Intelligent Power Control for IoT Networks with Semi-Grant-Free NOMA

Muhammad Fayaz, Wenqiang Yi, Yuanwei Liu, Arumugam Nallanathan

Queen Mary University of London, UK

January 18, 2021

Introduction and Background

- Grant-Free (GF) Transmission
  - Users send traffic in arrive and go manner
  - Signalling overhead and latency reduced
  - Grant-free usage examples could be: Power meters, health care
  - Massive access to each sub-channel (Decoding failure)
- Grant-Based (GB) Transmission
  - User subject to handshaking and processing
  - Example: A sensor for health care and critical monitoring subject to overhead
- Semi-Grant-Free (SGF) Transmission
  - Hybrid version of GF and GB Transmission
  - Sub-channel occupied by GB users is always available to GF users
  - Signal overhead reduced by circumventing handshake process

Motivation

- Existing works adapt fixed power control (FPC) strategy with closed-loop thus additional overhead
- Traditional optimization methods, open-loop power control is arduous owing to received power level difference requirement
- In contrast, machine learning (ML) methods can be used to solve NP-hard optimization problems more efficiently [5]

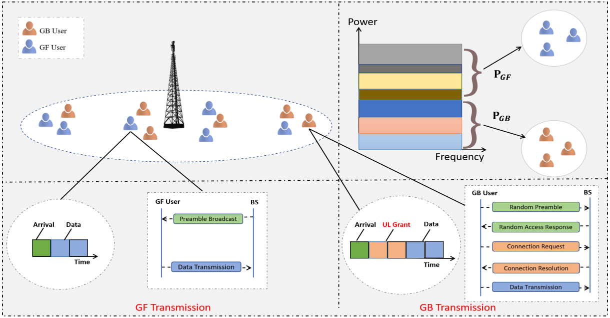
Contributions

- Distributed transmit power control through MA-DRL technique
- Formulation of throughput optimization problem as MDP
- Solving the formulated problem using MA-DRL framework
- Transmit power pool (PP) design for each resource block (RB) for open-loop distributed power control (DPC)
- Comparison with benchmark SFG-NOMA with fixed power control (FPC)
- Comparison with pure GF-NOMA

System Model

- **GF users: U** $= \{1, 2, ...N_{GF}\}$
- **GB users: V** $= \{1, 2, ...N_{GB}\}$
- **Channel gains** of i-th GB and GF users with transmit powers P_i^{GB} and P_j^{GF} :

$$h_i^{GB} = |h_i|^2 (r_{i,GB})^{-\alpha}$$
$$h_j^{GF} = |h_j|^2 (r_{j,GF})^{-\alpha}$$



System Model

- Conventional GB transmission provides limited connectivity and more capacity for IoT scenario than required
- Extra capacity can be utilized to enhance connectivity via GF transmission hence forming SGF transmission
- Channel reserved by GB user via GB protocol
- Available to all GF users to access via GF protocol
- Massive connectivity can be supported hence spectrally efficient

- Precisely, in SGF-NOMA scheme GB and GF users share the same RB for uplink transmission.
- Thus, the combined information received at the BS in a time slot t

$$y(t) = \sum_{m=1}^{M} \sum_{i=1}^{N_{GB,m}} \sqrt{P_{m,i}^{GB}(t)} h_{m,i}^{GB}(t) x_{m,i}(t) + \sum_{m=1}^{M} \sum_{j=1}^{N_{GF,m}} \sqrt{P_{m,j}^{GF}(t)} h_{m,j}^{GF}(t) x_{m,j}(t) + n_0(t)$$

- GB users has the highest priority (e.g, a sensor for healthcare monitoring)
- BS always decodes the GB user in the first stage of SIC
- BS turns to decode GF user with second strongest received power
- Analysis with 1 GB user in an RB, constraint given as:

$$P_{m,1}^{GB} h_{m,1}^{GB}(t) \geq P_{m,1}^{GF} h_{m,1}^{GF}(t) \geq P_{m,2}^{GF} h_{m,2}^{GF}(t) \cdots\cdots \geq P_{m,N_{GF}}^{GF} h_{m,N_{GF}}^{GF}(t)$$

