CS361 Questions: Week 3

Name: **Zhenyu Zhu**

Date: **6/24/2014**

EID: **cike**

CS login: **zhenyu**

Email: **zhu_zhenyu@utexas.edu**

HW: **#3**

CS361 Questions: Week 3

Lecture 34

1. Why is it impossible to transmit a signal over a channel at an average rate greater than C/h?

   Because C is the capacity of the channel, and h is the most efficient possible encoding of the language, C/h is the maximum symbol per sec a channel can have given a perfect encoding and a noiseless channel. There is also a small fraction of error involve. If transmit a signal over this C/h rate, then either there could be another better encoding than h (contradict with definition of entropy) or a bigger capacities (contradict with the giving condition C).

2. How can increasing the redundancy of the coding scheme increase the reliability of transmitting a message over a noisy channel?

   Because when you increase the redundancy of the coding scheme, then you need to send more information over the channel. The more information you send, it is easier to get an approximate answer to reduce the uncertainty on the receiver side. You can also repeated send the same information within the redundancy to improve reliability over noisy channel. No matter how noisy the channel is, information always can be sent through.

Lecture 35

1. If we want to transmit a sequence of the digits 0-9. According to the zero order models, what is the entropy of the language?

   h = -($\log_2 1/10$) // zero order model assume all digits are equal likely to appear

2. What are reasons why computing the entropy of a natural language is difficult?

   Because a natural language is not “Zero-Memory”, every symbol has interrelated relations to each other. There are many factors and many different ways to calculate the probability appearance of each symbol, such as examples in the lecture 35. Also a natural language can have different interpretation over the same context, they all contain significant amount of redundancy and might have specific text involved that can’t be compute by the average entropy.

3. Explain the difference between zero, first, second and third order models.

   The zero-order model is based on the assumption that all characters are equal likely to appear. The first-order model is based on the assumption that all symbols are independent of each other, and calculated entropy is based on the probability achieved on such assumption. The second-order model is calculation entropy based on probability of digrams (two letter combinations). And the third-order model entropy is calculated base on trigrams (three letter combinations).

Lecture 36

1. Why are prior probabilities sometimes impossible to compute?

   Because there might be different versions of the uncertainty a receiver can have, depends on the question we ask. Just as the example in the lecture 36 slides 2, there are different information content of a messages need to be send base on different state of knowledge of the receiver.

2. Why is the information content of a message relative to the state of knowledge of an observer?

   Because we are trying to reduce the uncertainty within the receiver, without knowing exactly what the receiver knows about certain question, we can’t come up with the proper encoding and efficient entropy.

3. Explain the relationship between entropy and redundancy.

   The difference between the efficiency of the encoding and the entropy is a measure of redundancy in the encoding. Entropy can be used to measure the amount of “redundancy” in the encoding. If the information content of a message is equal to the length of the encoding message, then there is no redundancy. If you find an encoding with efficiency matching the entropy and there is no redundancy.

Lecture 37

1. List your observations along with their relevance to cryptography about Captain Kidd's encrypted message.

   There are 21 different symbols used in the plaintext, which 6 pairs of special symbol and 9 decimal digits (0-9 without 7) used in this message. The numbers might have digram or trigram relation or patterns among each other, same with the pair of special symbols such as "(), []…" Some digit appears more than the rest such as number 8. It could be a direct simple encryption key applied to the plaintext since only 21 symbol detected. There could be substitution and transposition applied to the message, assuming English is the underlying language and the message was in 1800.

2. Explain why a key may be optional for the processes of encryption or decryption.

   Because encryption and decryption are just function which transform one text into another, both algorithms can perform simple substitution on their own. Depends on how you define the key, as in Caesar cipher, it could be a keyless cipher (shift amount is fixed) or with certain keys (if we choose the shift amount as key).

3. What effect does encrypting a file have on its information content?

   To hide and preserve the information content of a file, so that a receiver can extract the information content from the message later.

4. How can redundancy in the source give clues to the decoding process?

   The redundancy in the source might give the attacker some hints or leverage to decode the message, depends on the underlying language and characteristic of the underlying language.

Lecture 38

1. Rewrite the following in its simplest form: D (E (D (E (P)))).

   = D (E (D (C))) = D (E (P)) = D (C) = P

2. Rewrite the following in its simplest form: D (E (E (P, K_E), K_E), K_D).

   = D (E (C, K_E), K_D) = C

3. Why might a cryptanalyst want to recognize patterns in encrypted messages?

   To get some clues or hints about the scenario, without knowing the content of the message and by just analysis the patterns in the message.

