Guessing Cryptographic Secrets and Oblivious Distributed Guessing

Serdar Boztaş

School of Mathematical and Geospatial Sciences

RMIT University

August 2014
Monash University

Outline

1. Introduction
   - Problem Statement
   - Our Contribution
2. Guessing, Predictability and Entropy
   - Definitions
   - Guessing by one attacker
   - Limited Resource Guessing
   - Power and Memory Constrained Guessor Minimizing Failure Probability
   - Multiple Memory Constrained Oblivious Guessors
3. Conclusions

Outline

1. Introduction
   - Problem Statement
   - Our Contribution
2. Guessing, Predictability and Entropy
   - Definitions
   - Guessing by one attacker
   - Limited Resource Guessing
   - Power and Memory Constrained Guessor Minimizing Failure Probability
   - Multiple Memory Constrained Oblivious Guessors
3. Conclusions

Outline

1. Introduction
   - Problem Statement
   - Our Contribution
2. Guessing, Predictability and Entropy
   - Definitions
   - Guessing by one attacker
   - Limited Resource Guessing
   - Power and Memory Constrained Guessor Minimizing Failure Probability
   - Multiple Memory Constrained Oblivious Guessors
3. Conclusions

Problem Statement

- Let X be an unknown discrete random variable with distribution $\mathbb{P}$ and taking values in $\mathcal{X}$ which is finite or countable. X could represent an unknown key, IV, or password for a cryptosystem, or an unknown quantity of information security value.

- To model problems of interest, we assume that the *guessor* is not all-powerful and can only ask atomic questions (e.g., query keys/passwords) regarding singletons in $\mathcal{X}$. This corresponds to submitting the password and seeing if the login is successful or not.

- We assume that a sequence of questions of the form

  Is $X = x$*?*

  are posed until the first YES answer determines the value of the random variable X.

Problem Statement

- Let X be an unknown discrete random variable with distribution $\mathbb{P}$ and taking values in $\mathcal{X}$ which is finite or countable. X could represent an unknown key, IV, or password for a cryptosystem, or an unknown quantity of information security value.

- To model problems of interest, we assume that the *guessor* is not all-powerful and can only ask atomic questions (e.g., query keys/passwords) regarding singletons in $\mathcal{X}$. This corresponds to submitting the password and seeing if the login is successful or not.

- We assume that a sequence of questions of the form

  Is $X = x$*?*

  are posed until the first YES answer determines the value of the random variable X.

Problem Statement

- Let X be an unknown discrete random variable with distribution $\mathbb{P}$ and taking values in $\mathcal{X}$ which is finite or countable. X could represent an unknown key, IV, or password for a cryptosystem, or an unknown quantity of information security value.

- To model problems of interest, we assume that the *guessor* is not all-powerful and can only ask atomic questions (e.g., query keys/passwords) regarding singletons in $\mathcal{X}$. This corresponds to submitting the password and seeing if the login is successful or not.

- We assume that a sequence of questions of the form

  Is $X = x$*?*

  are posed until the first YES answer determines the value of the random variable X.

Problem History

- The link between guessing and entropy was popularized by James L. Massey in the early 1990s. *If X has high entropy is it hard to Guess? Is Shannon entropy the right measure?*
- The problem of bounding the expected number of guesses in terms of Rényi entropies was investigated by Erdal Arikan in the context of sequential decoding. Arikan used the Hölder Inequality to obtain his bound.
- John Pliam independently investigated the relationship between entropy, "guesswork" and security.
- Boztas improved Arikan's bound and presented other tighter bounds for specific cases.
- The concept of "guessing entropy" has (i) been adopted by NIST as a measure of password strength, and (ii) also applied by others to graphical passwords.

Problem History

- The link between guessing and entropy was popularized by James L. Massey in the early 1990s. *If X has high entropy is it hard to Guess? Is Shannon entropy the right measure?*
- The problem of bounding the expected number of guesses in terms of Rényi entropies was investigated by Erdal Arikan in the context of sequential decoding. Arikan used the Hölder Inequality to obtain his bound.
- John Pliam independently investigated the relationship between entropy, "guesswork" and security.
- Boztaş improved Arikan's bound and presented other tighter bounds for specific cases.
- The concept of "guessing entropy" has (i) been adopted by NIST as a measure of password strength and (ii) also applied by others to graphical passwords.

Problem History

- The link between guessing and entropy was popularized by James L. Massey in the early 1990s. *If X has high entropy is it hard to Guess? Is Shannon entropy the right measure?*
- The problem of bounding the expected number of guesses in terms of Rényi entropies was investigated by Erdal Arikan in the context of sequential decoding. Arikan used the Hölder Inequality to obtain his bound.
- John Pliam independently investigated the relationship between entropy, "guesswork" and security.
- Boztaş improved Arikan's bound and presented other tighter bounds for specific cases.
- The concept of "guessing entropy" has (i) been adopted by NIST as a measure of password strength; and (ii) also applied by others to graphical passwords.

Problem History

- The link between guessing and entropy was popularized by James L. Massey in the early 1990s. *If X has high entropy is it hard to Guess? Is Shannon entropy the right measure?*
- The problem of bounding the expected number of guesses in terms of Rényi entropies was investigated by Erdal Arikan in the context of sequential decoding. Arikan used the Hölder Inequality to obtain his bound.
- John Pliam independently investigated the relationship between entropy, "guesswork" and security.
- Boztaş improved Arikan's bound and presented other tighter bounds for specific cases.
- The concept of "guessing entropy" has (i) been adopted by NIST as a measure of password strength; and (ii) also applied by others to graphical passwords.

