PANDAA:
Physical Arrangement Detection of Networked Devices through Ambient-Sound Awareness

Zheng Sun (CMU)                        Aveek Purohit (CMU)
Kaifei Chen (USTC)                    Shijia Pan (USTC)
Trevor Pering (independent)       Pei Zhang (CMU)

Ubicomp full paper, Sept 21st 2011 (Best Demo Award, too)
The Problem
Potential Applications
“Swipe-and-send”
Potential Applications
“Swipe-and-send”
Potential Applications
“Swipe-and-send”
Potential Applications
“Swipe-and-send”
Potential Applications
“Swipe-and-send”
Meetings - Intuitive Content Sharing
Requirements

• Must be accurate (sub-meter).

• Work on off-the-shelf devices, minimal requirement of specialized hardware.

• Non-intrusive, automated operation and maintenance.
## Related Work (Indoor Localization)

<table>
<thead>
<tr>
<th>Method</th>
<th>Desired Sub-meter Accuracy</th>
<th>Requirement of Specialized Hardware</th>
<th>Non-intrusive</th>
</tr>
</thead>
<tbody>
<tr>
<td>WiFi signal strength range/fingerprint</td>
<td>X</td>
<td>Low</td>
<td>✓</td>
</tr>
<tr>
<td>Ultrasound-RF</td>
<td>✓</td>
<td>High</td>
<td>✓</td>
</tr>
<tr>
<td>Audible chirp ranging</td>
<td>✓</td>
<td>Low</td>
<td>X</td>
</tr>
<tr>
<td>Ambient sound ranging (PANDAA)</td>
<td>✓</td>
<td>Low</td>
<td>✓</td>
</tr>
<tr>
<td>Problem</td>
<td>Related work</td>
<td>Proposed approach</td>
<td>Evaluation</td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
</tbody>
</table>

Friday, November 18, 11
PANDAA

A microphone

Wireless connection
PANDAA
PANDAA

Indoor ambient sounds:
- a door closing
- a barking dog
- human talk
- coughs
- hand claps
- a ringing phone
- finger snaps
- ......
PANDAA

Indoor ambient sounds:

- a door closing
- a barking dog
- human talk
- coughs
- hand claps
- a ringing phone
- finger snaps
- ......
Ambient Sound Processing Pipeline

1. Pre-processing
   (described in the paper)

   - Microphone
   - Framing
     Impulsive Sound Extraction

2. Pairwise Distance Estimation
   (described in the paper)

   - TDoA Estimation
   - Two-tier TDoA Aggregation
   - Compensation of Pairwise TDoA Errors
     Arrangement Detection

3. Arrangement Detection
   (this talk)

   - Relative Device Locations (Output)

Ambient Sound (Input)

server

devices

TDoA: Time Difference of Arrival
Time difference of sound arrivals (TDoA) can be expressed as

$$\Delta t_{BA} = \frac{(d_{SB} - d_{SA})}{\text{speed of sound}}$$

computable through matching of impulsive peaks

what we want!

a constant
Estimate Distances Between Devices

Given one source $S$, we have one lower bound of $d_{AB}$

$$|d_{SB} - d_{SA}| \leq d_{AB}$$

\[\text{a lower bound of } d_{AB}\]
Successive Estimation of $d_{AB}$

Given multiple sources, we have overlapped bounds of $d_{AB}$
Successive Estimation of $d_{AB}$

Given multiple sources, we have overlapped bounds of $d_{AB}$
Successive Estimation of $d_{AB}$

Given multiple sources, we have overlapped bounds of $d_{AB}$
Successive Estimation of $d_{AB}$

Given multiple sources, we have overlapped bounds of $d_{AB}$

unknown value
estimated value

Friday, November 18, 11
Successive Estimation of $d_{AB}$

Given multiple sources, we have overlapped bounds of $d_{AB}$.

The maximal lower bound will get closer and closer to the actual $d_{AB}$.
A Problem!
Different Sound Source Locations

- $S_1$: Good
- $S_2$: Good
- $S_3$: Not so good
- $S_4$: Bad
Compensate for Pairwise Errors

- As \#devices increases, estimation accuracy can be improved
  - A sound source may be **bad** for one particular device pair, but **good** for others.

Only two devices A, B
S is not good for estimating the distance between
A and B
Compensate for Pairwise Errors

- As \#devices increases, estimation accuracy can be improved
  - A sound source may be \textbf{bad} for one particular device pair, but \textbf{good} for others.

Only two devices A, B
S is not good for estimating the distance between A and B

If we have 2 more devices in the network
S is not good for estimating the distance between A-B, but is good for A-C, B-C, and C-D
<table>
<thead>
<tr>
<th>Problem</th>
<th>Related work</th>
<th>Proposed approach</th>
<th>Evaluation</th>
<th>Discussion</th>
</tr>
</thead>
</table>

Friday, November 18, 11
Meeting-room Experiments

- 8x6m² meeting room
- Eight nodes (orange dots on the floor plan)
- 100 locations to generate ambient sound (grid intersections)
Ambient Sound Used In Experiments

95-second audio at each source location (the “grid”) using loudspeaker

<table>
<thead>
<tr>
<th>Types</th>
<th>Durations (s)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cough</td>
<td>32</td>
<td>12 coughs from 6 individuals (2 males and 4 females)</td>
</tr>
<tr>
<td>Conversation</td>
<td>21</td>
<td>Between a male and a female</td>
</tr>
<tr>
<td>Music #1</td>
<td>21</td>
<td>“Billie Jean”</td>
</tr>
<tr>
<td>Music #2</td>
<td>21</td>
<td>“The Sound of Silence”</td>
</tr>
</tbody>
</table>
Impulsive Sound Event Detection

- Averagely 1 event/cough; for other types, 1 event/sec.
- Effective to extract impulsive sound from all four sound types.
- Detection rate is high to generate sufficient events for arrangement detection.
Estimated Locations vs. Ground-truths

\[ \text{Source: 1} \quad \text{Error: 1.1789m} \]

- \( \times \): Sound sources
- \( + \): Ground truths
- \( \diamond \): Estimated locations
Estimated Locations vs. Ground-truths

- **x**: Sound sources
- **+**: Ground truths
- **◊**: Estimated locations

#Source: 2  Error: 1.041m
Estimated Locations vs. Ground-truths

- **X**: Sound sources
- **+**: Ground truths
- **◊**: Estimated locations

#Source: 6  Error: 0.29679m
Location Errors vs. #Sound Sources

- Accuracy becomes stable after 6 sources.
- Average ultimate accuracy is 0.17m.
- Naturally impulsive sound work the best.

- Fluctuate due to erroneous TDoA estimates caused by ambient noise, echos, non-line-of-sight.
Conclusions

• **Novel approach** - prove that using *ambient sound* in physical arrangement detection is possible.

• PANDAA achieves **0.17m** accuracy in the meeting-room experiments given uniformly distributed sound sources.