Cooperative Wireless Edge Caching with Relay Selection

Constantinos Psomas*, Gan Zheng†, and Ioannis Krikidis*

*Department of Electrical and Computer Engineering, University of Cyprus, Cyprus

†Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, UK

e-mail: psomas@ucy.ac.cy, g.zheng@lboro.ac.uk, krikidis@ucy.ac.cy

***Abstract*—Relay selection is a simple yet effective means to improve the reliability and coverage of wireless cooperative networks. However, it suffers from inefficient use of the available bandwidth resources. This paper introduces the use of content caching at relays in order to tackle this problem and improve the performance of relay selection. Three cache placement schemes are considered: one based on the most popular files, a uniform based caching and a hybrid scheme of the two. We analytically derive their outage performance as well as their diversity order and coding gain. Numerical results demonstrate the substantial performance gains of these schemes over the traditional optimal relay selection approach without caching capabilities.**

I. Introduction

Relay selection is a low-complexity yet effective means to reap the benefits of cooperative communications [1], [2]. In traditional half-duplex relay-assisted communications, the selected relays first receive the information from the source, then forward it to the destination, so extra bandwidth resources are needed. To solve this problem, the most promising approach in the literature is full-duplex relaying, which allows the relays to receive the information from the source and forward it to the destination simultaneously over the same frequency band [3]. However, the main drawback of full-duplex relaying is the self-interference which not only requires complex analog and/or digital cancellation techniques to mitigate it, but it may also lead to loss of diversity gain compared to the half-duplex relaying [4]. Recently, wireless edge caching has received much interest due to its ability to tackle the wireless resource scarcity in content centric networks [5], [6]. The main idea is that by caching popular contents at the access points and/or the user devices, any requests of those contents can be served locally. Therefore, conventional issues such as high backhaul bandwidth requirements and excess delay can be significantly alleviated. Various caching algorithms have been studied in [6] and [7], which assume error-free transmissions over the wireless medium.

Applying this concept to relay-assisted cooperative communications, relays can store the popular contents and serve the destination's requests directly rather than receiving and forwarding them from the source. As a result, similar benefits to full-duplex relaying can be achieved since the overall transmission time can be reduced approximately by half. Moreover, performance degradation due to the self-interference does not apply in this case and thus the system efficiency can be further improved. Despite the obvious benefits emerging from caching at relays, very few works exist in this area. For instance, it is shown in [8], that caching at the relays can create more Multiple-Input Multiple-Output (MIMO) broadcast opportunities such that degrees of freedom (DoF) gains can be achieved. Despite its advantages, relay caching also brings new design challenges. First, although the cost of storage is getting lower, in reality, relays have limited caching capacity and cannot store all requested contents; second, caching design is not an isolated issue but is coupled with the transmission schemes and affected by channel fading. Therefore, caching placement, i.e., which contents are stored at which relays, needs to be jointly designed with the transmission schemes. Existing caching schemes such as caching the most popular contents (MPC) at all relays to achieve the signal diversity gain and uniform caching (UC) at different relays to achieve the largest content diversity gain cannot achieve the optimal caching performance for wireless networks. To balance the signal diversity gain and caching diversity gain, joint caching placement and transmission adaption have been studied in base station cooperation and small cell networks [9], [10]. However, to the best of our knowledge, no existing work considers caching as a potential solution to tackle the inefficiency of traditional relay selection schemes and its performance gain is not well understood, and this is the main motivation of the current paper.

In this paper, we introduce caching as a new solution to improve the performance of the optimal amplify-and-forward (AF) relay selection scheme and analyze the outage performance of different caching placement solutions, including MPC, UC and a hybrid scheme. The hybrid scheme combines the benefits of MPC and UC, by caching some most popular contents at all relays and then uniformly stores some less popular contents in the rest of the storage space. Our analysis reveals that the MPC and hybrid caching scheme can achieve full diversity gain and substantial coding gains over the traditional relay selection. Moreover, the optimized hybrid scheme demonstrates superior performance against the other schemes. Simulation results verify our analysis and show that caching can indeed significantly improve the relay

This work was supported in part by the Research Promotion Foundation, Cyprus under the project KOINA/ERANETMED/1114/03 and by the UK Engineering and Physical Sciences Research Council (EPSRC) project EP/N007840/1.

selection performance when compared with the traditional relay selection, even when a small amount of storage is available at the relays.

