Traffic Adaptive Transmission Schemes for the Internet of Things

Chowdhury Saleha Ferdowsy, Zied Bouida, Mohamed Ibnkahla
Systems and Computer Engineering Department,
Carleton University, Ottawa, Canada
Emails: {chowdhury.ferdowsy, zied.bouida, mohamed.ibnkahla}@carleton.ca

***Abstract*—The growing popularity of the Internet of Things' (IoT) applications comes with new challenges for wireless communications. Indeed, wireless transmission systems should more efficiently support heterogeneous traffic from diverse types of information sources. In this paper, we propose a set of traffic-oriented transmission schemes for a massive MIMO system that adapts the number of antennas based on three types of IoT traffic (i) energy sensitive, (ii) throughput sensitive, and (iii) highly reliable traffic. We jointly consider the uplink and downlink transmission for every IoT traffic and our energy efficient model reveals the optimum number of antennas that ensure each traffic's Quality of Service (QoS) when communicating with a certain number of IoT nodes. Numerical results are shown in terms of average transmit power, spectral efficiency, average area throughput and energy efficiency. The results demonstrate that the performance is improved with the number of nodes which ensures the scalability of the IoT network.**

***Index Terms*—Energy efficiency, heterogeneous traffic, Internet of Things, Quality of Service (QoS), Massive MIMO.**

I. Introduction

Massive connectivity is a key requirement for future wireless cellular networks that aim to support numerous Internet-of-Things (IoT) applications [1], [2]. In such a scenario, a cellular base station (BS) may be required to connect to a large number of IoT devices, but unlike traditional broadband service, IoT connectivity has requirements and constraints that are significantly diverse [3]–[6]. Therefore, the BS needs to adapt its transmission strategies and system dimensions with heterogeneous IoT applications. While there has been significant development in digital wireless transmission technologies over the last few years, the one aimed at multi-user antenna technology which is called Massive MIMO (MaMIMO) (also known as scalable MIMO, very large multi-user MIMO [7]–[11]) holds the potential to support transmission techniques that are suitable for various IoT traffic.

Over the last few years, massive MIMO has gained significant attention in the literature. A comprehensive overview and survey on massive MIMO can be found in [12]. Massive MIMO in IoT application is well investigated in [4]. Standard EE-maximization techniques for millimetre wave based massive MIMO have been explored in [13]. Authors in [14] proposed an EE maximization problem for URLLC in massive MIMO that optimizes transmit power, bandwidth, and the number of active antennas while taking the end to end delay and overall packet loss into consideration. In [15], the authors provided in-depth analysis for energy efficient Massive MIMO systems to find the optimal number of antenna, terminals and power allocation. Their results showed that the number of antennas and terminals can be in the same order of magnitude to maintain EE. In [16] an adaptive antenna selection scheme in a multi-cell Massive MIMO scenario was proposed where the BS activates the number of antennas depending on the daily load variations. Authors in [17] have optimized the number of BS antennas and transmit power in the DL of a Massive MIMO system under effective SINR constraints. Based on their result they conclude that the constraints on the transmission rate can be satisfied by optimizing the number of antennas. In [18] the authors have analyzed the UL and DL performance in terms of achievable rate in multicell massive MIMO system taking imperfect channel estimation, pilot contamination, antenna correlation, and path loss into consideration. An OFDM based Massive MIMO system for improving energy efficiency of IoT application has been studied in [19] where power control has been adopted in UL and peak-to-average power ratio reduction has been applied in DL transmission for reducing the device's power consumption.

Motivated by the above contributions, we have formulated an energy efficient optimization problem for single-cell MaMIMO for three different IoT traffic using different QoS constraints. In particular, we have classified IoT traffic in three categories energy sensitive, throughput sensitive, and highly reliable traffics. The energy sensitive traffic is defined as the traffic that can tolerate low data rate and reliability but very sensitive to power consumption. A use case of energy sensitive traffic is massive Machine Type (mMTC) applications. Throughput sensitive traffic requires high data rate and moderate reliability which basically represents enhanced mobile broadband (eMBB) service. Finally, the highly reliable traffic represents the traffic that requires high reliability in mission-critical applications. The objective is to find the optimal number of antennas that need to be activated to ensure diverse QoS while maximizing the energy efficiency. Therefore, our optimization problem aims at finding the number of active antennas which will allow to adapt the transmission scheme with individual traffic.

