
This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Optimization of hybrid cache placement for collaborative relaying

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1109/LCOMM.2016.2623307>

PUBLISHER

IEEE

VERSION

VoR (Version of Record)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) licence. Full details of this licence are available at: <http://creativecommons.org/licenses/by/3.0/>

LICENCE

CC BY 3.0

REPOSITORY RECORD

Zheng, Gan, Himal A. Suraweera, and Ioannis Krikidis. 2019. "Optimization of Hybrid Cache Placement for Collaborative Relaying". figshare. <https://hdl.handle.net/2134/23448>.

Optimization of Hybrid Cache Placement for Collaborative Relaying

Gan Zheng, *Senior Member, IEEE*, Himel A. Suraweera, *Senior Member, IEEE*,
and Ioannis Krikidis, *Senior Member, IEEE*

Abstract—Traditional wireless multi-hop relaying systems suffer from inefficient use of bandwidth resources. This letter studies the use of content caching at distributed relays to tackle this problem and improve the performance of collaborative relaying. We propose a hybrid caching scheme that is jointly optimized with the transmission schemes to achieve a fine balance between the signal cooperation gain and the caching diversity gain. The optimization problem of cache placement to minimize the outage probability is studied and is shown to be convex. Numerical results demonstrate significant outage performance gains over traditional relaying without caching.

Index Terms—Collaborative relaying, wireless caching, outage probability.

I. INTRODUCTION

COLLABORATIVE relaying is an effective means to improve the performance of wireless communications systems. However, the traditional half-duplex relaying leads to inefficient use of system resources. Over the last decade, significant research efforts on relaying have been focused on the investigation of cooperative techniques that overcome this limitation [1]. Towards this direction, several techniques such as cognitive-radio cooperation [2], full-duplex (FD) transmission [3], [4], and buffer-aided relaying [5], have been reported.

Motivated by the recent advances in wireless edge caching [6], [7], we aim to explore caching to improve the efficiency of relaying systems. Relays, if equipped with caching capabilities to store popular content, can serve the destination directly without first fetching them from the source. The main benefit of relay caching is that it reduces the overall transmission time approximately by half, which is similar to the FD transmission and thus substantially improves the efficiency.

Edge caching has been recently introduced to wireless networks as a promising solution to reduce the reliance on backhaul. However, caching in wireless networks brings new challenges because of the channel fading effects and the sophisticated transmission schemes required. Therefore, caching schemes must be jointly designed with transmission schemes to reap the benefits of caching. Commonly used caching schemes such as caching the most popular

content (MPC) to achieve the cooperation gain and caching different contents to achieve the largest content diversity (LCD) gain are no longer optimal for wireless networks. A proper balance between these two gains needs to be achieved. To this end, several recent works have studied joint caching placement and transmission adaption in base station cooperation and small cell networks [8]–[11]. However, application of caching schemes and corresponding performance benefits in relay networks are still not well understood, and this motivates the current work.

In this letter, we study caching as a novel solution to combat the half-duplex constraint of traditional relaying. Our main contribution is to propose and optimize a hybrid caching scheme together with relay clustering that strikes a balance between the signal cooperation gain achieved by the MPC and the caching diversity gain by the LCD to minimize the outage probability in cooperative relaying systems. Similar to [9]–[11] that consider threshold-based caching schemes, we divide files into different groups adopting different caching strategies according to the content popularity and relay storage limits. However, a fundamental difference from existing work is that caching at individual relays may not always outperform the traditionally collaborative relaying without cache that involves all relays. To tackle this new challenge, we propose relay clustering such that cooperative relays within each formed cluster will always lead to lower outage probability than the traditional collaborative relaying by all relays. We then optimize the number of the most popular files that are stored at all relays to minimize the outage probability. We derive a closed-form solution for a special case and provide useful insights on the impact of system parameters. The proposed caching scheme demonstrates superior performance over the traditional transmission without caching, the existing MPC and LCD caching schemes as well as the work in [10], and achieves a similar performance as the multi-threshold caching solution proposed in [11] with much lower complexity.

