Information Security:

Theory vs. Reality

Exercise 3

due date: 12 February 2012

Tel Aviv University

0368-4474-01, Winter 2011-2012

Lecturer: Eran Tromer

Teaching Assistant: Roei Schuster (royschuster@mail.tau.ac.il)

Question 1

1. In class we have shortly discussed the following "traitor tracing" scheme for movies:
    - Purpose: Sell both Marge and Homer a movie (letting each one of them hold a copy of it), then be able to know in retrospect which one of them distributed it (e.g. on a P2P filesharing software)Individual movie frames can have two versions, almost indistinguishable to humans but very easily distinguished by automatic inspection.
    - Marge gets the movie with one version of one of the frames, Homer gets the other version of the frame.

   Bart&Nobles, a large media distribution corporation, wishes to sell many movies to a many people and "trace traitors". Propose a similar scheme to the one above, which fulfills Bart&Nobles' need: Every client has a unique ID and can buy any movie. Examination of an illegally distributed copy will reveal the ID of the owner to whom it was sold.
2. Bart&Nobles doesn't want to burn a special DVD for every client. Propose a method to encrypt the information in the disc, so that the decrypted information will be watermarked with the decrypter's identity, as in the scheme you've proposed in sec. 1. Assume clients can receive if necessary, upon registration to Bart&Noble, some additional data.
3. If Flanders can register more than one Bart&Nobles account, and receive multiple IDs, how can he bypass this mechanism and distribute the full movie (without missing frames), without it being traced back to him?
4. In class we have learned a multicast unidirectional encryption protocol which enables secret-key revocation:
   For N receivers, a full binary tree containing at least N leaves is constructed. Every node in the tree is assigned a random "node key". Every receiver is assigned a (unique) leaf in the full binary tree, and holds the node keys corresponding to the nodes in the path from the root to this leaf.
   Multicast messages are encrypted using a random "media key", under a symmetric cipher. The media key is then encrypted under each node key corresponding to a node in the set of nodes EFFECTIVE, and this encrypted information is broadcast along with the encrypted message.
   Initialization: EFFECTIVE = {root }, REVOKED = {}.
   When revoking access to multicast messages from a receiver (corresponding to a leaf), all the nodes in the path from the leaf to the root are added to REVOKED. The subgraph induced by the nodes outside REVOKED is a graph with several connected components which are full binary trees. EFFECTIVE becomes the set of all roots of these trees.
   In our case, every Bart&Nobles buyer is a receiver and every DVD movie is a message. Whenever somebody is identified as having illegally distributed a copy (using the scheme proposed in sec 1), he is revoked and cannot decrypt any more new CDs. [1]
   Joe the biker watches movies with very high volume and with a huge open window. This disturbs the peace in a quiet suburban neighborhood. One day all the neighbors gathered and decided to "finish the biker". What can Joe's neighbors, equipped with a high resolution camera, do to stop Joe from watching any new DVDs?
5. Flanders believes in freedom of information and hates copyrights. Once he has cheated the mechanism (using your answer in sec. 3), he wants to destroy it completely. Recall that every time a new traitor is discovered, EFFECTIVE is updated, and every new DVD will have its media key

[1] This protocol is layered on top of the "traitor tracing" protocol, but both are unrelated and can exist individually. Any distributed (broadcast) data can be distributed under any one of these protocols, none of them or both of them.

encrypted on-disk under every key corresponding to a node in EFFECTIVE. How can Flanders exploit this to make it harder to create new CDs under the protocol?

Question 2

MrTwig Corporation develops hardware and software. One specific line of hardware devices is designed to run only proprietary MrTwig code – code digitally signed by MrTwig's private key. When the device is powered on, the device's CPU reads the code and accompanying digital signature from a flash memory chip, and verifies the signature using MrTwig's public key (which is stored inside the CPU in non-volatile memory). If verification passes, the CPU begins execution of the code; otherwise it displays an error message and hangs (by entering an infinite loop).
Describe 5 potential ways to attack MrTwig's device, making it run malicious code. For each, specify:

1. The attack approach
2. The resources required from the attacker to break a single device
3. The additional resources required from the attacker for every additional device to be broken
4. A potential countermeasure MrTwig may apply

Question 3

Attached to the exercise is an executable file[2] (for 32 bit Linux). A secret key of 256 bits is hidden inside the file. The key was generated by a pseudorandom number generator. Write code which, exploiting the high entropy of key bytes, locates the secret key and outputs its index (or a list of up to 5 candidates). Assume the key is byte-aligned. Your code does not have to consider anything about the file's structure or the program that it contains. Your code can be in any programming language[3]. Submit the code, and its output (the index list).
Hint: Digraphs are pairs of consecutive characters (bytes). Digraph statistics (e.g., their histogram) turns out to be an excellent way to characterize many natural data sequences, such as human languages and machine code. For most natural sources, digraph histograms show very low entropy, i.e., of the 65536 possible byte pairs, a few occur much more frequently than the rest.

[2] A compiled version of an arbitrarily -selected open-source project - https://github.com/evandar/t2x . But you don't have to know that in order to solve the exercise.
[3] But again, Python's warmly recommended.