Complexity and Cryptography

Thomas Zeugmann

Hokkaido University Laboratory for Algorithmics

https://www-alg.ist.hokudai.ac.jp/~thomas/COCRB/

Lecture 15: More Cryptographic Protocols



The coin flipping problem is the problem which initiated the whole area (see M. Blum (1981)).

The coin flipping problem is the problem which initiated the whole area (see M. Blum (1981)).

Blum described the scenario as follows. Suppose Alice and Bob and are going to get a divorce. They already live in cities far apart of each other and they don't want to see each other again. For deciding who will obtain the new car, they have agreed to flip a coin. Of course, they don't like to make their choice, say choosing head, and then hearing from the other end of the phone: "I am flipping the coin,..., the outcome is tail. I am so sorry for you."

The coin flipping problem is the problem which initiated the whole area (see M. Blum (1981)).

Blum described the scenario as follows. Suppose Alice and Bob and are going to get a divorce. They already live in cities far apart of each other and they don't want to see each other again. For deciding who will obtain the new car, they have agreed to flip a coin. Of course, they don't like to make their choice, say choosing head, and then hearing from the other end of the phone: "I am flipping the coin,..., the outcome is tail. I am so sorry for you."

So, we do not only have to realize the coin flip but also a method for verifying its outcome by the other party.

Flipping a Coin

Threshold Schemes

End

Flipping a Coin per Telephone II

Question

How can we attack this problem?

End

Flipping a Coin per Telephone II

Question

How can we attack this problem?

Let us have a look at the proposals made.

Question

How can we attack this problem?

Let us have a look at the proposals made.

The first proposal was made by Blum and Micali. Let X be a finite set of numbers containing as much even numbers as odd ones, and let $f: X \rightarrow Y$ be a one-way function.

Furthermore, assume Alice and Bob have agreed to use f. Then, the following protocol is used.

Proposal 1 (Blum/Micali)

- Step 1: Alice chooses at random an element $x \in X$, computes y = f(x) and sends y to Bob.
- Step 2: Bob guesses whether or not x is even or odd and sends his guess to Alice.
- Step 3: Alice tells Bob whether or not his guess was right and proves her claim by sending x to Bob, too.
- Step 4: Bob verifies Alice's claim by computing f(x) and comparing it to y.

Proposal 1 (Blum/Micali)

- Step 1: Alice chooses at random an element $x \in X$, computes y = f(x) and sends y to Bob.
- Step 2: Bob guesses whether or not x is even or odd and sends his guess to Alice.
- Step 3: Alice tells Bob whether or not his guess was right and proves her claim by sending x to Bob, too.
- Step 4: Bob verifies Alice's claim by computing f(x) and comparing it to y.

At first glance, this protocol looks good. But we are already warned. So, let us ask if a participant of this protocol can cheat.

We assume that f is indeed a one-way function. This is a good place to see why we have required one-way functions to be injective. If not, there could be two numbers x and x' such that x is even, x' is odd and f(x) = f(x').

We assume that f is indeed a one-way function. This is a good place to see why we have required one-way functions to be injective. If not, there could be two numbers x and x' such that x is even, x' is odd and f(x) = f(x').

But still, the definition of one-way function does not imply that we cannot compute that last bit of x. If we could, we already have to whole information needed.

We assume that f is indeed a one-way function. This is a good place to see why we have required one-way functions to be injective. If not, there could be two numbers x and x' such that x is even, x' is odd and f(x) = f(x').

But still, the definition of one-way function does not imply that we cannot compute that last bit of x. If we could, we already have to whole information needed.

This is a subtle point. So, we cannot prove anything about the protocol above.

We assume that f is indeed a one-way function. This is a good place to see why we have required one-way functions to be injective. If not, there could be two numbers x and x' such that x is even, x' is odd and f(x) = f(x').

But still, the definition of one-way function does not imply that we cannot compute that last bit of x. If we could, we already have to whole information needed.

This is a subtle point. So, we cannot prove anything about the protocol above.

Therefore, Blum (1981) proposed the following more advanced Protocol *CF* for flipping a coin per telephone.

