Spectrum-efficient D2D-based puncturing algorithm for eMBB and URLLC coexisting networks

Yuhong Zhu[1], Kun Jiang[2], Linlin Zhao[3], Xuefen Chi[4] and Shuang Wang[5]
{yhzhu@jlu.edu.cn [1], jiangkun17@ mails.jlu.edu.cn[2], zhaoll13@mails.jlu.edu.cn[3]}

College of Communication Engineering, Jilin University, Changchun 130012, China

Abstract. In this paper, we propose a spectrum-efficient device-to-device (D2D)-based puncturing algorithm to multiplex enhanced mobile broadband (eMBB) and ultra-reliable low-latency communication (URLLC) traffic. The users with URLLC traffic communicate directly in D2D mode by reusing the resource blocks (RBs) allocated to cellular user with eMBB traffic. First, we derive the achievable transmission rate of cellular link that suffers from time-varying interference caused by different D2D links in two-timescale scheduling slot structure. Additionally, we calculate the achievable transmission rate of D2D link considering the short-packet nature of URLLC traffic. Then, the delay and reliability quality of service (QoS) of URLLCs are studied based on the effective capacity theory. Finally, the puncturing algorithm is formulated as a throughput maximization problem subject to the QoS constraints of eMBB and URLLC traffic. The transmitting power of eMBB users and URLLC users, as well as the resource reusing pattern are jointly optimized.

Keywords: eMBB (enhanced mobile broadband); URLLC (ultra-reliable low-latency communication); D2D (device-to-device) based puncturing algorithm; effective capacity.

1 Introduction

Enhanced mobile broadband (eMBB) and ultra-reliable low-latency communication (URLLC) are two typical types of services in the fifth generation (5G). eMBB is supposed to support 0.1-1 Gbit/s user experience rates, and URLLCs require the end-to-end delay below 1 ms with the probability of at least 99.999% [1]. Though the URLLC traffic is sporadic, it should be supported by large bandwidth as similar as the eMBB traffic. Due to scarcity of bandwidth resource, multiplexing eMBB and URLLC traffic, which refers to transmitting the diverse traffic on the same bandwidth resources, may be a feasible approach.

For eMBB and URLLC multiplexing scenarios, 3GPP RAN WG1 developed a two-timescale scheduling slot structure. The transmission time interval (TTI) of eMBB is one millisecond, while the TTI of URLLC traffic is shorter than one millisecond [2]. In the two-timescale scheduling slot structure, the puncturing-type multiplexing schemes were proposed, where the URLLC traffic prefers to be scheduled immediately once it arrives even if the eMBB traffic is transmitting for guaranteeing the delay quality of service (QoS) of URLLC traffic. In the case of a single URLLC user, some researchers focused on mitigating the rate loss of the punctured eMBB traffic through joint scheduling [3] or rotating constellations [4], rather than satisfying the QoS of URLLC traffic. In [5], a multi-user-punctured scheduler was developed for achieving a balance between the spectral efficiency of eMBB and average latency of URLLC. However, guaranteeing the average delay is not enough for URLLC traffic, the delay and reliability QoS of which can be represented by the statistical delay QoS including the delay bound and delay bound violation probability.

Device-to-device (D2D) technique is a viable way to reduce the end-to-end delay of URLLC. In the existing works, both the cellular links and the D2D links were utilized to carry URLLC traffic [6-8]. Only the spatial diversity gain of D2D was exploited to boost reliability of URLLCs. In the conventional D2D communications, resource allocation schemes have been studied to maximize spatial multiplexing. However, they are not suitable for the case of eMBB and URLLC multiplexing. In the two-timescale scheduling slot structure, since the TTI of URLLC is smaller than the one of eMBB, the resource of a cellular link carrying eMBB traffic may be selectively multiplexed by several different D2D links with URLLC traffic. Therefore, the interference suffered by one transmission of a cellular link is time-varying and hard to be evaluated.

