Joint Resource Allocation and Link Association of URLLC Puncturing eMBB Traffic in Multicell Networks

Phong-Binh Tran

Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan

September 5, 2022

Supervised by Chair Professor Jang-Ping Sheu.

Scenario

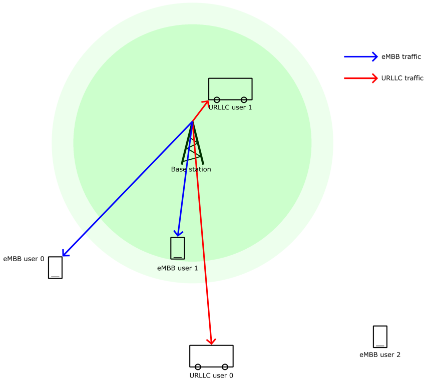

Figure: Singlecell model

- There are approximately 750 people per 1000 square meters living in suburban area[1][3].
- These are potential eMBB users, who surf the Web, watching videos, and download data.
- During work hours and at night, only a few self-driving vehicles operate that employ URLLC utilities.
  - Uplink transmission (whose bandwidth is independent from that of downlink) is used to upload the vehicles' observations e.g. camera images, sensors data, etc. to the cloud for navigation processing.
  - Downlink transmission accounts for the automobiles' control messages.

[1]Example

System

- Homogeneous base stations, mmWave, downlink transmission, OFDMA, multiple-input eMBB and single-input URLLC users.
- Saturated eMBB traffic [5]: Each eMBB user has infinite amount of data to be served.
- Strict URLLC constraint: Each URLLC has an amount of data required to be served within a minislot.
- The system aims to maximize eMBB total average rate and fairness while satisfying URLLC demands.

Poor Edge Service

- Since mmWave is extremely vulnerable to path loss, URLLC reliability is not guaranteed.
- Similarly, eMBB users at cell edges experience low throughput.

- URLLC multicell and eMBB multiconnectivity are prominent candidates to mitigate this issue.

Multicell

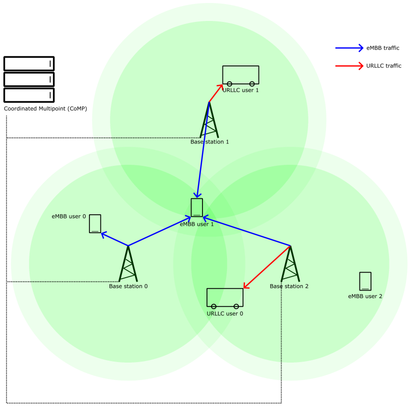

Figure: Multicell model

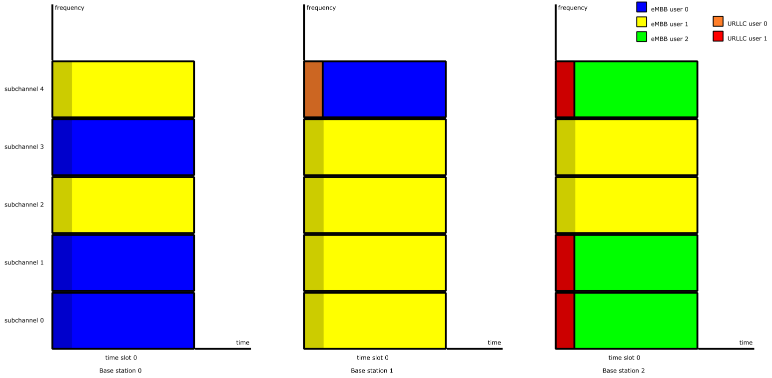

Figure: Multicell framework

Multicell Co-channel Interference

- Since the base stations are homogeneous i.e. use the same frequency band, there exists 3 types of interference:
  - eMBB-eMBB inteference e.g. at subchannel 3, signal from base station 0 to eMBB user 0 interferes with that from base station 1 to eMBB user 1.
  - eMBB-URLLC interference e.g. at subchannel 0, signal from base station 0 to eMBB user 0 interferes with that from base station 2 to URLLC user 1.
  - URLLC-URLLC interference e.g. at subchannel 4, signal from base station 1 to URLLC user 0 interferes with that from base station 2 to URLLC user 1.

