Indoor Localization Accuracy Estimation from Fingerprint Data

Artyom Nikitin\textsuperscript{1}

Christos Laoudias\textsuperscript{2}
Georgios Chatzimilioudis\textsuperscript{2}
Panagiotis Karras\textsuperscript{3}
Demetrios Zeinalipour-Yazti\textsuperscript{2,4}

\textsuperscript{1}Skoltech, 143026 Moscow, Russia
\textsuperscript{2}University of Cyprus, 1678 Nicosia, Cyprus
\textsuperscript{3}Aalborg University, 9220 Aalborg, Denmark
\textsuperscript{4}Max-Planck-Institut für Informatik, 66123 Saarbrücken, Germany
1 Motivation

2 Background

3 Our Solution

4 Experiments

5 Conclusions
Motivation: Indoor Localization

- Indoor Navigation Services spread widely.
- Applications: localization, marketing, warehouse optimization, guides, games, etc.
Motivation: **Indoor Localization**

- Indoor Navigation Services spread widely.
- Applications: localization, marketing, warehouse optimization, guides, games, etc.
- Different sources of data: cellular, Wi-Fi, BT, magnetic field of the Earth, light, sound, etc.
Motivation: **Accuracy Estimation**

- Important to estimate the accuracy of localization.
Motivation: **Accuracy Estimation**

- Important to estimate the accuracy of localization.
- Online: important for the end-user (Google Maps, CONE).
Motivation: **Accuracy Estimation**

- Important to estimate the accuracy of localization.
- **Online:** important for the end-user (Google Maps, CONE).
- **Offline:** important for the service provider.
  - Provide quality guarantees.
  - Perform decision making.
Outline

1 Motivation
2 Background
3 Our Solution
4 Experiments
5 Conclusions
Background: Localization Approaches

Modeling

- Known APs positions
- Known data model, e.g., Path Loss: \[ L = 10n \log_{10}(d) + C \]
Background: Localization Approaches

Modeling + Fingerprinting

- Known APs positions
- Known data model, e.g., Path Loss: $L = 10^n \log_{10}(d) + C$
- Known pre-collected fingerprints (position + readings)
Background: Localization Approaches

Fingerprinting

- Known APs positions
- Known data model, e.g., Path Loss: \( L = 10n \log_{10}(d) + C \)
- Known pre-collected fingerprints (position + readings)
Background: **Accuracy Estimation**

**Existing solutions**
- **Heuristics**: e.g., fingerprint density, cluster & merge, etc.
  - + Do not require models
  - − No theoretical guarantees
- **Theoretical**: e.g., use *Cramer-Rao Lower Bound* (CRLB)
  - + Provide theoretical guarantees
  - − Model is required

**Our goal:**
- + No model required
- + Provide guarantees via CRLB
Background: **Accuracy Estimation**

Common theoretical approach for offline accuracy estimation:

1. Measurements are random, e.g., Gaussian.
Common theoretical approach for offline accuracy estimation:

1. Measurements are random, e.g., Gaussian.
2. From the known information estimate the likelihood, i.e., the probability \( p(m|r) \) of measuring \( m \) at \( r \).
Background: **Accuracy Estimation**

Common theoretical approach for offline accuracy estimation:

1. Measurements are random, e.g., Gaussian.
2. From the known information estimate the *likelihood*, i.e., the probability $p(m|r)$ of measuring $m$ at $r$.
3. From the *likelihood* calculate *Cramer-Rao Lower Bound* (CRLB) on the variance of any unbiased estimator of $r$. 
Common theoretical approach for offline accuracy estimation:

1. Measurements are random, e.g., Gaussian.
2. From the known information estimate the \textit{likelihood}, i.e., the probability $p(m|r)$ of measuring $m$ at $r$.
3. From the \textit{likelihood} calculate \textit{Cramer-Rao Lower Bound} (CRLB) on the variance of any unbiased estimator of $r$. 
Common theoretical approach for offline accuracy estimation:
1. Measurements are random, e.g., Gaussian.
2. From the known information estimate the likelihood, i.e., the probability $p(m|r)$ of measuring $m$ at $r$.
3. From the likelihood calculate Cramer-Rao Lower Bound (CRLB) on the variance of any unbiased estimator of $r$. 

