UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Computer Science, Electrical Engineering and Mathematics

Communications Engineering
Prof. Dr.-Ing. Reinhold Häb-Umbach

PARAMETER ESTIMATION AND CLASSIFICATION OF CENSORED GAUSSIAN DATA WITH APPLICATION TO WIFI INDOOR POSITIONING

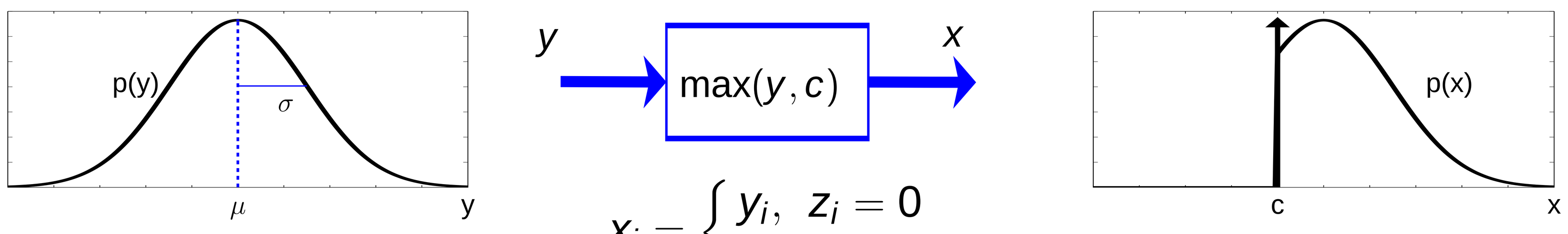
Manh Kha Hoang and Reinhold Haeb-Umbach

University of Paderborn, Germany, {hoang, haeb}@nt.uni-paderborn.de, http://www-nt.uni-paderborn.de/

Introduction

- RSSI fingerprint based WiFi indoor positioning
  - Limited sensitivity of WiFi chips: $-100\,\text{dBm} \leq \text{RSSI} \leq -30\,\text{dBm}$
  - Leads to censored, i.e., clipped observations
- EM algorithm for parameter estimation of censored Gaussian data:
  - Unbiased
  - Efficient (achieves the Cramer-Rao Bound (CRB))
- Optimal classification/localization in the presence of censored data

Parameter Estimation



$$x_i = \begin{cases} y_i, & z_i = 0 \\ c, & z_i = 1 \end{cases}$$

$y_i, i = 1, ..., N$: complete data

$x_i, i = 1, ..., N$: observed data

- Re-estimation formulas:

$$\mu^{(\kappa+1)} = \frac{1}{N} \cdot \frac{I_1(\theta^{(\kappa)})}{I_0(\theta^{(\kappa)})} \sum_{i=1}^{N} z_i + \frac{1}{N} \sum_{i=1}^{N} (1 - z_i) x_i$$

$$(\sigma^2)^{(\kappa+1)} = \left[\frac{I_2(\theta^{(\kappa)})}{I_0(\theta^{(\kappa)})} - 2\mu^{(\kappa)} \frac{I_1(\theta^{(\kappa)})}{I_0(\theta^{(\kappa)})} + \left(\mu^2\right)^{(\kappa)} \right] \frac{1}{N} \sum_{i=1}^{N} z_i + \frac{1}{N} \sum_{i=1}^{N} (1 - z_i) \left(x_i - \mu^{(\kappa)}\right)^2$$

where

- θ: (μ, σ^2)
- κ: iteration index
- $I_j(\theta^{(\kappa)}) = \int_{-\infty}^{c} y^j \mathcal{N}\left(y; \theta^{(\kappa)}\right) dy$
- c: clipping threshold

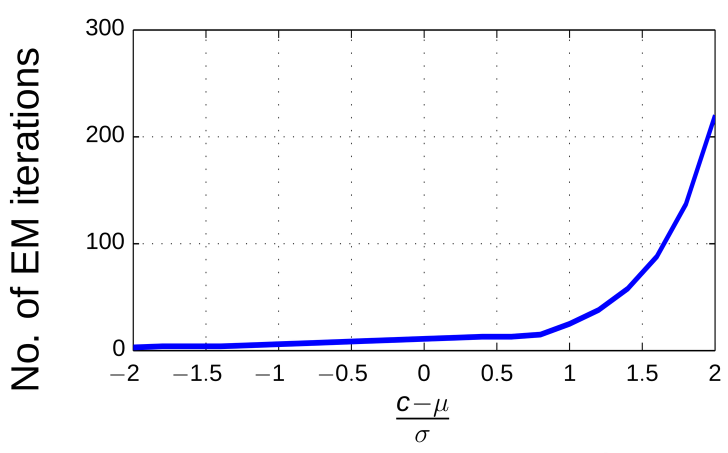
Properties of Estimates

- Unbiasedness:
  - Let $\tilde{\mu}^{(\kappa+1)} = \mu^{(\kappa+1)} - \mu$; $(\tilde{\sigma}^2)^{(\kappa+1)} = (\sigma^2)^{(\kappa+1)} - \sigma^2$

