

# Randomized Frameproof Codes for Content Protection

N. Prasanth Anthapadmanabhan  
(joint work with Alexander Barg)

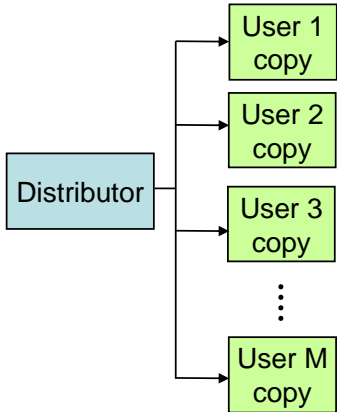
University of Maryland, College Park

First Annual School of Information Theory, 2008

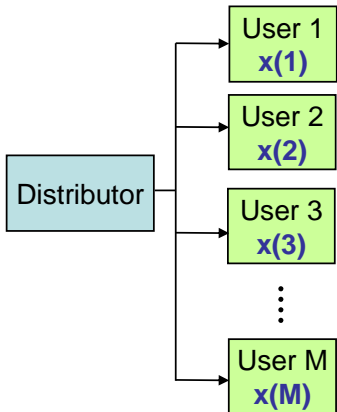
## Objective

To design a scheme to protect copyrighted content (esp. **software**) against piracy.

# Problem statement

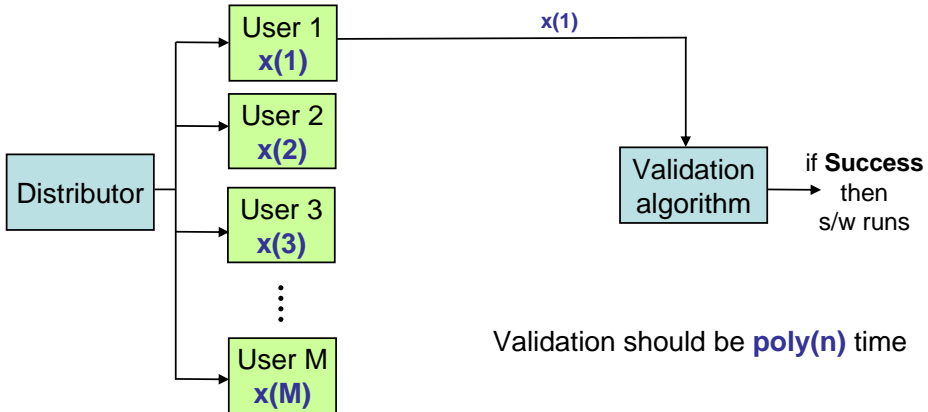


# Problem statement

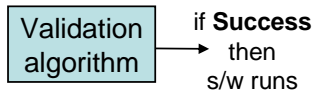
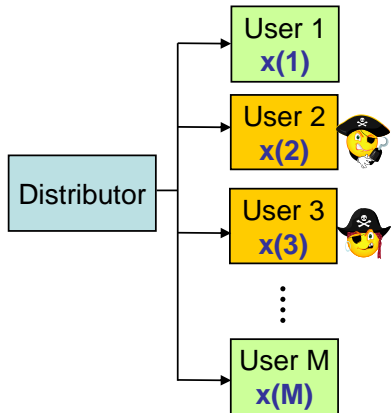