II. SYSTEM MODEL

Consider a wireless channel with a source S , a destination D and K ($K > 1$) relays $\{\mathcal{R}_k\}_{k=1}^K$ each with a single antenna, and there is no direct link between S and D . The source and relays' transmit powers are P_s and P_r , respectively. The $S - \mathcal{R}_k$ and $D - \mathcal{R}_k$ channels are denoted as h_k and g_k , and the corresponding Euclidean distances are l_k and d_k , respectively. Without loss of generality, the relays are ordered by the distance $\{d_k\}$ to D , i.e., \mathcal{R}_1 is the closest one to D . All noise components at the receive nodes follow a Gaussian distribution with zero mean and unit variance.

A. Caching Model

Let $\mathcal{X} = \{1, 2, \dots, N\}$ denote a set of N files of the same size S to be randomly requested. File n has popularity

Manuscript received September 2, 2016; revised October 10, 2016; accepted October 24, 2016. Date of publication November 7, 2016; date of current version February 9, 2017. This work was supported in part by the UK Engineering and Physical Sciences Research Council (EPSRC) project EP/N007840/1, and the Research Promotion Foundation, Cyprus, under the project KOINA/ERANETMED/1114/03. The associate editor coordinating the review of this letter and approving it for publication was N. Pappas.

G. Zheng is with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, LE11 3TU, U.K. (e-mail: g.zheng@lboro.ac.uk).

H. A. Suraweera is with the Department of Electrical and Electronic Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka (e-mail: himal@ee.pdn.ac.lk).

I. Krikidis is with the Department of Electrical and Computer Engineering, University of Cyprus, 1678 Nicosia, Cyprus (e-mail: krikidis@ucy.ac.cy).

Digital Object Identifier 10.1109/LCOMM.2016.2623307

profile f_n , where f_n is sorted in the descending order and $\sum_{n=1}^N f_n = 1$. We assume that f_n follows Zipf distribution [12], which is commonly used to model the content popularity. With Zipf distribution, the request probability of the n -th most popular file is given by $f_n = \frac{n^{-\gamma}}{\sum_{m=1}^N m^{-\gamma}}$, where γ is the Zipf parameter. A large γ means that the requests concentrate on the most popular high-rank files.

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., $CK < N$. Storing the same files at multiple relays will allow them to collaboratively transmit the content, while storing different files will maximize the hit probability that a file can be found at relays. Based on these observations, we employ a group-based caching strategy to balance the cooperation gain and the content diversity gain. To be specific, the content files are divided into three groups:

- Group 1: $\mathcal{N}_1 = \{1, \dots, M\}$. The M most popular files are stored at all K relays.
- Group 2: $\mathcal{N}_2 = \{M + 1, \dots, I(L, M)\}$, where L is the number of clusters. For convenience, we have defined $I(l, M) \triangleq l(C - M) + M$. Relays first form L clusters $\{C_1, \dots, C_L\}$ (details will be elaborated later in Section II.C) and those in the same cluster store the same content. Each cluster takes turns to store the next most popular distinct $(C - M)$ files to fill in its relays' storage space. In general, cluster C_l stores files in $\mathcal{N}_{2l} = \{I(l, M) + M - C + 1, \dots, I(l, M)\}$.
- Group 3: $\mathcal{N}_3 = \{M + L(C - M) + 1, \dots, N\}$. The rest $(N - M - L(C - M))$ less popular files are not cached at relays and must be forwarded from S to D by relays using the traditional non-caching transmission.

B. Signal Model

Suppose that each file of size S needs to be delivered in T seconds, so we can define the rate requirement $R_F \triangleq \frac{S}{T}$. Depending on which group a file belongs to, the relays will use different strategies to transmit it. The outage probability can be determined as follows:

- Group 1: When a file in Group \mathcal{N}_1 is requested, all K relays form collaborative beamforming to send it to D using the maximum ratio transmission principle. The outage probability is given by:

$$p^{G1} = \text{Prob} \left(\log_2 \left(1 + P_r \sum_{k=1}^K |g_k|^2 \right) < R_F \right). \quad (1)$$

- Group 2: When a file in Group \mathcal{N}_2 is requested, the corresponding relay cluster that stores it directly transmits it to D using maximum ratio transmission. The outage probability for those files served by C_l is expressed as:

$$p_{C_l}^{G2} = \text{Prob} \left(\log_2 \left(1 + P_r \sum_{k \in C_l} |g_k|^2 \right) < R_F \right). \quad (2)$$

