 3.4.



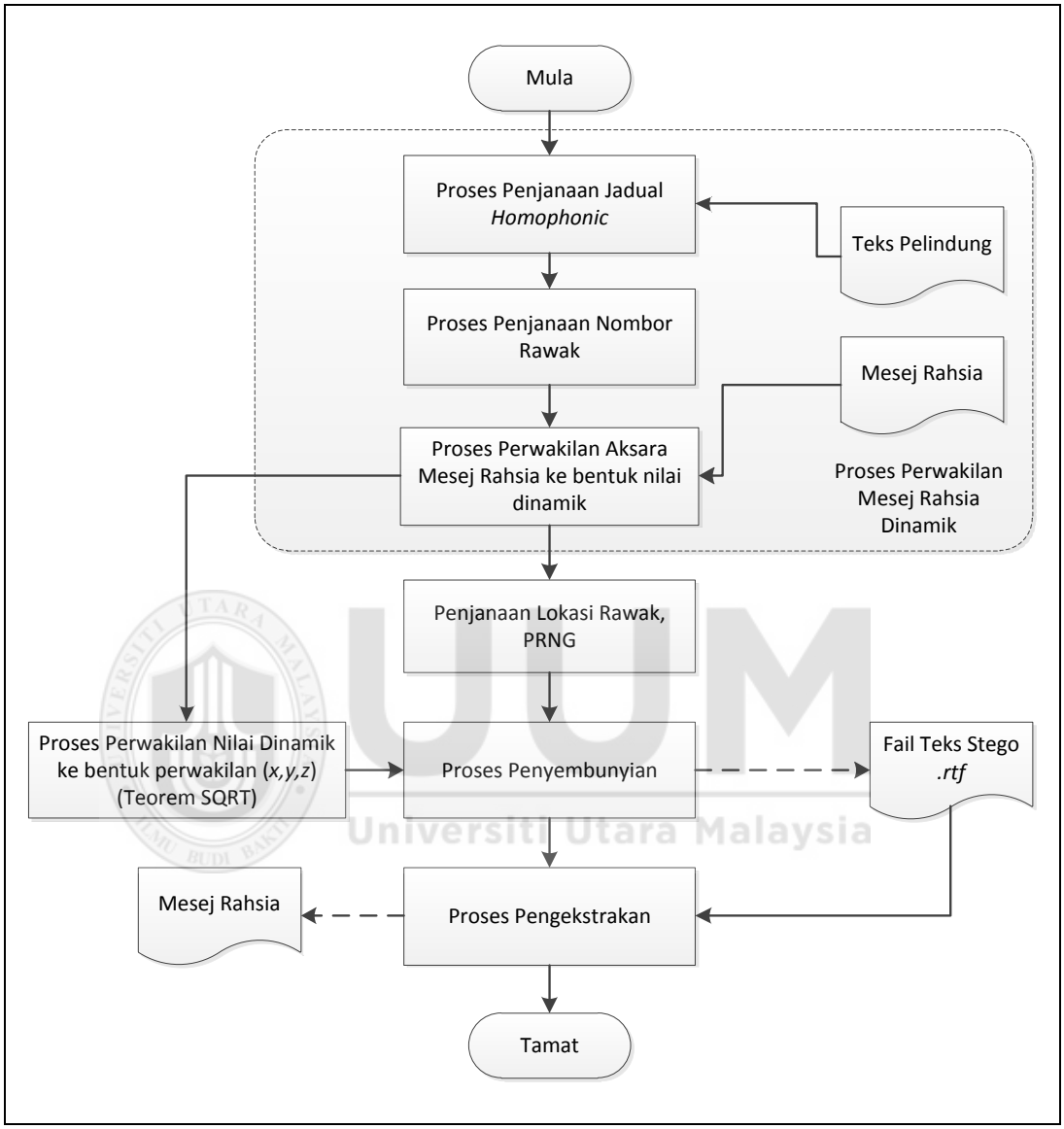
Rajah 3.4 Model Am Proses Penyembunyian

Teks stego merupakan objek akhir yang dihasilkan selepas melalui proses penyembunyian dan akan dinilai prestasinya selepas proses pengekstrakan.

3.1.2.3 Proses Pembangunan Model Tiga Dimensi Steganografi Teks

Proses Pembangunan Model Tiga Dimensi Steganografi Teks melibatkan proses tiga proses utama iaitu proses perwakilan mesej rahsia dinamik, proses penyembunyian dan proses pengekstrakan seperti yang ditunjukkan dalam Rajah 3.5. Proses perwakilan mesej rahsia dinamik melibatkan tiga proses iaitu proses penjanaan jadual

homophonic, proses penjanaan nombor rawak dan proses perwakilan mesej rahsia ke bentuk nilai dinamik.



Rajah 3.5 Proses Reka bentuk Dan Pembangunan Model

Hasil akhir proses perwakilan mesej rahsia dinamik merupakan satu set mesej rahsia yang diwakilkan dalam bentuk perwakilan nilai integer. Proses penyembunyian mesej rahsia dilakukan dengan mendapatkan lokasi penyembunyian secara rawak. Aksara pada lokasi rawak akan diformatkan dengan warna RGB yang telah diwakilkan dengan perwakilan (*x,y,z*). Akhir sekali proses pengekstrakan dilakukan untuk

mengekstrak semula mesej rahsia yang disembunyikan. Keseluruhan proses ini diterangkan dengan terperinci pada Bab 4.

3.1.2.4 Penjanaan Jadual *Homophonic* Sifer

Setiap aksara mesej rahsia sama ada berulang atau sebaliknya perlu diwakilkan dengan pelbagai nilai bagi membolehkan ia bersifat dinamik. Kajian ini telah menjana Jadual *Homphonic* sifer berdasarkan kepada kedudukan aksara di dalam sesuatu dokumen teks pelindung. Jadual 3.5 menunjukkan sebahagian daripada kandungan jadual *Homophonic* sifer yang telah dijana berdasarkan kepada teks pelindung pada Lampiran E.

Jadual 3.5

Jadual Homophonic Sifer

A:
6,19,26,32,47,49,52,57,79,122,125,126,133,143,161,168,189,243,264,268,271,284,295,336,
343,354,362,365,376,420,438,483,488,493,495,532,536,568,578,587,594,601,611,612,622,63
0,689,693,739,743,750,762,769,776,781,793,805,810,815,819,821,823,836,864,906,910,926,
928,930,937,953,962,964,987,1006,1021,1026,1054,1069,1073,1082,1101,1126,1128,1131,113
8,1169,1185,1189,1194,1199,1212,1219,1226,1236,1241,1245,1261,1273,1276,1313,1338,1347
,1363,1378,1386,1407,1433,1436,1444,1450,1459,1501,1512,1548,1578,1580,1589,1606,1615,
1620,1624,1648,1653,1655,1667,1670,1681,1685,1699,1704,1718,1728,1749,1751,1757,1765,1
773,1777,1783,1794,1798,1801,1820,1824,1829,1832,1837,1845,1850,1860,1872,1878,1880,18
87,1893,1904,1913,1915,1919,1926,1933,1948,1963,1972,2015,2024,2027,2040,2050,2053,205
4,2062,2066,2088,2098,2122,2135,2140,2146,2148,2158,2165,2175,2176,2178,2183,2203,220
B:
56,69,100,121,167,172,308,593,676,713,756,835,869,916,952,1042,1091,1102,1109,1175,134
6,1385,1413,1426,1432,1469,1560,1705,1793,1808,1859,1924,1964,2026,2049
C:
13,17,82,118,129,148,236,273,296,301,404,416,429,477,480,501,514,555,563,597,637,655,6
67,670,679,733,772,806,829,843,876,945,996,1051,1086,1116,1141,1235,1238,1246,1247,125
6,1259,1262,1263,1303,1332,1420,1458,1490,1499,1541,1605,1614,1617,1625,1626,1635,1638
,1656,1657,1678,1726,1733,1775,1896,1906,1934,1978,1985,1988,2037,2067,2090,2134,2163
D:
46,128,181,212,269,298,299,329,337,358,363,543,552,621,700,708,741,768,771,792,831,931
,932,934,948,960,971,975,978,1009,1013,1016,1071,1090,1133,1151,1155,1158,1244,1250,12
75,1278,1290,1360,1443,1453,1508,1536,1557,1582,1584,1623,1629,1669,1672,1720,1721,176
2,1805,1847,1876,2017,2094,2102,2160,2167,2170
E:
3,9,15,22,28,31,67,73,81,83,99,105,110,116,130,140,147,164,180,196,205,213,234,237,262
,281,297,300,305,335,350,352,357,361,378,388,414,424,428,433,441,454,469,472,476,479,4
86,491,502,504,527,539,544,551,567,577,584,598,608,620,652,666,669,673,684,696,701,710
,722,727,731,755,760,764,780,785,791,802,807,830,834,839,842,851,887,909,933,940,941,9
44,947,951,969,976,995,1002,1008,1014,1025,1043,1047,1050,1053,1062,1064,1067,1076,108
4,1092,1105,1137,1140,1156,1168,1176,1181,1198,1209,1222,1243,1258,1286,1289,1316,1341
,1345,1359,1362,1384,1431,1442,1454,1483,1485,1488,1493,1507,1521,1526,1530,1540,1544,
1551,1570,1572,1593,1596,1604,1611,1613,1622,1637,1645,1693,1707,1712,1722,1725,1739,1
741,1760,1761,1769,1792,1804,1809,1812,1819,1836,1842,1852,1858,1864,1866,1869,1875,18
91,1895,1900,1918,1925,1950,1960,1968,1977,1980,1984,1987,1993,1997,2012,2020,2033,206
1,2071,2073,2085,2097,2115,2126,2142,2150,2152,2173,2187,2191,2195,2202
F:
78,253,391,437,447,489,508,634,663,752,856,896,992,1017,1038,1044,1060,1080,1160,1192,
1220,1231,1309,1339,1368,1403,1505,1538,1553,1562,1601,1698,1715,1716,1770,1789,1855,1
945,2077,2107,2132,2210,

Berdasarkan Jadual 3.5, aksara „a“ dalam mesej rahsia boleh diwakilkan dengan pelbagai nilai sama ada 32, 365, 930, 2066 atau sebagainya dan begitu juga

dengan aksara-aksara lain. Nilai ini akan ditukarkan ke bentuk perwakilan 3D (*x,y,z*) menggunakan teorem SQRT yang diperkenalkan dan seterusnya dipadankan dengan warna RGB pada aksara teks pelindung lokasi rawak terpilih

3.1.2.5 Penjanaan Nombor Rawak

PRNG merupakan satu pendekatan yang banyak digunakan untuk menjana nombor rawak disebabkan kepantasan serta mempunyai hubungan berulang (*recurrence relation*) di mana nilai baru dijana berdasarkan kepada nilai sebelumnya. PRNG menjana jujukan output yang dikira berdasarkan benih awalan (*initial seeds*) dan seterusnya menerbitkan output berdasarkan kepada formula berikut:

$$S[i+1] = S[i] * A + B \; mod \; m; i = 0,1,2,3, \ldots \quad (3.2)$$

di mana

$$S[i], A, B \quad \in \quad \{0,1,2,3 \ldots \ldots m-1\}$$

m : pemalar Integer

Nombor rawak, *S* yang dijana akan berada di dalam julat berikut:

$$S = \{s_i | 1 \; \leq i \; \leq m\}$$

Nombor Pseudorawak telah digunakan di dalam steganografi imej untuk mendapatkan lokasi piksel secara rawak seperti yang dijalankan oleh Bailey dan Curran, (2006). Srikanth, Mehta, Yadav, Singh, dan Singhal (2017) menggunakan pendekatan nombor pseudorawak untuk menjana kekunci bagi proses penyulitan. Nombor rawak boleh digunakan untuk proses penyembunyian mesej rahsia pada lokasi rawak. Elmahi et al.

(2017) di dalam kajiannya telah menggunakan persamaan 3.2 untuk mendapatkan nombor rawak bagi menyembunyikan mesej rahsia di dalam teks pelindung.

3.1.2.6 Data Kajian

Data kajian merupakan komponen penting dalam penanda aras teknik steganografi. Kajian ini menggunakan set data teks yang merangkumi pelbagai tekstur dan saiz berdasarkan kepada ciri-ciri yang dicadangkan oleh Osman, Din, Zalizam, Muda dan Omar (2013). Set data yang digunakan di dalam kajian ini adalah daripada Reuters-21578 merupakan set data yang diguna pakai oleh beberapa penyelidik (Din & Amphawan, 2015; Rodríguez, Pesado, Merlino, & Martínez, 2018) dan mencakupi kebanyakan kajian-kajian di dalam bidang koleksi pengkategorian teks (Aghdam et al., 2009; Maiti & Samanta, 2010).

Penyelidikan terdahulu menggunakan saiz teks pelindung dan mesej rahsia yang pelbagai di dalam pengujian yang dijalankan. Jadual 3.6 menunjukkan saiz teks pelindung dan mesej rahsia yang digunakan oleh beberapa penyelidik di dalam kajian lepas.

Jadual 3.6

Saiz Teks Pelindung Dan Mesej Rahsia Bagi Kajian Lepas

Penyelidik	Saiz Mesej Rahsia (bait)	Saiz Teks Pelindung (bait)
(Tang & Chen, 2013)	43	336
(Saniei & Faez, 2013)	30,40,50,60,80,100	>3000
(Stojanov et al., 2014)	46	2252
(Singh et al., 2014)	4	38
(Singh & Diwakar, 2014)	26	220
(Wang & Li, 2014)	-	228

(Al-Asadi & Bhaya, 2016)	2 bit	7
(R. Kumar et al., 2016)	198	847
(Malik et al., 2016)	198	847
(Malik et al., 2017)	198	847
(Ramakrishnan et al., 2017)	500	1779
(Ishtiaq et al., 2017)	100,500,1000,5000	9821
(Kouser et al., 2017)	800	2640
(Naqvi et al., 2018)	198	847
(Al-Azzawi, 2018)	34	202

Berdasarkan kepada Jadual 3.6, pelbagai saiz mesej rahsia dan teks pelindung digunakan oleh penyelidik di dalam kajian yang dijalankan. Justeru itu kajian ini memilih beberapa fail daripada set Reuters-21578 yang menghampiri saiz di atas untuk mewakili teks tersembunyi dan teks pelindung untuk tujuan pengujian.

3.1.2.6.1 Data Mesej Rahsia

Kajian ini telah memilih pelbagai saiz mesej rahsia daripada set data Reuters-21578 yang menghampiri saiz mesej rahsia yang digunakan oleh penyelidik lepas seperti yang ditunjukkan di dalam Jadual 3.7.

Jadual 3.7

Kod dan Saiz Mesej Rahsia

Kod SM	Saiz SM (Aksara)
SM1	28
SM2	40
SM3	67
SMe	101
SMe	116
SMe	159
SMe1	166*
SMe2	183*
SM4	267
SMe3	283*
SM5	531
SMe4	660*
SM6	834
SMe5	1262*
SMe6	1958*
SMe7	2185*

Jadual 3.7 di atas menunjukkan 16 saiz mesej rahsia (SM) dalam pelbagai saiz digunakan di dalam kajian ini. Mesej rahsia tambahan (SMe) merupakan mesej rahsia yang mempunyai saiz menghampiri saiz teks pelindung bertujuan untuk menentukan kapasiti maksimum mesej rahsia yang boleh disembunyikan. Menurut Mahajan (2014) kapasiti penyembunyian adalah merujuk kepada jumlah maksimum mesej rahsia yang boleh disembunyikan ke dalam sesuatu teks pelindung. Oleh itu, beberapa mesej tambahan yang menghampiri saiz teks pelindung telah ditambah ke dalam sampel data kajian bertujuan untuk mendapatkan kapasiti penyembunyian maksimum mesej rahsia yang boleh disembunyikan ke dalam sesuatu teks pelindung.

Jadual 3.8

Isi Kandungan Mesej Rahsia

Kod SM	Saiz SM (Aksara)	Mesej Rahsia
SM1	28	ArrivedAtFebruaryTwentySeven
SM2	40	Theimportanceandamountofdatahaveincrease
SM3	67	TheCzechNationalBankbalanceofpaymentfigures forthefirsthalfoftheyear
SMe	101	Fassiosaiditwouldbemoreusefultoputfringecou ntrieslikeSlovakiaintothefoldwheremorepress urecanbeapplied
SMe	116	TheCzechNationalBankbalanceofpaymentfigures forthefirsthalfoftheyearwhichwillgivecleare rpictureofthebalanceofservices
SMe1	166*	Behindusingacovertextistohidethepresenceofs ecretmessagesthepresenceofembeddedmessagesi ntheresultingstegotextcannotbeeasilydiscove redbyanyoneexceptttheintendedrecipient
SMe2	183*	SeveralanalystsslashedratingsonCascadecitin gconcernthatgrowthratesforitscoreframerelay switchingbusinessmaybeslowingtoindustryleve lsofthirtytofiftypercentfromitspriorhundred percentplus

Jadual 3.8 menunjukkan sebahagian daripada kandungan fail mesej rahsia yang digunakan di dalam kajian ini manakala keseluruhan fail yang digunakan dapat dilihat pada Lampiran F dengan saiz maksimum fail mesej rahsia ialah 2185 aksara.

3.1.2.6.2 Data Teks Pelindung

Sebanyak 42 set data terlatih (*trained data*) daripada set data Reuters-21578 yang terdiri daripada pelbagai saiz telah dipilih secara rawak untuk dijadikan sebagai teks pelindung di dalam kajian ini. Fail yang digunakan bersaiz antara 1Kb hingga 4Kb yang menghampiri saiz fail yang digunakan dalam penyelidikan lepas seperti yang ditunjukkan dalam Jadual 3.9.

Jadual 3.9

Kepelbagaian Saiz Teks Pelindung

Nama Fail	Kod Fail	Bil Aksara	Saiz (Kb)	Nama Fail	Kod Fail	Bil Aksara	Saiz (Kb)
289-2015-Abaas*	CT180	180	1	26311newsML	CT1845	1845	3
120683newsML	CT243	243	1	285403newsML	CT1885	1885	3
104778newsML	CT376	376	1	344834newsML	CT2041	2041	3
173005newsML	CT412	412	1	355589newsML	CT2150	2150	3
235003newsML	CT452	452	1	33758newsML	CT2153	2153	3
183674newsML	CT452	452	1	198232newsML	CT2210	2210	3
134683newsML	CT464	464	1	186199newsML	CT2304	2304	3
11270newsML	CT667	667	1	33299newsML	CT2358	2358	3
201116newsML	CT793	793	1	383585newsML	CT2360	2360	3
181034newsML	CT821	821	1	361792newsML	CT2716	2716	3
296368newsML	CT837	837	1	188988newsML	CT2608	2608	4
148786newsML	CT887	887	2	118448newsML	CT2638	2638	4
225609newsML	CT1086	1086	2	241719newsML	CT2796	2796	4
191925newsML	CT1139	1139	2	293661newsML	CT2853	2853	4
174426newsML	CT1311	1311	2	223842newsML	CT2858	2858	4
161715newsML	CT1515	1515	2	133171newsML	CT2900	2900	4
156181newsML	CT1627	1627	2	112125newsML	CT3056	3056	4
231106newsML	CT1669	1669	2	144035newsML	CT3129	3129	4
245782newsML	CT1700	1700	2	250680newsML	CT3267	3267	4
28928newsML	CT1874	1874	2	200749newsML	CT3305	3305	4
146644newsML	CT1962	1962	2				
234712newsML	CT1972	1972	2				
250167newsML	CT2785	2785	2				

3.1.2.7 Pemilihan Mesej Rahsia dan Teks Pelindung

Secara amnya, berdasarkan kepada Jadual 3.6, saiz fail mesej rahsia dan teks pelindung yang digunakan di dalam kajian lepas masing-masing berada di antara julat 26 hingga 800 aksara dan 202 hingga 2640 aksara. Ringkasan saiz kedua-dua fail tersebut ditunjukkan dalam Jadual 3.10.

Jadual 3.10

Saiz Mesej Rahsia dan Teks Pelindung yang Digunakan

Kajian Lepas		**Kajian Yang Dijalankan**	
Saiz Mesej Rahsia	Saiz Teks Pelindung	Saiz Mesej Rahsia	Saiz Teks Pelindung
26	220	28	180
32	304	40	243
43	336	67	376
46	2252	67	2153
198	847	267	837
500	1779	531	1700
800	2640	834	2638

Justeru itu, kajian ini menggunakan fail teks tersembunyi dan teks pelindung yang hampir sama seperti dengan kajian lepas dengan masing-masing bersaiz antara 28 hingga 834 dan 180 hingga 2638 aksara seperti yang ditunjukkan di dalam Jadual 3.10. Walau bagaimanapun, beberapa mesej rahsia yang menghampiri saiz teks pelindung ditambah bertujuan untuk mendapatkan kapasiti maksimum mesej yang boleh disembunyikan di dalam teks pelindung seperti yang ditunjukkan pada Jadual 3.11.

Jadual 3.11

Saiz Mesej Rahsia Tambahan

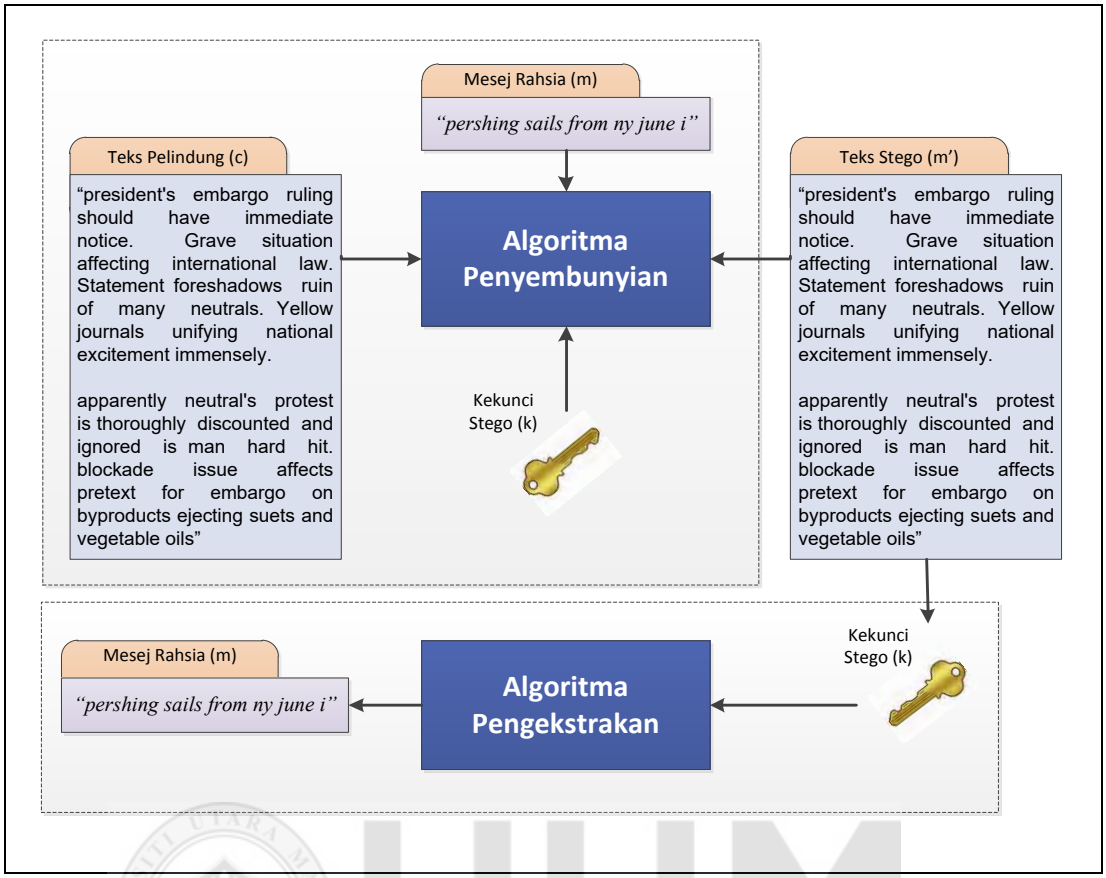
Teks Pelindung	**Mesej Rahsia Tambahan**
180	116
243	183
376	283
837	660
1700	1262
2153	1958
2638	2185

Jadual 3.11 menunjukkan beberapa saiz mesej rahsia tambahan yang menghampiri teks pelindung (melebihi saiz mesej kajian lepas) ditambah ke dalam kajian ini

bertujuan untuk mendapatkan kapasiti penyembunyian tertinggi berbanding dengan kajian lepas.

3.1.3 Fasa 3 : Implementasi

Fasa Implementasi melibatkan dua proses iaitu proses penyembunyian dan proses pengekstrakan terhadap teks stego yang dijana. Rajah 3.6 menunjukkan model penyembunyian mesej rahsia dan pengekstrakan teks stego yang melibatkan dua proses utama iaitu Proses Penyembunyian dan Proses Pengekstrakan. Proses penyembunyian melibatkan pertukaran mesej rahsia ke bentuk perwakilan (*x,y,z*) dan dipetakan ke bentuk RGB dan seterusnya disembunyikan pada lokasi rawak terpilih untuk menjana teks stego. Teks stego yang dihasilkan di ekstrak semula untuk mendapatkan mesej yang disembunyikan. Kedua-dua proses yang terlibat ini dihuraikan dengan terperinci pada Bab 4.



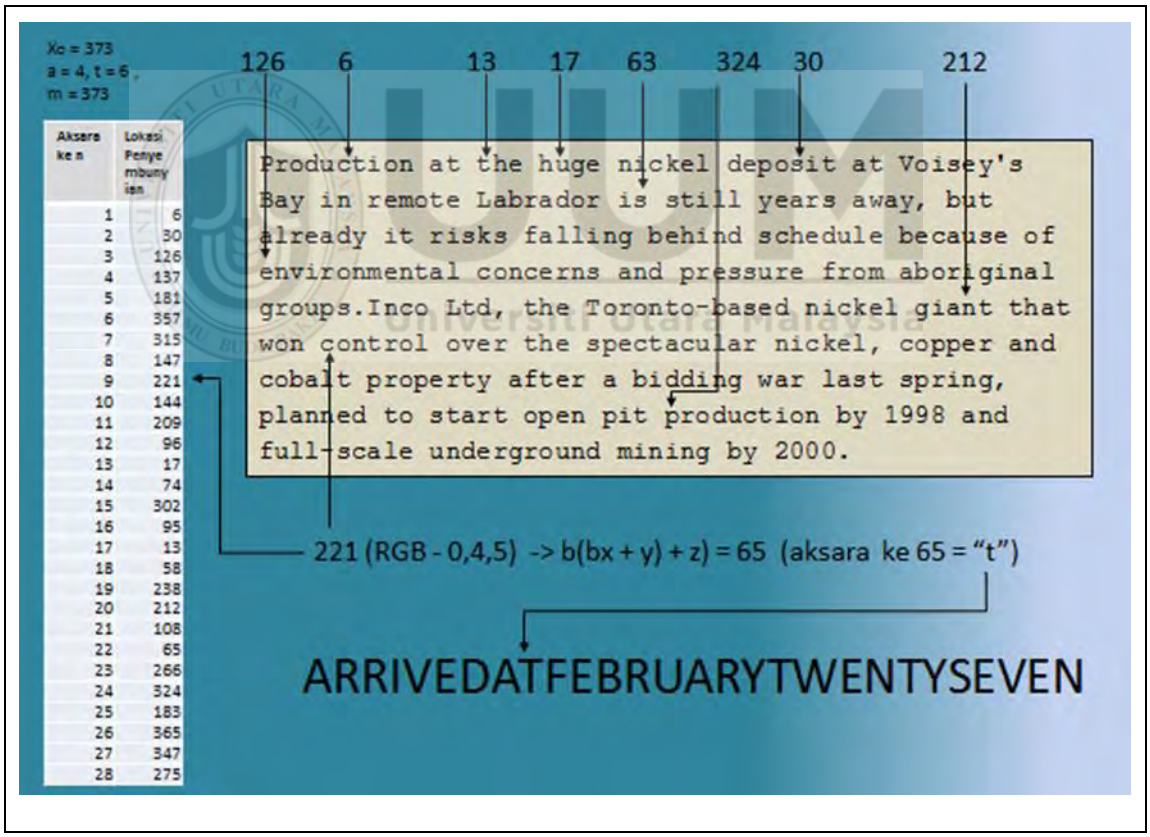
Rajah 3.6 Model Penyembunyian Dan Pengekstrakan Teks Stego

Microsoft Office (MS Word, MS Excel dan MS Access) merupakan aplikasi yang mengandungi model objek (*object model*) yang mempunyai ciri-ciri yang bersifat boleh dibaca atau boleh ditulis atau kedua-duanya sekali (Ishtiaq et al., 2017) dan hampir 80% syarikat enterprise menggunakan platform Microsoft Word di dalam produktiviti pekerjaan dan kolaborasi (R. Kumar, Malik, Singh, Kumar, & Chand, 2017). Menurut Ishtiaq et al., (2017), perisian Microsoft Visual Basic atau Peralatan Visual Studio Tools untuk Ms Office (*Visual Studio Tools for Office - VSTO*) menyediakan ciri-ciri yang boleh melakukan pengubahsuaian terhadap atribut (*bold, italic, underline, color* dsb) sesuatu aksara dan manipulasi terhadap atribut tersebut tidak menjejaskan isi kandungan fail dan sesuai digunakan di dalam steganografi.

Justeru itu, kajian ini menggunakan platform Microsoft Visual Basic 6.0 untuk melakukan proses penyembunyian dan proses pengekstrakan.

3.1.3.1 Proses Penyembunyian Dan Pengekstrakan

Input daripada Fasa 2 di atas digunakan untuk mencapai objektif kedua kajian. Proses penyembunyian dan pengekstrakan dilakukan berdasarkan kepada algoritma yang dihasilkan pada Fasa 2. Kedua-dua algoritma penyembunyian dan pengekstrakan tersebut diterangkan secara terperinci pada sub-topik 4.11 dan 4.12. Rajah 3.7 di bawah menunjukkan proses pengekstrakan untuk mesej yang disembunyikan.



Rajah 3.7 Proses Pengekstrakan Teks Stego

Berdasarkan kepada Rajah 3.7, mesej rahsia adalah bersaiz 28 aksara dan disembunyikan pada lokasi 6, 30, 126, 137, 181, 357, 315, 147, 221, 144 dan

seterusnya. Sebagai contoh, aksara mesej rahsia ke-9 disembunyikan pada lokasi ke 221 yang merupakan aksara *c* di dalam teks pelindung. Nilai RGB pada aksara tersebut akan di ekstrak (RGB = 0,4,5). Seterusnya nilai tersebut (0,4,5) dipadankan dengan `x = 0,y = 4` dan `z = 5` dan digantikan ke dalam persamaan 2.7 menggunakan nilai `b = 15`, iaitu:

```
=>(b*(bx +y) + z)

=>(15*(15*0 + 4) + 5)

= 15 * (4) + 5

= 65
```

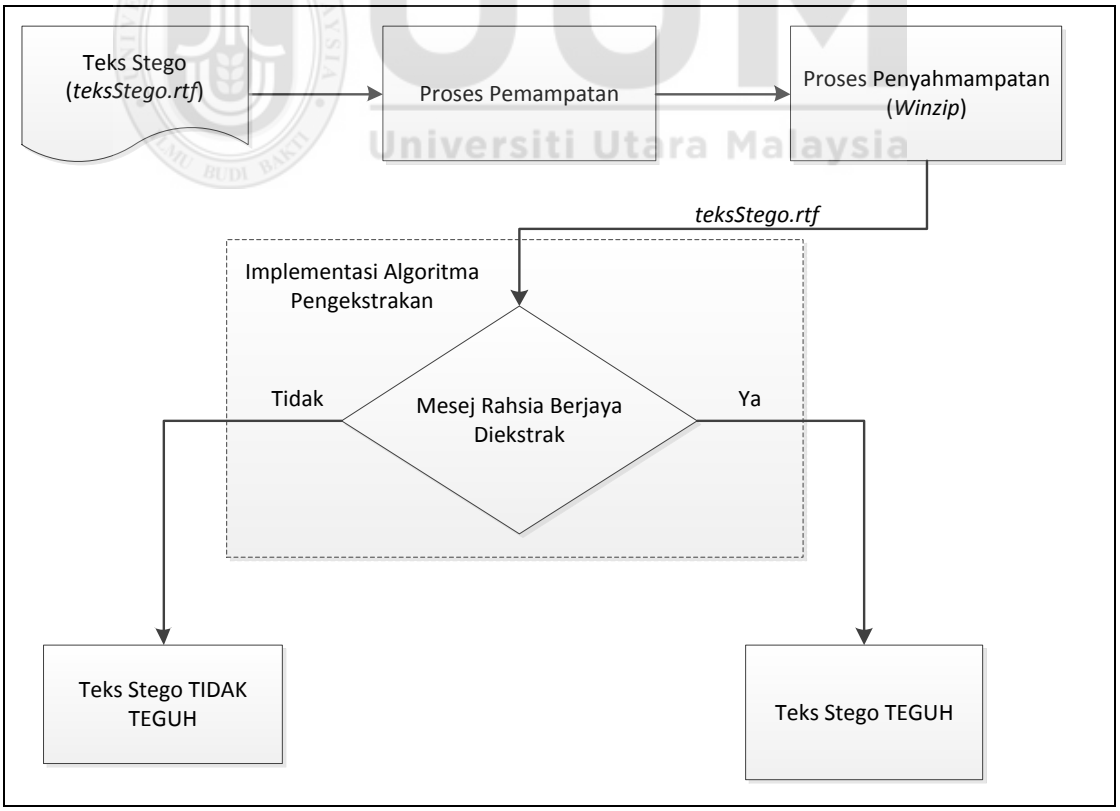
Seterusnya aksara pada lokasi ke 65 di dalam teks pelindung merupakan aksara "`t`" dan merupakan aksara ke-9 di dalam mesej tersembunyi.