In this paper, the spatial multiplexing gain of D2D technology is exploited to multiplex the URLLC traffic and the eMBB traffic. To guarantee the end-to-end delay QoS of URLLC, the users with URLLC

This work was supported by Jilin Provincial Science and Technology Department Key Scientific and Technological Project (No.20190302031GX), Changchun Scientific and Technological Innovation Double Ten Project (No.18SS010), National Natural Science Foundation of China (No.61671210), Jilin Province Development and Reform Commission Project (No.2017C046-3), and [illegible] (No.61801191).

EAI MOBIMEDIA 2020, August 27-28, Harbin, People's Republic of China

DOI 10.4108/eai.27-8-2020.2294705

traffic employ the D2D mode to communicate directly via reusing the resource blocks (RBs) allocated to the cellular users with eMBB traffic. The statistical delay QoS in effective capacity theory is adopted to characterize the reliability and delay QoS of URLLCs. Based on the effective capacity theory, we propose a novel D2D-based puncturing algorithm with the aim of achieving high system throughput while satisfying the heterogeneous QoS of eMBB and URLLC traffic. The main contributions of this paper are as follows.

(1) In the D2D-based multiplexing scenario, we calculate the average power of time-varying interference caused by multiple D2D transmissions and derive the achievable transmission rate of the cellular link. Then we derive the achievable transmission rate of D2D link considering short-packet transmission feature of URLLCs.

(2) We derive the effective capability of D2D link with URLLC traffic. Based on the effective capacity theory, we quantify the constraints on the transmission rate of D2D link considering the statistical delay QoS requirement of the random and sporadic URLLC traffic.

(3) The D2D-based puncturing algorithm is formulated as a throughput maximization problem subject to the rate requirement of eMBB traffic and the statistical delay QoS requirement of URLLC traffic. The problem is solved by the particle swarm optimization (PSO) algorithm.

2 System model

In this paper, M cellular links carrying eMBB traffic and N D2D links carrying URLLC traffic are considered to coexist in a single cell of wireless network. $\mathcal{M}=\{1, 2,\ldots, M\}$ and $\mathcal{N}=\{1, 2,\ldots, N\}$ denote the sets of cellular users and D2D links respectively. The frequency division duplex (FDD) is utilized. In this paper, the D2D links reuse the UL resources of cellular users. The BS is assumed to know the channel statement information (CSI) of cellular users and D2D links. The channel gain between the cellular user i and the BS is denoted by $g_i^{b,c}$. The channel gain between the transmitter of D2D link j and the BS is denoted by $g_j^{b,d}$. The channel gain between the cellular user i and the receiver of D2D link j is denoted by $g_{i,j}^{c,d}$. The channel gain of the D2D link j is denoted by $g_j^{d,d}$. The pathloss model for D2D links is $P=148+40\log_{10}(d/1000)$ and that for cellular links is $P=128.1+37.6\log_{10}(d/1000)$ [9]. And Rayleigh fading model for small-scale part is also considered.

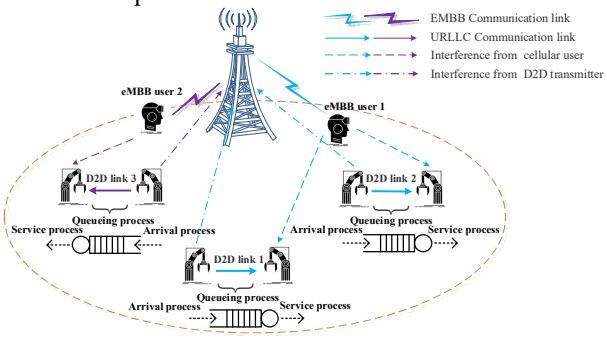

Fig. 1 The communications scenario

In this paper, the TTI of eMBB traffic is 1 ms, while the TTI of URLLC traffic is ωt_s, and $\omega t_s < 1$ ms. ω denotes the number of OFDM symbols occupied by each D2D link. t_s denotes the duration of one OFDM symbol. An example of the system model is depicted in the Fig.1. where the D2D link 1 and D2D link 2 multiplex the UL resources of eMBB user 1. The transmissions of these two D2D links occupy the same subcarriers but different symbols.

