Deep Learning Based PRACH load estimation for future mMTC scenarios

Luciano Miuccio, Daniela Panno, Salvatore Riolo

GTTI Workshop

Wireless Intelligence: From Reconfigurable Surfaces to Edge/Cloud Communications

March 26th, 2021

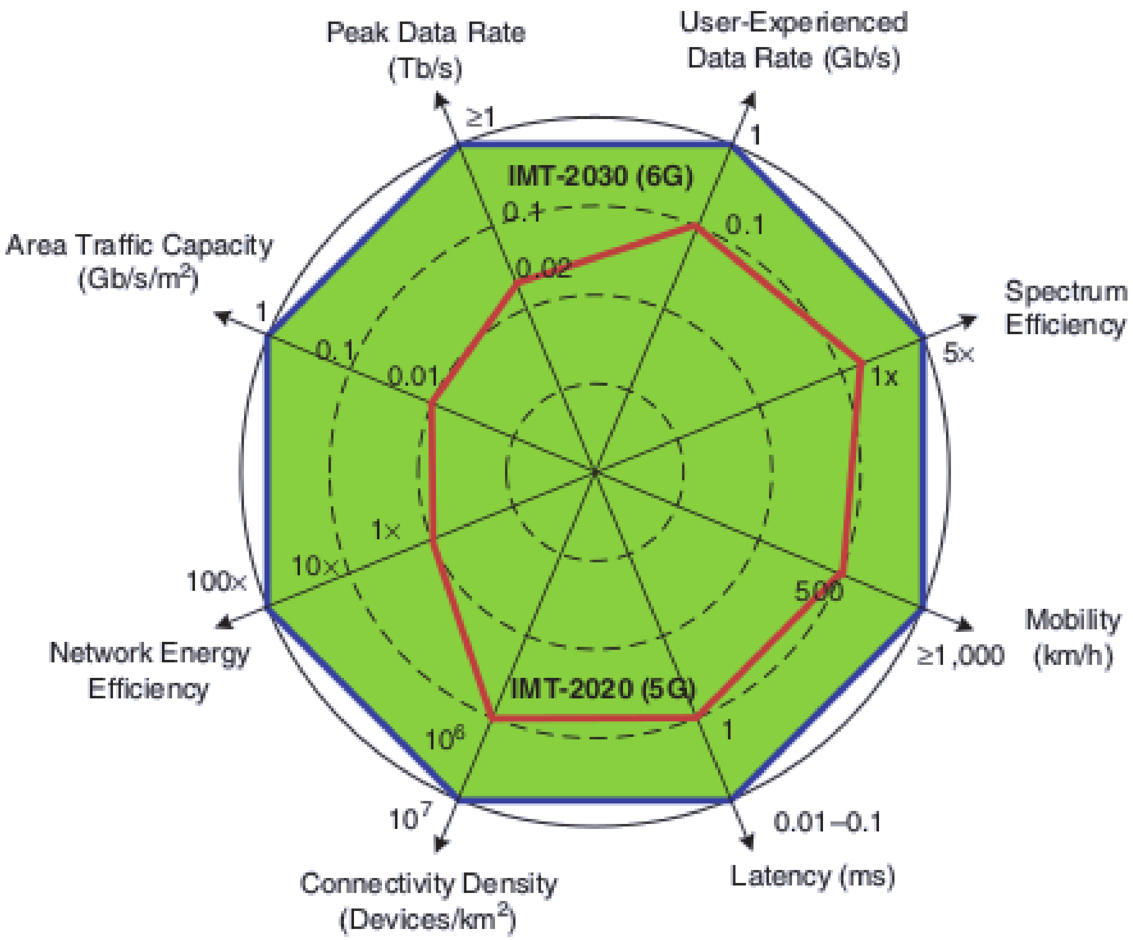
Università di Catania

DIEEI

Dipartimento di Ingegneria Elettrica, Elettronica e Informatica

(Department of Electrical, Electronics and Computer Engineering)

5G and 5GB/6G requirements comparison



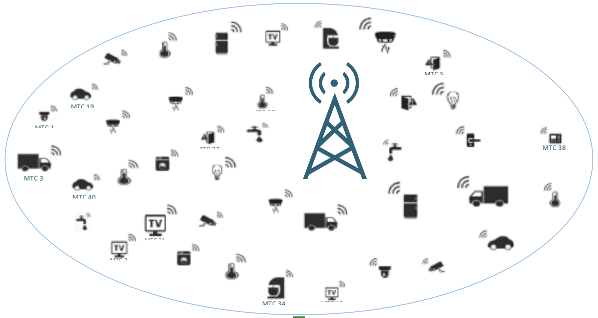
- **Connectivity Density (Devices/km²):** From 10^6 to 10^7.
- **Spectrum efficiency:** 5x in 6G with respect to 5G.
- **Network Energy Efficiency:** 100x in 6G compared to 5G.

Usage of tailored random access control schemes for avoiding network collapse.

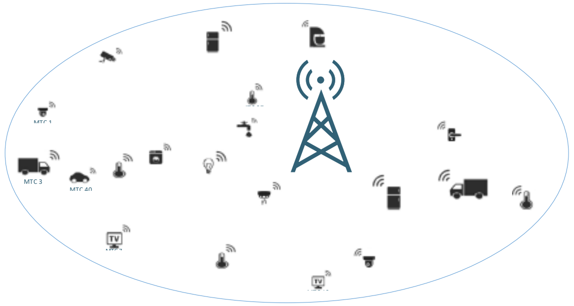
Z. Zhang *et al.*, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," in *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28-41, Sept. 2019.