3.1.3.2 Penjanaan Teknik Yang Dinamik

Pada fasa ini, teknik penempatan lokasi rawak digunakan untuk menyembunyikan mesej rahsia pada lokasi-lokasi tertentu secara rawak. Selain itu, algoritma untuk menjana kepelbagaian perwakilan terhadap mesej rahsia dibangunkan untuk memastikan aksara berulang dan aksara tunggal mesej rahsia dapat diwakilkan dengan pelbagai nilai yang dinamik. Proses perwakilan mesej rahsia yang dinamik dijelaskan dengan terperinci pada Bab 4. Hasil bagi fasa ini merupakan satu algoritma penyembunyian dan pengekstrakan serta pembentukan formula yang lebih bersifat dinamik bagi menjana teks stego.

3.1.4 Fasa 4 : Penilaian Prestasi

Fasa Penilaian Prestasi melibatkan penilaian terhadap teks stego yang dihasilkan berdasarkan kepada tiga ukuran iaitu kapasiti, keteguhan dan ketakbolehkelihatan yang dijelaskan secara terperinci pada sub-topik 2.3. Kapasiti mesej yang disembunyikan dan kapasiti penyembunyian tertinggi diukur menggunakan persamaan 2.2. Sementara itu, ketakbolehkelihatan diukur menggunakan skor Jaro Wrinkler yang dihuraikan pada persamaan 2.3. Skor Jaro Wrinkler menentukan sama ada kandungan teks pelindung dan teks stego adalah sama atau sebaliknya. Nilai skor bersamaan 1 menunjukkan terdapat kesamaan atau tiada perubahan isi kandungan berlaku antara teks pelindung dan teks stego. Rajah 3.8 menunjukkan proses menilai keteguhan teks stego yang dijana.



Rajah 3.8 Proses Penilaian Keteguhan Terhadap Teks Stego

Keteguhan diukur menggunakan proses pemampatan dan penyahmampatan berdasarkan kepada kejayaan untuk mengekstrak semula mesej rahsia sebelum dan selepas pemampatan menggunakan perisian *Winzip* seperti yang ditunjukkan pada Rajah 3.6. Teks stego yang dapat di ekstrak semula sepenuhnya tanpa sebarang kehilangan mesej rahsia adalah teguh dan sebaliknya ia dikatakan tidak teguh.

3.2 Ringkasan

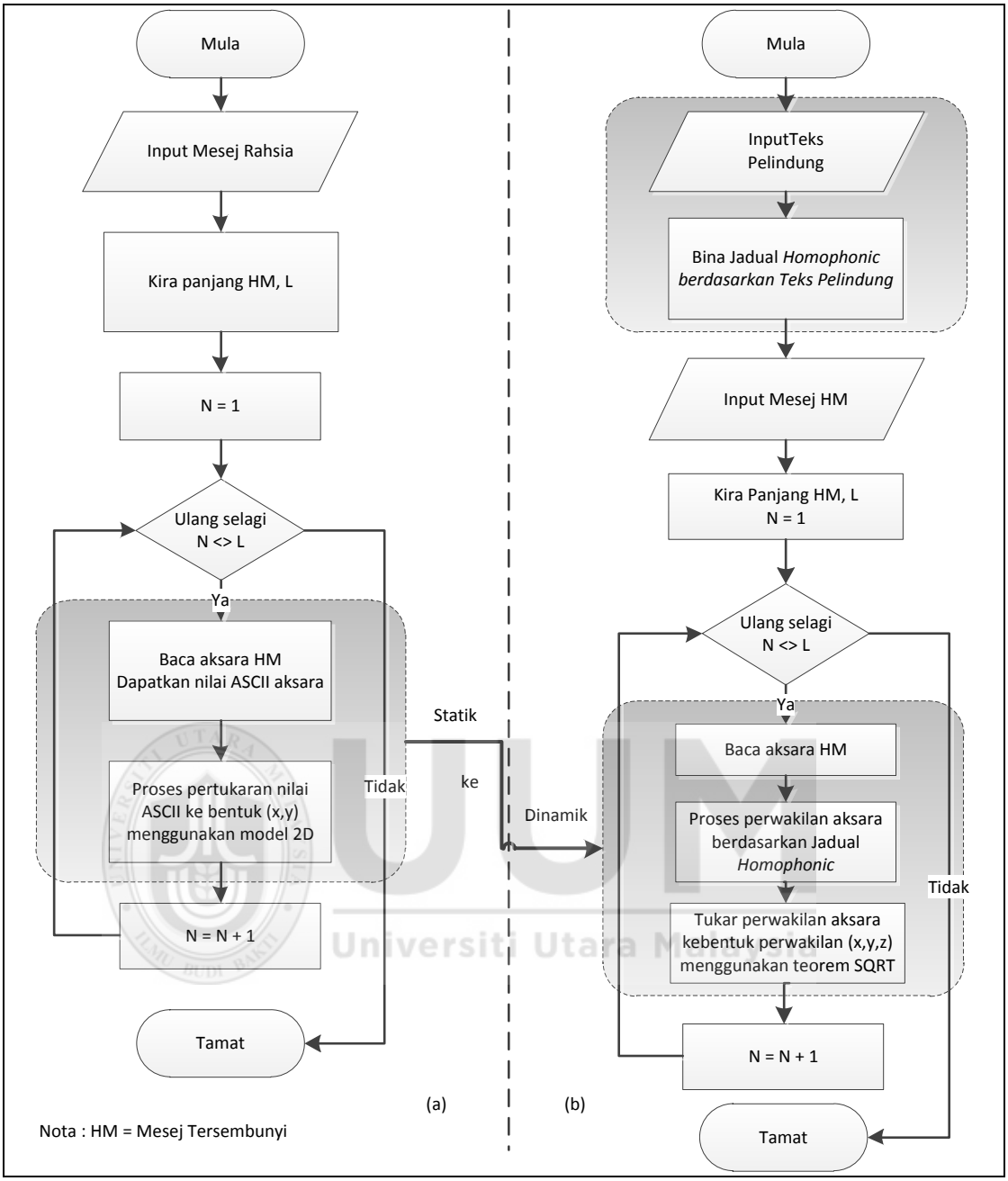
Bab ini menerangkan metodologi kajian yang berfungsi sebagai tunjang utama kepada kajian. Secara keseluruhannya, komponen metodologi kajian ini adalah berasaskan kepada penyelidikan eksperimental yang dibahagikan kepada empat fasa iaitu Fasa Awalan, Fasa Reka bentuk dan Pembangunan Model, Fasa Implementasi dan Fasa Penilaian Prestasi. Aksara berulang mesej rahsia diwakilkan dengan lebih dinamik menggunakan formula yang diperkenalkan pada subtopik 4.2 Semasa penyediaan data, pemilihan saiz mesej rahsia dilakukan dengan membahagikannya kepada dua bahagian iaitu berdasarkan kepada data kajian lepas dan data tambahan. Data kajian lepas digunakan untuk membanding keputusan yang dihasilkan di dalam kajian yang dijalankan berbanding dengan kajian lepas manakala data tambahan digunakan untuk menentukan kapasiti maksimum mesej rahsia yang boleh disembunyikan di dalam teks pelindung. Output yang dihasilkan pada setiap fasa telah berjaya memenuhi objektif kajian yang ditetapkan dalam penyelidikan ini.

BAB EMPAT
REKA BENTUK MODEL PENYEMBUNYIAN DAN IMPLIMENTASI

Bab ini menghuraikan proses penyembunyian dan pengekstrakan mesej rahsia berdasarkan kepada Reka bentuk dan Pembangunan Model serta Implementasi pada Fasa 2 dan Fasa 3. Huraian berkaitan teorem SQRT yang diterbitkan dari teorem QRT diterangkan di awal bab. Selain itu, bab ini juga menerangkan dengan terperinci penjanaan nombor rawak untuk menentukan lokasi penyembunyian serta teknik perwakilan aksara berulang dan aksara tunggal bagi mesej rahsia. Penerangan terperinci berkaitan dengan teorem SQRT, perwakilan warna RGB, teknik penempatan secara rawak, algoritma penyembunyian serta algoritma pengekstrakan diterangkan secara terperinci dalam bab ini.

4.1 Reka bentuk Perwakilan Statik Ke Perwakilan Dinamik

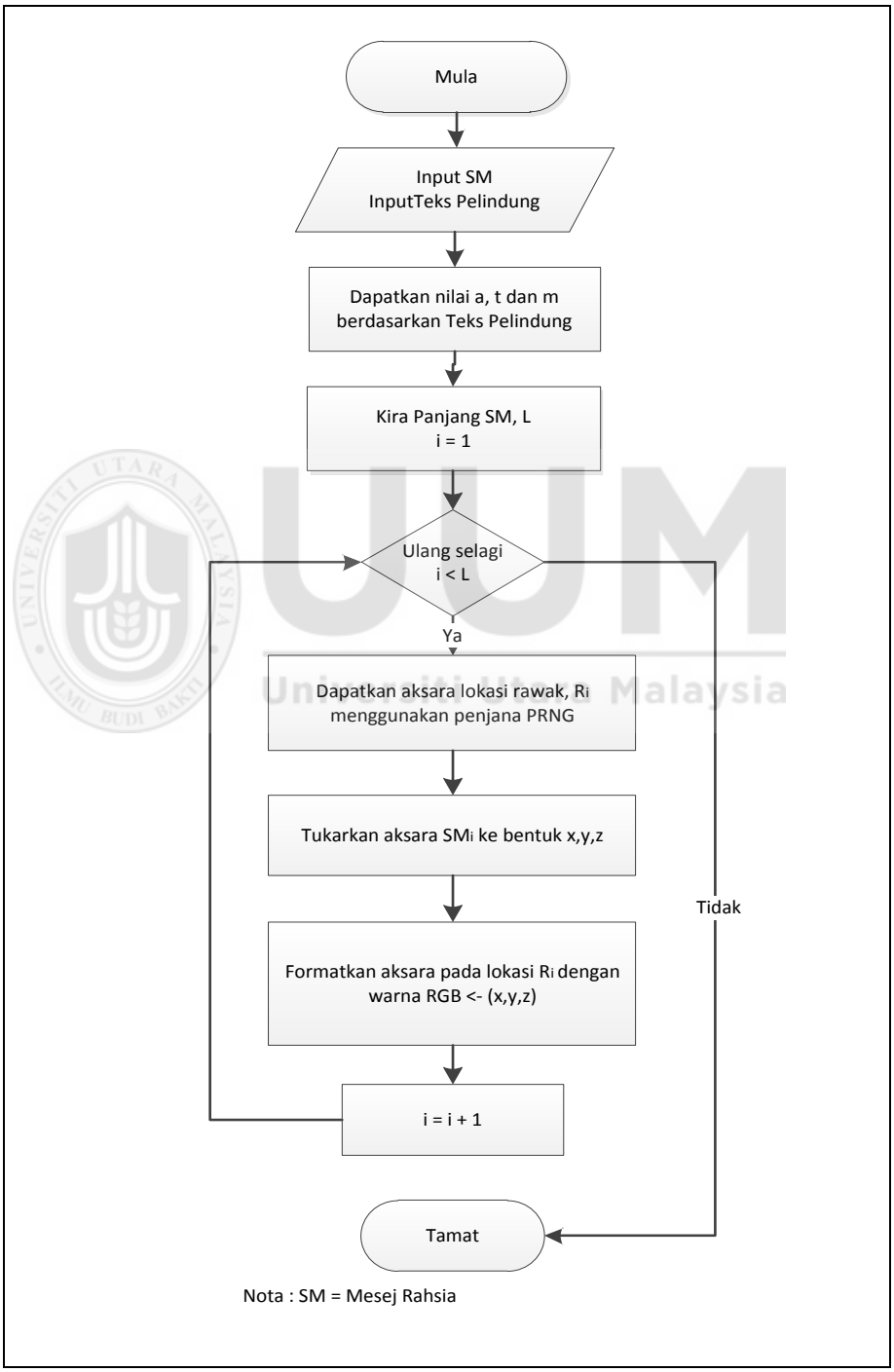
Bahagian ini menjelaskan reka bentuk kajian lepas dan reka bentuk kajian yang dibangunkan. Berdasarkan kepada kajian lepas yang dijalankan oleh Fuad et al. (2014), Mandal et al. (2014) serta Mandal, Chatterjee dan Chakraborty (2019) yang berasaskan kepada perwakilan dua dimensi, setiap aksara mesej rahsia ditukarkan ke bentuk ASCII dan seterusnya menukarkan nilai ASCII tersebut ke bentuk koordinat (x,y) seperti yang ditunjukkan dalam Rajah 4.1 (a). Kelemahan teknik ini ialah setiap aksara berulang mesej rahsia diwakili dengan perwakilan yang sama atau statik seperti yang ditunjukkan di dalam Jadual 3.4. Penambahbaikan dilakukan untuk mewakilkan aksara mesej rahsia berulang dan tunggal dengan perwakilan yang lebih bersifat dinamik seperti yang ditunjukkan di dalam Rajah 4.1 (b).



Rajah 4.1 Penambahbaikan Teknik Perwakilan Statik (a) ke Dinamik (b)

Rajah 4.1 (b) menunjukkan penjanaan Jadual *Homophonic* berdasarkan kepada teks pelindung dan perwakilan aksara mesej rahsia ke bentuk dinamik. Proses ini melibatkan beberapa langkah di mana langkah pertama ialah menjana Jadual *Homophonic* untuk mewakilkan aksara mesej rahsia berulang dan tunggal ke bentuk dinamik seperti yang diterangkan pada subtopik 4.3. Seterusnya, perwakilan aksara

tersebut akan ditukar ke bentuk perwakilan (*x,y,z*) menggunakan teorem SQRT. Setelah proses perwakilan ke bentuk dinamik selesai dilakukan, maka proses penyembunyian akan dilaksanakan. Rajah 4.2. menunjukkan proses penyembunyian mesej rahsia pada lokasi rawak terpilih.



Rajah 4.2 Proses Penyembunyian Mesej Pada Lokasi Rawak

Rajah 4.2 menunjukkan proses penyembunyian mesej rahsia pada lokasi rawak yang dihuraikan secara terperinci pada sub-topik 4.3 hingga 4.8. Nilai kekunci a, t dan m diperoleh berdasarkan kandungan teks pelindung seperti yang diterangkan pada sub-topik 4.7. Aksara teks pelindung pada lokasi rawak terpilih akan diformatkan dengan warna RGB yang dipadankan dengan nilai *x, y, z* seperti yang diterangkan pada sub-topik 4.3.

4.2 Terbitan Teorem Baki Hasil Bahagi Pringkat Kedua

Teorem Baki Hasil Bahagi Peringkat Kedua (*Second Quotient Remainder Theorem - SQRT*) merupakan teorem yang diterbitkan daripada pembahagian kali kedua terhadap Teorem Baki Hasil Bahagi (*Quotient Remainder Theorem - QRT*). Teorem *SQRT* ini digunakan untuk menukarkan sesuatu nilai ke bentuk perwakilan 3D atau (*x,y,z*) dan kedua-duanya diterangkan secara terperinci pada subtopik 4.2.1 dan 4.2.2.

4.2.1 Teorem Baki Hasil Bahagi

Teorem QRT telah digunakan di dalam kajian video steganografi sebagai medium pelindung yang dilakukan oleh Chandini dan Ganesh Kumar (2018) untuk menukarkan nilai ASCII kepada perwakilan nombor menggunakan formula berikut:

$$Divisor\ (d) * Quotient\ (q) + Remainder\ (r) = ASCII\ Message\ (m) \qquad (4.1)$$

Formula tersebut telah digunakan untuk mewakilkan mesej tersembunyi yang disulitkan dalam bentuk *d, q* dan *r* sebelum melakukan proses penyembunyian dan pengekstrakan menggunakan teknik padanan corak piksel. Berdasarkan kepada persamaan 4.1, jika $d = b$, dan $m = v$, maka

$$\frac{v}{b} \quad = q\ baki\ r \qquad (4.2)$$

Teorem ini menyatakan bahawa, apabila sesuatu nombor (v) dibahagikan dengan sesuatu nilai (b), maka hasil bahagi (q) dan baki (r) akan diperolehi. Teorem ini boleh diterjemahkan ke bentuk persamaan matematik seperti yang ditunjukkan pada persamaan 4.2. Persamaan 4.2 boleh ditulis semula di dalam bentuk persamaan matematik seperti di bawah

$$v = bq + r \tag{4.3}$$

di mana

v : Nombor Integer
b : Pembahagi (*divisor*)
q : Nilai hasil bahagi (*quotient*)
r : Baki hasil bahagi (*remainder*)

Persamaan 4.3 digunakan sebagai asas untuk menerbitkan teorem *SQRT* yang digunakan untuk menukarkan sesuatu nilai ke bentuk perwakilan *x,y,z* yang diterangkan secara terperinci pada subtopik 4.2.2.

4.2.2 Teorem Baki Hasil Bahagi Peringkat Kedua - SQRT

Berdasarkan kepada persamaan 4.3, jika nilai q dilakukan pembahagian kali kedua menggunakan nilai pembahagi b yang sama, maka persamaan berikut akan diperolehi.

$$q = bq'' + r'' \tag{4.4}$$

Jika digantikan persamaan 4.4 ke dalam persamaan 4.3, maka persamaan v boleh ditulis sebagai;

$$\begin{aligned} v &= bq + r \\ &= b(bq'' + r') + r \end{aligned} \tag{4.5}$$

Persamaan 4.5 dapat dibuktikan seperti berikut. Andaikan $v = 1968$, dan $b = 15$.

$$\begin{aligned} v &= bq + r \\ &= 15(131) + 3 \\ &= 15[15(8) + 11] + 3 \rightarrow b(bq'' + r'') + r \\ &= 1968 \text{ (\#terbukti)} \end{aligned}$$

Jika $q'' = x$; $r'' = y$ dan $r = z$, maka persamaan 4.5 boleh ditulis seperti di bawah.

$$v = b(bx + y) + z \tag{4.6}$$

Persamaan 4.6 diterbitkan dengan membuat pembahagian kali kedua terhadap nilai q pada persamaan 4.3 menggunakan nilai b yang sama. Justeru itu teorem pada persamaan 4.6 ini dikenali sebagai ***Second Quotient Remainder Theorem – SQRT*** yang digunakan untuk menukarkan sesuatu nilai ke bentuk perwakilan (x,y,z) atau 3D dan dipadankan dengan nilai warna RGB.

Teorem S*QRT* digunakan untuk mewakilkan aksara mesej rahsia dalam bentuk perwakilan 3D atau (x,y,z) seperti yang diterangkan di atas. Nilai maksimum bagi x, y dan z bergantung kepada nilai maksimum, v_{max} di mana $v_{max} = b^3 + b^2 - 1$. Nilai b menentukan sama ada nilai (x,y,z) yang diperolehi berada di dalam julat yang dikehendaki atau sebaliknya. Contoh, jika $b = 15$, maka nilai maksimum, v_{max} ialah $(15^3 + 15^2 - 1) = 3599$. Oleh itu, jika nilai $v < (b^3 + b^2 - 1)$, maka setiap nilai x,y dan z yang diperolehi tidak akan melebihi nilai b. Sebaliknya, Jika nilai $v > (b^3 + b^2 - 1)$, akan mengakibatkan sekurang-kurangnya salah satu nilai x atau y atau z yang dihasilkan akan melebihi nilai b. Contoh, jika $v = 3602$ (melebihi nilai v_{max}) dan $b = 15$, maka nilai x, y dan z yang dihasilkan ialah $(16,0,2)$, di mana nilai x melebihi nilai b. Oleh itu, di dalam kajian ini nilai $b = 15$ dipilih bertujuan untuk memastikan nilai (x,y,z) yang dihasilkan tidak melebihi julat (15,15,15) yang akan digunakan

untuk mewakilkan warna RGB untuk semasa proses penyembunyian. Namun begitu, teorem SQRT masih lagi boleh diaplikasikan di dalam kajian berkaitan dengan mengubah nilai *b* sekiranya julat *x,y*,dan *z* tidak menjadi pertimbangan penyelidik.

4.3 Perwakilan Rawak Aksara Mesej Rahsia Berdasarkan Jadual *Homophonic*

Setiap aksara mesej rahsia diwakilkan dengan nilai rawak agar ia bersifat lebih dinamik terutama bagi membolehkan aksara mesej rahsia berulang dapat diwakilkan dengan pelbagai nilai yang berbeza. Perwakilan berbentuk dinamik dilakukan dengan menjana Jadual *Homophonic* berdasarkan frekuensi aksara yang terdapat di dalam teks pelindung seperti yang dijelaskan pada sub-topik 2.8.2. Jadual *Homophonic* (seperti Lampiran E – Langkah 3) dapat mewakilkan aksara mesej rahsia dengan pelbagai nilai berbeza terutama bagi aksara yang berulang. Penggunaan teknik rawakan berjaya menghasilkan pelbagai set mesej rahsia seperti yang ditunjukkan di dalam Jadual 4.1.

Jadual 4.1

Perwakilan Mesej Rahsia Berdasarkan Jadual Homophonic

	M	**E**	**E**	**T**	**Y**	**O**	**U**	**A**	**T**	**T**	**E**	**N**
C1	503	1968	81	1098	1590	287	1782	568	1477	682	1866	496
C2	1284	577	995	239	338	1032	1283	1131	1884	1914	1809	1953
C3	450	947	1864	11	1167	1534	1143	1313	366	1644	504	1691
C4	1285	297	1968	1981	1683	617	292	264	1884	293	995	595
C5	576	1014	361	1784	309	386	215	1681	607	340	1105	522
C6	730	1604	539	1922	794	1714	200	125	76	1055	1209	824

Jadual 4.1 menunjukkan aksara mesej rahsia diwakilkan dengan pelbagai perwakilan yang dilabelkan dengan C1 hingga C6 di mana penjanaannya bergantung pada Jadual *Homophonic*. Faktor yang paling penting di dalam perwakilan ini ialah setiap aksara berulang dan tunggal bagi sesuatu set perwakilan mesej rahsia diwakilkan dengan pelbagai nilai. Sebagai contoh, aksara berulang "E" bagi perwakilan set C1

diwakilkan dengan pelbagai nilai yang berbeza iaitu 1968, 81 dan 1866. Di samping itu, aksara tak berulang juga dapat diwakilkan dengan pelbagai nilai dengan wujudnya Jadual *Homophonic* seperti yang diwakilkan oleh set C1 hingga C6. Sebagai contoh, aksara "Y" boleh diwakilkan dengan pelbagai nilai sama ada 1550, 338, 1167, 1683, 309 atau 794. Perwakilan nilai yang pelbagai ini membolehkan mesej rahsia dapat diwakilkan dengan nilai yang lebih bersifat dinamik.

4.4 Perwakilan Aksara Mesej Rahsia Dalam Bentuk *x,y, z*

Perwakilan nilai yang berbeza bagi setiap aksara mesej rahsia dilakukan secara rawak berdasarkan Jadual *Homophonic* yang dijana seperti yang dijelaskan pada sub-topik 4.4. Teknik SQRT digunakan untuk menukarkan mesej rahsia ke bentuk perwakilan (*x,y,z*) seperti yang dijelaskan dalam sub-topik 4.3. Jadual 4.2 menunjukkan aksara mesej rahsia bagi set C1 (rujuk Jadual 4.1) yang telah ditukarkan ke bentuk perwakilan (*x,y,z*) menggunakan teknik SQRT.

Jadual 4.2

Perwakilan Nilai Rawak Dalam Format x, y dan z

	M	E	E	T	Y	O	U	A	T	T	E	N
	503	1968	81	1098	1590	287	1782	568	1477	682	1866	496
x	2	8	0	4	7	1	7	2	6	3	8	2
y	3	11	5	13	1	4	13	7	8	0	4	3
z	8	3	6	3	0	2	12	13	7	7	6	1

Berdasarkan kepada Jadual 4.2 di atas, didapati bahawa aksara berulang E diwakili dengan tiga perwakilan (*x, y, z)* yang berbeza iaitu (8,11,3), (0,5,6) dan (8,4,6). Ini menunjukkan aksara yang sama dapat diwakilkan dengan pelbagai nilai berbeza yang akan dipadankan dengan nilai warna RGB semasa proses penyembunyian.

4.5 Pertukaran Nilai *x,y,z* Ke Nilai Asal

Nilai asal *v,* dapat diperolehi semula dengan menggantikan nilai *x*, *y* dan *z* ke dalam persamaan 4.6. Sebagai contoh, nilai asal bagi perwakilan (*x,y,z*) = *(8,11,3)* boleh diperolehi semula dengan menggantikan nilai tersebut ke dalam persamaan 4.6 seperti berikut:

$$v = b(bx + y) + z$$
$$= 15(15x8 + 11) + 3$$
$$= 15(131) + 3$$
$$= 1968$$

Oleh itu, nilai asal v = `1968` telah dibuktikan sebagai perwakilan kepada (`8,11,3`) dalam format (*x,y,z*) yang mewakili aksara E.

4.6 Penjanaan Lokasi Rawak

Penyembunyian mesej rahsia terhadap teks pelindung dilakukan berdasarkan kepada lokasi rawak. Aksara teks pelindung pada lokasi rawak terpilih akan diformatkan dengan warna RGB yang dipadankan dengan nilai *x,y,z* seperti yang dijelaskan pada sub-topik 4.5. Nilai lokasi rawak dijana menggunakan penjana nombor rawak (PNR) berikut:

$$x_{n+1} = (ax_n + t) \,\%\, m \qquad (4.7)$$

di mana

x_{n+1} : lokasi jujukan rawak ke n ; $n = 0,1,2,3.....$

a dan t : pemalar integer dinamik untuk penjanaan set rawak

m : nombor perdana yang paling hampir dengan bilangan aksara di dalam teks pelindung

Nilai permulaan, x_0 merupakan nilai permulaan bagi proses penjanaan nombor rawak. Di dalam kajian ini bilangan aksara yang terdapat di dalam teks pelindung (di tolak 3) dijadikan sebagai nilai permulaan, x_0 penjanaan nombor rawak. Tiga lokasi terakhir teks pelindung digunakan untuk menyimpan nilai a, t dan saiz teks pelindung (bilangan aksara). Nilai a dan t merupakan kekunci utama untuk menghasilkan nombor rawak. Menurut Srikanth et al. (2017), untuk mendapatkan nilai rawak yang maksimum, maka nilai *m* perlulah sebesar yang mungkin. Justeru itu, dalam kajian ini nombor perdana yang paling hampir dengan bilangan aksara teks pelindung (saiz teks pelindung) dipilih sebagai nilai m yang bertujuan untuk mendapatkan nilai julat rawak yang terbesar.

Sebagai contoh, jika nilai a = 13; t = 9, m = 2207 dan saiz teks pelindung ialah 2212, maka sebahagian senarai lokasi rawak boleh diperolehi seperti berikut:

$x_1 = (ax_0 + t) \% m$
$= (13(2209) + 9) \% 2207$
$= (28717 + 9) \% 2207$
$= 28726 \% 2207$
$= 35$

$X_2 = (13(\underline{35}) + 9) \% 2207$
$= (455 + 9) \% 2207$
$= 464$

$X_3 = (13(\underline{464}) + 9) \% 2207$
$= (6032 + 9) \% 2207$
$= 1627$

$X_4 = (13(\underline{1627}) + 9) \% 2207$
$= (21151 + 9) \% 2207$
$= 1297$

* Nota : Dalam sampel kajian ini, m = 2207 merupakan nombor perdana paling hampir dengan 2209 di mana; 2209 = saiz teks pelindung - 3

Berdasarkan kepada nilai rawak di atas, aksara teks pelindung pada lokasi `35,464,1297` akan diformatkan dengan warna RGB masing-masing `(2,3,8), 8,11,3),(0,5,6)` yang diperolehi pada sub-topik 4.4. Proses di atas berulang sehingga akhir aksara mesej rahsia.

4.7 Kekunci Persamaan Jujukan Lokasi Rawak

Nilai `a` dan `t` di dalam persamaan 4.7 merupakan kekunci untuk penjanaan nombor rawak dan disembunyikan pada lokasi kedua terakhir di dalam teks stego yang dihasilkan. Nilai `a` dan `t` boleh digantikan dengan sebarang nilai berbentuk integer. Aksara teks pelindung pada lokasi tersebut akan diformatkan dengan warna RGB dalam format `(0,a,t)` untuk menyimpan nilai `a` dan `t`. Berdasarkan teks pelindung di dalam Lampiran E, nilai `a` dan `t` masing-masing disimpan pada aksara "T" yang diwakili oleh warna G dan B.

4.8 Saiz Mesej Rahsia

Saiz atau panjang mesej rahsia disembunyikan pada lokasi terakhir teks stego. Nilai ini disimpan di dalam format RGB dan perlu ditukarkan menggunakan persamaan 4.6 untuk mendapatkan nilai sebenar. Sebagai contoh, aksara terakhir bagi teks stego dalam Lampiran E ialah simbol noktah "." . Perwakilan nilai RGB pada lokasi ini akan ditukar ke bentuk nombor yang mewakili saiz mesej rahsia menggunakan persamaan 4.6.

4.9 Perwakilan Nilai Warna RGB

Sistem warna RGB menghasilkan warna dengan menggabungkan nilai warna Merah (*Red*), Hijau (*Green*) dan Biru (*Blue*). Nilai RGB diwakilkan di dalam format RGB bermula dari nilai (0,0,0) hingga (255,255,255). Nilai (0,0,0) mewakili nilai hitam dan (255,255,255) mewakili nilai putih, manakala sebahagian warna-warna lain diwakili dengan perwakilan RGB seperti yang ditunjukkan pada Rajah 4.3 di bawah.

Warna	R,G,B (Grey)	Warna	R,G,B (Pelbagai)
	0,0,0		16,160,100
	1,1,1		20,120,100
	3,3,6		25,125,100
	5,5,7		0,138,0
	8,10,11		0,255,0
	10,8,11		1,50,138
	10,10,10		24,222,232
	13,13,13		255,0,0
	15,15,15		255,255,255

Rajah 4.3 Sebahagian Perwakilan Warna RGB

Kajian steganografi berasaskan warna telah dijalankan oleh Singh dan Diwakar (2014) menggunakan dua warna RGB iaitu (0,0,0) dan (1,1,1) untuk mewakili bit 0 dan 1 masing-masing. Kapasiti penyembunyian menggunakan teknik ini agak terhad di mana lapan aksara teks pelindung diperlukan untuk menyembunyikan satu bait kekunci mesej rahsia. Kekurangan tersebut telah dilakukan penambahbaikan dengan menggunakan 16 variasi warna yang terdiri daripada nilai RGB bermula dari (0,0,0), (1,1,1), (2,2,2), (3,3,3) hingga (15,15,15) untuk mewakilkan kekunci mesej rahsia. Setiap variasi warna dapat menyembunyikan 4 bit mesej rahsia (berdasarkan kepada Jadual Pemetaan RGB). Namun, kapasiti penyembunyian masih rendah di mana dua

aksara teks pelindung diperlukan untuk menyembunyikan 1 bait aksara tersembunyi selain teknik penyembunyian adalah secara berjujukan.

Nilai RGB (0,0,0) hingga (15,15,15) menghasilkan skala warna yang hampir gelap berbanding dengan nilai-nilai lain seperti yang ditunjukkan di dalam Rajah 4.3. Berdasarkan kepada Rajah 4.3, nilai RGB di antara (0,0,0) hingga (15,15,15) menunjukkan perubahan warna gelap (*grey*) sukar dibezakan oleh sistem visual manusia berbanding perubahan warna yang lebih cerah (Mokrzycki & Tatol, 2011) dan memenuhi model piawaian *Commission International de „Eclairage - CIELAB* (Al-Azzawi, 2018). Model CIELAB digunakan untuk mengukur perbezaan warna yang relevan dengan sistem visual manusia. Oleh itu, di dalam kajian ini, nilai 15 dijadikan asas untuk pertukaran sesuatu nilai kepada perwakilan nilai warna RGB bertujuan untuk memastikan perubahan warna RGB tidak menjejaskan kualiti teks stego yang dijana serta tidak dapat dikesan oleh sistem visual manusia.

4.10 Algoritma Proses Penyembunyian

Lokasi rawak yang dijana pada sub-topik 4.7 digunakan untuk menyembunyikan aksara mesej rahsia dengan memformatkan warna RGB aksara pada lokasi tersebut. Algoritma untuk proses penyembunyian ditunjukkan di dalam Rajah 4.4.

