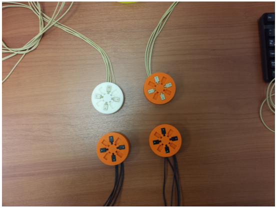
Towards Wireless Acoustic Sensor Networks for Location Estimation and Counting of Multiple Speakers in Real-life Conditions

Anastasios Alexandridis Nikolaos Stefanakis Athanasios Mouchtaris

Institute of Computer Science – Foundation for Research and Technology–Hellas
Computer Science Department – University of Crete

ICASSP 2017
March 5–9, New Orleans, USA

Wireless Acoustic Sensor Networks (WASNs)



4 $\mathcal{N}_4$

3 $\mathcal{N}_3$

Each sensor is
a microphone array

Fusion center
for localization

$\mathcal{N}_1$ 1

$\mathcal{N}_2$ 2

F

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Real-life deployment

- Microphone array positioning
  - need to be placed near walls in order not to pose restrictions on speakers' activities

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Real-life deployment

- Microphone array positioning
  - need to be placed near walls in order not to pose restrictions on speakers' activities

Real speakers' characteristics

- Directivity pattern
- Spatial volume
- Orientation

$\neq$

- Omni-directional
- Point sources

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Location estimation & counting using DOA estimates

Real-life deployment

- Microphone array positioning
  - need to be placed near walls in order not to pose restrictions on speakers' activities

Real speakers' characteristics

- Directivity pattern
- Spatial volume
- Orientation

$\neq$

- Omni-directional
- Point sources

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Location estimation & counting using DOA estimates

Real-life deployment

- Microphone array positioning
  - need to be placed near walls in orde n speakers' activities

"Reflection-aware" DOA estimation

Real speakers' characteristics

- Directivity pattern
- Spatial volume
- Orientation

$\neq$

- Omni-directional
- Point sources

Challenges in speaker localization

WASN

- Computational complexity
- Bandwidth usage
- Synchronization between nodes (microphone arrays)

Location estimation & counting using DOA estimates

Real-life deployment

- Microphone array positioning
  - need to be placed near walls in orde … n speakers' activities