Access control schemes for 5GB/6G

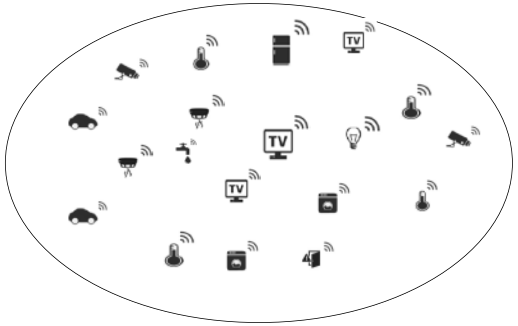
Access Class Barring (ACB) schemes



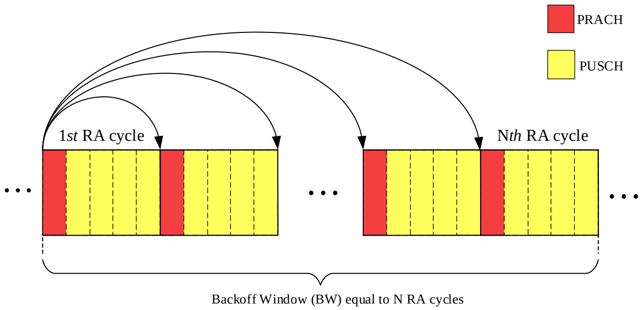
Reduction of the number of attempting MTC devices barring the access to some of them.



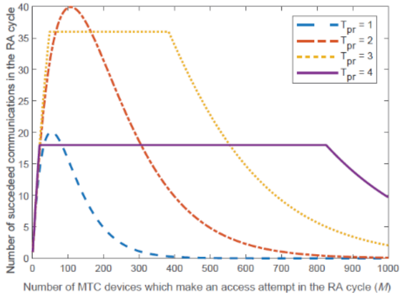
Dynamic Backoff Indicator schemes



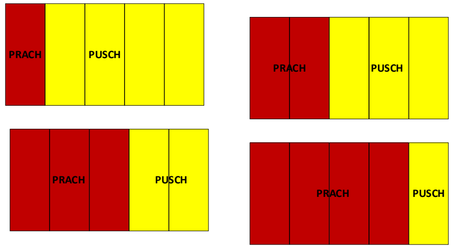
Spread the access requests of the MTC devices in backoff state inside a proper backoff window.



Dynamic Uplink Resource Dimensioning schemes

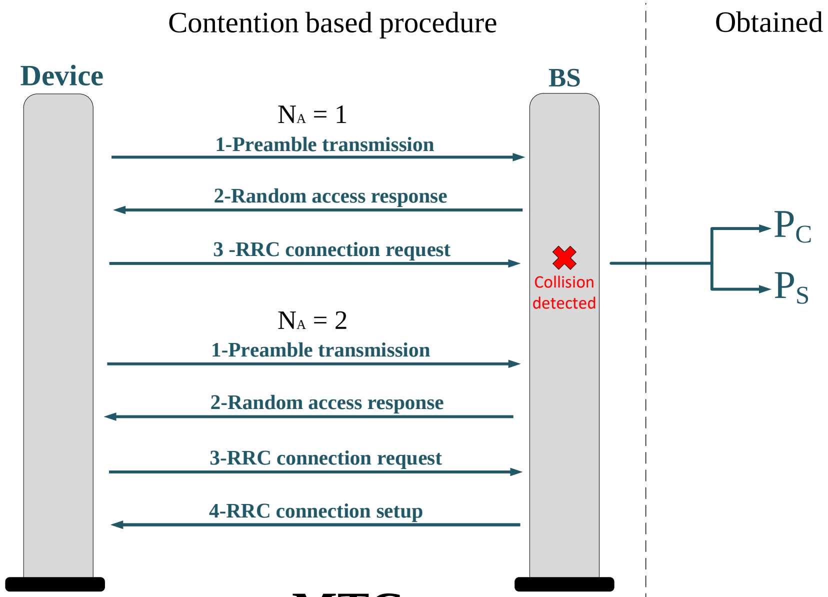


Optimal Dimensioning of the uplink resources between access and transmission resources.

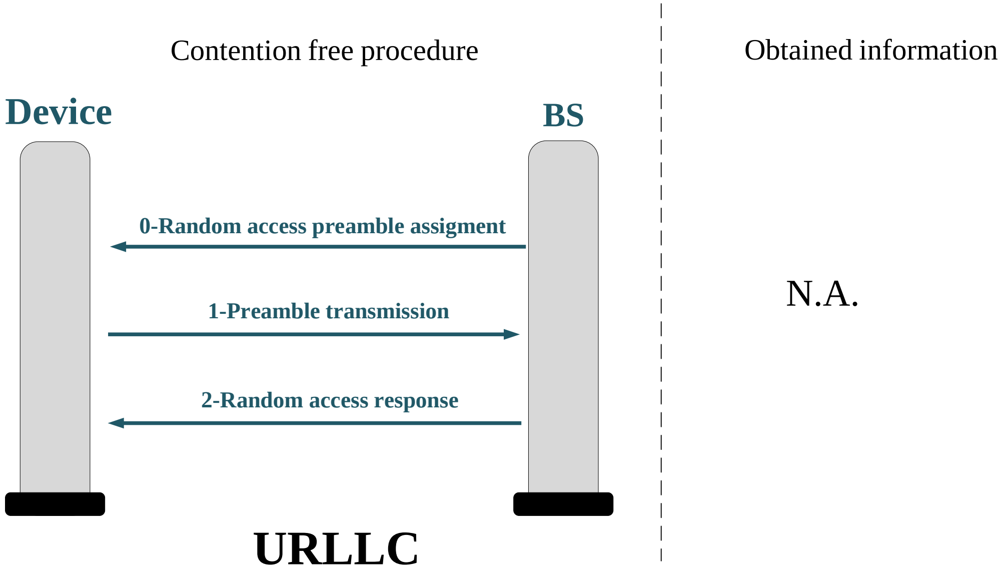


To function properly, these schemes need to have information about the current number of attempting access requests (M)