Algoritma : Penyembunyian Mesej Rahsia

Input : Mesej Rahsia - ***msgTersembunyi***
: Teks Pelindung - ***teksPelindung***
Output : Teks Stego - ***teksStego.rtf***

```
1   i = 0
2   saizMsgTersembunyi  = length(MsgTersembunyi)
3   saizTeksPelindung = length(TeksPelindung)
4   Baca teksPelindung dan jadual Homophonic
5   Tukarkan aksara msgTersembunyi kepada perwakilan nombor rawak,v menggunakan jadual Homophonic
6   Dapatkan nilai nombor perdana, m paling hampir dengan nilai saizTeksPelindung
7   Dapatkan nilai a dan t secara rawak
8   Dapatkan nilai permulaan, x_0 = saizTeksPelindung - 3
9   while i < > saizMsgTersembunyi do
10      Baca perwakilan aksara pertama mesej rahsia, v_i
11      Tukarkan perwakilan, v_i ke bentuk RGB (sub-topik 4.2)
12              R = x
13              G = y
14              B = z
15      Tentukan nilai lokasi rawak (sub-topik 4.6)
16          x_{n+1} = (a x_n + t) % m
17
18      Formatkan warna aksara pada lokasi ke- x_{n+1} dengan nilai RGB
19      i = i + 1
20  end while
21  Formatkan warna aksara pada lokasi terakhir dengan nilai saizMsgTersembunyi
22  Formatkan warna aksara pada lokasi kedua terakhir dengan nilai a dan t
23  teksStego = save(teksPelindung)
```

Rajah 4.4 Algoritma Proses Penyembunyian Mesej Rahsia

Berdasarkan kepada algoritma pada Rajah 4.4, panjang mesej rahsia dan teks pelindung ditentukan terlebih dahulu. Seterusnya, jadual *Homophonic* dijana berdasarkan kepada teks pelindung seperti yang dijelaskan pada sub-topik 4.4. Setiap aksara mesej rahsia diwakilkan dengan satu nilai secara rawak berdasarkan kepada jadual *Homophonic*. Nilai tersebut akan ditukarkan ke bentuk perwakilan *x*, *y* dan *z* (sub-topik 4.2) dan dipadankan dengan nilai warna RGB (R ← *x*; G ← *y*; B ← *z*). Langkah seterusnya ialah mengenal pasti lokasi rawak untuk proses penyembunyian berdasarkan kepada persamaan 4.7. Aksara pada lokasi tersebut akan diformatkan dengan nilai warna RGB dan proses ini akan berulang sehingga akhir aksara mesej rahsia. Output yang dihasilkan pada proses penyembunyian ini merupakan teks stego yang disimpan (*save*) dalam format *Rich Text Format* (*.rtf)*. Format fail *.rtf* digunakan bertujuan untuk membolehkan sesuatu aksara diformatkan dengan warna RGB dan

pernah digunakan di dalam kajian lepas (Singh et al., 2014;Baykara, Das & Tuna, 2017; Win & Oo, 2018). Fail berformat *.rtf* boleh dibaca oleh kebanyakan aplikasi pangkalan data, klien e-mel dan sistem pengoperasian seperti Unix, Machintosh dan Windows berbanding dengan format *.doc*. Ia merupakan satu produk yang bersifat pertukaran lintas-platform (*cross-platform interchange*) serta menyokong format fail teks (*text file formating*).

4.11 Algoritma Proses Pengekstrakan

Teks stego (*.rtf*) yang dihasilkan pada sub-topik 4.10 akan di ekstrak untuk mendapatkan semula mesej yang disembunyikan. Proses pengekstrakan dilakukan berdasarkan kepada algoritma yang ditunjukkan pada Rajah 4.5 di bawah.

Algoritma : Pengekstrakan Mesej Rahsia
Input : Teks Stego – ***teksStego.rtf***
Output : Mesej Rahsia - ***msgTersembunyi***

```
1   b = 15
2   Baca fail teksStego.rtf
3   saizTeksStego = length(teksStego.rtf)
4   Dapatkan panjang mesej rahsia
5       xyz = extract(warna RGB pada lokasi terakhir teksStego)
6       x = R
7       y = G
8       z = B
9       msgLen = b(bx + y) + z
10  Dapatkan nilai kekunci a dan t
11      at = extract(warna RGB pada lokasi kedua terakhir teksStego)
12      a = G
13      t = B
14  Kira nilai m,
15      m = nilai nombor perdana paling hampir dengan nilai saizTeksStego.
16  Dapatkan nilai permulaan, x0 untuk nombor rawak,
17    x0 = saizTeksStego − 3     ; 2 aksara terakhir telah digunakan
18  n = 0
19  while n <= msgLen
20  |   x(n+1) = (a x(n) + b) % m
21  |   Dapatkan nilai (x,y,z) berdasarkan nilai RGB pada lokasi x(n+1)
22  |   Tukarkan nilai (x,y,z) kepada nilai sebenar v,
23  |       v = b(bx+y) + z
24  |       aksTersembunyi = Banding nilai v dengan Jadual Homophonic
25  |   msgTersembunyi(n) = aksTersembunyi
26  |   n++
    end while
```

Rajah 4.5 Algoritma Proses Pengekstrakan Teks Stego

Rajah 4.5 menunjukkan algoritma untuk proses pengekstrakan terhadap fail teks stego yang dihasilkan. Langkah pertama ialah mendapatkan saiz teks stego, *saizTeksStego* berdasarkan kepada fail input teks stego, *teksStego.rtf*. Seterusnya panjang teks tersembunyi, *msgLen* dikenal pasti dengan mengekstrak nilai RGB pada aksara terakhir *teksStego.rtf*. Nilai kekunci `a` dan `t` diperolehi dengan mengekstrak nilai RGB pada lokasi kedua terakhir *teksStego.rtf* manakala nilai *m* diperolehi berdasarkan kepada nombor perdana yang paling hampir dengan nilai saiz *saizTeksStego*. Nilai permulaan untuk lokasi rawak diperolehi berdasarkan nilai saiz *teksStego-3*. Seterusnya, warna RGB aksara pada lokasi ini akan di ekstrak dan ditukarkan ke nilai sebenar, v dan dibandingkan dengan Jadual *Homophonic* untuk mendapatkan aksara sebenar mesej rahsia. Proses mendapatkan lokasi rawak dan mengekstrak warna RGB pada lokasi ini akan berulang sehingga *msgLen*.

4.12 Ringkasan

Secara keseluruhan, bab ini menjelaskan Teorem *SQRT* yang diterbitkan daripada teorem *QRT*. Aksara mesej rahsia diwakilkan dengan nilai rawak berdasarkan kepada Jadual *Homophonic* sifer yang dijana. Proses penyembunyian bermula dengan menukarkan nilai rawak ini ke bentuk perwakilan (*x,y,z*) menggunakan formula yang dijelaskan yang pada sub-topik 4.3 dan dipadankan dengan warna RGB. Seterusnya aksara lokasi rawak akan ditentukan untuk memformatkan aksara tersebut dengan nilai warna RGB. Proses pengekstrakan dilakukan dengan mengenal pasti lokasi rawak menggunakan persamaan 4.7 dan mengekstrak warna RGB pada lokasi tersebut menggunakan persamaan 4.6. Seterusnya, nilai yang dihasilkan dibandingkan dengan Jadual *Homophonic* sifer untuk mengenal pasti aksara mesej rahsia. Proses

penyembunyian dan pengekstrakan dilakukan berdasarkan kepada algoritma yang ditunjukkan pada Rajah 4.4 dan Rajah 4.5.

BAB LIMA
HASIL DAN PERBINCANGAN

Bab ini menerangkan hasil dan perbincangan yang diperolehi selepas menjalankan kajian. Perbandingan terhadap hasil yang diperolehi dibandingkan dengan kajian lepas bagi menjelaskan output yang dihasilkan. Semua fail mesej rahsia dan fail teks pelindung yang digunakan telah di analisis sebelum melakukan proses penyembunyian dan pengekstrakan. Fail teks stego yang dihasilkan di analisis dan dibincangkan di dalam bab ini. Selain itu, proses penilaian terhadap kapasiti penyembunyian, keteguhan dan ketakbolehkelihatan teks stego dilakukan untuk mengenal pasti prestasi teks stego yang dihasilkan berbanding dengan kajian lepas.

5.1 Hasil Analisis Teks Pelindung Dan Mesej Rahsia

Teks pelindung dan mesej rahsia yang digunakan di dalam kajian ini dianalisis untuk mengenal pasti taburan aksara di dalam kedua-dua fail tersebut. Hasil analisis diterangkan pada sub-topik 5.1.1 dan 5.1.2. Fail teks pelindung dilabelkan dengan kod CT*n* (*n* ialah bilangan aksara di dalam fail) untuk memudahkan perbincangan seperti yang ditunjukkan dalam Jadual 5.1 di bawah.

5.1.1 Hasil Analisis Aksara Teks Pelindung

Analisis terhadap 43 teks pelindung telah dilakukan untuk mengenal pasti taburan aksara di dalam setiap fail. Secara keseluruhannya, fail teks pelindung yang diuji adalah bersaiz antara 1Kb hingga 4Kb di mana kandungan aksara didalamnya mengandungi 180 hingga 3305 aksara. Kandungan aksara di dalam setiap fail adalah

berbeza mengikut fail. Jadual 5.1 menunjukkan analisis peratus kewujudan dan ketiadaan aksara di dalam teks pelindung yang digunakan.

Jadual 5.1

Peratus Ketiadaan Dan Kewujudan Aksara di dalam Teks Pelindung

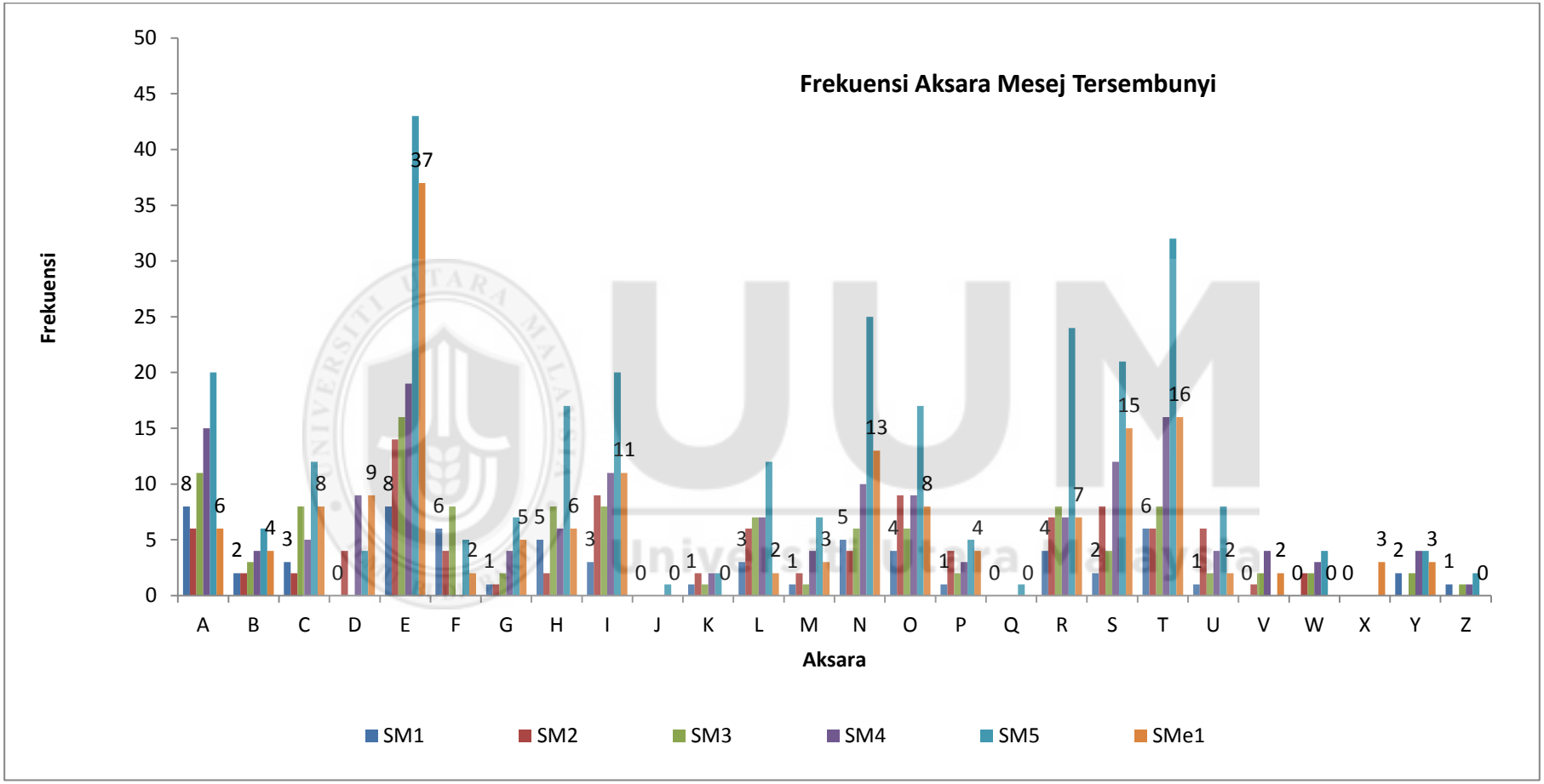
Saiz Fail	Fail Teks Pelindung	Ketiadaan Aksara		Peratus Kewujudan Aksara	Saiz Fail	Fail Teks Pelindung	Ketiadaan Aksara		Peratus Kewujudan Aksara
		Aksara	**%**				**Aksara**	**%**	
1 Kb	CT180	J,Q,X	11.54	88.46	2 Kb	CT887	Q,Z	7.69	92.31
	CT243	J,K,Q,Z	15.38	84.62		CT1086	J,Q,Z	11.54	88.46
	CT376	J,Q,X,Z	15.38	84.62		CT1139	Q	3.85	96.15
	CT412	J,Q	7.69	92.31		CT1311	Q,Z	7.69	92.31
	CT452	J,Z	7.69	92.31		CT1515	-	0.00	100.00
	CT452	Z	3.85	96.15		CT1627	Z	3.85	96.15
	CT464	J,Z	7.69	92.31		CT1669	Q	3.85	96.15
	CT667	Q,Z	7.69	92.31		CT1700	Q,Z	7.69	92.31
	CT793	Z	3.85	96.15		CT1874	-	0.00	100.00
	CT821	-	0.00	100.00		CT1962	Z	3.85	96.15
	CT837	X	3.85	96.15		CT1972	-	0.00	100.00
	Purata (%)		**6.29**	**93.70**		**Purata (%)**		**4.55**	**95.45**
3 Kb	CT1845	-	0.00	100.00	4 Kb	CT2608	-	0.00	100.00
	CT1885	-	0.00	100.00		CT2638	Z	3.85	96.15
	CT2041	-	0.00	100.00		CT2785	Z	3.85	96.15
	CT2150	-	0.00	100.00		CT2796	-	0.00	100.00
	CT2153	-	0.00	100.00		CT2853	-	0.00	100.00
	CT2210	Q	3.85	96.15		CT2858	Z	3.85	96.15
	CT2304	-	0.00	100.00		CT2900	-	0.00	100.00
	CT2358	-	0.00	100.00		CT3056	-	0.00	100.00
	CT2360	-	0.00	100.00		CT3129	-	0.00	100.00
	CT2716	Z	3.85	96.15		CT3267	-	0.00	100.00
						CT3305	-	0.00	100.00
	Purata (%)		**0.77**	**99.23**		**Purata (%)**		**1.05**	**98.95**

Berdasarkan kepada Jadual 5.1 didapati bahawa fail teks pelindung bersaiz 2Kb dan ke bawah mempunyai purata ketiadaan aksara hampir 5.42% berbanding teks pelindung bersaiz lebih dari 3Kb mempunyai purata ketiadaan aksara hanya 0.91%. Bagi fail bersaiz kurang dari 2Kb, didapati bahawa aksara J,Q,X dan Z paling kurang

muncul di dalam fail tersebut. Walau bagaimanapun, aksara Q dan Z masih kurang muncul di dalam fail bersaiz melebihi 3Kb tetapi dengan kadar yang sangat rendah. Secara keseluruhannya, aksara Q dan Z paling kurang muncul di dalam kesemua fail teks pelindung. Sebaliknya, analisis menunjukkan hampir keseluruhan fail teks pelindung mengandungi aksara A-Z (tidak termasuk aksara J,Q,X dan Z) dengan peratus kewujudan aksara melebihi 93%. Secara keseluruhannya, fail yang bersaiz 3Kb ke atas mempunyai peratus kewujudan aksara paling tinggi dengan purata 99%.

5.1.2 Hasil Analisis Mesej Rahsia

Analisis terhadap enam fail mesej rahsia yang dipilih pada sub-topik 3.7 telah dilakukan untuk mengenal pasti taburan aksara yang muncul di dalam fail tersebut. Rajah 5.1 menunjukkan analisis kekerapan aksara terhadap fail mesej rahsia tersebut.



Rajah 5.1 Kekerapan Aksara Dalam Mesej Rahsia

Rajah 5.1 menunjukkan kekerapan kewujudan aksara bagi enam mesej rahsia SM1,SM2,SM3,SM4, SM5 dan SMe1. Berdasarkan kepada Rajah 5.1 didapati bahawa setiap fail mempunyai aksara berulang dengan aksara E merupakan aksara yang paling kerap berulang di dalam kesemua fail dengan kemunculan tertinggi ialah di dalam fail SM5 melebihi 40 kali ulangan. Bagi fail yang bersaiz paling kecil (SM1), didapati masih terdapat aksara berulang iaitu aksara A, B,C, E, F, N, T dan sebagainya. Label data dalam Rajah 5.1 menunjukkan bahawa fail SMe1 mempunyai 21 aksara berulang (kecuali aksara J, K, Q, X dan Z) dengan aksara E merupakan aksara yang paling kerap berulang dengan kekerapan sebanyak 37 kali kemunculan. Aksara T, A, I, N dan R merupakan 5 aksara seterusnya yang paling kerap muncul di dalam kebanyakan dokumen. Sebaliknya, aksara J, K, Q, X dan Z merupakan aksara yang paling kurang muncul di dalam mesej rahsia dengan peratusan paling rendah iaitu kurang dari 1%. Analisis menunjukkan bahawa fail bersaiz kecil masih mempunyai aksara berulang berbanding fail yang bersaiz besar yang mempunyai aksara berulang yang lebih tinggi.

Jadual 5.2

Purata Keberulangan Aksara Mesej Rahsia

Aksara	Purata Aksara Berulang	Aksara	Purata Aksara Berulang
E	22.8%	U	3.8%
T	14.0%	B	3.5%
A	11.0%	G	3.3%
N	10.5%	M	3.0%
I	10.3%	P	3.2%
S	10.3%	Y	2.5%
R	9.5%	W	1.8%
O	8.8%	V	1.5%
H	7.3%	K	1.3%
C	6.3%	Z	0.8%
L	6.2%	X	0.5%
D	4.3%	J	0.2%
F	4.2%	Q	0.2%

Jadual 5.2 menunjukkan purata keberulangan aksara di dalam keenam-enam fail mesej rahsia. Berdasarkan kepada jadual di atas, didapati aksara E merupakan aksara yang paling kerap muncul dengan purata keberulangan sebanyak 20% diikuti dengan aksara T, A, N, I dan S dengan purata keberulangan masing-masing ialah 14.0%, 11.0%, 10.5%, 10.3%, 10.3%. Jadual tersebut juga menunjukkan bahawa aksara J, Q, X dan Z merupakan aksara yang paling kurang muncul di dalam sesuatu dokumen dengan kemunculan masing-masing ialah 0.2%,0.2%, 0.5% dan 0.8%.

Kajian yang dijalankan oleh Ramakrishnan, Thandra, dan Srinivasula (2017) mendapati bahawa, aksara J, Q, X dan Z merupakan aksara yang paling kurang muncul di dalam sesuatu dokumen dengan peratusan paling rendah iaitu kurang dari 0.03% dan kemunculan tertinggi ialah ruang kosong sebanyak 20% dan diikuti aksara E, T, A, O, I, S dengan kadar masing-masing ialah 9.6%, 7.6%, 6.8%, 6.3% , 5.4% dan 5.05%. Oleh itu, mesej rahsia pada Lampiran F dijadikan sebagai data kajian ketika menjalankan pengujian kerana ia selari dengan kajian yang dijalankan oleh Ramakrishnan et al. (2017).

5.2 Padanan Lokasi Jujukan Rawak dan Mesej Rahsia

Proses mendapatkan nilai lokasi rawak dilakukan menggunakan persamaan 4.7 yang bergantung kepada kekunci `a`, `t` dan `m`. Jadual 5.3 menunjukkan beberapa eksperimen yang dijalankan untuk menjana lokasi rawak menggunakan nilai kekunci `a,t,m` dan pelbagai fail teks pelindung (CT) berdasarkan set mesej rahsia C1 dari Jadual 4.1.

Jadual 5.3

Lokasi Jujukan Rawak Untuk Mesej Rahsia

	M	E	E	T	Y	O	U	A	T	T	E	N
v	503	1968	0081	1098	1590	0287	1782	0568	1477	0682	1866	0496
x	2	8	0	4	7	1	7	2	6	3	8	2
y	3	11	5	13	1	4	13	7	8	0	4	3
z	8	3	6	3	0	2	12	13	7	7	6	1
E1	*Lokasi Jujukan Rawak, a = 13, t = 9, saiz Fail CT = 2210*											
	35	464	1627	1297	1421	826	1919	679	8	113	1478	1567
E2	*Lokasi Jujukan Rawak, a = 11, t = 6, saiz Fail CT = 2210*											
	6	72	196	492	798	1248	1369	1718	1729	1823	2162	2163
E3	*Lokasi Jujukan Rawak, a = 11, t = 6, saiz Fail CT = 1885*											
	34	251	1770	1129	400	934	914	774	1673	450	1284	1485
E4	*Lokasi Jujukan Rawak, a = 7, t = 14, saiz Fail CT = 1885*											
	35	259	1827	1529	1322	1752	1004	1405	454	1313	1689	563
E5	*Lokasi Jujukan Rawak, a = 5, t = 5, saiz fail CT = 821*											
	5	30	155	780	1698	1874	547	533	463	113	570	648
E6	*Lokasi Jujukan Rawak, a = 9, t = 7, saiz Fail CT = 821*											
	70	637	63	574	307	337	607	604	577	334	580	12
E7	*Lokasi Jujukan Rawak, a = 9, t = 4, saiz Fail CT = 412*											
	4	40	364	8	76	279	61	144	73	252	227	2
E8	*Lokasi Jujukan Rawak, a = 11, t = 5, saiz Fail CT = 412*											
	5	60	256	367	361	295	387	172	261	13	148	406

Jadual 5.3 menunjukkan fail teks pelindung yang sama dapat menghasilkan jujukan lokasi rawak yang berbeza dan bergantung pada nilai kekunci `a` dan `t`. Pelbagai lokasi rawak boleh dijana menggunakan teks pelindung yang sama, tetapi ia bergantung kepada nlai `a` dan `t`. Contohnya, bagi teks pelindung bersaiz 2210 (E1 dan E2), lokasi jujukan rawak yang dijana untuk mesej rahsia "MEETYOUATTEN" dengan nilai kekunci `a = 13, t = 9` ialah `35,464, 1627,1297,1421,826,1919,8,1478,1567` manakala kekunci `a = 11` dan `t=6` menjana lokasi `6,72,196,492,798,1248,1369,1718,1729,1823, 2162,2163` menggunakan fail yang sama.

Keputusan yang sama juga ditunjukkan dalam beberapa eksperimen lain yang menghasilkan pelbagai lokasi berbeza menggunakan fail pelindung dengan mesej

rahsia yang sama tetapi dengan nilai a dan t yang berbeza seperti yang ditunjukkan dalam eksperimen (E3,E4), (E5,E6) dan (E7,E8). Selain itu, eksperimen juga menunjukkan lokasi rawak yang berbeza dapat dihasilkan bagi mesej yang sama menggunakan fail CT yang berbeza (2210, 1885, 821,412).

Setiap aksara mesej rahsia dipadankan pada lokasi rawak dan diformatkan dengan nilai RGB berdasarkan kepada nilai x, y dan z yang diperolehi. Berdasarkan kepada Jadual 5.3, pada eksperimen E1, aksara "M" yang diwakilkan dengan nilai (2,3,8) disembunyikan pada aksara ke 35 dalam teks pelindung. Merujuk kepada Lampiran E, aksara ke 35 (huruf „h') dalam teks pelindung akan diformatkan dengan nilai RGB (2,3,8) dan proses ini akan berulang sehingga ke aksara terakhir mesej rahsia. Secara kesimpulannya, pelbagai lokasi rawak boleh dijana berdasarkan peraturan berikut;

1. Teks pelindung yang sama tetapi menggunakan nilai kekunci a dan t berbeza.
2. Teks pelindung yang berbeza menggunakan kekunci a dan t yang sama.

5.3 Proses Penyembunyian

Selepas nilai lokasi rawak dikenal pasti, proses penyembunyian dilakukan dengan memformatkan aksara pada lokasi tersebut dengan nilai RGB yang sepadan. Setiap fail teks pelindung akan menjana lokasi nombor rawak yang berbeza bergantung pada nilai a dan t seperti yang dijelaskan pada sub-topik 4.6.

Jadual 5.4

Perwakilan Mesej Rahsia dan Lokasi Jujukan Rawak Berdasarkan Teks Pelindung Bersaiz 180 aksara

Teks Pelindung	Eksperimen	Perwakilan Mesej Rahsia	Lokasi Rawak
CT180 (180 aksara)	E1	SM1 – 28 63 132 128 174 117 65 153 24 161 166 28 98 146 145 167 170 17 163 90 32 173 83 17 160 32 117 72 168 (X,Y,Z): (0,4,3)(0,8,12)(0,8,8)(0,11,9)(0,7,12)(0,4,5)(0,10,3)(0,1,9)(0,10,11)(0,11,1)(0,1,13)(0,6,8)(0,9,11)(0,9,10)(0,11,2)(0,11,5)(0,1,2)(0,10,13)(0,6,0)(0,2,2)(0,11,8)(0,5,8)(0,1,2)(0,10,10)(0,2,2)(0,7,12)(0,4,12)(0,11,3)	Kekunci : a = 2; t = 10; m = 173 : 20 50 110 57 124 85 7 24 58 126 89 15 40 90 17 44 98 33 76 162 161 159 155 147 131 99 35 80 170
	E2	SM1 - 28 151 54 64 122 117 143 66 162 163 154 72 137 82 50 172 132 103 83 90 120 29 102 57 87 120 117 2 152 (X,Y,Z): (0,10,1)(0,3,9)(0,4,4)(0,8,2)(0,7,12)(0,9,8)(0,4,6)(0,10,12)(0,10,13)(0,10,4)(0,4,12)(0,9,2)(0,5,7)(0,3,5)(0,11,7)(0,8,12)(0,6,13)(0,5,8)(0,6,0)(0,8,0)(0,1,14)(0,6,12)(0,3,12)(0,5,12)(0,8,0)(0,7,12)(0,0,2)(0,10,2)	Kekunci : a = 9; t = 10; m = 173 : 55 159 57 4 46 78 20 17 163 93 155 21 26 71 130 142 77 11 109 126 106 99 36 161 75 166 120 52 132
	E3	SM2 - 40 141 84 129 69 21 52 125 89 177 63 71 114 38 107 173 153 97 61 179 145 180 74 53 112 73 176 163 107 84 172 117 42 4 135 121 146 60 97 25 72 (X,Y,Z): (0,9,6)(0,5,9)(0,8,9)(0,4,9)(0,1,6)(0,3,7)(0,8,5)(0,5,14)(0,11,12)(0,4,3)(0,4,11)(0,7,9)(0,2,8)(0,7,2)(0,11,8)(0,10,3)(0,6,7)(0,4,1)(0,11,14)(0,9,10)(0,12,0)(0,4,14)(0,3,8)(0,7,7)(0,4,13)(0,11,11)(0,10,13)(0,7,2)(0,5,9)(0,11,7)(0,7,12)(0,2,12)(0,0,4)(0,9,0)(0,8,1)(0,9,11)(0,4,0)(0,6,7)(0,1,10)(0,4,12)	Kekunci : a = 13; t = 7; m = 173 : 72 78 156 132 166 89 126 88 113 92 165 76 130 140 97 57 56 43 47 99 83 48 112 79 169 128 114 105 161 24 146 2 33 90 139 84 61 108 27 12 163
	E4	SM3 – 67 104 84 28 114 175 88 144 140 173 162 59 4 75 152 131 36 98 107 29 108 137 63 36 151 180 114 76 169 112 52 63 17 77 96 29 163 112 124 130 86 64 60 5 111 125 132 15 142 65 112 91 147 127 156 142 63 100 166 105 154 59 142 38 57 113 24 14 (X,Y,Z): (0,6,14)(0,5,9)(0,1,13)(0,7,9)(0,11,10)(0,5,13)(0,9,9)(0,9,5)(0,11,8)(0,10,12)(0,3,14)(0,0,4)(0,5,0)(0,10,2)(0,8,11)(0,2,6)(0,6,8)(0,7,2)(0,1,14)(0,7,3)(0,9,2)(0,4,3)(0,2,6)(0,10,1)(0,12,0)(0,7,9)(0,5,1)(0,11,4)(0,7,7)(0,3,7)(0,4,3)(0,1,2)(0,5,2)(0,6,6)(0,1,14)(0,10,13)(0,7,7)(0,8,4)(0,8,10)(0,5,11)(0,4,4)(0,4,0)(0,0,5)(0,7,6)(0,8,5)(0,8,12)(0,1,0)(0,9,7)(0,4,5)(0,7,7)(0,6,1)(0,9,12)(0,8,7)(0,10,6)(0,9,7)(0,4,3)(0,6,10)(0,11,1)(0,7,0)(0,10,4)(0,3,14)(0,9,7)(0,2,8)(0,3,12)(0,7,8)(0,1,9)(0,0,14)	Kekunci : a = 3; t = 11; m = 173 : 26 89 105 153 124 37 122 31 104 150 115 10 41 134 67 39 128 49 158 139 82 84 90 108 162 151 118 19 68 42 137 76 66 36 119 22 77 69 45 146 103 147 106 156 133 64 30 101 141 88 102 144 97 129 52 167 166 163 154 127 46 149 112 1 14 53 170 2

Jadual 5.4 di atas menunjukkan pengujian yang dijalankan ke atas tiga mesej rahsia yang berbeza saiz SM1, SM2 dan SM3 terhadap teks pelindung yang bersaiz 180

aksara. Eksperimen E1 dan E2 menunjukkan mesej rahsia yang sama (SM1) dapat diwakilkan dengan perwakilan yang berbeza dan disembunyikan pada fail yang sama tetapi pada lokasi yang berbeza (Lampiran H). Selain itu, eksperimen juga menunjukkan mesej rahsia yang pelbagai dapat disembunyikan pada teks pelindung yang sama (CT180) tetapi pada lokasi yang berbeza seperti yang ditunjukkan di dalam Eksperimen E2, E3, E4.

Sesuatu mesej rahsia boleh diwakilkan dengan pelbagai nilai berdasarkan kepada teks pelindung atau mempunyai hubungan satu-ke-banyak (*one-to-many*). Selain itu, pelbagai perwakilan mesej rahsia boleh dihasilkan menggunakan pelbagai teks pelindung atau dikatakan mempunyai hubungan banyak-ke-banyak (*many-to-many*). Hubungan kedua-duanya dapat ditunjukkan di dalam Jadual 5.3 dan Jadual 5.4. Kesimpulannya sesuatu mesej rahsia dapat diwakilkan dengan pelbagai perwakilan dan boleh disembunyikan pada pelbagai lokasi menggunakan teks pelindung yang sama.