II. System Model

Consider a clustered network with a source $\mathtt{S}$, a destination $\mathtt{D}$ and a cluster of K $(K > 1)$ relays $\{\mathtt{R}_k\}_{k=1}^{K}$, each with a single antenna. A two-hop AF protocol is employed and we assume that no direct link between $\mathtt{S}$ and $\mathtt{D}$ exists due to strong path attenuation. The source and relays' transmit powers are P_s and P_r, respectively. The $\mathtt{S}-\mathtt{R}_k$ and $\mathtt{D}-\mathtt{R}_k$ channels are denoted as h_k and g_k, and the corresponding distances are l_k and d_k, respectively. The channels are modelled as $h_k = \sqrt{l_k^{-\alpha}}\tilde{h}_k$ and $g_k = \sqrt{d_k^{-\alpha}}\tilde{g}_k$, where $\tilde{h}_k, \tilde{g}_k$ are complex Gaussian random variables with zero mean and unit variance and $\alpha > 2$ is the path loss exponent. Without loss of generality, the relays are ordered by the distance $\{d_k\}$ to $\mathtt{D}$, i.e., $\mathtt{R}_1$ is the closest one to $\mathtt{D}$. All noise components at the receiving nodes follow a Gaussian distribution with zero mean and unit variance. Therefore, the received signal-to-noise (SNR) via relay k is

$$\Gamma_k = \frac{P_s|h_k|^2 P_r|g_k|^2}{1 + P_s|h_k|^2 + P_r|g_k|^2}, \tag{1}$$

where $|h_k|^2$ and $|h_k|^2$ are exponential random variables with parameters l_k^{α} and d_k^{α} respectively.

Each relay $\mathtt{R}_k$ is assumed to employ caching storage and assist the source $\mathtt{S}$ with the transmission of the destination's requested files. Time is considered to be slotted and at each time slot, one relay from the cluster is selected to assist the source $\mathtt{S}$. We consider different caching placement schemes with a single relay selection policy which are presented in the next section. For comparison purposes, we also consider the traditional relay selection policy without caching. Other relay selection policies, such as the max-min relay selection scheme [1], can be considered but are out of the scope of this paper. Moreover, each policy selects a single relay but this work can be extended to take into account multiple relay selection.

The optimal traditional relay selection policy chooses the relay that achieves the maximum instantaneous capacity. In mathematical terms, this is written as $k^* = \arg_k \max \Gamma_k$ where Γ_k is given by (1). Therefore, the outage performance of the traditional optimal relay selection is

$$p^{\mathrm{TRAD}} = \text{Prob}\left(\frac{1}{2}\log_2\left(1 + \Gamma_{k^*}\right) \le R_F\right), \tag{2}$$

where R_F is the target threshold rate.

III. Caching Model and Placement Schemes

In this section, we present the considered caching model and the placement schemes for the cached files.

A. Caching Model

Let $\mathcal{N} = \{1, 2, \dots, N\}$ denote a set of N files of the same size S, each of which is randomly requested based on a popularity distribution. The n-th file has popularity profile f_n, where f_n is sorted in descending order with $\sum_{n=1}^{N} f_n = 1$. We assume that f_n follows Zipf distribution [11], which is commonly used to model the content popularity. According to the Zipf distribution, the request probability of the n-th most popular file is given by

$$f_n = \frac{n^{-\gamma}}{\sum\limits_{m=1}^{N} m^{-\gamma}}, \tag{3}$$

where γ is the Zipf parameter. A large γ means that the requests concentrate on the most popular high-rank files whereas a small γ depicts a heavy-tailed popularity distribution.

Suppose each relay has a limited storage capacity (in the unit of number of files) $C \ll N$. Because of the storage limit, relays cannot store all files but need to choose judiciously which files to store, i.e., $CK < N$. The requested file is obtained straight from the relays if it's contained in the caching storage; otherwise, the file is sent from $\mathtt{S}$ to $\mathtt{D}$ via the relays in the traditional way.

