Digital Walkie-Talkie Identification scheme based on Sparse Representation with Multiple features

Presenter: M.S. Student / Kiwon Yang
Advisor: Professor / Heung-No Lee

GIST, Dept. of Electrical Engineering and Computer Science
INFONET Lab.

Gwangju Institute of Science and Technology
Contents

I. Introduction
II. Contribution
III. Research Process
IV. Results
V. Comparison
VI. Appendix
VII. Reference
I. Introduction

- **Motivation**
  - For the efficient support in the electronic warfare, the ability of exact detection and analysis on the enemy’s transmitter is necessary.
  - In the Internet of Things (IoT) network, the technique for identifying the access of the counterfeit transmitter is needed.

- Identification of the radio transmitters using the transmitted signals is called **RF (Radio frequency) fingerprinting**.
I. Introduction

- Proposed system

- A **feature** is a sample vector cultivated from the transmitted RF signals and bears unique information about the pertinent device.

- **Goal**
  - We want to check if the performance will be increased or not once **multiple features** – rising transient feature, falling transient feature, and sync feature – are used **simultaneously**.
I. Introduction

- The feature is occurred by
  - Element characteristic
  - A part design such as filter, amplifier etc.
  - PCB material, soldering etc.

- The feature types

I. Introduction

- Related works
  - Merchant et al [3]
    - Convolutional neural network
  - Peng et al [4]
    - Differential constellation trace, carrier frequency offset, and 2 features of the error on I/Q domain
  - Patel et al [5]
    - Random forest and AdaBoost
II. Contribution

- The **combination** of **rising transient feature**, **falling transient feature**, and **sync feature** has not been used in the previous studies.
  - Falling transient feature has not been used.
  - We show that the performance of the proposed scheme is improved when more feature is included.
  - There are no experiments on RF fingerprinting with multiple features based on sparse representation-based classification algorithm (SRC).
III. Research Process – Signal acquisition

- The digital walkie-talkie models

Two models follow DMR standard

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer</td>
<td>MOTOROLA</td>
<td>HYTERA</td>
</tr>
<tr>
<td>Frequency</td>
<td>UHF (CH1 : 423.1875MHz)</td>
<td>UHF (CH1 : 423.1875MHz)</td>
</tr>
<tr>
<td># of devices</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
III. Research Process – Signal acquisition

- Digital Mobile Radio standard [6]
  - 2-slot Time-division multiple access (TDMA) method
  - 4 level frequency shift-keying modulation

![Diagram showing signal acquisition process with data, sync signal, and steady-state signal](image-url)
III. Research Process – Signal acquisition

- Procedure for signal acquisition
  - In LOS environment, the receiver gets the transmitted signal.
III. Research Process – Feature extraction

- Threshold method to extract interest signals

![Amplitude vs Samples plots for Rising transient signal, Falling transient signal, and Sync signal]
III. Research Process – Feature extraction

- Main lobe extraction
  - Since main lobe occupies most of the energy of each signal part, the main lobe is used as a feature.

![Main lobe of rising transient signal](image1)
![Main lobe of falling transient signal](image2)
![Main lobe of sync signal](image3)

---

**Main lobe of rising transient signal**

**Main lobe of falling transient signal**

**Main lobe of sync signal**
III. Research Process – Feature concatenation

- The extracted features – rising transient feature, falling transient feature, and sync feature – are **concatenated**.

- In the system,

\[
u = As,
\]

the concatenated features for training data are arranged to the columns of \( A \) and the feature for test data is put into \( u \).
III. Research Process – SRC

- Sparse representation-based classification scheme
  - In the underdetermined system $\mathbf{y} = \mathbf{D}s$, $s$ has infinite cases of solutions and SRC finds a sparse solution $s$.
  - The condition to find the sparse solution $s$ is sensitive to the mutual correlation between columns of $\mathbf{D}$.

\[
\text{Class} = \arg \min_{l \in \{1, \ldots, L\}} \|\mathbf{y} - \mathbf{D}^{(l)}s^{(l)}\|_2
\]
III. Research Process – SRC & Test

- Sparse representation-based classification scheme
  - Principal component analysis (PCA) removes correlations among columns of $A$.
  - $u = As$ is changed to $y = Ds$ by PCA.
  - The sparse solution $s$ is obtained by basis pursuit algorithm
    $$\min_s ||s||_1 \text{ subject to } y = Ds.$$  
  - The class of test data is output from SRC

- Test
  - We used 5 cross validation technique.
  - Fifty data were captured per a digital walkie-talkie.
IV. Results

- When **additional feature** is included, **the performance of SRC is improved**.
- The accuracy recorded **98.75%**.
- Falling transient signal could have unique information for RF fingerprinting.

