

An Attack on Black-box Traitor Tracing Schemes

Jeff Jianxin Yan¹ and Yongdong Wu²

¹ Computer Laboratory, University of Cambridge
Email: Jeff.Yan@cl.cam.ac.uk

² Kent Ridge Digital Labs (KRDL), Singapore

A traitor tracing scheme traces the source of keys used in pirate decoders for sensitive or proprietary data, such as pay-TV programs. If a pirate decoder is tamper-resistant, a tracer might be unable to efficiently extract a key used by the box. Without the key, it is impossible for the tracer to identify any traitor. So a very practical and important property of a traitor tracing scheme is to support black-box tracing, where a decoder is treated as a black box, and its embedded key can be deduced by testing how it decrypts chosen ciphertexts.

Boneh and Franklin [Crypto'99] proposed a deterministic traitor tracing scheme that claimed to catch all traitors while not accusing any innocent users as long as the number of traitors is at or below a threshold k . Their scheme supports black-box tracing. In this talk, firstly, a novel pirate decoder P_3 will be presented to defeat the BF scheme. In our design (Fig. 1.), any input ciphertext will be multiplexed to three decoding logic that will solve S_1, S_2 and S_3 independently just like stand-alone decoders. These values are then input to a comparator, which will output the equaled value when any two of them are equal (e.g. S_1 when $S_1 = S_2 \neq S_3$), but output random bits when $S_1 \neq S_2 \neq S_3$. P_3 works well as a legitimate decoder in normal operation, while neither the single-key nor arbitrary pirate tracing algorithm presented by Boneh and Franklin can identify all keys used by P_3 , and actually both algorithms catch none of the traitors with a high probability. The reason that the Boneh-Franklin scheme fails is that the designers failed to consider that it is possible for a pirate box to fool the tracer by detecting and then maliciously responding to tracing inquiries.

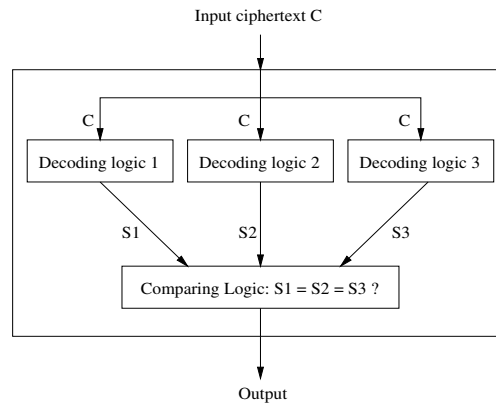


Figure 1. The structure of a pirate decoder P_3

The construction of P_3 shows a simple way for a pirate decoder to automatically distinguish some tracing traffic from normal ciphertext traffic. Traffic analysis based on a comparator as in Fig. 1 enables a pirate to disturb and defeat all other published black-box tracing methods such as Chor et al (Crypto'94) and Pfitzmann (IH'96). In this sense, traitor tracing researchers appear to have used too idealistic threat models, which unavoidably lead to failures when an “intelligent” pirate box is used by a hacker.