Problem History

- The link between guessing and entropy was popularized by James L. Massey in the early 1990s. *If X has high entropy is it hard to Guess? Is Shannon entropy the right measure?*
- The problem of bounding the expected number of guesses in terms of Rényi entropies was investigated by Erdal Arikan in the context of sequential decoding. Arikan used the Hölder Inequality to obtain his bound.
- John Pliam independently investigated the relationship between entropy, "guesswork" and security.
- Boztaș improved Arikan's bound and presented other tighter bounds for specific cases.
- The concept of "guessing entropy" has (i) been adopted by NIST as a measure of password strength; and (ii) also applied by others to graphical passwords.

Our Contribution

- In this talk we first focus on a *Single Attacker Guessing* an unknown random variable X.
- In this simple form, the problem is easier to state and analyze, and we revisit proofs of the early results in estimating the *average number of guesses* to determine X.
- This is the quantity called "guessing entropy" by NIST. A related quantity defined by Pliam, which specifies the minimal number of guesses required to succeed with a given probability in guessing X is also of interest.

Our Contribution

- In this talk we first focus on a *Single Attacker Guessing* an unknown random variable X.
- In this simple form, the problem is easier to state and analyze, and we revisit proofs of the early results in estimating the *average number of guesses* to determine X.
- This is the quantity called "guessing entropy" by NIST. A related quantity defined by Pliam, which specifies the minimal number of guesses required to succeed with a given probability in guessing X is also of interest.

Our Contribution

- In this talk we first focus on a *Single Attacker Guessing* an unknown random variable X.
- In this simple form, the problem is easier to state and analyze, and we revisit proofs of the early results in estimating the *average number of guesses* to determine X.
- This is the quantity called “guessing entropy” by NIST. A related quantity defined by Pliam, which specifies the minimal number of guesses required to succeed with a given probability in guessing X is also of interest.

Our Contribution

- Consider a single guessor. He can guess X in order of decreasing probability. Clearly this minimizes the expected number of guesses. *How is this related to the entropy of X?*

- It is tempting to have a number of different guessors working in parallel in trying to determine X, but tricky to make this practical and scalable if they have to keep track of what each other is guessing–consider guessors entering and leaving the group performing the search.

- Moreover the computational power of each participant (thus the rate at which they can implement the guessing mechanism) can vary a great deal. These factors make the study of *Oblivious* Distributed Guessing of interest.

Our Contribution

- Consider a single guessor. He can guess X in order of decreasing probability. Clearly this minimizes the expected number of guesses. *How is this related to the entropy of X?*
- It is tempting to have a number of different guessors working in parallel in trying to determine X, but tricky to make this practical and scalable if they have to keep track of what each other is guessing–consider guessors entering and leaving the group performing the search.
- Moreover the computational power of each participant (thus the rate at which they can implement the guessing mechanism) can vary a great deal. These factors make the study of *Oblivious* Distributed Guessing of interest.

Our Contribution

- Consider a single guessor. He can guess X in order of decreasing probability. Clearly this minimizes the expected number of guesses. *How is this related to the entropy of X?*

- It is tempting to have a number of different guessors working in parallel in trying to determine X, but tricky to make this practical and scalable if they have to keep track of what each other is guessing–consider guessors entering and leaving the group performing the search.

- Moreover the computational power of each participant (thus the rate at which they can implement the guessing mechanism) can vary a great deal. These factors make the study of *Oblivious* Distributed Guessing of interest.

Definitions

- A *guessing strategy* can be represented by a function $G : \mathcal{X} \to \{1, 2, \ldots\}$ where $G(k)$ equals the time index of the question **Is** $X = k$**?**.

- Clearly, G must be invertible on its range $\{1, 2, \ldots\}$ since only one element may be probed at any given time by a guessor. Since the answers to the queries **Is** $X = k$**?** are noiseless, it is enough to ask the above question *exactly once* for each $k \geq 1$. Hence the mapping G must be one-to-one and onto.

- Assuming that the guessor knows $\mathbb{P}$ she is interested in minimizing–an increasing function of–the number of questions required to determine X. Formally, she wants to minimize a positive moment $\mathbb{E}[G^\rho]$ (mostly $\rho = 1$ is of interest) where

$$\mathbb{E}[G^\rho] = \sum_{x \in \mathcal{X}} \mathbb{P}(x) G(x)^\rho = \sum_{k \geq 1} k^\rho \mathbb{P}(G^{-1}(k)).$$

Definitions

- A *guessing strategy* can be represented by a function $G : \mathcal{X} \to \{1, 2, \dots\}$ where $G(k)$ equals the time index of the question **Is** $X = k$**?**.
- Clearly, G must be invertible on its range $\{1, 2, \dots\}$ since only one element may be probed at any given time by a guessor. Since the answers to the queries **Is** $X = k$**?** are noiseless, it is enough to ask the above question *exactly once* for each $k \geq 1$. Hence the mapping G must be one-to-one and onto.
- Assuming that the guessor knows $\mathbb{P}$ she is interested in minimizing–an increasing function of–the number of questions required to determine X. Formally, she wants to minimize a positive moment $\mathbb{E}[G^\rho]$ (mostly $\rho = 1$ is of interest) where

$$\mathbb{E}[G^\rho] = \sum_{x \in \mathcal{X}} \mathbb{P}(x) G(x)^\rho = \sum_{k \geq 1} k^\rho \mathbb{P}(G^{-1}(k)).$$