4. How might properties of language be of use to a cryptanalyst?

   Depends on the property, it can be used as a hint to figure the encrypted message. For example, the frequency of symbol in the English, we can use this try to substitute with the most appear symbol in the encrypted message with letter “e”.

Lecture 39

1. Explain why an encryption algorithm, while breakable, may not be feasible to break?

   Because even with modern computer’s clock speed, it is still going to take a very long time to calculate all the possibilities of encipherments. There is a feasible programmer to do that, but the time needed to find a solution is infeasible and unnecessary.

2. Why, given a small number of plaintext/ciphertext pairs encrypted under key K, can K be recovered by exhaustive search in an expected time on the order of 2^{n-1} operations?

   Yes, assuming the key K here a n-bit string, you will have 2^n possible keyspace, And to take a linear search on this space will take half of the 2^n operations, $2^n/2 = 2^{n-1}$.

3. Explain why substitution and transposition are both important in ciphers.

   To accomplish both confusion and diffusion during encryption steps.

4. Explain the difference between confusion and diffusion.

   The main difference is the encrypted text stay in the same place in confusion but it is moved around to a different place in diffusion during an encryption step.

5. Is confusion or diffusion better for encryption?

   Both are important to encryption, the combination of two is needed to provide a strong encryption algorithm.

Lecture 40

1. What is the difference between monoalphabetic and polyalphabetic substitution?

   Depends on where the symbol occurs in the plaintext and how the exchange is performed. If the exchanged of symbol is done uniformly on the plaintext, then it is a monoalphabetic cipher, if different substitutions are made depending on where in the plaintext the symbol occurs, it’s polyalphabetic substitution.

2. What is the key in a simple substitution cipher?

   The key is however you specify the mapping of the alphabet; it could be a table or another scheme that exhibit the mapping.

3. Why is there k! Mappings from plaintext to ciphertext alphabets in simple substitution?

   Because simple substitution is an injection (1-1 mapping) of the alphabet into itself or another alphabet, and there are total k! Possible permutations of an alphabet of k characters.

4. What is the key in the Caesar Cipher example?

   A table of 26 x 1, where the column is a “y” alphabet by adding “fixed distance” to the row alphabet “x”.

5. What is the size of the keyspace in the Caesar Cipher example?

   It depends on how may letter of string given, if just one character as in the example, the keyspace should be 26 alphabets.

6. Is the Caesar Cipher algorithm strong?

   Probably not, since both the plaintext and ciphertext are in English, it has regularity property (such as symbol frequency), which can be used by attacker to simplify the decryption process.

7. What is the corresponding decryption algorithm to the Vigenere ciphertext example?

   It is the reverse process of finding ciphertext. First look up the ciphertext character within the Vigenere Tableau, then look at matching key phrase character by either row or column (since table is symmetric), then the plaintext character should be the correspond column/row character.

Lecture 41

1. Why are there 17576 possible decryptions for the “xyy” encoding on slide 3?

   Because the first question only gives us a string of 3 characters, did not specify it is a simple substitution or polyalphabetic substitution. So the space of possible decryption is 26 x 26 x 26 for 3 characters, which equal 17576.

2. Why is the search space for question 2 on slide 3 reduced by a factor of 27?

   Because more information was given, (simple substitution tells us 1-1 mapping), and since we know the given “xyy” string only has two different characters, so we have 26 x 25, which equals 650. And divide this number by 17576, we have a factor of 27.

3. Do you think a perfect cipher is possible? Why or why not?

   Yes in theoretically, no in practical. Such as the one-time pad is a theoretically perfect cipher, but in reality it has key distribution problem.

Lecture 42

1. Explain why the one-time pad offers perfect encryption.

   Because under one-time pad scheme, even if you knew the encryption algorithm and the ciphertext, you as an analyst still cannot reduce the search space. As in example provided to us on slides 4, every possible plaintext could be the pre-image of the ciphertext under a plausible key.

2. Why is it important that the key in a one-time pad be random?

   Because if the key is not random, and if analyst know some property of the key in advance, he can work backwards with the given ciphertext to reduce the search space of the encrypted message. Therefore one-time pad will not be a perfect cipher by definition.

3. Explain the key distribution problem.

   Since both sender and receiver have to agree on this same length key they can use in the algorithm. If they can send the key securely over the covert channel, then they can just sending the plaintext securely without using the key. So basically the problem is how to send this key to the other end of the channel securely.

Lecture 43

1. What is a downside to using encryption by transposition?

   It is not a very strong cipher on its own. Since the transposition cipher only reorder characters, but doesn’t replace them. If the underlying language is English, then first-order entropy is preserved. Analysis on letter frequency on the characters in the ciphertext can help identify the characters used in the plaintext.