Contention based procedure vs Contention free procedure for traffic load estimation



mMTC



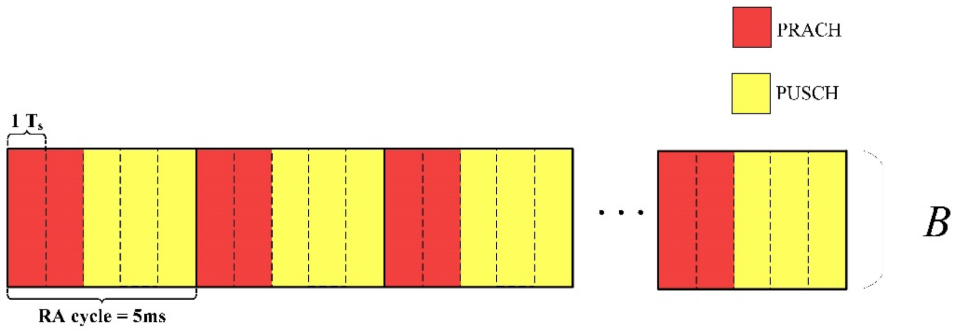
URLLC

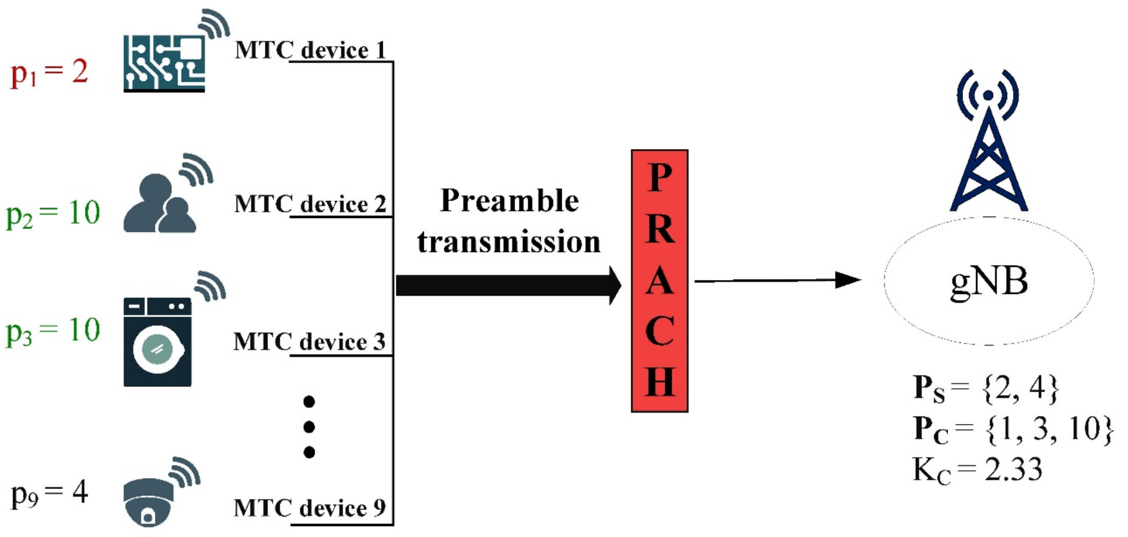
For the estimation of the traffic load in massive MTC scenarios the usage of a Contention based-RA procedure provides important information:

1. the number of collided preambles (P_C);

2. the number of succeeded preambles (P_S).

Traffic load (M) estimation problem formulation





$\mathbf{P_T} = [p_1, p_2, \ldots, p_9] = [2, 10, 10, 1, 3, 3, 1, 10, 4]$

$$\widetilde{M} = f(P_S, P_C) = P_S + P_C \left(\frac{1}{P_C} \sum_{p \in \mathbf{P_C}} M_p\right) \quad (1)$$

Having $M = 9$MTC devices that generates the reported sets $\mathbf{P_C}$ and $\mathbf{P_S}$, with cardinality P_C and P_S, the estimated number of MTC devices by using (1) is $\widetilde{M} = 2 + \frac{1}{3}2 + \frac{1}{3}2 + \frac{1}{3}3 = 8.9\overline{9}$. The M_p value represents the number of MTC devices that selected preamble p.

M_p cannot be known, so it has to be estimated by means of the value $\boldsymbol{K_C}$, that was termed "**collision coefficient**".

$$K_C = \frac{1}{P_C} \sum_{p \in \mathbf{P_C}} M_p \quad (2)$$

Traffic load estimation methods available in literature

1. **Empirical relation with very low complexity, proposed in [14]**

$$\tilde{M} = P_S + 2 \cdot 1.97^{\frac{P_C}{L}} P_C$$

[14] L. Miuccio, D. Panno and S. Riolo, "Dynamic Uplink Resource Dimensioning for Massive MTC in 5G Networks Based on SCMA," *European Wireless 2019; 25th European Wireless Conference*, Aarhus, Denmark, 2019, pp. 1-6.

2. **Mathematical relation with huge complexity, proposed in [10]**

$$Pr\{P_S, P_C \mid M, L\} = \frac{1}{L^M}\binom{M}{P_S}\binom{L}{P_S}P_S!\binom{L-P_S}{L-P_S-P_C}\sum_{i=0}^{P_C}(-1)^i\binom{P_C}{i}\sum_{w=0}^{i}\binom{M-P_S}{w}\binom{i}{w}w!\,(P_C-i)^{M-P_S-w}$$

$$\tilde{M} = \arg\max_{M'}\left(Pr\{P_S, P_C \mid M', L\}\right)$$

[10] G. Lin, S. Chang and H. Wei, "Estimation and Adaptation for Bursty LTE Random Access," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2560-2577, April 2016.

3. **Mathematical relation with low complexity, working only for $M \geq L$, proposed in [11]**

$$\tilde{r} = \min\left\{1, r\left(1 + \frac{[P_C - L(1-2e^{-1})]e}{2L}\right)^{-1}\right\} \Longrightarrow \tilde{M} = \frac{L}{\tilde{r}}$$

[11] S. Duan, V. Shah-Mansouri, Z. Wang and V. W. S. Wong, "D-ACB: Adaptive Congestion Control Algorithm for Bursty M2M Traffic in LTE Networks," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9847-9861, Dec. 2016.