B. Cache Placement Schemes

We first consider two simple placement schemes: storing the same files and storing different files through the caching storage of the relays. The former allows $\mathtt{D}$ to choose the best relay which caches the most popular C contents, while the latter maximizes the hit probability that a file can be found at relays without resorting to $\mathtt{S}$.

- Most popular content (MPC) caching. All relays store the same C most popular contents. For content files with index $n, 1 \le n \le C$, $\mathtt{D}$ will select the relay with the best $\mathtt{R}-\mathtt{D}$ channel, i.e., $k^* = \arg_k \max |g_k|^2$. Hence, its outage is given by

  $$p^{\mathrm{MPC}} = \text{Prob}\left(\log_2(1 + P_r|g_{k^*}|^2) \le R_F\right), \tag{4}$$

  where $|g_{k^*}|^2$ is an exponentially distributed random variable with parameter $d_{k^*}^{\alpha}$. Here, the prelog factor is 1 since communications are established in one channel use. The remaining content files which are not cached at the relays, will be obtained by the source and the relay selection employed, in this case, is the traditional one.
- Uniform caching (UC). Relay $\mathtt{R}_k$ stores the files with index $n, (k-1)C + 1 \le n \le kC$, based on the relay distance information, and no file is cached in more than one relay in order to achieve the maximal caching diversity. As such, the n-th file is stored at $\mathtt{R}_{\lceil n/C \rceil}$. In this case, $\mathtt{D}$ simply selects the relay which possesses the required file without considering the channel conditions. The outage probability of the UC scheme is

  $$p_k^{\mathrm{UC}} = \text{Prob}\left(\log_2(1 + P_r|g_k|^2) \le R_F\right), \tag{5}$$

  where $|g_k|^2$ is an exponentially distributed random variable with parameter d_k^{α}. Similarly to MPC, any file not cached will be acquired from the source using the traditional relay selection policy.

Finally, we consider a hybrid caching placement scheme. We can see that MPC provides the best selection diversity but can only cache up to C files in total, while UC can achieve the maximal content diversity gain by caching the KC most popular content but the relay who stores a specific file may not provide the best link. Based on these observations, we employ a group-based hybrid caching strategy to balance the relay selection diversity gain and the content diversity gain. Specifically, the content files are divided into three groups:

- Group 1: $\mathcal{N}_1 = \{1, \ldots, M\}$. The M most popular files are stored at all K relays where M is a parameter to optimize. For those files, the outage probability is p^{MPC}.
- Group 2: $\mathcal{N}_2 = \{M+1, \ldots, I(K, M)\}$. Relays $\mathtt{R}_1, \ldots, \mathtt{R}_K$ take turns and each stores the next most popular distinct $(C-M)$ files to fill in its storage space. In general, $\mathtt{R}_k$ stores the files in $\mathcal{N}_{2k} = \{I(k,M)+M-C+1, \ldots, I(k,M)\}$, where $I(k,M) \triangleq k(C-M)+M$. For the files in $\mathcal{N}_{2k}$, the outage probability is p_k^{UC}.
- Group 3: $\mathcal{N}_3 = \{I(K,M)+1, \ldots, N\}$. The remaining $N - I(K,M)$ less popular files are not cached at the relays and must be forwarded from $\mathtt{S}$ to $\mathtt{D}$ through the relays using the traditional non-caching relay selection policy. The outage, in this case, is p^{TRAD}.

It is obvious that when $M = C$ and $M = 0$, the hybrid scheme reduces to the MPC and the UC scheme, respectively.

IV. Performance Analysis

In this section, we analyze the outage performance of both the non-caching and caching schemes given in Sections II and III. In order to simplify the notation, define $\lambda_{s,k} \triangleq \frac{1}{P_s l_k^{-\alpha}}$ and $\lambda_{r,k} \triangleq \frac{1}{P_r d_k^{-\alpha}}$.

A. Traditional Scheme (No Caching)

Proposition 1. *The outage probability of the traditional relay selection is given by*

$$p^{\text{TRAD}} = \prod_{k=1}^{K} \left(1 - 2e^{-(\lambda_{s,k}+\lambda_{r,k})\tau} \sqrt{\lambda_{s,k}\lambda_{r,k}\tau(1+\tau)} \times K_1\left(2\sqrt{\lambda_{s,k}\lambda_{r,k}\tau(1+\tau)}\right)\right), \quad (6)$$

where $\tau \triangleq 2^{2R_F} - 1$ and $K_1(\cdot)$ is the first-order modified Bessel function of the second kind.