The rest of the paper is organized as follows: Section II describes the system model and assumptions. Section III illustrates the problem formulation. Section IV presents the simulation results and discussions. Section V concludes the paper.

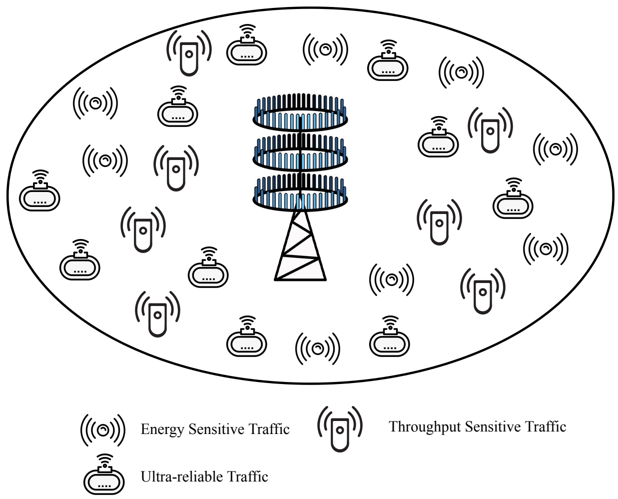

Fig. 1. Single-cell IoT network with large number of nodes generating energy sensitive, throughput sensitive and highly reliable traffic.

II. System Model

We consider an IoT link where the BS uses M antennas to communicate simultaneously with a set of K single antenna nodes where $M > K$. All the nodes in this scenario are running diverse applications which is either energy sensitive application, throughput sensitive application, or highly reliable IoT application as shown in Fig. 1. We assume that all nodes are grouped/clustered based on their traffic type and each cluster of traffic is scheduled in round-robin fashion. At any transmission, the active set of nodes send its traffic information to the BS through the control channel so that the BS can select the modulation mode and activate its antenna according to the traffic characteristics. The BS and nodes are assumed to be perfectly synchronized and operate according to the TDD mode which consists of three phases called uplink training, uplink data transmission, and downlink data transmission. In the uplink training period, all nodes transmit mutually orthogonal pilot sequences to the BS for channel estimation. After that, the nodes send their data to the BS in the uplink data transmission period. In the downlink data transmission phase, the BS transmits data to the active nodes by utilizing the same uplink pilot signals since both uplink and the downlink share the same frequency channel. This phenomenon is also known as channel reciprocity of TDD protocol. Here, we assume the uplink transmission ratio, $\epsilon^u = 0.6$ and downlink transmission ratio, $\epsilon^d = 0.4$.

A. Channel Model and Linear Processing

The channel between the BS and k_{th} active node includes both large scale path loss and small scale block fading i.e the channel is constant for one coherence block and then updated independently from the circular-symmetric complex Gaussian distribution. The channel matrix $\mathbf{H} = [\mathbf{h1}, \mathbf{h2}, ...\mathbf{h_k}] \in \mathbb{C}^{M\times K}$ denotes instantaneous propagation channel of all nodes. The BS is assumed to have knowledge about large scale fading and estimates the small scale fading in each coherence interval. We denote the uplink linear receive combining matrix by $\mathbf{G} = [\mathbf{g1}, \mathbf{g2}, ...\mathbf{g_k}] \in \mathbb{C}^{M\times K}$ and precoding matrix by $\mathbf{A} = [\mathbf{a1}, \mathbf{a2}, ...\mathbf{a_k}] \in \mathbb{C}^{M\times K}$ and set $\mathbf{G} = \mathbf{A}$. We consider Zero Forcing (ZF) as linear processing scheme which gives $\mathbf{G} = \mathbf{H}(\mathbf{H^H H})^{-1}$.