Lecture 44

1. Is a one-time pad a symmetric or asymmetric algorithm?

   Symmetric algorithm, since one-time pad use the same key for encryption and decryption.

2. Describe the difference between key distribution and key management.

   Key distribution is about (confidentiality) how to convey the key secretly and securely to those who need them to establish secure communication. Key management is about (integrity and availability) how to preserve the already known keys' safety and availability.

3. If some one gets a hold of Ks, can he or she decrypt S's encrypted messages? Why or why not?

   It depends on if this is a symmetric or asymmetric system. In a symmetric system, if they get a hold of Ks, then he or she can decrypt S's encrypted messages. In an asymmetric system, they cannot decrypt S's encrypted messages because two different sets of keys are used for encryption and decryption.

4. Are symmetric encryption systems or public key systems better?

   It depends on the use. Symmetric encryption system might be better for computers and public key system might be better for human agent in the field. Also depends on how many users involved in the secure system. If it is a pairwise secure system, a symmetric encryption system is better. If it is an n to 1 secure system, then public key system is better.

Lecture 45

1. Why do you suppose most modern symmetric encryption algorithms are block ciphers?

   Because stream ciphers are often malleable encryption algorithm, and most modern block-structured ciphers are non-malleable.

2. What is the significance of malleability?

   That being able to manipulate ciphertext with predictable effect on the plaintext, which is bad for encryption cipher.

3. What is the significance of homomorphic encryption?

   It is malleable by design with a specific (might be different) algebraic operation perform on plaintext and ciphertext.

Lecture 46

1. Which of the 4 steps in AES uses confusion and how is it done?

   SubBytes is the step in AES uses confusion. In subByte, the value of each byte is used as index into a 256 elements lookup table and replace byte by the value stored at the location of the table.

2. Which of the 4 steps in AES uses diffusion and how is it done?

   ShiftRow is the step in AES uses diffusion. It used the row number as the number of bytes that need to shift to left in that row. So row 0 shift left 0 byte, where row 1 shift left 1 byte, etc.

3. Why does decryption in AES take longer than encryption?

   Because in decryption step of AES, you have to multiplied the column of state with a much more complicated fixed integer 4x4 array, which take a lot longer time than the multiplication of very simple integer matrix at encryption steps.

4. Describe the use of blocks and rounds in AES.

   The block is the basic unit that gets operates on, and the rounds are loops (10,12,14) with 4 operations on the unit within each loop.

   The block in AES has a fix size of 128 bits. It is arranged as a 4x4 array of bytes, which called “state”. This “state” is modified in place in each round. Every round perform exact same 4 steps (operations) on “state” in AES. These steps are subBytes, shiftRows, mixColumns and addRoundKey. The number of rounds depend the length of keys selected for AES, which can varies between 10, 12 or 14.

5. Why would one want to increase the total number of Rounds in AES?

   Because when the key length increased, we need to increase the rounds to maintain confusion and diffusion goals. So we can maintain the complexity of the encryption, not making it weaker.

Lecture 47

1. What is a disadvantage in using ECB mode?

   Identical blocks in the plaintext yields identical blocks in the ciphertext, leaves too much regularity in the ciphertext

2. How can this flaw be fixed?

   One way we can do is to randomize blocks before they are encrypted, such as XOR with some previous ciphertext block or seed (IV) to “chain” the blocks in CBC.

3. What are potential weaknesses of CBC?

   Such as observed changes and content leak. Observed changes are for attack to spot the first block that changed in ciphertext, and content leak is by compare two identical ciphertext block, attacker can derive information about two plaintext blocks.

4. How is key stream generation different from standard block encryption modes?

   The output format is different. Block encryption mode output is still in block form and contained store information that can be recoverable from. Key stream generation mode acts like a pseudorandom number generator, which creates “appear random” stream of bits in reproducible fashion.

Lecture 48

1. For public key systems, what must be kept secret in order to ensure secrecy?

   The private key used for decryption must be kept secret.

2. Why are one-way functions critical to public key systems?

   Because the basis of any public key system is the identification of a one-way function: easily computed, but difficult to invert without additional information.

3. How do public key systems largely solve the key distribution problem?

   Because the public key can be distribute openly (no secret) without have to worry about convey this information over a secure channel, which is the core problem of key distribution.

4. Simplify the following according to RSA rules: $\{\{\{P\}_K{}^{-1}\}_K\}_K{}^{-1}$.

   $\{P\}_{K^{-1}}$

5. Compare the efficiency of asymmetric algorithms and symmetric algorithms.

   Asymmetric algorithms are much less efficient than symmetric algorithm, a public key encryption (asymmetric algorithm) may take 10,000 times as long to perform as a symmetric encryption.