Proof. The cumulative distribution function (CDF) of the optimal relay SNR is $\Gamma_{k^*} = \max_{1\le k \le K} \Gamma_k$. By using higher order statistics, the CDF can be evaluated as $F_{\Gamma_{k^*}}(\tau) = \prod_{k=1}^{K} F_{\Gamma_k}(\tau)$ where $F_{\Gamma_k}(\tau)$ is given by [12]

$$F_{\Gamma_k}(\tau) = 1 - 2e^{-(\lambda_{s,k}+\lambda_{r,k})\tau} \sqrt{\lambda_{s,k}\lambda_{r,k}\tau(1+\tau)} \times K_1\left(2\sqrt{\lambda_{s,k}\lambda_{r,k}\tau(1+\tau)}\right),$$

and the result follows. □

In the high SNR regime, i.e. when $P = P_s, P_r \to \infty$, the above expression can be simplified as

$$p_{\infty}^{\text{TRAD}} = \prod_{k=1}^{K} \left(\frac{l_k^{\alpha}}{P} + \frac{d_k^{\alpha}}{P} \right) \tau \quad (7)$$

$$\to O(1/P^K),$$

which follows from the approximations $1 - e^{-x} \approx x$ and $K_1(x) \approx 1/x$ for small x. From (7), we can see that the traditional scheme achieves diversity gain equal to the number of available relays K. Therefore, its coding gain [14, Eq. (3.158)] is

$$G^{\text{TRAD}} = \tau^K \prod_{k=1}^{K} \left(l_k^{\alpha} + d_k^{\alpha} \right). \quad (8)$$

B. MPC Scheme

Proposition 2. *The outage probability of the MPC cache placement scheme is given by*

$$\Pi^{\text{MPC}} = \prod_{k=1}^{K} \left(1 - e^{-\lambda_{r,k}\tau'} \right) \sum_{n=1}^{C} f_n + p^{\text{TRAD}} \sum_{n=C+1}^{N} f_n, \quad (9)$$

where $\tau' \triangleq 2^{R_F} - 1$ and p^{TRAD} is given by (6).

Proof. The probability that $\mathtt{D}$ obtains the requested file from a relay is given by $\sum_{n=1}^{C} f_n$, as all relays in the MPC scheme store the same C most popular files. In this case, $\mathtt{D}$ selects the relay with the best $\mathtt{R}-\mathtt{D}$ channel and so from (4) the outage probability is given by

$$p^{\text{MPC}} = \prod_{k=1}^{K} \text{Prob}\left(P_r |g_k|^2 \le \tau' \right) = \prod_{k=1}^{K} \left(1 - e^{-\lambda_{r,k}\tau'} \right), \quad (10)$$

which is derived by using higher order statistics and the fact that $P_r|g_k|^2$ is exponentially distributed with parameter $\lambda_{r,k}$. On the other hand, if the requested file is not cached, the outage probability is given by p^{TRAD} and thus the result follows. □

In this case, when $P = P_s, P_r \to \infty$, (9) converges to

$$\Pi_{\infty}^{\text{MPC}} = \prod_{k=1}^{K} \frac{d_k^{\alpha}}{P} \tau' \sum_{n=1}^{C} f_n + \prod_{k=1}^{K} \left(\frac{l_k^{\alpha}}{P} + \frac{d_k^{\alpha}}{P} \right) \tau \sum_{n=C+1}^{N} f_n \quad (11)$$

$$\to O(1/P^K),$$

using the same approximations as before. As with the traditional scheme, the MPC scheme provides diversity gain equal to K. The coding gain for this scheme is

$$G^{\text{MPC}} = \tau'^K \prod_{k=1}^{K} d_k^{\alpha} \sum_{n=1}^{C} f_n + \tau^K \prod_{k=1}^{K} \left(l_k^{\alpha} + d_k^{\alpha} \right) \sum_{n=C+1}^{N} f_n \quad (12)$$

$$\le G^{\text{TRAD}},$$

where G^{TRAD} is given by (8); the inequality holds as $\tau^K \prod_{k=1}^{K} (l_k^{\alpha} + d_k^{\alpha}) \ge \tau'^K \prod_{k=1}^{K} d_k^{\alpha}$ and $\sum_{n=1}^{C} f_n + \sum_{n=C+1}^{N} f_n = 1$. Therefore, the coding gain gap is

$$G^{\text{TRAD}} - G^{\text{MPC}} = \sum_{n=1}^{C} f_n \left(\tau^K \prod_{k=1}^{K} (l_k^{\alpha} + d_k^{\alpha}) - \tau'^K \prod_{k=1}^{K} d_k^{\alpha} \right). \quad (13)$$