- A viable solution might be 5G Non-orthogonal Multiple Access (NOMA) with Successive Interference Cancellation (SIC).
- Inspired by Low Energy Adaptive Clustering Hierarchy (LEACH), a modified Code Division Multiple Access (CDMA) scheme is proposed to tackle the problem[2].

[2] Example

- Our scheme works well with the often small number of base stations.
- Our scheme encompasses URLLC multiconnectivity via joint transmission.
- This hence introduces a joint CDMA/OFDMA scheme.

Spectrum Inefficiency

- Dedicated URLLC bandwidth wastes spectral resources significantly in multicell systems.
  - If 2 subchannels of each base station are dedicated to URLLC traffic, then we would have 6 subchannels sitting idle for most of the time in the aforementioned scenario.

- This problem can be addressed by leveraging URLLC superposition/puncturing scheme.
- URLLC superposition scheme employs 5G NOMA SIC, whose performance equals to puncturing when the considered eMBB and URLLC users have the same channel gain.
- URLLC puncturing scheme is discussed here.

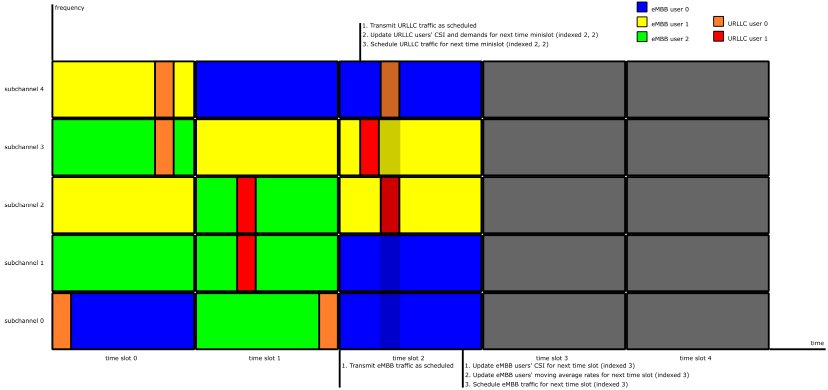

Figure: Singlecell framework

Stochastic Problem

$$
\begin{aligned}
\underset{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}}{\text{maximize}} \quad & \sum_u \ln \bar{R}_u && (1a)\\
\text{subject to} \quad & \sum_u \alpha_{u,n,s,l} \leq 1 \quad \forall n \forall s \forall l, && (1b)\\
& \alpha_{u,n,s,l} \in \{0,1\} \quad \forall u \forall n \forall s \forall l, && (1c)\\
& \sum_s \delta_{v,n,m,s} \leq 1 \quad \forall v \forall n \forall m, && (1d)\\
& \beta_{v,u,n,m,s,l} \leq \delta_{v,n,m,s} \forall v \forall u \forall n \forall m \forall s \forall l, && (1e)\\
& \delta_{v,n,m,s} \in \{0,1\} \quad \forall v \forall n \forall m \forall s, && (1f)\\
& \sum_v \beta_{v,u,n,m,s,l} \leq \alpha_{u,n,s,l} \forall u \forall n \forall m \forall s \forall l, && (1g)\\
& R_{v,n,m} \geq R^{dm}_{v,n,m} \quad \forall v \forall n \forall m, && (1h)\\
& \beta_{v,u,n,m,s,l} \in \{0,1\} \quad \forall v \forall u \forall n \forall m \forall s \forall l && (1i)
\end{aligned}
$$

- The system maximizes eMBB traffic's total average rate and fairness (1a).
- For each time slot, the system
  - schedules a subchannel to at most one eMBB user (1b).
  - either un-schedules or schedules a subchannel to an eMBB user (1c).