3 D2D-based puncturing algorithm for eMBB And URLLC multiplexing scenario

3.1 Analysis on the achievable transmission rates of eMBB and URLLC Traffic

In this paper, $\rho_{i,j}$ is defined to reflect the resources reusing relationship between the cellular link and the D2D link. If the RBs of the cellular user i are reused by the D2D link j, $\rho_{i,j}=1$; otherwise, $\rho_{i,j}=0$. In

our scheme, we regulate that each D2D link shares at most one cellular user's RBs, i.e., $\sum_i \rho_{i,j} \le 1, \forall\, j \in \mathcal{N}$. And $\sum_j \rho_{i,j} \le \lambda, \forall\, i \in \mathcal{M}$. λ is defined as the maximum number of D2D links that are allowed to multiplex on the RBs allocated to one cellular user. We assume that the data in one transmission block of eMBB traffic is jointly coded and decoded. Hence, the average power of interference caused by multiple D2D links is considered in deriving the transmission rate of cellular communication. The average interference is given as

$$N_i^{AVG} = \omega t_s \sum_j \rho_{i,j} P_d g_j^{b,d} / T_f, \lambda \cdot \omega \le 14 \tag{1}$$

$\lambda \cdot \omega \le 14$ holds for the normal cyclic prefix (CP) condition that one TTI of eMBB traffic contains 14 OFDM symbols.

According to the average power of the time-varying interference, the SINR of the cellular user *i* is given as

$$\gamma_i^c = \frac{P_c g_i^{b,c}}{N_0 + N_i^{AVG}} = \frac{P_c g_i^{b,c}}{N_0 + \omega t_s \sum_j \rho_{i,j} P_d g_j^{b,d} / T_f}, \lambda \cdot \omega \le 14 \tag{2}$$

where N_0 denotes noise power of wireless channel. P_c denotes the transmission power of cellular users. Since $\sum_i \rho_{i,j} \le 1$, the SINR of D2D link *j* is given as

$$\gamma_j^d = \frac{\sum_i \rho_{i,j} P_d g_j^{d,d}}{N_0 + \sum_i \rho_{i,j} P_c g_{i,j}^{c,d}} \tag{3}$$

According to Shannon formula, the achievable transmission rate of cellular users *i* can be calculated as

$$R_i^c = B_0 \log_2(1 + \gamma_i^c) \tag{4}$$

where B_0 denotes the bandwidth resource that is allocated to the cellular user *i*.

Because the short data packets with finite block-length channel codes are transmitted in URLLCs, the block error rate (BLER) of the short packet transmission is inevitable. For a given block-length *m* and BLER φ_j, the achievable transmission rate of D2D link *j* is approximated by [10]

$$R_j^d = B_o \left\{ \log_2(1 + \gamma_j^d) - \sqrt{\frac{(2 + \gamma_j^d)\gamma_j^d (\log_2 e)}{(1 + \gamma_j^d)^2 m}} f_Q^{-1}(\varphi_j) \right\} \tag{5}$$

where $f_Q^{-1}(x)$ denotes the inverse function of the Gaussian Q-function. According to our multiplexing scheme, the achievable sum-rate of the system is given as

$$R_o = \sum_i \left\{ R_i^c + \sum_j \rho_{i,j} R_j^d \right\} = B_o \sum_i \left\{ \log_2\left(1 + \gamma_i^c\right) + \sum_j \rho_{i,j} \left\{ \log_2\left(1 + \gamma_j^d\right) - \sqrt{\frac{\left(2 + \gamma_j^d\right)\gamma_j^d \left(\log_2 e\right)}{\left(1 + \gamma_j^d\right)^2 m}} f_Q^{-1}(\varphi_j) \right\} \right\} \tag{6}$$

3.2 Statistical delay QoS analysis of URLLCs

In this paper, the arrival process and the service process of dynamic queuing system are assumed to be stationary ergodic stochastic processes. Let $Q_j(\infty)$ denotes the stationary queue length of the transmitter of D2D link *j*. $Q_j^{\max}$ denotes the threshold of queue length. On the basis of large deviation principle (LDP), the relationship holds as follow [11]

$$-\lim_{Q_j \to \infty} \frac{\log\left(\Pr\left\{Q_j(\infty) > Q_j^{\max}\right\}\right)}{Q_j^{\max}} = \theta_j \tag{7}$$

where θ_j is QoS exponent of the URLLC traffic carried by the D2D link *j*. The larger the θ_j is, the more stringent the QoS requirement is.