Definitions

- A *guessing strategy* can be represented by a function $G : \mathcal{X} \to \{1, 2, \ldots\}$ where $G(k)$ equals the time index of the question **Is** $X = k$**?**.
- Clearly, G must be invertible on its range $\{1, 2, \ldots\}$ since only one element may be probed at any given time by a guessor. Since the answers to the queries **Is** $X = k$**?** are noiseless, it is enough to ask the above question *exactly once* for each $k \geq 1$. Hence the mapping G must be one-to-one and onto.
- Assuming that the guessor knows $\mathbb{P}$ she is interested in minimizing–an increasing function of–the number of questions required to determine X. Formally, she wants to minimize a positive moment $\mathbb{E}[G^\rho]$ (mostly $\rho = 1$ is of interest) where

$$\mathbb{E}[G^\rho] = \sum_{x \in \mathcal{X}} \mathbb{P}(x) G(x)^\rho = \sum_{k \geq 1} k^\rho \mathbb{P}(G^{-1}(k)).$$

Definitions

- The Rényi entropy of order α of X is defined as

$$H_\alpha(X) = \frac{\log\left(\sum_{X \in \mathcal{Y}} \mathbb{P}(X)^\alpha\right)}{1-\alpha} \qquad \alpha \in [0,1) \cup (1,\infty),$$

  and is a generalization of the Shannon entropy

$$H(X) = -\sum_{X \in \mathcal{X}} \mathbb{P}(X) \log(\mathbb{P}(X))$$

  and obeys $\lim_{\alpha \to 1} H_\alpha(X) = H(X)$ as well as being strictly decreasing in α unless X is uniform on its support.

- Tsallis and other entropies also connected with Rényi entropy. Most entropies lack one or more of the nice properties of Shannon entropy, but can be useful in special settings.

Definitions

- The Rényi entropy of order α of X is defined as

$$H_\alpha(X) = \frac{\log\left(\sum_{X\in\mathcal{Y}} \mathbb{P}(X)^\alpha\right)}{1-\alpha} \qquad \alpha \in [0,1)\cup(1,\infty),$$

  and is a generalization of the Shannon entropy

$$H(X) = -\sum_{X\in\mathcal{X}} \mathbb{P}(X)\log(\mathbb{P}(X))$$

  and obeys $\lim_{\alpha\to 1} H_\alpha(X) = H(X)$ as well as being strictly decreasing in α unless X is uniform on its support.
- Tsallis and other entropies also connected with Rényi entropy. Most entropies lack one or more of the nice properties of Shannon entropy, but can be useful in special settings.

Guessing by one attacker

- Guess every value of X one by one in order of decreasing probability, when the distribution $\mathbb{P}(x)$ is known.

Theorem

(Arikan) *For all $\rho \geq 0$, a guessing algorithm for X obeys the lower bound*

$$\mathbb{E}[G(X)^\rho] \geq \frac{[\sum_{k=1}^M P_X(x_k)^{1/(1+\rho)}]^{1+\rho}}{(1+\ln M)^\rho},$$

while an optimal guessing algorithm for X satisfies the upper bound

$$\mathbb{E}[G(X)^\rho] \leq \left[\sum_{k=1}^{M} P_X(x_k)^{1/(1+\rho)}\right]^{1+\rho}.$$

Guessing by one attacker

- Guess every value of X one by one in order of decreasing probability, when the distribution $\mathbb{P}(x)$ is known.

Theorem

(Arikan) *For all $\rho \geq 0$, a guessing algorithm for X obeys the lower bound*

$$\mathbb{E}[G(X)^\rho] \geq \frac{[\sum_{k=1}^{M} P_X(x_k)^{1/(1+\rho)}]^{1+\rho}}{(1+\ln M)^\rho},$$

while an optimal guessing algorithm for X satisfies the upper bound

$$\mathbb{E}[G(X)^\rho] \leq \left[\sum_{k=1}^{M} P_X(x_k)^{1/(1+\rho)}\right]^{1+\rho}.$$

Guessing by one attacker

- Arikan's bounds give

$$\frac{\left[\sum_{k=1}^{M} \sqrt{P_X(x_k)}\right]^2}{(1+\ln M)} \leq \mathbb{E}[G(X)] \overset{(a)}{\leq} \left[\sum_{k=1}^{M} \sqrt{P_X(x_k)}\right]^2$$

where (a) applies to the optimal guessing sequence.

- Boztaş's improved upper bound gives

$$\mathbb{E}[G(X)] \leq \frac{1}{2}\left[\sum_{k=1}^{M} \sqrt{P_X(x_k)}\right]^2 + \frac{1}{2} = 2^{H_{1/2}(X)-1} + \frac{1}{2}$$

for a more general class of guessing sequences. These provide an operational definition of Rényi entropy of order 1/2.