![Graph showing Gaussian distributions at x and x + Δx](image)
How to find the *likelihood*?
Background: **Accuracy Estimation**

**Modeling.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \) (FSPL)
- Position \( x_{AP} \) of the AP
- Noise
Background: **Accuracy Estimation**

**Modeling.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2 \), \( C = 20 \log_{10} \frac{4\pi}{\lambda} \) (FSPL)
- Position \( x_{AP} \) of the AP
- Noise

1. Predict using model
Background: **Accuracy Estimation**

**Modeling.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \) (FSPL)
- Position \( x_{AP} \) of the AP
- Noise

1. Predict using model
2. Estimate noise
Background: **Accuracy Estimation**

**Modeling.** We know:

- Model, e.g., Path Loss: \[ L = 10n \log_{10}(|x - x_{AP}|) + C \]
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \) (FSPL)
- Position \( x_{AP} \) of the AP
- Noise

1. Predict using model
2. Estimate noise
3. Compare to measurements
Background: **Accuracy Estimation**

**Modeling + Fingerprinting.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \)
- Position \( x_{AP} \) of the AP
- Noise
Background: Accuracy Estimation

Modeling + Fingerprinting. We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \)
- Position \( x_{AP} \) of the AP
- Noise

1. Assume parametric model
Background: **Accuracy Estimation**

**Modeling + Fingerprinting.** We know:

- Model, e.g., Path Loss: \( L = 10^n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} P_0 \)
- Position \( x_{AP} \) of the AP
- Noise

1. Assume parametric model
2. Get fingerprints
Background: **Accuracy Estimation**

**Modeling + Fingerprinting.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \)
- Position \( x_{AP} \) of the AP
- Noise

1. Assume parametric model
2. Get fingerprints
3. Estimate parameters

![Diagram showing path loss and fingerprints](image)
Background: **Accuracy Estimation**

**Modeling + Fingerprinting.** We know:

- Model, e.g., Path Loss: \( L = 10n \log_{10}(|x - x_{AP}|) + C \)
- Model parameters, e.g., \( n = 2, C = 20 \log_{10} \frac{4\pi}{\lambda} \)
- Position \( x_{AP} \) of the AP
- Noise

1. Assume parametric model
2. Get fingerprints
3. Estimate parameters
4. Estimate noise
Background: **Accuracy Estimation**

**Pure Fingerprinting**
- No model provided.
Background: **Accuracy Estimation**

**Pure Fingerprinting**

- No model provided.
- Data is too complex, e.g., ambient magnetic field:
  - vector field = direction + magnitude;
  - predictable outdoors;
  - perturbed indoors by metal constructions and electrical equipment.
Background: **Accuracy Estimation**

**Pure Fingerprinting**

- No model provided.
- Data is too complex, e.g., ambient magnetic field:
  - vector field = direction + magnitude;
  - predictable outdoors;
  - perturbed indoors by metal constructions and electrical equipment.
Outline

1 Motivation

2 Background

3 Our Solution

4 Experiments

5 Conclusions
Our solution: Goal

- Pure fingerprinting approach

- Arbitrary data sources

- \( FM = \{(r_i, m_i) : i = 1, N, r_i \in \mathbb{R}^{d_r}, m_i \in \mathbb{R}^{d_m}\} \)
  - \( m_i \) - \( d_m \)-dimensional vector of measurements at location \( r_i \).

- Given the FM, assign to any location a navigability score.

- Visualize navigability scores to assist INS deployer.
**ACCES framework**

1. **Interpolation:**
   \[ \text{FM} + \text{Gaussian Process Regression (GPR)} \Rightarrow \text{likelihood} \]

2. **CRLB:**
   \[ \text{Likelihood} + \text{CRLB} \Rightarrow \text{lower bound on localization error} \]

3. **Lower bound on localization error \Rightarrow navigability score**
   - Theoretical bound on localization error
   - Assume that behaves similar to real error
Our Solution: **Interpolation**

**Gaussian Process Regression:**

- **Input:** fingerprint map of measurements
- **Output:** Gaussian likelihood $p(m|r)$ of measuring $m$ at $r$
  (prediction + uncertainty)

**Intuition:**
- measurements are Gaussian random variables
- spatial correlation: close are correlated, far are not
Our Solution: **Interpolation**

Gaussian Process Regression:

- **Input:** fingerprint map of measurements
- **Output:** Gaussian likelihood $p(m|r)$ of measuring $m$ at $r$ (prediction + uncertainty)
- **Intuition:**
  - measurements are Gaussian random variables
  - spatial correlation: close are correlated, far are not
- **Properties:**
  - models arbitrary noisy data
  - captures FM’s spatial sparsity

Nuances:
- parameters tuning is required (kernel, length scale, etc.)
- assume normality condition (does not directly work for NLOS)
- directly applicable only to scalar data $\Rightarrow$ assume independence
- computationally expensive $\Rightarrow$ clustering
Our Solution: **Interpolation**

