Understanding and Modeling of WiFi Signal Based Human Activity Recognition

Wei Wang†, Alex X. Liu†‡, Muhammad Shahzad‡, Kang Ling†, Sanglu Lu†

†Nanjing University, ‡Michigan State University

September 8, 2015
Motivation

- WiFi signals are available almost everywhere and they are able to monitor surrounding activities.
Problem Statement

WiFi based Activity Recognition

- Using commercial WiFi devices to recognize human activities.

Advantages

✓ Work in dark
✓ Better coverage
✓ Less intrusive to user privacy
✓ No need to wear sensors
WiFi based Activity Recognition

- Using commercial WiFi devices to recognize human activities.

Advantages
- ✓ Work in dark
- ✓ Better coverage
- ✓ Less intrusive to user privacy
- ✓ No need to wear sensors
WiFi based Activity Recognition

- Using commercial WiFi devices to recognize human activities.

Advantages

✓ Work in dark
✓ Better coverage
✓ Less intrusive to user privacy
✓ No need to wear sensors
Problem Statement

WiFi based Activity Recognition

- Using commercial WiFi devices to recognize human activities.

Advantages

✓ Work in dark
✓ Better coverage
✓ Less intrusive to user privacy
✓ No need to wear sensors
Problem Statement

WiFi based Activity Recognition

- Using commercial WiFi devices to recognize human activities.

Advantages

✓ Work in dark
✓ Better coverage
✓ Less intrusive to user privacy
✓ No need to wear sensors
OFDM PHY Basics

- Guard Band
- Sub-carrier spacing
- DC Carrier
- Sub-carrier
- Effective Channel Band-width
- Channel Band-width

802.11a OFDM Physical Parameters

- 52 subcarriers
- 48 Data, 4 Pilot (BPSK), 1 Null
- OBW 16.6 MHz
- BW 20 MHz
- One Subcarrier = 1 constellation point
- 1 OFDM symbol = 52 subcarriers
- 1 OFDM Burst = one or more OFDM symbols
Commercial Off-the-Shelf Cards provide 30 sub-carriers CSI measurement taken every frame.

Every \( h \) entry - CFR (Channel Frequency Response) - amplitude and phase information of each sub-carrier for an antenna stream at time \( t \).
Challenges

- Measurement from commercial devices are noisy and have unpredictable carrier frequency offsets
- Needs robust and accurate models to extract useful information from measurements

Noise - electromagnetic interference; internal state changes (power and rate adaptation)

CFO - Channel Frequency Offset - 802.11n 5GHz channel - sub-carrier frequency can drift by up-to 100 kHz from central frequency
Key observations

- Multipaths contain both static component and dynamic component
- Each path has different phase
- Phases determine the amplitude of the combined signal
Understanding Multipath

Motivation  Modeling  Design  Experiments  Conclusions

Sender

Receiver

Wall

Reflected by body

Reflected by wall

LoS path

$d_k(0)$

Static component

Dynamic Component

Combined

$\text{LoS path}$

$\text{I}$

$\text{Q}$
Understanding Multipath

Motivation  Modeling  Design  Experiments  Conclusions

Sender

LoS path

d_k(t)

Reflected by wall

Reflected by body

Wall

Receiver

Dynamic Component

Static component

Combined

Combined component

LoS path

Reflected by body

Qui

Combined component

Dynamic Component
Understanding Multipath

Motivation  Modeling  Design  Experiments  Conclusions

Sender  Receiver  Wall  Reflected by body  Reflected by wall  LoS path

\[ d_k(t) \]

Combined  Dynamic Component  Static component

\[ I \quad Q \]
Understanding Multipath

Interpreting CSI amplitude

- Phases of paths are determined by path length
- Path length change of one wavelength gives phase change of $2\pi$
- **Frequency of amplitude change can be converted to movement speed**
How accurate is it?

- Wave length → 5 ~ 6 cm in 5 GHz band
- Steel plate of diameter 30 cm moving along a straight line

CSI Waveform for a movement with 0.8 m path length change

Measurement of path length change
- Ground truth determined using laser rangefinder
- Hilbert Transform used to find phase change and then multiply by wavelength
  = path length change = 1/2 mov. distance
How accurate is it?

- Wave length → $5 \sim 6\text{cm}$ in 5 GHz band
- Steel plate of diameter 30 cm moving along a straight line

CSI amplitude changes are close to sinusoids

Average distance measurement error of 2.86 cm
CSI-Speed Model

How robust is it?