- Group 3: Files in Group \mathcal{N}_3 are not cached at any relay, and a traditional transmission scheme is used where two-hop amplify-and-forward protocol using collaborative

beamforming is employed. In the first phase, S transmits to the relays and in the second phase, all K relays forward the received signal to D using the amplify-and-forward collaborative beamforming. Each relay R_k will compensate the composite channel phases based on the local channel station information h_k and g_k , but not necessarily uses full power and the optimal power control solution is given in [13]. We denote the outage probability of this traditional transmission as p^{G3} .

C. Relay Clustering for Files in Group \mathcal{N}_2

From the above description, we can easily see that $p^{G1} < p^{G3}$, i.e., files in Group \mathcal{N}_1 will experience less outage than those in Group \mathcal{N}_3 because of caching and only a single hop is needed. However, for files in Group \mathcal{N}_2 , there may be only a small number of relays in some clusters, and if caching placement is not carefully optimized, the outage probability could be higher than that of the traditional transmission scheme for files in Group \mathcal{N}_3 where all relays are used and in this case, caching is not worth doing. To remedy this situation, we propose a relay clustering scheme such that caching files in Group \mathcal{N}_2 will lead to lower outage probability than transmitting them using the traditional transmission scheme for files in Group \mathcal{N}_3 . The clustering scheme is described as follows.

- 1) Suppose the outage probability of a single R_k is $p_k = \text{Prob}(\log_2(1 + P_r |g_k|^2) < R_F)$. Since we sort the relays according to their distances to D in an ascending order, it holds that $p_k \leq p_{\tilde{k}}, \forall k < \tilde{k}$. We can always find an integer $k' \in [1, K]$ such that $p_k \leq p^{G3}, \forall k \leq k'$ and $p_k > p^{G3}, \forall k > k'$, or in other words, the individual transmission of relays in the set $S \triangleq \{R_{k'+1}, \dots, R_K\}$ will lead to higher outage than the traditional transmission scheme without caching. Instead of individual transmissions, these relays can form clusters to improve their outage performance. As a by-product, the relay $R_k, k \in \{R_1, \dots, R'_k\}$ forms a cluster C_k of its own whose outage probably is less than or equal to p^{G3} .
- 2) The next cluster $C_{k'+1}$ is formed by adding one relay from the set S each time in the reverse order to the cluster ($R_{k'}$ is the first one to be added) until the cluster's outage $p_{C_{k'+1}}^{G2}$ is no greater than p^{G3} .
- 3) Continue the above step until each relay in S is allocated to a cluster. If the remaining relays cannot form a cluster that has lower probability no greater p^{G3} , they will be added to the last formed cluster.
- 4) After all clusters are formed, they are sorted in the ascending order according to the outage probability, i.e., $p_{U_l}^{G2} \leq p_{U_m}^{G2}, \forall l < m$. Then relays in each cluster will cache the same next less popular $(C - M)$ files, as described in Subsection II.A.

Fig. 1(a) shows an example of the proposed clustering and cache placement. The caching schemes in [10] and [11] are also illustrated in Fig. 1(b) and Fig. 1(c), respectively, for comparison. We observe that the multi-threshold caching in [11] allows to store different number of copies of files according to their popularity, so it is a finer partitioning of the

	C1	C2	C3	C4
R1	1	2	3	4
R2	1	8	9	10
R3	1	5	6	7
R4	1	5	6	7

(a)

	C1	C2	C3	C4
R1	1	2	3	4
R2	1	5	6	7
R3	1	8	9	10
R4	1	11	12	13

(b)

	C1	C2	C3	C4
R1	1	2	3	6
R2	1	2	3	7
R3	1	4	5	8
R4	1	4	5	9

(c)

Fig. 1. (a) An example of the proposed clustering and caching results for $K = 4, C = 4, M = 1$. $\{R_1\}$, $\{R_3, R_4\}$ and $\{R_2\}$ form three clusters, respectively. The cluster $\{R_3, R_4\}$ has lower outage probability than $\{R_2\}$. (b) The caching scheme in [10]. (c) The caching scheme in [11] with two thresholds $M_1 = 1$ and $M_2 = 2$, where M_1 most popular files are stored at all 4 relays and the next $2M_2$ most popular files have 2 copies each.

cache space than [10], but its complexity is $O(C^H)$ where H is the number of thresholds. Our proposed caching optimization involves a single threshold and the worst-case complexity is $O(C)$. In addition, it is proved to be convex as shown in Section III, so the complexity is much lower than [11].