- For each time minislot, the system
    - associates at most one base station to a URLLC user (1d).
    - schedules a subchannel to a URLLC user only if it associates the corresponding base station to the user (1e).
    - either un-associates or associates a base station to a URLLC user (1f).
    - schedules a subchannel to at most one URLLC user, and punctures the subchannel for a URLLC user only if it schedules the subchannel to the corresponding eMBB user (1g)[3].
    - serves demands of URLLC users without delays (1h).
    - employs URLLC puncturing scheme instead of superposition (1i).

[3] Proof in supplementary

- Do note that current eMBB users' rate models for URLLC puncturing scheme in the literature are mostly non-linear [2].
- Whilst proposed linear models are either intractable [6] or inappropriate [1] for discrete subchannel scheduling with multiple URLLC users.

eMBB Problem

$$\underset{\boldsymbol{\alpha}}{\text{maximize}} \quad \sum_{u} \ln \bar{R}_u \tag{2a}$$

$$\text{subject to} \quad \sum_{u} \alpha_{u,n,s,l} \leq 1 \qquad \forall n \forall s \forall l, \tag{2b}$$

$$\alpha_{u,n,s,l} \in \{0,1\} \forall u \forall n \forall s \forall l \tag{2c}$$

Relaxed eMBB Problem

$$\underset{\boldsymbol{\alpha}'}{\text{maximize}} \quad \sum_{u} \ln \bar{R}'_u \tag{3a}$$

$$\text{subject to} \quad \sum_{u} \alpha'_{u,n,s,l} \leq 1 \forall n \forall s \forall l, \tag{3b}$$

$$\alpha'_{u,n,s,l} \geq 0 \forall u \forall n \forall s \forall l \tag{3c}$$

Gradient Problem

$$\underset{\boldsymbol{\alpha}'_{n_0}}{\text{maximize}} \quad \sum_{u} \frac{r'_{u,n_0}}{\tilde{r}'_{u,n_0}} \tag{4a}$$

$$\text{subject to} \quad \sum_{u} \alpha'_{u,n_0,s,l} \leq 1 \forall s \forall l, \tag{4b}$$

$$\alpha'_{u,n_0,s,l} \geq 0 \forall u \forall s \forall l \tag{4c}$$

- The relaxed moving average rate of eMBB user is defined based on exponential moving average (EMA) as

$$\forall u\colon\ \tilde{r}'_{u,n} = \begin{cases} \frac{1}{\mathfrak{n}} \sum_{s,l} \frac{1}{\mathfrak{u}} \mathfrak{r}_{u,n,s,l} & n = 0 \\ (1-\epsilon)\, \tilde{r}'_{u,n-1} + \epsilon r'_{u,n-1} & n = 1, \ldots, \mathfrak{n}-1 \end{cases} \left[\frac{bits}{slot}\right]. \tag{5}$$

- The initial value of which is defined by the feasible policy $\hat{\alpha}'$ for the relaxed eMBB problem where[4]

$$\forall u\colon \forall s\colon \forall l\colon \hat{\alpha}'_{u,n,s,l} = \begin{cases} \frac{1}{\mathfrak{u}} & n = 0 \\ 0 & n = 1, \ldots, \mathfrak{n} - 1 \end{cases}. \qquad (6)$$

[4] Proof in supplementary

- Given a policy $\hat{\alpha}$ where for $n = 0, \ldots, \mathfrak{n} - 1$, $\hat{\alpha}_n$ is a basic optimal point for the n^{th} gradient problem, then $\hat{\alpha}$ is an asymptotically optimal policy for the eMBB problem[5].

[5] Proof in supplementary

- However, simplex methods run in exponential time in worst-case scenarios.

- The following policy is asymptotically optimal with respect to the eMBB problem[6]:

$$\forall n\colon \forall s\colon \forall l\colon \hat{\alpha}_{u,n,s,l} = \begin{cases} 1 & u = \min_{\hat{u}} \arg\max_{\hat{u}} \frac{\mathfrak{r}_{\hat{u},n,s,l}}{\tilde{r}'_{\hat{u},n}} \\ 0 & \text{otherwise} \end{cases} . \tag{7}$$

[6] Proof in supplementary

- This is a direct generalization of the proportional fairness algorithm [4] in multiconnectivity-based networks.