Protocol CF

- Step 1: Alice chooses two huge primes p and q sends their product n = pq to Bob.
- Step 2: Bob chooses randomly a number s from $\{1, ..., \lfloor \frac{n}{2} \rfloor\}$. Furthermore, he computes $z = s^2 \mod n$ and sends z to Alice.
- Step 3: Alice computes the four discrete roots $\pm x$ and $\pm y$ of z modulo n. Let x' be the smaller number of $x \mod n$ and $-x \mod n$ and let y' be defined analogously.
- Step 4: Alice looks for the smallest bit position i in which x' and y' differ. Then she guesses one of these numbers and communicates her guess to Bob by telling him: "The ith bit of your number is 0" and "The ith bit of your number is 1," respectively.
- Step 5: Bob tells Alice whether or not her guess was correct (if it was correct, she wins).
- Step 6: Bob sends his number s to Alice.
- Step 7: Alice tells Bob the factorization of n.

Example 1

Here we keep p and q small for being able to follow the computations.

- (1) Alice chooses p = 5 and q = 13, and sends 65.
- (2) Bob chooses 21, computes $51 \equiv 21^2 \mod 65$, and sends 51.
- (3) Alice computes ± 21 and ± 31 , chooses x' = 21 = 10101 and y' = 31 = 11111.
- (4) She guesses "The ith bit of your number is 1."
- (5) Bob tells her that she is wrong.
- (6) Bob sends 21 to Alice.
- (7) Alice sends 5 and 13 to Bob.

This protocol looks more complex than the previous one. It is also not obvious whether or not it is correct and fair. Thus, we have do to the following.

This protocol looks more complex than the previous one. It is also not obvious whether or not it is correct and fair. Thus, we have do to the following.

- Analyze the Protocol *CF* carefully.
- Output: Check whether or not it is fair and secure.
- Finding out whether or not it can be executed efficiently.

This protocol looks more complex than the previous one. It is also not obvious whether or not it is correct and fair. Thus, we have do to the following.

- Analyze the Protocol *CF* carefully.
- Output: Check whether or not it is fair and secure.
- § Finding out whether or not it can be executed efficiently.

If Alice is making her guess randomly and if Bob is choosing his number s indeed randomly from the set $\{1, \ldots, \lfloor \frac{n}{2} \rfloor\}$, then the probability that Alice wins is clearly 1/2.

End 0

Flipping a Coin per Telephone VII

Furthermore, it is not a good idea for Bob not to choose his number randomly (provided the protocol is executed repeatedly), since a certain preference for some numbers would offer Alice a possibility to possibly increase her chance of winning.

Furthermore, it is not a good idea for Bob not to choose his number randomly (provided the protocol is executed repeatedly), since a certain preference for some numbers would offer Alice a possibility to possibly increase her chance of winning.

So, the most important question we have to study right here is whether or not Bob can possibly cheat, if he is changing s *after* having sent *z* to Alice.

Furthermore, it is not a good idea for Bob not to choose his number randomly (provided the protocol is executed repeatedly), since a certain preference for some numbers would offer Alice a possibility to possibly increase her chance of winning.

So, the most important question we have to study right here is whether or not Bob can possibly cheat, if he is changing s *after* having sent *z* to Alice.

In order to avoid being detected as cheater, Bob should possess x' as well as y'.

Taking into account that

$$\begin{array}{rcl} (\mathbf{x}')^2 &\equiv& z \mod n \\ (\mathbf{y}')^2 &\equiv& z \mod n \text{ , we get} \\ (\mathbf{x}')^2 - (\mathbf{y}')^2 &\equiv& 0 \mod n \text{ .} \end{array}$$

Furthermore, we have $x' \neq y' \mod n$. This clearly implies $x' - y' \neq 0 \mod n$. Additionally, it is not hard to see that we also have $x' + y' \neq 0 \mod n$. Thus, putting it all together we directly arrive at

 $(x')^2 - (y')^2 \ \equiv \ (x' - y')(x' + y') \equiv 0 \ \text{mod} \ n \, .$

Taking into account that

$$\begin{array}{rcl} (\mathbf{x}')^2 &\equiv& z \mod n \\ (\mathbf{y}')^2 &\equiv& z \mod n \text{ , we get} \\ (\mathbf{x}')^2 - (\mathbf{y}')^2 &\equiv& 0 \mod n \text{ .} \end{array}$$

Furthermore, we have $x' \neq y' \mod n$. This clearly implies $x' - y' \neq 0 \mod n$. Additionally, it is not hard to see that we also have $x' + y' \neq 0 \mod n$. Thus, putting it all together we directly arrive at

$$(x')^2 - (y')^2 \equiv (x' - y')(x' + y') \equiv 0 \mod n$$
.