C. UC Scheme

Proposition 3. *The outage probability of the UC cache placement scheme is given by*

$$\Pi^{\text{UC}} = \sum_{k=1}^{K} \left(1 - e^{-\lambda_{r,k}\tau'}\right) \sum_{n=(k-1)C+1}^{kC} f_n + p^{\text{TRAD}} \sum_{n=KC+1}^{N} f_n, \tag{14}$$

where $\tau' \triangleq 2^{R_F} - 1$ and p^{TRAD} is given by (6).

Proof. In this case, D connects to the relay that has cached the requested file. Therefore, from (5) and conditioning on k we have

$$p_k^{\text{UC}} = \text{Prob}\left(P_r|g_k|^2 \leq \tau'\right) = 1 - e^{-\lambda_{r,k}\tau'}, \tag{15}$$

as $P_r|g_k|^2$ is an exponential random variable with parameter $\lambda_{r,k}$. The rest of the proof follows similarly to the one of Proposition 2, the difference being that the outage probability of the R − D channel is taken over all K relay channels and their cached files. □

Similarly to above, we derive the diversity of the UC scheme. For $P = P_r, P_s \to \infty$, (14) gives

$$\Pi_{\infty}^{\text{UC}} = \sum_{k=1}^{K} \frac{d_k^{\alpha}}{P}\tau' \sum_{n=(k-1)C+1}^{kC} f_n + p_{\infty}^{\text{TRAD}} \sum_{n=KC+1}^{N} f_n, \tag{16}$$

where p_{∞}^{TRAD} is given by (7). It's clear that (16) is a sum of two terms of different orders, 1 and K. As the one with the smallest order dominates the other, the UC scheme achieves a diversity order of one.

D. Hybrid Caching Scheme

The performance of the hybrid scheme depends on the parameter M which specifies the weight given to each placement scheme. Hence, the outage probability of the hybrid scheme is given by

$$\Pi^{\text{HY}}(M) = p^{\text{MPC}} \sum_{n=1}^{M} f_n + p^{\text{TRAD}} \sum_{n=I(K,M)+1}^{N} f_n + \sum_{k=1}^{K} p_k^{\text{UC}} \sum_{n=I(k,M)+M-C+1}^{I(k,M)} f_n, \tag{17}$$

where p^{TRAD}, p^{MPC} and p^{UC} are given by (6), (15) and (15) respectively. The optimal M^* can be found by exhaustive search over $M, 1 \leq M \leq C$, and its complexity is $\mathcal{O}(C)$.

In the high SNR regime, the above expression simplifies to

$$\Pi_{\infty}^{\text{HY}}(M) = \frac{1}{P^K}\left(\tau'^K \prod_{k=1}^{K} d_k^{\alpha} \sum_{n=1}^{M} f_n + \tau^K \prod_{k=1}^{K} (l_k^{\alpha} + d_k^{\alpha}) \sum_{n=I(K,M)+1}^{N} f_n\right) + \frac{\tau'}{P} \sum_{k=1}^{K} d_k^{\alpha} \sum_{n=I(k,M)+M-C+1}^{I(k,M)} f_n. \tag{18}$$