III. OPTIMIZATION OF CACHE PLACEMENT

It can be seen that the parameter M uniquely determines the cache placement, transmission schemes and the ultimate system performance. Our proposed caching design is to identify the optimal value of M , to minimize the outage probability, which is expressed as

$$\begin{aligned} \bar{p} &= p^{G1} \sum_{n=1}^M f_n + p^{G3} \sum_{n=I(L,M)+1}^N f_n \\ &+ \sum_{l=1}^L p_{U_l}^{G2} \sum_{n=I(l,M)+M-C+1}^{I(l,M)} f_n \\ &= p^{G1} \sum_{n=1}^M f_n + p^{G3} \left(1 - \sum_{n=1}^{I(L,M)} f_n \right) \\ &+ \sum_{l=1}^L p_{U_l}^{G2} \left(\sum_{n=1}^{I(l,M)} f_n - \sum_{n=1}^{I(l,M)+M-C} f_n \right). \end{aligned} \quad (3)$$

For reference, traditional relaying without caching has a fixed outage probability p^{G3} .

Observe that \bar{p} is a complicated function of M because M is an integer variable and its value dynamically determines the numbers of files in different groups. To find the optimal M that minimizes the above probability, one can use an exhaustive search with a complexity of $O(N)$ but it gives no insight on the impact of system parameters. This motivates us to study its analytical property and derive the closed-form solution below.

Our first step to tackle the difficulty of optimizing M is to invoke the following result of approximating the sum of Zipf probabilities [9]:

$$\sum_{n=1}^m f_n \approx \frac{m^{1-\gamma} - 1}{N^{1-\gamma} - 1}. \quad (4)$$

Next to deal with the discrete variable M , we introduce a new continuous variable $\rho \triangleq \frac{M}{C} \in [0, 1]$. Once we have the optimal ρ^* , the optimal M can be found as $M^* = \lceil C\rho^* \rceil$. Then we can simplify the outage probability expression in (3) as

$$\begin{aligned} \bar{p} &\approx p^{G1} \frac{M^{1-\gamma} - 1}{N^{1-\gamma} - 1} + p^{G3} \left(1 - \frac{I(L, M)^{1-\gamma} - 1}{N^{1-\gamma} - 1} \right) \\ &+ \sum_{l=1}^L p_{U_l}^{G2} \left(\frac{(I(l, M))^{1-\gamma}}{N^{1-\gamma} - 1} - \frac{(I(l, M) + M - C)^{1-\gamma}}{N^{1-\gamma} - 1} \right) \\ &= \frac{C^{1-\gamma}}{N^{1-\gamma} - 1} y(\rho) + \text{constant (not a function of } \rho), \end{aligned} \quad (5)$$

where $y(\rho) \triangleq \left((p^{G1} - p_{U_1}^{G2})\rho^{1-\gamma} + \sum_{l=1}^{L-1} (p_{U_l}^{G2} - p_{U_{(l+1)}}^{G2})(l + (1-l)\rho)^{1-\gamma} + (p_{U_L}^{G2} - p^{G3})(L + (1-L)\rho)^{1-\gamma} \right) / (N^{1-\gamma} - 1)$. $y(\rho)$ is still a complex polynomial function of ρ . Below we study its convexity and derive the closed-form solution for a special case.

Theorem 1: $y(\rho)$ and \bar{p} in (5) are convex functions of ρ .

Proof: See the Appendix.

Because $y(\rho)$ is convex, we can use more efficient numerical algorithms, such as the bisection search, to find the optimal ρ .

Based on Theorem 1, we have the closed-form solution for a special case summarized in the following corollary.

