Meet-up 2019 | Doctorants & Industrie

ONLINE EXPONENTIAL LEARNING FOR BEAM-ALIGNMENT IN MULTI-USER MILLIMETER WAVE SYSTEMS

Irched CHAFAA[1,2]

E. Veronica Belmega [2], Mérouane Debbah [1,3]

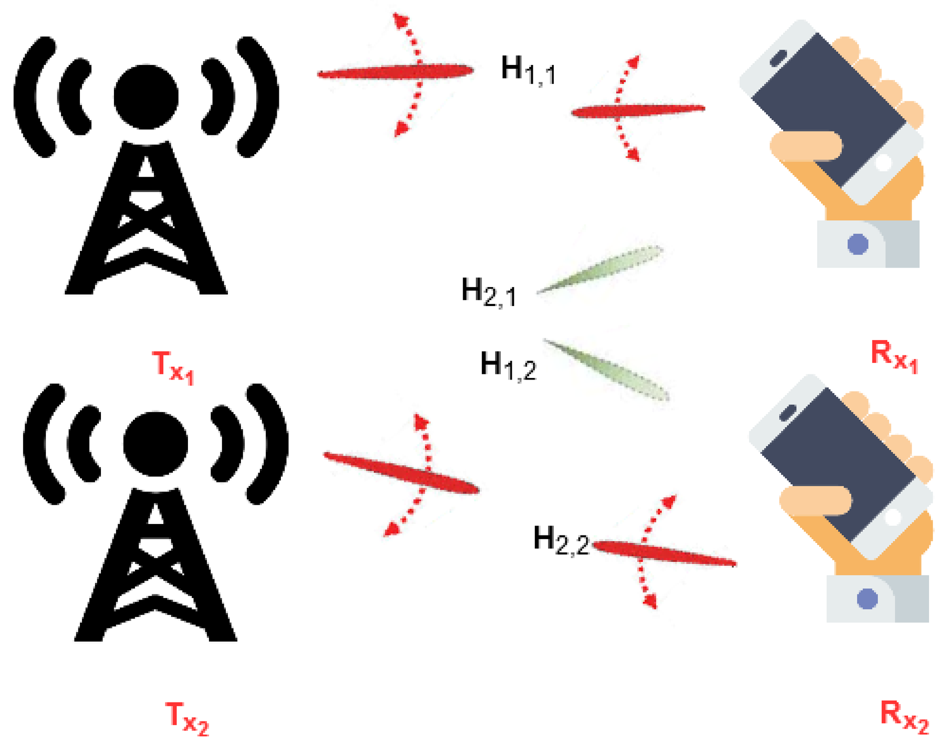
[1] L2S, UMR 8506, CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France
[2] ETIS, UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, Cergy, France
[3] Mathematical and Algorithmic Sciences Lab, Huawei France R&D, Paris, France

1 CONTEXT

Objective: design a distributed beam-alignment policy to support mobility in mmWave systems

- **Application:** mobile mmWave (virtual reality, vehicular communications, wireless HDTV, WLAN, ...)
- **Challenges:** coverage and reliability (severe propagation loss, channel's sparsity, blockage,...). latency (frequent adjusting of beam-directions, large training delay)
- **Tools:** multi-armed bandits (MAB), exponential weights algorithm (EXP3) [1]

2 PROBLEM FORMULATION



- Beamforming vectors $\{f_1, f_2, ..., f_A\}$ and $\{w_1, w_2, ..., w_A\}$ are used at Tx$_k$ and Rx$_k$, $k \in \{1,2\}$
- Received signal at Rx$_k$
  $y_k = \mathbf{w}_j^H \mathbf{H}_{k,k}\mathbf{f}_i s_k + \mathbf{w}_j^H \mathbf{n}_k + \mathbf{w}_j^H \mathbf{H}_{z,k}\mathbf{f}_z s_z,\ z \neq k$
- Channel model [2]
  $\mathbf{H}_{z,k} = \sqrt{\frac{M_T M_R}{\rho_{z,k}}} \sum_{q=1}^{Q} \alpha_q\, \mathbf{a}_{R_k}(\theta_q)\, \mathbf{a}_{T_z}(\phi_q)^H e^{j2\pi\nu_q t}$
- SINR at Rx$_k$ $\quad SINR_k = \frac{|\mathbf{w}_j^H \mathbf{H}_{k,k}\mathbf{f}_i|^2 P_{tr}}{N_0 + |\mathbf{w}_j^H \mathbf{H}_{z,k}\mathbf{f}_z|^2 P_{tr}}$
- System Outage probability
  $P_{out}^{sys} \triangleq P[SINR_1 < \xi_1, SINR_2 < \xi_2]$
- Sum-rate $\ \mathcal{R} \triangleq \sum_{k=1}^{2} \log_2(1 + SINR_k)$
- Research question: Find optimal beam-directions using the MAB framework:
  - Each chosen beam-direction (i_t at Tx or j_t at Rx) represents an action
  - A reward is assigned to the chosen actions

$$r(t) \triangleq r_k(i_t, j_t) \triangleq \begin{cases} 1, \text{ if } SINR_k \geqslant \xi_k, \\ 0, \quad \text{otherwise.} \end{cases}$$

3 PROPOSED BEAM-ALIGNMENT POLICY

Parameters $\eta > 0$, $\beta \geq 1$ and $\gamma \in (0, 1]$
Initialization: Set $G_\ell(0) = 0$, $p_{N,\ell}(0) = 1/A$ and $N \in \{$**Tx**, **Rx**$\}$
Repeat for $t = 1, 2, ..., \mathcal{T}$