The statistical delay QoS requirement of the URLLC traffic carried by the D2D link *j* is defined as $(D_j^{\max}, \varepsilon_j)$, where $D_j^{\max}$ and ε_j denote the delay bound and the maximum allowable delay violation

probability, respectively. Let $D_j(\infty)$ denote the stationary delay of the packets transmitting on the D2D link *j*. The statistical delay QoS requirement holds for $\Pr\{D_j(\infty) > D_j^{\max}\} \le \varepsilon_j$. From [12], the approximated function of delay violation probability is given as

$$\Pr\{D_j(\infty) > D_j^{\max}\} \approx \delta_j \exp\{-\theta_j E_j^B(\theta_j) D_j^{\max}\} \tag{8}$$

δ_j denotes the probability of the event that the buffer of the transmitter of D2D link *j* is non-empty, and $\delta_j \le 1$. Then

$$\Pr\{D_j(\infty) > D_j^{\max}\} \le \exp\{-\theta_j E_j^B(\theta_j) D_j^{\max}\} \le \varepsilon_j \tag{9}$$

$E_j^B(\theta)$ denotes the effective bandwidth of URLLC traffic carried by the D2D link *j*. When the average arrival rate is α_j, the effective bandwidth for a Poisson process is

$$E_j^B(\theta_j) = \frac{\ln(1/\varepsilon_j)}{D_j^{\max} \ln\left[1 + \frac{\ln(1/\varepsilon_j)}{\alpha_j D_j^{\max}}\right]} \tag{10}$$

The effective capacity characterizes the maximum constant arrival rate that a system could support so as to satisfy the statistical delay QoS requirement [13]. The transmissions of D2D links are regarded as service process in this paper. The effective capacity of D2D link *j* is given as

$$E_j^C(\theta_j) = -\frac{1}{\theta_j} \ln\left[E\left(e^{-\theta_j R_j^d}\right)\right] \tag{11}$$

For URLLCs, the delay bound is much smaller than the coherence time of wireless channel in typical scenarios except high mobility scenarios where the velocity is larger than 100 km/h. Hence, the channel is referred to as quasi-static fading channel, the random fading coefficients of which keep constant in the duration of transmitting one data packet of URLLC traffic. Then, the effective capacity of D2D link *j* carrying URLLC traffic is derived as

$$E_j^C(\theta_j) = B_o \left\{ \log_2(1+\gamma_j^d) - \sqrt{\frac{(2+\gamma_j^d)\gamma_j^d (\log_2 e)}{(1+\gamma_j^d)^2 m}} f_Q^{-1}(\varphi_j) \right\} \tag{12}$$

According to [14], for statistical independent arrival and service processes, the statistical delay QoS requirement characterized by θ_j can be guaranteed, if and only if the effective capacity is greater than effective bandwidth, i.e.,

$$E_j^C(\theta_j) \ge E_j^B(\theta_j) \tag{13}$$

3.3 Problem Formulation

In this paper, we formulate the spectrum-efficient D2D-based puncturing algorithm as a system throughput maximization problem with multiple constraints including the statistical delay QoS constraints of URLLCs, which is formulated as

$$\max_{P_c, P_d, \rho_{i,j}} \sum_i \left\{ R_i^c + \sum_j \rho_{i,j} R_j^d \right\} \tag{14}$$

$$\text{s.t.} \quad \sum_i \rho_{i,j} \le 1, \rho_{i,j} \in \{0,1\}, \forall j \in \mathcal{N} \tag{14.a}$$

$$\sum_j \rho_{i,j} \le \lambda, \rho_{i,j} \in \{0,1\}, \forall i \in \mathcal{M} \tag{14.b}$$