Corollary 1: When the outage probabilities of relay clusters are the same, i.e., $p^{G2} \triangleq p_{C_l}^{G2}, l = 1, \dots, L$, a file request from Group \mathcal{N}_2 is served by a random relay cluster, and the approximately optimal cache placement solution is given by $M^* = \lceil C\rho^* \rceil$, where

$$\begin{aligned} \rho^* &= \min(\rho^\dagger, 1), \\ \text{where } \rho^\dagger &\triangleq \frac{L}{\left(\frac{p^{G2} - p^{G1}}{(p^{G3} - p^{G2})(L-1)} \right)^{-\frac{1}{\gamma}} + L - 1}. \end{aligned} \quad (6)$$

Proof: According to Theorem 1, the optimal ρ^* is the critical point ρ^\dagger that corresponds to zero first-order derivative, which leads to (6). In Section IV, $p_{C_l}^{G2}$ is determined by relay clustering, so the assumption of $p^{G2} = p_{C_l}^{G2}, \forall l$ is unnecessary.

Remarks: From (6), we can draw useful insights on the impact of system parameters. It is noted that ρ^* critically depends on the ratio of $\frac{p^{G2} - p^{G1}}{p^{G3} - p^{G2}}$, but not only on each individual outage probabilities like the solution derived in [10]. When $\frac{p^{G2} - p^{G1}}{p^{G3} - p^{G2}} < L - 1$, ρ^* is increasing with γ . This happens at the medium signal-to-noise ratio (SNR) regime, where both cooperation gain and caching diversity are needed. When $\frac{p^{G2} - p^{G1}}{p^{G3} - p^{G2}} > L - 1$, e.g., at the low SNR regime where $p^{G1} \ll \{p^{G2}, p^{G3}\}$, $\rho^* = 1$ is optimal which means to store the most popular content is the best strategy in this case.

IV. NUMERICAL RESULTS

In this section numerical results are presented to evaluate the performance of the proposed caching scheme, and illustrate the impact of key system parameters. We assume there are $K = 8$ relays without loss of generality. The channels are modelled

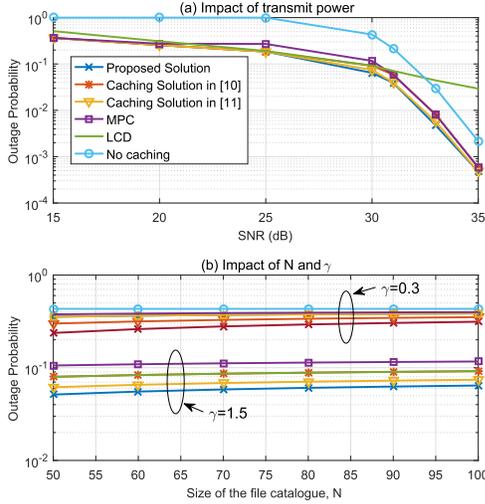


Fig. 2. Comparison of outage probabilities of different caching schemes and impact of system parameters.

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 $\alpha = 2.5$ is the path loss exponent. The $S - R_k$ distance is $l_k = (K - k + 1)/2$, and the $R_k - D$ 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 catalogue with $N = 100$ files and Zipf parameter $\gamma = 1.5$, unless otherwise specified. The required data rate is $R_F = 4$ bps/Hz. The cache capacity per relay is $C = 5$ files. The following benchmarks are used for comparison: 1) the traditional relaying without caching [13]; 2) the MPC based caching, 3) the LCD based caching, 4) the cache placement in [10] and 5) the multi-threshold-based cache placement in [11]. Note that [10] and [11] study different systems and design objectives than those in this letter; we only consider their cache placement schemes for comparison.

In Fig. 2(a), the outage probabilities against the SNR are provided. The proposed scheme offers more than 10 dB SNR gain over traditional relaying in the low to medium SNR regimes. The proposed solution brings significant gains over both the MPC and LCD because it can balance the cooperation gain and caching diversity gain. In the high SNR regime, the MPC scheme achieves near-optimal performance. The LCD is the worst caching scheme because when the content popularity is more concentrated, uniform caching will lose the signal cooperation gain. The proposed caching scheme outperforms that in [10] in the high SNR regime by up to 1 dB, and also achieves a similar performance as [11] with reduced complexity. The impact of the file catalogue size N and the Zipf parameter γ on the outage probabilities is shown in Fig. 2(b), when SNR is 30 dB. It is observed that the proposed placement achieves consistent gain over existing schemes for the range of the considered parameters.