This is possible if and only if

$$\begin{array}{rcl} gcd(n,x'+y') &=& p \ or \\ gcd(n,x'+y') &=& q \ . \end{array}$$

Thus, if Bob is able to cheat he is able to factorize n, too. Therefore, the security of our second protocol is based on the difficulty to factorize. We summarize our knowledge by the following theorem.

Thus, if Bob is able to cheat he is able to factorize n, too. Therefore, the security of our second protocol is based on the difficulty to factorize. We summarize our knowledge by the

following theorem.

Theorem 1

The Protocol CF is secure provided factoring is difficult.

Thus, if Bob is able to cheat he is able to factorize n, too. Therefore, the security of our second protocol is based on the difficulty to factorize. We summarize our knowledge by the following theorem.

Theorem 1

The Protocol CF is secure provided factoring is difficult.

We have elaborated this point here in some more detail, since it also shows why Alice is sending just one bit in Step 4 and not x' or y'.

So, it remains to show that the protocol can be executed efficiently.

Since Alice is knowing the factorization of n, it suffices to argue that Alice can efficiently compute discrete square roots modulo a prime. Again, we refer to Lecture 5, where we studied Berlekamp's (1970) procedure for taking discrete square roots modulo a prime. This algorithm is a Las Vegas method and has an expected running time that is polynomially bounded in the length of the input a and the modulus p.

Since Alice is knowing the factorization of n, it suffices to argue that Alice can efficiently compute discrete square roots modulo a prime. Again, we refer to Lecture 5, where we studied Berlekamp's (1970) procedure for taking discrete square roots modulo a prime. This algorithm is a Las Vegas method and has an expected running time that is polynomially bounded in the length of the input a and the modulus p.

So, the whole Protocol *CF* relies on the assumption that factoring is difficult.

Since Alice is knowing the factorization of n, it suffices to argue that Alice can efficiently compute discrete square roots modulo a prime. Again, we refer to Lecture 5, where we studied Berlekamp's (1970) procedure for taking discrete square roots modulo a prime. This algorithm is a Las Vegas method and has an expected running time that is polynomially bounded in the length of the input a and the modulus p.

So, the whole Protocol *CF* relies on the assumption that factoring is difficult.

Next, we look at another problem of high practical relevance, i.e., how to share a secret.

The problem we want to consider goes back to Liu (1968) who stated it as follows.

Eleven scientists are working on a secret project. They wish to lock up the documents in a cabinet so that the cabinet can be opened if and only if six or more of the scientists are present. What is the smallest number of keys to the locks each scientists must carry?

The problem we want to consider goes back to Liu (1968) who stated it as follows.

Eleven scientists are working on a secret project. They wish to lock up the documents in a cabinet so that the cabinet can be opened if and only if six or more of the scientists are present. What is the smallest number of keys to the locks each scientists must carry?

Shamir (1979) showed that the smallest solution comprises 462 locks at all and 252 keys per scientist.

Let us assume that all keys and all locks have numbers printed on them. The locks are numbered by pairwise different numbers, and without loss of generality we shall assume that they are numbered $\ell = 1, ..., n$.

Let us assume that all keys and all locks have numbers printed on them. The locks are numbered by pairwise different numbers, and without loss of generality we shall assume that they are numbered $\ell = 1, ..., n$.

A key with number ℓ can open the lock with number m if and only if $\ell = m$.

Let us assume that all keys and all locks have numbers printed on them. The locks are numbered by pairwise different numbers, and without loss of generality we shall assume that they are numbered $\ell = 1, ..., n$.

A key with number ℓ can open the lock with number m if and only if $\ell = m$.

Now, one can show:

Claim 1. *The smallest number of locks needed is* $\binom{n}{k-1}$ *and the smallest number of keys needed is* $\binom{n-1}{k-1}$.

Let us assume that all keys and all locks have numbers printed on them. The locks are numbered by pairwise different numbers, and without loss of generality we shall assume that they are numbered $\ell = 1, ..., n$.

A key with number ℓ can open the lock with number m if and only if $\ell = m$.

Now, one can show:

Claim 1. *The smallest number of locks needed is* $\binom{n}{k-1}$ *and the smallest number of keys needed is* $\binom{n-1}{k-1}$.