URLLC Problem

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0}, \boldsymbol{\delta}_{n_0}}{\text{maximize}} \quad & \sum_{u} \frac{r'_{u,n_0}}{\tilde{r}'_{u,n_0}} && \text{(8a)} \\
\text{subject to} \quad & \sum_{s} \delta_{v,n_0,m,s} \leq 1 \qquad \forall v \forall m, && \text{(8b)} \\
& \beta_{v,u,n_0,m,s,l} \leq \delta_{v,n_0,m,s} \forall v \forall u \forall m \forall s \forall l, && \text{(8c)} \\
& \delta_{v,n_0,m,s} \in \{0,1\} \quad \forall v \forall m \forall s, && \text{(8d)} \\
& \sum_{v} \beta_{v,u,n_0,m,s,l} \leq \hat{\alpha}_{u,n_0,s,l} \forall u \forall m \forall s \forall l, && \text{(8e)} \\
& R_{v,n_0,m} \geq R^{dm}_{v,n_0,m} \; \forall v \forall m, && \text{(8f)} \\
& \beta_{v,u,n_0,m,s,l} \in \{0,1\} \quad \forall v \forall u \forall m \forall s \forall l && \text{(8g)}
\end{aligned}
$$

Puncturing Problem

$$
\begin{aligned}
\underset{\boldsymbol{\beta}_{n_0,m_0},\,\boldsymbol{\delta}_{n_0,m_0}}{\text{minimize}} \quad & \sum_{v,u,s,l} \frac{\mathfrak{r}_{u,n_0,s,l}}{\tilde{r}'_{u,n_0}} \beta_{v,u,n_0,m_0,s,l} && (9a)\\
\text{subject to} \quad & \sum_{s} \delta_{v,n_0,m_0,s} \le 1 \qquad \forall v, && (9b)\\
& \beta_{v,u,n_0,m_0,s,l} \le \delta_{v,n_0,m_0,s} \forall v \forall u \forall s \forall l, && (9c)\\
& \delta_{v,n_0,m_0,s} \in \{0,1\} \quad \forall v \forall s, && (9d)\\
& \sum_{v} \beta_{v,u,n_0,m_0,s,l} \le \hat{\alpha}_{u,n_0,s,l} \quad \forall u \forall s \forall l, && (9e)\\
& r_{v,n_0,m_0} \ge \mathfrak{r}^{dm}_{v,n_0,m_0} \quad \forall v, && (9f)\\
& \beta_{v,u,n_0,m_0,s,l} \in \{0,1\} \quad \forall v \forall u \forall s \forall l && (9g)
\end{aligned}
$$

- Given a time slot policy $(\hat{\boldsymbol{\beta}}_{n_0}, \hat{\boldsymbol{\delta}}_{n_0})$ where for $m = 0, \ldots, \mathfrak{m}-1$, $(\hat{\boldsymbol{\beta}}_{n_0,m}, \hat{\boldsymbol{\delta}}_{n_0,m})$ solves the m^{th} puncturing problem, then $(\hat{\boldsymbol{\beta}}_{n_0}, \hat{\boldsymbol{\delta}}_{n_0})$ is an optimal point for the URLLC problem[7].

[7] Proof in supplementary

Scheduling Problem

$$\begin{aligned}
\underset{\boldsymbol{\gamma}_{n_0,m_0},\,\boldsymbol{\delta}_{n_0,m_0}}{\text{minimize}} \quad & \sum_{v,s,l} c_{n_0,s,l}\gamma_{v,n_0,m_0,s,l} && && \text{(10a)}\\
\text{subject to} \quad & \sum_{s} \delta_{v,n_0,m_0,s} \leq 1 && \forall v, && \text{(10b)}\\
& \gamma_{v,n_0,m_0,s,l} \leq \delta_{v,n_0,m_0,s} && \forall v \forall s \forall l, && \text{(10c)}\\
& \delta_{v,n_0,m_0,s} \in \{0,1\} && \forall v \forall s, && \text{(10d)}\\
& \sum_{v} \gamma_{v,n_0,m_0,s,l} \leq 1 && \forall s \forall l, && \text{(10e)}\\
& \sum_{s,l} \gamma_{v,n_0,m_0,s,l}\mathfrak{r}_{v,n_0,m_0,s} \geq \mathfrak{r}^{dm}_{v,n_0,m_0} && \forall v, && \text{(10f)}\\
& \gamma_{v,n_0,m_0,s,l} \in \{0,1\} && \forall v \forall s \forall l && \text{(10g)}
\end{aligned}$$