Guessing by one attacker

- Arikan's bounds give

$$\frac{\left[\sum_{k=1}^{M}\sqrt{P_X(x_k)}\right]^2}{(1+\ln M)} \leq \mathbb{E}[G(X)] \overset{(a)}{\leq} \left[\sum_{k=1}^{M}\sqrt{P_X(x_k)}\right]^2$$

where (a) applies to the optimal guessing sequence.

- Boztaş's improved upper bound gives

$$\mathbb{E}[G(X)] \leq \frac{1}{2}\left[\sum_{k=1}^{M}\sqrt{P_X(x_k)}\right]^2 + \frac{1}{2} = 2^{H_{1/2}(X)-1} + \frac{1}{2}$$

for a more general class of guessing sequences. These provide an operational definition of Rényi entropy of order $1/2$.

Limited Resource Guessing

- Consider a set of guessors attacking multiple targets, whose passwords are assumed to come from the same distribution $\mathbb{P}(x)$.
- *Given $\mathbb{P}(x)$, how should the attacker(s) choose a distribution $\mathbb{Q}(x)$ in order to optimize some performance criterion, when all the guessor(s) draw random sequential guesses from $\mathbb{Q}(x)$?*
- In general the guessor(s) should work in parallel, independently.

Limited Resource Guessing

- Consider a set of guessors attacking multiple targets, whose passwords are assumed to come from the same distribution $\mathbb{P}(x)$.
- *Given* $\mathbb{P}(x)$, *how should the attacker(s) choose a distribution* $\mathbb{Q}(x)$ *in order to optimize some performance criterion, when all the guessor(s) draw random sequential guesses from* $\mathbb{Q}(x)$?
- In general the guessor(s) should work in parallel, independently.

Limited Resource Guessing

- Consider a set of guessors attacking multiple targets, whose passwords are assumed to come from the same distribution $\mathbb{P}(x)$.
- *Given* $\mathbb{P}(x)$*, how should the attacker(s) choose a distribution* $\mathbb{Q}(x)$ *in order to optimize some performance criterion, when all the guessor(s) draw random sequential guesses from* $\mathbb{Q}(x)$?
- In general the guessor(s) should work in parallel, independently.

Limited Memory Single Guessor

- Consider a single guessor who is memory constrained and won't keep track of past guesses, but knows the distribution $\mathbb{P}$ which the opponent uses to draw a single value X from $\mathcal{X}$.
- Define $G = \min\{k : X_k = X\}$ as a random variable which denotes the number of guesses before she is successful in exposing X. The guessor generates i.i.d. guesses $X_1, X_2, \ldots,$ from $\mathcal{X}$ according to a distribution $\mathbb{Q}(x)$ with the goal of minimizing $\mathbb{E}[G]$.
- Note that $G = k$ with probability $\sum_{x \in \mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^{k-1}\mathbb{Q}(x)$, where $k \geq 1$, by a success-fail argument. This is because

$$\mathbb{P}(G = k) = \sum_{x \in \mathcal{X}} \mathbb{P}(X = x)\mathbb{P}(G = k \mid X = x)$$

and we can use the geometric distribution with success probability $\mathbb{Q}(x)$.

Limited Memory Single Guessor

- Consider a single guessor who is memory constrained and won't keep track of past guesses, but knows the distribution $\mathbb{P}$ which the opponent uses to draw a single value X from $\mathcal{X}$.
- Define $G = \min\{k : X_k = X\}$ as a random variable which denotes the number of guesses before she is successful in exposing X. The guessor generates i.i.d. guesses $X_1, X_2, \ldots,$ from $\mathcal{X}$ according to a distribution $\mathbb{Q}(x)$ with the goal of minimizing $\mathbb{E}[G]$.
- Note that $G = k$ with probability $\sum_{x\in\mathcal{X}} \mathbb{P}(x)(1-\mathbb{Q}(x))^{k-1}\mathbb{Q}(x)$, where $k \geq 1$, by a success-fail argument. This is because

$$\mathbb{P}(G = k) = \sum_{x\in\mathcal{X}} \mathbb{P}(X = x)\mathbb{P}(G = k \mid X = x)$$

and we can use the geometric distribution with success probability $\mathbb{Q}(x)$.

Limited Memory Single Guessor

- Consider a single guessor who is memory constrained and won't keep track of past guesses, but knows the distribution $\mathbb{P}$ which the opponent uses to draw a single value X from $\mathcal{X}$.
- Define $G = \min\{k : X_k = X\}$ as a random variable which denotes the number of guesses before she is successful in exposing X. The guessor generates i.i.d. guesses $X_1, X_2, \ldots,$ from $\mathcal{X}$ according to a distribution $\mathbb{Q}(x)$ with the goal of minimizing $\mathbb{E}[G]$.
- Note that $G = k$ with probability $\sum_{x\in\mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^{k-1}\mathbb{Q}(x)$. where $k \geq 1$, by a success-fail argument. This is because

$$\mathbb{P}(G = k) = \sum_{x\in\mathcal{X}} \mathbb{P}(X = x)\mathbb{P}(G = k \mid X = x)$$

and we can use the geometric distribution with success probability $\mathbb{Q}(x)$.

Limited Memory Single Guessor

If we apply Lagrange multipliers with the Lagrangian

$$J = \mathbb{E}[G] + \lambda(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1) = \sum_{x \in \mathcal{X}} \frac{\mathbb{P}(x)}{\mathbb{Q}(x)} + \lambda(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1),$$

we can actually show that $\mathbb{E}[G]$ is minimized when we choose

$$\mathbb{Q}(x) \propto \sqrt{\mathbb{P}(x)}$$

which means that the distribution $\mathbb{Q}(x)$ should be "flatter" than $\mathbb{P}(x)$.