B. Transmit Power

We have adopted equal power control method to guarantee equal bit rate to the K nodes that belongs to the same traffic. Let $\mathbf{P}^{(u)} = \text{diag}(p_1^u, p_2^u, ..., p_k^u)$ denote the uplink transmitted power by node i for $i = 1, 2, ...K$. It follows that the uplink power allocation vector $\mathbf{p}^u = (p_1^u, p_2^u, ..., p_k^u)^T$and downlink power allocation vector $\mathbf{p}^d = (p_1^d, p_2^d, ..., p_k^d)^T$ must satisfy the condition [20]

$$\mathbf{p}^u = \sigma^2(\mathbf{D}^u)^{-1}\mathbf{1}_K, \quad (1)$$

$$\mathbf{p}^d = \sigma^2(\mathbf{D}^d)^{-1}\mathbf{1}_K, \quad (2)$$

where σ^2 denotes the noise variance at the k_{th} node in (Joule/symbol) and the (l,k)th element of both $(\mathbf{D}^{(d)}, \mathbf{D}^{(u)}) \in \mathbb{C}^{K\times K}$ is given by

$$[\mathbf{D}^{(u)}]_{l,k} = \begin{cases} \frac{|\mathbf{g_k^H h_k}|^2}{2(R/B-1)\|\mathbf{g_k}\|^2} & \text{for} \quad k = l, \\ -\frac{|\mathbf{g_k^H h_l}|^2}{\|\mathbf{g}_k\|^2} & \text{for} \quad k \neq l. \end{cases} \quad (3)$$

$$[\mathbf{D}^{(d)}]_{l,k} = \begin{cases} \frac{|\mathbf{h_k^H a_k}|^2}{2(R/B-1)\|\mathbf{a_k}\|^2} & \text{for} \quad k = l, \\ -\frac{|\mathbf{h_k^H a_l}|^2}{\|\mathbf{a}_l\|^2} & \text{for} \quad k \neq l. \end{cases} \quad (4)$$

Due to the duality of TDD mode, the total uplink and downlink transmit power are the same except for the power amplifier efficiency of nodes and the BS and the fraction of transmission. The transmit power consumption P_T is given as follows where the uplink and downlink parameters are denoted using superscripts [15].

$$P_T^u = \frac{B\epsilon^u}{\eta^K}\mathbb{E}\{\mathbf{1}_K^T\mathbf{p}_k^u\}. \quad (5)$$

$$P_T^d = \frac{B\epsilon^d}{\eta^{BS}}\sum_{k=1}^{K}\mathbb{E}\{p_k^d\}. \quad (6)$$

C. Achievable Rate

The average achievable rate, R_T (bit/sec) for both UL and DL with ZF linear processing under the condition $M \geq K+1$ is given by [15]

$$R_T = \left(1 - \frac{(\tau^u + \tau^d)K}{U}\right) B\log(1+\gamma) \quad (7)$$

Here B is the transmission bandwidth of the channel, the term $1 - \frac{(\tau^u+\tau^d)K}{U}$ accounts for the pilot overhead and γ is the average SNR. To ensure reliability, we need a minimum SNR to guarantee specific BER and modulation mode. Here we have adopted Adaptive Quadrature Amplitude Modulation (M-QAM) for determining transmission rate under certain reliability. In this case, the required SNR for a target BER β and constellation size 2^n is approximated as [21]

$$\gamma_{T_n} = -\frac{2}{3}\ln(5\beta)(2^n-1); \quad n = 0,1,...N, \tag{8}$$

where n is the modulation mode. As we can see from (8), for a specific modulation mode, the traffic that demands lower BER to ensure high reliability will require higher SNR.

D. Circuit Power

We define P_{cp} as the total circuit power spent by the BS and IoT nodes. Circuit power consumption of the BS consists of a fixed part (e.g., control signal, backhaul, DC-DC conversion) and a dynamic part which depends on the number of active antennas M, number of nodes K, and the channel gain. We have adopted the circuit power model as [15] where the total circuit power consumed in total UL and DL transmission is given by

$$P_{cp} = \alpha_1 R_T + \alpha_2 + K P_{tc}^{K} + M P_{tc}^{BS} + P_{CE} + P_{lp}, \tag{9}$$

where α_1 accounts for coding, decoding, and backhaul power consumption and R_T is calculated using (7). The second term α_2 is the constant power consumption for site cooling, control signaling, and frequency synthesis at the BS. The term P_{tc} represents the transceiver power consumption by the nodes and BS denoted by the superscript. P_{CE} stands for the BS channel estimation's power consumption and is given by