Working region determination

With L equal to the number of available preambles if the number of attempting devices is too high it is not possible to perform an estimation on the basis of P_C and P_S, <u>since the association is not unique</u>.

$$\widetilde{M} = f(P_S, P_C)$$

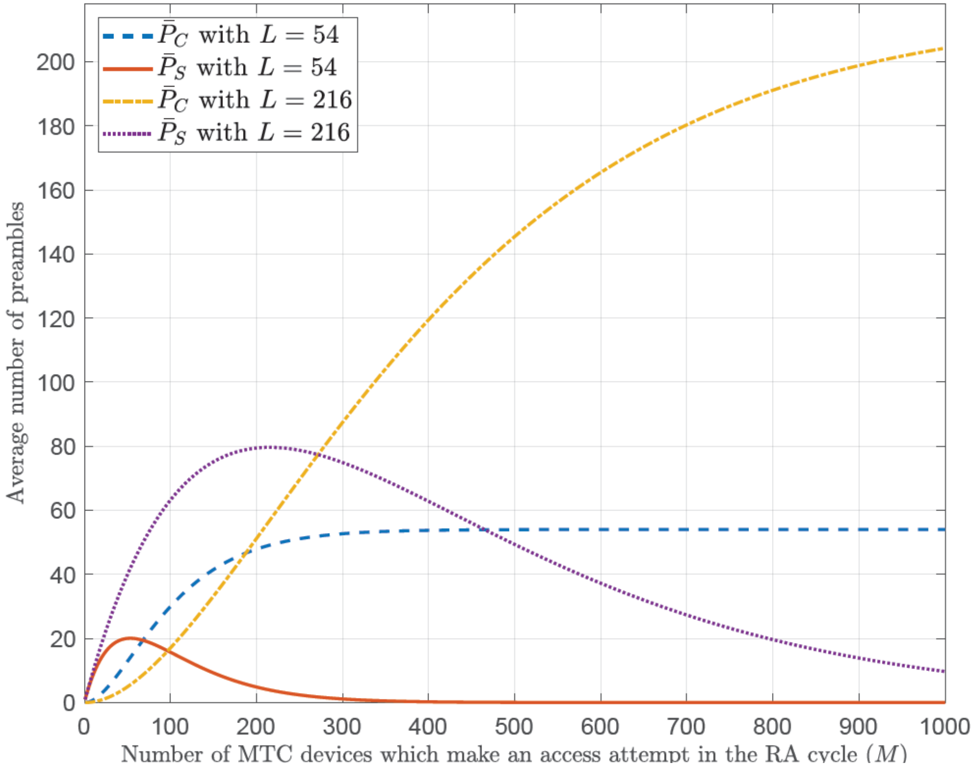
There is the need to set the maximum number of supported MTC devices for each $L = L_0 T_{pr}$, with $T_{pr} = \{1,2,3,4\}$.

Pmf of P_C: $$p_{P_C}[k] = \binom{L}{k}\left[1-\left(1-\frac{1}{L}\right)^M - M\left(\frac{1}{L}\right)\left(1-\frac{1}{L}\right)^{M-1}\right]^k \left[1-\left(\frac{1}{L}\right)^M + M\left(\frac{1}{L}\right)\left(1-\frac{1}{L}\right)^{M-1}\right]^{L-k}.$$

We set: $\Pr\{P_C = L_0 T_{pr}\} \leq \gamma = 10^{-6}$

M^h_{max} is the supported maximum value of M, for $h = T_{pr}$. Then,

$$M^h_{max} = g(L_0 h, \gamma)$$



$$\bar{P}_S = M\left(1 - {}^1/_L\right)^{M-1}$$

$$\bar{P}_C = L\left[1-\left(1-\frac{1}{L}\right)^M - M\left(\frac{1}{L}\right)\left(1-\frac{1}{L}\right)^{M-1}\right]$$

T_{pr} = number of Time slots dedicated to PRACH.

Derivation of the dataset

On the basis of the **determined working region**, by simulations we created the dataset with the following parameters.
For each configuration $T_{pr} \in \{1,2,3,4\}$ and $M \in M_{max}^{T_{pr}}$:

Inputs: P_C, P_S, and T_{pr}

Output: $\widetilde{K}_C$

$\boldsymbol{N_{sim}}$: 10000

Total points in the dataset: 21790000

Percentage of points in the training set: 90% of the dataset → Training set: $\mathbf{S_{TR}}$

Percentage of points in the test set: 10% of the dataset → Test set: $\mathbf{S_{TE}} \supset \boldsymbol{S}_{\mathbf{TE}}^{\boldsymbol{h}}$

Trained DNN

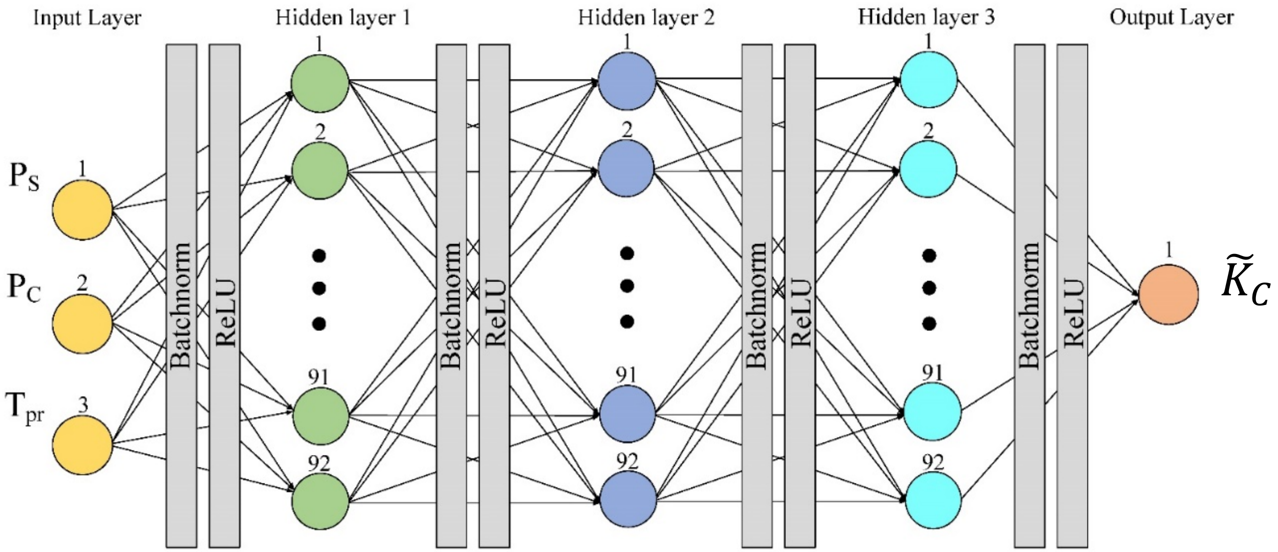
Loss function: Mean Square Error (MSE) → $\sum_{i=1}^{S_{TR}} (y_i - \tilde{y}_i)^2$