- SINR for i-th GB user on sub-channel m in slot t is:

$$\gamma_{m,i}^{GB}(t) = \frac{P_{m,i}^{GB} h_{m,i}^{GB}(t)}{\sum_{j=1}^{N_{GF}} P_{m,j}^{GF} h_{m,j}^{GF}(t) + n_0^2}, \tag{1}$$

- SINR of j-th GF user can be expressed as

$$\gamma_{m,j}^{GF}(t) = \frac{P_{m,j}^{GF} h_{m,j}^{GF}(t)}{\sum_{j'=j+1}^{N_{GF}} P_{m,j'}^{GF} h_{m,j'}^{GF}(t) + n_0^2}. \tag{2}$$

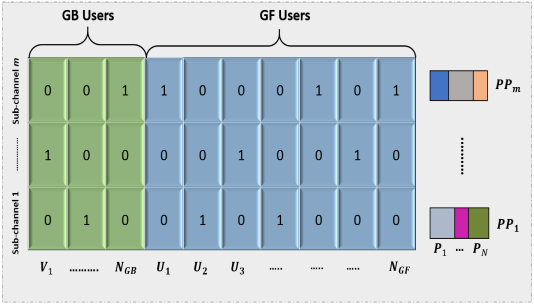
- To guarantee SIC and maintain QoS of GB users following constraint applied:

$$R_{m,i}^{GB}(t) = B_s \log_2(1 + \gamma_{m,i}^{GB}(t)) \geq \tau, \tag{3}$$

τ is the required target data rate. B_s is the bandwidth of sub-channel obtained as $B_s = B/M$, and B is total bandwidth.

- Let Q represent the number of GF users share the same sub-channel with GB user
- The set of Q_m GF users selecting the m-th sub-channel can be expressed as $\mathbf{S_q^m} \subset \mathbf{U} = \{q : 0 \leq q \leq N_{GF}\}$
- To control GF users transmit power and to restrict interference for maintaining uplink QoS of GB users, the central BS has three steps:
  - It calculates the interference threshold ϕ_m that the connected GB user can tolerate on sub-channel m.
  - It maps the threshold ϕ_m to each sub-channel, $\phi_m \xrightarrow{\mathcal{M}} m$.
  - It broadcasts the $\phi = [\phi_1, \phi_2, \cdots\cdots, \phi_m]$ and the PP corresponds to each sub-channel to GF users
- BS broadcasts above information via control signal

- GF users randomly select a transmit power level independently from one of the available PP and transmit on the sub-channel associated with that PP
- To generate PP for each sub-channel, the proposed algorithm can learn a policy that guides all IoT users to adjust their transmit power under some practical constraints



GF and GB users sharing the same sub-channel and PP associated to each sub-channel

Problem Formulation

- The sum rate of the SGF-NOMA IoT networks can be maximized by selecting the sub-channel and optimal transmit power
- By optimizing $\mathbf{P}_{m,j}^{GF}(t)$ and $\mathbf{K}_{m,j}(t)$ for GF users, long-term throughput can be maximized.
- Data rate of a GB IoT user $i \in GB$ over sub-channel m is

$$R_{m,i}^{GB}(t) = B_s \log_2(1 + \gamma_{m,i}^{GB}(t)), \tag{4}$$

- Data rate of GF IoT user $j \in GF$ sharing a sub-channel m with GB user i can be expressed as

$$R_{m,j}^{GF}(t) = B_s \log_2(1 + \gamma_{m,j}^{GF}(t)), \tag{5}$$

Based on (4) and (5) the cumulative capacity can be given as

$$\bar{C} = \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{i=1}^{N_{t,i}} R_{m,j}^{GF}(t) + \sum_{t=1}^{T} \sum_{i=1}^{M} \sum_{j=1}^{N_{t,i}} R_{m,i}^{GB}(t). \tag{6}$$