Theorem

The distribution $\mathbb{Q}$ which minimizes the expected number of guesses for single guessor targeting X with distribution $\mathbb{P}$ is

$$\mathbb{Q}(x) = \frac{\sqrt{\mathbb{P}(x)}}{\sum_{y \in \mathcal{X}} \sqrt{\mathbb{P}(y)}}$$

Limited Memory Single Guessor

- Easy to check the Lagrange multipliers give minimum.
- Note that if we choose $\mathbb{Q}(x) = \mathbb{P}(x)$ for all $x \in \mathcal{X}$ which may look like an attractive choice, we obtain $\mathbb{E}[G] = |\mathcal{X}|$ which is surprisingly high.
- What is the minimum value of the expectation which the guessor using Proposition 1 achieves? It is

$$\begin{aligned}\mathbb{E}[G] &= \sum_{x \in \mathcal{X}} \frac{\mathbb{P}(x)}{\mathbb{Q}(x)} = \sum_{y \in \mathcal{X}} \sqrt{\mathbb{P}(y)} \sum_{x \in \mathcal{X}} \frac{\mathbb{P}(x)}{\sqrt{\mathbb{P}(x)}} \\ &= \left[\sum \sqrt{\mathbb{P}(x)}\right]^2 = 2^{H_{1/2}(X)}\end{aligned}$$

which provides a new *operational definition* of Rényi entropy of order 1/2 relating it *exactly* to oblivious guessing.

Limited Memory Single Guessor

- Easy to check the Lagrange multipliers give minimum.
- Note that if we choose $\mathbb{Q}(x) = \mathbb{P}(x)$ for all $x \in \mathcal{X}$ which may look like an attractive choice, we obtain $\mathbb{E}[G] = |\mathcal{X}|$ which is surprisingly high.
- What is the minimum value of the expectation which the guessor using Proposition 1 achieves? It is

$$\begin{aligned} \mathbb{E}[G] &= \sum_{x\in\mathcal{X}} \frac{\mathbb{P}(x)}{\mathbb{Q}(x)} = \sum_{y\in\mathcal{X}} \sqrt{\mathbb{P}(y)} \sum_{x\in\mathcal{X}} \frac{\mathbb{P}(x)}{\sqrt{\mathbb{P}(x)}} \\ &= \left[\sum \sqrt{\mathbb{P}(x)}\right]^2 = 2^{H_{1/2}(X)} \end{aligned}$$

which provides a new *operational definition* of Rényi entropy of order 1/2 relating it *exactly* to oblivious guessing.

Limited Memory Single Guessor

- Easy to check the Lagrange multipliers give minimum.
- Note that if we choose $\mathbb{Q}(x) = \mathbb{P}(x)$ for all $x \in \mathcal{X}$ which may look like an attractive choice, we obtain $\mathbb{E}[G] = |\mathcal{X}|$ which is surprisingly high.
- What is the minimum value of the expectation which the guessor using Proposition 1 achieves? It is

$$\begin{aligned}\mathbb{E}[G] &= \sum_{x\in\mathcal{X}} \frac{\mathbb{P}(x)}{\mathbb{Q}(x)} = \sum_{y\in\mathcal{X}} \sqrt{\mathbb{P}(y)} \sum_{x\in\mathcal{X}} \frac{\mathbb{P}(x)}{\sqrt{\mathbb{P}(x)}} \\ &= \left[\sum \sqrt{\mathbb{P}(x)}\right]^2 = 2^{H_{1/2}(X)}\end{aligned}$$

which provides a new *operational definition* of Rényi entropy of order $1/2$ relating it *exactly* to oblivious guessing.

Power and Memory Constrained Guessor Minimizing Failure Probability

- Now the guesses are still i.i.d. from $\mathbb{Q}(x)$ but the guessor (e.g., a sensor net node) decides *ahead of time* that she will only use $L \in \mathbb{N}$ guesses. We aim to find the $\mathbb{Q}(x)$ which minimizes the failure probability in L guesses, namely

$$P_{fail}(L) = \sum_{x \in \mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^L.$$

- This yields the Lagrangian

$$\begin{aligned} J &= P_{fail}(L) + \lambda(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1) \\ &= \sum_{x \in \mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^L + \lambda(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1). \end{aligned}$$

Power and Memory Constrained Guessor Minimizing Failure Probability

- Now the guesses are still i.i.d. from $\mathbb{Q}(x)$ but the guessor (e.g., a sensor net node) decides *ahead of time* that she will only use $L \in \mathbb{N}$ guesses. We aim to find the $\mathbb{Q}(x)$ which minimizes the failure probability in L guesses, namely

$$P_{fail}(L) = \sum_{x \in \mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^L.$$

- This yields the Lagrangian

$$\begin{aligned} J &= P_{fail}(L) + \lambda\Big(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1\Big) \\ &= \sum_{x \in \mathcal{X}} \mathbb{P}(x)(1 - \mathbb{Q}(x))^L + \lambda\Big(\sum_{x \in \mathcal{X}} \mathbb{Q}(x) - 1\Big). \end{aligned}$$

Power and Memory Constrained Guessor Minimizing Failure Probability

- The Lagrangian leads to the conditions

$$\frac{\partial J}{\partial \mathbb{Q}(x)} = -L\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-1} = -\lambda, \qquad \forall x \in \mathcal{X}$$

- Considering the Lagrangian and observing that L is constant, we have

$$\mathbb{Q}(x) = 1 - (\mu/\mathbb{P}(x))^{1/(L-1)}$$

for some positive constant $\mu = \lambda/L$.