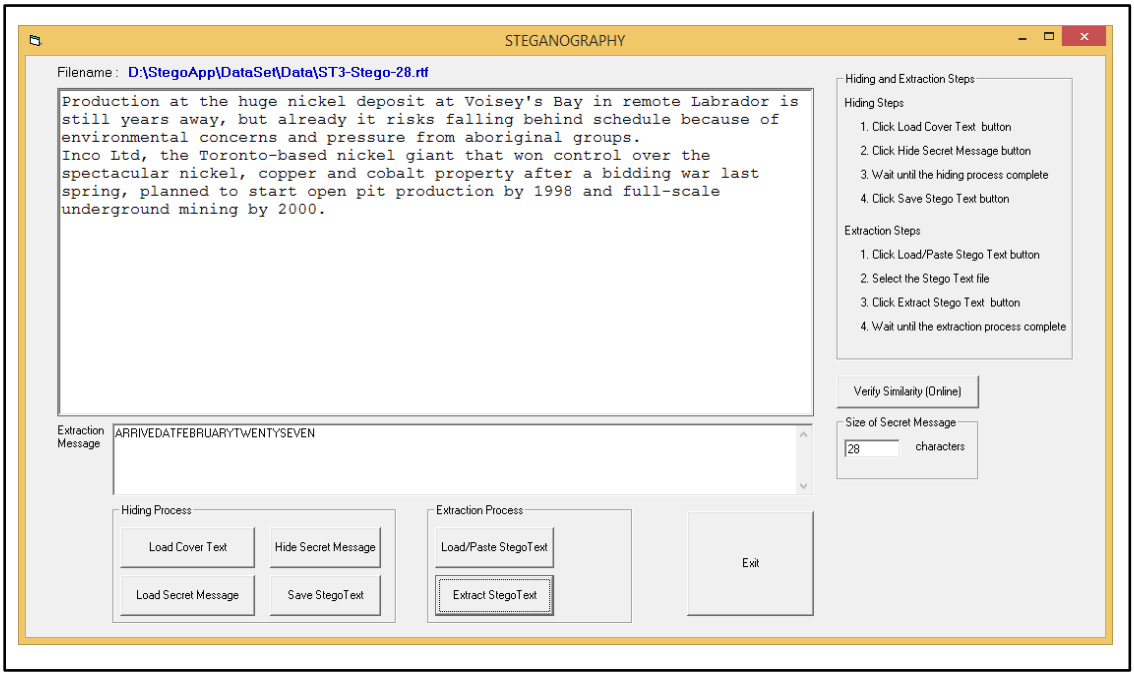
5.4 Proses Pengekstrakan

Mesej rahsia yang disembunyikan perlu di ekstrak semula untuk memastikan mesej yang disembunyikan dapat diperolehi kembali. Langkah pertama dilakukan dengan mengenal pasti nilai `a` dan `t` terlebih dahulu. Nilai-nilai ini diperolehi dari fail teks stego (*.rtf*) yang dihasilkan dengan mengekstrak nilai RGB pada lokasi saiz(*rtf*)-1 dan saiz(*rtf*)-2 fail tersebut. Proses pengekstrakan bermula dengan menggantikan nilai `a` dan `t` ke dalam persamaan 4.7 untuk mendapatkan lokasi permulaan penyembunyian. Aksara pertama pada lokasi yang diperolehi akan diekstrak warna RGB dan ditukar ke nilai asal menggunakan persamaan 4.6. Nilai yang diperolehi akan dipadankan dengan

Jadual *Homophonic* sifer untuk mendapatkan aksara sebenar dan proses ini akan berulang sehingga tamat.

5.5 Antara Muka Sistem Penyembunyian Dan Pengekstrakan

Antara muka Sistem Penyembunyian dan Pengekstrakan dibangunkan menggunakan Perisian Microsoft Visual Basic 6.0. Di samping itu, bahasa Java digunakan untuk menjana lokasi penyembunyian rawak berdasarkan kepada aksara mesej rahsia dan nilai `a,b` dan `t`. Proses penyembunyian dan pengekstrakan dilakukan menggunakan antara muka seperti ditunjukkan pada Rajah 5.2 selepas proses penjanaan lokasi rawak berjaya dilakukan. Teks pelindung dimuatkan ke dalam sistem dengan menekan butang "*Load Cover Text*" dan ikuti butang "*Load Secret Text*". Seterusnya, proses penyembunyian dilakukan dengan menekan butang "*Hide Secret Message*" dan "*Save Stego Text*". Fail teks stego yang telah rerjaya dijana akan disimpan dalam format *.rtf (rich text file)* dan dapat dibuka menggunakan perisian Microsoft Word. Proses pengekstrakan dilakukan dengan memuatkan fail teks stego (*.rtf*) yang telah dijana dan seterusnya menekan butang "*Extract Stego Text*" untuk memaparkan mesej yang disembunyikan.



Rajah 5.2 Antara muka untuk Proses Penyembunyian dan Pengekstrakan

Rajah 5.2 menunjukkan mesej rahsia "ARRIVEATFEBRUARYTWENTYSEVEN" telah diekstrakkan dari fail teks stego CT376. Pengujian menunjukkan proses pengekstrakan berjaya mengekstrak semula 100 peratus mesej yang disembunyikan di dalam teks stego yang dijana. Rajah 5.3 menunjukkan sampel fail teks pelindung manakala Rajah 5.4 menunjukkan fail teks stego yang dihasilkan selepas melalui proses penyembunyian mesej di atas.

```
Production at the huge nickel deposit at Voisey's Bay in
remote Labrador is still years away, but already it risks
falling behind schedule because of environmental concerns
and pressure from aboriginal groups.Inco Ltd, the Toronto-
based nickel giant that won control over the spectacular
nickel, copper and cobalt property after a bidding war last
spring, planned to start open pit production by 1998 and
full-scale underground mining by 2000.
```

Rajah 5.3 Fail Teks Pelindung

```
Production at the huge nickel deposit at Voisey's Bay in
remote Labrador is still years away, but already it risks
falling behind schedule because of environmental concerns
and pressure from aboriginal groups.Inco Ltd, the Toronto-
based nickel giant that won control over the spectacular
nickel, copper and cobalt property after a bidding war
last spring, planned to start open pit production by 1998
and full-scale underground mining by 2000.
```

Rajah 5.4 Fail Teks Stego

Berdasarkan teks stego yang dihasilkan pada Rajah 5.4, didapati bahawa tiada sebarang perbezaan yang jelas berlaku di antara teks pelindung dan teks stego. Selain itu fail teks stego yang dihasilkan juga menunjukkan tiada sebarang perubahan berlaku terhadap isikandungan fail tersebut.

5.6 Hasil Penyembunyian dan Pengekstrakan

Proses penyembunyian telah dilakukan terhadap tujuh teks pelindung pelbagai saiz menggunakan pelbagai saiz mesej rahsia seperti yang ditunjukkan di dalam Jadual 5.5. Beberapa saiz teks tersembunyi tambahan telah digunakan bertujuan untuk menentukan kapasiti penyembunyian tertinggi bagi setiap teks pelindung.

Jadual 5.5

Hasil Penyembunyian Dan Pengekstrakan Pelbagai Saiz Mesej Rahsia dan Teks Pelindung

Teks Pelindung (Aksara)	Mesej Rahsia	Fail Teks Stego	Kejayaan Penyembunyian Mesej	Kejayaan Pengekstrakan Mesej	Aksara Yang Gagal Disembunyikan
CT180 (1Kb)	SM28	ST-180-28	100%	100%	
	SM40	ST-180-40	100%	100%	
	SM67	ST-180-67	100%	100%	
	SM116 *	ST-180-116	100%	100%	

	SM267				
	SM531				
	SM834				
	SM28	ST-243-28	100%	100%	
	SM40	ST-243-40	100%	100%	
	SM67	ST-243-67	97.01%	97.01%	K,Z
CT243	SM116*	ST-243-116	98.27%	98.27%	K,Z
(1Kb)	SM183*	ST-243-183	100%	100%	
	SM267		-	-	
	SM531		-	-	
	SM834		-	-	
	SM28	ST-376-28	100%	100%	
	SM40	ST-376-40	100%	100%	
	SM67	ST-376-67	98.51%	98.51%	Z
CT376	SM267	ST-376-267	98.88%	98.88%	J,Q,Z
(1Kb)	SM283*	ST-376-238	100%	100%	
	SM531		-	-	
	SM834		-	-	
	SM28	ST-837-28	100%	100%	
	SM40	ST-837-40	100%	100%	
	SM67	ST-837-67	100%	100%	
CT837	SM267	ST-837-267	100%	100%	
(1Kb)	SM531	ST-837-531	99.81%	99.81%	X
	SM660*	ST-837-660	100%	100%	
	SM834	ST-837-834	-	-	
	SM28	ST-1700-28	100%	100%	
	SM40	ST-1700-40	100%	100%	
	SM67	ST-1700-67	98.51%	98.51%	Z
CT1700	SM267	ST-1700-267	98.88%	98.88%	Z(2),Q
(2Kb)	SM531	ST-1700-531	100%	100%	
	SM834	ST-1700-834	100%	100%	
	SM1262*	ST-1700-1262	100%	100%	
	SM28	ST-2153-28	100%	100%	
	SM40	ST-2153-40	100%	100%	
	SM67	ST-2153-67	100%	100%	
CT2153*	SM267	ST-2153-267	100%	100%	
(3Kb)	SM531	ST-2153-531	100%	100%	
	SM834	ST-2153-834	100%	100%	
	SM1262*	ST-2153-1262	100%	100%	
	SM1958	ST-2153-1958	100%	100%	
	SM28	ST-2638-28	100%	100%	
	SM40	ST-2638-40	100%	100%	
	SM67	ST-2638-67	98.51%	98.51%	Z
CT2638	SM267	ST-2638-267	99.25%	99.25%	Z (2)
(4Kb)	SM531	ST-2638-531	100%	100%	
	SM834	ST-2638-834	100%	100%	
	SM1262*	ST-2638-1262	100%	100%	
	SM1958*	ST-2638-1958	100%	100%	
	SM2185*	ST-2638-2185	100%	100%	

*Teks pelindung dan mesej rahsia tambahan

Jadual 5.5 menunjukkan kepelbagaian mesej rahsia yang disembunyikan di dalam pelbagai teks pelindung. Berdasarkan kepada Jadual 5.5, hampir 80% mesej yang disembunyikan berjaya diekstrakkan semula sepenuhnya. Baki sebanyak 20% gagal diekstrakkan semula disebabkan ketiadaan aksara di dalam teks pelindung. Walau bagaimanapun, mesej rahsia yang gagal disembunyikan sepenuhnya di dalam teks pelindung tertentu masih dapat disembunyikan menggunakan teks pelindung yang berlainan. Berdasarkan kepada Jadual 5.5, didapati beberapa tidak wujud di dalam beberapa teks pelindung yang menyebabkan berlakunya kegagalan proses penyembunyian. Sebagai contohnya, walaupun mesej rahsia dalam SM67 gagal disembunyikan di dalam teks pelindung CT243, CT376 ,CT1700 dan CT2368 disebabkan ketakwujudan aksara Z dalam fail tersebut, namun mesej tersebut masih boleh disembunyikan menggunakan teks pelindung CT180, CT837 dan CT2153. Secara amnya, kesemua fail teks stego yang berjaya diekstrakkan sepenuhnya diringkaskan di dalam Jadual 5.6.

Jadual 5.6

Fail Teks Stego Yang Berjaya Di ekstrak Sepenuhnya

Fail Teks Pelindung	**Mesej Rahsia**	**Fail Teks Stego**
CT180	SM28	ST-180-28
	SM40	ST-180-40
	SM67	ST-180-67
	SM116 *	ST-180-116
CT243	SM28	ST-243-28
	SM40	ST-243-40
	SM183*	ST-243-183
CT376	SM28	ST-376-28
	SM40	ST-376-40
	SM283*	ST-376-283
CT837	SM28	ST-837-28
	SM40	ST-837-40
	SM67	ST-837-67

	SM267	ST-837-267
	SM660*	ST-837-660
CT1700	SM28	ST-1700-28
	SM40	ST-1700-40
	SM531	ST-1700-531
	SM834	ST-1700-834
	SM1262*	ST-1700-1262
CT2153	SM28	ST-2153-28
	SM40	ST-2153-40
	SM67	ST-2153-67
	SM267	ST-2153-267
	SM531	ST-2153-531
	SM834	ST-2153-834
	SM1262*	ST-2153-1262
	SM1958	ST-2153-1958
CT2638	SM28	ST-2638-28
	SM40	ST-2638-40
	SM531	ST-2638-531
	SM834	ST-2638-834
	SM1262*	ST-2638-1262
	SM1958*	ST-2638-1958
	SM2185*	ST-2638-2185

Berdasarkan kepada Jadual 5.6, sebanyak 35 fail teks stego berjaya diekstrakkan 100 peratus mesej rahsia. Justeru itu, fail-fail pada Jadual 5.6 digunakan untuk proses penyembunyian dan perbincangan seterusnya.

5.7 Pelbagai Perwakilan Nilai Rawak Bagi Mesej Rahsia

Mesej yang disembunyikan diwakilkan dengan nilai rawak sebelum ditukarkan kepada bentuk (*x,y,z*). Nilai rawak bagi mesej rahsia dihasilkan berdasarkan kepada Jadual *Homophonic* seperti yang dijelaskan dalam Bab 2. Jadual 5.7 di bawah menunjukkan mesej rahsia "`ArrivedAtFebruaryTwentySeven`" diwakilkan dengan pelbagai nilai rawak serta dipadankan dengan pelbagai lokasi rawak yang dihasilkan berdasarkan teks pelindung bersaiz 180 aksara.

Jadual 5.7

Kepelbagaian Lokasi Rawak Bagi Mesej Rahsia Yang Sama

Eksperimen	SM/ Lokasi Rawak	Perwakilan Rawak SM1 dan Lokasi Penyembunyian Rawak
E1	SM1	63 14 82 37 117 76 133 176 92 111 129 98 54 157 107 64 103 19 90 118 29 163 57 22 159 117 143 168
	Lokasi Rawak	: a = 10; t = 4; m = 173 54 25 81 122 13 134 133 123 23 61 95 89 29 121 3 34 171 157 17 1 14 144 60 85 162 67 155 170 147
E2	SM1	63 14 54 69 117 72 73 131 139 111 60 137 147 155 167 54 57 19 80 159 8 55 57 58 143 117 96 152
	Lokasi Rawak	: a = 12; t = 6; m = 173 66 106 67 118 38 116 14 1 18 49 75 41 152 100 168 119 50 87 12 150 76 53 123 98 144 4 54 135 69
E3	SM1	107 89 132 44 117 81 73 176 177 166 60 98 14 155 107 132 57 92 90 143 168 26 57 87 65 117 67 173
	Lokasi Rawak	: a = 14; t = 7; m = 173 77 47 146 148 3 49 1 21 128 69 108 135 167 96 140 64 38 20 114 46 132 125 27 39 34 137 22 142 92
E4	SM1	: 172 14 54 134 117 113 1 63 55 154 164 98 14 145 63 64 17 94 80 113 180 15 17 39 110 117 67 173
	Lokasi Rawak	: a = 6; t = 9; m = 173 39 70 83 161 110 150 44 100 90 30 16 105 120 37 58 11 75 113 168 152 56 172 3 27 171 170 164 128 85
E5	SM1	24 146 132 99 117 32 66 131 115 111 159 137 146 50 24 14 103 163 80 88 168 102 103 49 96 117 42 8
	Lokasi Rawak	: a = 2; t = 10; m = 173 20 50 110 57 124 85 7 24 58 126 89 15 40 90 17 44 98 33 76 162 161 159 155 147 131 99 35 80 170

Berdasarkan Jadual 5.7, eksperimen E1 hingga E5 menunjukkan mesej rahsia SM1 diwakilkan dengan pelbagai nilai serta dapat disembunyikan di dalam pelbagai lokasi rawak menggunakan teks pelindung yang sama. Pelbagai lokasi rawak dijana berdasarkan kepada nilai `a` dan `t` yang berbeza menggunakan teks pelindung yang sama. Ini membuktikan bahawa mesej rahsia boleh disembunyikan di dalam pelbagai lokasi yang berbeza menggunakan teks pelindung yang sama.

5.8 Kapasiti Penyembunyian dan Kadar Ralat Bit (BER)

Mesej rahsia yang gagal di ekstrak ditentukan menggunakan ukuran yang dikenali sebagai *Bit Error Rate* (BER). BER merupakan nisbah (peratusan) di antara jumlah bit yang tidak dapat di ekstrak berbanding dengan jumlah bit mesej rahsia (Abuadbba, Khalil, Ibaida, & Atiquzzaman, 2015; RaSMan et al., 2017). BER boleh ditakrifkan seperti formula pada persamaan 5.1.

$$BER = A = \frac{B_{err}}{B_{total}} \times 100\% \quad (5.1)$$

Semakin kecil nilai BER yang menghampiri sifar, maka pembetulannya semakin baik. Nilai BER = 0 menandakan tiada sebarang kehilangan mesej berlaku selepas proses pengekstrakan seperti di dalam kajian yang dilakukan oleh Rahman et al. (2017) ke atas enam set mesej rahsia. Berdasarkan Jadual 5.6, didapati terdapat beberapa fail teks stego yang tidak dapat di ekstrak sepenuhnya disebabkan ketakwujudan aksara tersebut di dalam teks pelindung. Oleh itu, nilai *BER* dapat dikira berdasarkan persamaan 6.1. Jadual 5.8 menunjukkan nilai *BER* dan kapasiti penyembunyian terhadap eksperimen yang dijalankan.

Jadual 5.8

Kapasiti Penyembunyian Dan Bit Error Rate (BER)

Teks Pelindung (Aksara)	SM (Aksara)	Pengekstrakan Mesej (Teks Stego)	Kapasiti Penyembunyian	BER
CT180* (1kb)	ST-180-28	100%	15.56%	0.00%
	ST-180-40	100%	22.22%	0.00%
	ST-180-67	100%	37.22%	0.00%
	ST-180-116	100%	64.44%	0.00%
	ST-180-267	-	-	-
	ST-180-531	-	-	-

	ST-180-834	-	-	
CT243 (1Kb)	ST-243-28	100%	11.5%	0.00%
	ST-243-40	100%	16.5%	0.00%
	ST-243-67	97.01%	27.6%	2.99%
	ST-243-116	98.27%	47.7%	1.73%
	ST-243-183	100%	75.3%	0.00%
	ST-243-267	-	-	-
	ST-243-531	-	-	-
	ST-243-834	-	-	-
CT376 (1Kb)	ST-376-28	100%	7.4%	0.00%
	ST-376-40	100%	10.6%	0.00%
	ST-376-67	98.51%	17.8%	1.49%
	ST-376-267	98.88%	71.0%	1.12%
	ST-376-283	100%	75.3%	0.00%
	ST-376-531	-	-	-
	ST-376-834	-	-	-
CT837 (1Kb)	ST-837-28	100%	3.3%	0.00%
	ST-837-40	100%	4.8%	0.00%
	ST-837-67	100%	8.0%	0.00%
	ST-837-267	100%	31.9%	0.00%
	ST-837-531	99.81%	63.4%	0.19%
	ST-837-660	100%	78.9%	0.00%
	ST-837-834	-	-	-
CT1700 (2Kb)	ST-1700-28	100%	1.6%	0.00%
	ST-1700-40	100%	2.4%	0.00%
	ST-1700-67	98.51%	3.9%	1.49%
	ST-1700-267	98.88%	15.7%	1.12%
	ST-1700-531	100%	31.2%	0.00%
	ST-1700-834	100%	49.1%	0.00%
	ST-1700-1262	100%	74.2%	0.00%
CT2153* (3Kb)	ST-2153-28	100%	1.3%	0.00%
	ST-1700-40	100%	1.9%	0.00%
	ST-1700-67	100%	3.1%	0.00%
	ST-1700-267	100%	12.4%	0.00%
	ST-1700-531	100%	24.7%	0.00%
	ST-1700-834	100%	38.7%	0.00%
	ST-1700-1262	100%	58.6%	0.00%
	ST-1700-1958	100%	90.9%	0.00%
CT2638 (4Kb)	ST-2638-28	100%	1.1%	0.00%
	ST-2638-40	100%	1.5%	0.00%
	ST-2638-67	98.51%	2.5%	1.49%

ST-2638-267	99.25%	10.1%	0.75%
ST-2638-531	100%	20.1%	0.00%
ST-2638-834	100%	31.6%	0.00%
ST-2638-1262	100%	47.8%	0.00%
ST-2638-1958	100%	74.2%	0.00%
ST-2638-2185	100%	82.8%	0.00%

Kapasiti penyembunyian tertinggi dalam sesuatu teks pelindung yang berjaya di ekstrak semula sepenuhnya merupakan kapasiti penyembunyian maksimum mesej rahsia yang boleh disembunyikan dalam sesuatu teks pelindung. Jadual 5.9 menunjukkan kapasiti tertinggi penyembunyian bagi setiap fail teks pelindung.

Jadual 5.9

Kapasiti Penyembunyian Maksimum

Teks Pelindung (Bait)	SM Maksimum (Bait)	Kapasiti Penyembunyian (%)
CT180	116	64.4%
CT243	183	75.3%
CT376	283	75.3%
CT837	660	78.9%
CT1700	1262	74.2%
CT2153	1958	90.9%
CT2638	2185	82.8%
	Purata	**77.4%**

Jadual 5.9 menunjukkan peratus kapasiti penyembunyian mesej rahsia yang maksimum dapat disembunyikan bagi setiap teks pelindung. Peratus kapasiti penyembunyian terendah ialah 64.4%, manakala peratus kapasiti penyembunyian tertinggi ialah 90.9% menggunakan fail masing-masing CT180 dan CT2153. Secara keseluruhannya, lebih 60% mesej rahsia dapat disembunyikan dalam teks stego yang dihasilkan dengan purata peratus penyembunyian ialah 77.4%.

Beberapa data kajian lepas telah dikumpulkan bagi mengukur prestasi kapasiti penyembunyian kajian lepas berbanding dengan kajian yang dijalankan. Jadual 5.10 menunjukkan peratus kapasiti penyembunyian kajian dijalankan berbanding kajian lepas yang menggunakan teknik warna RGB dan sebaliknya

Jadual 5.10

Peratus Kapasiti Penyembunyian Kajian Yang Dijalankan Berbanding Kajian Lepas

Kajian Lepas		Kajian Yang Dijalankan		Peratus Peningkatan Kapasiti (%)
Saiz Teks Pelindung	Kapasiti %	Saiz teks Pelindung	Kapasiti %	
CT202	25.50	CT180	64.4	38.90
CT220	11.81	CT243	75.3	63.49
CT336	50.00	CT376	75.3	25.30
CT847	18.34	CT837	78.9	60.56
CTX1779*	28.11	CT1700	74.2	46.09
CT2252	95.60	CT2153	90.9	-4.70
CTX2640*	23.25	CT2638	82.8	59.55
Purata	**36.09**	**Purata**	**77.40**	**41.31**

*Tidak berasaskan warna RGB

Berdasarkan kepada Jadual 5.10, secara purata kapasiti penyembunyian meningkat sebanyak 41.31% berbanding dengan kajian lepas. Ini menunjukkan teknik yang digunakan berjaya meningkatkan prestasi kapasiti penyembunyian dengan kapasiti penyembunyian tertinggi ialah 90.9%. Rajah 5.5 menunjukkan kapasiti penyembunyian menggunakan warna RGB yang dijalankan oleh penyelidikan lepas berbanding dengan kajian yang dijalankan.

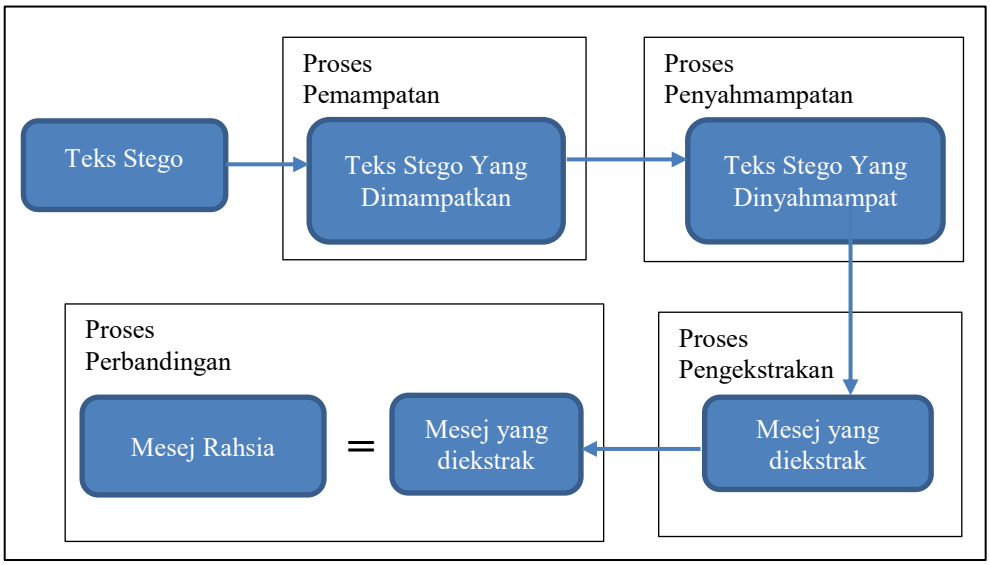
Rajah 5.5 Kapasiti Penyembunyian Teknik Warna RGB Yang Dijalankan Berbanding Teknik Warna RGB Kajian Lepas

Kesimpulannya, kapasiti penyembunyian menggunakan teknik warna RGB yang dicadangkan berjaya melakukan penyembunyian dengan purata penyembunyian 77.4% berbanding teknik RGB terkini yang dijalankan oleh Al-Azzawi, (2018) dengan kapasiti penyembunyian hanya 25.5% seperti yang ditunjukkan di dalam Rajah 5.5.

5.9 Keteguhan

Teks stego yang dihasilkan telah melalui proses pemampatan dan penyahmampatan bertujuan untuk menentukan keteguhan sesuatu teks stego yang dihasilkan. Teks stego yang dijana akan dimampatkan dan dinyahmampatkan menggunakan teknik pemampatan *lossless* yang sesuai digunakan untuk data berbentuk teks. Teks stego ini seterusnya di ekstrak semula untuk memastikan mesej yang disembunyikan dapat di ekstrak semula sepenuhnya. Penggunaan warna RGB Rajah 5.6 menunjukkan proses

menilai keteguhan teks stego yang melalui tiga peringkat iaitu pemampatan, penyahmampatan dan pengekstrakan semula terhadap teks stego yang dinyahmampat.



Rajah 5.6 Proses Menilai Keteguhan Teks Stego

Terdapat dua jenis teknik pemampatan iaitu *lossy* dan *lossless* di mana teknik *lossy* digunakan untuk pemampatan imej tetapi tidak untuk teks manakala teknik *lossless* melibatkan teks (Kavitha, 2016). Teknik penyembunyian mesej rahsia yang digunakan dalam kajian ini berdasarkan kepada teks warna RGB. Oleh itu teknik pemampatan *lossless* digunakan disebabkan ia melibatkan teks. Kesemua saiz fail teks stego yang melalui proses pemampatan dan penyahmampatan direkodkan seperti ditunjukkan di dalam Jadual 5.11. Peratus pengekstrakan direkodkan bertujuan untuk memastikan mesej yang disembunyikan dapat diekstrakan semula sepenuhnya atau sebaliknya. Nilai pengekstrakan 100% menunjukkan mesej yang disembunyikan dapat di ekstrak semula tanpa sebarang kehilangan data.

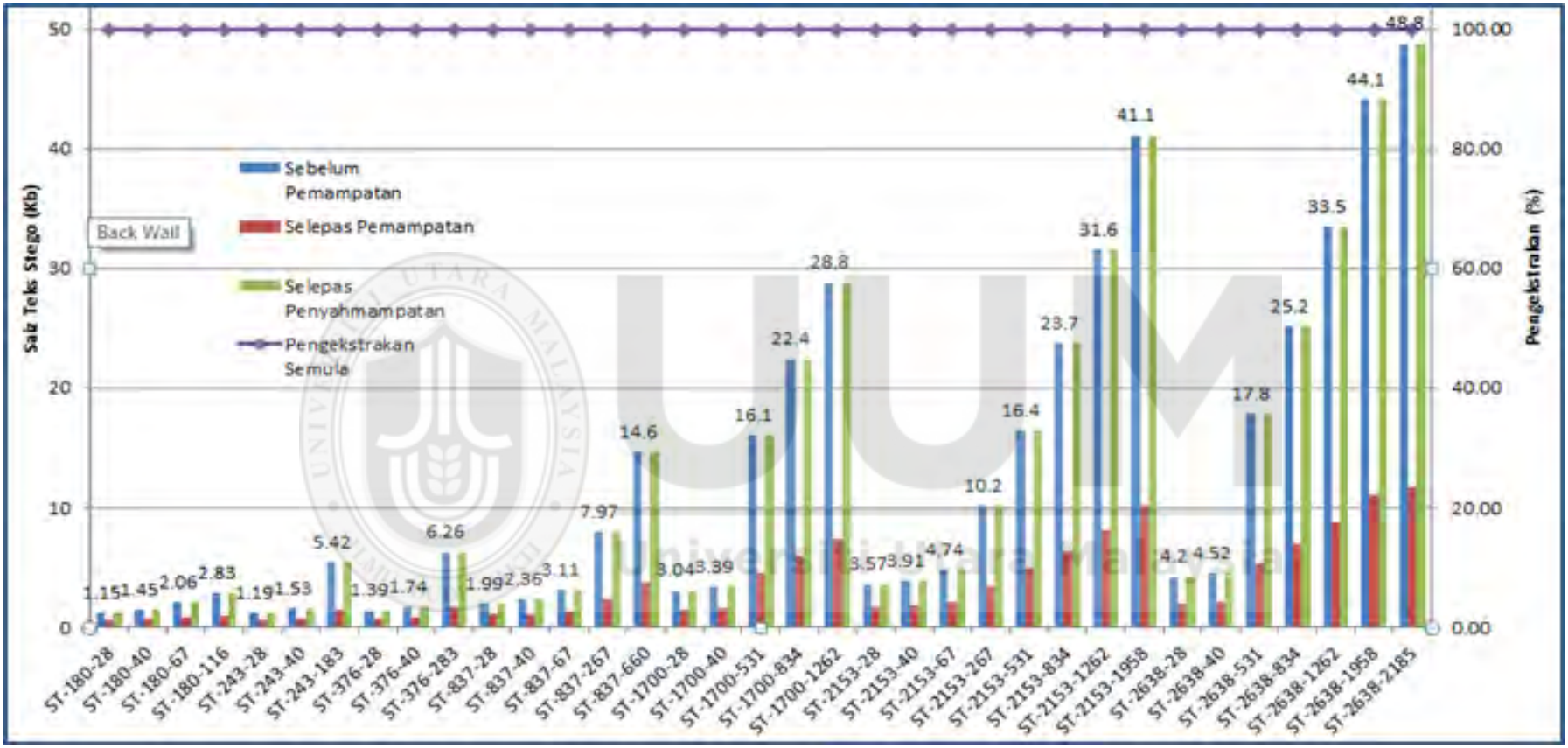
Jadual 5.11

Pengekstrakan Mesej Rahsia Selepas Pemampatan

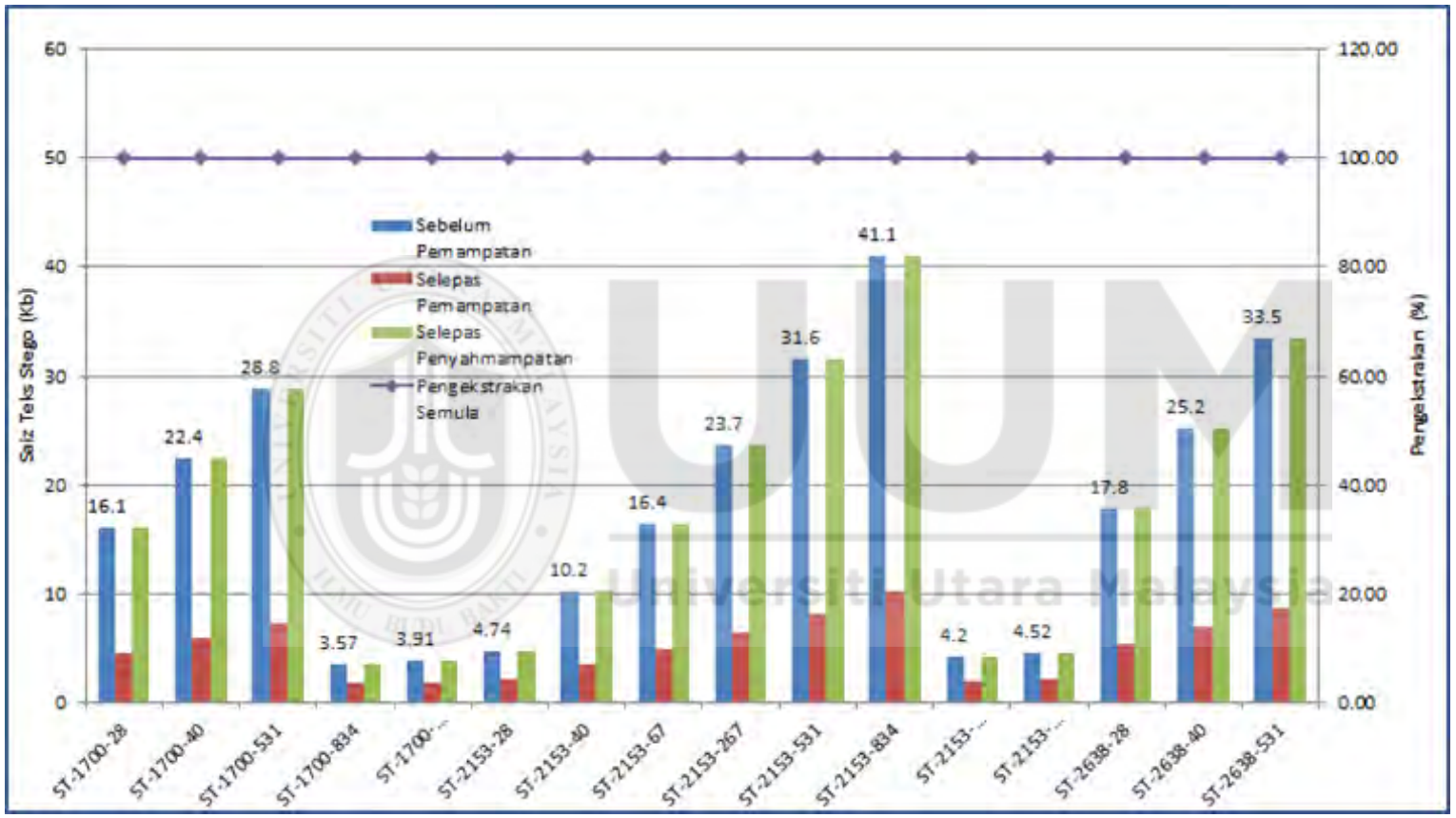
Teks Stego	**Saiz Fail Teks Stego**			Pengekstrakan Mesej Rahsia selepas Penyahmampatan
	Sebelum Pemampatan (KB)	Selepas Pemampatan (KB)	Selepas Penyahmampatan (KB)	
ST-180-28	1.15	0.5556	1.15	100%
ST-180-40	1.45	0.6328	1.45	100%
ST-180-67	2.06	0.7617	2.06	100%
ST-180-116	2.83	0.9092	2.83	100%
ST-243-28	1.19	0.6221	1.19	100%
ST-243-40	1.53	0.6875	1.53	100%
ST-243-183	5.42	1.47	5.42	100%
ST-376-28	1.39	0.7275	1.39	100%
ST-376-40	1.74	0.8232	1.74	100%
ST-376-283	6.26	1.73	6.26	100%
ST-837-28	1.99	1.05	1.99	100%
ST-837-40	2.36	1.14	2.36	100%
ST-837-67	3.11	1.33	3.11	100%
ST-837-267	7.97	2.37	7.97	100%
ST-837-660	14.6	3.73	14.6	100%
ST-1700-28	3.04	1.50	3.04	100%
ST-1700-40	3.39	1.61	3.39	100%
ST-1700-531	16.1	4.57	16.1	100%
ST-1700-834	22.4	5.89	22.4	100%
ST-1700-1262	28.8	7.30	28.8	100%
ST-2153-28	3.57	1.72	3.57	100%
ST-2153-40	3.91	1.83	3.91	100%
ST-2153-67	4.74	2.06	4.74	100%
ST-2153-267	10.2	3.42	10.2	100%
ST-2153-531	16.4	4.91	16.4	100%
ST-2153-834	23.7	6.39	23.7	100%
ST-2153-1262	31.6	8.15	31.6	100%
ST-2153-1958	41.1	10.2	41.1	100%
ST-2638-28	4.20	1.98	4.20	100%
ST-2638-40	4.52	2.11	4.52	100%
ST-2638-531	17.8	5.34	17.8	100%
ST-2638-834	25.2	6.95	25.2	100%
ST-2638-1262	33.5	8.72	33.5	100%
ST-2638-1958	44.1	11.1	44.1	100%
ST-2638-2185	48.8	11.7	48.8	100%

Jadual 5.11 menunjukkan saiz fail teks stego sebelum dan selepas pemampatan serta selepas penyahmampatan. Selepas proses pemampatan, saiz fail teks stego menjadi semakin kecil berbanding fail teks stego sebelum pemampatan. Walau bagaimanapun , selepas proses penyahmampatan, saiz teks stego kembali normal dan kesemua mesej

rahsia dapat di ekstrak semula 100% tanpa sebarang kehilangan data. Ini menunjukkan bahawa teks stego yang dihasilkan teguh terhadap proses pemampatan seperti yang ditunjukkan di dalam Rajah 5.7 (a) manakala Rajah 5.7 (b) menunjukkan keratan sebahagian daripada Rajah 5.7(a).