Accuracy rate of the proposed method

<table>
<thead>
<tr>
<th></th>
<th>4 BD-358</th>
<th>4 SL1M</th>
<th>4 BD-358</th>
<th>4 SL1M</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy rate (Minimum number of PC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TR(R)</td>
<td>88% (24)</td>
<td>82% (48)</td>
<td>90.5% (45)</td>
<td></td>
</tr>
<tr>
<td>TR(F)</td>
<td>87.5% (45)</td>
<td>90% (12)</td>
<td>92.25% (13)</td>
<td></td>
</tr>
<tr>
<td>TR(R + F)</td>
<td>93% (49)</td>
<td>92% (20)</td>
<td>95.5% (63)</td>
<td></td>
</tr>
<tr>
<td>Sync</td>
<td>99% (45)</td>
<td>83.5% (22)</td>
<td>93.75% (86)</td>
<td></td>
</tr>
<tr>
<td>TR(R + F) + Sync</td>
<td><strong>99% (44)</strong></td>
<td><strong>98.5% (22)</strong></td>
<td><strong>98.75% (21)</strong></td>
<td></td>
</tr>
</tbody>
</table>

R: Rising, F: Falling, PC: Principal components
IV. Results

- The cluster on each class is distinctly formed when the concatenated feature is used.

SL1M transient (rising+falling) features

SL1M sync features

SL1M transient (rising+falling) + sync features
IV. Results

- It is noticeable that **the highest accuracy rate** is recorded even though **the less number of training data** is used relatively.

- The comparison experiment is necessary for more accurate comparison.

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of Devices</th>
<th>Experiment condition</th>
<th>Accuracy rate</th>
<th>Number of training data per a device</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The proposed method</strong></td>
<td>8 (Digital walkie-talkies)</td>
<td>1m LOS</td>
<td>98.75%</td>
<td>40</td>
</tr>
<tr>
<td>Patel et al. [3]</td>
<td>4 (Zigbee devices)</td>
<td>12 dB</td>
<td>Higher than 90%</td>
<td>1500</td>
</tr>
<tr>
<td>Peng et al. [4]</td>
<td>54 (Zigbee devices)</td>
<td>1-3m LOS</td>
<td>96%</td>
<td>1 (template feature)</td>
</tr>
<tr>
<td>Merchant et al. [5]</td>
<td>7 (Zigbee devices)</td>
<td>28 dB</td>
<td>92.29%</td>
<td>900</td>
</tr>
</tbody>
</table>
V. Conclusion

- We proposed the RF fingerprinting scheme based on SRC with multiple features.
- As a feature, the main lobes of rising transient signal, falling transient signal, and sync signal were used simultaneously.
- When many features were used as concatenation, the accuracy rate was increased.
- The accuracy rate of the proposed method recorded 98.75%.
- As a future work, we need to study on RF fingerprinting scheme based on SRC with the various features besides the used features.
- The paper on this study is under revision.
Thank you
VI. Appendix

- Deep learning for RF device fingerprinting [3]
  - They used the convolutional neural network (CNN) for the RF fingerprinting.
  - They collected 7,000 data from 7 devices.
  - Each full dataset of 7,000 transmissions was randomly partitioned into 80% training data, 10% validation data, and 10% testing data.
  - The overall correct identification rate is 92.29%
VI. Appendix

- Hybrid RF fingerprint extraction and device classification scheme [4]
  - They used 4 features simultaneously – differential constellation trace figure, carrier frequency offset, and 2 features of the error on I/Q domain.
  - The experiment is performed on the total 54 Zigbee devices.
  - They did the experiment on the 4 environment.
    - Line of sight / Non line of sight
    - Line of sight after 18 month with same receiver and different receivers
VI. Appendix

- Improving zigbee device network authentication using ensemble decision tree classifiers [5]
  - The used RF DNA features contain information on variance, skewness, and kurtosis, within a preamble response.
  - They showed the result of ‘Random Forest’ and ‘Multi-class AdaBoost’ for RF fingerprinting.
  - The top-ranked 25 variables selected by Variable Importance (VI) metric built in Random Forest classifier on 4 zigbee devices are used.
VI. Appendix

Principal components analysis [7]

- PCA is a method to project the original data onto the new space on the variance.
- Let $\mathbf{u} = \mathbf{A}s$, where $\mathbf{A}$ is the training data matrix and $\mathbf{u}$ is the test data.
- A covariance matrix of $\mathbf{A}$ is eigen-decomposed as,
  \[
  (\mathbf{A} - \mathbf{m}1)(\mathbf{A} - \mathbf{m}1)^T = \mathbf{W}\Lambda\mathbf{W}^T
  \]
  where $\mathbf{m} = \frac{1}{N}\sum_{n=1}^{N} n$ th columns of $\mathbf{A}$, $\mathbf{1}_N = [1 \ 1 \ \cdots \ 1]$.
- The eigen-vectors of the covariance matrix are orthonormal.
- The eigen-value matrix $\Lambda$ is proportional to the variance of $\mathbf{A}$,
  \[
  \mathbf{W}^T(\mathbf{A} - \mathbf{m}1)(\mathbf{A} - \mathbf{m}1)^T\mathbf{W} = \Lambda.
  \]
- Let the eigen-values $\lambda_1, \lambda_2, \ldots \lambda_n$ of the eigen-value matrix $\Lambda$ be rearranged in order of the sizes.
- Let the eigen-vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_n$ of the eigen-vector matrix $\mathbf{W}$ be also rearranged by the eigen-values.
VI. Appendix