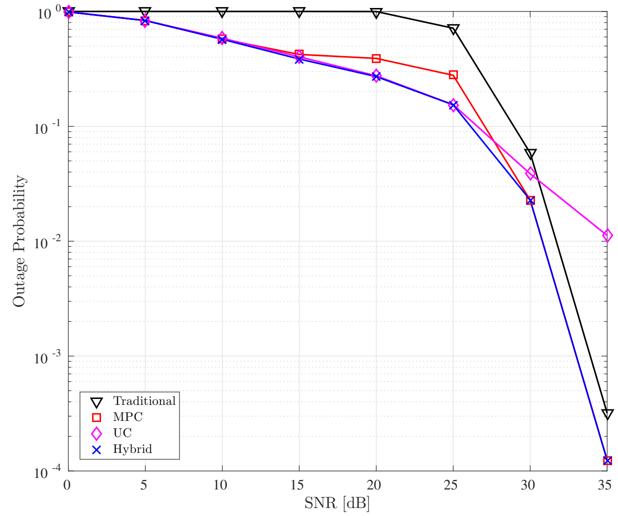

Fig. 1. Comparison of outage probability for the different schemes; $K = 6$, $\gamma = 0.9$, $R_F = 2.5$ bps/Hz; solid lines depict the theoretical results.

It's clear that the first term in the above expression, drops to zero at a faster rate compared to the second term. As such, in this case, the optimal $M^* = C$ which corresponds to the MPC scheme. Therefore, we can conclude that the hybrid scheme has diversity of order K and coding gain equal to (12).

V. Numerical Results

In this section, numerical results are presented to evaluate the performance of the considered schemes and illustrate the benefits of caching for relay selection as well as the impact of the key system parameters. We assume there are $K = 6$ relays and that $\alpha = 2.5$. The $\text{S} - \text{R}_k$ distance is $l_k = (K - k + 1)/2$, and the $\text{R}_k - \text{D}$ distance is $d_k = k/2$. The source SNR is defined as P_s and we assume the source has much higher transmit power than the relays, i.e., $P_s = 5P_r$. We consider a content catalog with $N = 1000$ files and the data rate threshold is $R_F = 2.5$ bps/Hz. Unless otherwise stated, the cache capacity per relay is $C = 100$ files, so the overall relay storage is 60% of the total catalog size.

In Figs. 1 and 2, the outage probability against the source SNR is provided with Zipf parameters $\gamma = 0.9$ and $\gamma = 2$, respectively. Firstly, note that there is a perfect match between the theoretical results (solid lines) and the simulation results (markers) which validate our analytical expressions. As can be seen from both figures, all the caching schemes offer more than a 10 dB SNR gain over the traditional relay selection without caching, in the low-to-medium SNR regimes. It can also be observed, that both the hybrid and MPC caching schemes preserve the full diversity order as the traditional relay selection, while the UC loses the diversity. By comparing the MPC and UC schemes, we can see that for $\gamma = 0.9$, the UC scheme outperforms the MPC scheme in the low-to-medium SNR regime as expected, since for low Zipf parameters the requests become more uniform and so the UC scheme performs well due to its content diversity. On the other hand, for $\gamma = 2$, the MPC is clearly superior since for high Zipf parameters

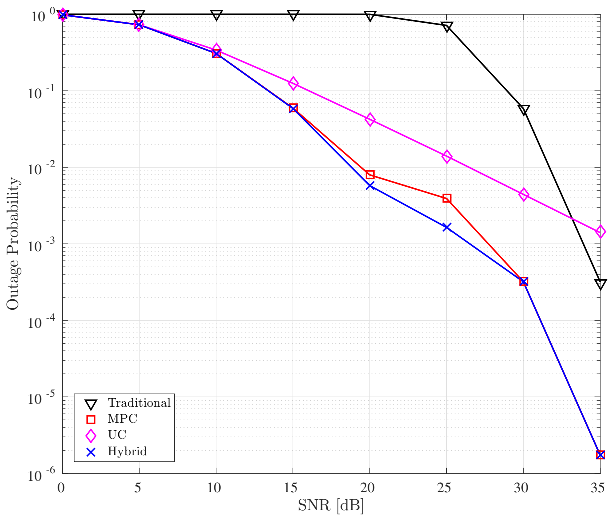

Fig. 2. Comparison of outage probability with different schemes; $K = 6$, $\gamma = 2$, $R_F = 2.5$ bps/Hz; solid lines depict the theoretical results.