Gaussian Process Regression:

- **Input**: fingerprint map of measurements
- **Output**: Gaussian likelihood $p(m|r)$ of measuring $m$ at $r$ (prediction + uncertainty)

**Intuition:**
- measurements are Gaussian random variables
- spatial correlation: close are correlated, far are not

**Properties:**
- models arbitrary noisy data
- captures FM’s spatial sparsity

**Nuances:**
- parameters tuning is required (kernel, length scale, etc.)
- assume normality condition (does not directly work for NLOS)
- directly applicable only to scalar data $\Rightarrow$ assume independence
- computationally expensive $\Rightarrow$ clustering
Our Solution: **Interpolation**

**Gaussian Process Regression (1-D example)**

The diagram illustrates Gaussian Process Regression for a 1-D example. It shows the initial data points, noisy data, the prediction, and the uncertainty. The x-axis represents the input variable, and the y-axis represents the output variable.
Cramer-Rao Lower Bound:

- **Input:** likelihood $p(m|r)$ of measuring $m$ at $r$
- **Output:** smallest RMSE achievable by any unbiased estimator of $r$

**Intuition:**
- likelihood carries information about the distribution
- distribution does not vary locally $\Rightarrow$ degradation
Our Solution: CRLB

Cramer-Rao Lower Bound:

- **Input**: likelihood $p(m|r)$ of measuring $m$ at $r$
- **Output**: smallest RMSE achievable by any unbiased estimator of $r$
- **Intuition**:
  - likelihood carries information about the distribution
  - distribution does not vary locally $\Rightarrow$ degradation
- **Properties**:
  - theoretical
  - easily found for unbiased estimators
Our Solution: CRLB

Cramer-Rao Lower Bound:

- **Input:** likelihood $p(m|r)$ of measuring $m$ at $r$
- **Output:** smallest RMSE achievable by any unbiased estimator of $r$
- **Intuition:**
  - likelihood carries information about the distribution
  - distribution does not vary locally $\Rightarrow$ degradation
- **Properties:**
  - theoretical
  - easily found for unbiased estimators
- **Nuances:**
  - an underestimation of the real error $\Rightarrow$ we care about qualitative behavior
  - analytical representation depends on GPR parameters $\Rightarrow$ we involve numerical methods
Our Solution: CRLB

CRLB: error of any unbiased location estimator is bounded as

\[ \text{RMSE} \geq \sqrt{\text{tr}(I^{-1}(r))}, \]

where \( I(r) \) is a Fisher Information Matrix:

\[
I(r) = -\mathbb{E} \left( \frac{\partial^2 \log p(m|r)}{\partial r_i \partial r_j} \right)
\]
Our Solution: CRLB

CRLB: error of any unbiased location estimator is bounded as

$$RMSE \geq \sqrt{tr(I^{-1}(r))},$$

where $I(r)$ is a *Fisher Information Matrix*:

$$I(r) = -\mathbb{E} \left( \frac{\partial^2 \log p(m|r)}{\partial r_i \partial r_j} \right)$$

From GPR:

$$m|\mathbf{r} \sim \mathcal{N}(\mu(\mathbf{r}), \Sigma(\mathbf{r}))$$

Thus,

$$I(\mathbf{r}) = \frac{1}{2} \sum_{k=1}^{d_m} \left[ (\sigma_k^2 + \mu_k^2) H(\sigma_k^{-2}) + H(\mu_k^2 \sigma_k^{-2}) - 2 \mu_k H(\mu_k \sigma_k^{-2}) + 2 H(\log \sigma_k) \right]$$
Outline

1 Motivation
2 Background
3 Our Solution
4 Experiments
5 Conclusions
Experiments: Data

- UJIIndoorLoc-Mag database
- 8 corridors over 260 $m^2$ lab
- 40,159 discrete captures
- Magnetometer readings
- Measurements along the corridors $\Rightarrow$ 1-D data

IEEE MDM 2017 | Nikitin, Laoudias, Chatzimilioudis, Karras, Zeinalipour
Experiments: Algorithms

- **Real accuracy**: RMSE via WkNN
Experiments: Algorithms

- **Real accuracy**: RMSE via WkNN
- **Naïve approach**: Fingerprint Spatial Sparsity Indicator:
  
  $$FSSI(r) = \min_{i \in 1, N} ||r - r_i||$$

  considers only distance between measurements
Experiments: Algorithms

- **Real accuracy**: RMSE via WkNN
- **Naïve approach**: Fingerprint Spatial Sparsity Indicator:
  \[ FSSI(r) = \min_{i \in 1, N} \| r - r_i \| \]
  considers only distance between measurements
- **ACCES**: our solution
Experiments: Metrics

\( DQRelSim(X, Y) \) - behavior similarity of sequences \( X = X_i, \ Y = Y_i \) for 1-D case.

