Dynamic Bayesian Optimization for Improving the Performance of Cellular Networks

Author: Emanuele Mengoli
Host Institution: École Polytechnique Fédérale de Lausanne (EPFL)
EPFL Supervisors: Patrick Thiran, Anthony Bardou
École Polytechnique supervisor: Juan Antonio Cordero Fuertes
Internship Period: 15/04/2024 - 05/09/2024
Internship course code: INF690

September 5, 2024

Table of Contents

1. Cellular Network Framework
   - Introduction
   - Resource Sharing Mechanism (RSM) choice
2. Problem formulation
   - Propagation model
   - Objective Function
   - Scheduling
3. Power Allocation Strategies
   - Introduction to Bayesian Optimization (BO)
   - Power Allocation Strategies
4. UEs modilization
   - UEs Generation
   - UEs Mobility
5. Experimental Results
   - Objective function comparison
   - W-DBO's hyper-parameters comparison
6. Conclusion
   - Current limitations
   - Future extensions
   - Outcomes

Table of Contents

1 Cellular Network Framework
- Introduction
- Resource Sharing Mechanism (RSM) choice

2 Problem formulation
- Propagation model
- Objective Function
- Scheduling

3 Power Allocation Strategies
- Introduction to Bayesian Optimization (BO)
- Power Allocation Strategies

4 UEs modilization
- UEs Generation
- UEs Mobility

5 Experimental Results
- Objective function comparison
- W-DBO's hyper-parameters comparison

6 Conclusion
- Current limitations
- Future extensions
- Outcomes

Problem Introduction

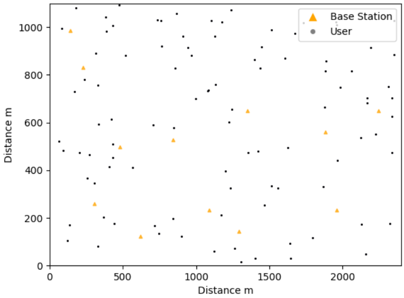


Figure: Example of Dynamic Wireless Cellular Network. BSs position are retrieved from [1].

Key elements

1. Parisian Base Stations (BSs) data. [1]
2. Choose Resource Sharing Mechanism (RSM).
3. Power Allocation strategy:
   - Define parametric obj function to be optimized.
4. Model User Equipments (UEs):
   - Placement distribution.
   - Birth & Death process.
   - Mobility models.
5. Border effect filtering.

RSM choice

Non-Orthogonal Multiple Access (NOMA)

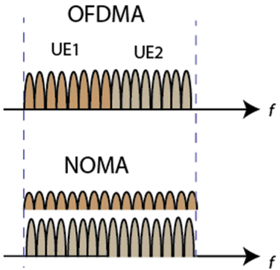


Figure: Spectrum sharing for two users: NOMA vs. OFDMA. Image source: [2].

Key points

- Full spectrum sharing.
- 2 service region partioning per BS.
- $\mathbf{p}_i = \left(p_i^{(\text{in})}, p_i^{(\text{out})}\right)$
- $P_- \leqslant \|\mathbf{p}_i\|_1 \leqslant P_+$ (BS power interval)
- $p^{(\text{in})} \leqslant p^{(\text{out})}$.
- Successive Interference Cancellation (SIC):
  - Inner UEs decode the signal for outer UEs first and cancel their interference.
  - Outer UEs decode only their own signal and treat the inner UEs' signal as noise.
- Literature contribution [2, 3, 4, 5].