$$P_c \le P_c^{\max}, P_d \le P_d^{\max} \tag{14.c}$$

$$\gamma_i^c \ge \gamma_c^{\min}, \forall i \in \mathcal{M} \tag{14.d}$$

$$E_j^C(\theta_j) \ge E_j^B(\theta_j), \forall j \in \mathcal{N} \tag{14.e}$$

Constraints in (14.a) restrict that each D2D link only reuse at most one cellular user's bandwidth resource. Constraints in (14.b) allow each cellular user's RBs to be multiplexed by at most λ D2D link(s). Constraints in (14.c) limit the transmitting power of cellular users and D2D links. Constraints in (14.d) guarantee the SINR of cellular link which reflect to the rate constraints of eMBB traffic. Lastly, the

constraints in (14.e) hold when the statistical delay QoS requirement of URLLC traffic is satisfied. In this paper, we employ the PSO algorithm with penalty function to cope with the optimization problem.

4. Simulation Results

In our simulations, M (=10) cellular users with eMBB traffic and $2N$ (=10~40) URLLC terminals (communicating via D2D links) are uniformly and randomly distributed in a cell (with radius of 50 m and 100 m respectively). The maximal transmit power $P_d^{\max}$ of D2D transmitter is set to be 0.1 w. The required minimal SINR of cellular user is 0.2. The delay and reliability QoS requirement of URLLC traffic is 0.8 ms and 1-10^{-5} respectively. The numerical results are averaged over 500 times.

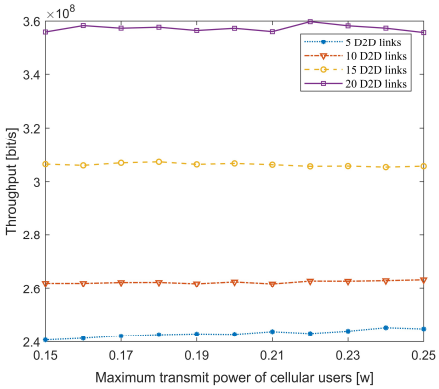

Fig.2 The throughput versus the maximum transmitting power of eMBB users.

Fig. 2 shows the system throughput of our D2D-based puncturing algorithm versus the allowable maximum transmitting power of cellular users. It shows that the system throughput is unaffected by the maximum transmitting power in case of more (i.e., $N > 5$) D2D links existing. When $N = 5$, the system throughput increases lightly along with the maximum transmitting power of cellular users. On the one hand, a larger transmitting power of cellular users with eMBB traffic can elevate the transmission rate of cellular link. On the other hand, a larger transmitting power also causes more interference to D2D links with URLLC traffic. When the number of D2D links is small, the larger transmitting power is taken by our algorithm for the cellular user to improve their transmission rates. When there are more D2D links, a moderate transmitting power is obtained by our algorithm for reducing interference and satisfying the statistical delay QoS of URLLCs.

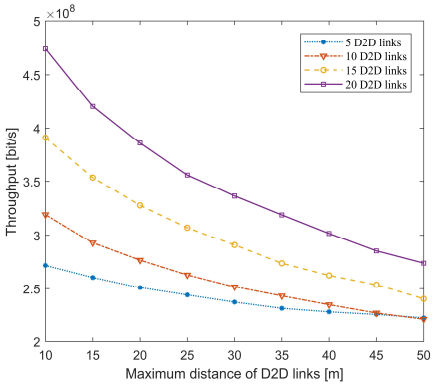

Fig. 3 The throughput versus the maximum distance of D2D links.

Fig. 3 shows the system throughput of our D2D-based puncturing algorithm in the scenarios with different maximum communication distances of the D2D links. From Fig. 3, we can find that increasing the maximum communication distance of the D2D link makes the system throughput decrease. Further, as Fig. 3 shown, the decreasing rate of the throughput curve for the case of 5 D2D links is obviously lower than others. In case of existing 5 D2D links, the D2D transmissions cause less interference than other cases. Hence, for the case of $N = 5$, increasing the maximum distance for the D2D link does not result in a dramatic reduction in system throughput. Fig. 4 depicts the system throughput versus the average arrival packet rate

of URLLC traffic in each D2D link. It shows that packet arrival rate has little effect on the system throughput when the arrival rate is smaller than 0.1. It is indicated that our puncturing algorithm is robust to the variation of the average arrival rates.