$$P_{CE} = \frac{B\,K^2}{U}\left(\frac{2\,M\tau^u}{\Psi_{BS}} + \frac{4\tau^d}{\Psi_K}\right), \tag{10}$$

where Ψ_{BS} and Ψ_K denote the computational efficiency of the BS and the K nodes, respectively. Finally, P_{lp} accounts for the power consumed for linear processing. For ZF precoding, P_{lp} is given by

$$\begin{aligned} P_{lp} = {} & \frac{2\,B\,M\,K}{\Psi_{BS}}\left(1 - \frac{(\tau^u+\tau^d)K}{U}\right) \\ & + \frac{B}{U}\left(\frac{K^3}{3\Psi_{BS}} + \frac{3M\,K^2 + M\,K}{\Psi_{BS}}\right) \end{aligned} \tag{11}$$

It is important to note that the circuit power consumption is a function of number of antennas (M) and number of nodes (K) which means more antennas and nodes will consume more power. Therefore, while designing an energy efficient transmission scheme for the IoT traffic we need to optimize the number of antennas depending on the number of active nodes in the network to better utilize the resources. The description of each parameter of circuit power consumption is shown in Table I.

III. Problem Formulation

In this section, we aim at finding the optimal SNR and the corresponding number of antennas to select the transmission scheme for each of the heterogeneous IoT traffics. The optimal SNR under different traffic constraints allows the selection of the modulation scheme ensuring low power consumption for energy sensitive traffic, high bit rate for throughput sensitive traffic, and better error performance for the highly reliable traffic. Let us denote the BER for energy, throughput, and reliability sensitive traffics by β_{E}, β_{T}, and β_{R}, respectively. For energy sensitive traffic, we enable communication with the lowest modulation mode. On the other side, for throughput sensitive traffic the communication takes place in the highest SNR region. Finally, for the highly reliable traffic, the operating region is set to the lowest modulation mode satisfying BER of β_{R}.

TABLE I

Reference Parameter

Parameter	Description
Coding,Decoding and backhaul, $\alpha 1$	1.15×10^{-9} W/(Gb/s)
Static Power Consumption, $\alpha 2$	20 W
Transceiver chain at BS, P_{tc}^{BS}	1W
Transceiver chain of nodes, P_{tc}^{K}	0.1W
Downlink transmission, ϵ^d	0.4
Uplink Transmission, ϵ^u	0.6
Pilot length at uplink and downlink : τ^u, τ^d	1
Computational Efficiency at BS, Ψ_{BS}	12.8 Gflops/W
Computational Efficiency at nodes, Ψ_K	5 Gflops/W
Power amplifier efficiency at BS, η^{BS}	0.39
Power amplifier efficiency at nodes, η^K	0.3

The optimal system parameter should maintain the energy efficiency (EE) of the overall wireless system for every traffic. Therefore, we have adopted the EE as the cost function which is computed as the ratio of average sum rate (in bit/sec) and the average total power consumption (in Joules). The optimization problem to maximize the overall EE in a multi-user setting for both uplink and downlink is formulated as follows

$$(M^*, \gamma^*) = \text{argmax EE} = \frac{\sum_{k=1}^{K} R_T(\gamma)}{(P_T^u + P_T^d) + P_{cp}(M,K,\gamma)}, \tag{12}$$

Subject to

$$\gamma \leq \gamma_{T_E} \quad \text{for energy sensitive traffic,} \tag{13a}$$
$$\gamma \geq \gamma_{T_T} \quad \text{for throughput sensitive traffic,} \tag{13b}$$
$$\gamma \geq \gamma_{T_R} \quad \text{for highly reliable traffic,} \tag{13c}$$

where $\gamma_{T_E}, \gamma_{T_T}$ and γ_{T_R} are given by

$$\gamma_{T_E} = \gamma_{T_2} \quad \text{for } \beta_E, \tag{14a}$$
$$\gamma_{T_T} = \gamma_{T_4} \quad \text{for } \beta_T, \tag{14b}$$
$$\gamma_{T_R} = \gamma_{T_1} \quad \text{for } \beta_R, \tag{14c}$$

where γ_{T_2}, γ_{T_4} and γ_{T_1} are calculated using equation (10). Note that equation 15(a) restricts the rate transmission to lowest modulation mode so that the low power node use less energy. Equation 15(b) ensures that when the BS will serve the throughput sensitive nodes, it will serve with high modulation mode since the optimal SNR is set to achieve high spectral efficiency. Finally, equation 15(c) sets the modulation threshold for a lower BER to ensure high reliability for highly reliable traffic and the constraints is applied for low constellation size.