"Reflection-aware" DOA estimation

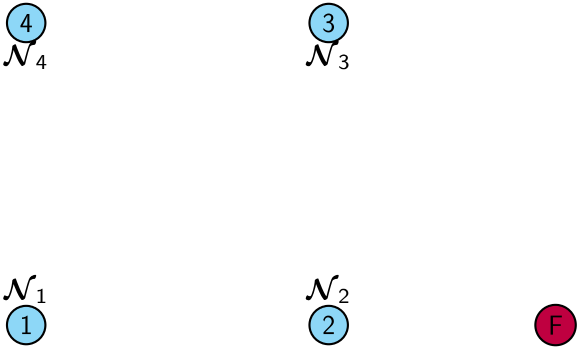
Real speakers' characteristics

- Directivity pattern
- Spatial volume
- Orientation

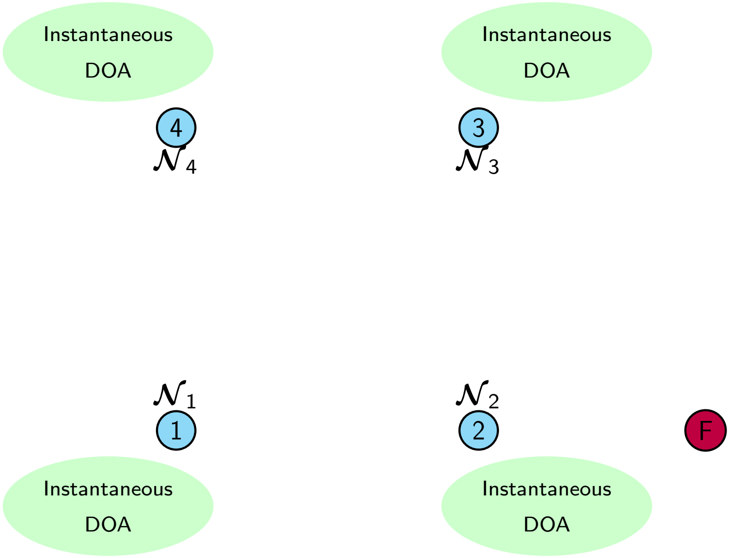
$\neq$

Real recorded dataset with real speakers

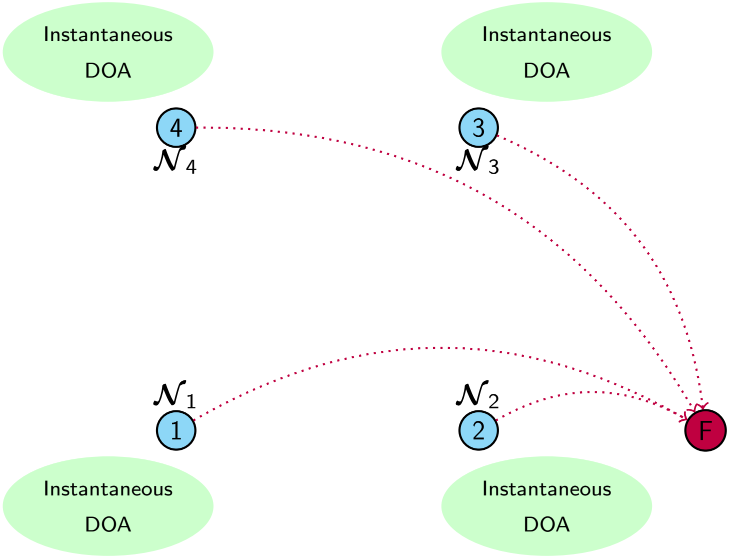
System Overview



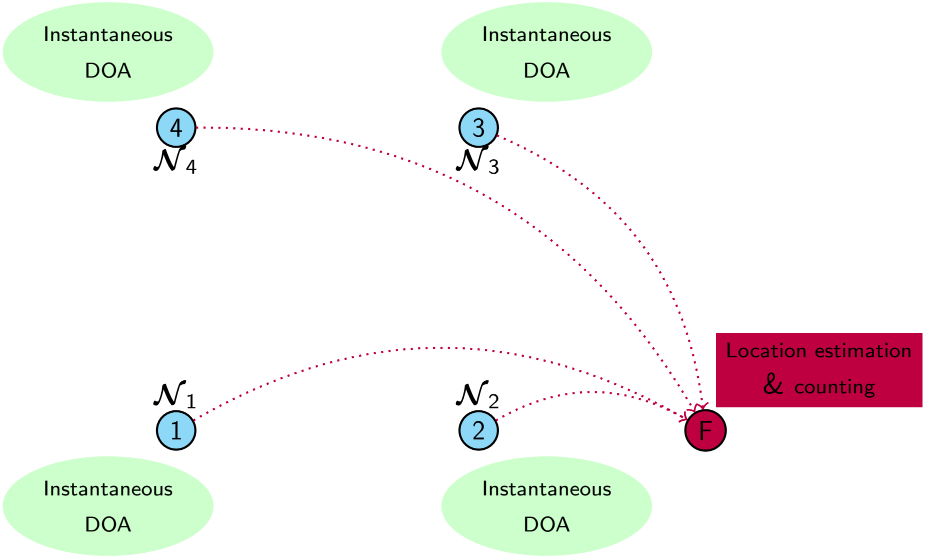
System Overview



System Overview



System Overview



“Reflection-aware” propagation model

[Stefanakis 2016]

Typical propagation model

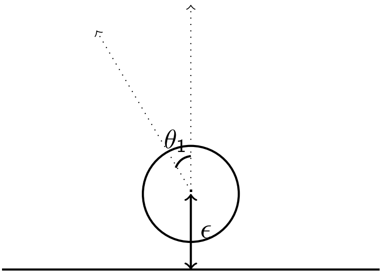
$$a_m(\omega, \theta) = e^{jkR\cos(\phi_m - \theta)}$$