Table of Contents

1 Cellular Network Framework
- Introduction
- Resource Sharing Mechanism (RSM) choice

2 Problem formulation
- Propagation model
- Objective Function
- Scheduling

3 Power Allocation Strategies
- Introduction to Bayesian Optimization (BO)
- Power Allocation Strategies

4 UEs modilization
- UEs Generation
- UEs Mobility

5 Experimental Results
- Objective function comparison
- W-DBO's hyper-parameters comparison

6 Conclusion
- Current limitations
- Future extensions
- Outcomes

Problem formulation (A)

Propagation model

Recall **NOMA power constraint** for efficient SIC → Define the **parameter space** of the Problem:

$$\underbrace{\left(p_i^{(\text{in})}, p_i^{(\text{out})}\right)}_{\mathbf{p}_i} \rightarrow \underbrace{\left(p_i^{(\text{in})} + p_i^{(\text{out})}, \frac{p_i^{(\text{out})}}{p_i^{(\text{in})} + p_i^{(\text{out})}}\right)}_{\mathbf{x}_i} \tag{1}$$

$$\mathcal{D} = \bigtimes_{w=1}^{n} \mathcal{D}^{(w)} = \left([P_-, P_+] \times \left[\frac{1}{2}, 1\right)\right)^n \tag{2}$$

SINR - inner UE j and outer UE j', BS i:

$$\gamma_{i,j}^{(\text{in})} = \frac{g_{i,j} p_1^{(i)}}{WN + \sum_{i' \in \mathcal{N}_i \setminus \{i\}} g_{i',j}(p_1^{(i')} + p_2^{(i')})} \tag{3}$$

$$\gamma_{i,j'}^{(\text{out})} = \frac{g_{i,j'} p_2^{(i)}}{WN + g_{i,j'} p_1^{(i)} + \sum_{i' \in \mathcal{N}_i \setminus \{i\}} g_{i',j'}(p_1^{(i')} + p_2^{(i')})} \tag{4}$$

N noise power density (dBm/Hz), W signal bandwidth, $g_{i,j}$ radio link gain (computed through standard **log-distance path-loss model**).

Problem formulation (B)

Objective Function

Downlink Shannon capacity for UE j is:

$$c^{(j)} = c(\gamma_{i,j}) = W_j \cdot \log_2(1 + \gamma_{i,j}). \tag{5}$$

Objective

Goal: The natural objective of the model is to maximize the overall downlink capacity, yielding the highest Quality of Service (QoS) for the end-users.
How? $F_\alpha(c_t)$ (Eq. (6))

Include the notion of α-**fairness** for scalar mapping:

$$F_\alpha(c_t) = \begin{cases} \sum_{i=1}^{m} \log c_t^{(j)} & \text{if } \alpha = 1, \\ \sum_{j=1}^{m} \frac{\left(c_t^{(j)}\right)^{1-\alpha}}{1-\alpha} & \text{otherwise}, \end{cases} \tag{6}$$

where $c_t^{(j)}$ denotes $c^{(j)}(\mathbf{x}_t)$.

Problem formulation (C)

Scheduling

Optimal Scheduling $s_j^{(k)}$, $j \in \mathcal{A}^{(k)}$, w.r.t the k-th network partition ($k = 1, \cdots, 2n$).

$$s_j^{(k)} = \begin{cases} \mathbb{1}_{j=j*} & \text{if } \alpha = 0, \\ \dfrac{\left(c_t^{(j)}\right)^{(1-\alpha)/\alpha}}{\sum_{i \in \mathcal{A}^{(k)}} \left(c_t^{(i)}\right)^{(1-\alpha)/\alpha}} & \text{otherwise,} \end{cases} \tag{7}$$

Thus, the **objective function** $f_\alpha^{(i)} : \mathcal{D}^{(\mathcal{N}_i)} \to \mathcal{D}$, where $\mathcal{D}^{(\mathcal{N}_i)} = \times_{i' \in \mathcal{N}_i} \mathcal{D}^{(i')}$:

$$f_\alpha\left(\mathbf{x}_t\right) = \sum_{i=1}^{n} \sum_{\mathcal{A}^{(k)} \in \mathcal{A}_i} \sum_{j \in \mathcal{A}^{(k)}} F_\alpha\left(\mathbf{s}^{(k)} \odot \mathbf{c}^{(k)}\left(\mathbf{x}_t^{(\mathcal{N}_i)}\right)\right) \tag{8}$$

Where $\mathcal{N}_i$ denotes the neighbourhood of BS i.

Optimization

How to optimize $f_\alpha\left(\mathbf{x}_t\right)$ (Eq. (8)) ? $\to$ **Power Allocation Strategies**.