Fingerprints  $x(1), \dots, x(M)$  are **binary** vectors of length  $n$

# Problem statement

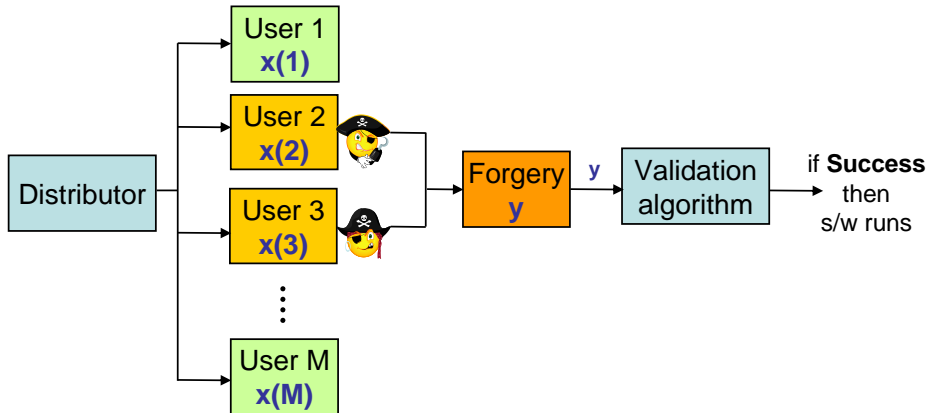


# Problem statement

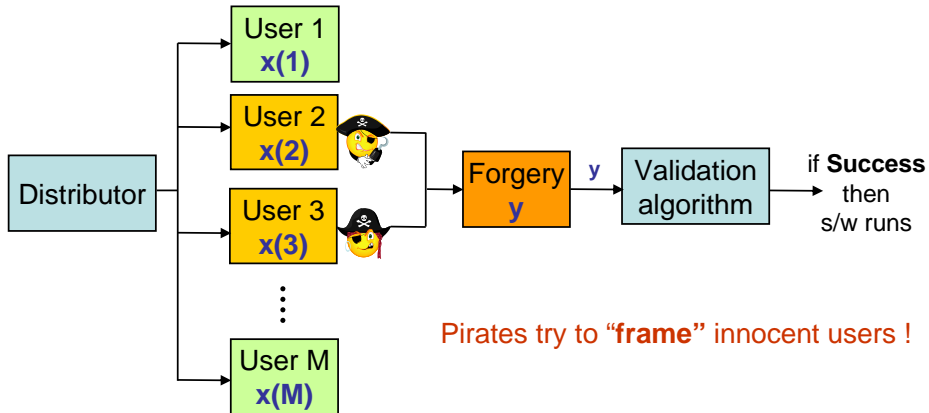


Suppose users 2 and 3 are “pirates”

# Problem statement



# Problem statement



# Problem statement

What are the rules for the pirates?

# Problem statement

What are the rules for the pirates?

Marking assumption [Boneh-Shaw '98]

Pirates **detect** fingerprint positions by finding differences in their copies.  
They make **changes only in the detectable positions**.

Example:

$x_1$	:	110001000
$x_2$	:	100011010
$y$	:	110011000

# Problem statement

What are the rules for the pirates?

Marking assumption [Boneh-Shaw '98]

Pirates **detect** fingerprint positions by finding differences in their copies.  
They make **changes only in the detectable positions**.

Example:

$x_1$	:	110001000
$x_2$	:	100011010
$y$	:	110011000

**Envelope:**  $\mathcal{E}(x_1, \dots, x_t)$

Set of all possible forgeries for  $t$  pirates with fingerprints  $x_1, \dots, x_t$ .

# Problem statement

What can the distributor do?

# Problem statement

What can the distributor do?

- Distributor uses a **randomized code**  $\mathcal{C}$ , i.e., picks a code at random from a family of codes  $\{C_k\}$ ,  $|C_k| = M$ .

Rate  $R = \log_2 M/n$ .

# Problem statement

What can the distributor do?

- Distributor uses a **randomized code**  $\mathcal{C}$ , i.e., picks a code at random from a family of codes  $\{C_k\}$ ,  $|C_k| = M$ .

Rate  $R = \log_2 M/n$ .

- **Validation algo:**  
Checks whether fingerprint is present in current codebook.  
Preferably polynomial-time complexity.

# Problem statement

What can the distributor do?

- Distributor uses a **randomized code**  $\mathcal{C}$ , i.e., picks a code at random from a family of codes  $\{C_k\}$ ,  $|C_k| = M$ .

Rate  $R = \log_2 M/n$ .