V. CONCLUSIONS

This letter studied the optimization of caching schemes at distributed relays to improve the performance of cooperative communications. We optimized a hybrid caching scheme that can achieve the desired tradeoff between the cooperation gain

and caching diversity gain. Numerical results demonstrate substantial outage performance gains of caching over traditional relaying without caching, and the superiority of the proposed caching solution over existing cache placement schemes.

APPENDIX: PROOF OF THE CONVEXITY OF $y(\rho)$ IN (5)

The first and second order derivatives of $y(\rho)$, are respectively,

$$\begin{aligned} \frac{y'(\rho)}{c} &= (p^{G1} - p_{U_1}^{G2})\rho^{-\gamma} \\ &\quad + (p_{U_L}^{G2} - p^{G3})(1-L)(L + (1-L)\rho)^{-\gamma} \\ &\quad + \sum_{l=1}^{L-1} (p_{U_l}^{G2} - p_{U_{(l+1)}}^{G2})(1-l)(l + (1-l)\rho)^{-\gamma}, \text{ and} \\ \frac{y''(\rho)}{\gamma c} &= -(p^{G1} - p_{U_1}^{G2})\rho^{-\gamma-1} \\ &\quad - (p_{U_L}^{G2} - p^{G3})(1-L)^2(L + (1-L)\rho)^{-\gamma-1} \\ &\quad - \sum_{l=1}^{L-1} (p_{U_l}^{G2} - p_{U_{(l+1)}}^{G2})(1-l)^2(l + (1-l)\rho)^{-\gamma-1}, \end{aligned} \quad (7)$$

where $c \triangleq \frac{1-\gamma}{N^{1-\gamma}-1} > 0$. We already know that $p^{G1} \leq p_{U_1}^{G2}, p_{U_1}^{G2} \leq p_{U_{(l+1)}}^{G2}$ and $p_{U_L}^{G2} \leq p^{G3}$, so $y''(\rho) > 0$, therefore $y(\rho)$ is convex. This completes the proof.

REFERENCES

- [1] Z. Ding, I. Krikidis, B. Rong, J. S. Thompson, C. Wang, and S. Yang, "On combating the half-duplex constraint in modern cooperative networks: Protocols and techniques," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 20–27, Dec. 2012.
- [2] I. Krikidis, B. Rong, and A. Ephremides, "Network-level cooperation for a multiple-access channel via dynamic decode-and-forward," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7759–7770, Dec. 2011.
- [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, Sep. 2011.
- [4] Y. Deng, K. J. Kim, T. Q. Duong, M. Elkashlan, G. K. Karagiannidis, and A. Nallanathan, "Full-duplex spectrum sharing in cooperative single carrier systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 2, pp. 68–82, Mar. 2016.
- [5] N. Zlatanov and R. Schober, "Buffer-aided half-duplex relaying can outperform ideal full-duplex relaying," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 479–482, Mar. 2013.
- [6] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femto-caching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 142–149, Apr. 2013.
- [7] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.
- [8] J. Song, H. Song, and W. Choi, "Optimal caching placement of caching system with helpers," in *Proc. ICC*, London, U.K., Jun. 2015, pp. 1825–1830.
- [9] M. Taghizadeh, K. Micinski, C. Ofria, E. Torng, and S. Biswas, "Distributed cooperative caching in social wireless networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 6, pp. 1037–1053, Jun. 2013.
- [10] Z. Chen, J. Lee, T. Q. S. Quek, and M. Kountouris. (2016). "Cooperative caching and transmission design in cluster-centric small cell networks." [Online]. Available: <https://arxiv.org/abs/1601.00321>
- [11] W. C. Ao and K. Psounis, "Distributed caching and small cell cooperation for fast content delivery," in *Proc. MobiHoc*, Hangzhou, China, Jun. 2015, pp. 127–136.
- [12] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: Evidence and implications," in *Proc. IEEE INFOCOM*, New York, NY, USA, Mar. 1999, pp. 126–134.
- [13] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2499–2517, Jun. 2009.