Family Name: .....

First Name: .....

Section: .....

Security Protocols and Applications (Part 1) — Solutions

Final Exam

June 25th, 2010

Duration: 3:00

This document consists of 4 pages.

Instructions

Electronic devices and documents are *not* allowed.

This exam contains 2 *independent* parts.

Answers for each part must be written on its separate sheet.

Answers can be either in French or English.

Questions of any kind will certainly *not* be answered. Potential errors in these sheets are part of the exam.

You have to put your full name on the first page and have all pages *stapled*.

1 On Plaintext-Dependent Decryption in Secure Channels

This exercise is inspired by the article "Plaintext Recovery Attacks Against SSH" by Albrecht, Watson, and Paterson published in IEEE Symposium on Security and Privacy 2009, IEEE Press, 2009.

We use some parameters w, B, a, b, c, d, d'. Considering a set of *words* $Z = \{0,1\}^w$ of 2^w elements, a finite sequence $x \in Z^*$ has a length (in words) denoted by $|x|$. (Namely, the bitlength is $w|x|$.) We assume a binary encoding function V_n mapping an integer $m \in \{0, 1, \ldots, B_n\}$ to an element $V_n(m)$ of Z^n so that it can be decoded unambiguously by a function V_n^{-1}. For instance, we may consider binary encoding with $B_n = 2^{wn} - 1$. We use a block cipher Enc with blocks of 2^d words in CBC mode (with secret initial vector K_3) and a message authentication code MAC of $2^{d'}$ words. We assume a secure communication channel which is considered as a continuous stream from A to B based on some secret keys K_1, K_2, and K_3. To send a new message x such that $|x| \leq 2^B$ from Alice to Bob, Alice waits until messages in the queue have been sent. Then, x is first transformed into a payload

$$y = V_a(b + |x| + |\mathsf{pad}_x|)\|V_b(|\mathsf{pad}_x|)\|x\|\mathsf{pad}_x$$

where pad_x denotes the padding for message x such that $|\mathsf{pad}_x| \geq c$ and $|y|$ is multiple of 2^d. The exact way that pad_x is constructed is unimportant. Then, it is transformed into

$$z = \mathsf{Enc}_{K_1}(y)\|\mathsf{MAC}_{K_2}(\mathsf{header}\|y)$$

where header contains some extra protocol information which is not important here. Practically, the stream is split into packets which are sent sequentially in an asynchronous channel. For applications, we will assume $aw - B = 14$.

1. In the case of AES and openSSH, what are the values of w, a, b, c, d, an B?

   Words are bytes, $w = 8$, $a = 4$, $b = 1$, $c = 4$, and $d = 4$. We deduce $B = aw - (aw - B) = 18$.

2. Recall how the CBC mode works.

   Assume that $|y_i| = 2^d$ for all i and $y = y_1\| \cdots \|y_n$, then $z_i = C_{K_1}(y_i \oplus z_{i-1})$ with z_0 set to some secret initial value K_3.

3. Assuming that Bob receives z', explain the algorithm to extract x' from z' such that $x' = x$ when $z' = z$. In this exercise, we assume that errors in extraction are immediately notified but that there are no differences between the types of error.

   The first block z_1 of z is decrypted to y_1 then y_1 is parsed to $m\|n\| \cdots$ where $|m| = a$ and $|n| = b$. Then, m and n are decoded to integers (we take the same notation for the decoded numbers). If $m > 2^B$ or 2^d does not divide m, or $n < c$, an error is returned. Otherwise, what follows is decrypted on the fly until m more words are received. This decrypts to x. Then, more words are received to reach a total of $a + n + 2^{d'}$. The stream is parsed into $z_1\| \cdots \|z_\ell\|z_{\ell+1}$ where $|z_i| = 2^d$ for $i \leq \ell$ and $|z_{\ell+1}| = 2^{d'}$. The string $z_1\| \cdots \|z_\ell$ is decrypted into y' and the MAC $z_{\ell+1}$ is verified. If $z_{\ell+1} \neq \mathsf{MAC}_{K_2}(\mathsf{header}\|y')$, an error is returned. Otherwise, y' is parsed into $m\|n\|x'\|\mathsf{pad}$ such that $|\mathsf{pad}| = n$ and x' is returned.

4. If an adversary sends a random block as a leading packet of z', what is the probability p that no error is returned?

> *It is the probability that $m \leq 2^B$ and 2^d divides m and $n \geq c$ for random a-word and b-word m and n, which is $p = 2^{B-d-aw}(1 - c2^{-wb})$. In our case, this is $p \approx 2^{-18}$.*

5. Show how an adversary can decrypt $aw - B$ bits of information of a payload block y_i form z with probability p^{-1}.

> *Assume that we have cut the stream such that $z_{\ell-1}$ is the end of some MAC value so that z_ℓ is the first ciphertext block of a new payload. Send $z'_\ell = z_i$ instead of z_ℓ. Bob will decrypt this first block to $z_{i-1} \oplus y_i \oplus z_{\ell-1}$. With probability p^{-1}, the block is accepted and it means that the first a words of this decodes to an integer lower than 2^B. Using z_{i-1} and $z_{\ell-1}$, this gives $aw - B$ bits of y_i.*

6. To thwart the previous attack, could we have $|x|$ put at the end instead? Why?

> *No, because we have no clue when the payload ends so we need a way to get the length.*

7. Could we have $|x|$ sent in clear instead? Why?

> *We could have the length of the payload sent in clear but we want to hide the length of the message x. If we send the length in clear, it shall be added in the header for authentication.*

8. Could we have $z = \mathsf{Enc}_{K_1}(y \| \mathsf{MAC}_{K_2}(\mathsf{header} \| y))$ instead? Why?

> *Yes, but it does not change anything regarding the attack.*

9. Could we have $\mathsf{MAC}_{K_2}(\mathsf{header} \| y)$ checked before the length instead? Why?

> *Yes but we still need to know when the payload ends to check the MAC so it does not change anything regarding the attack.*

10. Could we have $V_a(b + |x| + |\mathsf{pad}_x|)$ authenticated in a separate way instead? Why?

> *Yes. It defeats the attack. However, it adds $2^{d'}$ more words in the payload.*

11. What would you propose as a countermeasure?

> *The adopted solution consists of switching to CTR mode with initial counter set to a secret K_3 instead of CBC. Another countermeasure consists of waiting until 2^B words before issuing the error, even when the error comes from the MAC verification. Maybe an algorithm-independent solution would consists of authenticating the lengths in a separate way as suggested in the previous question.*

Any attempt to look at
the content of these pages
before the signal
will be severely punished.

Please be patient.