The proof can be found in the script.

Furthermore, it should be mentioned that $\binom{n}{k}$ becomes maximal for k = n/2, if n is even, and for k = (n + 1)/2, if k is odd. Moreover, $\binom{n}{n/2} = O(4^{n/2})$, and thus it grows exponentially in n.

Thus, a (k, n) threshold scheme, for being practically applicable, has to give up the intuitive appealing idea of locks and keys. Instead, we shall look for different possibilities to share a secret.

Furthermore, it should be mentioned that $\binom{n}{k}$ becomes maximal for k = n/2, if n is even, and for k = (n + 1)/2, if k is odd. Moreover, $\binom{n}{n/2} = O(4^{n/2})$, and thus it grows exponentially in n.

Thus, a (k, n) threshold scheme, for being practically applicable, has to give up the intuitive appealing idea of locks and keys. Instead, we shall look for different possibilities to share a secret.

For doing this, let us first give the general definition of a shared secret and of a (k, n) threshold scheme given by Shamir (1978). Assume we have n persons P_1, \ldots, P_n and a secret datum D which we want to divide into n pieces D_1, D_2, \ldots, D_n .

Definition 2

We say that n participants k-divide a secret, where $1 < k \le n$ provided the following 3 conditions are satisifed.

- (1) Each participant P_i possesses an information D_i which is not known to any other participant P_j , $i \neq j$ for $j \in \{1, ..., n\}$.
- (2) The knowledge of any k or more of the D₁, D₂,..., D_n pieces allows us to compute the whole datum D easily,
- (3) the knowledge of any k 1 or fewer D_1, D_2, \dots, D_n pieces leaves D completely undetermined.

A set $\{D_1, ..., D_n\}$ satisfying (2) and (3) is said to be a (k, n) threshold scheme.

The pieces D_i of information are referred to as a share. The example at the beginning of this chapter provides at least evidence that (k, n) threshold schemes can be realized.

The pieces D_i of information are referred to as a share. The example at the beginning of this chapter provides at least evidence that (k, n) threshold schemes can be realized.

However, while the idea of using locks is intuitively appealing we still have to outline how to simulate them by appropriately chosen problems that meet the wanted complexity theoretic requirements in Items (2) and (3). Furthermore, we aim to find a simulation such that the the number of simulated "locks" does no longer grow exponentially.

The pieces D_i of information are referred to as a share. The example at the beginning of this chapter provides at least evidence that (k, n) threshold schemes can be realized.

However, while the idea of using locks is intuitively appealing we still have to outline how to simulate them by appropriately chosen problems that meet the wanted complexity theoretic requirements in Items (2) and (3). Furthermore, we aim to find a simulation such that the the number of simulated "locks" does no longer grow exponentially.

The following construction is based on Mignotte's threshold sequences.

A sequence $m_1 < \cdots < m_n$ of pairwise relatively prime and positive numbers is said to be a (k, n) threshold sequence if

```
\mathfrak{m}_1 \cdot \mathfrak{m}_2 \cdot \cdots \cdot \mathfrak{m}_k > \mathfrak{m}_n \cdot \mathfrak{m}_{n-1} \cdot \cdots \cdot \mathfrak{m}_{n-k+2} (A)
```

A sequence $m_1 < \cdots < m_n$ of pairwise relatively prime and positive numbers is said to be a (k, n) threshold sequence if

$$\mathfrak{m}_1 \cdot \mathfrak{m}_2 \cdot \cdots \cdot \mathfrak{m}_k > \mathfrak{m}_n \cdot \mathfrak{m}_{n-1} \cdot \cdots \cdot \mathfrak{m}_{n-k+2}$$
(A)

Now, suppose, we have a (k, n) threshold sequence. We set

 $M = m_1 \cdot m_2 \cdot \dots \cdot m_k \text{, and}$ $N = m_n \cdot m_{n-1} \cdot \dots \cdot m_{n-k+2}.$

The secret is then any number D satisfying N \leqslant D \leqslant M. Now, the pieces for each participant are defined by

 $D_i = D \mod m_i$,

that is, P_i obtains D_i and nothing else, i = 1, ..., n.

Thus, Condition (1) of Definition 2 is fulfilled by construction.