- CFR amplitude - linear combination of the reflected paths and the speeds of path length change
  - Linear combination of multipath do not change frequency
  - Robust over different multipath conditions and movement directions (88% accuracy with 0-0.61-1 separation)

![Speed distribution of different activities in different environments](image)

Speed distribution of different activities in different environments
Activities are characterized by

- Movement speeds
- Change in movement speeds
- Speeds of different body components
CSI-Activity Model

- Denoise CSI Values (PCA based)
- Use time-frequency analysis to extract features (DWT)
- Use HMM to characterize the state transitions of movements

Walking

Falling

Sitting down
CSI-Activity Model

- Build one HMM model for each activity
- Determine states based on observations in waveform patterns
- State durations and relationships are captured by transition probabilities
System Architecture

- CSI measurement collection
- Noise reduction
- Activity data collection
- Online monitoring
  - Activity detection and segmenting
  - Feature extraction
  - HMM based activity recognition
  - Monitoring records
- Model generation
  - HMM training
  - HMM Model
Data Collection

$N \times M \times 30$ CSI streams

30 subcarriers

$N \times M \times 30$ CSI streams
Noise Reduction

Correlation of CSI on different subcarriers

- Subcarriers only differ slightly in wavelength
- Subcarriers have the same set of paths, with different phases
- Principal Component Analysis (PCA) to filter noises

312.5kHz

Wave length = 5.150214 cm

Wave length = 5.149662 cm

Frequency
Correlation in CSI Streams

Correlation of CSI on different subcarriers

- Noise in principal component 1 is discarded, next 5 are kept

Phase changes by $2\pi$

Noises present in all streams

- CSI "peaks" are red, "valleys" are blue
Noise Reduction

Combines $N \times M \times 30$ subcarriers using PCA to detect time-varying correlations in signal

Original

Low-pass filter

2nd PCA Component
Real-time Recognition

- Activity detection
  - Use both the signal variance and correlation to detect presence of activities

- Feature extraction
  - Time-frequency analysis (DWT)

- HMM model building
  - Eight activities
    - Walking, running, falling, brushing teeth, sitting down, opening refrigerator, pushing, boxing
  - More than 1,400 samples from 25 persons as the training set
Evaluation Setup

- Commercial hardware with no modification
  - Transmitter: NetGEAR JR6100 Wireless Router
  - Receiver: Thinkpad X200 with Intel 5300 NIC
- A single communicating pair is enough to monitor 450 $m^2$ open area
- Measurement on UDP packets sent between the pair
- Sampling rate 2,500 samples per second
## Evaluation Results

### Activity recognized

<table>
<thead>
<tr>
<th>True activity</th>
<th>R</th>
<th>W</th>
<th>S</th>
<th>O</th>
<th>F</th>
<th>B</th>
<th>P</th>
<th>T</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Walking</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Sitting</td>
<td>0.000</td>
<td>0.000</td>
<td>0.947</td>
<td>0.030</td>
<td>0.011</td>
<td>0.000</td>
<td>0.012</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Opening</td>
<td>0.000</td>
<td>0.005</td>
<td>0.150</td>
<td>0.803</td>
<td>0.042</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Falling</td>
<td>0.000</td>
<td>0.010</td>
<td>0.041</td>
<td>0.010</td>
<td>0.939</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Boxing</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Pushing</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Brushing</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Empty</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

- Ten-fold validation accuracy: **96.5%**
- Detects human movements at **14** meters
- Real-time recognition on laptops
- Packet sending rate / CSI sampling rate can be as low as **800 frames per second**
Motivation  Modeling  Design  Experiments  Conclusions

Evaluation on Robustness

- Models are robust to environment changes
- Train once, apply to different scenarios
- Training use database collected in lab with different users
- Test in with users not in the training set
  - Open lobby
  - Apartment (NLOS)
  - Small office

![Diagram showing experimental locations and layouts](image-url)
Consistent performance in unknown environments, with more than 80% average accuracy
Conclusions

- CSI measurements contains fine-grained movement informations

- CSI-Speed model
  quantifies the correlation between CSI value dynamics and human movement speeds

- CSI-Activity model
  quantifies the correlation between the movement speeds of different human body parts and a specific human activity

- Our models are robust to environment changes
Thank you!

Questions?