Table of Contents

1 Cellular Network Framework
- Introduction
- Resource Sharing Mechanism (RSM) choice

2 Problem formulation
- Propagation model
- Objective Function
- Scheduling

3 Power Allocation Strategies
- Introduction to Bayesian Optimization (BO)
- Power Allocation Strategies

4 UEs modilization
- UEs Generation
- UEs Mobility

5 Experimental Results
- Objective function comparison
- W-DBO's hyper-parameters comparison

6 Conclusion
- Current limitations
- Future extensions
- Outcomes

Bayesian Optimization (BO)

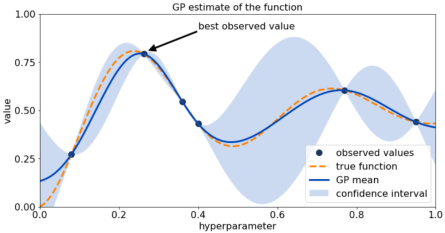


Figure: Example of BO procedure using GP. Image source [6].

How?

Input: objective function f, acquisition function φ, time horizon T

Init dataset $\mathcal{D} = \varnothing$ and surrogate model $\mathcal{G} = \mathcal{GP}(\mu(\mathbf{x}), k((\mathbf{x}), (\mathbf{x}')))$ where $\mu_0(\mathbf{x}) = 0$ w.l.o.s.

for $t \in [1, T]$ **do**

$\mu_t(\mathbf{x}) = \mathbf{k}(\boldsymbol{x}, \mathbf{X})^\top \mathbf{K}^{-1}\mathbf{z}$
$\sigma_t^2(\mathbf{x}) = \sigma_0^2(\mathbf{x}) - \boldsymbol{k}(\boldsymbol{x}, \mathbf{X})^\top \mathbf{K}^{-1}\mathbf{k}(\boldsymbol{x}, \mathbf{X})$

where $\mathbf{k}(\mathbf{x}, \mathbf{X}) = (k(\mathbf{x}, \mathbf{x}_i))_{i \in [1,t]}$, $\mathbf{K} = (k(\mathbf{x}_i, \mathbf{x}_j))_{i,j \in [1,t]}$, $\mathbf{X} = (\boldsymbol{x}_i)_{i=1}^t$, $\mathbf{z} = (z_i)_{i=1}^t$.

Compute $x_t = \arg\max_{x \in \mathcal{M}} \varphi(x; \mathcal{G})$ *, being $\mathcal{M}$ the domain

Observe $z_t = f(x_t)$ and add (x_t, z_t) to $\mathcal{D}$

end for

* e.g. GP-UCB [7]:
$\varphi_t(\mathbf{x}) = \mu_t(\mathbf{x}) + \beta_t^{1/2}\sigma_t(\mathbf{x})$

Key Components

- **Objective function**: $f : \mathcal{S} \times \mathcal{T} \to \mathbb{R}$, s.t. $\mathcal{S} \subseteq \mathbb{R}^d$, $\mathcal{T} \subseteq \mathbb{R}$.
- **Surrogate Model**: $\mathcal{G} \sim \mathcal{GP}(\mu(\mathbf{x}, t), k((\mathbf{x}, t), (\mathbf{x}', t')))$
  - **Init**: Mean $\mu(\mathbf{x}, t) = 0$, Covariance $k((\mathbf{x}, t), (\mathbf{x}', t'))$
  - Let $k : \mathcal{S} \times \mathcal{T} \times \mathcal{S} \times \mathcal{T} \to \mathbb{R}_+$
- **Kernel Decomposition**:

$$k((\mathbf{x}, t), (\mathbf{x}', t')) = \lambda k_S(\|\mathbf{x} - \mathbf{x}'\|, l_S) k_T(|t - t'|, l_T)$$

  - **Spatial Kernel**: $k_S(\cdot, l_S)$, s.t. $k_S : \mathbb{R}_+ \to [0, 1], l_S > 0$
  - **Temporal Kernel**: $k_T(\cdot, l_T)$, s.t. $k_T : \mathbb{R}_+ \to [0, 1], l_T > 0$
  - **Lengthscale parameters**: λ, l_S, l_T; $\lambda > 0$