"Reflection-aware” propagation model

$$a_m(\omega, \theta) = e^{jkR\cos(\phi_m - \theta)} e^{jk\epsilon\cos\theta} + h e^{jkR\cos(\phi_m - \pi + \theta)} e^{-jk\epsilon\cos\theta}$$

ϵ: distance from the wall

$h \in [0, 1]$: Image-Source reflective gain



N. Stefanakis, A. Mouchtaris, “Direction of arrival estimation in front of a reflective plane using a circular microphone array,” *European Signal Processing Conference (EUSIPCO)*, 2016

“Reflection-aware DOA estimation

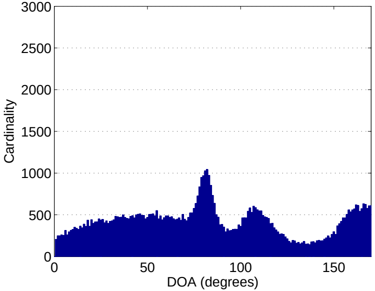
- Minimum Variance Distortionless Response (MVDR) beamformer

  $\hat{\theta}(\omega, \tau)$: the DOA where the MVDR beamformer response is maximized

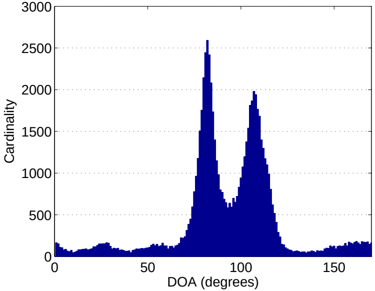
“Reflection-aware DOA estimation

- Minimum Variance Distortionless Response (MVDR) beamformer

  $\hat{\theta}(\omega, \tau)$: the DOA where the MVDR beamformer response is maximized



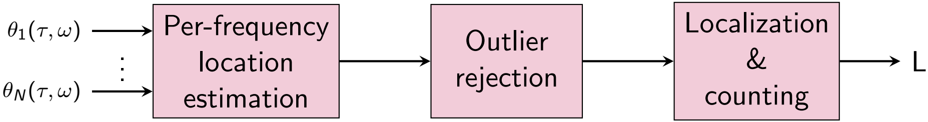
Classical model



“Reflection-aware” model

Location estimation & counting

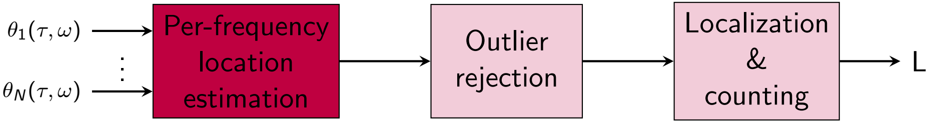
[Alexandridis 2015]

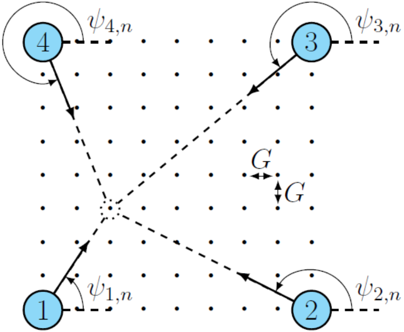


A. Alexandridis, A. Mouchtaris, "Multiple sound source location estimation and counting in a wireless acoustic sensor network," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015.

Location estimation & counting

[Alexandridis 2015]



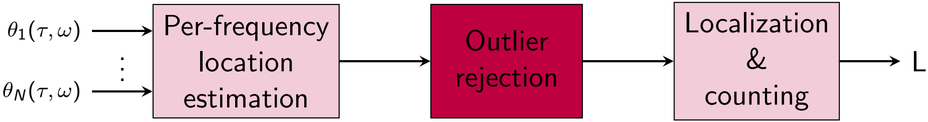


$$n^* = \arg\min_n \sum_{m=1}^{M} \left[A(\hat{\theta}_m, \psi_{m,n}) \right]^2$$

A. Alexandridis, A. Mouchtaris, "Multiple sound source location estimation and counting in a wireless acoustic sensor network," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015.