(a) Keseluruhan Teks Stego



(b) Pembesaran Sebahagian Teks Stego

Rajah 5.7 Perbandingan Sebelum/Selepas Pemampatan, Selepas Penyahmampatan dan Pengekstrakan Semula

Rajah 5.7 menunjukkan saiz teks stego yang dihasilkan selepas melalui proses pemampatan dan penyahmampatan. Rajah 5.7(b) menunjukkan sebahagian keratan daripada Rajah 5.7(a) dan didapati bahawa tiada perbezaan berlaku terhadap saiz fail teks stego yang dihasilkan sebelum pemampatan dan selepas penyahmampatan. Kesemua mesej rahsia yang disembunyikan berjaya di ekstrak semula sepenuhnya selepas melalui proses pemampatan dan penyahmampatan. Sebagai contoh, fail ST-1700-531 sebelum dan selepas penyahmampatan mempunyai saiz yang sama iaitu 28.8Kb. Fail teks stego yang telah melalui proses penyahmampatan menunjukkan tidak berlaku sebarang kehilangan data dan ini menunjukkan bahawa keteguhan teks stego yang dihasilkan dapat dibuktikan.

5.10 Ketakbolehkelihatan

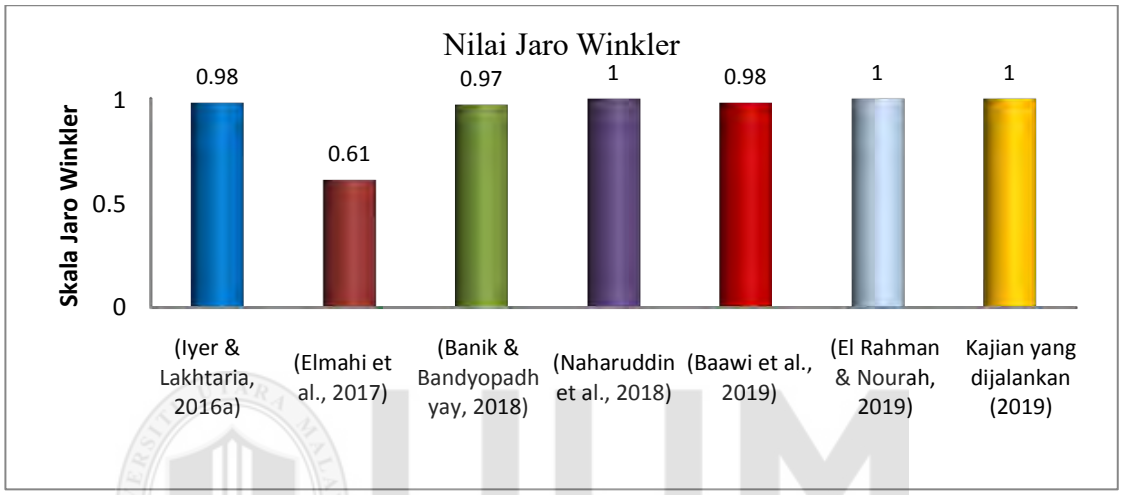
Skor Jaro Winkler digunakan untuk membanding kesamaan antara teks pelindung dan teks stego yang dihasilkan. Eksperimen seterusnya dijalankan terhadap semua fail yang berjaya menyembunyikan 100% mesej rahsia berdasarkan kepada Jadual 5.11. Skor Jaro Winkler yang dihasilkan dalam kajian ini ditunjukkan di dalam Jadual 5.12 di bawah.

Jadual 5.12

Skor Jaro Winkler Bagi Teks Stego yang Dihasilkan

Teks Pelindung (Aksara)	SM (Aksara)	Fail Teks Stego	Nilai skor Jaro Winkler
CT180	28	ST-180-28	1
	40	ST-180-40	1
	67	ST-180-67	1
	116 *	ST-180-116	1
CT243	28	ST-243-28	1
	40	ST-243-40	1
	183*	ST-243-183	1
CT376	28	ST-376-28	1
	40	ST-376-40	1
	283*	ST-376-283	1
CT837	28	ST-837-28	1
	40	ST-837-40	1
	67	ST-837-67	1
	267	ST-837-267	1
	660*	ST-837-660	1
CT1700	28	ST-1700-28	1
	40	ST-1700-40	1
	531	ST-1700-531	1
	834	ST-1700-834	1
	1262*	ST-1700-1262	1
CT2153	28	ST-2153-28	1
	40	ST-2153-40	1
	67	ST-2153-67	1
	267	ST-2153-267	1
	531	ST-2153-531	1
	834	ST-2153-834	1
	1262*	ST-2153-1262	1
	1958*	ST-2153-1958	1
CT2638	28	ST-2638-28	1
	40	ST-2638-40	1
	531	ST-2638-531	1
	834	ST-2638-834	1
	1262*	ST-2638-1262	1
	1958*	ST-2638-1958	1
	2185*	ST-2638-2185	1

Jadual 5.12 menunjukkan skor Jaro Winkler bagi semua teks stego yang dihasilkan bersamaan dengan 1. Nilai skor ini menunjukkan bahawa tiada perbezaan berlaku antara teks pelindung dan teks stego yang dihasilkan (Lampiran G). Sementara itu, Rajah 5.8 menunjukkan perbandingan skor Jaro Winkler bagi beberapa kajian lepas dan kajian yang dijalankan.



Rajah 5.8 Perbandingan Skor Jaro Winkler Berbanding Kajian Lepas

Rajah 5.8 menunjukkan enam skor Jaro Winkler bagi kajian terkini dan didapati bahawa hanya kajian yang dijalankan oleh Naharuddin et al. (2018) dan El Rahman dan Nourah (2019) menunjukkan skor Jaro Winkler bernilai 1 berbanding dengan kajian lain. Rajah 5.8 juga menunjukkan, nilai Jaro Winkler bagi kajian yang dijalankan bernilai 1 yang menunjukkan terdapat kesamaan antara teks pelindung dan teks stego. Kesimpulannya, prestasi ketakbolehkelihatan teks stego yang dihasilkan adalah sama dengan yang dilakukan oleh Naharuddin et al., (2018) dan El Rahman, (2019).

5.11 Masa Dan Kompleksiti

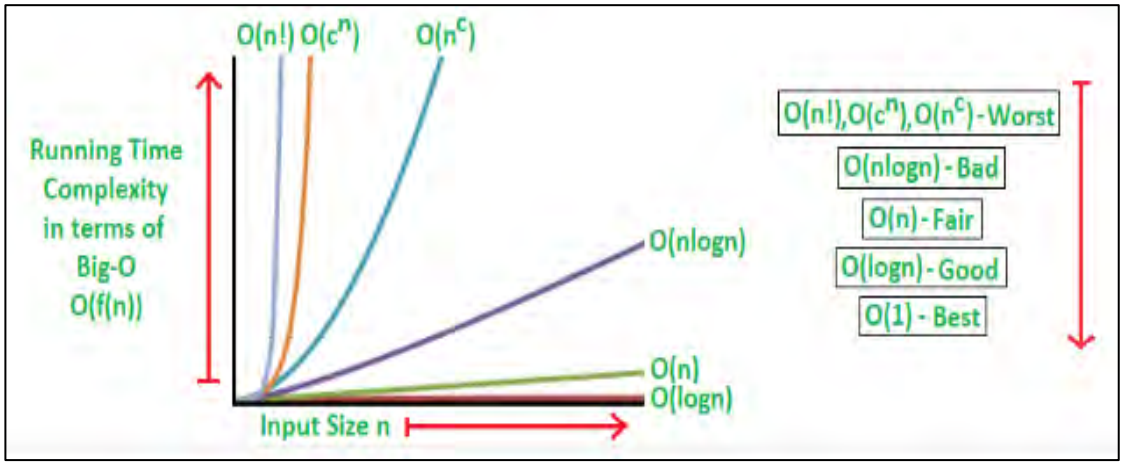
Kecekapan sesuatu algoritma bergantung kepada kecekapan masa dan kecekapan ingatan yang boleh digunakan untuk mengukur prestasi algoritma (Chandel, 2012). Teknik steganografi yang digunakan melibatkan proses penyembunyian dan proses pengekstrakan. Kecekapan proses penyembunyian bergantung kepada algoritma yang digunakan di dalam sesuatu teknik. Algoritma penyembunyian perlulah memastikan prestasi masa penyembunyian adalah pantas serta tidak kompleks. Oleh itu, dalam kajian ini, fasa penyembunyian melibatkan proses penjanaan jadual Homophonic sifer dan proses penentuan lokasi penyembunyian rawak di mana masing-masing melibatkan lokasi rawak aksara dalam teks pelindung dan fungsi nombor rawak PRNG. Penggunaan PRNG bertujuan untuk menggantikan teknik lokasi jujukan bagi meningkatkan prestasi teknik yang digunakan di samping proses memformatkan warna RGB menggunakan formula SQRT yang diperkenalkan. Kompleksiti algoritma terhadap masa boleh diukur menggunakan notasi Big O seperti yang ditunjukkan di dalam Jadual 5.13 bermula daripada paling baik ke paling lemah.

Jadual 5.13

Urutan Kompleksiti Notasi Big O

Nama	Notasi Big O	Penerangan	Contoh
Pemalar *Constant*	O(1)	Umpukkan pemboleh ubah	Nombor genap atau ganjil
Algoritma Logaritma *Logarithmic Algorithm*	O(*log n*)	Masa larian berkembang secara logaritmik mengikut *n*	Carian Binari *Binary Search.*
Algoritma Linear *Linear Algorithm*	*O(n)*	Masa larian berkembang secara langsung dengan *n*.	Carian Linear *Linear Search.*
Algoritma Superlinear *Superlinear Algorithm*	*O(n log n)*	Masa larian tumbuh mengikut *n*	Isih *Heap,* Isih *Merge* *Heap Sort, Merge Sort*
Algoritma Polinomial *Polynomial Algorithm*	$O(n^c)$	Masa larian berkembang lebih cepat daripada algoritma di atas berdasarkan *n*	Isih *Gelembung,* Isih *Pilihan* *Bubble Sort, Selection Sort*
Algoritma Exponen *Exponential Algorithm*	$O(c^n)$	Masa larian berkembang lebih pantas daripada algoritma polinomial berdasarkan *n*.	*Tower of Hanoi.*
Algoritma Faktorial *Factorial Algorithm*	*O(n!)*	Masa larian berkembang paling cepat dan tidak sesuai digunakan untuk nilai kecil *n*.	Algoritma *Brute force Search* untuk masalah *Traveling Salesman*

Menurut Wikipedia, PRNG menjana nombor rawak secara *Linear Congruence Generator* di mana nombor pseudorawak dijana berdasarkan persamaan linear kepingan (*piecewise linear equation*) yang tidak berterusan. Di dalam bidang matematik, persamaan linear kepingan akan membentuk graf berbentuk segmen garis-lurus (Stanle & William, 2004). Secara amnya, kajian yang dijalankan menggunakan teknik penyembunyian secara linear (PRNG) yang tergolong di dalam notasi Big O = *O(n)* di mana masa larian berkembang secara langsung dengan *n* kepada Jadual 5.13 . Hubungan kompleksiti dan masa ditunjukkan di dalam Rajah 5.9 di bawah.



Rajah 5.9 Kompleksiti Notasi Big O Antara Masa dan Saiz Input

Oleh itu, berdasarkan kepada Rajah 5.9, hubungan antara kompleksiti dan masa bagi teknik yang digunakan adalah sederhana, O(*n*).

5.12 Ringkasan

Hasil yang diperolehi daripada kajian menggunakan teknik penyembunyian warna RGB menunjukkan purata kapasiti penyembunyian meningkat sebanyak 41.31% berbanding kajian lepas. Selain itu, teks stego yang dihasilkan berjaya dibuktikan keteguhannya apabila proses pemampatan dan penyahmampatan tidak menjejaskan mesej rahsia yang disembunyikan dalam teks stego yang dihasilkan. Kesemua teks stego yang dihasilkan menunjukkan nilai Jaro Winkler yang dihasilkan bersamaan dengan 1 yang menunjukkan tiada perbezaan berlaku antara teks pelindung dan teks stego yang dihasilkan. Secara keseluruhannya, prestasi teks stego yang dihasilkan berjaya ditingkatkan. Selain itu, teknik yang dibangunkan ini menghasilkan teks stego yang pelbagai untuk sesuatu mesej rahsia, dengan menggunakan satu teks pelindung sahaja. Ini berbeza dengan teknik-teknik lain yang hanya menghasilkan satu teks stego menggunakan mesej rahsia dan teks pelindung yang sama.

BAB ENAM
KESIMPULAN DAN KERJA MASA DEPAN

Bab ini merumuskan keseluruhan kajian yang telah dijalankan. Perkara yang dibincangkan dalam bab ini merangkumi ringkasan penyelidikan, pencapaian objektif, sumbangan kajian, limitasi kajian dan kajian masa hadapan.

6.1 Ringkasan Penyelidikan

Steganografi merupakan kaedah untuk menyembunyikan mesej di dalam pelbagai media seperti teks, imej, audio, video dan sebagainya yang dikenali sebagai media pelindung. Matlamat utama steganografi adalah untuk memelihara mesej yang dihantar agar kewujudannya di dalam media pelindung tidak disedari oleh pihak ketiga. Kajian ini memilih media teks sebagai media pelindung dan hasil akhir yang dihasilkan daripada proses steganografi dalam kajian dikenali sebagai teks stego.

Proses penyembunyian dalam kajian ini merangkumi beberapa langkah. Pertama, penjanaan Jadual *Homophonic* berdasarkan kepada teks pelindung yang dipilih. Langkah kedua ialah mewakilkan aksara mesej rahsia dengan nilai ASCII berdasarkan kepada Jadual *Homophonic* yang dijana. Langkah ketiga ialah dengan menukarkan nilai ASCII mesej rahsia ke bentuk perwakilan 3D (*x,y,z*) menggunakan formula SQRT. Tujuan perwakilan ini adalah untuk memetakan nilai (*x,y,z*) kepada warna RGB aksara teks pelindung yang dipilih secara rawak menggunakan nombor pseudorawak. Seterusnya, warna RGB aksara pada lokasi terpilih akan ditukarkan dengan nilai (*x,y,z*) yang dihasilkan pada langkah ketiga di atas. Akhir sekali, proses pengekstrakan dilakukan bertujuan untuk mendapatkan semula mesej yang disembunyikan bagi mengesahkan proses steganografi berjaya dilaksanakan.

Prestasi teks stego yang dihasilkan diukur dengan tiga parameter utama iaitu kapasiti, keteguhan dan ketakbolehkelihatan. Secara ringkasnya, kajian yang dihasilkan ini telah mencapai objektif yang telah ditetapkan seperti yang diterangkan pada sub-topik berikutnya.

6.1.1 Pencapaian Pertama

Objektif pertama kajian ialah untuk menentukan perwakilan aksara mesej rahsia yang berulang agar mempunyai nilai yang dinamik dengan menjana Jadual *Homophonic* sifer serta penggunaan nombor rawak. Kajian yang dijalankan ini berjaya membuktikan aksara mesej rahsia yang berulang berjaya diwakilkan dengan pelbagai nilai yang dinamik menggunakan jadual *homophonic*. Jadual ini dijana secara rawak menggunakan penjana rawak PRNG berdasarkan kepada lokasi aksara yang terdapat dalam teks pelindung. Penjanaan jadual ini bertujuan untuk memetakan aksara mesej rahsia yang berulang dengan pelbagai nilai. Hasil kajian menunjukkan bahawa mesej rahsia boleh diwakilkan dengan pelbagai nilai dinamik yang menggambarkan kepelbagaian perwakilan mesej rahsia seperti yang ditunjukkan pada Jadual 6.4.

6.1.2 Pencapaian Kedua

Objektif kedua kajian ialah untuk menentukan aksara mesej rahsia diwakilkan dengan perwakilan tiga dimensi untuk dipetakan kepada warna RGB. Penggunaan teorem hasil bahagi (QRT) dapat mewakilkan sesuatu nombor dalam bentuk perwakilan dua dimensi iaitu (x,y). Penambahbaikan dilakukan dengan melaksanakan pembahagian kali kedua (SQRT) terhadap nilai hasil bahagi QRT bagi

membolehkan nilai tersebut diwakilkan dalam bentuk tiga dimensi (*x,y,z*). Proses penukaran ke bentuk perwakilan tiga dimensi berjaya dilakukan dan diterangkan pada sub-topik 4.2 hingga 4.5. Perwakilan tiga dimensi ini seterusnya dipetakan kepada perwakilan warna RGB bertujuan untuk melakukan penyembunyian pada aksara yang dipilih secara rawak dalam teks pelindung. Lokasi penyembunyian dilakukan secara rawak menggunakan fungsi nombor pseudorawak menggunakan kekunci tertentu seperti yang diterangkan pada sub-topik 4.6. Proses penyembunyian telah berjaya dilaksanakan sepenuhnya akan menghasilkan teks stego di dalam format *rtf*.

6.1.3 Pencapaian Ketiga

Objektif ketiga kajian ialah untuk menilai prestasi teks stego yang dijana berdasarkan kepada kapasiti, keteguhan dan ketakbolehkelihatan teks stego yang dihasilkan berbanding dengan teknik-teknik lain. Kajian ini melibatkan dua proses utama iaitu proses penyembunyian dan proses pengekstrakan di mana output yang dihasilkan ialah teks stego. Teks stego yang dihasilkan dinilai prestasinya berdasarkan kepada tiga ukuran utama iaitu kapasiti, keteguhan dan ketakbolehkelihatan.

6.1.3.1 Ukuran Pertama : Kapasiti

Teknik yang dicadangkan berjaya meningkatkan kapasiti mesej rahsia hampir 41.31% (Jadual 5.10) dengan purata kapasiti penyembunyian ialah 77.4% seperti yang ditunjukkan di dalam Jadual 5.9. Kapasiti penyembunyian meningkat sebanyak 38.9% berbanding teknik terkini yang diperkenalkan oleh Al-Azzawi, (2018). Kapasiti penyembunyian dikira berdasarkan kepada

formula yang diperkenalkan di dalam bidang steganografi seperti yang dijelaskan pada sub-topik 2.3. Proses pengekstrakan ke atas teks stego dilakukan untuk memastikan mesej yang disembunyikan berjaya di ekstrak semula sepenuhnya tanpa sebarang kehilangan data.

6.1.3.2 Ukuran Kedua : Keteguhan

Teks stego yang dihasilkan telah diuji keteguhannya melalui proses pemampatan dan penyahmampatan. Hasil kajian menunjukkan kesemua mesej rahsia berjaya di ekstrak semula 100% selepas melalui proses pemampatan dan penyahmampatan seperti yang ditunjukkan pada Jadual 6.8. Kejayaan mengekstrak semula mesej yang disembunyikan di dalam semua fail teks stego yang dijana menunjukkan tidak berlaku sebarang kehilangan data semasa melalui proses pemampatan dan penyahmampatan.

6.1.3.3 Ukuran Ketiga : Ketakbolehkelihatan

Prestasi ketakbolehkelihatan teks stego yang dihasilkan diukur menggunakan skala Jaro Winkler seperti yang diterangkan pada sub-topik 2.3. Hasil kajian menunjukkan kesemua teks stego yang dihasilkan menghasilkan nilai skor 1 yang menandakan tidak berlaku sebarang perbezaan antara teks pelindung dan teks stego yang dihasilkan.

6.2 Sumbangan Kajian

Secara umumnya sumbangan kajian ini dapat dibahagikan kepada dua aspek iaitu: daripada aspek teori dan praktikal.

6.2.1 Sumbangan 1 : Penjanaan Jadual *Homophonic* Yang Fleksibel

Jadual *Homophonic* dijana berdasarkan kepada frekuensi aksara di dalam teks pelindung bertujuan untuk mewakilkan kepelbagaian nilai untuk aksara mesej rahsia terutamanya bagi aksara berulang. Jadual yang dijana ini lebih fleksibel (*flexible*) berbanding jadual lain kerana ia dijana berdasarkan kepada teks pelindung yang digunakan. Kajian sebelum ini menunjukkan setiap aksara mesej rahsia diwakilkan dengan satu perwakilan sahaja sebelum digunakan untuk proses penyembunyian. Ini berbeza dengan penjanaan Jadual *Homophonic* yang dicadangkan di dalam kajian ini di mana aksara mesej rahsia boleh diwakilkan dengan pelbagai nilai yang bersifat lebih dinamik kerana pemetaannya bersifat satu-ke-banyak (*one-to-many*) berbanding teknik lain yang lebih bersifat satu-ke-satu (*one-to-one*). Penjanaan Jadual Homophonic yang dicadangkan ini dapat membantu proses mewakilkan aksara mesej rahsia dengan perwakilan yang lebih dinamik.

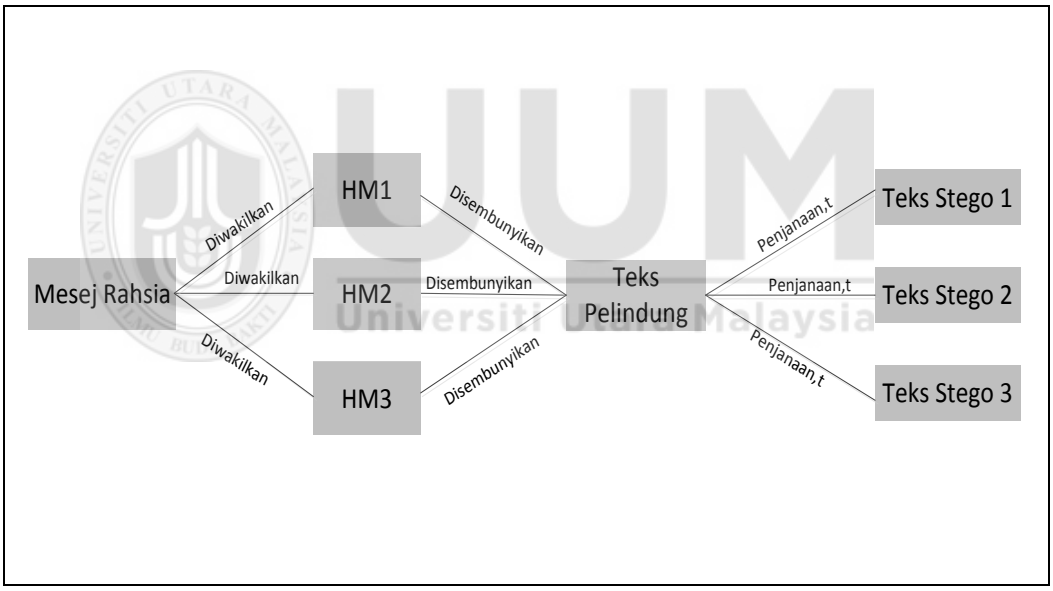
6.2.2 Sumbangan 2 : Teknik Penyembunyian Lokasi Rawak

Kebanyakan teknik penyembunyian di dalam kajian lepas adalah berdasarkan kepada lokasi jujukan. Antaranya ialah penyembunyian pada setiap huruf pertama sesuatu perkataan atau huruf pertama sesuatu ayat atau setiap ruang kosong antara perkataan. Pendekatan ini memudahkan teknik steganalisis untuk membuat pengesanan lokasi penyembunyian. Oleh itu, kajian ini mencadangkan lokasi penyembunyian secara rawak menggunakan formula nombor rawak pseudo seperti yang diterangkan pada sub-topik 3.2.1.8. Gabungan teknik penyembunyian berasaskan lokasi rawak dan perwakilan aksara mesej rahsia secara dinamik merupakan teknik yang belum diterokai oleh penyelidik pada masa kini dan boleh diimplementasikan di dalam

pelbagai bahasa serta boleh digunakan untuk medium-medium lain seperti imej, audio, video dan sebagainya.

6.2.3 Sumbangan 3 : Kepelbagaian Mesej Rahsia di dalam Teks Pelindung

Mesej rahsia berbeza kebiasaannya memerlukan teks pelindung yang berbeza untuk melakukan proses penyembunyian. Namun teknik yang digunakan di dalam kajian ini membolehkan pelbagai mesej rahsia dapat disembunyikan di dalam teks pelindung yang sama menggunakan hubungan satu-ke-banyak (*one-to-many)* seperti yang ditunjukkan pada Rajah 6.1 di bawah.



Rajah 6.1 Hubungan Mesej Rahsia dan Teks Pelindung

Mesej yang sama dapat disembunyikan menggunakan teks pelindung yang berbeza dan mesej yang berbeza dapat disembunyikan pada teks pelindung yang sama.

6.2.4 Sumbangan 4 : Perwakilan Mesej Rahsia dalam Bentuk 3D

Model 2D merupakan perwakilan aksara mesej rahsia yang ditukarkan ke bentuk (*x,y*) yang diperkenalkan oleh penyelidik lepas. Model 2D ini mewakilkan aksara mesej rahsia dalam julat yang terhad. Oleh itu, satu formula telah diperkenalkan di dalam kajian ini yang dikenali sebagai teorem SQRT yang diterbitkan berdasarkan kepada teorem asas QRT bertujuan untuk mewakilkan aksara mesej rahsia dalam bentuk yang dinamik dan diwakilkan dalam bentuk perwakilan tiga dimensi. Perwakilan (*x,y,z*) ini dapat memetakan aksara mesej rahsia kepada warna RGB. Di samping itu, teorem SQRT yang diperkenalkan boleh diimplementasikan di dalam steganografi menggunakan medium-medium lain seperti imej, audio, video dan sebagainya.

6.2.5 Sumbangan 5 : Perwakilan Warna RGB

Penyembunyian berdasarkan warna RGB dengan julat antara (0,0,0) hingga (15,15,15) berjaya meningkatkan prestasi penyembunyian mesej rahsia bagi mengelakkan berlakunya kecurigaan warna. Sebanyak 3375 (b^3, b = 15) perwakilan warna (gelap) dapat diwakilkan menggunakan formula yang diperkenalkan, namun perwakilan untuk warna selain warna tersebut dapat dilakukan dengan melakukan perubahan terhadap pemalar *b*. Jumlah perwakilan ini bergantung kepada pemboleh ubah *b* yang digunakan di dalam formula tersebut yang bersifat berkadar langsung dengan nilai *b*. Hasil implementasi model ini ialah satu notasi geometri berbentuk (*x,y,z*) di mana perwakilan ini berpotensi tinggi digunakan di dalam steganografi imej kerana nilai yang dihasilkan dapat dipetakan ke bentuk warna RGB. Formula yang diterbitkan boleh digunakan untuk medium-medium lain terutamanya untuk mewakilkan aksara

berulang dan tunggal dalam bentuk yang lebih dinamik berbanding perwakilan statik yang digunakan sebelum ini.

6.2.6 Sumbangan 6 : Kepelbagaian Bidang

Secara praktikalnya, teknik penyembunyian menggunakan gabungan perwakilan 3D dan nombor rawak telah berjaya meningkatkan kapasiti penyembunyian. Oleh itu, dengan kapasiti penyembunyian yang tinggi, banyak maklumat rahsia dapat disembunyikan di dalam teks stego yang dijana. Justeru itu, teknik yang dicadangkan ini boleh digunakan untuk penyembunyian maklumat rahsia di dalam pelbagai bidang tertentu seperti ketenteraan (kod rahsia), perbankan (kata kunci), pemakanan (status halal), perindustrian (ketulenan produk) dan sebagainya seperti penggunaan untuk menentukan ketulenan dokumen, penyembunyian kod rahsia, penyembunyian maklumat status halal dan sebaginya dapat dilakukan.

6.3 Limitasi

Kajian ini hanya memfokuskan kepada penyembunyian aksara mesej rahsia ke dalam aksara teks pelindung dengan memetakan perwakilan nilai aksara mesej rahsia dengan warna RGB. Beberapa limitasi kajian telah dikenal pasti di dalam kajian ini, antaranya:

i. Setiap aksara mesej rahsia perlu wujud di dalam teks pelindung bagi membolehkan proses penyembunyian dapat dilakukan sepenuhnya. Ketiadaan aksara mesej rahsia di dalam teks pelindung akan menyebabkan proses penyembunyian gagal dilaksanakan.

ii. Kajian ini hanya memfokuskan kepada mesej rahsia yang mengandungi aksara A-Z (termasuk huruf kecil) sahaja tanpa mengambil kira aksara lain. Penggunaan simbol atau nombor akan menyebabkan proses penyembunyian gagal dilaksanakan.

iii. Julat perwakilan nilai RGB ialah di antara (0,0,0) hingga (15,15,15) sahaja. Julat perwakilan melebihi nilai tersebut boleh menyebabkan warna aksara jelas kelihatan dan akan mendorong kepada kecurigaan teks stego yang dijana.

6.4 Kajian Masa Depan

Kajian ini menyembunyikan aksara mesej rahsia pada lokasi rawak di dalam teks pelindung di mana setiap aksara mesej rahsia perlu wujud di dalam teks pelindung. Jadual *Homophonic* dijana berdasarkan kepada lokasi aksara di dalam teks pelindung seperti yang dijelaskan pada sub-topik 2.8.2. Terdapat keadaan di mana aksara mesej rahsia tidak wujud di dalam teks pelindung yang menyebabkan mesej rahsia gagal untuk disembunyikan di dalam teks pelindung seperti yang ditunjukkan di dalam Jadual 5.5. Oleh itu, kajian masa hadapan boleh diteruskan dengan mencari satu kaedah yang boleh menyembunyikan setiap aksara mesej rahsia tanpa perlu bergantung kepada kewujudan setiap aksara di dalam teks pelindung. Penyembunyian peringkat bit mungkin boleh digunakan untuk menyelesaikan masalah ini.