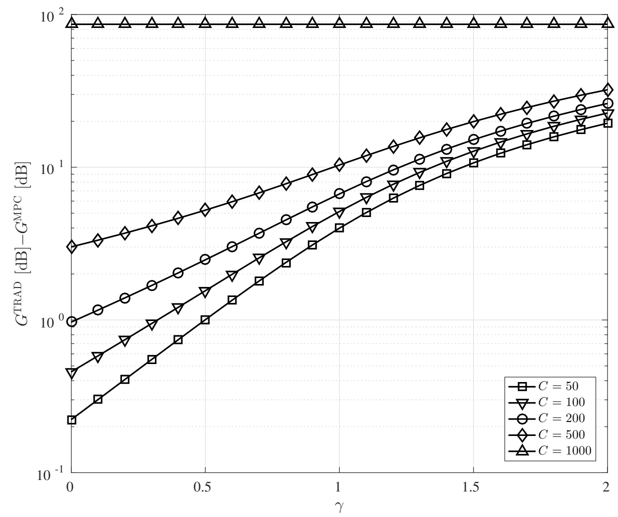

Fig. 3. Coding gain gap between the traditional and MPC schemes; $K = 6$, $R_F = 2.5$ bps/Hz.

the requests focus on the most popular files. In the high SNR regime, MPC performs better than UC irrespective of the Zipf parameter due to the achieved full diversity; both schemes perform equally well for very low SNR values. Furthermore, the hybrid scheme can achieve a fine balance between the two schemes. For small γ, the hybrid scheme employs the UC scheme and switches to MPC for high SNR values. For large γ, the hybrid scheme mainly follows the MPC scheme but combines both MPC and UC for intermediate SNR values to take advantage of the file diversity. Fig. 3 depicts the coding gain gap between the traditional relay selection scheme and the MPC scheme, given by (13), for different values of γ and C. First, observe that the MPC outperforms the traditional scheme and the coding gain increases with γ; this is expected, since the larger γ is the better the MPC performs. Note that these observations are in line with the outage curves in Figs. 1 and 2. Moreover, the coding gain gap increases as the storage space C increases since a larger storage space will increase the file diversity of the MPC scheme. The upper bound is given when the caching capacity C at each relay is equal to the total catalog size, i.e. $C = N = 1000$. At this point, full file diversity can be achieved for any value of the Zipf parameter γ. Finally, even a small caching capacity (e.g., $C = 50$) can lead to significant coding gains for contents with skewed popularity.

VI. Conclusions

This paper studied the employment of cache placement schemes at relays in order to improve the performance of relay selection in cooperative communications. We presented analytical results for the outage performance, diversity and coding gains of the considered cache placement schemes, namely MPC, UC and hybrid caching. Numerical results demonstrated the significant performance gains, in terms of reducing the outage probability, by employing caching schemes at the relays. A future extension of this work is to consider a coding-based caching scheme.

References

[1] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[2] A. S. Ibrahim, A. K. Sadek, W. Su, and K. J. Ray Liu, "Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2814–2827, July 2008.

[3] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sept. 2011.

[4] I. Krikidis, H. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 4524–4535, Dec. 2012.

[5] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 142–149, Apr. 2013.

[6] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[7] K. Shanmugam, N. Golrezaei, A. Dimakis, A. Molisch, and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.

[8] W. Han, A. Liu and V. Lau, "Degrees of Freedom in Cached MIMO Relay Networks," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 3986–3997, Aug. 2015.

[9] S. H. Chae and W. Choi, "Caching placement in stochastic wireless caching helper networks: Channel selection diversity via caching," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6626–6637, Oct. 2016.

[10] X. Peng, J.-C. Shen, J. Zhang, and K. B. Letaief, "Backhaul-aware caching placement for wireless networks," in *Proc. IEEE Global Commun. Conf.*, San Diego, CA, Dec. 2015, pp. 1–6.

[11] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: Evidence and implications," in *Proc. IEEE Int. Conf. Comput. Commun.*, New York, NY, Mar. 1999, pp. 126–134.

[12] T. A. Tsiftsis, G. K. Karagiannidis, P. T. Mathiopoulos, and S. A. Kotsopoulos, "Nonregenerative dual-hop cooperative links with selection diversity," *EURASIP J. Wireless Commun. Netw.*, vol. 2006, pp. 1–8, Apr. 2006.

[13] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Elsevier, 2007.

[14] D. Tse and P. Viswanath, *Fundamentals of wireless communications*. Cambridge University Press, 2004.