W-DBO

Spatio-Temporal BO (2/2)

Key Components

- **W-DBO characteristic**:
  - **MLE Estimation**: Infer (λ, l_S, l_T)
  - **Wasserstein Distance**:

$$W_2\left(\mathcal{GP}_{\mathcal{D}}, \mathcal{GP}_{\tilde{\mathcal{D}}}\right) = \left(\oint_{\mathcal{S}} \int_{t_0}^{\infty} W_2^2\left(\mathcal{N}_{\mathcal{D}}(\mathbf{x}, t), \mathcal{N}_{\tilde{\mathcal{D}}}(\mathbf{x}, t)\right) \, d\mathbf{x} \, dt\right)^{\frac{1}{2}}$$

Measures deviation between two GPs, differing by one removed observation. $\mathcal{D}$ is the original dataset, and $\tilde{\mathcal{D}}$ is the dataset with one observation removed.

Power Allocation Strategies

Random Power Pick

At each optimization round t the vector $\mathbf{x}_{\text{rndm}}^{(t)} \sim \mathcal{U}(\mathcal{D})$.

Constant Power Pick

$\mathbf{x}_{\text{const}} \sim \mathcal{U}(\mathcal{D})$ is sampled just once, during the first round.

W-DBO [8]

- W-DBO instances define different power level suggestions for a common neighborhood: $\mathbf{x}_{\text{wdbo},i}^{(t)} \in \mathcal{D}^{(i)}$, $\forall i \in \mathcal{N}$.
- Median is computed to determine BS power levels: $\mathbf{p}_i = \text{MEDIAN}_{i' \in \mathcal{N}_i}(\mathbf{p}_{i'})$, $\forall i \in \mathcal{N}$.
- Power levels $\mathbf{p}_i$ are retrieved via the bijection in Eq. (1).

Table of Contents

1. Cellular Network Framework
   - Introduction
   - Resource Sharing Mechanism (RSM) choice
2. Problem formulation
   - Propagation model
   - Objective Function
   - Scheduling
3. Power Allocation Strategies
   - Introduction to Bayesian Optimization (BO)
   - Power Allocation Strategies
4. UEs modilization
   - UEs Generation
   - UEs Mobility
5. Experimental Results
   - Objective function comparison
   - W-DBO's hyper-parameters comparison
6. Conclusion
   - Current limitations
   - Future extensions
   - Outcomes

UEs Generation

Placement

- Arbitrary number of UEs can be initialized.
- UEs distributed uniformly over the grid or placed using:
  - Poisson point process [9]
  - Log-Gaussian Cox process [10]
  - Radial-basis distribution (details in Technical Report Appendix A.1.1)

Birth & Death Process

- M/M/1 queue [11, 12]

UEs Mobility

Mobility - Trajectory Based Motion

- **O/D Traffic-flow data available**:
- Custom Probabilistic Mobility Model (details in Technical Report Appendix A.2.1)
- **Cluster-based Motion**:
  - hybrid-GMM (details in Technical Report Appendix A.2.2)
- **User-based Motion**:
  - Biased Random Walk
  - Random Waypoint (RPW) [13]
  - Lévy flight [14]
  - Truncated Lévy flight [15, 16, 17]
  - Brownian motion

Our testbed

For our tesbed we used **hybrid-GMM**, **Biased Random Walk**, **RPW**.
How these models work?

Mobility models description

RPW

RPW Model is used in simulations to model the movement of nodes. It involves:

- Moving towards a randomly selected waypoint (Uniformly sampled).
- Pausing at waypoints (Pause time uniformly sampled).
- Selecting new waypoints after pausing (Uniformly sampled).

Biased Random Walk

Biased Random Walk limits erratic movement typical of Random Walks, by introducing a biased direction.

hybrid-GMM

hybrid-GMM combines Gauss-Markov Mobility model and RPW. Modelize transiction:

- Temporal-Spatial correlated movement (GMM contribution)
- Pausing time during motion
- Destinations uniformly sampled *

* In our implementation to define way-points, we considered the BS areal density (computed through DBSCAN algorithm [18]).