- Select action ℓ_t with probability

$$p_{N,\ell_t}(t) = (1-\gamma)\frac{\exp(\eta\, G_{\ell_t}(t-1))}{\sum_{v=1}^{A}\exp(\eta\, G_v(t-1))} + \frac{\gamma}{A}$$

- Update the reward $r_{\ell_t}(t)$
- Update the cumulative rewards

$$\begin{cases} G_\ell(t) = G_\ell(t-1) + \hat{r}_\ell(t) & \ell \in \{\ell_t\} \bigcup V_{\ell_t} \\ \mathrm{G}_\ell(t) = G_\ell(t-1) & \text{otherwise} \end{cases}$$

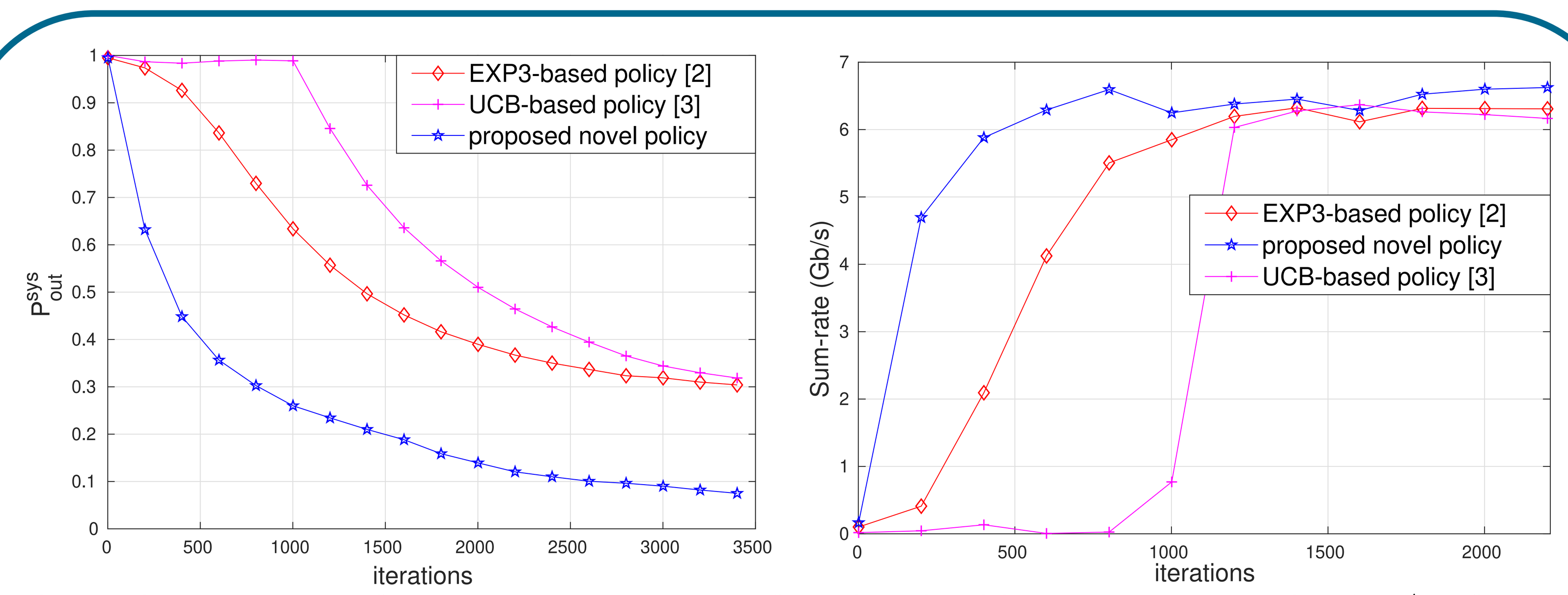
- Update the distribution $p_N(t)$

- Estimated reward vector:

$$\hat{r}_\ell(t) = \begin{cases} \dfrac{(-1)^{r(t)+1}\,\beta^{r(t)}}{1 - r(t) + (-1)^{r(t)+1}\,\hat{p}_{N,\ell}(t)} & if\ \ell = \ell_t \\ \dfrac{\beta\, r(t)}{\hat{p}_{N,\ell}(t)} & if\ \ell \in V_{\ell_t} \\ 0 & otherwise \end{cases}$$

- Neighboring beams of ℓ_t: $V_{\ell_t} \triangleq \{\ell_t - 1, \ell_t + 1\}$

4 SIMULATION RESULTS



Simulations setup: $28GHz$ carrier frequency, $1GHz$ bandwidth, maximum speed $30km/h$, $Q = 1$, $A = 32$, $M_T = 32$, $M_R = 4$, $\xi_1 = \xi_2 = 6dB$ and $\beta = 10$

5 FUTURE WORK

Use tools from the Deep Reinforcement Learning framework and compare both approaches

References

[1] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *36th Annual Symposium on Foundations of Computer Science,*, pages 322–331. IEEE, 1995.

[2] Irched Chafaa, E Veronica Belmega, and Mérouane Debbah. Adversarial multi-armed bandit for mmwave beam alignment with one-bit feedback. *12th EAI International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, 2019.

[3] Jun-Bo Wang, Ming Cheng, Jin-Yuan Wang, Min Lin, Yongpeng Wu, Huiling Zhu, and Jiangzhou Wang. Bandit inspired beam searching scheme for mmwave high-speed train communications. *arXiv preprint arXiv:1810.06150*, 2018.

Ecole doctorale STIC, Université Paris-Saclay
CentraleSupélec
Irched chafaa- irched.chafaa@ensea.fr