Problem Formulation

$$\underset{\mathbf{P}_t,\mathbf{K}_t}{\text{maximize}} \quad \bar{C} \tag{7a}$$

$$\text{s.t.} \quad P_{m,1}^{GB} h_{m,1}^{GB}(t) \geq P_{m,1}^{GF} h_{m,1}^{GF}(t) \geq P_{m,2}^{GF} h_{m,2}^{GF}(t) \cdots \geq P_{m,N_{GF}}^{GF} h_{m,N_{GF}}^{GF}(t), \tag{7b}$$

$$\sum_{m=1}^{M} P_{m,j\in GF}(t) \leq P_{max}, \ \forall m, \forall t, \tag{7c}$$

$$\sum_{m=1}^{M} k_{i,j\in GF}(t) \leq 1, \quad \forall j, \forall t, \tag{7d}$$

$$N_{G\in(GB,BF),m}(t) \geq 2, \quad \forall m, \forall t, \tag{7e}$$

$$\sum_{m=1}^{M} R_{m,i}^{GB}(t) \geq \tau, \ \forall i, \forall t, \tag{7f}$$

$$\sum R_{m,j}^{GF}(t) \geq \bar{\tau}, \ \forall j, \forall t, \tag{7g}$$

MA-DRL Framework for SGF-NOMA Networks

- Multiple agents jointly explore the environment.
- In the SGF-NOMA scenario can be modelled as a Markov decision process (MDP) problem.
- Each GF IoT user acts as an agent and interacts with the wireless environment.
- Main elements of the MDP explain below
  - **State space S** : We represent the sum-rate of GF users as current state $s_j(t) \in S_j$ in TS t.

$$S_j = \{R_{1,1}^{GF}(t), R_{2,1}^{GF}(t), \cdots R_{m,j}^{GF}(t), \cdots, R_{M,N_{GF}}^{GF}(t)\} \tag{8}$$

  - **Action Space A** : Action of GF user j consists of transmit power and sub-channel selection.

$$A_j(t) = \{1, 2, \cdots pm, \cdots, PM\}. \tag{9}$$

  - **Reward Engineering** Re : At each TS, every agent receives the system throughput as a reward signal.

$$r_j(t) = \begin{cases} \bar{C}, & \text{if } R^{GF}(t+1) \geq R^{GF}(t) \ , \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

MA-DRL Framework for SGF-NOMA Networks

- We define Q function $Q_j^{\pi}(s_j(t), a_j(t))$ associated with policy π as the expected cumulative discounted reward for each agent j after taking action a_j in state s_j i.e.,

$$Q_j^{\pi}(s_j, a_j) = \mathbb{E}^{\pi}\Big[Re(t)\Big|s_j(t) = s, a_j(t) = a\Big], \tag{11}$$

where Re is the long-term accumulated and discounted reward and calculated as

$$Re = \sum_{k=0}^{K} \alpha^k r^{(t+k+1)}, \qquad 0 < \alpha \leq 1, \tag{12}$$

where α, k and K represents the discount factor, epoch and maximum epoch respectively.

- If an agent j, observe a state $s_j(t)$, perform action $a_j(t)$, receiving a reward $r(t)$ and the next state $s_j(t+1)$, then its corresponding Q-value in the Q-table is updated as

$$Q(s_j(t), a_j(t)) \leftarrow r(t) + \alpha \max_{a_j \in A_j} Q(s_j(t+1), a_j). \tag{13}$$

- Once the optimal Q function $Q^*(s, a)$ is obtained, every agent finds an optimal policy π^* that maximize the reward.
- The classic Q learning maintain a Q-table to record Q-values for each state-action pair.
- However, for IoT scenario, the size of Q-table increases with the increasing number of state-action spaces that makes Q-learning expensive in terms of memory and computation.
- Deep Q learning overcome the this problem by combining the Q learning with Deep Neural Network (DNN) with weights θ for Q function approximation $Q(s, a; \theta)$
- In MA-DRL each agent consists of primary (online) network, target network and a replay memory.

MA-DRL Framework for SGF-NOMA Networks

- In a TS t, each agent j input the current state s_j to the DQN and output Q-values corresponding to all actions.
- The agent select action with the highest Q-value and obtain an experience in the form of a tuple $\big(s_j(t), a_j(t), r(t), s_j(t+1)\big)$ and store it to replay memory
- To update the weights θ of the target Q-network, a mini-batch of data is randomly sampled from the replay memory to produce target Q-value as

$$y_j(t) = r(t) + \alpha \max_{a_j(t+1) \in A_j} Q(s_j(t+1), a_j(t+1); \theta). \tag{14}$$