Activation function: Rectified Linear Unit (ReLU)

Number of hidden layers: 3

Number of nodes for each hidden layer: 92

Additional features for preventing vanishing and exploding gradients problems: **Xavier initialization** and **batch normalization**.



$\tilde{M} = f(P_S, P_C) = P_S + \tilde{K}_C P_C$

Key Performance Indicators for K_C estimation methods comparison

$$SS_{res} = \sum_{i=1}^{S_{TE}^h} (\tilde{y}_i^h - y_i^h)^2$$

$$\bar{y}^h = \frac{1}{S_{TE}^h} \sum_{i=1}^{S_{TE}^h} y_i^h$$

$$SS_{tot} = \sum_{i=1}^{S_{TE}^h} (y_i^h - \bar{y}^h)^2$$

- The **Coefficient of Determination**:

$$R_h^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

The better the estimation fits the data in comparison to the horizontal straight line (the null $\tilde{y}^h = \bar{y}^h$ hypothesis), the closer the value of R_h^2 is to 1.

- The **Root Mean Square Error (RMSE)** for the set $\boldsymbol{S_{\mathrm{TE}}^h}$:

$$RMSE_h = \sqrt{\frac{1}{S_{TE}^h} \sum_{i=1}^{S_{TE}^h} (\tilde{y}_i^h - y_i^h)^2}$$

The lower $RMSE_h$, the better the accuracy of the estimation for $h = T_{pr}$.

Performance evaluation of the estimate in a single RA cycle

Average value of $RMSE_h$ and R_h^2 values achieved in the estimation of K_C for the considered methods

	DNN_K-based		Proposed in [10]		Proposed in [11]		Empirical method	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
$T_{pr} = 1$	0.2180	0.7941	0.9302	-1.8897	0.2369	0.7548	0.2286	0.7770
$T_{pr} = 2$	0.1803	0.9202	1.2512	-5.1714	/	/	0.2017	0.8779
$T_{pr} = 3$	0.1651	0.9507	1.4521	-4.6340	/	/	0.2589	0.8402
$T_{pr} = 4$	0.1556	0.9644	1.5999	-4.3127	/	/	0.3328	0.7882

Average value of $RMSE_h$ values achieved in the estimation of M for the considered methods

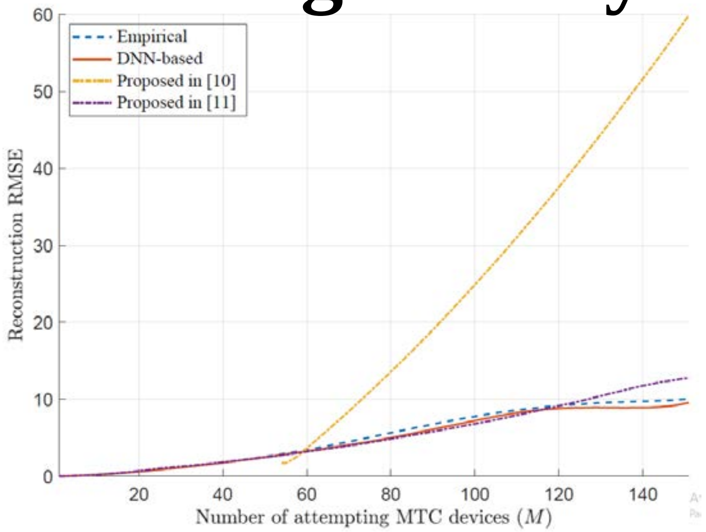
	DNN_K-based	Proposed in [10]	Proposed in [11]	Empirical method
$T_{pr} = 1$	5.80	32.64	6.47	6.23
$T_{pr} = 2$	12.60	104.41	/	14.60
$T_{pr} = 3$	19.30	193.70	/	34.15
$T_{pr} = 4$	25.81	294.41	/	62.93

$$\widetilde{M} = f(P_S, P_C) = P_S + \widetilde{K}_C P_C \quad (1)$$

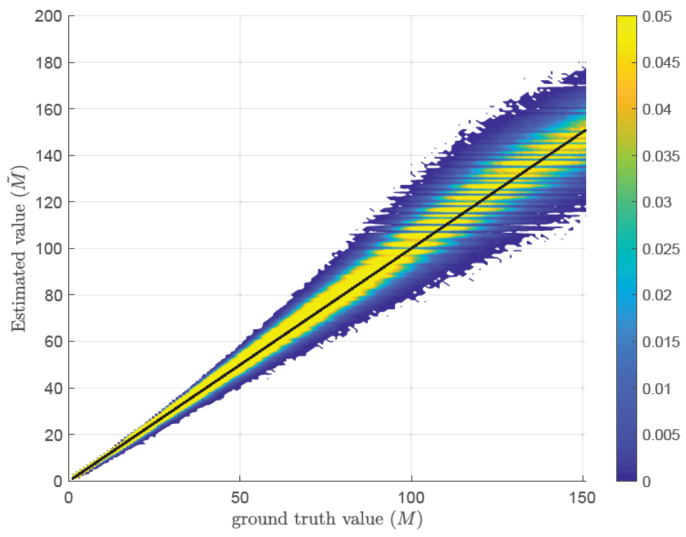
Note: "/" means that the complexity was so huge that cannot be computed by means of standard processors.