Kajian yang dijalankan ini hanya memfokuskan kepada aksara mesej rahsia yang terdiri daripada aksara A-Z (termasuk huruf kecil) sahaja. Ini disebabkan penyembunyian dilakukan pada peringkat aksara dan bergantung sepenuhnya kepada kandungan aksara teks pelindung. Oleh itu satu kajian secara menyeluruh boleh dilakukan dengan melibatkan aksara mesej rahsia yang mengandungi nombor dan

aksara lain tanpa perlu bergantung sepenuhnya kepada teks pelindung atau dengan menggabungkan teknik kriptografi terhadap mesej rahsia sebelum ditukarkan kepada perwakilan rawak.

Kajian ini dapat mewakilkan setiap aksara mesej rahsia dengan pelbagai nilai termasuk aksara yang berulang dan sebaliknya seperti yang ditunjukkan di dalam Jadual 4.1. Perwakilan nilai rawak dan penggunaan Jadual *Homophonic* dapat menjana pelbagai set mesej rahsia. Namun. kajian ini hanya memilih salah satu set perwakilan tersebut secara rawak. Oleh itu bagi memantapkan lagi prestasi steganografi, satu kajian boleh dilakukan untuk memilih satu set perwakilan mesej rahsia yang terbaik di kalangan set perwakilan yang dijana dengan menggunakan teknik-teknik kecerdasan buatan seperti algoritma genetik, koloni semut atau sebagainya. Trend penggunaan algoritma genetik dalam media lain seperti imej, audio dan video yang sedang digunakan pada masa ini untuk menyembunyikan mesej rahsia dan berkemungkinan tinggi dapat meningkat prestasi teks stego yang dijana.

RUJUKAN

Abuadbba, A., Khalil, I., Ibaida, A., & Atiquzzaman, M. (2015). Resilient to Shared Spectrum Noise Scheme for Protecting Cognitive Radio Smart Grid Readings − BCH Based Steganographic Approach. *Ad Hoc Networks*, *0*(November), 1–17. https://doi.org/10.1016/j.adhoc.2015.11.002

Adesina, A. O., Nyongesa, H. O., & Agbele, K. K. (2010). Digital Watermarking : A State-of-the-Art Review. In *IST-Africa 2010 Conference Proceedings* (pp. 1–8).

Agarwal, M. (2013). Text Steganographic Approaches: A Comparison. *International Journal of Network Security & Its Applications*, *5*(1), 91–106. https://doi.org/10.5121/ijnsa.2013.5107

Agarwal, M. (2017). Making It Unseen-Espionaging Data in WhatsApp Emoticons. In *3rd IEEE International Conference on "Computational Intelligence and Communication Technology"* (pp. 1–4).

Agath, A., Sidpara, C., & Upadhyay, D. (2018). Critical Analysis of Cryptography and Steganography. *International Journal of Scientific Research in Science, Engineering and Technology*, *4*(2), 1–6.

Aghdam, M. H., Ghasem-Aghaee, N., & Basiri, M. E. (2009). Text feature selection using ant colony optimization. *Expert Systems with Applications*, *36*(3 PART 2), 6843–6853. https://doi.org/10.1016/j.eswa.2008.08.022

Ahvanooey, M. T., Li, Q., Hou, J., Mazraeh, H. D., & Zhang, J. (2018). AITSteg : An Innovative Text Steganography Technique for Hidden Transmission of Text Message via Social Media. *IEEE Access*, *20*, 1. https://doi.org/10.1109/ACCESS.2018.2866063

Ahvanooey, M. T., Li, Q., Hou, J., Rajput, A. R., & Yini, C. (2019). Modern Text Hiding , Text Steganalysis , and Applications : A Comparative Analysis. *Entropy*, *21*(4), 1–29. https://doi.org/10.3390/e21040355

Ahvanooey, M. T., Li, Q., Shim, H. J., & Huang, Y. (2018). A Comparative Analysis of Information Hiding Techniques for Copyright Protection of Text Documents. *Security and Communication Networks*, (April). https://doi.org/10.1155/2018/5325040

Akotoye, F. X. K. (2017). A Text Steganographic System Based on Word Length Entropy Rate. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, *5*(3), 71–76.

Al-Asadi, S. A., & Bhaya, W. (2016). Text Steganography in Excel Documents Using Color and Type of Fonts. *Research Journal of Applied Sciences*, *11*(10), 1054–1059.

Al-Azzawi, A. F. (2018). A Multi-Layer Hybrid Text Steganography For Secret Communication Using Word Tagging and RGB Color. *International Journal of Network Security & Its Applications*, *10*(6), 1–12. https://doi.org/10.5121/ijnsa.2018.10601

Alanazi, H. O., Zaidan, a. a., Zaidan, B. B., Jalab, H. a., & AL-Ani, Z. K. (2010).

New Classification Methods for Hiding Information into Two Parts: Multimedia Files and Non Multimedia Files. *JJournal of Computing*, *2*(3), 144–151. Retrieved from http://arxiv.org/abs/1003.4084

Ali, A. A., & Saad, H. A.-S. (2013). New Text Steganography Technique by using Mixed-Case Font. *International Journal of Computer Applications*, *62*(3), 6–9.

Almuhammadi, S., & Al-shaaby, A. (2017). A Survey on Recent Approaches of Combining Cryptography And Steganography. *Computer Science & Information Technology (CS & IT)*, 63–74.

Alotaibi, R. A., & Elrefaei, L. A. (2018). Improved capacity Arabic text watermarking methods based on open word space. *Journal of King Saud University - Computer and Information Sciences*, *30*(2), 236–248. https://doi.org/10.1016/j.jksuci.2016.12.007

Aman, M., Khan, A., Ahmad, B., & Kouser, S. (2017). A Hybrid Text Steganography Approach Utilizing Unicode Space Characters and Zero-Width Character. *International Journal on Information Technologies & Security*, *9*(1), 85–100.

Amirthrajan, R., & Rayappan, J. B. . (2013). Steganography Time to Time: A Review. *Research Journal of Information Technology*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Archana, S., Judice, A. A., & Kaliyamurthie, K. P. (2013). A Novel Approach on Image Steganographic Methods for Optimum Hiding Capacity. *International Journal of Engineering and Computer Science*, *2*(2), 378–385.

Arya, A., & Soni, S. (2018). A Literature Review on Various Recent Steganography Techniques. *International Journal on Future Revolution in Computer Science & Communication Engineering*, *4*(January), 143–149.

Atoum, M. S. (2018). Steganography and Watermarking : Review. *International Journal of Science and Research (IJSR)*, *7*(6), 2017–2019. https://doi.org/10.21275/ART20183644

Baawi, S. S., Mokhtar, M. R., & Sulaiman, R. (2017). New Text Steganography Technique Based On A Set Of Two-Letter Words. *Journal of Theoretical and Applied Information Technology*, *95*(22), 6247–6255.

Baawi, S. S., Mokhtar, M. R., & Sulaiman, R. (2018). A Comparative Study on The Advancement of Text Steganography Techniques in Digital Media. *ARPN Journal of Engineering and Applied Sciences*, *13*(5), 1854–1863.

Baawi, S. S., Mokhtar, M. R., & Sulaiman, R. (2019). Enhancement of Text Steganography Technique Using Lempel-Ziv-Welch Algorithm and Two-Letter Word Technique. In *3rd International Conference of Reliable Information and Communication Technology* (pp. 525–537). Springer International Publishing. https://doi.org/10.1007/978-3-319-99007-1

Babu, K. R. (2010). A Survey on Cryptography and Steganography Methods for Information Security. *International Journal*, *12*(2), 13–17. https://doi.org/10.5120/1660-2235

Bailey, K., & Curran, K. (2006). An Evaluation of Image Based Steganography Methods using Visual Inspection and Automated Detection Techniques.

Multimed Tools Appl Ication, *31*(3), 55–88. https://doi.org/10.1007/s11042-006-0008-4

Banik, B. G., & Bandyopadhyay, S. K. (2018). Novel Text Steganography Using Natural Language Processing and Part-of-Speech Tagging Novel Text Steganography Using Natural Language Processing and. *IETE Journal of Research*, *2063*. https://doi.org/10.1080/03772063.2018.1491807

Baykara, M., Das, R., & Tuna, G. (2017). A Novel Symmetric Encryption Algorithm and its Implementation Muhammet BAYKARA 1 , Resul DAŞ 1 , Gürkan TUNA 2 1. *Turkish Journal of Science & Technology*, *12*(1), 5–9.

Bennett, K. (2004). *Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text*. Retrieved from https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2004-13.pdf

Bhat, D., Prabhu, S., & Renuka, A. (2017). Information Hiding through Dynamic Text Steganography and Cryptography. In *International Conference on Advances in Computing, Communications and Informatics* (pp. 1826–1831).

Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2010). A Novel Approach of Secure Text Based Steganography Model using Word Mapping Method. *International Journal of Computer and Information Engineering*, *4*(2), 96–103.

Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2011). A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier. *Journal of Global Research in Computer Science*, *2*(4), 44–54.

Bhattacharyya, S., Indu, P., Dutta, S., Biswas, A., & Sanyal, G. (2011). Text Steganography using CALP with High Embedding Capacity. *Journal of Global Research in Computer Science*, *2*(4), 29–36.

Bhaya, W., Rahma, A. M., & Al-nasrawi, D. (2013). Text Steganography Based on Font Type in MS-Word Documents. *Journal of Computer Science*, *9*(7), 898–904. https://doi.org/10.3844/jcssp.2013.898.904

Bloom, J. A., Cox, I. J., Linnartz, J. P. M. G., Kalker, T., Miller, M. L., & Traw, C. B. S. (1999). Copy Protection for DVD Video. In *Proceedings of the IEEE*. https://doi.org/10.1109/5.771077

Böhme, R. (2010). Principles of Modern Steganography and Steganalysis. In *Advanced Statistical Steganalysis* (pp. 11–77). Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-14313-7_2

Chandini, B., & Ganesh Kumar, M. . (2018). A Novel Architecture for Video Steganography Using Pixel Pattern Matching. *International Journal of Computer Science and Mobile Computing*, *7*(6), 16–22.

Chaudhary, S., & Dave, M. (2016). Text Steganography Based on Feature Coding Method. In *International Conference on Advances in Information Communication Technology & Computing* (pp. 5–8).

Chaudhary, S., Dave, M., & Sanghi, A. (2016). Review of Linguistic Text Steganographic Methods. *International Journal on Recent and Innovation Trends in Computing and Communication*, *4*(7), 377–381.

Chaundhary, S., Dave, M., & Sanghi, A. (2016). Aggrandize Text Security and Hiding Data through Text Steganography. In *Conference: 2016 IEEE 7th Power India International Conference (PIICON)*.

Conklin, W., & William, A. (2011). *CompTIA security+ exam guide (exam SYO-301) electronic resource all in one* (3rd ed.). New York: McGraw-Hill.

Dasgupta, K., Mondal, J. K., & Dutta, P. (2013). Optimized Video Steganography Using Genetic Algorithm (GA). *Procedia Technology*, *10*, 131–137. https://doi.org/10.1016/j.protcy.2013.12.345

David, K. (1967). *The Codebreakers – The Story of Secret Writing*.

Dhavare, A., Low, R. M., & Stamp, M. (2013). Efficient Cryptanalysis of Homophonic Substitution Ciphers. *Cryptologia*, *37*(3), 37–41. https://doi.org/10.1080/01611194.2013.797041

Din, R., & Amphawan, A. (2015). Performance Analysis on Text Steganalysis Method Using A Computational Intelligence Approach. In *International Conference on Electrical Engineering, Computer Science and Informatics* (pp. 19–20).

Din, R., Ani, Z. C., & Samsudin, A. (2012). A Formulation of Conditional States on Steganalysis Approach. *WSEAS Transactions on Mathematics*, *11*(3), 173–182.

Din, R., & Samsudin, A. (2009). Digital Steganalysis : Computational Intelligence Approach. *International Journal*, *3*(1), 161–170.

Din, R., & Utama, S. (2018). Analysis Review of Feature-Based Method in Term of Verification and Validation Performance. *Journal of Telecommunication, Electronic and Computer Engineering*, *10*(2–4), 173–177.

Dobriyal, P., Yadav, J., & Jain, J. (2015). A Review on Text Based Steganography. *The International Journal Research Publications*, *4*(3), 44–50.

Douglas, M., Bailey, K., Leeney, M., & Curran, K. (2018). An overview of steganography techniques applied to the protection of biometric data. *Multimed Tools Appl*, *77*(13), 17333–17373.

Dulera, S., Jinwala, D., & Dasgupta, A. (2011). Experimenting with the Novel Approaches in Text Steganography. *International Journal of Network Security & Its Applications*. https://doi.org/10.5121/ijnsa.2011.3616

El Rahman, S. A. (2019). Text Steganography Approaches Using Similarity of English Font Styles. *International Journal of Software Innovation*, *7*(3), 29–50. https://doi.org/10.4018/IJSI.2019070102

Elmahi, M. Y., Wahbi, T. M., & Sayed, M. H. (2017). Text Steganography Using Compression and Random Number Generators. *International Journal of Computer Applications Technology and Research*, *6*(6), 259–263.

Fateh, M., & Rezvani, M. (2018). An e-mel-based high capacity text steganography using repeating characters. *International Journal of Computers and Applications*, *0*(0), 1–7. https://doi.org/10.1080/1206212X.2018.1517713

Fuad, M. N., Informatika, J. T., & Informasi, F. T. (2014). Steganografi Berbasis Model Matematika dari Kode ASCII dalam Directed Graph. Retrieved from

https://www.academia.edu/17713522/Steganografi_Berbasis_Model_Matematika_dari_Kode_ASCII_dalam_Directed_Graph

Gongshen, L., Xiaoyun, D., Bo, S., & Meng, K. (2013). A Text Information Hiding Algorithm Based on Alternatives. *Journal of Software*, *8*(8), 2072–2079. https://doi.org/10.4304/jsw.8.8.2072-2079

Grigas, G., & Juškevičienė, A. (2018). Letter Frequency Analysis of Languages Using Latin Alphabet. *International Linguistics Research*, *1*(1), 18–31.

Gupta, R., Gupta, S., & Singhal, A. (2014). Importance and Techniques of Information Hiding: A Review. *International Journal of Computer Trends and Technology*, *9*(5), 260–265. https://doi.org/10.14445/22312803/IJCTT-V9P149

Hamdan, A. M., & Hamarsheh, A. (2017). AH4S : an algorithm of text in text steganography using the structure of omega network. *Security and Communication Networks*, *9*(18). https://doi.org/10.1002/sec.1752

Hana‟a, M. S. (2008). A Natural Language Steganography Technique for Text Hiding Using LSB‟s. *Enginnering and Technology*, *26*. Retrieved from http://www.uotechnology.edu.iq/tec_magaz/volume262008/No3/Researches/6.pdf

Herodotus. (1992). *The Histories*. London, England: J. M. Dent & Sons, Ltd.

Hmood, A. K., Jalab, H. A., Kasirun, Z. M., Zaidan, B. B., & Zaidan, A. A. (2010). On the Capacity and Security of Steganography Approaches-An Overview.pdf. *Journal of Applied Sciences*, *10*(16), 1825–1833.

Htet, M., & Phyo, S. W. (2016). A Novel Text Steganographic Technique Using Specific Alphabets. *Journal of Computer Science*, *2*(1), 1–11.

Ingemar, J. C., Matthew, L. M., Jeffrey, A. B., Fridrich, J., & Kalker, T. (2008). *Digital Watermarking and Steganography* (Second). Burlington, USA: Morgan Kaufmann Publishers.

Iranmanesh, V., Wei, H. J., Dao-ming, S. L., & Arigbabu, O. A. (2015). On using Emoticons and Lingoes for Hiding Data in SMS. In *2015 International Symposium on Technology Management and Emerging Technologies (ISTMET)* (pp. 103–107).

Ishtiaq, S., Khan, A., & Samim, B. A. (2017). Novel Content Independent Steganographic Method for Microsoft Office. *INTERNETWORKING INDONESIA JOURNAL*, *9*(2).

Iyer, S. S. (2017). Practical Evaluation and Comparative Study of Text Steganography Algorithms. *International Journal of Advance Engineering and Research Development*, *3*(4), 277–283.

Iyer, S. S., & Lakhtaria, K. (2016a). Clustering Algorithm for Text Steganography. *International Journal of Advanced Research in Computer and Communication Engineering*, *5*(3), 74–77. https://doi.org/10.17148/IJARCCE

Iyer, S. S., & Lakhtaria, K. (2016b). New Robust and Secure Alphabet Pairing Text Steganography Algorithm. *International Journal of Current Trends in Engineering & Research (IJCTER)*, *2*(7), 15–21.

Jassim, F. A. (2013). A Novel Steganography Algorithm for Hiding Text in Image using Five Modulus Method. *International Journal of Computer Applications*, *72*(17), 39–44.

Johnson, N. F., & Jajodia, S. (1998). Exploring steganography: Seeing the unseen. *IEEE Computer*, *31*. https://doi.org/10.1109/MC.1998.4655281

Jose, P. G., Chatterjee, S., Patodia, M., Kabra, S., & Nath, A. (2016). Hash and Salt based Steganographic Approach with Modified LSB Encoding. *International Journal of Innovative Research in Computer and Communication Engineering*, *4*(6), 2257–2263. https://doi.org/10.15680/IJIRCCE.2016.

Joseph, P., & Vishnukumar, S. (2015). A study on steganographic techniques. In *Global Conference on Communication Technologies, GCCT 2015* (pp. 206–210). https://doi.org/10.1109/GCCT.2015.7342653

Joshi, K. (2018). A New Approach of Text Steganography Using ASCII Values. *International Journal of Engineeering Research & Technology*, *7*(5), 490–493.

Kant, C., Nath, R., & Chaudhary, S. (2008). Biometrics Security using Steganography. *International Journal of Security*, *2*, 1–5.

Kataria, S., Singh, B., Kumar, T., & Shekhawat, H. S. (2013). PDAC (Parallel Encryption with Digit Arithmetic of Cover Text) Based Text Steganography. In *International Conference on Advances In Computer Sciences* (pp. 175–182).

Kaur, M., Gupta, S., Sandhu, P. S., & Kaur, J. (2010). A Dynamic RGB Intensity Based Steganography Scheme. *International Journal of Information and Communication Engineering*, *4*(7), 1170–1173.

Kavitha, P. (2016). A Survey on Lossless and Lossy Data Compression Methods. *International Journal of Computer Science & Engineering Technology (IJCSET)*, *7*(3), 110–114.

Khadim, U., Khan, A., Ahmad, B., & Khan, A. (2015). Information Hiding in Text to Improve Performance for Word Document. *International Journal of Technology and Research*.

Khairullah, M. (2009). A Novel Text Steganography System Using Font Color of the Invisible Characters in Microsoft Word Documents. *Second International Conference on Computer and Electrical Engineering*, 482–484. https://doi.org/10.1109/ICCEE.2009.127

Khairullah, M. (2019). A Novel Steganography Method using Transliteration of Bengali Text. *Journal of King Saud University - Computer and Information Sciences*, *31*(3), 348–366. https://doi.org/10.1016/j.jksuci.2018.01.008

Khan, A. (2015). Robust Textual Steganography. *Journal of Science (JOS)*, *4*(4), 426–434. Retrieved from www.worldsciencepublisher.org

Kingslin, S., & Kavitha, N. (2015). Evaluative Approach towards Text Steganographic Techniques. *Indian Journal of Science and Technology*, *8*(November), 1–8. https://doi.org/10.17485/ijst/2015/v8i2

Kingslin, S., & Saraswathi, V. (2015). Providing security to Online Shopping using credit cards through Text Steganography Techniques. *International Journal of*

Advanced Research in Computer and Communication Engineering, *4*(6), 106–108. https://doi.org/10.17148/IJARCCE.2015.4624

Kodituwakku, S. R., & Amarasinghe, U. S. (2014). Comparison of Lossless Data Compression Algorithms for Text Data. *Indian Journal of Computer Science and Engineering*, *1*(4), 416–425.

Koley, S., & Mandal, K. K. (2016). A Novel Approach of Secret Message Passing Through Text Steganography. In *International Conference on "Signal Processing, Communication, Power and Embedded System (SCOPES-2016)*.

Koley, S., & Mandal, K. K. (2017). Number System Oriented Text Steganography in Various Language for Short Messages. In J. K. Mandal, P. Dutta, & S. Mukhopadhyay (Eds.), *Computational Intelligence, Communications, and Business Analytics: International Conference* (1st ed., Vol. 1, pp. 552–566). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-10-6427-2

Kothari, L., Thakkar, R., & Khara, S. (2017). Data hiding on web using combination of Steganography and Cryptography. In *International Conference on Computer, Communications and Electronics* (pp. 448–452).

Kouser, S. (2016). A Novel Content-Based Feature Extraction Approach : Text Steganography. *International Journal of Computer Science and Information Security (IJCSIS), 14*(12), 916–922.

Kouser, S., Khan, A., & Qamar, E. (2017). A Novel Feature Extraction Approach : Capacity Based Zero-Text Steganography. *International Journal on Information Technologies & Security*, *9*(3), 85–98.

Krishnan, R. B., Thandra, P. K., & Baba, M. S. (2017). An overview of text steganography. In *2017 4th International Conference on Signal Processing, Communication and Networking, ICSCN 2017* (pp. 0–5). https://doi.org/10.1109/ICSCN.2017.8085643

Kumar, A., & Pooja, K. (2010). Steganography- A Data Hiding Technique. *International Journal of Computer Applications*, *9*(7), 19–23.

Kumar, K. A., Pabboju, S., & Desai, M. N. S. (2014). Advance Text Steganography Algorithms : an Overview. *International Journal of Research and Applications*, *1*(1), 31–35.

Kumar, R., Malik, A., Singh, S., & Chand, S. (2016). A high capacity E-mel based text steganography scheme using Huffman compression. *3rd International Conference on Signal Processing and Integrated Networks (SPIN) A*, 53–56.

Kumar, R., Malik, A., Singh, S., Kumar, B., & Chand, S. (2017). A space based reversible high capacity text steganography scheme using font type and style. *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2016*, 1090–1094. https://doi.org/10.1109/CCAA.2016.7813878

Kumar, T., Abhinav, P., Jyoti, K., & Nehra, M. S. (2014). Development of Crossover and Encryption Based Text Steganography (CEBTS) Technique. In S. Sengupta, K. Das, & G. Khan (Eds.), *Emerging Trends In Computing and Communication*

(Vol. 298, pp. 103–122). New York: Springer New Delhi. https://doi.org/10.1007/978-81-322-1817-3

Lee, I., & Tsai, W. (2010). A new approach to covert communication via PDF files. *Signal Processing*, *90*(2), 557–565. https://doi.org/10.1016/j.sigpro.2009.07.022

Lee, M. Y., Iranmanesh, V., & Quiroz, J. C. (2015). A New Approach to SMS Steganography using Mathematical Equations. In *International Conference on Computer Applications & Technology*.

Lwin, T., & Phyo, S. W. (2014). Information Hiding System using Text and Image Steganography. *International Journal of Scientific Engineering and Technology Research*, *3*(10), 1972–1977.

Mahajan, P. (2014). Steganography : A Data Hiding Technique. *International Journal of Advanced Research in Computer Science and Software Engineering*, *4*(11), 759–763.

Mahajan, S., & Singh, A. (2012). A Review of Methods and Approach for Secure Stegnography. *International Journal of Advanced Research in Computer Science and Software Engineering*, *2*(10), 67–70.

Mahato, S., Khan, D. A., & Yadav, D. K. (2017). A Modified Approach to Data Hiding in Microsoft Word Documents by Change-Tracking Technique. *Journal of King Saud University Computer and Information Sciences*.

Mahato, S., Yadav, D. K., & Khan, D. A. (2014). A Novel Approach to Text Steganography Using Font Size of Invisible Space Characters in Microsoft Word Document. In *Intelligent Computing, Networking, and Informatics. Advances in Intelligent Systems and Computing* (pp. 1047–1054). Springer New Delhi.

Maiti, S., & Samanta, D. (2010). Clustering Web Search Results to Identify Information Domain. In *Emerging Trends In Computing and Communication* (pp. 328–335). https://doi.org/10.1080/19393555.2010.514890

Majumder, A., & Changder, S. (2013). A Novel Approach for Text Steganography: Generating Text Summary Using Reflection Symmetry. *Procedia Technology*, *10*, 112–120. https://doi.org/10.1016/j.protcy.2013.12.343

Malalla, S., & Shareef, F. R. (2017). A Novel Approach for Arabic Text Steganography Based on the " BloodGroup " Text Hiding Method. *Engineering, Technology & Applied Science Research*, *7*(2), 1482–1485.

Malik, A., Sikka, G., & Verma, H. K. (2016). A High Capacity Text Steganography Scheme Based on LZW Compression and Color Coding. *Engineering Science and Technology, an International Journal*, 4–11. https://doi.org/10.1016/j.jestch.2016.06.005

Malik, A., Sikka, G., & Verma, H. K. (2017). A High Capacity Text Steganography Scheme Based on Huffman Compression and Color Coding. *Journal of Information and Optimization Sciences, 38*(5), 647–664. https://doi.org/10.1080/02522667.2016.1197572

Mandal, K. K., Chatterjee, S., & Chakraborty, A. (2019). *Applying Encryption Algorithm on Text Steganography Based on Number Applying Encryption Algorithm on Text Steganography Based on Number System*. Springer Singapore.

https://doi.org/10.1007/978-981-13-8687-9

Mandal, K. K., Jana, A., & Agarwal, V. (2014). A New Approach of Text Steganography Based on Mathematical Model of Number System. In *International Conference on Circuit, Power and Computing Technologies* (pp. 1737–1741).

Mandal, K. K., Koley, S., & Dhar, S. (2016). A Mathematical Model for Secret Message Passing Using Steganography. In *IEEE International Conference on Computational Intelligence and Computing Research* (Vol. , pp. 1–6).

Mandal, P. C. (2012). An Extensive Review of Current Trends in Steganalysis. *Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *1*(7), 215–220.

Maniriho, P., & Ahmad, T. (2017). A Data Hiding Approach Using Enhanced-RDE in Grayscale Images. In *2017 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA)* (pp. 35–40). IEEE.

Memon, J. A., Khowaja, K., & Kazi, H. (2008). EVALUATION OF STEGANOGRAPHY FOR URDU / ARABIC TEXT . *Journal of Theoretical and Applied Information Technology*, 232–237.

Mihaela, L. (2011). Survey of the Use of Steganography over the Internet, *15*(2), 153–164.

Mokrzycki, W., & Tatol, M. (2011). Color difference Delta E - A survey. *Machine Graphics and Vision*, (April 2011), 383–411.

Moraldo, H. H. (2012). An Approach for Text Steganography Based on Markov Chains. *Proceedings of the 4th Workshop de Seguridad Informatica,(WSI"12)*, (4), 26–39. Retrieved from http://scisweb.ulster.ac.uk/~kevin/papers/textstego/Text stego based on markov chains.pdf%5Cnhttps://github.com/hmoraldo/markovTextStego

Mulunda, C. K., & Wagacha, P. W. (2013). Genetic Algorithm Based Model in Text Steganography. *African Journal of Information System*, *5*(4).

Nagarhalli, T. P., Bakal, J. W., & Jain, N. (2016). A Survey of Hindi Text Steganography. *International Journal of Scientific & Engineering Research*, *7*(3), 55–61.

Naharuddin, A., Wibawa, A. D., & Sumpeno, S. (2018). A High Capacity and Imperceptible Text Steganography Using Binary Digit Mapping on ASCII Characters. In *International Seminar on Intelligent Technology and Its Applications (ISITIA)* (pp. 287–292). IEEE.

Naqvi, N., Abbasi, A. T., Hussain, R., Khan, M. A., & Ahmad, B. (2018). Multilayer Partially Homomorphic Encryption Text Steganography (MLPHE-TS): A Zero Steganography Approach. *Wireless Personal Communications*, *103*(2). https://doi.org/10.1007/s11277-018-5868-1

Odeh, A., Elleithy, K., Faezipour, M., & Abdelfattah, E. (2015). Highly efficient novel text steganography algorithms. *2015 IEEE Long Island Systems, Applications and Technology Conference, LISAT 2015*.

https://doi.org/10.1109/LISAT.2015.7160209

Odeh, A., & Khaled, E. (2013). Steganography in Text by Merge ZWC and Space Character. In *28th International Conference on Computers and Their Applications, CATA-2013, At Honolulu, Hawaii, USA*.

Osman, B., Din, R., & Idrus, M. R. (2015). Capacity Performance of Steganography Method in. *ARPN Journal of Engineering and Applied Sciences*, *10*(3), 1345–1351.

Osman, B., Din, R., Zalizam, T. M., & Omar, M. N. (2013). A Performance of Embedding Process for Text Steganography Method. *6th WSEAS World Congress: Applied Computing Conference (ACC '13)*, 115–119.

Osman, B., Yasin, A., & Omar, M. N. (2016). An Analysis of Alphabet-based Techniques in Text Steganography. *Journal of Telecommunication, Electronic and Computer Engineering*, *8*(10), 109–115.

Patel, P., & Patro, S. P. (2017). Analysis of Information Security through Crypto-Stenography with Reference to E-Cipher Methods. *International Journal of Advanced Research in Computer and Communication Engineering*, *6*(11), 332–336. https://doi.org/10.17148/IJARCCE.2017.61158

Pawar, S. S., & Kakde, V. (2014). Review on Steganography for Hiding Data. *International Journal of Computer Science and Mobile Computing*, *3*(4), 225–229.

Por, L. Y., & Delina, B. (2008). Information Hiding : A New Approach in Text Steganography. In *Intrenational Conference On Applied Computer and Applied Computational Science* (pp. 689–695).

Por, L. Y., & Delina, B. (2008). WhiteSteg: A New Scheme in Information Hiding Using Text Steganography. *7th WSEAS Int. Conf. on Applied Computer & Applied Computational Science (ACACOS)*, *7*(6), 689–695. Retrieved from http://www.wseas.us/e-library/conferences/2008/hangzhou/acacos/116-586-634.pdf

Por, L. Y., Wong, K., & Chee, K. O. (2012). UniSpaCh: A text-based data hiding method using Unicode space characters. *Journal of Systems and Software*, *85*(5), 1075–1082. https://doi.org/10.1016/j.jss.2011.12.023

Potdar, V. M., Han, S., & Chang, E. (2005). Dictionary Module and UDC : Two new approaches to Enhance Embedding Capacity of a Steganographic Channel. *2005 3rd IEEE International Conference on Industrial Informatics*, (August).

Pujari, A. A., & Shinde, S. S. (2016). Data Security using Cryptography and Steganography. *IOSR Journal of Computer Engineering (IOSR-JCE)*, *18*(4), 130–139. https://doi.org/10.9790/0661-180405130139

Rafat, K. F., & Hussain, M. J. (2017). Secure Text Steganography for Microsoft Word Document. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, *11*(6), 722–727.

Rahman, M. S., Khalil, I., Yi, X., & Dong, H. (2017). Highly Imperceptible and Reversible Text Steganography Using Invisible Character based Codeword Highly Imperceptible and Reversible Text Steganography Using Invisible

Character based Codeword. In *Pacific Asia Conference on Information System (PACIS)*.

Ramakrishnan, B. K., Thandra, P. K., & Srinivasula, A. V. S. M. (2017). Text steganography : A Novel Character-Level Embedding Algorithm Using Font Attribute. In *Security And Communcation Networks*. https://doi.org/10.1002/sec.1757

Rani, M., & Singh, V. (2016). A Survey On Lossless Text Data Compression Techniques. *International Journal of Advanced Research in Computer Engineering & Technology*, *5*(6), 1–5.

Rasmi, A., & Mohanapriya, M. (2016). An Extensive Survey of Data Hiding Techniques. *Indian Journal of Science and Technology*, *9*(28), 1–7. https://doi.org/10.17485/ijst/2016/v9i28/90457

Rauf, A., Rose, H., Jamal, & Nur Hafizah. (2014). Feasibility of Text Visualization in Text Steganalysis. In *New Trends in Software Methodologies, Tools and Techniques* (pp. 103–115).

Ray, R., Sanyal, J., Das, D., & Nath, A. (2012). A New Challenge of Hiding any Encrypted Secret Message Inside any Text / ASCII File or in MS Word File : RJDA Algorithm. In *International Conference on Communication System and Network Technologies* (pp. 889–893). https://doi.org/10.1109/CSNT.2012.191

Rodríguez, J. M., Pesado, P., Merlino, H. D., & Martí nez, R. G. (2018). Evaluation of Open Information Extraction Methods Using Reuters-21578 Database. In *International Conference on Machine Learning and Soft Computing* (pp. 87–92). https://doi.org/10.1145/3184066.3184099

Roslan, N. A., Mahmod, R., Udzir, N. U. R. I., & Zurkarnain, Z. A. (2014). Primitive Structural Method for High Capacity. *Journal of Theoretical and Applied Information Technology*, *67*(2), 373–383.