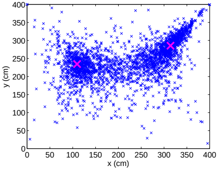
Location estimation & counting

[Alexandridis 2015]

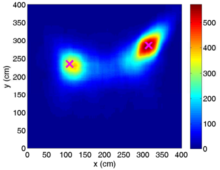


➥ Outlier rejection rule

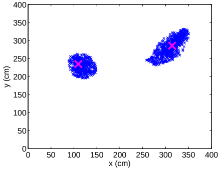
Remove the location estimates whose cardinality in the histogram is less than q times the maximum cardinality



Block estimates



Histogram

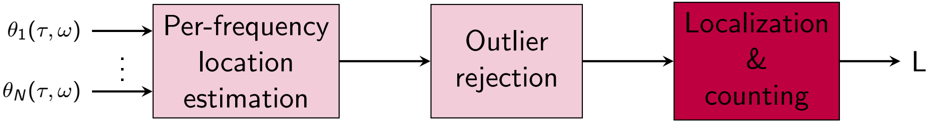


Outlier rejection result

A. Alexandridis, A. Mouchtaris, "Multiple sound source location estimation and counting in a wireless acoustic sensor network," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015.

Location estimation & counting

[Alexandridis 2015]



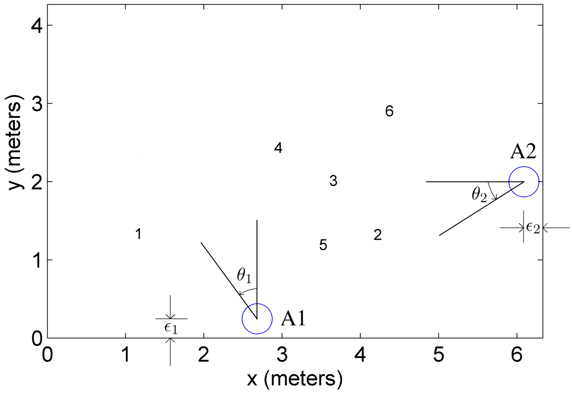
➥ Bayesian K-means objective is to find the number of clusters and the clustering assignment

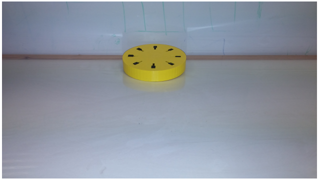
Basic principles of Bayesian K-means

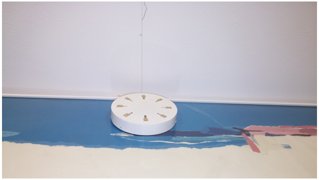
- Initialize with $C = 1$ cluster
- Split a cluster
- Merge two clusters
- Perform split/merge operations until the cost function is no longer decreased

A. Alexandridis, A. Mouchtaris, "Multiple sound source location estimation and counting in a wireless acoustic sensor network," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015.

The dataset







Speakers:

- (M01) Loc. 1–3
- (M02) Loc. 4–6

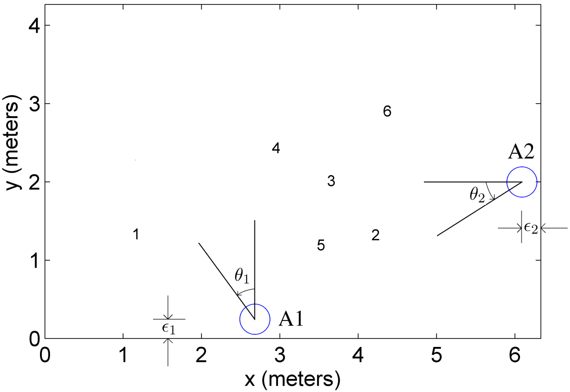
Uniform circular arrays (8 mics, 5 cm radius):