Eq. (12) is a constrained optimization problem where we want to find optimal number of antenna M for a given number of nodes K. Since we are considering ZF linear processing, we need to set $M \geq K+1$. At first, for any number of scheduled nodes (K) bearing identical traffic, we apply equal power allocation method to guarantee an equal SNR which is feasible for all nodes. After that, we find the optimal feasible SNR (γ^*) by gradient decent method that satisfies the corresponding traffic's SNR constraint in (13) and maximizes EE in the objective function (12). This optimization is executed for every node (K) with all possible combination of antennas (M). We therefore, use brute force method to search over all the values of M that gives maximum EE value for that particular number of nodes (K). The number of antennas (M) that globally maximizes EE, is the optimum M^*. After we find the optimum number of antennas and corresponding power allocation, we determine the modulation scheme that needs to be applied for the individual IoT traffic. This way in every transmission, the proposed traffic adaptive transmission scheme adapts the number of antennas and modulation mode based on the traffic type. The detailed algorithm is shown in Algorithm 1.

Algorithm 1 QoS constrained joint optimization

Input : K
Output : M, γ
Initialization : $\gamma = 0.1$
for $K = 1$ to K_{max} **do**
  for $M = K + 1$ to M_{max} **do**
    for $i = 1$ to $iterations$ **do**
      generate channel realization for K nodes and M antennas using small scale channel fading and large scale path loss attenuation
    end for
    apply equal power allocation technique on all channel realization to find a feasible and equal γ to K nodes by using eq. (1) and (3)
    find the optimum γ by gradient descend method that is feasible for K nodes, maximizes EE in eq. (12) and satisfies the constraints in (13)
  end for
  find the optimum (M^*, γ^*) that globally maximizes EE by using brute force method
  find modulation mode supported by the corresponding γ^*
end for
return M^* , γ^*

IV. Numerical results

This section provides numerical results using Monte Carlo simulations with random user locations and small scale fading for optimizing EE under QoS constraints. We consider a single-cell scenario where the BS is equipped with 126 antennas and serving 50 nodes. The simulation parameters are given in Table II and were performed using Matlab.

Fig. 2 illustrates the optimal antenna-node ratio (M/K) i.e, the number of antennas that should be proportionally increased with the number of nodes to maximize EE for all types of traffic under consideration. For every traffic, it is achieved by finding the optimal SNR region under the constraints in eq. (13) and then finding the number of antennas (M) that will maximize the objective function in (12). Therefore the difference in the number of antennas is due to different operating region constraints (i.e, the SNR constraints) applied to the individual traffic. Since the SNR constraints are set to be higher for the throughput sensitive and highly reliable traffics for higher throughput and reliability respectively, they require more antennas than the energy sensitive traffic. Fig. 2 depicts that the number of antennas to serve 40 energy sensitive nodes is 58 whereas 76 and 81 antennas need to be active for serving the same number of highly reliable and throughput sensitive nodes, respectively.

Fig. 3 illustrates the average UL transmit power per IoT node for different K (using the corresponding optimal M) that maximizes EE under QoS constraint. We see that the throughput sensitive traffic and highly reliable traffic need to operate with high power consumption. This is because throughput sensitive traffic needs to spend more power for high data rate whereas highly reliable traffic requires that to ensure high reliability. We also observe that for all types of traffic, the transmit power consumption per node decreases with the number of nodes K. This is the result of array

TABLE II
Simulation Parameter

Simulation Parameter	Value
Maximum number of antenna, M_{max}	126
Maximum number of nodes, K_{max}	50
Number of Iterations	100
Cell radious, d_{max}	250 m
Minimum distance, d_{min}	35 m
Path loss at Distance d	$\frac{10^{-3.53}}{\lVert d\rVert^{3.76}}$
Transmission Bandwidth, B	20 MHz
Channel coherence interval, U	1800 symbols
Total Noise Power	-96 dBm
Energy and throughput sensitive BER, $\beta_E = \beta_T$	10^{-3}
BER of highly reliable traffic, β_R	10^{-6}