- Principal components analysis [7]
  - Since the eigen-vectors of the covariance matrix are orthonormal and the eigen-value matrix \( \Lambda \) is proportional to the variance of \( A \), the eigen-vectors can be basis for the creating the new space on the variance of \( A \).
  - The training data matrix \( A \) and the test data \( u \) is transformed to the new space by \( D = W^T(A - m1) \) and \( y = W^T(u - m) \).
  - PCA removes correlations among columns of \( A \).
  - Also, PCA can remove the size of the columns of \( A \).
VI. Appendix

- $L_p$ norms [8]
  \[ ||x||_p = \left( \sum_i |x_i|^p \right)^{\frac{1}{p}} \]

- Uniqueness of sparse solution ($L_1$) [8]
  - Suppose $y = Ds_0$ with
    \[ ||s_0||_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu(D)} \right), \]
  where $\mu(D) = \max_{1 \leq k, j \leq m, k \neq j} \frac{|a_k^Ta_j|}{||a_k||_2 \cdot ||a_j||_2}$
  Then $s_0$ is the unique optimal solution to
  Minimize $||s||_1$ subject to $y = Ds$.

- If the function $f$ has a second derivative that is non-negative (positive) over an interval, the function is convex (strictly convex) over that interval. [9]
VI. Appendix

- $L_p$ norms level sets

- Basis pursuit algorithm [10]
  - The mathematical optimization problem of the form
    $$\min_{s} ||s||_1 \text{ subject to } y = Ds.$$  
    - To solve the problem, ‘Primal-Dual Barrier method’ is used.
VI. Appendix

- **Primal-Dual Barrier method** [11]
  - A certain class of algorithms that solve linear and nonlinear convex optimization problems.
  - Consider the dual pair for Linear programming problem
    \[
    \min c^T x \quad s.t. \quad Ax = b, \quad x \geq 0, \quad \min b^T \lambda \quad s.t. \quad A^T \lambda + s = c, \quad s \geq 0
    \]
  - The Karush-Kuhn-Tucker conditions for both equation are
    \[
    \begin{align*}
    A^T \lambda + s & = c \\
    Ax & = b \\
    x & \geq 0 \\
    s & \geq 0 \\
    x^{(i)} s^{(i)} & = 0, \quad 1 \leq i \leq n
    \end{align*}
    \]
  - Let \( s = (s^{(1)}, s^{(2)}, \ldots, s^{(n)}) \), \( S = \text{diag}(s) \), and \( e = (1,1,\ldots,1) \). We can rewrite the constraints into
    \[
    \tilde{F}(x, \lambda, s) = \begin{bmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe \end{bmatrix} = 0
    \]
VI. Appendix

- Primal-Dual Barrier method [11]
  
  - We relax the last constraint $x^{(i)}_S^{(i)} = 0$ to $x^{(i)}_S^{(i)} = \mu$ and obtain
    $$F(x, \lambda, s) = \begin{bmatrix} A^T\lambda + s - c \\ Ax - b \\ XSe - \mu e \end{bmatrix} = 0$$

  - The Jacobian will be
    $$J = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}$$

  and the Newton’s method read
    $$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_\lambda \\ d_s \end{bmatrix} = \begin{bmatrix} -A^T\lambda - s + c \\ b - Ax \\ -XSe + \mu e \end{bmatrix}$$

  - Solve
    $$\min B(x_k, \mu_k) = c^T x - \mu \sum_{i=1}^{n} \log x_i, \mu > 0$$
VI. Appendix

- Primal-Dual Barrier method [11]

Algorithm 2 Primal-Dual Newton Barrier Method for LP

1: \( \mu_0 \leftarrow 1, \rho \in (0, 1) \)
2: Generate \((x_0, \lambda_0, s_0), \text{s.t. } x_0 > 0, s_0 > 0\)
3: for \( k = 1, 2, 3, \ldots \) do
4: \[ \mu_k \leftarrow \rho \mu_{k-1} \]
5: Solve
\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S_{k-1} & 0 & X_{k-1}
\end{bmatrix}
\begin{bmatrix}
d_X \\
d_\lambda \\
d_s
\end{bmatrix}
= -
\begin{bmatrix}
A^T \lambda_{k-1} + s_{k-1} - c \\
Ax_{k-1} - b \\
X_{k-1}S_{k-1}e - \mu_k e
\end{bmatrix}
\] (21)
6: Solve
\[
\min_{\alpha > 0} B(x_k, \mu_k) \]
\[\text{s.t. } (x_k, \lambda_k, s_k) = (x_{k-1}, \lambda_{k-1}, s_{k-1}) + \alpha (d_X, d_\lambda, d_s)\]
7: \((x_k, \lambda_k, s_k) \leftarrow (x_{k-1}, \lambda_{k-1}, s_{k-1}) + \alpha (d_X, \lambda_k, s_k)\)
8: Check stop criterion.
VII. Reference


VII. Reference