- The puncturing cost of subchannel is defined as

$$\forall n\colon \forall s\colon \forall l\colon\ c_{n,s,l} = \max_{u} \frac{\mathfrak{r}_{u,n,s,l}}{\tilde{r}'_{u,n}}. \tag{11}$$

- The scheduling problem is equivalent to the puncturing problem[8].

[8] Proof in supplementary

- The scheduling problem is independent from α.

Homogeneous Problem

$$
\begin{aligned}
\underset{\gamma_{n_0,m_0},\, \boldsymbol{\delta}_{n_0,m_0}}{\text{minimize}} \quad & c_{n_0}^{mn} \sum_{v,s,l} \gamma_{v,n_0,m_0,s,l} && \text{(12a)}\\
\text{subject to} \quad & \sum_{s} \delta_{v,n_0,m_0,s} \leq 1 \qquad \forall v, && \text{(12b)}\\
& \gamma_{v,n_0,m_0,s,l} \leq \delta_{v,n_0,m_0,s} \forall v \forall s \forall l, && \text{(12c)}\\
& \delta_{v,n_0,m_0,s} \in \{0,1\} \qquad \forall v \forall s, && \text{(12d)}\\
& \sum_{v} \gamma_{v,n_0,m_0,s,l} \leq 1 \qquad \forall s \forall l, && \text{(12e)}\\
& \sum_{s,l} \gamma_{v,n_0,m_0,s,l} \geq \frac{\mathfrak{r}_{v,n_0,m_0}^{dm}}{\mathfrak{r}_{v,n_0,m_0}^{mx}} \quad \forall v, && \text{(12f)}\\
& \gamma_{v,n_0,m_0,s,l} \in \{0,1\} \qquad \forall v \forall s \forall l && \text{(12g)}
\end{aligned}
$$

- The minimum puncturing cost is defined as

$$\forall n\colon\ c_n^{mn} = \min_{s,l} c_{n,s,l}. \tag{13}$$

- The maximum rate of URLLC user is defined as

$$\forall n\colon \forall m\colon \forall v\colon\ \mathfrak{r}_{v,n,m}^{mx} = \max_{s} \mathfrak{r}_{v,n,m,s}. \tag{14}$$

- $c_{n_0}^{mn} \sum_v \left\lceil \frac{\mathfrak{r}_{v,n_0,m_0}^{dm}}{\mathfrak{r}_{v,n_0,m_0}^{mx}} \right\rceil$ is the optimal value for the homogeneous problem.

- The nearest association algorithm defines the following point:

$$\forall v\colon \hat{\delta}_{v,n_0,m_0,s} = \begin{cases} 1 & s = \min_{\hat{s}} \arg\max_{\hat{s}} \mathfrak{r}_{v,n_0,m_0,\hat{s}} \\ 0 & \text{otherwise} \end{cases},$$

$$\forall s\colon \forall v\colon \hat{\gamma}_{v,n_0,m_0,s,l} = \begin{cases} 1 & \begin{aligned} &\sum_{\hat{v}=0}^{v-1} \left\lceil \frac{\mathfrak{r}^{dm}_{\hat{v},n_0,m_0}}{\mathfrak{r}^{mx}_{\hat{v},n_0,m_0}} \right\rceil \hat{\delta}_{\hat{v},n_0,m_0,s} \\ &\leq l \\ &\leq \left(\sum_{\hat{v}=0}^{v} \left\lceil \frac{\mathfrak{r}^{dm}_{\hat{v},n_0,m_0}}{\mathfrak{r}^{mx}_{\hat{v},n_0,m_0}} \right\rceil \hat{\delta}_{\hat{v},n_0,m_0,s} \right) - 1 \end{aligned} \\ 0 & \text{otherwise} \end{cases}.$$

- The nearest association algorithm is a $\frac{c_{n_0}^{mx}}{c_{n_0}^{mn}}$-approximation algorithm for the scheduling problem[9].