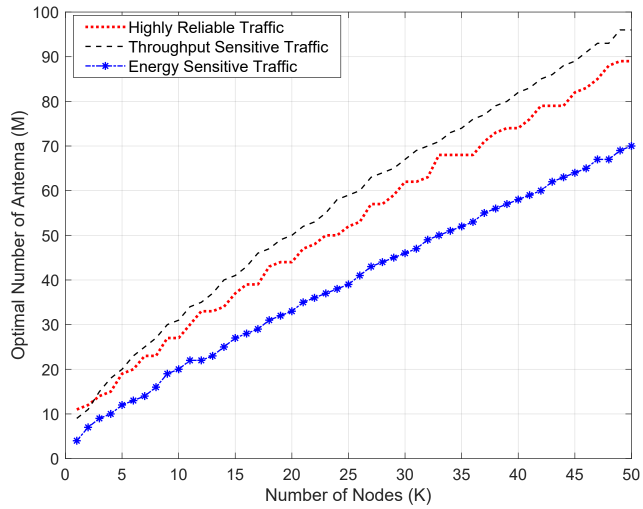

Fig. 2. Optimal number of antennas at EE maximizing point.

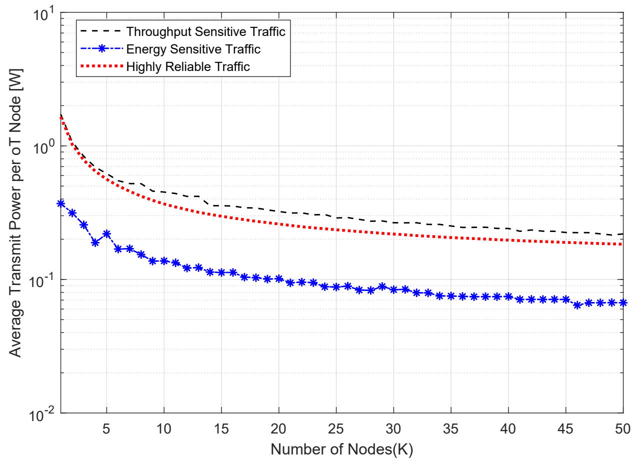

Fig. 3. Average transmit power per node at EE-maximization solution.

gain which means the proportional increase in the number of antennas with the number of nodes allows the array to collect more energy from the desired signal. As a result, the higher array gain (results from additional antennas in Fig. 2) improves SNR and allows the nodes to use less power in transmission. Fig. 3 also shows the variation of power level for different traffics. For example, when 40 throughput sensitive and highly reliable nodes need to transmit data, they will be consuming 236 mW and 197 mW power, respectively. In contrast, the same number of energy sensitive nodes will use 72 mW. This is again because of the different SNR regimes of different traffics.

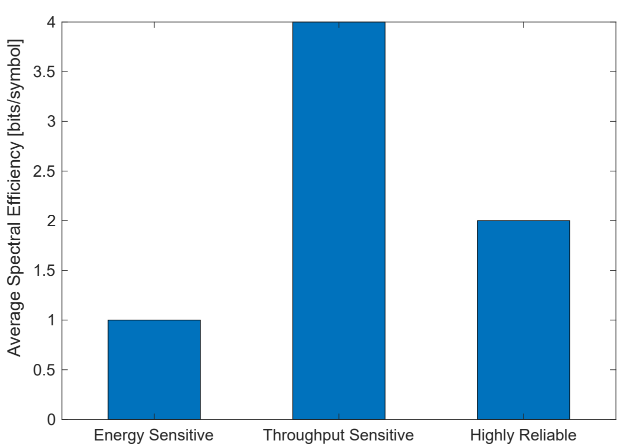

Fig. 4. Achievable spectral efficiency for different IoT traffic.

