2015

Improved Lower Bounds for Coded Caching

Hooshang Ghasemi
Iowa State University, ghasemi@iastate.edu

Aditya Ramamoorthy
Iowa State University, adityar@iastate.edu

Improved Lower Bounds for Coded Caching

Hooshang Ghasemi and Aditya Ramamoorthy
Dept. of Electrical & Computer Eng. Iowa State University, Ames, IA 50011
Emails: {ghasemi, adityar}@iastate.edu

***Abstract*—Content delivery networks often employ caching to reduce transmission rates from the central server to the end users. Recently, the technique of coded caching was introduced whereby coding in the caches and coded transmission signals from the central server are considered. Prior results in this area demonstrate that (a) carefully designing placement of content in the caches and (b) designing appropriate coded delivery signals allow for a system where the delivery rates can be significantly smaller than conventional schemes. However, matching upper and lower bounds on the transmission rates have not yet been obtained. In this work, we derive tighter lower bounds on coded caching rates than were known previously. We demonstrate that this problem can equivalently be posed as one of optimally labeling the leaves of a directed tree. Several examples that demonstrate the utility of our bounds are presented.**

I. Introduction

Content distribution over the Internet is an important problem and is the core business of several enterprises such as Youtube, Netflix, Hulu etc. One commonly used technique to facilitate content delivery is caching [1], whereby relatively popular server content is stored in local cache memory at the end users. When files are requested by the users, the cached content is first used to serve them. The remainder of the content is obtained from the server. This reduces the number of bits transmitted from the server on average. In conventional approaches to caching, coding in the content of the cache and/or coding in the transmission from the server are typically not considered.

The work of [2] introduced the problem of coded caching, where there is a server with N files and K users each with a cache of size M. The users are connected to the server by a shared link. In each time slot each user requests one of the N files. There are two distinct phases in coded caching. In the *placement phase*, the content of caches is populated; this phase should not depend on the actual user requests (which are assumed to be arbitrary). In the *delivery phase*, the server transmits a signal of rate R over the shared link that serves to satisfy the demands of each of the users. The work of [2] demonstrates that a carefully designed placement scheme and a corresponding delivery scheme achieves a rate that is significantly lower than conventional caching. The work of [2] also shows that their achievable rate is within a factor of 12 of the cutset lower bound for all values of N, K and M. There have been some subsequent contributions in this area. Decentralized coded caching where the placement phase is such that each user stores a random portion of each file was investigated in [3]. Schemes where the popularity of the files are taken into account appeared in [4]. A hierarchical scheme where there are multiple levels of caches was considered in [5].

In this work our main contribution is in developing improved lower bounds on the rate for the coded caching problem. The computation of this lower bound can be posed as a labeling problem on a directed tree. This paper is organized as follows. Section II contains the problem formulation, Section III develops the proposed lower bound. Section IV presents several examples of the proposed lower bound and compares it with prior results.

II. Problem Formulation

Let $[m] = \{1, \dots, m\}$. The coded caching problem can be formally described as follows. Let $\{W_n\}_{n=1}^N$ denote N independent random variables (representing the files) each uniformly distributed over $[2^F]$. The i-th user requests the file W_{d_i}, where $d_i \in [N]$. A (M, R) system consists of

- caching functions, $Z_i = \phi_i(W_1, \dots, W_N)$.
- Encoding functions $\varphi_{d_1,\dots,d_K}(W_1, \dots, W_N)$, so that the delivery phase signal $X_{d_1,\dots,d_K} = \varphi_{d_1,\dots,d_K}(W_1, \dots, W_N)$.
- Decoding functions for the k-th user $\mu_{d_1,\dots,d_K;k}(X_{d_1,\dots,d_K}, Z_k)$, $k = 1, \dots, K$ so that decoded file $\hat{W}_{d_1,\dots,d_K;k} = \mu_{d_1,\dots,d_K;k}(X_{d_1,\dots,d_K}, Z_k)$.

The probability of error is defined as

$$\max_{(d_1,\dots,d_K)\in[N]^K} \max_{k\in[K]} P(\hat{W}_{d_1,\dots,d_K;k} \neq W_{d_k}).$$

Definition 1: The pair (M, R) is said to be achievable if for $\epsilon > 0$, there exists a file size F large enough so that there exists a (M, R) caching scheme with probability of error at most ϵ. We define

$$R^\star(M) = \inf\{R : (M, R) \text{ is achievable}\}.$$

In this work we are interested in tight lower bounds on $R^\star(M)$.

