Equiangular Tight Frame Fingerprinting Codes

Dustin G. Mixon

Program in Applied and Computational Mathematics
Princeton University

May 25, 2011

Joint work with:
- Christopher Quinn, Negar Kiyavash (Urbana-Champaign)
- Matthew Fickus (Air Force Institute of Technology)
Introduction to digital fingerprints

- Mark an $N$-dimensional signal $s$ and issue to $M > N$ users
- The $m$th user is given $s + f_m$
- Users $\mathcal{K}$ forge the host signal:
  \[ y = \sum_{k \in \mathcal{K}} \alpha_k (s + f_k) + \epsilon \]
- Goal: Identify the culprits
To identify culprits, isolate the combination of fingerprints:

$y - s = \sum_{k \in \mathcal{K}} \alpha_k f_k + \epsilon = F\alpha + \epsilon$

We want to recover the support of $\alpha$ given $y - s = F\alpha + \epsilon$

In the noiseless case, we have

$y - s = F\alpha$

Compressed sensing recovers $\alpha$ by assuming support sparsity

If we assume $|\mathcal{K}|$ is small, we can use CS to find $\mathcal{K}$
Introduction to compressed sensing

**Definition**

We say $F$ satisfies the $(K, \delta)$-restricted isometry property (RIP) if for every $K$-sparse vector $x$,

$$(1 - \delta)\|x\|_2^2 \leq \|Fx\|_2^2 \leq (1 + \delta)\|x\|_2^2.$$ 

**Theorem (Candès-Tao, 2005)**

Suppose $F$ is $(2K, \delta)$-RIP for some $\delta < \sqrt{2} - 1$. Then for every $K$-sparse vector $x$,

$$x = \arg \min \|\hat{x}\|_1 \text{ subject to } \hat{x} \in F^{-1}(Fx).$$

- Moral: If $F$ is RIP, we can find $\alpha$ in the noiseless case using linear programming.
What about the noise?

- If $\epsilon$ is small, linear programming can still recover $\alpha$
- Focused detection with RIP fingerprints is resilient to noise in the equal-weights case:

$$G_m^{(K)} := \left\{ \frac{1}{|K|} \sum_{k \in K} f_k : m \in K \subseteq \{1, \ldots, M\}, \ |K| \leq K \right\}$$

$$-G_m^{(K)} := \left\{ \frac{1}{|K|} \sum_{k \in K} f_k : m \not\in K \subseteq \{1, \ldots, M\}, \ |K| \leq K \right\}$$

**Theorem (Mixon-Quinn-Kiyavash-Fickus, 2010)**

Suppose fingerprints $F = [f_1, \ldots, f_M]$ are $(K, \delta)$-RIP. Then

$$\text{dist}(G_m^{(K)}, -G_m^{(K)}) \geq \sqrt{\frac{1-\delta}{K(K-1)}}.$$
Gaussian random entries can give RIP with high probability, but checking RIP is NP-hard

Theorem (Gershgorin, 1931)
For each eigenvalue $\lambda$ of a $K \times K$ matrix $[a_{ij}]$, there is an $i \in \{1, \ldots, K\}$ such that $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$.

Take $F = [f_1, \ldots, f_M]$ to have unit-norm columns
Define worst-case coherence $\mu := \max_{i \neq j} |\langle f_i, f_j \rangle|$
For each $K$, the smallest $\delta$ for which $F$ is $(K, \delta)$-RIP is

$$\delta_{\min} = \max_{\mathcal{K} \subseteq \{1, \ldots, M\}} \left\| F_{\mathcal{K}}^* F_{\mathcal{K}} - I_K \right\|_2 \leq (K - 1)\mu$$

Therefore, $F$ is $(K, \delta)$-RIP for $\delta \geq (K - 1)\mu$
What sort of fingerprints are RIP?

**Theorem (Welch, 1974)**

For any $N \times M$ matrix with unit-norm columns, $$\mu \geq \sqrt{\frac{M-N}{N(M-1)}}.$$ 

- Equiangular tight frames (ETFs) achieve the Welch bound.
- ETFs are $(K, \delta)$-RIP for $$\delta^2 \geq \frac{(K-1)^2(M-N)}{N(M-1)}.$$ 

- ETFs are state-of-the-art deterministic RIP constructions.
- Various methods construct ETFs (e.g., google Steiner ETFs).
- ETFs appear particularly well-suited as fingerprinting codes.

D.G. Mixon, C. Quinn, N. Kiyavash, M. Fickus

Equiangular Tight Frame Fingerprinting Codes
Denote $z := \sum_{k \in K} \alpha_k f_k + \epsilon$

Test statistic: $T_m(z) := \frac{1}{\gamma^2} \langle z, f_m \rangle$

Given a threshold $\tau$, we decide $m$ is guilty if $T_m(z) \geq \tau$

False-positive probability:

$$P_I(F, m, \tau, K, \alpha) := \text{Prob}[T_m(z) \geq \tau \mid m \text{ not guilty}]$$

False-negative probability:

$$P_{II}(F, m, \tau, K, \alpha) := \text{Prob}[T_m(z) < \tau \mid m \text{ guilty}]$$
Analysis of focused detection

- Worst-case false-positive probability:
  \[ P_I(F, \tau, \alpha) := \max_{\mathcal{K}} \max_{m \notin \mathcal{K}} P_I(F, m, \tau, \mathcal{K}, \alpha) \]

- We wish to catch at least one colluder (the most vulnerable)

- Worst-case false-negative probability:
  \[ P_{II}(F, \tau, \alpha) := \max_{\mathcal{K}} \min_{m \in \mathcal{K}} P_{II}(F, m, \tau, \mathcal{K}, \alpha) \]

**Theorem (Mixon-Quinn-Kiyavash-Fickus, 2010)**

Suppose fingerprints \( F = [f_1, \ldots, f_M] \) form an ETF. Then

\[
P_I(F, \tau, \alpha) \leq Q\left[ \frac{\gamma}{\sigma} (\tau - \mu) \right],
\]

\[
P_{II}(F, \tau, \alpha) \leq Q\left[ \frac{\gamma}{\sigma} \left( (1 + \mu) \max_{k \in \mathcal{K}} \alpha_k - \mu \right) - \tau \right],
\]

where \( Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-u^2/2} du. \)
Analysis of focused detection

- $P_1(F, \tau, \alpha) \leq Q\left[\frac{\gamma}{\sigma}(\tau - \mu)\right]$
  - Upper bound is independent of $\alpha$
  - Interpretation: Colluders cannot intentionally frame an innocent user

- $P_\Pi(F, \tau, \alpha) \leq Q\left[\frac{\gamma}{\sigma}\left((1 + \mu)\max_{k \in \mathcal{K}} \alpha_k - \mu\right) - \tau\right]$
  - Upper bound is maximized when $\alpha_k = \frac{1}{|\mathcal{K}|}$ for every $k \in \mathcal{K}$
  - Interpretation: Colluders have best chance of not being detected with equal weights
Compressed sensing ideas (like RIP) are helpful for digital fingerprint design and identifying culprits
ETFs are state-of-the-art deterministic RIP matrices
Focused detection works well with ETF fingerprints
Future work: Evaluate other methods to identify culprits with ETF fingerprints
Dustin G. Mixon
dmixon@princeton.edu