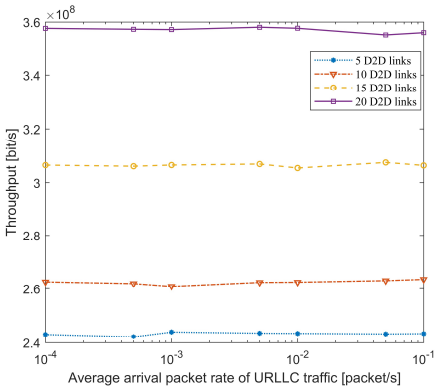

Fig. 4 The throughput versus the average arrival packet rate of URLLC traffic.

5. Conclusions

In this paper, we proposed a spectrum-efficient D2D-based puncturing algorithm for eMBB and URLLC traffic co-existence networks, where D2D links with URLLC traffic reused the bandwidth resources allocated to the cellular link carrying eMBB traffic. Our novel proposed puncturing algorithm relied on the effective capability theory to provision the heterogeneous QoS guarantees of eMBB and URLLC traffic, and achieved high throughput of system at the same time.

References

[1] Study on New Radio Access Technology (Release 14) V14.0.0, document TR 38.801, 3GPP, Mar. 2017.

[2] Chairman's notes 3GPP: 3GPP TSG RAN WG1 Meeting 88bis, Available at http://www.3gpp.org/ftp/TSG RAN/WG1 RL1/TSGR1 88b/Report/, April 2017

[3] A. Anand, G.: De Veciana, S. Shakkottai. Joint Scheduling of URLLC and eMBB Traffic in 5G Wireless Networks. Proc. IEEE INFOCOM. pp. 1970–1978(2018).

[4] Z. Wu, F. Zhao, X. Liu.: Signal space diversity aided dynamic multiplexing for eMBB and URLLC traffics. IEEE ICCC. pp. 1396–140 (2018).

[5] A.A. Esswie, K.I. Pedersen.: Multi-User Preemptive Scheduling for Critical Low Latency Communications in 5G Networks. Proc. IEEE ISCC. pp. 136–141(2018).

[6] C. She, C. Yang.: Available Range of Different Transmission Modes for Ultra-Reliable and Low-Latency Communications. IEEE VTC Spring. pp. 1-5(2017).

[7] B. Singh, Z. Li, M.A. Uusitalo.: Flexible resource allocation for device-to-device communication in fdd system for ultra-reliable and low latency communications. Proc. IEEE RTUWO. pp. 186-191(2017).

[8] L. Liu, W. Yu.: A D2D-Based Protocol for Ultra-Reliable Wireless Communications for Industrial Automation. IEEE Trans. Wireless Commun. pp. 5045–5058(2018).

[9] L. Su, Y. Ji, P. Wang, F. Liu.: Resource allocation using particle swarm optimization for D2D communication underlay of cellular networks, IEEE WCNC. pp. 129–133(2013).

[10] Y. Polyanskiy, H.V. Poor, S. Verdú.: Channel coding rate in the finite blocklength regime. IEEE Trans. Inform. Theory. pp. 2307–2359(2010).

[11] X. Zhang, J. Wang.: Statistical QoS-driven power adaptation for distributed caching based mobile offloading over 5G wireless networks. IEEE INFOCOM WKSHPS. pp. 486–491(2018).

[12] C. She, C. Yang, T.Q.S. Quek.: Cross-Layer Optimization for Ultra-Reliable and Low-Latency Radio Access Networks. IEEE Trans. Wireless Commun. pp. 127–141(2018).

[13] D. Wu, R. Negi.: Effective capacity: A wireless link model for support of quality of service. IEEE Trans. Wireless Commun. pp. 630–643(2003).

[14] L. Zhao, X. Chi, S. Yang.: Optimal ALOHA-Like Random Access with Heterogeneous QoS Guarantees for Multi-Packet Reception Aided Visible Light Communications. IEEE Trans. Wireless Commun. pp. 7872–7884(2016).