- The second derivative is

$$\frac{\partial^2 J}{\partial \mathbb{Q}(x)^2} = L(L-1)\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-2}$$

and if we assume the non-degeneracy condition $0 < \mathbb{Q}(x) < 1$ for all $x \in \mathcal{X}$ and $L > 1$ we conclude it is positive

Power and Memory Constrained Guessor Minimizing Failure Probability

- The Lagrangian leads to the conditions

$$\frac{\partial J}{\partial \mathbb{Q}(x)} = -L\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-1} = -\lambda, \qquad \forall x \in \mathcal{X}$$

- Considering the Lagrangian and observing that L is constant, we have

$$\mathbb{Q}(x) = 1 - (\mu/\mathbb{P}(x))^{1/(L-1)}$$

for some positive constant $\mu = \lambda/L$.

- The second derivative is

$$\frac{\partial^2 J}{\partial \mathbb{Q}(x)^2} = L(L-1)\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-2}$$

and if we assume the non-degeneracy condition $0 < \mathbb{Q}(x) < 1$ for all $x \in \mathcal{X}$ and $L > 1$ we conclude it is positive.

Power and Memory Constrained Guessor Minimizing Failure Probability

- The Lagrangian leads to the conditions

$$\frac{\partial J}{\partial \mathbb{Q}(x)} = -L\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-1} = -\lambda, \qquad \forall x \in \mathcal{X}$$

- Considering the Lagrangian and observing that L is constant, we have

$$\mathbb{Q}(x) = 1 - (\mu/\mathbb{P}(x))^{1/(L-1)}$$

  for some positive constant $\mu = \lambda/L$.

- The second derivative is

$$\frac{\partial^2 J}{\partial \mathbb{Q}(x)^2} = L(L-1)\mathbb{P}(x)(1-\mathbb{Q}(x))^{L-2}$$

  and if we assume the non-degeneracy condition $0 < \mathbb{Q}(x) < 1$ for all $x \in \mathcal{X}$ and $L > 1$ we conclude it is positive.

Power and Memory Constrained Guessor Minimizing Failure Probability

Thus we have a minimum for $P_{fail}(L)$. The normalization condition can be shown to yield

$$\mu = \left(\frac{|\mathcal{X}| - 1}{\sum_{x \in \mathcal{X}} \mathbb{P}(x)^{-1/(L-1)}} \right)^{L-1},$$

thus proving:

Theorem

If the attacker is restricted to a fixed number of $L \geq 2$ guesses, her optimal oblivious strategy is to generate L i.i.d. guesses from the following distribution

$$\mathbb{Q}(x) = 1 - \left[\frac{|\mathcal{X}| - 1}{\sum_{y \in \mathcal{X}} \left(\mathbb{P}(x)/\mathbb{P}(y)\right)^{-1/(L-1)}} \right], \qquad \forall x \in \mathcal{X}$$

Multiple Memory Constrained Oblivious Guessors

- Consider $v \geq 2$ guessors working in parallel, each drawing i.i.d. guesses from $\mathbb{Q}(x)$, but not coordinating their guesses. If they collectively work at a rate v times the rate of the single guessor, then

$$\left\lfloor \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rfloor \leq \mathbb{E}_{\mathbb{Q}}[G_v] \leq \left\lceil \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rceil$$

  where $\mathbb{E}_{\mathbb{Q}}[G_v]$ denotes the expected number of guesses when v guessors each use $\mathbb{Q}(x)$.

- How should we optimize $\mathbb{Q}(x)$ once v is fixed?
- Drop the subscript $\mathbb{Q}$ from the expectations and note that

$$P[G_v = k] = Pr[G \in [(k-1)v + 1, kv] \cap \mathbb{Z}^+].$$

Multiple Memory Constrained Oblivious Guessors

- Consider $v \geq 2$ guessors working in parallel, each drawing i.i.d. guesses from $\mathbb{Q}(x)$, but not coordinating their guesses. If they collectively work at a rate v times the rate of the single guessor, then

$$\left\lfloor \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rfloor \leq \mathbb{E}_{\mathbb{Q}}[G_v] \leq \left\lceil \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rceil$$

  where $\mathbb{E}_{\mathbb{Q}}[G_v]$ denotes the expected number of guesses when v guessors each use $\mathbb{Q}(x)$.

- How should we optimize $\mathbb{Q}(x)$ once v is fixed?

- Drop the subscript $\mathbb{Q}$ from the expectations and note that

$$P[G_v = k] = Pr[G \in [(k-1)v + 1, kv] \cap \mathbb{Z}^+].$$

Multiple Memory Constrained Oblivious Guessors

- Consider $v \geq 2$ guessors working in parallel, each drawing i.i.d. guesses from $\mathbb{Q}(x)$, but not coordinating their guesses. If they collectively work at a rate v times the rate of the single guessor, then

$$\left\lfloor \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rfloor \leq \mathbb{E}_{\mathbb{Q}}[G_v] \leq \left\lceil \frac{\mathbb{E}_{\mathbb{Q}}[G]}{v} \right\rceil$$

  where $\mathbb{E}_{\mathbb{Q}}[G_v]$ denotes the expected number of guesses when v guessors each use $\mathbb{Q}(x)$.