[9] Proof in supplementary

- The maximum puncturing cost is defined as

$$\forall n \colon c_n^{mx} = \max_{s,l} c_{n,s,l}. \tag{15}$$

- The opportunistic nearest association algorithm defines the following point:

$$\forall v\colon \hat{\delta}_{v,n_0,m_0,s} = \begin{cases} 1 & s = \min_{\hat{s}} \arg\max_{\hat{s}} \mathfrak{r}_{v,n_0,m_0,\hat{s}} \\ 0 & \text{otherwise} \end{cases},$$

$$\forall s\colon \forall v\colon \hat{\gamma}_{v,n_0,m_0,s,l} = \begin{cases} 1 & \begin{aligned} &\left(\sum_{\hat{v}=0}^{v-1} \left\lceil \frac{\mathfrak{r}^{dm}_{\hat{v},n_0,m_0}}{\mathfrak{r}^{mx}_{\hat{v},n_0,m_0}} \right\rceil \hat{\delta}_{\hat{v},n_0,m_0,s}\right) + 1 \\ &\leq \mathbf{card}\left\{\hat{l} \,\middle|\, c_{n_0,s,l} > c_{n_0,s,\hat{l}} \vee \left(c_{n_0,s,l} = c_{n_0,s,\hat{l}} \wedge l > \hat{l}\right)\right\} \\ &\leq \sum_{\hat{v}=0}^{v} \left\lceil \frac{\mathfrak{r}^{dm}_{\hat{v},n_0,m_0}}{\mathfrak{r}^{mx}_{\hat{v},n_0,m_0}} \right\rceil \hat{\delta}_{\hat{v},n_0,m_0,s} \end{aligned} \\ 0 & \text{otherwise} \end{cases}.$$

- The opportunistic nearest association algorithm is an optimal algorithm for the scheduling problem in single-cell networks[10].

[10] Proof in supplementary

Contributions

References

[1] Arjun Anand, Gustavo de Veciana, and Sanjay Shakkottai. "Joint Scheduling of URLLC and eMBB Traffic in 5G Wireless Networks". In: *IEEE/ACM Transactions on Networking* 28.2 (2020), pp. 477–490. DOI: 10.1109/TNET.2020.2968373.

[2] Anupam Kumar Bairagi et al. "Coexistence Mechanism Between eMBB and URLLC in 5G Wireless Networks". In: *IEEE Transactions on Communications* 69.3 (2021), pp. 1736–1749. DOI: 10.1109/TCOMM.2020.3040307.

[3] Richard Florida. *How Should We Define the Suburbs?* 2019. URL: https://www.bloomberg.com/news/articles/2019-06-12/why-we-need-a-standard-definition-of-the-suburbs (visited on 07/13/2022).

[4] Harold J. Kushner and Philip A. Whiting. "Asymptotic Properties of Proportional-Fair Sharing Algorithms". In: *40th Annual Allerton Conference on Communication, Control, and Computing*. 2002, pp. 1051–1059.

[5] Alexander L. Stolyar. "On the Asymptotic Optimality of the Gradient Scheduling Algorithm for Multiuser Throughput Allocation". In: *Operations Research* 53.1 (2005), pp. 12–25. DOI: 10.1287/opre.1040.0156.

[6] Hao Yin, Lyutianyang Zhang, and Sumit Roy. "Multiplexing URLLC Traffic Within eMBB Services in 5G NR: Fair Scheduling". In: *IEEE Transactions on Communications* 69.2 (2021), pp. 1080–1093. DOI: 10.1109/TCOMM.2020.3035582.