Roy, S., & Manasmita, M. (2011). A novel approach to format based text steganography. *Proceedings of the 2011 International Conference on Communication, Computing & Security - ICCCS "11*, 511. https://doi.org/10.1145/1947940.1948046

Roy, S., & Venkateswaran, P. (2013). A Text based Steganography Technique with Indian Root. *Procedia Technology*, *10*, 167–171. https://doi.org/10.1016/j.protcy.2013.12.349

Sabri, R. S., Din, R., & Mustapha, A. (2018). Analysis Review on Performance Metrics for Extraction Schemes in Text Steganography. *Indonesian Journal of Electrical Engineering and Computer Science*, *11*(2), 761–767. https://doi.org/10.11591/ijeecs.v11.i2.pp761-767

Salomon, D. (2003). *Data Privacy and Security* (1st ed.). New York: Springer New York. https://doi.org/10.1007/978-0-387-21707-9

Samer, T. A., Zaid, N. K., Sami, K. K., & Zainab, N. K. (2015). Improve Capacity in Text in Text Steganography. *European Academic Research*, *II*(12), 15049–15062.

Saniei, R., & Faez, K. (2013). The capacity of arithmetic compression based text

steganography method. *Iranian Conference on Machine Vision and Image Processing, MVIP*, *3*(6), 38–42. https://doi.org/10.1109/IranianMVIP.2013.6779946

Saraswathi, V., & Kingskin, S. (2014). Different Approaches to Text Steganography: A Comparison. *International Journal of Research in Management and Technology*, *9359*(11), 124–127.

Satir, E., & Isik, H. (2012a). A compression based text steganography method. *Multimedia Tools and Applications*, *70*(3), 2085–2110. https://doi.org/10.1007/s11042-012-1223-9

Satir, E., & Isik, H. (2012b). A Huffman compression based text steganography method. *Multimedia Tools and Applications*, *70*(3), 2085–2110. https://doi.org/10.1007/s11042-012-1223-9

Shirali-shahreza, M. H. (2006). A New Approach to Persian / Arabic Text Steganography. In *Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science* (pp. 1–6).

Shivani, Yadav, V. K., & Batham, S. (2015). A Novel Approach of Bulk Data Hiding using Text Steganography. *Procedia Computer Science*, *57*, 1401–1410. https://doi.org/10.1016/j.procs.2015.07.457

Sidhu, A. S., & Garg, E. M. (2014). Research Paper on Text Data Compression Algorithm using Hybrid Approach. *International Journal of Computer Science and Mobile Computing*, *3*(12), 1–10.

Simmons, G. J. (1984). *The Prisoners" Problɯ and the Subliminal Channel*. Springer US. https://doi.org/1007/978-1-4684-4730-9_5

Singh, H., & Diwakar, A. (2014). An Intelligent Approach for Secure Data Transmission using Text Steganography. In *International Congress on Computer, Electronics, Electrical, and Coomunication Engineering* (Vol. 59, pp. 13–17). https://doi.org/10.7763/IPCSIT.2014.V59.3

Singh, H., Diwakar, A., & Upadhyaya, M. S. (2014). A Novel Approach to Text Steganography. In *International Congress on Computer, Electronics, Electrical, and Communication Engineering* (Vol. 59, pp. 7–12). https://doi.org/10.7763/IPCSIT.2014.V59.2

Singh, H., Singh, P. K., & Saroha, K. (2009). A Survey on Text Based Steganography. In *Proceedings of the 3rd National Conference; INDIACom-2009* (Vol. 3, pp. 332–335).

Singh, S., & Singh, A. (2013). A Review on the Various Recent Steganography Techniques. *International Journal of Computer Science and Network*, *2*(6), 142–156.

Sloan, T., & Hernandez-castro, J. (2015). Forensic Analysis of Video Steganography Tools PrePrints PrePrints. *PeerJ Computer Science*, (May), 1–14.

Srikanth, P., Mehta, A., Yadav, N., Singh, S., & Singhal, S. (2017). Encryption and Decryption Using Genetic Algorithm Operations and Pseudorandom Number. *International Journal of Computer Science and Network*, *6*(3), 455–459.

Stephen, M. J., Reddy, P., Naidu, D., Sonali, S., & Heymaraju. (2012). More Secured Text Transmission with Dual Phase Message Morphing Algorithm. In *Proceedings of the International Conference on Information Systems Design and Intelligent Applications* (pp. 845–852).

Stojanov, I., Mileva, A., & Stojanovi, I. (2014). A New Property Coding in Text Steganography of Microsoft Word Documents A New Property Coding in Text Steganography of Microsoft Word Documents. In *SECURWARE 2014 : The Eighth International Conference on Emerging Security Information, Systems and Technologies*.

Sumathi, C. P., Santanam, T., & Umamaheswari, G. (2013). A Study of Various Steganographic Techniques Used for Information Hiding. *International Journal of Computer Science & Engineering Survey*, *4*(6), 9–25. https://doi.org/10.5121/ijcses.2013.4602

Sunariya, U., Din, R., & Mahmudin, M. (2016). Critical Analysis on Steganography Technique in Text Domain. In *Proceedings of the 5th International Cryptology and Information Security Conference 2016 (CRYPTOLOGY2016)*.

Tan, Y., Tan, W., & Guo, X. (2013). Integrated Lossy and Lossless Compression with LSB Insertion Technique in Steganography. In *International Conference on Digital Image Processing* (Vol. 8878, pp. 18–22). https://doi.org/10.1117/12.2031061

Tang, X., & Chen, M. (2013). Design And Implementation Of Information Hiding System Based On RGB. In *IEEE-3rd International Conference on Consumer Electronics, Communications and Networks* (pp. 217–220).

Tiwari, R. K., & Sahoo, G. (2011). A Novel Methodology for Data Hiding in PDF Files. *Information Security Journal: A Global Perspective*, *20*(1), 45–57. https://doi.org/10.1080/19393555.2010.544703

Tutuncu, K., & Hassan, A. A. (2015). New Approach in E-mail Based Text Steganography. *International Journal of Intelligent Systems and Applications in Engineering*, *3*(2), 54–56.

Vennice, M. G., Rao, T., Swapna, M., & Sasi, K. J. (2012). Hiding the Text Information using Stegnography. *International Journal of Engineering Research and Applications*, *2*(1), 126–131.

Wang, F., Huang, L., Chen, Z., Yang, W., & Miao, H. (2013). A Novel Text Steganography by Context-based Equivalent Substitution. *IEEE International Conference on Signal Processing, Communications and Computing, ICSPCC 2013*, 1–6. https://doi.org/10.1109/ICSPCC.2013.6663950

Wang, X., & Li, H. (2014). Research on Information Hiding Method Based on Word Text. *Advanced Materials Research*, *930*, 2815–2818. https://doi.org/10.4028/www.scientific.net/AMR.926-930.2815

Win, T. Z., & Oo, A. W. (2018). Data Hiding Process with Indexes Based on Whitespace Method. *International Journal of Engineering Trends and Applications (IJETA)*, *5*(4), 32–35.

Xiang, L., Sun, X., Luo, G., & Xia, B. (2014). Linguistic Steganalysis Using the

Features Derived From Synonym Frequency. *Multimedia Tools and Applications*, *71*, 1893–1911. https://doi.org/10.1007/s11042-012-1313-8

Xiao, C., Zhang, C., & Zheng, C. (2018). FontCode : Embedding Information in Text Documents Using Glyph Perturbation. *ACM Transactions on Graphics, 37*(2).

Yang, H., & Cao, X. (2010). Linguistic Steganalysis Based on Meta Features and Immune Mechanism. *Chinese Journal of Electronics*, *19*(4), 661–666. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-78649888557&partnerID=40&md5=a2c67101faac65278c0dcc7d844e41e5

Zaidan, B. B., Zaidan, A. ., Al-Frajat, A. ., & Jalab, A. . (2010). On the Differences between Hiding Information and Cryptography Techniques-An Overview.pdf. *Journal of Applied Sciences*, *10*(15), 1650–1655.

Zhang, J., Huang, H., Wang, L., Lin, H., & Gao, D. (2017). Coverless Text Information Hiding Method Using the Frequent Words Hash. *International Journal of Network Security*, *19*(6), 1016–1023. https://doi.org/10.6633/IJNS.201711.19(6).18

Zielinska, E., Mazurczyk, W., & Szczypiorski, K. (2012). The Advent of Steganography in Computing Environments. *arXiv*, *abs/1202.5*. Retrieved from http://dblp.uni-trier.de/db/journals/corr/corr1202.html#abs-1202-5289

Zielińska, E., Mazurczyk, W., & Szczypiorski, K. (2014, March). Trends in Steganography. *Communications of the ACM*, *57*(3), 86–95. https://doi.org/10.1145/2566590.2566610

Lampiran A

Peratus Kekerapan Aksara di dalam Sesuatu Dokumen

Letter	Frequency of letter	Number assigned	Letter	Frequency of letter	Number assigned
E	11.1607 %	15	M	3.0129 %	7
A	8.4966 %	14	H	3.0034 %	7
R	7.5809 %	13	G	2.4705 %	6
I	7.5448 %	13	B	2.0720 %	5
O	7.1635 %	12	F	1.8121 %	4
T	6.9509 %	11	Y	1.7779 %	4
N	6.6544 %	11	W	1.2899 %	3
S	5.7351 %	10	K	1.1016 %	3
L	5.4893 %	10	V	1.0074 %	3
C	4.5388 %	9	X	0.2902 %	2
U	3.6308 %	8	Z	0.2722 %	2
D	3.3844 %	8	J	0.1965 %	1
P	3.1671 %	7	Q	0.1962 %	0

Lampiran B

Perbezaan Antara Teks Pelindung dan Teks Stego (CEBTS, Kumar et al.,2014)

Newspapermen make news. The just published Jobs Rated Report by careercast.com has ousted the lumberjack from the position of worst profession. Taking its place is the newspaper reporter. Says the survey: " Lumberjack is a job that has lost its lustre dramatically over the past five years and is expected to plummet even further by 2020". In other words, like the village blacksmith and the town crier, the traditional journalist won't be left with a profession after a few years. According to the Jobs Rated Report, a good job is one that has demand and scope, low stress levels and a great work environment. 'The best jobs of 2013 are: actuary, biomedical engineer, audiologist and financial planner. The worst jobs of 2013 are: reporter (news-paper), lumberjack, enlisted military personnel, actor and oil rig worker. All these careers have very high stress levels though some may have quite supportive work environments. The Career-Cast report is into its 25th edition now. Last year the best job was software engineer. That's in decline phase, there has been a drop in hiring outlook. The other ousted professions are HR manager and dental hygienist. At the bottom end, the dairy farmer has made an exit. A great job does not have to be glamorous. Topping the list for women is the soldier, which is in the dumps at Career-Cast. Men like yoga teachers and they like women working in sports and recreation. Nobody wants the insurance agents or curiously enough, people working in consultancy and strategy. Somewhere or the other, they have slipped up on strategy; they need a Sudhir Kakkar for consultancy. In India, most lists are derivative: some HR expert sets himself up as an authority and extrapolates from a survey conducted elsewhere. But here is a survey from apex chamber Assocham polled B-school grads. In Tier II and Tier III towns, the vast majority 85 percent would prefer to join a public sector undertaking (PSU). The principal attraction, says Assocham, is "100% stability". It's the old story: once you get a PSU job, mediocrity takes over fro meritocracy. In the metros, however, it was a different story. Nearly, 90 percent of B-school students said that leading corporate

Teks Pelindung

Newspapermen make news. Says the survey: " Lumberjack is a job that has lost its lustre dramatically over the past five years and is expected to plummet even further by 2020". The best jobs of 2013 are: actuary, biomedical engineer, audiologist and financial planner. The other ousted professions are HR manager and dental hygienist. A great job does not have to be glamorous. Topping the list for women is the soldier, which is in the dumps at Career-Cast. In India, most lists are derivative: some HR expert sets himself up as an authority and extrapolates from a survey conducted elsewhere.

Teks Stego

Lampiran C

Kekunci Stego Menggunakan Teknik Pemetaan Aksara ASCII
(Naharuddin et al., 2018)

121112511211131222622112651111122411221213213241112214151426121121
111141152426224212112111412261111122111223232441142111212126111143
121311115311511251131534213212113132612312311121126253214352342151
333141111131111111213121111211532232223253554312311111113265232113
122112261513251122121513421421211323165132221243312111112511122122
121113412421212313212131211641521516423113121231261113111222213514
412312121113161231232641112121411213216213311121121355331241122231
111113533111122611311111114235143113222121112632112221331211221213
211212132115341231122231316111222121223111212262114121421131554432
411211311112231151112412223126213111211332155443141112231111113111
113312111112512115124113113412422212323263522133122112261131211123
523554421211111221111121212114632112613124525111621121311651111126
411113111116435132111111226211211323421411222323161321412123226252
121122422155341131122212111221124321131261521211112253154412312111
111211312312111216152512232212315441142121212311513224311113226 21
152111112535544114111211211612111122133121321125211112211115543213
112223231621322311151231364231131211111625222231151115544113221121
131111111326421123113121231212512223115311113412322212111213535412
121121211211522253113412411112121263121122231212112621232114122 13
553321311221111112122111112243222112615212423212131534123222121112
121313525123311131211122221122253154432311122121211113322411221 22
621242113221115543114122223211135111141212111115122111211515133

Lampiran D

Ringkasan Kajian Lepas

Bil	Penulis	Tahun	Teknik	Jumlah bit Penyembunyian	Kapasiti Penyembunyian	Kelebihan	Kelemahan
1	Por L. Y., Wong K. & Chee K.O.	2012	UniSpaCh Menggunakan aksara kosong Unicode yang berbeza - ruang kosong antara perkataan dan antara ayat. - Ruang kosong pada hujung baris dan hujung perenggan.	2 bit mesej rahsia diwakili dengan setiap ruang kosong.	20%	Memanipulasi aksara kosong Unicode Tidak memampar simbol janggal „‟ atau „→ “ apabila di analisis menggunakan DASH	1. Berlaku perubahan terhadap CT 2. Hanya boleh digunakan untuk dokumen Ms Word sahaja 3. Tidak teguh (sepeti yang dinyatakan di dalam kajian tersebut)
2	Majumder, A. & Changder, S.	2013	Simetri Refleksi Membahagikan aksara kepada 4 kumpulan berdasarkan kepada atribut refleksi simetri melintang atau menegak sesuatu abjad. Menyembunyikan SM pada aksara pertama setiap ayat berdasarkan kepada atribut di atas.	2 bit mesej rahsia dipadankan dengan 1 aksara	0.41%	1. Menjana teks stego yang baru.	1. Menjana teks stego yang berbeza dengan teks pelindung yang mendorong kepada kecurigaan. 2. Menyembunyikan 2 bit pada setiap abjad pertama ayat tidak sesuai untuk mesej yang panjang. 3. Kapasiti mesej yang disembunyikan adalah rendah

3	Mulunda C. K., Wagacha P.W. & Adede A.O.	2013	GATS Menyembunyikan HM pada setiap aksara ke *6i*. Nilai ASCII pada lokasi *2n* akan ditukarkan dengan nilai ASCII SM. *(i=1,2,... length(SM))* CT dijana menggunakan teknik GA. Setiap nilai ASCII SM digantikan dengan nilai ASCII CT pada lokasi ke *6i*.	8 bit SM disembunyikan pada 1 aksara CT.	16.6%	1. Menggunakan teknik penyulitan dan penyahsulitan terhadap SM 2. Tempoh penyembunyian yang singkat	1. Teks stego yang dihasilkan adalah dalam bentuk nilai ASCII. 2. Menjana teks stego yang berbeza dengan teks pelindung yang boleh mendorong kepada kecurigaan. 3. Menyembunyikan mesej berdasarkan konsep LSB yang menyebabkan perubahan minimum terhadap teks pelindung 4. Saiz teks stego bergantung kepada saiz mesej.
4	Roy S. & Venkateswaran P.	2013	Taburan Frekuensikuansi Abjad Membahagikan aksara tersembunyi kepada kumpulan 4 bit dan dipadankan dengan Jadual Peratusan Frekuensikuansi. Teks stego dijana berdasarkan aksara yang dipilih di atas.	4 bit SM dipadankan dengan 1 aksara	5.7%	1. Menjana teks stego tanpa teks pelindung. 2. Perwakilan aksara yang pelbagai tetapi terhad.	1. Pembentukan struktur ayat yang sukar kerana pembentukan ayat bergantung kepada perwakilan aksara. 2. Menyembunyikan mesej pada huruf pertama setiap aksara. 3. Kapasiti penyembunyian rendah 4. Kapasiti penyembunyian rendah

5	Kumar, et al.,	2014	CEBTS CT dijana dlm bentuk *ciphertext*. Menambahkan aksara SM yang telah dilakukan proses XOR ke dalam CT pada lokasi-*5i* (*i=1,2,3... length(SM)*)	8 bit SM (1 aksara) ditambah ke dalam CT yang telah dijana secara rawak.	24.4%	1. Menjana CT dalam bentuk ciphertext yang dijana secara rawak. 2. Teknik ini teguh terhadap pemformatan teks dan *retyping*.	1. Menjana teks stego yang berbeza dengan teks pelindung yang mendorong kepada kecurigaan. 2. Kekunci stego dijana dan disembunyikan ke dalam teks stego akan menyebabkan berlakunya perubahan terhadap teks stego.
6	Mandal et al.	2014	Memperkenalkan skima Model Matematik Sistem Nombor (MMSN) Menukar aksara SM (ASCII) ke bentuk koordinat (x,y) menggunakan formula ((x*(x+1)/2) + y)	8 bit SM (1 aksara) memerlukan 1 aksara CT	Tiada maklumat kapasiti penyembunyian kerana penyelidik hanya memperkenalkan skema	1. Menggunakan teknik penyembunyian dengan menukarkan nilai ASCII SM ke bentuk kordinat (x,y). 2. Tidak menggunakan CT sebagai asas untuk penyembunyian mesej rahsia.	1. Mewakilkan aksara A-Z dengan 26 koordinat (x,y) sahaja. 2. Aksara SM yang berulang diwakilkan dengan nilai yang sama. 3. ST sukar dijana berdasarkan kepada koordinat. 4. Julat koordinat yang dihasilkan hanya tertumpu pada kawasan tertentu sahaja.
7	Fuad M. N.	2014	Model Matematik Sistem Nombor Mewakilkan setiap abjad dengan koordinat (Menambahbaik formula Mandal)	8 bit SM memerlukan 4 bait	Hanya memperkenalkan skema pertukaran nilai ASCII	1. Julat penyembunyian lebih luas daripada teknik Mandal.	1. Menambahbaik formula Mandal. 2. Julat kordinat yang yang dihasilkan masih tidak menyeluruh. 3. Aksara SM yang berulang diwakilkan dengan nilai yang sama.

8	Mandal, Koley, & Dhar	2016	Menambahbaik model MMSN dengan menambahkan simbol (++,-+,--,+-) pada setiap koordinat yang dihasilkan	8 bit SM (1 aksara) memerlukan 1 aksara CT	Tiada maklumat kapasiti penyembunyian kerana penyelidik hanya memperkenalkan skema	4. Setiap 2 nilai (x,y) berada pada lokasiyang berbeza. 5. Julat kordinat lebih luas berbanding teknik sebelumnya.	1. Mewakilkan aksara A-Z dengan 26 koordinat (x,y) sahaja. 2. Aksara SM yang berulang diwakilkan dengan nilai yang sama. 3. CT sukar dijana berdasarkan kepada koordinat. 4. Julat koordinat yang dihasilkan hanya tertumpu pada kawasan tertentu (masih terhad).
9	Htet M. & Phyo S.W	2016	Menyembunyikan bit SM pada setiap perkataan pada lokasi (1^{st},2^{nd} , 2^{nd} akhir ,terakhir)	1 bit SM memerlukan 1 bait	3.51%	1. Menyembunyikan SM pada huruf pertama, kedua, kedua terakhir atau terakhir bagi sesuatu perkataan. 2. Boleh digunakan dalam format .txt 3. Tiada sebarang perubahan terhadap CT.	1. Menyembunyiakn SM pada lokasi berjujukan. 2. Satu bit disembunyikan pada satu aksara. 3. Saiz kekunci stego bergantung kepada saiz SM. 4. Kapasiti penyembunyian rendah kerana 1 bit disembunyikan pada 1 aksara. 5. Tidak sesuai untuk mesej yang panjang 6. Lokasi penyembunyian secara jujukan.
10	Malik A, Sikka G. &Verma H. K.	2016	Pemampatan LZW dan Pengkodan Warna	8 bit mesej rahsia diwakili oleh 1 aksara teks pelindung	13.43%	1. Memampatkan mesej rahsia sebelum proses penyembunyian. 2. Mengurangkan kompleksiti pengkomputan.	1. Perubahan warna pada teks stego amat ketara dan boleh menimbulkan kecurigaan. 2. Tidak sesuai dddiiigunakkkan untuk dokumen berbentuk *plain*.

11	Mandal & Koley	2016	Mengimplimentasikan skema MMSN yang telah diperkenalkan Menukar aksara SM (ASCII) ke bentuk koordinat (x,y) menggunakan formula (M (M + 1)) / 2 + N)	8 bit SM (1 aksara) memerlukan 1 aksara CT.	1.4%	1. Setiap nilai (x,y) berada pada lokasiyang berbeza dalam CT. 2. Mudah untuk proses pengekstrakkan kerana penyembunyian diwakilkan dengan nilai numerik (hari dan bulan).	1. Mewakilkan aksara A-Z dengan 26 koordinat (x,y) sahaja. 2. Sekurang-kurangnya 4 bait perlu ditambah ke dalam CT untuk menyembunyikannya dalam format tarikh *dd/mm/yyyy* 3. Aksara SM yang berulang diwakilkan dengan nilai yang sama. 4. Teks stego sukar dijana berdasarkan kepada format *dd/yy.* 5. Kapasiti penyembunyian rendah
12	Koley & Mandal	2017	Mengimplimentasikan skema MMSN yang telah diperkenalkan Menukar aksara SM (ASCII) kepada Octal dan seterusnya menukarkannya ke bentuk koordinat (x,y) menggunakan formula (M (M + 1)) / 2 + N)	8 bit SM (1 aksara) memerlukan 1 aksara CT.	5.75%	1. Setiap nilai (x,y) berada pada lokasiyang berbeza dalam CT.	1. Mewakilkan aksara A-Z dengan 26 koordinat (x,y) sahaja. 2. Sekurang-kurangnya 4 bait perlu ditambah ke dalam CT untuk menyembunyikannya dalam format (xx,yy). 3. Aksara SM yang berulang diwakilkan dengan nilai yang sama. 4. Kapasiti penyembunyian rendah

13	Kauser S. & Khan Aihab	2017	Kumpulan aksara dibahagikan kepada 8 kumpulan berdasarkan kepada gabungan ciri-ciri aksara seperti kebolehan menulis dalam satu arah atau sebaliknya, mempunyai garisan melintang atau menegak, mempunyai bentuk cengkung atau sebaliknya	3 bit boleh disembunyikan pada 1 aksara teks pelindung	23.25%	1. Mudah melakukan proses penyembunyian dengan melakukan proses pemetaan bit SM dengan kumpulan yang berkaitan	1. Saiz kekunci hampir 3 kali ganda bilangan aksara SM 2. Kapasiti penyembunyian masih rendah
14	Fateh & Rezvani	2018	Menyembunyikan aksara SM dengan menjana alamat e-mel berdasarkan kepada CT	4 bit boleh disembunyikan pada 1 aksara pada alamat e-mel	10.6%	1. Tidak dihadkan kepada bahasa-bahasa tertentu. 2. Menyembunyikan aksara pada alamat e-mel.	1. Alamat e-mel yang dijana tidak sah dan boleh menimbulakn kecurigaan. 2. Bilangan alamat e-mel kan bertambah bila saiz SM bertambah.
15	Naharuddin et al.	2018	Satu bit aksara SM akan dipetakan dengan lokasi yang sepadan di dalam CT untuk menjana kekunci stego.	1 bit SM memerlukan 1 aksara CT yang dipadankan	14.2%	1. Mudah melakukan proses penyembunyian dengan melakukan proses pemetaan bit SM dengan kumpulan yang berkaitan	1. Saiz kekunci stego yang dijana berkadar terus dengan saiz SM. 2. Kapasiti penyembunyian masih rendah
16	Baawi et al.	2019	Menyuntik simbol Unicode tak bercetak ke dalam CT pada lokasi permulaan, pertengahan, hujung, hujung perkataan atau berasingan	4 bit disembunyikan pada setiap 2 aksara	12.02%	1.	1. Kapasiti penyembunyian masih rendah

Lampiran E

Teks Pelindung dan Jadual *Homophonic*

Langkah 1 : Menginputkan mesej rahsia yang hendak disembunyikan, *m*

***m* = "MEETYOUATTEN"**

Langkah 2 : Memilih teks pelindung, *ct*

1	The Prague Stock Exchange hit a new year-high on Thursday as major banking
63	issues broke out of a recent slump to spur the bourse higher.
113	Komercni Banka and Ceska Sporitelna, the country's two largest banks both posted
182	strong gains to help push the PX50 index up 3.3 points, or 0.58 percent, to a
244	1996 high of 574.9.
260	Overall, advancing issues narrowly outpaced decliners by 320 to 298, with 192
326	holding steady. Total volume remained steady at 916,194 shares on turnover of
392	648.3 million crowns.
411	Komercni, a likely recipient of attention from investors looking to enter the
477	Czech market after an announcement of its inclusion into the Morgan Stanley
541	index, jumped 99 crowns to close at 2,439.
576	Meanwhile, savings bank Ceska Sporitelna also posted a strong gain of 16 crowns
643	to 351.
649	"The inclusion of the Czech Republic in the Morgan Stanley index should help
713	boost issues like Komercni," said Jan Sykora of the brokerage Wood and Company.
780	Earlier in the day, Sporitelna CEO Jaroslav Klapal announced the bank expects
846	1996 net profit to total 2.1 billion crowns, while gross profit will hit at
908	least 6.1 billion.
924	Klapal added that he expected the bank would pay a higher dividend in 1996 than
989	its five crown per share dividend for last year.
1029	"Gross profit, before the creation of reserves and payment of taxes could be
1093	slightly above 6.1 billion crowns," Klapal said. "We are counting on dividends
1160	for this year will be higher than that of last year."
1203	Sporitelna's 1995 after-tax profit, calculated according to Czech accounting
1271	standards, plummetted to 263 million crowns from a previous 980 million crowns
1338	after the bank provisioned heavily for risky loans.
1382	The bank's 1995 gross profit was 9.17 billion crowns but the bank assigned
1444	almost an identical sum, 9.13 billion, to its reserves covering "classified"
1510	loans -- those whose likelihood of recovery ranges from doubtful to hopeless.
1576	Klapal did not say whether his forecasts were calculated according to Czech or
1642	international accounting standards, which vary mainly in terms of allowable
1708	write-offs and depreciation costs.
1739	Even though analysts agreed that the financial situation of the bank was
1800	markedly better this year than it was last year, they said shares of the bank
1863	were overvalued.
1878	Analysts say they are closely watching what appears to be a strong acquisition
1944	of shares in Sporitelna by the usually secretive Czech investment group
2005	Motoinvest and the rival bank Investicni a Postovni Banka a.s.
2058	"The attractiveness of (Sporitelna) could greatly drop if Motoinvest took a more
2127	significant share in management," said Richard Podpiera analyst at the
2188	investment house Atlantk FT.

Langkah 3 : Penjanaan Jadual *Homophonic* - berdasarkan lokasi aksara dalam teks pelindung

```
A:
6,19,26,32,47,49,52,57,79,122,125,126,133,143,161,168,189,243,264,268,271,284,295,3
36,343,354,362,365,376,420,438,483,488,493,495,532,536,568,578,587,594,601,611,612,
622,630,689,693,739,743,750,762,769,776,781,793,805,810,815,819,821,823,836,864,906
,910,926,928,930,937,953,962,964,987,1006,1021,1026,1054,1069,1073,1082,1101,1126,1
128,1131,1138,1169,1185,1189,1194,1199,1212,1219,1226,1236,1241,1245,1261,1273,1276
,1313,1338,1347,1363,1378,1386,1407,1433,1436,1444,1450,1459,1501,1512,1548,1578,15
80,1589,1606,1615,1620,1624,1648,1653,1655,1667,1670,1681,1685,1699,1704,1718,1728,
1749,1751,1757,1765,1773,1777,1783,1794,1798,1801,1820,1824,1829,1832,1837,1845,185
0,1860,1872,1878,1880,1887,1893,1904,1913,1915,1919,1926,1933,1948,1963,1972,2015,2
024,2027,2040,2050,2053,2054,2062,2066,2088,2098,2122,2135,2140,2146,2148,2158,2165
,2175,2176,2178,2183,2203,2206,
B:
56,69,100,121,167,172,308,593,676,713,756,835,869,916,952,1042,1091,1102,1109,1175,
1346,1385,1413,1426,1432,1469,1560,1705,1793,1808,1859,1924,1964,2026,2049,
C:
13,17,82,118,129,148,236,273,296,301,404,416,429,477,480,501,514,555,563,597,637,65
5,667,670,679,733,772,806,829,843,876,945,996,1051,1086,1116,1141,1235,1238,1246,12
47,1256,1259,1262,1263,1303,1332,1420,1458,1490,1499,1541,1605,1614,1617,1625,1626,
1635,1638,1656,1657,1678,1726,1733,1775,1896,1906,1934,1978,1985,1988,2037,2067,209
0,2134,2163,
D:
46,128,181,212,269,298,299,329,337,358,363,543,552,621,700,708,741,768,771,792,831,
931,932,934,948,960,971,975,978,1009,1013,1016,1071,1090,1133,1151,1155,1158,1244,1
250,1275,1278,1290,1360,1443,1453,1508,1536,1557,1582,1584,1623,1629,1669,1672,1720
,1721,1762,1805,1847,1876,2017,2094,2102,2160,2167,2170,
E:
3,9,15,22,28,31,67,73,81,83,99,105,110,116,130,140,147,164,180,196,205,213,234,237,
262,281,297,300,305,335,350,352,357,361,378,388,414,424,428,433,441,454,469,472,476
,479,486,491,502,504,527,539,544,551,567,577,584,598,608,620,652,666,669,673,684,69
6,701,710,722,727,731,755,760,764,780,785,791,802,807,830,834,839,842,851,887,909,9
33,940,941,944,947,951,969,976,995,1002,1008,1014,1025,1043,1047,1050,1053,1062,106
4,1067,1076,1084,1092,1105,1137,1140,1156,1168,1176,1181,1198,1209,1222,1243,1258,1
286,1289,1316,1341,1345,1359,1362,1384,1431,1442,1454,1483,1485,1488,1493,1507,1521
,1526,1530,1540,1544,1551,1570,1572,1593,1596,1604,1611,1613,1622,1637,1645,1693,17
07,1712,1722,1725,1739,1741,1760,1761,1769,1792,1804,1809,1812,1819,1836,1842,1852,
1858,1864,1866,1869,1875,1891,1895,1900,1918,1925,1950,1960,1968,1977,1980,1984,198
7,1993,1997,2012,2020,2033,2061,2071,2073,2085,2097,2115,2126,2142,2150,2152,2173,2
187,2191,2195,2202,
F:
78,253,391,437,447,489,508,634,663,752,856,896,992,1017,1038,1044,1060,1080,1160,11
92,1220,1231,1309,1339,1368,1403,1505,1538,1553,1562,1601,1698,1715,1716,1770,1789,
1855,1945,2077,2107,2132,2210,
```