Thus, Condition (1) of Definition 2 is fulfilled by construction. *Claim* 1. *Condition* (2) *is satisfied. Proof.* Let D_{i_1}, \ldots, D_{i_k} be any subset of k elements from $\{D_1, \ldots, D_n\}$. By the Chinese remainder theorem, the system

 $x \equiv D_i \mod m_i \text{,} \qquad i \in \{i_1, \dots, i_k\}$

has a uniquely determined solution \widehat{D} modulo $\prod_{j=1}^{k} \mathfrak{m}_{i_j}$.

Thus, Condition (1) of Definition 2 is fulfilled by construction.

Claim 1. *Condition* (2) *is satisfied*.

Proof. Let D_{i_1}, \ldots, D_{i_k} be any subset of k elements from $\{D_1, \ldots, D_n\}$. By the Chinese remainder theorem, the system

$$x \equiv D_i \mod m_i , \qquad i \in \{i_1, \dots, i_k\}$$

has a uniquely determined solution \widehat{D} modulo $\prod_{j=1}^{k} \mathfrak{m}_{i_j}$. By the definition of a (k, n) threshold sequence and the choice of D we obtain

$$D \leqslant M = \mathfrak{m}_1 \cdot \mathfrak{m}_2 \cdot \cdots \cdot \mathfrak{m}_k \leqslant \prod_{j=1}^n \mathfrak{m}_{i_j}$$
.

Thus, Condition (1) of Definition 2 is fulfilled by construction.

Claim 1. *Condition* (2) *is satisfied*.

Proof. Let D_{i_1}, \ldots, D_{i_k} be any subset of k elements from $\{D_1, \ldots, D_n\}$. By the Chinese remainder theorem, the system

$$x \equiv D_i \mod m_i$$
, $i \in \{i_1, \dots, i_k\}$

has a uniquely determined solution \widehat{D} modulo $\prod_{j=1}^{k} m_{i_j}$. By the definition of a (k, n) threshold sequence and the choice of D we obtain

$$D \leqslant M = \mathfrak{m}_1 \cdot \mathfrak{m}_2 \cdot \cdots \cdot \mathfrak{m}_k \leqslant \prod_{j=1}^{\kappa} \mathfrak{m}_{i_j}.$$

Since $\widehat{D} \equiv D_i \equiv D \mod m_i$ for all $i \in \{i_1, \dots, i_k\}$, we see that any k participants of the secret sharing scheme can compute the secret datum D. This proves (2).

Claim 2. Condition (3) is satisfied.

Claim 2. Condition (3) is satisfied.

Proof. Let $\{D_{i_1}, \ldots, D_{i_{k-1}}\}$ be any subset of k-1 elements from $\{D_1, \ldots, D_n\}$. Again, we may apply the Chinese remainder theorem, and obtain

$$\widehat{\mathsf{D}} = e_{\mathfrak{i}_1} \cdot \mathsf{D}_{\mathfrak{i}_1} + \dots + e_{\mathfrak{i}_{k-1}} \cdot \mathsf{D}_{\mathfrak{i}_{k-1}} \mod \prod_{j=1}^{k-1} \mathfrak{m}_{\mathfrak{i}_j} \tag{1}$$

Claim 2. Condition (3) is satisfied.

Proof. Let $\{D_{i_1}, \ldots, D_{i_{k-1}}\}$ be any subset of k-1 elements from $\{D_1, \ldots, D_n\}$. Again, we may apply the Chinese remainder theorem, and obtain

$$\widehat{D} = e_{i_1} \cdot D_{i_1} + \dots + e_{i_{k-1}} \cdot D_{i_{k-1}} \mod \prod_{j=1}^{k-1} m_{i_j}$$
 (1)

Obviously, (1) is the congruence containing all information we have. Nevertheless, (1) leaves many possibilities for D. Therefore, we continue by estimating the number of possibilities.