- **Validation algo:**  
Checks whether fingerprint is present in current codebook.  
Preferably polynomial-time complexity.
- Code family is public. **Only selection of  $k$  is secret!**

# Frameproof codes

When can a set of pirates  $U$  frame an innocent user?

$$\underbrace{\mathcal{E}(C_k(U))}_{\text{Envelope of pirates}} \cap \underbrace{(C_k \setminus C_k(U))}_{\text{Fingerprints of innocent users}} \neq \emptyset.$$

# Frameproof codes

When can a set of pirates  $U$  frame an innocent user?

$$\underbrace{\mathcal{E}(C_k(U))}_{\text{Envelope of pirates}} \cap \underbrace{(C_k \setminus C_k(U))}_{\text{Fingerprints of innocent users}} \neq \emptyset.$$

## Definition

$\mathcal{C}$  is  *$t$ -frameproof with  $\varepsilon$ -error* if for any set  $U$  of at most  $t$  pirates

$$\underbrace{\Pr\{\mathcal{E}(C(U)) \cap (C \setminus C(U)) \neq \emptyset\}}_{\text{Prob. of framing}} \leq \varepsilon.$$

# Frameproof codes

When can a set of pirates  $U$  frame an innocent user?

$$\underbrace{\mathcal{E}(C_k(U))}_{\text{Envelope of pirates}} \cap \underbrace{(C_k \setminus C_k(U))}_{\text{Fingerprints of innocent users}} \neq \emptyset.$$

## Definition

$C$  is  *$t$ -frameproof with  $\varepsilon$ -error* if for any set  $U$  of at most  $t$  pirates

$$\underbrace{\Pr\{\mathcal{E}(C(U)) \cap (C \setminus C(U)) \neq \emptyset\}}_{\text{Prob. of framing}} \leq \varepsilon.$$

**Objective:** Construct frameproof codes with  $\text{poly}(n)$  complexity validation.

# Frameproof codes with poly-time validation: Key ideas

- **Linear** codes are good: can be validated in  $O(n^2)$  !!  
But...

# Frameproof codes with poly-time validation: Key ideas

- **Linear** codes are good: can be validated in  $O(n^2)$  !!  
But... **works only for  $t = 2$  pirates.**  
(In fact, can achieve  $R \approx 1/2$  for linear 2-frameproof codes).

# Frameproof codes with poly-time validation: Key ideas

- **Linear** codes are good: can be validated in  $O(n^2)$  !!  
But... **works only for  $t = 2$  pirates.**  
(In fact, can achieve  $R \approx 1/2$  for linear 2-frameproof codes).
- Frameproof codes with  $\text{poly}(n)$  validation for  $t > 2$ ?

# Frameproof codes with poly-time validation: Key ideas

- **Linear** codes are good: can be validated in  $O(n^2)$  !!  
But... **works only for  $t = 2$  pirates.**  
(In fact, can achieve  $R \approx 1/2$  for linear 2-frameproof codes).
- Frameproof codes with  $\text{poly}(n)$  validation for  $t > 2$ ?  
Use **code concatenation**.

# Polynomial-time validation for larger $t$

- Outer code: Reed-Solomon (**linear**) code with rate  $\approx 1/t$ .
- Inner code:  $t$ -**frameproof** with unrestricted complexity
- **Validation**: Exhaustive look-up at inner level. Parity-checks at outer level.

## Theorem

The concatenated code is  $t$ -**frameproof** with error prob.  $\exp(-\Omega(n))$ , validation complexity  $O(n^2)$  and rate  $\approx R_t/t$ .

where

$$R_t = \max_{p \in [0,1]} [-p^t \log_2 p - (1-p)^t \log_2(1-p)].$$

## Interesting problems

- Better rates?
- Upper bounds?

## If you're interested...

- N. P. Anthapadmanabhan, A. Barg, “Randomized frameproof codes: Fingerprinting plus validation minus tracing”, CISS 2008 (arXiv:0802.3419)
- D. Boneh, J. Shaw, “Collusion-secure fingerprinting for digital data”, IEEE Trans. Inform. Theory, vol. 44, no. 5, pp. 1897-1905, Sep. 1998.

Email: [nagarajp@umd.edu](mailto:nagarajp@umd.edu)