Table of Contents

1 Cellular Network Framework
- Introduction
- Resource Sharing Mechanism (RSM) choice

2 Problem formulation
- Propagation model
- Objective Function
- Scheduling

3 Power Allocation Strategies
- Introduction to Bayesian Optimization (BO)
- Power Allocation Strategies

4 UEs modilization
- UEs Generation
- UEs Mobility

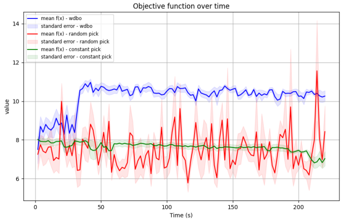
5 Experimental Results
- Objective function comparison
- W-DBO's hyper-parameters comparison

6 Conclusion
- Current limitations
- Future extensions
- Outcomes

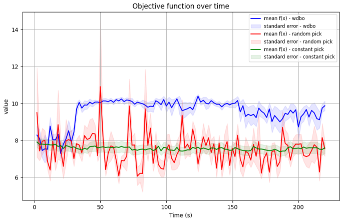
Experimental Results

Obj. Comparison (1/2)

Figure: W-DBO (blue), Random pick (red) and Constant pick (green). α-fairness equal to 1.



(a) Biased Random Walk

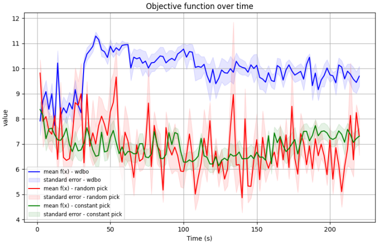


(b) Random Waypoint

Method	**App. W-DBO Time**	**App. Pick Time (Random,Constant)**
Biased Random Walk	85-90 mins	35-40 mins
Random-waypoint	85-90 mins	35-40 mins
Hybrid-GMM	85-90 mins	35-40 mins

Table: Computational Time Comparison - MacBook Pro with a 2.9 GHz Dual-Core Intel Core i5 processor and 16 GB of 2133 MHz LPDDR3 memory. - 2 concurrent threads.

Experimental Results
Obj. Comparison (2/2)

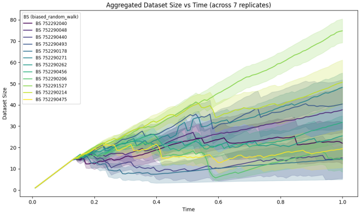


(a) Hybrid-GMM

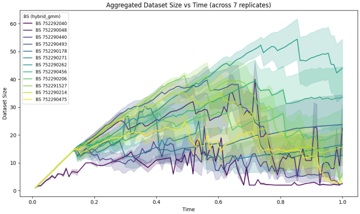
- The Hybrid-GMM model exhibits more dynamic objective function behavior than Biased Random Walk and RPW.
- This is due to moving cluster dynamics, impacting UE density.
- Objective function behavior depends on W-DBO's hyper-parameters, i.e., dataset size, thus temporal lengthscale.
- Spikes in the Random Pick curve result from a higher allocation density of UEs to out-of-border BSs.

Hyper-parameters

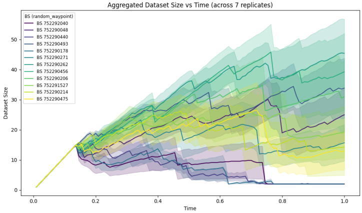
Comparison on disaggregated view - Dataset Size



(a) Biased random walk.



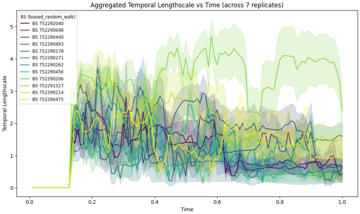
(b) Hybrid-GMM.