The biggest product of k - 1 numbers chosen from $\{m_1, \ldots, m_n\}$ is N. The smallest product of k numbers chosen from $\{m_1, \ldots, m_n\}$ is M. Since $D \equiv \widehat{D} \mod m_i$ for all $i \in \{i_1, \ldots, i_{k-1}\}$, we also have

$$\mathsf{D} \equiv \widehat{\mathsf{D}} \mod \prod_{j=1}^{k-1} \mathfrak{m}_{\mathfrak{i}_j} ,$$

(remember that the numbers m_i , $i \in \{1, ..., n\}$ are relatively prime), i.e., D and \widehat{D} differ by a multiple of $\prod_{j=1}^{k-1} m_{i_j}$. Consequently, one can try all

$$D=\widehat{D}+\prod_{j=1}^{k-1}m_{\mathfrak{i}_{j}},\qquad D=\widehat{D}+2\cdot\prod_{j=1}^{k-1}m_{\mathfrak{i}_{j}},\ldots,$$

This gives a lower bound of

$$\frac{M-N-1}{N}$$

many possibilities. Of course, the true value of D can be only determined, if one has an oracle for testing these possibilities. For example, it is well imaginable that one has only 3 trials to test D (like passwords).

This gives a lower bound of

$$\frac{M-N-1}{N}$$

many possibilities. Of course, the true value of D can be only determined, if one has an oracle for testing these possibilities. For example, it is well imaginable that one has only 3 trials to test D (like passwords).

Thus, it remains to show that one can always choose (k, n) threshold sequences in a way such that (M - N - 1)/N is large.

Let $\pi(x)$ be the number of all primes less than or equal to x. The prime number theorem is telling us that

$$\frac{x}{\ln x} < \pi(x) < \frac{5}{4} \cdot \frac{x}{\ln x}$$

for all $x \ge 114$,

i.e., there is constant c such that

$$\pi(\mathbf{x}) \leqslant \mathbf{c} \cdot \frac{\mathbf{x}}{\log \mathbf{x}} \, .$$

Let $\pi(x)$ be the number of all primes less than or equal to x. The prime number theorem is telling us that

$$rac{x}{\ln x} < \pi(x) < rac{5}{4} \cdot rac{x}{\ln x} \qquad ext{ for all } x \geqslant 114$$
 ,

i.e., there is constant c such that

$$\pi(\mathbf{x}) \leqslant \mathbf{c} \cdot \frac{\mathbf{x}}{\log \mathbf{x}} \; .$$

Now, let $\pi(n, \alpha)$ be the number of all primes in the interval (p_n^{α}, p_n) , where p_n is the n-th prime number and $\alpha \in (0, 1)$. Then, we can show the following lemma.

Lemma 1

Let $n \in \mathbb{N}$ with $n \ge 2$. For every k with $2 \le k \le n$ there are arbitrarily big numbers y such that

$$\pi\left(y,\frac{k^2-1}{k^2}\right) > n \; .$$

Lemma 1

Let $n \in \mathbb{N}$ with $n \ge 2$. For every k with $2 \le k \le n$ there are arbitrarily big numbers y such that

$$\pi\left(y,\frac{k^2-1}{k^2}\right) > n \; .$$

Before proving the lemma, we show how to get the desired result from it.

We choose y such that $\pi(y, \frac{k^2-1}{k^2}) > n$. That is, in the interval

$$(\mathfrak{p}_{y}^{(k^{2}-1)/k^{2}},\mathfrak{p}_{y}]$$

there are at least n prime numbers. Let m_1, \ldots, m_n be the first n primes in this interval.

Claim. m_1, \ldots, m_n form a (k, n) threshold sequence.

Claim. m_1, \ldots, m_n *form a* (k, n) *threshold sequence.* The condition $m_1 < m_2 < \cdots < m_n$ is obvious. Moreover,

$$\begin{split} M &= \prod_{i=1}^k \mathfrak{m}_i > \mathfrak{m}_1^k \geqslant \left(p_y^{\frac{k^2 - 1}{k^2}} \right)^k \\ &= p_y^{\frac{(k+1)(k-1)}{k}} > p_y^{k-1} \\ &\geqslant \mathfrak{m}_n \cdot \mathfrak{m}_{n-1} \cdot \cdots \cdot \mathfrak{m}_{n-k+2} = N \,, \end{split}$$

where the last inequality holds, since $\mathfrak{m}_n \leq \mathfrak{p}_y$ and thus $\mathfrak{m}_n^{k-1} \leq \mathfrak{p}_y^{k-1}$. This proves the claim.

Finally, we obtain

$$\frac{M-N}{N} \ge \frac{p_y^{\frac{k^2-1}{k}} - p_y^{k-1}}{p_y^{k-1}} = p_y^{\frac{k-1}{k}} - 1.$$

Consequently, one can start with a lower bound B for (M - N - 1)/N. Then one searches for p_z such that

$$\mathfrak{p}_z^{\frac{k-1}{k}} - 1 \geqslant B.$$

By the lemma above, then there exists a $y \ge z$ such that one can form the wanted (k, n)-threshold sequence from m_1, \ldots, m_n .