Fig. 4 compares the achievable SE in bits per symbol that allows us to select the modulation scheme for different IoT traffics. This result is achieved after finding the optimal number of antennas and optimal SNR. As mentioned, the proposed scheme applies an equal power control technique to guarantee a given feasible SNR to all nodes and then finds the optimal feasible SNR regime that maximizes EE in eq. (12) and satisfies QoS constraints in eq. (13). This way, the SNR is calculated for every possible combination of K and M but out of all the combinations, the optimum SNR that is found from the optimal antenna-node ratio is used for determining the modulation mode. Fig. 4 shows that our proposed energy efficient transmission scheme allows the lowest SE of only 1 bit per symbol for the energy sensitive traffic because of the rate constraint applied in the optimization problem. This also means that this traffic will be allowed to communicate with binary phase-shift keying (BPSK) using the BER of 10^{-3} to minimize power consumption as shown in Fig. 3. The optimal modulation mode for highly reliable traffic is found to be 2 bits per symbol which means the energy efficient modulation mode for this type of traffic is QPSK under low BER of 10^{-6}. As we have seen in Fig. 2 and Fig. 3, this higher reliability comes at the cost of high antenna selection and transmit power respectively. Finally, the energy efficient modulation mode for throughput sensitive traffic is found to be 16-QAM or 4 bits per symbol under the BER of 10^{-3} which is also an outcome of a high number of antennas and high power consumption.

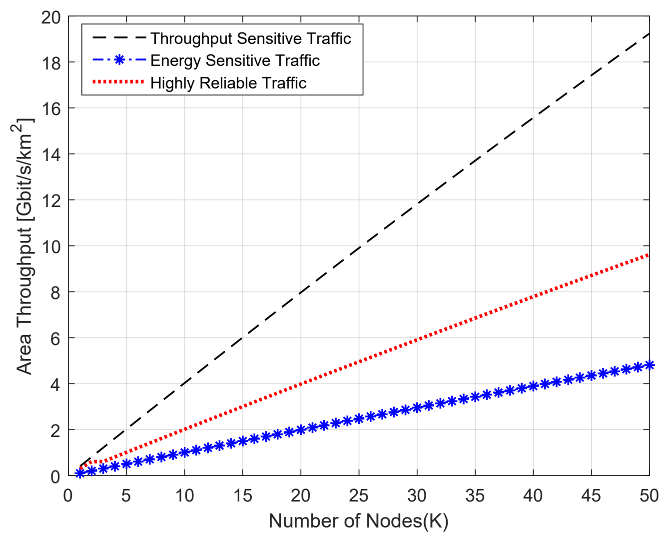

Fig. 5. Area throughput at EE maximizing solution.

Fig. 5 shows the area throughput in $(Gbit/s/km^2)$ that maximizes the EE for the different number of nodes K that will be served simultaneously. The EE maximizing throughput is achieved using a transmission bandwidth of 20 MHz which is typically used in the current LTE network. For all the considered traffic, a higher number of nodes results in a higher aggregated data rate. This is the impact of spatial multiplexing gain that results in a proportional increase in sum SE with the number of nodes. As expected, the throughput sensitive traffic has the highest throughput because of the high SE as shown in Fig. 4. The other two traffic is also following the same trend of SE. Fig. 5 also indicates that using only 40 antennas from BS as shown in Fig. 2, 25 energy sensitive nodes can be served simultaneously with a total data rate of 2.5 $Gbit/s$. On the other hand, the energy efficient solution in Fig. 2 yields that 52 and 60 antennas are sufficient for providing total data rate of 5 $Gbit/s$ and 10 $Gbit/s$ per kilometre of area, to the same (i.e, 25) number of highly reliable nodes and throughput sensitive nodes, respectively.

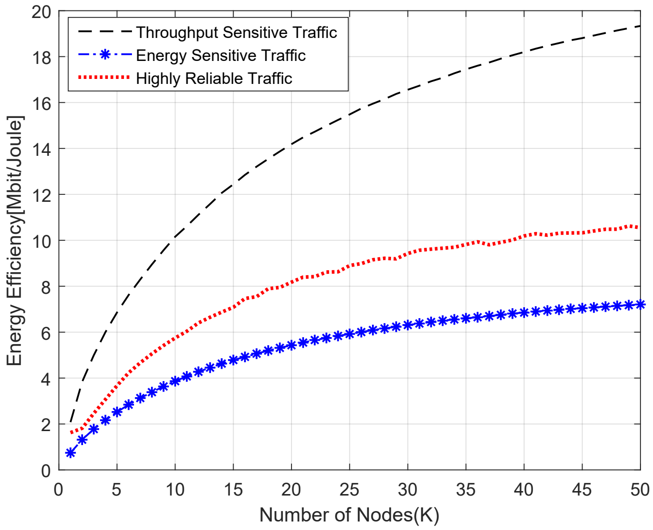

Fig. 6. Maximum energy efficiency using optimal number of antennas.