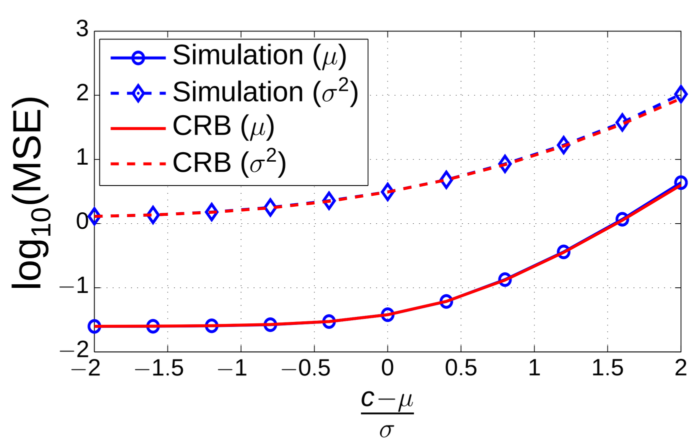
    It can be shown:

$$\begin{pmatrix} E\left[\tilde{\mu}^{(\kappa+1)}\right] \\ E\left[(\tilde{\sigma}^2)^{(\kappa+1)}\right] \end{pmatrix} \approx \begin{pmatrix} W_\mu & W_{\mu\sigma} \\ W_{\sigma\mu} & W_\sigma \end{pmatrix}^{\kappa+1} \begin{pmatrix} E\left[\tilde{\mu}^{(0)}\right] \\ E\left[(\tilde{\sigma}^2)^{(0)}\right] \end{pmatrix}$$

  - Eigenvalues $|\lambda_i(W)| < 1 \Longrightarrow$ unbiased estimates
- Convergence speed and precision of EM algorithm:



Convergence speed



Comparison of simulation with CRB

  - Number of iterations quickly rises once more than 50% of the data are clipped
  - Mean square error of the estimator: achieves the CRB

Classification of Censored Data

- Likelihood

$$p(\mathbf{y}|\ell_k) = p(y_1, ..., y_{N_{AP}}|\ell_k) = \prod_{j=1}^{N_{AP}} \mathcal{N}(y_j; \hat{\mu}_{j,\ell_k}, \hat{\sigma}^2_{j,\ell_k})$$

  - N_{AP}: number of access points
  - $\ell_k, k = 1, ..., K$: possible user locations
  - RSSI measurement: $x_j = \max(y_j, c)$

$$p(x_j|\ell_k) = \begin{cases} \mathcal{N}(x_j; \hat{\mu}_{j,\ell_k}, \hat{\sigma}^2_{j,\ell_k}), & \text{if } x_j > c \\ I_0(\hat{\mu}_{j,\ell_k}, \hat{\sigma}^2_{j,\ell_k}), & \text{if } x_j = c \end{cases}$$

- User positioning = classification

$$\hat{\ell} = \arg\max_{\ell_k} P(\ell_k|\mathbf{x}) = \arg\max_{\ell_k} \prod_{j=1}^{N_{AP}} p(x_j|\ell_k) P(\ell_k)$$

- Refinements:
  - Multiple observations per position: $\mathbf{x}_{1:S} = \mathbf{x}_1, ..., \mathbf{x}_S$
  - Weighted average over the set $\mathcal{P}$ of most likely positions

$$\hat{\ell} = \frac{1}{\sum_{k \in \mathcal{P}} P(\ell_k|\mathbf{x}_{1:S})} \sum_{k \in \mathcal{P}} \ell_k P(\ell_k|\mathbf{x}_{1:S})$$

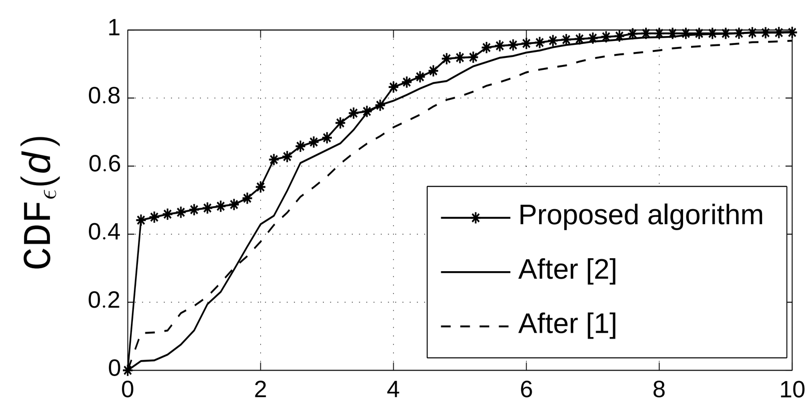
Experimental Results

- Classification on artificial data: 2-class problem, $N_{AP} = 5$, the percentage of clipped data was approximately 42%

Method	Error rate (%)
Plain training + plain recognition	30.7
EM training + plain recognition	26.9
EM training + censored recognition	22.5
1-nearest neighbor	36.8

- Classification on field data:

Floor plan (12 m x 30 m)



Distance d [m]

  - Proposed method outperforms the other w.r.t. accuracy
  - Low computational complexity

Conclusions

- EM algorithm delivers unbiased and efficient estimates, achieving the CRB
- An optimal classification method accounting for censored data was presented
- The effectiveness of the proposed algorithm for WiFi indoor positioning was demonstrated

[1] P. Bahl and V.N. Padmanabhan, "RADAR: An In-Building RF-Based User Location and Tracking System," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies.* IEEE, 2000, vol. 2, pp. 775–784.

[2] Nicolas Le Dortz, Florian Gain, and Per Zetterberg, "WiFi fingerprint indoor positioning system using probability distribution comparison," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2301–2304, March 2012.