Lecture 49

1. If one generated new RSA keys and switched the public and private keys, would the algorithm still work? Why or why not?

   Yes, assuming generate also publish the new public key after switching, since RSA is symmetric in the use of Keys.

2. Explain the role of prime numbers in RSA.

   Prime numbers are difficult to be factoring after some mathematic operations on them, and which is what RSA algorithm relies on. Prime number also can be used in any one-way function, which public key system is based on.

3. Is RSA breakable?

   Yes, by brute force or break the person who has the private key.

4. Why can no one intercepting $\{M\}_{Ka}$ read the message?

   Because only A has the private key to decrypt the message.

5. Why can't A be sure $\{M\}_{Ka}$ came from B?

   Because K_a is a public key of A and anyone can access this to encrypt the message.

6. Why is A sure $\{M\}_{Kb}-1$ originated with B?

   Because the message is encrypted with private key only B knows about.

7. How can someone intercepting $\{M\}_{Kb}-1$ read the message?

   Because since the message is encrypted with the private key of B, so anyone with the public key of B interception this message can decrypt the message, by definition of RSA.

8. How can B ensure authentication as well as confidentiality when sending a message to A?

   Can't in RSA or any public key system, use the nested encryption with symmetric encryption system and the public key system as in lecture 51 slide 5.

Lecture 50

1. Why is it necessary for a hash function to be easy to compute for any given data?

   So we can preserve the integrity of any given data using the hash value it computes.

2. What is the key difference between strong and weak collision resistance of a hash function.

   Strong collision resistance is trying to find any two messages that have the same hash value, where weak collision resistance has an assumption that the two messages are not equal to each other.

3. What is the difference between preimage resistance and second preimage resistance?

   Preimage resistance is given a hash value and function, trying to find any message that has that hash value, where second preimage resistance is given a function and a message, trying to find a second message (not identical with first message) that has the same hash value as the first message.

4. What are the implications of the birthday attack on a 128-bit hash value?

   That it will need to evaluate the function (1.25 * $\sqrt{2^{128}}$ = 1.25 * 2^{64}) different arguments on average before finding a collision.

5. What are the implications of the birthday attack on a 160-bit hash value?

   That it will need to evaluate the function (1.25 * $\sqrt{2^{160}}$ = 1.25 * 2^{80}) different arguments on average before finding a collision.

6. Why aren't cryptographic hash functions used for confidentiality?

   Because the hash function and its property is about protect the integrity and not the privacy of the data. You cannot use hash value to check if who can see the file, since there is an infinite number of collision for any hash value, it adds ambiguous and uncertainty when checking confidentiality.

7. What attribute of cryptographic hash functions ensures that message M is bound to H (M), and therefore tamper-resistant?

   Hash value or message digest.

8. Using RSA and a cryptographic hash function, how can B securely send a message to A and guarantee both confidentiality and integrity?

   Send the encrypted hash value of the message with the public key of A. $\{\{h(m)\}_{K_B^{-1}}\}_{K_A}$

Lecture 51

1. For key exchange, if S wants to send key K to R, can S send the following message: $\{\{K\}_{KS^{-}1}\}_{KR^{-}1}$? Why or why not?

   No, because any eavesdropper can use R's public key to decrypt the first outer layer, then use S's public key to decrypt the 2nd inner layer to get the suppose secret key K intended for R.

2. In the third attempt at key exchange on slide 5, could S have done the encryptions in the other order? Why or why not?

   Yes, you can send the K to R when encryption in other order, although any other eavesdropper other than R can all use S's public key to decrypt the outer layer, no one but R has the private key to decrypt the inner layer to retrieve K. So it's probably ok to send the message in this order.

3. Is $\{\{\{K\}_{KS}{}^{-1}\}_{KR}\}_{KS}$ equivalent to $\{\{K\}_{Ks}{}^{-1}\}_{KR}$?

   No, R can't get K in first encryption, but R can get K in second encryption.

4. What are the requirements of key exchange and why?

   It needs nested encryption of S and R's public/private key to achieve both confidentiality and authentication.

Lecture 52

1. What would happen if g, p and $g^a \bmod p$ were known by an eavesdropper listening in on a Diffie-Hellman exchange?

   Nothing will happen; eavesdropper can't compute the shared secret with given 3 values.

2. What would happen if a were discovered by an eavesdropper listening in on a Diffie-Hellman exchange?

   Then the channel is compromised, eavesdropper can use Bob's message and the values of a and p to find the secret key.

3. What would happen if b were discovered by an eavesdropper listening in on a Diffie-Hellman exchange?

   The channel is also compromised, he can use Alice's message with values of b and p to find out the secret key.