A. Preliminaries

Definition 2: Directed in-tree. A directed graph $\mathcal{T} = (V, A)$, is called a directed in-tree if there is one designated node called the root such that from any other vertex $v \in V$ there is exactly one directed path from v to the root.

The nodes in a directed in-tree that do not have any incoming edges are referred to as the leaves. The remaining nodes, excluding the leaves and the root are called internal nodes. Each node in a directed in-tree has at most one outgoing edge. We have the following definitions for a node $v \in V$.

$$out(v) = \{u \in V : (v, u) \in A\},$$
$$in(v) = \{u \in V : (u, v) \in A\}.$$

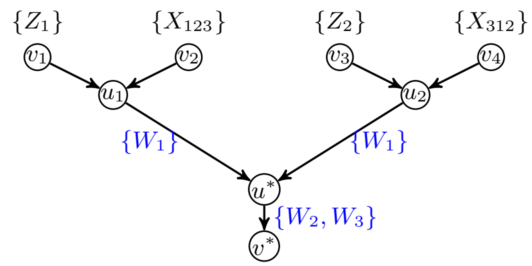


Fig. 1: Problem instance for Example 1. For clarity of presentation, only the $W_{new}(u)$ label has been shown on the edges.

In this work, we exclusively work with trees which are such that the in-degree of the root equals 1. There is a natural topological order in $\mathcal{T}$ whereby for nodes $u \in \mathcal{T}$ and $v \in \mathcal{T}$, we say that $u \succ v$ if there exists a sequence of edges that can be traversed to reach from u to v. This sequence of edges is denoted $path(u, v)$.

Let $D = \cup_{d_1 \in [N], \ldots, d_K \in [N]} \{X_{d_1,\ldots,d_K}\}$. Suppose that we are given a directed in-tree denoted $\mathcal{T}$, with ℓ leaves denoted $v_1, \ldots, v_\ell$. Furthermore, assume that each node $v \in \mathcal{T}$ is assigned a label, denoted $label(v_i)$, which is a subset of $\{W_1, \ldots, W_N\} \cup \{Z_1, \ldots, Z_K\} \cup D$. Moreover, we also specify $\mathbb{W}(v) \subseteq \{W_1, \ldots, W_N\}$, $\mathbb{Z}(v) \subseteq \{Z_1, \ldots, Z_K\}$ and $\mathbb{D}(v) \subseteq D$ so that $label(v) = \mathbb{W}(v) \cup \mathbb{Z}(v) \cup \mathbb{D}(v)$. In our formulation, the leaf nodes $v_i, i = 1, \ldots, \ell$ are such that $\mathbb{W}(v_i) = \emptyset$.

Definition 3: We say that a singleton source subset $\{W_i\}$ is recoverable from the pair $Z_j, X_{d_1,\ldots,d_K}$ if $d_j = i$. Similarly, for a given set of caches $Z' \subset \{Z_1, \ldots, Z_K\}$ and delivery phase signals $D' \subseteq D$, we define a set $Rec(Z', D') \subseteq \{W_1, \ldots, W_N\}$ to be the subset of the sources that can be recovered from pairs of the form (Z_i, X_J) where $Z_i \in Z'$ and J is a multiset of cardinality K with entries from $[N]$ such that $X_J \in D'$.

For nodes $u, v \in \mathcal{T}$, we let $\Delta(u, v) = Rec(\mathbb{Z}(u), \mathbb{D}(v))$. For a given node $u \in \mathcal{T}$, we define

$$W_{new}(u) = \Delta(u, u) \setminus \mathbb{W}(u), \tag{1}$$

i.e., $W_{new}(u)$ is the subset of sources that can be recovered from $(\mathbb{Z}(u), \mathbb{D}(u))$ that are distinct from $\mathbb{W}(u)$. A word about notation. We let the entropy of a set of random variables equal the joint entropy of all the random variables in the set. We also let $[x]^+ = \max(x, 0)$.

III. Lower Bound on $R^\star(M)$

Given a directed tree $\mathcal{T}$ and appropriate labels on its leaves $v_1, \ldots, v_\ell$, where we assume that $\mathbb{W}(v_i) = \emptyset$, for $i = 1, \ldots, \ell$, we claim that Algorithm 1 generates an inequality of the form $\alpha R^\star + \beta M \geq L(\alpha, \beta)$. We demonstrate this by means of the following example and defer the proof of the general case to the Appendix.