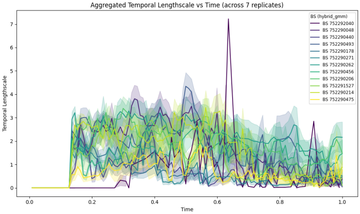
(c) Random waypoint.

Hyper-parameters

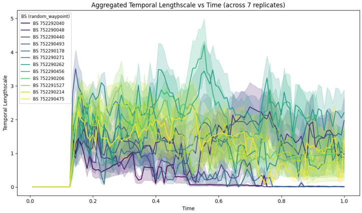
Comparison on disaggregated view - Temporal Lengthscale



(a) Biased random walk.



(b) Hybrid-GMM.



(c) Random waypoint.

Hyper-parameters Discussion

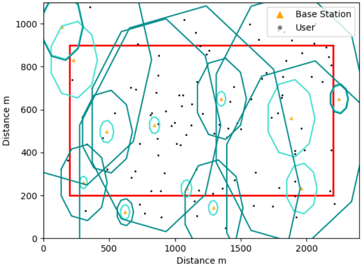
Key points

- Dataset size is driven by the temporal lengthscale curve. If it falls below the threshold (optimization frequency total simulation time $\simeq 1$), uncorrelated observations are considered uncorrelated, hence removed.
- The hybrid-GMM model shows oscillatory trends in dataset size, while RPW and Biased Random Walk models exhibit smoother trends.

Visual tool

Breathing cells visualization

(a) Interactive GIF-visualization. W-DBO power allocation strategy and Biased Random Walk mobility model.



(b) Interactive GIF-visualization. Const-pick power allocation strategy and RPW mobility model.

Figure: Full set of available visualization per motion model and power allocation strategy at NOMA Simulator - Visualizations.

Table of Contents

1 Cellular Network Framework
- Introduction
- Resource Sharing Mechanism (RSM) choice

2 Problem formulation
- Propagation model
- Objective Function
- Scheduling

3 Power Allocation Strategies
- Introduction to Bayesian Optimization (BO)
- Power Allocation Strategies

4 UEs modilization
- UEs Generation
- UEs Mobility

5 Experimental Results
- Objective function comparison
- W-DBO's hyper-parameters comparison

6 Conclusion
- Current limitations
- Future extensions
- Outcomes

Conclusion (1/2)

Current Limitations

- Small-scale simulations with simple motion dynamics did not fully justify the use of W-DBO.
- Simulation duration limited to 220 seconds due to computational constraints.

Future Extensions

- Expand simulations to larger-scale settings with over 30 BSs and more complex cluster-based motion dynamics.
- Integrate real urban mobility traces [19] for more realistic simulations.
- Extend simulation duration to over 30 minutes for better analysis of real-world scenarios.
- Implement a control procedure to exclude simulations with high UE density at outer-border BSs.
- Study a W-DBO variant with power adjustments constrained within a confidence interval.

Conclusion (2/2)

Outcomes

- Developed a simulator to test W-DBO algorithm for power allocation in a dynamic wireless cellular networks.
- Benchmarked W-DBO against Random Pick and Constant Pick strategies.
- Demonstrated that W-DBO consistently outperforms these simpler strategies.

Q&A

References I

France Government, "Arcep Mobile Data - Base station sites," https://files.data.gouv.fr/arcep_donnees/mobile/sites/2022_T3/2022_T3_sites_Metropole.csv, 2022.

R. C. Kizilirmak and H. K. Bizaki, "Non-orthogonal multiple access (noma) for 5g networks," *Towards 5G Wireless Networks-A Physical Layer Perspective*, vol. 83, pp. 83–98, 2016.

S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi, "Successive interference cancellation: A back-of-the-envelope perspective," in *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, 2010, pp. 1–6.

J.-M. Gorce, P. Mary, D. Anade, and J.-M. Kélif, "Fundamental limits of non-orthogonal multiple access (noma) for the massive gaussian broadcast channel in finite block-length," *Sensors*, vol. 21, no. 3, p. 715, 2021.

B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire, "Is noma efficient in multi-antenna networks? a critical look at next generation multiple access techniques," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1310–1343, 2021.