- How should we optimize $\mathbb{Q}(x)$ once v is fixed?
- Drop the subscript $\mathbb{Q}$ from the expectations and note that

$$P[G_v = k] = Pr[G \in [(k-1)v + 1, kv] \cap \mathbb{Z}^+].$$

Multiple Memory Constrained Oblivious Guessors

- We obtain

$$\mathbb{E}[G_v] = \sum_{x\in\mathcal{X}} \mathbb{P}(x)\mathbb{Q}(x) \sum_{k=0}^{\infty} (1+k)[(1-\mathbb{Q}(x))^v]^k \sum_{j=1}^{v} (1-\mathbb{Q}(x))^{j-1},$$

or

$$\mathbb{E}[G_v] = \sum_{x\in\mathcal{X}} \mathbb{P}(x)\mathbb{Q}(x) \sum_{k=0}^{\infty} (1+k)[(1-\mathbb{Q}(x))^v]^k \left[\frac{1-(1-\mathbb{Q}(x))^v}{\mathbb{Q}(x)}\right],$$

- Using generation functions yields

$$\mathbb{E}[G_v] = \sum_{x\in\mathcal{X}} \left(\frac{\mathbb{P}(x)}{1-(1-\mathbb{Q}(x))^v}\right).$$

and the Lagrangian is now

$$J_v = \mathbb{E}[G_v] + \lambda\Big(\sum_{x\in\mathcal{X}} \mathbb{Q}(x) - 1\Big)$$

Multiple Memory Constrained Oblivious Guessors

- We obtain

$$\mathbb{E}[G_v]=\sum_{x\in\mathcal{X}}\mathbb{P}(x)\mathbb{Q}(x)\sum_{k=0}^{\infty}(1+k)[(1-\mathbb{Q}(x))^v]^k\sum_{j=1}^{v}(1-\mathbb{Q}(x))^{j-1},$$

or

$$\mathbb{E}[G_v]=\sum_{x\in\mathcal{X}}\mathbb{P}(x)\mathbb{Q}(x)\sum_{k=0}^{\infty}(1+k)[(1-\mathbb{Q}(x))^v]^k\left[\frac{1-(1-\mathbb{Q}(x))^v}{\mathbb{Q}(x)}\right],$$

- Using generation functions yields

$$\mathbb{E}[G_v]=\sum_{x\in\mathcal{X}}\left(\frac{\mathbb{P}(x)}{1-(1-\mathbb{Q}(x))^v}\right).$$

and the Lagrangian is now

$$J_v=\mathbb{E}[G_v]+\lambda(\sum_{x\in\mathcal{X}}\mathbb{Q}(x)-1)$$

Multiple Memory Constrained Oblivious Guessors

- Differentiation indicates that the optimum distribution $\mathbb{Q}(x)$ satisfies

$$\frac{v(1-\mathbb{Q}(x))^{v-1}}{(1-(1-\mathbb{Q}(x))^{v})^{2}} \propto \frac{1}{\mathbb{P}(x)}.$$

Let $R(x) = 1 - \mathbb{Q}(x)$ which takes on values in $(0,1)$ but is not a probability distribution since $\sum_x R(x) = |\mathcal{X}| - 1$.

- Thus we have

$$\frac{(1-R(x)^{v})^{2}}{vR(x)^{v-1}} \propto \mathbb{P}(x)$$

and by considering the function $f(u) = \frac{(1-u^v)^2}{vu^{v-1}}$ on $(0,1)$ and its derivative

$$f'(u) = -\frac{(1-u^{v})[(v+1)u^{v}+v-1]}{vu^{v}}$$

we conclude that we have a minimum.

Multiple Memory Constrained Oblivious Guessors

- Differentiation indicates that the optimum distribution $\mathbb{Q}(x)$ satisfies
$$\frac{v(1-\mathbb{Q}(x))^{v-1}}{(1-(1-\mathbb{Q}(x))^v)^2} \propto \frac{1}{\mathbb{P}(x)}.$$
Let $R(x) = 1 - \mathbb{Q}(x)$ which takes on values in $(0,1)$ but is not a probability distribution since $\sum_x R(x) = |\mathcal{X}| - 1$.
- Thus we have
$$\frac{(1-R(x)^v)^2}{vR(x)^{v-1}} \propto \mathbb{P}(x)$$
and by considering the function $f(u) = \frac{(1-u^v)^2}{vu^{v-1}}$ on $(0,1)$ and its derivative
$$f'(u) = -\frac{(1-u^v)[(v+1)u^v + v - 1]}{vu^v}$$
we conclude that we have a minimum.

