CS 6431: Security and Privacy Technologies
Homework #5

Due: May 8, 2018

Instructors: Tom Ristenpart and Vitaly Shmatikov

Problem 1

Recall that padding oracle attacks arise when decryption leaks information about unauthenticated values.

1. Explain the role of the number 13 in the Lucky 13 paper.

2. Bellare and Namprempre in a paper in 2000 argue that one should build symmetric encryption via the encrypt-then-MAC construction: one first encrypts using a mode of operation such as CBC, and then applies a MAC to the resulting ciphertext using, e.g., HMAC. So a ciphertext is IV$\|C\|T$ where IV is the random initializaiton vector of CBC mode, C is the remaining portion of the CBC ciphertext, and T is the HMAC output. Here $\|$ denotes concatenation of strings.

   To build a secure channel out of this encrypt-then-MAC using CBC and HMAC, suppose we encode a plaintext message M of length ℓ bytes by $\langle \ell \rangle_4 \| M \| \langle P \rangle_1^P$ where:

   - $\langle x \rangle_m$ means encode in some canonical way the numerical value x as an m byte string.
   - $\|$ denotes concatenation of byte strings
   - P is the number of bytes needed to ensure that the encoded string is a multiple of the block-size n (in bytes) of the block-cipher underlying CBC. This is the same padding mechanism as used in TLS: so for example, if $\ell = 10$ and $n = 16$ bytes, then one appends the padding string "01 01".

A server receiving a streaming ciphertext sent over TCP proceeds as follows. Decrypt the first $2n$ bytes which here will be the IV of CBC mode and the first ciphertext block. Then, one decodes the first 4 bytes of plaintext to get the length ℓ of the plaintext, determines the number of further bytes of ciphertext blocks and tag bytes, and then proceeds by checking the tag (recomputing HMAC over IV$\|C$ and comparing with the received tag) and then ultimately decrypting the rest of the CBC ciphertext and releasing it to the application. If the MAC tag fails, the server sends an error message to the client.

Describe an attack against this scheme that recovers some amount of plaintext from a target ciphertext without violating the security of the block cipher or MAC underlying the mechanism.

Discuss the implications of your attack — why do Bellare and Namprempre's paper about the strong authenticated-encryption security of Encrypt-then-MAC not apply to the scheme?

Problem 2: Recall that in the ZMap paper a key feature and justification for its speed is the fact that it can send a single packet per destination IP address, port pair. In this question we focus on measuring efficacy via coverage, i.e., the fraction of hosts that respond to probes that ZMap indicates as indeed responding to probes (on the target port).

(a) Explain potential sources of coverage error in a horizontal scanner.

(b) Explain the mechanisms ZMap uses that mitigate such errors.

(c) Explain the ZMap author's methodology for measuring coverage.

(d) Discuss other approaches to coverage estimation.

(d) Describe how one could render hosts "ZMap-blind", i.e., responding to probes but not ZMap probes specifically.

Problem 3: List all assumptions underlying the analysis of the Witty worm in "Exploiting the Underlying Structure…" paper. How can a clever adversary exploit these assumptions to avoid detection and/or analysis?

Problem 4:

- Explain the similarities and differences in how hidden Markov models are used in the "Keyboard Acoustic Emanations Revisited" and "Phonotactic Reconstruction. . . ..." papers.
- Why is separating the stream into words more difficult in the case of phonotactic reconstruction?
- Does the use of METEOR scores to evaluate the quality of phonotactic reconstruction over- or under-estimates its success as an attack on privacy? Explain.