Performance evaluation of the estimate in a single RA cycle

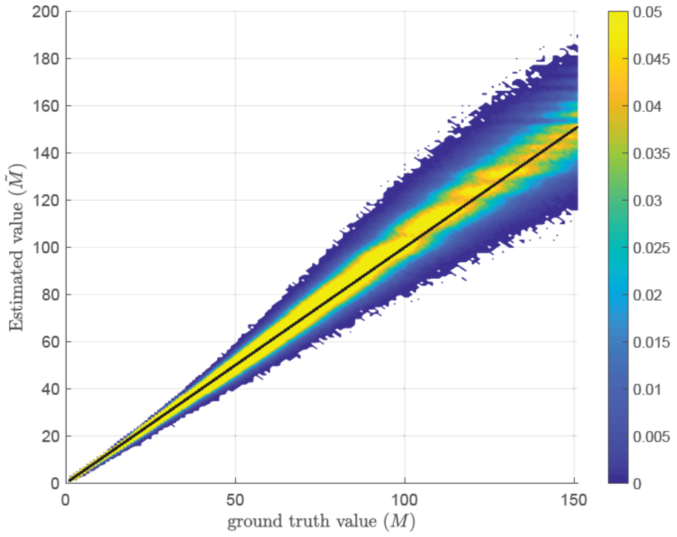
Comparison among the considered estimation methods for $T_{pr} = 1$ (i.e., $L = 54$), as M changes.



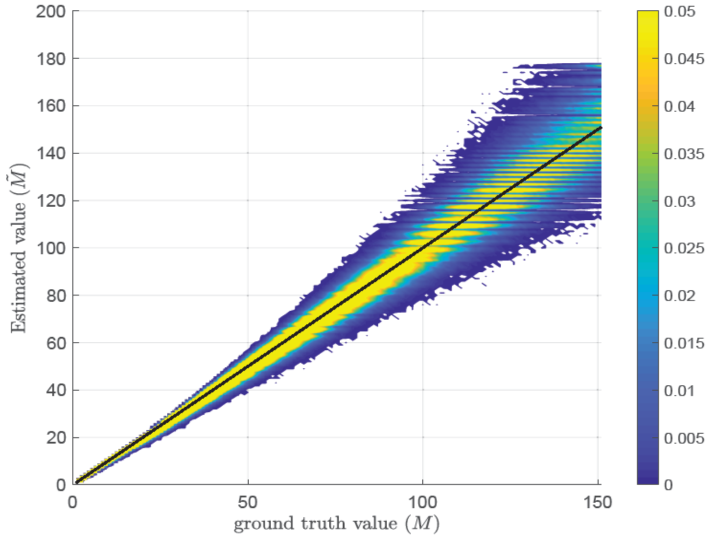
Density scatter plots of $\widetilde{M}$ vs M values with the $T_{pr} = 1$ dimensioning.



DNN-based method



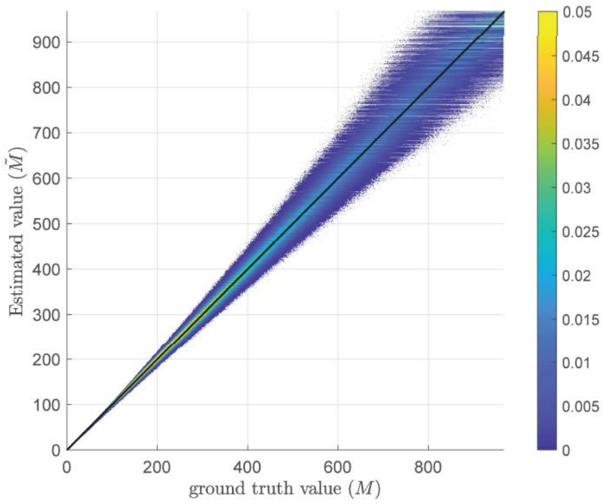
Empirical method



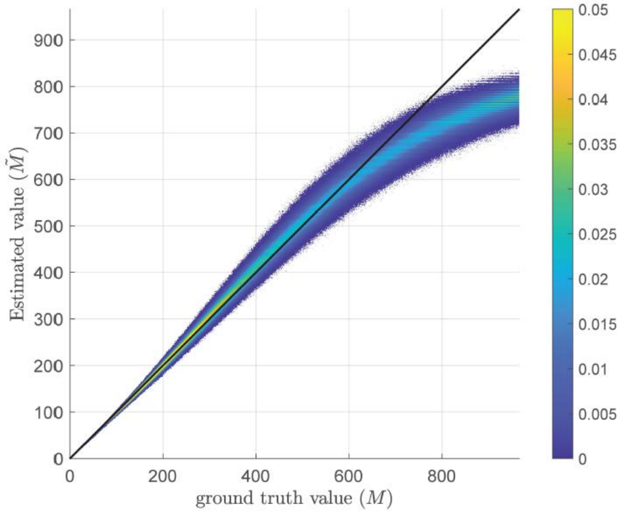
Method proposed in [11]

Performance evaluation of the estimate in a single RA cycle

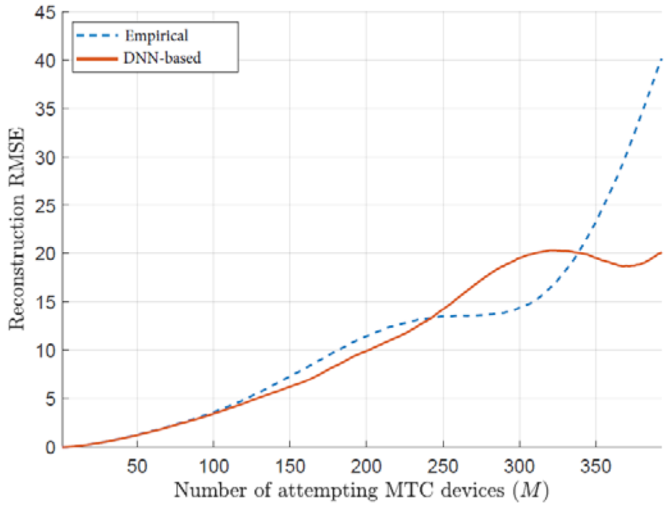
Comparison among the considered estimation methods for T_{pr} = {2, 3, 4} in terms of reconstruction RMSE, and density scatter plots of $\widetilde{M}$ vs M values with the T_{pr} = 4 (i.e., L = 216).