Multiple Memory Constrained Oblivious Guessors

Theorem

v oblivious memory constrained attackers wanting to minimize $\mathbb{E}[G_v]$ should generate i.i.d. guesses from

$$\mathbb{Q}(x) \;\propto\; \left[1 - f^{-1}(\mathbb{P}(x))\right].$$

For a distribution $\mathbb{P}$ for which the maximum probability is much smaller than one, we have

$$z = f(u) = (1-u^v)^2/(vu^{v-1}) \approx (1-2u)/v$$

giving $f^{-1}(z) \approx (1-vz)/2$ resulting in the fast approximation

$$\mathbb{Q}(x) = \frac{1+v\mathbb{P}(x)}{\sum_{y\in\mathcal{X}} 1+v\mathbb{P}(y)}.$$

Multiple Memory Constrained Oblivious Guessors

Theorem

v oblivious memory constrained attackers wanting to minimize $\mathbb{E}[G_v]$ should generate i.i.d. guesses from

$$\mathbb{Q}(x) \ \propto \ \left[1 - f^{-1}(\mathbb{P}(x))\right].$$

For a distribution $\mathbb{P}$ for which the maximum probability is much smaller than one, we have

$$z = f(u) = (1 - u^v)^2/(vu^{v-1}) \approx (1 - 2u)/v$$

giving $f^{-1}(z) \approx (1 - vz)/2$ resulting in the fast approximation

$$\mathbb{Q}(x) = \frac{1 + v\mathbb{P}(x)}{\sum_{y \in \mathcal{X}} 1 + v\mathbb{P}(y)}.$$

Conclusions

- Our results continue work on information theoretic problems in the context of guessing and prediction–with applications in the setting of security.
- We have provided an alternative but exact operational definition of Rényi entropy in terms of oblivious guessing.
- We have generalized the guessing framework to multiple guessors, in the regime where communication between guessors is expensive or undesirable, such as P2P networks
- Thank you for listening

Conclusions

- Our results continue work on information theoretic problems in the context of guessing and prediction–with applications in the setting of security.
- We have provided an alternative but exact operational definition of Rényi entropy in terms of oblivious guessing.
- We have generalized the guessing framework to multiple guessors, in the regime where communication between guessors is expensive or undesirable, such as P2P networks
- Thank you for listening

Conclusions

- Our results continue work on information theoretic problems in the context of guessing and prediction–with applications in the setting of security.
- We have provided an alternative but exact operational definition of Rényi entropy in terms of oblivious guessing.
- We have generalized the guessing framework to multiple guessors, in the regime where communication between guessors is expensive or undesirable, such as P2P networks
- Thank you for listening

Conclusions

- Our results continue work on information theoretic problems in the context of guessing and prediction–with applications in the setting of security.
- We have provided an alternative but exact operational definition of Rényi entropy in terms of oblivious guessing.
- We have generalized the guessing framework to multiple guessors, in the regime where communication between guessors is expensive or undesirable, such as P2P networks
- Thank you for listening

References

- E. Arikan; An Inequality on Guessing and Its Application to Sequential Decoding, *IEEE Transactions on Information Theory*, 42(1):99-105, 1996.
- E. Arikan and N. Merhav; Guessing subject to distortion, *IEEE Transactions on Information Theory*, 44(3):1041-1056, 1998.
- E. Arikan and N. Merhav; Joint Source-channel Coding and Guessing with Application to Sequential Decoding, *IEEE Transactions on Information Theory*, 44(5):1756-1769, 1998.
- S. Boztaş; Comments on 'An Inequality on Guessing and Its Application to Sequential Decoding', *IEEE Transactions Information Theory*, 43(6):2062-2063, 1997.
- S.S. Dragomir and S. Boztaş; Some Estimates of the Average Number of Guesses to Determine a Random Variable, *Proc. IEEE International Symposium on Information Theory*, 1997.

References (cont'd)

- S.S. Dragomir and S. Boztaş; Estimation of Arithmetic Means and Their Applications in Guessing Theory, *Mathematical and Computer Modelling*, 28(10):31-43, 1998.
- J. L. Massey; Guessing and entropy, *Proc. 1994 IEEE International Symposium on Information Theory*, p. 204, 1994.
- D. Malone, W.G. Sullivan; Guesswork and entropy, *IEEE Transactions Information Theory*, 50(3):525- 526, 2004.
- M. Feder and N. Merhav; Relations between entropy and Error Probability, *IEEE Transactions on Information Theory* 40(1):259-266, 1994.
- N. Merhav and E. Arikan; The Shannon Cipher System with a Guessing Wiretapper, it IEEE Transactions on Information Theory, 45(6):1860-1866, 1999.

References (cont'd)

- N. Merhav, R.M. Roth, E. Arikan; Hierarchical guessing with a fidelity criterion, *IEEE Transactions Information Theory*, 45(1):330-337, 1999.
- C.-E. Pfister, W.G. Sullivan; Rényi Entropy, Guesswork Moments, and Large Deviations, *IEEE Transactions on Information Theory*, 50(11):2794, 2004.
- J. O. Pliam; On the incomparability of Entropy and Marginal Guesswork in Brute-force Attacks, Proc. INDOCRYPT 2000, *Lecture Notes in Computer Science* 1977:67–79, 2000.
- R. Sundaresan; Guessing Under Source Uncertainty, *IEEE Transactions on Information Theory* 53(1): 269 - 287, 2007.
- M. K. Hanawal and R. Sundaresan; Randomised Attacks on Passwords, *Technical Report TR-PME-2010-11*, and R. Sundaresan; Guessing and Compression Subject to Distortion, *Technical Report TR-PME-2010-12*; Dept. ECE, Indian Institute of Science. http://www.pal.ece.iisc.ernet.in/PAM/docs/techreports/tech_rep10/