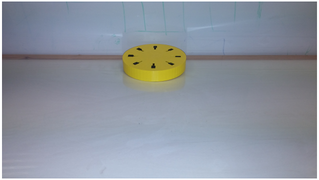
- (A01), 8.6 cm
- (A02), 8.2 cm

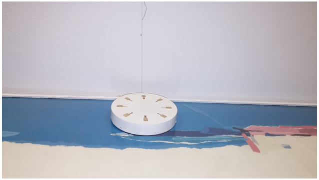
from the walls

- Arrays operated individually
- Synchronization by eye-inspection only

The dataset





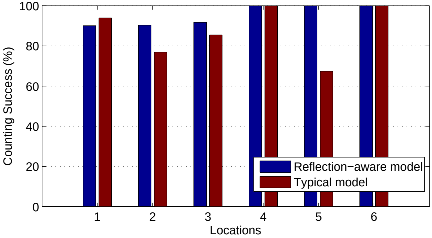


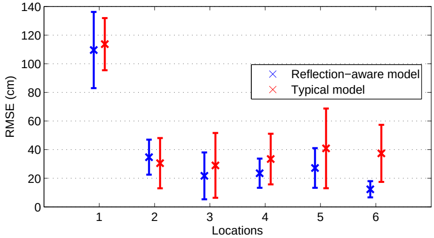
Available at:

`https://github.com/spl-icsforth/WASN-Recordings-OfficeRoom`

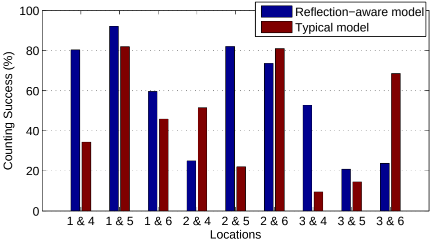
Results

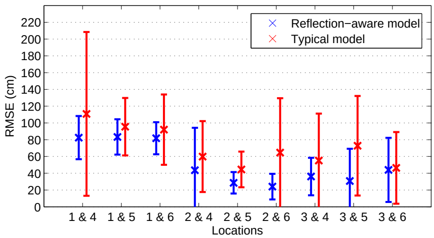
One Source





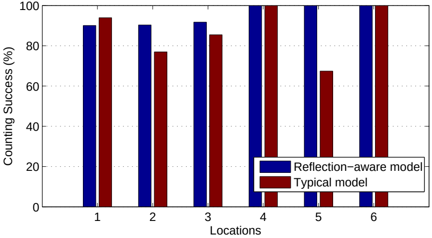
Two Sources

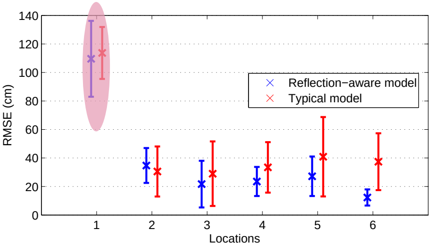




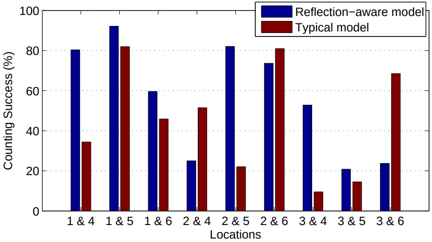
Results

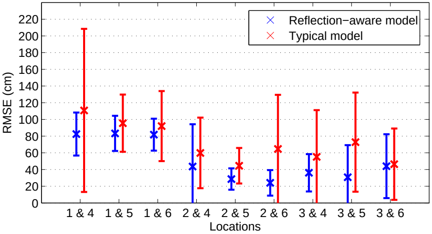
One Source





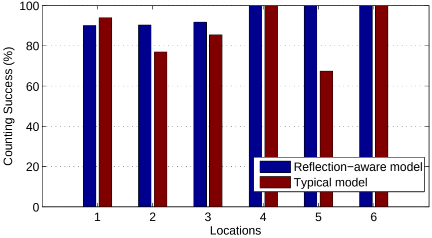
Two Sources

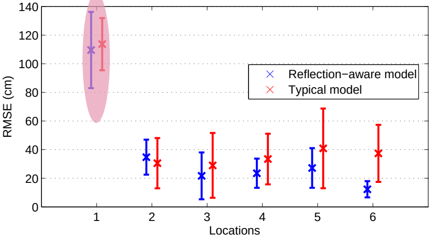




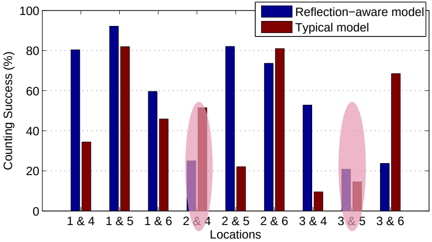
Results

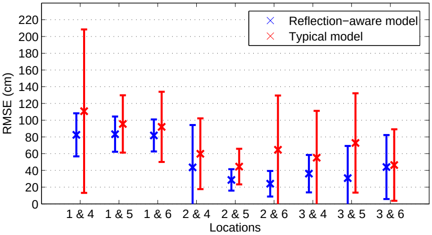
One Source





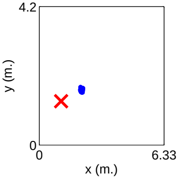
Two Sources



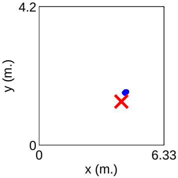


Results

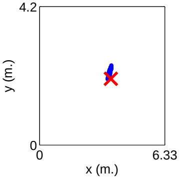
Location estimates, One Source



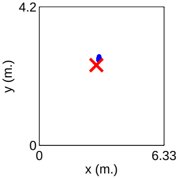
Loc. 1



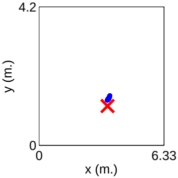
Loc. 2



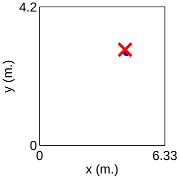
Loc. 3



Loc. 4



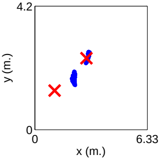
Loc. 5



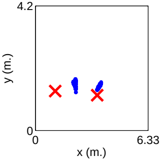
Loc. 6

Results

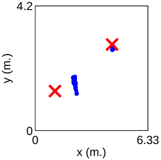
Location estimates, Two Sources



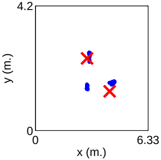
Loc. 1 & 4



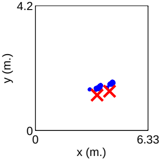
Loc. 1 & 5



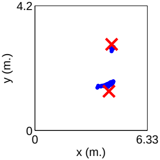
Loc. 1 & 6



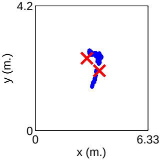
Loc. 2 & 4



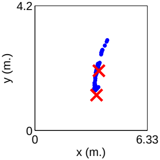
Loc. 2 & 5



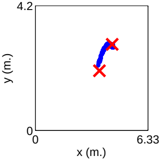
Loc. 2 & 6



Loc. 3 & 4



Loc. 3 & 5



Loc. 3 & 6

Conclusions

Location estimation & counting in real WASN and real-life conditions

- ✔ Integration of DOA-based location estimation with a "Reflection-aware" DOA estimator
- ✔ Publicly available dataset of real recorded signals in a 2-node WASN

`https://github.com/spl-icsforth/WASN-Recordings-OfficeRoom`

Things to remember!

- ✔ Evaluation with real speakers is necessary
- ✔ Evaluation across the entire localization area is important
- ✔ Performance may vary for different source configurations

Thank you!!

FORTH
Institute of Computer Science

Follow us...

http://www.spl.edu.gr

ICS-FORTH Audio Group

@spl_icsforth