G:
7,21,37,62,108,163,187,188,250,276,332,466,531,591,628,629,688,763,888,967,1030,109
6,1148,1179,1253,1270,1395,1440,1497,1550,1632,1664,1747,1758,1910,1932,2000,2095,2
129,2149,
H:
2,18,23,35,38,42,98,106,109,146,175,195,202,204,248,251,322,326,375,475,481,526,581
,651,665,671,683,704,709,754,790,833,884,903,936,939,950,965,968,986,1005,1049,1097
,1164,1177,1180,1184,1188,1260,1344,1361,1383,1430,1518,1523,1533,1567,1592,1595,15
98,1639,1676,1679,1744,1748,1764,1768,1791,1815,1823,1841,1849,1857,1890,1907,1912,
1947,1967,1989,2019,2060,2139,2164,2186,2198,
I:
24,36,60,63,107,120,138,190,210,222,249,274,277,303,320,330,355,398,401,418,422,430
,432,444,451,464,509,512,518,521,541,582,589,606,631,653,659,678,680,698,718,725,73
5,740,784,787,800,857,870,873,885,897,900,904,917,920,966,972,974,979,989,993,1010,
1012,1039,1056,1095,1110,1113,1132,1146,1152,1154,1165,1172,1178,1207,1232,1251,126
8,1297,1300,1318,1326,1329,1354,1356,1365,1372,1404,1414,1417,1439,1452,1457,1470,1
473,1479,1495,1504,1506,1528,1532,1583,1599,1630,1642,1650,1662,1677,1686,1690,1710
,1727,1730,1771,1776,1780,1785,1816,1826,1846,1908,1937,1939,1941,1952,1958,1982,19
90,2009,2022,2030,2036,2039,2048,2069,2083,2106,2112,2128,2131,2133,2143,2159,2162,
2172,2188,
J: 53,547,742,809,
K:
14,59,72,113,124,132,170,411,423,463,485,596,600,726,728,747,759,817,838,924,955,11
24,1349,1374,1388,1435,1529,1576,1796,1803,1862,2029,2052,2121,2209,
L:
87,141,160,197,265,266,289,302,328,344,347,399,400,421,425,460,515,538,564,583,609,
613,656,677,695,707,711,724,783,803,814,818,822,865,871,872,886,901,902,908,918,919
,925,929,959,1020,1089,1094,1099,1111,1112,1125,1129,1173,1174,1193,1210,1237,1240,
1282,1298,1299,1327,1328,1366,1376,1415,1416,1445,1460,1471,1472,1500,1510,1527,153
1,1564,1571,1577,1581,1616,1619,1654,1688,1700,1701,1706,1752,1778,1806,1831,1873,1
881,1897,1901,1961,1973,1974,2025,2086,2093,2100,2179,2205,
M:
51,89,115,349,353,397,413,450,482,503,528,549,576,685,730,774,1075,1284,1285,1296,1
312,1325,1446,1463,1556,1684,1695,1800,1996,2005,2108,2123,2145,2151,2194,
N:
20,27,40,58,61,84,119,123,127,142,151,169,186,191,211,223,238,272,275,283,304,331,3
56,381,385,403,408,417,434,442,446,452,465,470,494,496,497,500,505,513,520,522,533,
537,542,559,579,590,595,610,627,632,641,654,661,681,690,694,699,734,744,770,777,788
,804,824,825,828,837,850,875,880,922,954,977,980,988,1000,1015,1058,1070,1077,1115,
1120,1144,1147,1150,1157,1186,1211,1252,1266,1269,1274,1302,1307,1331,1336,1348,135
8,1379,1387,1419,1424,1434,1441,1451,1455,1475,1496,1513,1549,1585,1631,1643,1647,1
652,1660,1663,1668,1687,1691,1719,1732,1742,1750,1772,1774,1787,1795,1825,1861,1879
,1909,1931,1943,1953,1962,1991,1998,2010,2016,2028,2031,2038,2047,2051,2072,2087,21
13,2130,2136,2144,2147,2153,2177,2189,2196,2207,

O:
12,39,54,71,74,77,92,101,114,136,149,159,173,177,185,194,221,227,242,252,260,287,291,314,327,341,346,380,386,390,402,406,412,436,445,449,457,461,462,468,498,507,519,524,529,557,562,565,604,615,617,626,633,639,644,660,662,686,705,714,715,729,748,751,758,766,767,773,798,808,812,826,855,860,862,874,878,890,895,921,957,998,1018,1032,1037,1045,1057,1059,1079,1087,1103,1114,1118,1142,1149,1161,1191,1205,1230,1248,1255,1264,1292,1301,1305,1311,1319,1330,1334,1352,1357,1369,1377,1397,1402,1418,1422,1447,1474,1478,1491,1511,1519,1524,1534,1535,1537,1542,1555,1558,1566,1568,1586,1602,1627,1634,1640,1651,1658,1697,1702,1714,1731,1734,1745,1786,1788,1854,1867,1898,1923,1930,1942,1944,1956,2002,2006,2008,2042,2045,2076,2081,2091,2104,2109,2111,2119,2120,2124,2169,2199,
P:
4,90,94,135,176,198,199,206,216,220,233,294,431,550,603,616,674,712,775,797,820,841,853,893,927,943,961,1001,1035,1072,1127,1204,1228,1281,1314,1350,1400,1569,1579,1723,1916,1917,1955,2004,2041,2080,2105,2168,2171,
Q: 1935,
R:
5,33,44,55,70,80,96,103,111,117,137,153,162,184,228,235,263,285,286,306,351,377,384,389,405,415,427,448,458,473,484,492,530,556,605,625,638,672,687,732,749,757,761,782,786,799,811,854,877,889,894,970,997,1003,1007,1019,1027,1031,1036,1046,1052,1061,1065,1117,1139,1162,1170,1182,1200,1206,1223,1229,1249,1277,1304,1310,1315,1333,1342,1351,1370,1371,1396,1401,1421,1482,1486,1494,1539,1545,1547,1554,1597,1603,1612,1628,1641,1646,1671,1682,1694,1709,1724,1759,1802,1813,1821,1838,1851,1865,1870,1894,1920,1929,1949,1957,1979,2001,2021,2065,2082,2096,2103,2125,2141,2161,2166,2174,
S:
10,45,50,64,65,68,86,93,104,131,134,156,165,171,178,182,192,201,225,278,279,282,307,333,359,374,379,409,455,459,511,517,534,560,566,586,592,599,602,614,618,623,642,658,691,703,716,719,720,723,738,745,796,813,845,881,891,892,911,991,1004,1022,1033,1034,1063,1068,1085,1093,1121,1130,1159,1166,1195,1203,1214,1271,1279,1308,1321,1337,1355,1373,1380,1390,1398,1399,1408,1425,1437,1438,1448,1461,1481,1484,1489,1502,1503,1514,1520,1525,1552,1573,1574,1588,1600,1607,1609,1665,1673,1696,1717,1735,1737,1754,1756,1779,1799,1817,1830,1833,1844,1848,1853,1883,1885,1886,1899,1921,1927,1938,1946,1951,1954,1970,1976,1994,2013,2034,2043,2056,2074,2075,2079,2116,2127,2138,2157,2181,2192,2201,
T:
1,11,25,41,76,85,91,97,139,145,152,157,166,174,179,183,193,203,224,239,241,293,313,321,334,340,342,360,366,382,435,439,440,443,456,467,471,474,487,490,506,510,523,525,535,561,569,607,619,624,643,650,664,682,692,717,753,789,801,832,844,852,858,859,861,863,898,905,907,912,935,938,946,949,985,990,1023,1040,1048,1055,1078,1081,1098,1145,1163,1183,1187,1190,1196,1208,1221,1225,1233,1242,1254,1267,1272,1287,1288,1291,1340,1343,1382,1405,1428,1429,1449,1456,1477,1480,1517,1561,1565,1587,1594,1608,1621,1633,1644,1649,1661,1666,1692,1711,1729,1736,1743,1755,1763,1766,1767,1781,1784,1790,1810,1811,1814,1822,1827,1834,1840,1856,1884,1889,1905,1914,1922,1928,1940,1959,1966,1981,1995,1999,2007,2014,2018,2035,2044,2059,2063,2064,2068,2084,2099,2110,2117,2118,2137,2154,2182,2184,2185,2193,2197,2204,2208,2211,

```
U:
8,43,66,75,88,95,102,150,200,215,280,292,348,383,499,516,548,657,675,706,721,827,958
,1088,1143,1239,1265,1283,1320,1427,1462,1559,1563,1618,1659,1746,1782,1874,1936,196
9,1971,2003,2092,2200,
V:
261,270,345,387,453,588,816,973,994,1011,1066,1104,1153,1317,1353,1364,1487,1492,154
3,1680,1740,1868,1871,1983,1992,2011,2023,2032,2046,2070,2114,2190,
W:
29,158,288,319,407,558,580,640,765,879,883,899,956,999,1119,1136,1171,1306,1335,1406
,1423,1522,1591,1610,1675,1703,1708,1797,1828,1863,1903,1911,
X: 16,207,214,545,702,840,942,1083,1227,
Y:
30,48,154,290,309,338,364,426,540,697,746,778,794,963,1024,1074,1100,1167,1197,1367,
1375,1546,1590,1683,1689,1753,1807,1818,1835,1843,1882,1888,1892,1902,1965,1975,2101
,2180,
Z: 478,668,1257,1636,1986,
```

Langkah 4 : Penjanaan Kepelbagian Perwakilan Mesej Rahsia

	M	E	E	T	Y	O	U	A	T	T	E	N
C1	503	1968	81	1098	1590	287	1782	568	1477	682	1866	496
C2	1284	577	995	239	338	1032	1283	1131	1884	1914	1809	1953
C3	450	947	1864	11	1167	1534	1143	1313	366	1644	504	1691
C4	1285	297	1968	1981	1683	617	292	264	1884	293	995	595
C5	576	1014	361	1784	309	386	215	1681	607	340	1105	522
C6	730	1604	539	1922	794	1714	200	125	76	1055	1209	824
C7	1284	551	1431	1928	1689	1118	657	1624	1981	1480	441	632
C8	685	1864	947	487	746	1658	43	953	1242	1621	944	275

Lampiran F

Kepelbagaian Saiz Mesej Rahsia

KOD HM	SAIZ SM (Aksara)	MESEJ RAHSIA
SM1	28	ArrivedAtFebruaryTwentySeven
SM2	40	Theimportanceandamountofdatahaveincrease
SM3	67	TheCzechNationalBankbalanceofpaymentfiguresforthefirsthalfoftheyear
SMe	101	FassiosaiditwouldbemoreusefultoputfringecountrieslikeSlovakiaintothefoldwheremorepressurecanbeapplied
SMe	116	TheCzechNationalBankbalanceofpaymentfiguresforthefirsthalfoftheyearwhichwillgiveclearerpictureofthebalanceofservices
SMe	159	TheCzechcrownroseonThursdaytoitshighestlevelagainstitsdollarbasketbuoyedbycommentsbyPrimeMinisterVaclavKlausthatdevaluationwasnotneeddespiteawideningtradegap
SMe1	166*	Behindusingacovertextistohidethepresenceofsecretmessagesthepresenceofembeddedmessagesintheresultingstegotextcannotbeeasilydiscoveredbyanyoneexceptttheintendedrecipient
SMe2	183*	SeveralanalystsslashedratingsonCascadeciting concernthatgrowthratesforitscoreframerelayswitchingbusinessmaybeslowingtoindustrylevelsofthirtytofiftypercentfromitspriorhundredpercentplus
SM4	267	SalzmannwhosaidhecannotbycontractretirefromthebankforatleastonemoreyearrejectschargesthathismembershipintheSenatewhileheadingthecountrylargestbankinggroupwouldconstituteaconflictofinterestThereisnolegalrequirementintheCchecRepublicforlegislatorstosuspendtheirbusiness
SMe3	283*	SteganographyisnotanewareaItdatesbacktofifthcenturyBCHarpagususedharetosendhismessagebykillingitandhidingthemessageinsideitsbellyApersondisguisedashuntercarriedtheharetothedestinationAnotherincidentwasofKingDariusofSusaHistiaeuswasassignedthedutyofshavingtheheadofhismosttrustedslave

SM5	531	IntheresearchareaoftextsteganographyalgorithmsbasedonfontformathaveadvantagesofgreatcapacitygoodimperceptibilityandwideapplicationrangeHoweverlittleworkonsteganalysisforsuchalgorithmshasbeenreportedintheliteratureBasedonthefactthatthestatisticfeaturesoffontformatwillbechangedafterusingfontformatbasedsteganographicalgorithmswepresentanovelsupportvectormachinebasedsteganalysisalgorithmtodetectwhetherhiddeninformationexistsornotalgorithmcannotonlyeffectivelydetecttheexistenceofhiddeninformationbutalsoestimatethehiddeninformation
SMe4	660*	FrancesstateownedelectricityutilityElectricitedeFrancesaidonWednesdayithadmadeaeighthundredandfiftymillionfrancassetgainsellingitstenpercentstakeinSwedensSydkraftABtoNorwaysStatkraftSFItallowsustorealiseabigassetgainandthedealisalsothestartofanalliancewithStatkraftEDFchairmanEdmondAlphanderytoldanewsconferenceEDFofficialssaidthealliancewithStatkraftalsoastateownedutilitywouldmainlybeintheformofcooperationinhydroelectricprojectsinsoutheastAsiaEDFissellingclosetosevenmillionclassAandtwelvepointthreeninefivemillionclassCsharesinSydkraftABtoStatkraftforasumoftwopointonethreebillionfrancsThisallowstheNorwegianfirmtoincreaseitsstakeinSydkrafttofifteenpercent
SM6	834	BehindusingacovertextistohidethepresenceofsecretmessagesthepresenceofembeddedmessagesintheresultingstegotextcannotbeeasilydiscoveredbyanyoneexcepttheintendedrecipientIntheresearchareaoftextsteganographyalgorithmsbasedonfontformathaveadvantagesofgreatcapacitygoodimperceptibilityandwideapplicationrangeLittleworkonsteganalysisforsuchalgorithmshasbeenreportedintheliteratureBasedonthefactthatthestatisticfeaturesoffontformatwillbechangedafterusingfontformatbasedsteganographicalgorithmswepresentanovelsupportvectormachinebasedsteganalysisalgorithmtodetectwhetherhiddeninformationexistsornotThisalgorithmeffectivelydetectttheexistenceofhiddeninformationandalsoestimatethehiddeninformationlengthAsshownbyexperimentalresultsthedetectionaccuracyofouralgorithmreachesashighasninetyninepercentwhentheh iddeninformationlengthisatleastsixteenbits

SMe5	1262*	FrancesstateownedelectricityutilityelectricitedefrancesaidonwednesdayithadmadeaeighthundredandfiftymillionfrancassetgainsellingitstenpercentstakeinswedenssydkraftabtonorwaysstatkraftsfiallowsustorealiseabigassetgainandthedealisalsothestartofanalliancewithstatkraftedfchairmanedmondalphanderytoldanewsconferenceedfofficialssaidthealliancewithstatkraftalsoastateownedutilitywouldmainlybeintheformofcooperationinhydroelectricprojectsinsoutheastasiaedfissellingclosetosevenmillionclassandtwelvepointthreeninefivemillionclasscshareinsydkraftabtostatkraftforasumoftwopointonethreebillionfrancsthisallowstheNorwegianfirmtoincreaseitsstakeinsydkrafttofifteenpercentTheformaltransferofownershipwilltakeplaceinNovemberedftookitsstakeinsydkraftswedensbiggestpowerfirmtwoyearsagobutitchangeditsstrategyinaprilwhenitsNorthelecunitobtainedatwentyfivepercentstakeingraningeverkensabswedensnumbersixpowergroupitlaterraisedthestakeNortheleciajointventurewithskanskaitsbettertohaveabigstakeinasmallcompanythantohaveasmallstakeinabigcompanyalphanderysaidatsydkraftwedidnotevenhaveaseatontheboardatgraningewehaveasharereholderspactallowingustoeffectivelycontroltthecompanyheaddededfhasasharereholderspactwithswedensindustrialfamilyversteeghandjointlythetheownfiftyfourpercentofgraninge
SMe6	1958*	FrancesstateownedelectricityutilityelectricitedefrancesaidonwednesdayithadmadeaeighthundredandfiftymillionfrancassetgainsellingitstenpercentstakeinswedenssydkraftabtonorwaysstatkraftsfiallowsustorealiseabigassetgainandthedealisalsothestartofanalliancewithstatkraftedfchairmanedmondalphanderytoldanewsconferenceedfofficialssaidthealliancewithstatkraftalsoastateownedutilitywouldmainlybeintheformofcooperationinhydroelectricprojectsinsoutheastasiaedfissellingclosetosevenmillionclassandtwelvepointthreeninefivemillionclasscshareinsydkraftabtostatkraftforasumoftwopointonethreebillionfrancsthisallowstheNorwegianfirmtoincreaseitsstakeinsydkrafttofifteenpercentTheformaltransferofownershipwilltakeplaceinNovemberedftookitsstakeinsydkraftswedensbiggestpowerfirmtwoyearsagobutitchangeditsstrategyinaprilwhenitsNorthelecunitobtainedatwentyfivepercentstakeingraningeverkensabswedensnumbersixpowergroupitlaterraisedthestakeNortheleciajointventurewithskanskaitsbettertohaveabigstakeinasmallcompanythantohaveasmallstakeinabigcompanyalphanderysaidatsydkraftwedidno

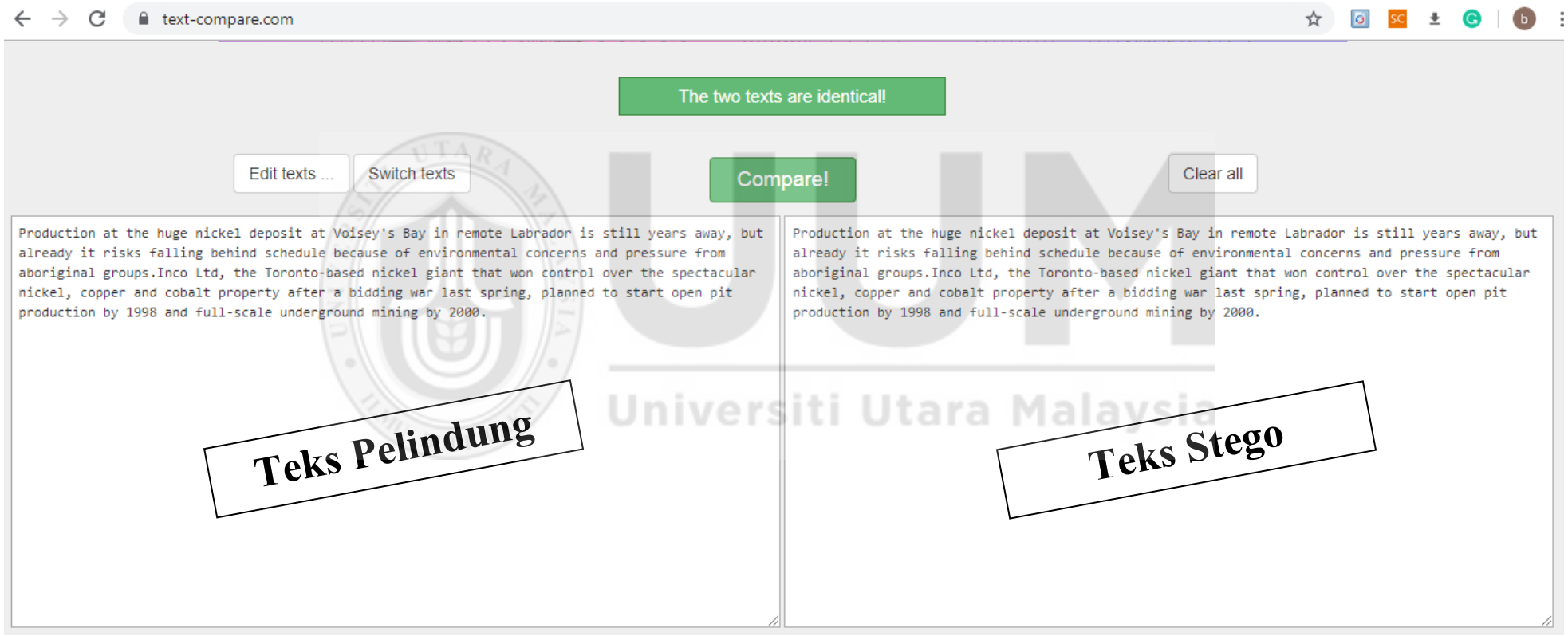
		tevenhaveaseatontheboardatGraningewehaveasha reholderspactallowingustoeffectivelycontrolt hecompanyheaddedEDFhasashareholderspactwithS wedensindustrialfamilyVersteeghandjointlythe yownfiftyfourpercentofGraningeWesawthestrate gicvalueoftheNordicmarketandwehavebeenableto cashinontheassetgainbecausethevalueofSydkraf tshareshasrisenalothesaidAlphanderyaformerec onomicsministernamedtoEDFinNovembertoreplace GillesMenagewhowasanappointeeoftheformerSoci alistgovernmentandanaidetoFrancoisMitterands aidthesalewasproofofEDFswillingnesstorealise anassetgainwhenthereisoneHedidnotexpectfurth ersalesfromEDFsinternationalportfolioTheFren chutilityinninetynietysixalonehasmadecommitm entstotallingfourpointsixbillionfrancsinfore ignstakesandAlphanderysaidEDFwouldcontinueit sexpansionmainlyintheMercosurareainLatinAmer icaandinsoutheastAsiaincludingChinaEDFobtain edastakeinBrazilsRiodeJaneiropowerutilityLig htandEdenorinArgentina
SMe7	2185*	ThefutureofFrenchelectronicsgroupThomsonSAco uldbedecidedlateronMondaywhenthegovernmentwi llsaywhetherAlcatelAlsthomortheLagardereGrou pecanbuythegroupSourcesclosetothegovernments aidthatthedecisionpreparedincloseconsultatio nwithPresidentJacquesChiraccouldbeannouncedl ateonMondayoratthelatestonWednesdayafteracab inetmeetingThetimingoftheannouncementcouldbe broughtforwardduetoaplannedjournalistsstrike onTuesdayovertaxincreasesAtstakeismorethanth eidentityoftheindustrialgroupthatcantakeanin debtedandlossmakingfirmoffthestateshandsThed ecisionalsohasconsequencesforofFrancesdefenc eindustryanditsconsumerelectronicssectorFren chnewspaperssuggestthatAlcatelAlsthomheadedb yformerarmsengineerSergeTchurukstandsabetter chanceofwinningthantheLagardereGroupeanditsf oreignalliesThomsonSAisseventysixpercentowne dbythestatewhiletwentypercentisheldbytelecom municationsoperatorFranceTelecomThomsonSAhas fullownershipofThomsonMultimediaoneoftheworl dsleadingtelevisionsetcompaniesaswellasafift yeightpercentstakeinThomsonCSFwhichisamongth etopfiveworlddefenceelectronicscompaniesThom sonSAmadeaninetyninetyfivelossofonepointthre esevenninebillionfrancsandhasdebtsofsometwen tyfivebillionfrancsofwhichfourteenbillionare atthemultimediaarmwhichmadealossmorethannone billionfrancsThomsonCSFreturnedtotheblackinn intynintyfivewithaprofitofonepointonebillion francsThomsonCSFhasatwentypercentstakeinsemi conductorgroupSGSThomsonMicroelectronicsNVBo

	thbiddershaveunveiledthemainlinesoftheirbusi nessplanforThomsonbutneitherhasdetailedthefi nancialelementsoftheirbidsFrenchnewspapersspe culatethestatemayhavetopumpbillionsoffrancs intoThomsonpriortothetransferLagardereGroupe ledbyitsfounderJeanLucLagardereisaconglomera tewithinterestsspanningfromthemediabusinesst oitsMatradefencebusinessLagardereplanstocont rolThomsonCSFwhileBritishAerospacePlcGECPlca ndperhapsDaimlerBenAerospaceGmbHcouldtakemin oritystakesThisdealinwhichLagarderewouldmerg eitsMatraDefenseEspacearmwithThomsonCSFwould beamajorrestructuringofEuropessprawlingdefen ceindustryLagarderehaslinedupSouthKoreasDaew ootorunThomsonMultimediaDaewooschairmansaidh esawsynergiesincomponentsandhehaspromisedtoc reatejobsinFranceAlcatelAlsthomunveileditsTh omsonplansonSeptembereighteen

SMe – Mesej rahsia tambahan yang mempunyai saiz yang hampir sama dengan saiz teks pelindung

Lampiran G

Perbandingan Persamaan Antara Teks Pelindung dan Teks Stego



Lampiran H

Perwakilan Mesej Rahsia dan Lokasi Jujukan Rawak Berdasarkan Teks Pelindung

Teks Pelindung	Eksperimen	Perwakilan Mesej Rahsia	Lokasi Rawak
CT180 (180 aksara)	E1	SM1 – 28 63 132 128 174 117 65 153 24 161 166 28 98 146 145 167 170 17 163 90 32 173 83 17 160 32 117 72 168 (X,Y,Z): (0,4,3)(0,8,12)(0,8,8)(0,11,9)(0,7,12)(0,4,5)(0,10,3)(0,1,9)(0,10,11)(0,11,1)(0,1,13)(0,6,8)(0,9,11)(0,9,10)(0,11,2)(0,11,5)(0,1,2)(0,10,13)(0,6,0)(0,2,2)(0,11,8)(0,5,8)(0,1,2)(0,10,10)(0,2,2)(0,7,12)(0,4,12)(0,11,3)	Kekunci : a = 2; t = 10; m = 173 Lokasi Rawak : 20 50 110 57 124 85 7 24 58 126 89 15 40 90 17 44 98 33 76 162 161 159 155 147 131 99 35 80 170
	E2	SM1- 28 151 54 64 122 117 143 66 162 163 154 72 137 82 50 172 132 103 83 90 120 29 102 57 87 120 117 2 152 (X,Y,Z): (0,10,1)(0,3,9)(0,4,4)(0,8,2)(0,7,12)(0,9,8)(0,4,6)(0,10,12)(0,10,13)(0,10,4)(0,4,12)(0,9,2)(0,5,7)(0,3,5)(0,11,7)(0,8,12)(0,6,13)(0,5,8)(0,6,0)(0,8,0)(0,1,14)(0,6,12)(0,3,12)(0,5,12)(0,8,0)(0,7,12)(0,0,2)(0,10,2)	Kekunci : a = 9; t = 10; m = 173 Lokasi Rawak : 55 159 57 4 46 78 20 17 163 93 155 21 26 71 130 142 77 11 109 126 106 99 36 161 75 166 120 52 132
	E3	SM2 - 40 141 84 129 69 21 52 125 89 177 63 71 114 38 107 173 153 97 61 179 145 180 74 53 112 73 176 163 107 84 172 117 42 4 135 121 146 60 97 25 72 (X,Y,Z): (0,9,6)(0,5,9)(0,8,9)(0,4,9)(0,1,6)(0,3,7)(0,8,5)(0,5,14)(0,11,12)(0,4,3)(0,4,11)(0,7,9)(0,2,8)(0,7,2)(0,11,8)(0,10,3)(0,6,7)(0,4,1)(0,11,14)(0,9,10)(0,12,0)(0,4,14)(0,3,8)(0,7,7)(0,4,13)(0,11,11)(0,10,13)(0,7,2)(0,5,9)(0,11,7)(0,7,12)(0,2,12)(0,0,4)(0,9,0)(0,8,1)(0,9,11)(0,4,0)(0,6,7)(0,1,10)(0,4,12)	Kekunci : a = 13; t = 7; m = 173 Lokasi Rawak : 72 78 156 132 166 89 126 88 113 92 165 76 130 140 97 57 56 43 47 99 83 48 112 79 169 128 114 105 161 24 146 2 33 90 139 84 61 108 27 12 163
	E4	SM3 – 67 104 84 28 114 175 88 144 140 173 162 59 4 75 152 131 36 98 107 29 108 137 63 36 151 180 114 76 169 112 52 63 17 77 96 29 163 112 124 130 86 64 60 5 111 125 132 15 142 65 112 91 147 127 156 142 63 100 166 105 154 59 142 38 57 113 24 14 (X,Y,Z):	Kekunci : a = 3; t = 11; m = 173 Lokasi Rawak : 26 89 105 153 124 37 122 31 104 150 115 10 41 134 67 39 128 49 158 139 82 84 90 108 162 151 118 19 68 42 137 76 66 36 119 22 77 69 45 146 103 147 106 156 133 64 30 101 141 88 102 144 97 129 52 167 166 163 154 127 46 149 112

		(0,6,14)(0,5,9)(0,1,13)(0,7,9)(0,11,10)(0,5,13)(0,9,9)(0,9,5)(0,11,8)(0,10,12)(0,3,14)(0,0,4)(0,5,0)(0,10,2)(0,8,11)(0,2,6)(0,6,8)(0,7,2)(0,1,14)(0,7,3)(0,9,2)(0,4,3)(0,2,6)(0,10,1)(0,12,0)(0,7,9)(0,5,1)(0,11,4)(0,7,7)(0,3,7)(0,4,3)(0,1,2)(0,5,2)(0,6,6)(0,1,14)(0,10,13)(0,7,7)(0,8,4)(0,8,10)(0,5,11)(0,4,4)(0,4,0)(0,0,5)(0,7,6)(0,8,5)(0,8,12)(0,1,0)(0,9,7)(0,4,5)(0,7,7)(0,6,1)(0,9,12)(0,8,7)(0,10,6)(0,9,7)(0,4,3)(0,6,10)(0,11,1)(0,7,0)(0,10,4)(0,3,14)(0,9,7)(0,2,8)(0,3,12)(0,7,8)(0,1,9)(0,0,14)	1 14 53 170 2
		SM3-67 177 140 129 144 175 38 114 142 8 24 139 116 179 173 167 100 137 107 152 108 98 167 36 24 8 144 85 47 166 78 162 17 77 38 29 59 111 4 70 86 64 164 25 154 79 158 26 84 72 111 37 158 18 94 142 162 100 166 105 154 115 140 72 17 113 63 54 (X,Y,Z) : (0,11,12)(0,9,5)(0,8,9)(0,9,9)(0,11,10)(0,2,8)(0,7,9)(0,9,7)(0,0,8)(0,1,9)(0,9,4)(0,7,11)(0,11,14)(0,11,8)(0,11,2)(0,6,10)(0,9,2)(0,7,2)(0,10,2)(0,7,3)(0,6,8)(0,11,2)(0,2,6)(0,1,9)(0,0,8)(0,9,9)(0,5,10)(0,3,2)(0,11,1)(0,5,3)(0,10,12)(0,1,2)(0,5,2)(0,2,8)(0,1,14)(0,3,14)(0,7,6)(0,0,4)(0,4,10)(0,5,11)(0,4,4)(0,10,14)(0,1,10)(0,10,4)(0,5,4)(0,10,8)(0,1,11)(0,5,9)(0,4,12)(0,7,6)(0,2,7)(0,10,8)(0,1,3)(0,6,4)(0,9,7)(0,10,12)(0,6,10)(0,11,1)(0,7,0)(0,10,4)(0,7,10)(0,9,5)(0,4,12)(0,1,2)(0,7,8)(0,4,3)(0,3,9)	Kekunci : a = 11; t = 12; m = 173 Lokasi Rawak : 67 57 120 121 132 80 27 136 124 165 97 41 117 88 115 66 46 172 1 23 92 159 31 7 89 126 14 166 108 162 64 24 103 107 151 116 77 167 119 110 11 133 91 148 83 60 153 138 146 61 164 86 93 170 152 127 25 114 55 98 52 65 35 51 54 87 104 118
	E5	SMe1- 116 102 27 85 114 175 28 121 140 126 131 74 124 125 29 30 36 137 131 135 108 98 167 100 176 126 121 60 47 166 78 63 17 77 113 135 177 111 4 70 155 128 159 49 154 105 64 102 84 72 154 33 89 56 92 27 151 36 166 47 111 55 95 88 103 148 97 14 90 93 37 121 95 80 33 36 100 70 33 117 113 144 100 143 167 147 120 128 12 124 114 163 86 158 148 169 154 163 142 113 137 97 100 30 48 144 28 125 154 68 113 64 117 44 114 42 39 (X,Y,Z): (0,6,12)(0,1,12)(0,5,10)(0,7,9)(0,11,10)(0,1,13)(0,8,1)(0,9,5)(0,8,6)(0,8,11)(0,4,14)(0,8,4)(0,8,5)(0,1,14)(0,2,0)(0,2,6)(0,9,2)(0,8,11)(0,9,0)(0,7,3)(0,6,8)(0,11,2)(0,6,10)(0,11,11)(0,8,6)(0,8,1)(0,4,0)(0,3,2)(0,11,1)(0,5,3)(0,4,3)(0,1,2)(0,5,2)(0,7,8)(0,9,0)(0,11,12)(0,7,6)(0,0,4)(0,4,10)(0,10,5)(0,8,8)(0,10,9)(0,3,4)(0,10,4)(0,7,0)(0,4,4)(0,6,12)(0,5,9)(0,4,12)(0,10,4)(0,2,3)(0,5,14)(0,3,11)(0,6,2)(0,1,12)(0,10,1)(0,2,6)(0,11,1)(0,3,2)(0,7,6)(0,3,10)(0,6,5)(0,5,13)(0,6,13)(0,9,13)(0,6,7)(0,0,14)(0,6,0)(0,6,3)(0,2,7)(0,8,1)(0,6,5)(0,5,5)(0,2,3)(0,2,6)(0,6,10)(0,4,10)(0,2,3)(0,7,12)(0,7,8)(0,9,9)(0,6,10)(0,9,8)(0,11,2)(0,9,12)(0,8,0)(0,8,8)(0,0,12)(0,8,4)(0,7,9)(0,10,13)(0,5,11)(0,10,8)(0,9,13)(0,11,4)(0,10,4)(0,10,13)(0,9,7)(0,7,8)(0,9,2)(0,6,7)(0,6,10)(0,2,0)(0,3,3)(0,9,9)(0,1,13)(0,8,5)(0,10,4)(0,4,8)(0,7,8)(0,4,4)(0,7,12)(0,2,14)(0,7,9)(0,2,12)(0,2,9)	Kekunci : a = 2; t = 12; m = 173 Lokasi Rawak : 22 56 124 87 13 38 88 15 42 96 31 74 160 159 157 153 145 129 97 33 78 168 2 16 44 100 39 90 19 50 112 63 138 115 69 150 139 117 73 158 155 149 137 113 65 142 123 85 9 30 72 156 151 141 121 81 1 14 40 92 23 58 128 95 29 70 152 143 125 89 17 46 104 47 106 51 114 67 146 131 101 41 94 27 66 144 127 93 25 62 136 111 61 134 107 53 118 75 162 163 165 169 4 20 52 116 71 154 147 133 105 49 110 59 130 99 37