Finally, we have to show Lemma 1 above.

Let k, n be arbitrarily fixed, and let $\alpha \in (0, 1)$. By the prime number theorem we know that $p_m = O(m \log m)$. Hence, $p_m^{\alpha} = O(m^{\alpha} (\log m)^{\alpha})$. Furthermore,

$$\pi(\mathfrak{m},\alpha)=\pi(\mathfrak{p}_{\mathfrak{m}})-\pi(\mathfrak{p}_{\mathfrak{m}}^{\alpha}).$$

We choose c_1 such that $\pi(p_m) \ge c_1 m$ and c_2 such that

$$\begin{aligned} \pi(\mathfrak{p}_{\mathfrak{m}}^{\alpha}) &\leqslant \quad c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\log(\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha})} \leqslant c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\log \mathfrak{m}^{\alpha}} \\ &= \quad c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\alpha \cdot \log \mathfrak{m}} \,. \end{aligned}$$

Finally, we have to show Lemma 1 above.

Let k, n be arbitrarily fixed, and let $\alpha \in (0, 1)$. By the prime number theorem we know that $p_m = O(m \log m)$. Hence, $p_m^{\alpha} = O(m^{\alpha} (\log m)^{\alpha})$. Furthermore,

$$\pi(\mathfrak{m}, \alpha) = \pi(\mathfrak{p}_{\mathfrak{m}}) - \pi(\mathfrak{p}_{\mathfrak{m}}^{\alpha}) .$$

We choose c_1 such that $\pi(p_m) \ge c_1 m$ and c_2 such that

$$\begin{aligned} \pi(\mathfrak{p}_{\mathfrak{m}}^{\alpha}) &\leqslant \quad c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\log(\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha})} \leqslant c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\log \mathfrak{m}^{\alpha}} \\ &= \quad c_{2} \cdot \frac{\mathfrak{m}^{\alpha}(\log \mathfrak{m})^{\alpha}}{\alpha \cdot \log \mathfrak{m}} \,. \end{aligned}$$

Thus, we obtain:

$$\begin{aligned} \pi(\mathbf{m}, \alpha) &= \pi(\mathbf{p}_{\mathbf{m}}) - \pi(\mathbf{p}_{\mathbf{m}}^{\alpha}) \geqslant c_{1} \cdot \frac{\mathbf{m} \log \mathbf{m}}{\log \mathbf{m}} - c_{2} \cdot \frac{\mathbf{m}^{\alpha} (\log \mathbf{m})^{\alpha}}{\alpha \cdot \log \mathbf{m}} \\ \geqslant & c \left(\frac{\mathbf{m} \log \mathbf{m}}{\log \mathbf{m}} - \frac{\mathbf{m}^{\alpha} (\log \mathbf{m})^{\alpha}}{\alpha \cdot \log \mathbf{m}} \right) \\ &= & \frac{c}{\log \mathbf{m}} \left(\mathbf{m} \log \mathbf{m} - \frac{\mathbf{m}^{\alpha} (\log \mathbf{m})^{\alpha}}{\alpha} \right) \\ &= & \frac{c \cdot \mathbf{m} \log \mathbf{m}}{\log \mathbf{m}} \left(1 - \frac{1}{\alpha \cdot \mathbf{m}^{1-\alpha} (\log \mathbf{m})^{1-\alpha}} \right) \\ &= & c \cdot \mathbf{m} \underbrace{\left(1 - \frac{1}{\alpha \cdot \mathbf{m}^{1-\alpha} (\log \mathbf{m})^{1-\alpha}} \right)}_{=:X} \end{aligned}$$

The expression X converges to 1 as m tends to infinity.

Consequently, for all sufficiently large values of \mathfrak{m} we see that $\pi(\mathfrak{m},\alpha)>\mathfrak{n}.$

Thus, setting $\alpha = (k^2 - 1)/k^2$, the lemma follows.

Consequently, for all sufficiently large values of m we see that $\pi(m, \alpha) > n$.

Thus, setting $\alpha = (k^2 - 1)/k^2$, the lemma follows.

So, for threshold schemes we were able to prove a very satisfactory result.

Thank you!