- **Construction:**
  - Difference Quotient \( \Rightarrow DQ(X) \) and \( DQ(Y) \)
  - DTW \( \Rightarrow \) optimally warped \( DQ(X)' \) and \( DQ(Y)' \) from Normalization

- **Values:**
  - Similar: 1, if \( X = Y + \text{const} \)
  - Dissimilar: 0, if either \( X \) or \( Y \) is constant
  - Opposite: \( -1 \), if \( X = -Y + \text{const} \)
Experiments: Settings

\( DQRelSim(ACCES, RMSE) \) vs \( DQRelSim(FSSI, RMSE) \)

1. **“Cut” scenario:**
   - Contiguous sequence of measurements is removed
   - \( \Leftrightarrow \) fingerprints were not collected

2. **“Flat” scenario:**
   - Contiguous sequence of measurements is made constant
   - \( \Leftrightarrow \) low signal variability

3. **“Sparse” scenario:**
   - Measurements are removed uniformly
   - \( \Leftrightarrow \) different frequency of fingerprint collection
“Cut” scenario: magnetic field magnitude

Experiments: Evaluation

IEEE MDM 2017 | Nikitin, Laoudias, Chatzimilioudis, Karras, Zeinalipour
Experiments: Evaluation

“Cut” scenario: similarity of RMSE, ACCES, FSSI

IEEE MDM 2017 | Nikitin, Laoudias, Chatzimilioudis, Karras, Zeinalipour
“Flat” scenario: magnetic field magnitude

Corridor 1

Corridor 2

Corridor 3

Corridor 4
“Flat” scenario: similarity of RMSE, ACCES, FSSI

Experiments: Evaluation

**Corridor 1**

\[ DQRelSim = 3.0 \times 10^{-01} \text{ vs } -1.3 \times 10^{-08} \]

**Corridor 2**

\[ DQRelSim = 2.7 \times 10^{-01} \text{ vs } -8.9 \times 10^{-07} \]

**Corridor 3**

\[ DQRelSim = 2.5 \times 10^{-01} \text{ vs } 8.7 \times 10^{-10} \]

**Corridor 4**

\[ DQRelSim = 2.3 \times 10^{-01} \text{ vs } -7.7 \times 10^{-08} \]
Experiments: Evaluation

“Sparse” scenario: behaviour of RMSE, ACCES

IEEE MDM 2017 | Nikitin, Laoudias, Chatzimilioudis, Karras, Zeinalipour
Conclusions

Summary:

- **ACCES** provides offline accuracy estimations and FM assessment.
- Does not consider the origin of the data.
- Applicable to pure fingerprinting.
- Shows reasonable correspondence to the real localization error behaviour.

Future work:
- Extensive experimental study with other data.
- Comparison to online accuracy estimation algorithms.
- Adding support for arbitrary models.
Conclusions

Summary:
- ACCES provides offline accuracy estimations and FM assessment.
- Does not consider the origin of the data.
- Applicable to pure fingerprinting.
- Shows reasonable correspondence to the real localization error behaviour.

Future work:
- Extensive experimental study with other data.
- Comparison to online accuracy estimation algorithms.
- Adding support for arbitrary models.
Advertisement: Anyplace

Anyplace Indoor Information Service

IEEE MDM 2017 | Nikitin, Laoudias, Chatzimilioudis, Karras, Zeinalipour
Anyplace Indoor Information Service

- Wi-Fi + IMU
- Optimize data usage
- 3 awards
Anyplace Indoor Information Service

- Wi-Fi + IMU
- Optimize data usage
- 3 awards
- Crowdsource-based
- Open-source
Advertisement: Anyplace

Anyplace Indoor Information Service

- Wi-Fi + IMU
- Optimize data usage
- 3 awards
- Crowdsource-based
- Open-source
- anyplace.cs.ucy.ac.cy
- github.com/dmsl/anyplace
Thank You!

Come to see our demo yesterday!
Summary:

- **ACCES** provides offline accuracy estimations and FM assessment.
- Does not consider the origin of the data.
- Applicable to pure fingerprinting.
- Shows reasonable correspondence to the real localization error behaviour.

Future work:

- Extensive experimental study with other data.
- Comparison to online accuracy estimation algorithms.
- Adding support for arbitrary models.