O. Kiss, "arxiv: Bayesian optimization for machine learning algorithms in the context of higgs searches at the cms experiment," Tech. Rep., 2019.

References II

N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting," *IEEE transactions on information theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

A. Bardou, P. Thiran, and G. Ranieri, "This too shall pass: Removing stale observations in dynamic bayesian optimization," 2024. [Online]. Available: https://arxiv.org/abs/2405.14540

F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks, Volume I - Theory*, ser. Foundations and Trends in Networking Vol. 3: No 3-4, pp 249-449, Baccelli, F., Blaszczyszyn, and B., Eds. NoW Publishers, 2009, vol. 1, stochastic Geometry and Wireless Networks, Volume II - Applications; see http://hal.inria.fr/inria-00403040. [Online]. Available: https://inria.hal.science/inria-00403039

J. Møller, A. R. Syversveen, and R. P. Waagepetersen, "Log gaussian cox processes," *Scandinavian journal of statistics*, vol. 25, no. 3, pp. 451–482, 1998.

L. Kleinrock, "Queuing systems, volume i: Theory," 1975.

A. Sankararaman and F. Baccelli, "Spatial birth–death wireless networks," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3964–3982, 2017.

A. Panisson, "Pymobility v0. 1—python implementation of mobility models," 2014.

References III

R. N. Mantegna, "Fast, accurate algorithm for numerical simulation of levy stable stochastic processes," *Physical Review E*, vol. 49, no. 5, p. 4677, 1994.

R. N. Mantegna and H. E. Stanley, "Stochastic process with ultraslow convergence to a gaussian: the truncated lévy flight," *Physical Review Letters*, vol. 73, no. 22, p. 2946, 1994.

——, "Scaling behaviour in the dynamics of an economic index," *Nature*, vol. 376, no. 6535, pp. 46–49, 1995.

B. Dybiec, E. Gudowska-Nowak, E. Barkai, and A. A. Dubkov, "Lévy flights versus lévy walks in bounded domains," *Physical Review E*, vol. 95, no. 5, p. 052102, 2017.

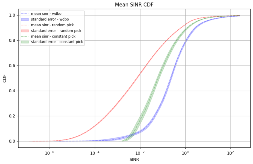
M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *kdd*, vol. 96, no. 34, 1996, pp. 226–231.

D. Zhang, J. Zhao, F. Zhang, and T. He, "Urbancps: A cyber-physical system based on multi-source big infrastructure data for heterogeneous model integration," in *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*, 2015, pp. 238–247.

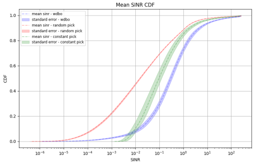
Experimental Results

SINR CDFs comparison

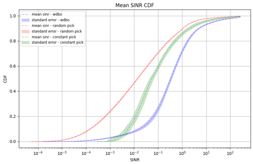
Figure: W-DBO (blue), Random pick (red), and Constant pick (green). α-fairness equal to 1.



(a) Biased Random Walk



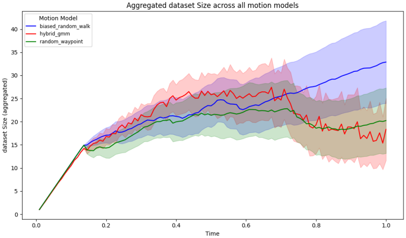
(b) Random Waypoint



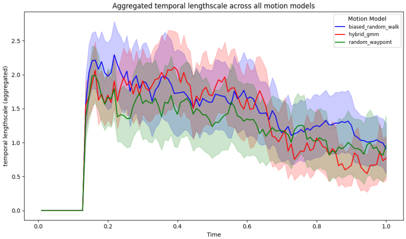
(c) Hybrid-GMM

Hyper-parameters

Comparison on aggregated view - 7 replicates and 12 BSs



(a) Aggregated dataset size metric. Biased Random Walk (blue), hybrid-GMM (red) and RPW (green).



(b) Aggregated temporal lengthscale metric. Biased Random Walk (blue), hybrid-GMM (red) and RPW (green).