- Using a variant of stochastic gradient descent (SGD), the primary Q-network can be trained by minimizing the loss function

$$Loss(\theta) = (y_j(t) - Q_j(t)(s_j(t), a_j(t); \theta))^2. \tag{15}$$

Numerical Results

Simulation Parameters

Cell radius	$1000m$
Power Levels	[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 0.9] W
Path loss exponent	$\alpha = 3.0$
AWGN (N_0)	-90dBm
No. of sub-channels	3
sub-channel Bandwidth	$B_s = 10$ KHz
User minimum data rate	10 bps/Hz
P_{max}	1 W
No. of Episodes	500
Layers	{Input, hidden layer1, hidden layer2, hidden layer3, output}
Neurons per layer	{250, 120, 64}
α, ϵ, Learning rate	0.9, 1.0, 0.001
Optimizer	Adam

- Fig. 1 depicts performance gain of proposed SGF-NOMA with DPC as compared to pure GF-NOMA and SGF with FPC
- Performance gain owes to allocation of distributed power to additional GF users
- GF users cause interference with in a tolerable threshold
- All GF users are allowed to transmit data creating strong interference on the sub-channels thus throughput reduction

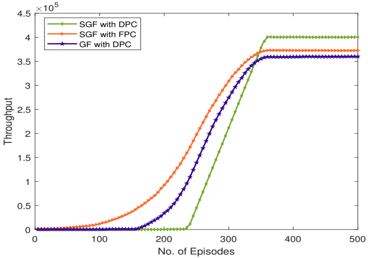
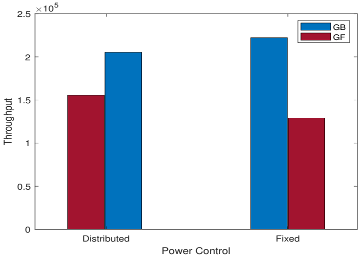


Figure 1: Performance Comparison

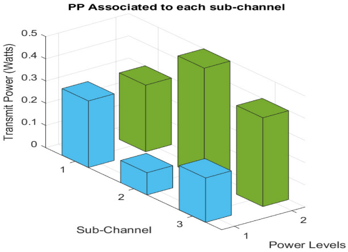
Numerical Results

DPC Impact on GB and GF users Throughput

- With DPC, throughput of GF users increased while that of GB decreased
- Increase and decrease owes to allocation of extra data rate of GB users to GF users
- GF users transmitting with same transmit power (FPC) diminished the throughput
- SGF-NOMA with FPC is unable to utilize extra capacity of GB users fairly and efficiently



- 03 sub-channels pre-occupied by GB users and available to GF users for uplink transmission utilized
- Optimal transmit power levels ascertained and mapped to corresponding sub-channels
- GF users select sub-channel and transmit power randomly from corresponding PP
- Proposed SGF-NOMA provides open-loop power control with less signalling overhead and latency



Conclusion

- We have proposed MA-DRL based SGF-NOMA algorithm with DPC to design and map PP to each sub-channel enabling open loop power control
- The MA-DRL based SGF-NOMA with DPC has been proved to provide better performance as compared to the algorithm with the FPC mechanism
- Additionally, the proposed scheme outperforms the pure GF networks in terms of throughput

- Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan and L. Hanzo, “Nonorthogonal Multiple Access for 5G and Beyond,” Proc. IEEE, vol. 105, pp. 2347–23811, 2017.
- Z. Ding, R. Schober, P. Fan and H. V. Poor, “Simple Semi-Grant-Free Transmission Strategies Assisted by Non-Orthogonal Multiple Access,” IEEE Trans. Commun., vol. 19, pp. 4464–4478, 2019.
- C. Zhang, Y. Liu, W. Yi, Z. Qin and Z. Ding, “Semi-Grant-Free NOMA: Ergodic Rates Analysis with Random Deployed Users,”IEEE Wireless Commun. Lett., pp. 1–1, 2020.
- W. Yi, W. Yu, Y. Liu, C. H. Foh, Z. Ding, and A. Nallanathan, “Multiple Transmit Power Levels based NOMA for Massive Machine-type Communications,”arXiv preprint arXiv:2011.12388, 2020.
- R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction, MIT press, 2018.

Thanks