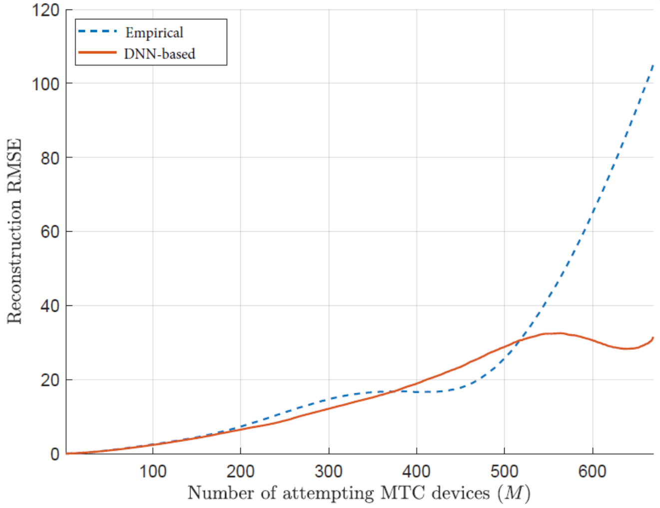
DNN-based method



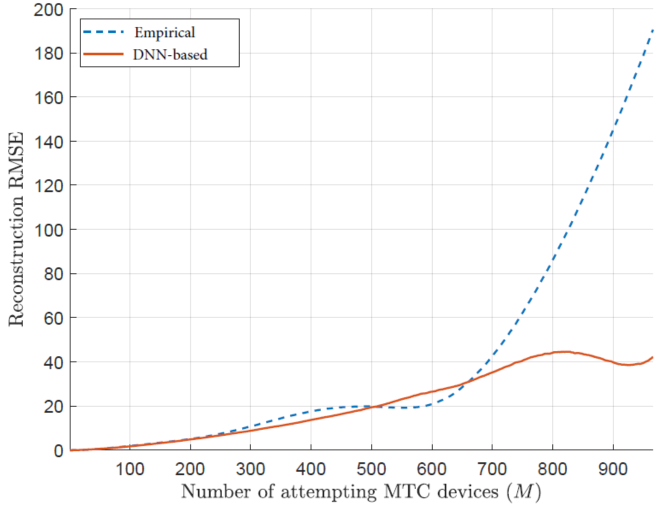
Empirical method



$T_{pr} = 2$ dimensioning



$T_{pr} = 3$ dimensioning



$T_{pr} = 4$ dimensioning

Evaluation of the estimation in a long term analysis applying the Dynamic Uplink Resource Dimensioning (DURD) scheme that uses an SCMA technique

Considering the available data transmission resources for the Sparse Code Multiple Access (SCMA) technique

$$DT_{max} = \left\lfloor \frac{\lfloor 72/Q \cdot 14 \rfloor \; (QK_{max})/S}{\lceil \theta_{max}/\log_2(I) \rceil} \right\rfloor (T_{ra} - T_{pr}),$$

and the average number of available access resource

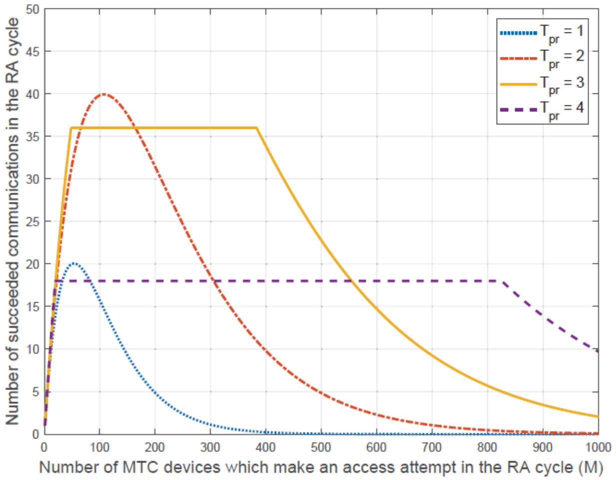
$$\bar{P}_S = M\left(1 - 1/L\right)^{M-1},$$

the average number of succeeded communications for each RA cycle is:

$$\bar{C}_S = \min(\bar{P}_S, DT_{max})$$

The DURD is applied for each RA cycle following this criterion:

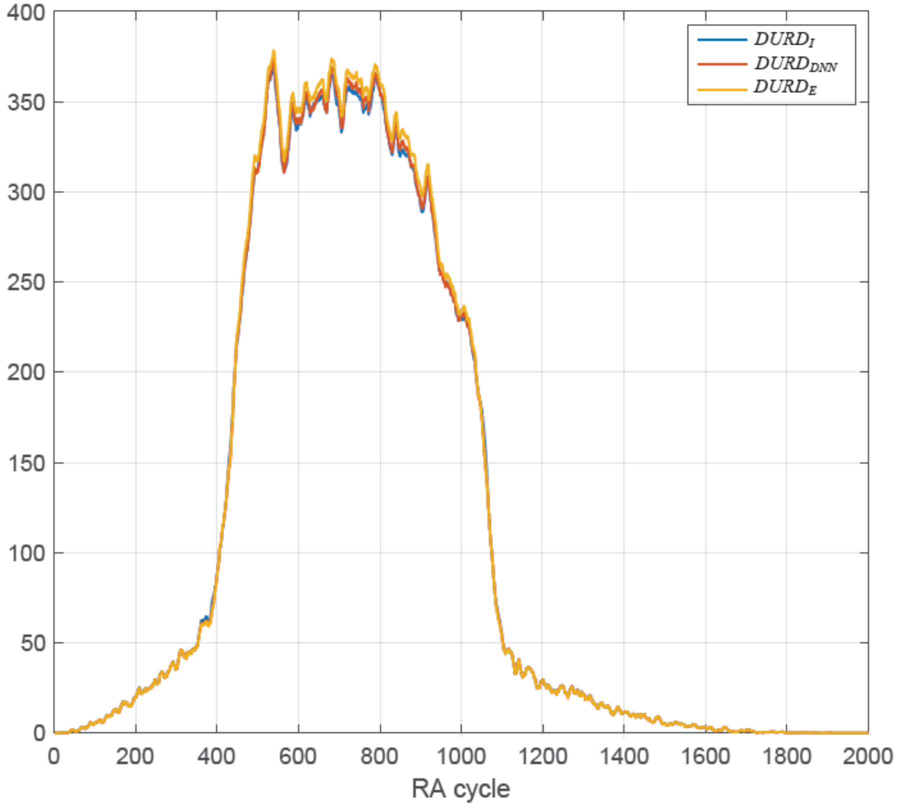
$$\begin{cases} T_{pr}^{j+1} = \underset{T_{pr} \in \{1,\dots,T_{ra}-1\}}{\arg\max} \left\{ \bar{C}_S(\tilde{M}^j) \right\} \\ M^{j+1} \cong \tilde{M}^j \end{cases}$$



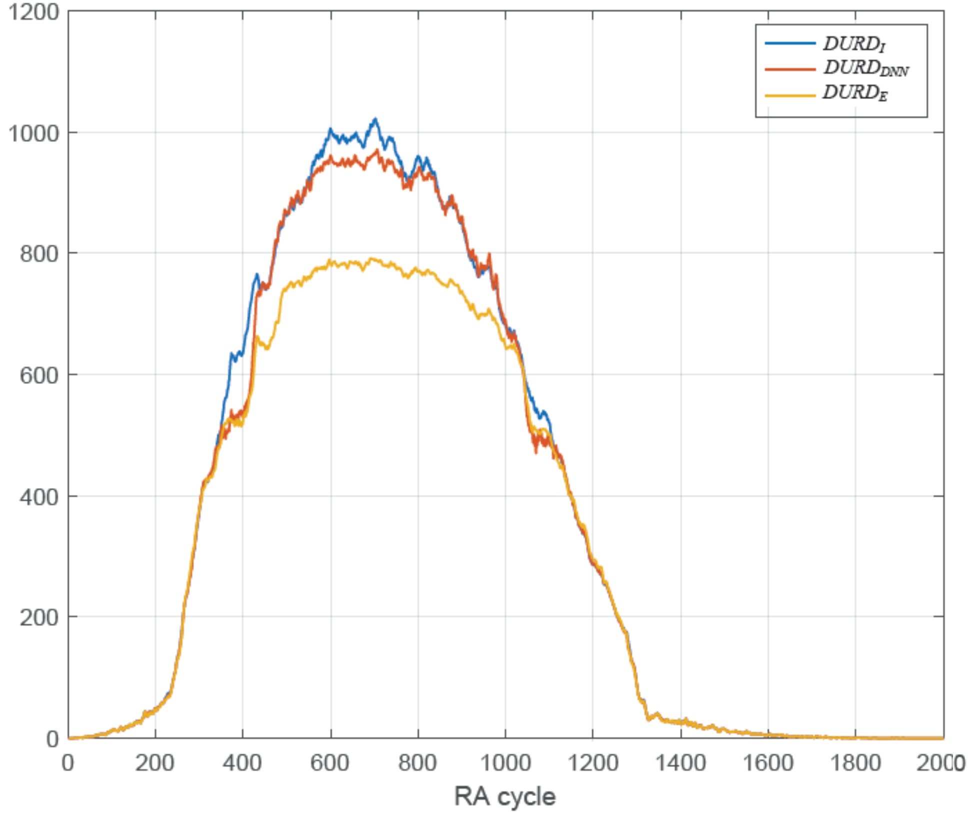
Symbol	Value
L_0	54
B	1.08 MHz
N_{MTC}	50000, 100000
M_A	10
θ_{max}	160 bits
α	3
β	4
$T_{Arrival}$	$10s$
T_{ra}	5 time slots
T_{sim}	$10s$ (2000 RA cycles)
Q	4
K_{max}	3
S	2
B_W	20ms
I	4

Performance evaluation of the estimate for a long term analysis

- Ideal amount of traffic load per RA cycle ($DURD_I$).
- Amount of traffic load per RA cycle estimated by means of empirical formula ($DURD_E$).
- Amount of traffic load per RA cycle estimated by means of the proposed DNN ($DURD_{DNN}$)



$N_{MTC} = 50000$



$N_{MTC} = 100000$

The estimate works well for all the considered cases, i.e., in medium and high load traffic conditions.

Conclusion & Future work

- Machine learning is bringing new tools for enabling network optimization in the view of 5GB and 6G.

- We analyzed the problem of estimating the traffic load on the basis of the PRACH information available at the gNB, we created both a proper dataset and an optimal DNN for addressing the problem, and benchmarked with good results our scheme with other schemes available in literature.

- Due to the importance of random access schemes and their required optimal functionality, **works in implementation** aim to apply the estimate outside the determined working region by means of more complex procedures based on Deep Learning.

Thank you for your attention!

Contact Info:

Luciano Miuccio, Daniela Panno, Salvatore Riolo

Email: luciano.miuccio@phd.unict.it, daniela.panno@dieei.unict.it, salvatore.riolo@unict.it