Fig. 6 shows the EE ($Mbit/J$) for different traffic under the QoS constraints. The EE for all types of traffic increases significantly with the number of nodes. This is attributed to the multiplexing and array gain of massive MIMO that increases with the number of nodes and results in high throughput as shown in Fig. 5 and low transmit power in Fig. 3, respectively. As expected, the throughput sensitive traffic shows the highest EE because the proposed scheme allows this traffic to operate in a high SNR region such that it can maintain higher throughput. As the throughput increases with the number of nodes, the EE also increases. The trend of highly reliable traffic depicts that although energy efficient solution enables this traffic to operate in high SNR, that SNR is utilized for ensuring low BER. Therefore, the rate is compromised and it shows lower EE compared to throughput sensitive traffic. Finally, the energy sensitive traffic shows the least EE as it transmits with a low bit rate which impacts the throughput. Although it achieves low power consumption, the ratio between the throughput and power consumption is lower than the other two traffics.

V. Conclusion

In this paper, we aim at designing a traffic-oriented energy efficient Massive MIMO system that can optimally scale the number of antennas with the number of active nodes in the network and adapt the transmission scheme with the traffic type. The results show that by adapting the number of antennas with proper modulation mode, the heterogeneous IoT traffic can be accommodated in common wireless transmission infrastructure. By taking advantage of massive MIMO technology, the performance of the proposed scheme is shown to be scalable and energy efficient to all IoT traffics.

References

[1] S. Li, L. Da Xu, and S. Zhao, "5G Internet of Things: A survey," *Journal of Industrial Information Integration*, vol. 10, pp. 1–9, 2018.

[2] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: physical and MAC-layer solutions," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 59–65, Sep. 2016.

[3] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct 2018.

[4] A.-S. Bana, E. De Carvalho, B. Soret, T. Abrão, J. C. Marinello, E. G. Larsson, and P. Popovski, "Massive MIMO for internet of things (IoT) connectivity," *Physical Communication*, vol. 37, p. 100859, 2019.

[5] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects," *IEEE Wireless Com.*, vol. 25, no. 3, pp. 124–130, JUNE 2018.

[6] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Com.*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[7] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[8] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, February 2016.

[9] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[10] E. G. Larsson and L. V. der Perre, "Massive MIMO for 5G," *IEEE 5G Tech Focus*, vol. 1, no. 1, March 2017.

[11] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[12] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[13] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: Opportunities and challenges," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 86–94, 2017.

[14] C. Sun, C. She, and C. Yang, "Energy-efficient resource allocation for ultra-reliable and low-latency communications," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[15] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal Design of Energy-Efficient Multi-User MIMO Systems:Is Massive MIMO the Answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[16] M. M. Hossain, C. Cavdar, E. Björnson, and R. Jäntti, "Energy saving game for Massive MIMO: Coping with daily load variation," *IEEE Trans. on Vehicular Tech.*, vol. 67, no. 3, pp. 2301–2313, March 2018.

[17] K. Senel, E. Björnson, and E. G. Larsson, "Adapting the number of antennas and power to traffic load: When to turn on massive MIMO?" in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018, pp. 1–6.

[18] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the ul/dl of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Com.*, vol. 31, no. 2, pp. 160–171, Feb 2013.

[19] B. M. Lee, "Improved energy efficiency of massive MIMO-OFDM in battery-limited iot networks," *IEEE Access*, vol. 6, pp. 38 147–38 160, 2018.

[20] S. U. Pillai, T. Suel, and Seunghun Cha, "The perron-frobenius theorem: some of its applications," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 62–75, March 2005.

[21] M.-S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer J. Wireless Communications*, vol. 13, pp. 119–143, May 2000.