Problem Set 5

This problem set is due *on paper,* in room 6-120 on *Thursday, November 6* at the beginning of class.

You are to work on this problem set in groups of three or four people. Problems turned in by individuals, pairs, pentuples, etc. will not be accepted. Be sure that all group members can explain the solutions. See Handout 1 (*Course Information*) for our policy on collaboration. If you do not have a group, seek partners by emailing `6.857-students-public@mit.edu`.

Homework must be typed! Each problem answer must appear on separate sheets of paper. Mark the top of each sheet with your names (alphabetically by last name), the course number (6.857), the problem set number and question, and the date. **Homework must be typed and clear.** We have provided templates for LaTeX and Microsoft Word on the course website.

Grading and Late Policy: Each problem is worth 10 points, except where noted. Late homework will not be accepted without prior approval. Homework should not be submitted by email except with prior approval. (*Somebody* from your group should be in class on the day that the homework is due.)

With the authors' permission, we will distribute our favorite solution to each problem as the "official" solution – this is your chance to become famous! If you do not wish for your homework to be used as an official solution, or if you wish that it only be used anonymously, please note this on your homework.

Problem 5-1. Watermarking

(a) Evaluate the following watermarking schemes. Consider the approximate density of marks (i.e., how many kilo/megabytes of data are required per mark, and how many marks one could expect to embed in the item in question), the efficiency of the marking and extraction procedures, whether the marks are perceptible (when comparing the marked version to the original), and whether the marks are resistant to detection and/or removal (ignoring the possibility of collusion). Limit each of your answers to 1/4 page.

- In a plain text document containing one of Shakespeare's plays, at the end of each line either a blank space is added (to mark a 0), or no space is added (to mark a 1).
- In a high-resolution, high-color photograph stored in a non-lossy format, for each pixel, either its blue color component is incremented (to mark a 0), or its red color component is incremented (to mark a 1).
- In a full-length motion picture, at every transition from one continuous shot to another, either the last frame of the ending shot is dropped (to mark a 0), or the first frame of the beginning shot is dropped (to mark a 1).
- In the source code for a 6.170 final project (written in Java), each loop is either written as a `for` loop (to mark a 0), or as a `while` loop (to mark a 1).
- In the executable file for Powerpoint, every `ADD X,Y` instruction is either converted to `SUB X,-Y` (to mark a 0), or left alone.

(b) Watermarks aren't always used to deter copying. Consider the following scenario: the MediaWare company (which is in both the hardware and media business) plans to freely give away a device that plays music or videos in some standard format. However, MediaWare only wants the device to play media that is sold by MediaWare. Describe how, using watermarking and some other cryptographic primitive we have discussed in this class, MediaWare can design its playback device and mark its media to achieve these goals.

Problem 5-2. Zero Knowledge (or is it?)

Polly Prover has published a 1024-bit RSA public key (n, e), and wants to prove to Vinnie Verifier that she knows the corresponding private key d. Their friend Zack Kauffman proposes the following proof system:

1. Vinnie chooses a random $r \pmod n$ and sends its encryption $r^e \pmod n$ to Polly.
2. Polly decrypts the message she receives, and sends the decrypted value to Vinnie.
3. Vinnie accepts if Polly's message equals r, otherwise he rejects.

(a) Argue that Zack's protocol is *complete* and *sound*. Do you think Vinnie needs to repeat the protocol several times in order to be convinced that Polly knows d? Explain why or why not.

(b) Argue that, if Vinnie is honest (i.e., he plays exactly according his prescribed instructions), then the protocol is *zero-knowledge*. Your argument should involve a simulator that produces a transcript of the protocol.

(c) Argue that, if Vinnie is *dishonest*, then the protocol is *not* zero-knowledge. In particular, you should describe how an adversarial Vinnie can learn something that he couldn't discover on his own.

Problem 5-3. More Watermarking!

The PIAA (Photography Industry Association of America) is fed up with people copying their valuable photos, and have decided to watermark their pictures using a collusion-secure fingerprinting code. Exactly *how* they are embedding marks into the photos is proprietary, however, you do know that they are using the basic Boneh-Shaw fingerprinting scheme to determine *what* marks to embed. (The scheme in question was presented in lecture on 10/28, and is described starting on page 7 in the paper "Collusion-Secure Fingerprinting for Digital Data," available on the class website.) You also know that they have set the parameters to $c = 5$ and $\epsilon = 2^{-20}$ (to avoid implicating innocent users).

However, because the images to be marked are so small, the PIAA has had to compromise and set $d = 100$. (Recall that the correct value of d is about $c^2 \log \frac{1}{\epsilon}$, which is almost 10 times as large as the value in use.)

A highly sought-after photograph is put up for sale by the PIAA, and a coalition of 5 users quickly purchases individual (marked) copies, which are available on the class website. (The files are in Extended PostScript format, which can be edited as plain-text and rendered in any PostScript viewer.) The coalition asks you to analyze the files and produce a new photograph that cannot be traced back to any of the coalition members.

The five images can be downloaded from the class website. They are encoded in EPS format, which is text-based and can be rendered by any PostScript-compatible graphics tool.

Here is the beginning of one of the EPS files:

```
%!PS-Adobe-3.0 EPSF-3.0
%%TemplateBox:0 0 405 628
%%BoundingBox: 0 0 506 785
%%PageOrigin:0 0
/width 405 def
/height 628 def
/pixwidth 506 def
/pixheight 785 def
/picstr width string def
/dopic {
gsave
width height 8
[width 0 0 height neg 0 height]
{currentfile picstr readhexstring pop}
image
grestore
} def
pixwidth pixheight scale
dopic
FFFFFFFFFFFFFFFFFFFFFFFF...
```

For this problem, you can ignore the header — all images that we provide and that you produce should have the same header. The string `FFFF...` is the actual black-and-white image itself, coded as 1-byte hexdecimal values. The first `FF` characters represent a white pixel, the second `FF` characters represent another white pixel to the right of the first pixel, and so on. A value of `00` represents a black pixel, and other values represent shades of gray on a continuous scale.

The end of the EPS file looks like this:

```
...FFFFFFFFFFFF
%%Trailer
showpage
```

The `showpage` command causes the PostScript printer to eject the current page. Thus, if you simply send one of these EPS files to a PostScript printer (or view it with GhostScript), the image will print.

(a) Analyze the different copies, and determine which user (among users $0, \ldots, c$) is *not* a member of the coalition. In fact, there is not enough information to determine the user definitively; you can only narrow it down to two possible users. Explain why, and identify which two are the possible innocent users.

(b) Construct a new version of the photograph that has essentially the same contents as the others, but which cannot be traced back to any of the colluders. You must obey the Marking Assumption in creating your new version; that is, if all the files have the same value at some position, then your new version must also have that value at that position.

Go to the class website and upload your group's image; only one person from your group needs to do this. (The website will ask for a PIN; use the PIN that you were given for Problem Set 4.) The website will run the tracing algorithm (as given in the Boneh-Shaw paper) and record the results, which will also be displayed to you.

Be sure to upload at least one image! You will be graded based on the last image uploaded. Your group will receive full credit if the algorithm does not implicate any of the colluders (i.e., if it implicates only the innocent user, or no users at all). Otherwise, points will be deducted based on how many of the colluding users are implicated.

(c) For 2 points of extra credit, identify and research the person in the PIAA photograph, and write a short, interesting anecdote about him.