Example 1: Consider a system with $N = K = 3$, and the directed in-tree $\mathcal{T}$ with labeling: $label(v_1) = Z_1$, $label(v_2) = X_{123}$, $label(v_3) = Z_2$ and $label(v_4) = X_{312}$ (see Fig. 1). It can be observed that an application of Algorithm 1 gives us

---

Algorithm 1 Labeling Algorithm

Input: $\mathcal{T} = (V, A)$ with leaves $v_1, \ldots, v_\ell$ and $\{label(v_i)\}_{i=1}^{\ell}$, such that $\mathbb{W}(v_i) = \emptyset, i = 1, \ldots, \ell$.
Initialization:
for $i \leftarrow 1, \ldots \ell$ **do**
  $W_{new}(v_i) = \Delta(v_i, v_i)$.
  $x_{(v_i, out(v_i))} = W_{new}(v_i)$.
  $y_{(v_i, out(v_i))} = |W_{new}(v_i)|$.
end for
while there exists an unlabeled edge **do**
  Pick an unlabeled node $u \in V$ s.t. all edges in $in-edge(u)$ are labeled.
  $\mathbb{W}(u) = \cup_{v \in in(u)} \mathbb{W}(v) \cup W_{new}(v)$.
  $\mathbb{Z}(u) = \cup_{v \in in(u)} \mathbb{Z}(v)$.
  $\mathbb{D}(u) = \cup_{v \in in(u)} \mathbb{D}(v)$.
  Entropy-label(u): $H(\mathbb{Z}(u) \cup \mathbb{D}(u) | \mathbb{W}(u))$.
  $W_{new}(u) = \Delta(u, u) \setminus \mathbb{W}(u)$.
  $x_{(u, out(u))} = W_{new}(u)$.
  $y_{(u, out(u))} = |W_{new}(u)|$.
end while
Output: $L = \sum_{e \in A} y_e$.

---

the inequality $2R^\star + 2M \geq 4$. This can be justified as follows.

$$\begin{aligned}
&2R^\star F + 2MF \geq H(Z_1, X_{123}) + H(Z_2, X_{312}) \\
&= I(W_1; Z_1, X_{123}) + H(Z_1, X_{123}|W_1), \\
&+ I(W_1; Z_2, X_{312}) + H(Z_2, X_{312}|W_1) \\
&\overset{(a)}{\geq} F(1-\epsilon) + F(1-\epsilon) + H(Z_1, Z_2, X_{123}, X_{312}|W_1) \\
&= 2F(1-\epsilon) + I(W_2, W_3; Z_1, Z_2, X_{123}, X_{312}|W_1), \\
&+ H(Z_1, Z_2, X_{123}, X_{312}|W_1, W_2, W_3) \\
&\overset{(b)}{\geq} 2F(1-\epsilon) + 2F(1-\epsilon) = 4F(1-\epsilon),
\end{aligned}$$

where inequality (a) holds since conditioning reduces entropy and e.g., $I(W_1; Z_1, X_{123}) \geq F - \epsilon F$ (by Fano's inequality). The other inequality can be shown to hold in a similar manner. Similarly, inequality (b) holds by the independence of the W_i's and Fano's inequality. This holds for arbitrary $\epsilon > 0$ and F large enough. Dividing throughout by F we have the required result.

It can be observed that at each internal node, certain cache signals and delivery phase signals *meet*, e.g. Z_1 and X_{123} meet at node u_1 in Fig. 1. The outgoing edge of an internal node is labeled by the *new* files that are recovered at the node, e.g., at u_1 the signals Z_1 and X_{123} recover the file W_1. We call a file new if it has not been recovered upstream of a given node. It can be seen that this labeling is in one to one correspondence with inequality (a) in Example 1 above. In a similar manner at u^* one can recover all the files $W_1, \ldots, W_3$; however only the set $\{W_2, W_3\}$ is labeled on edge (u^*, v^*) as W_1 was recovered upstream. This intuition is formalized in the Appendix (Lemma 2) where it is shown that a valid lower bound is always obtained when applying Algorithm 1.

Definition 4: Problem Instance. Consider a given tree $\mathcal{T}$ with leaves $v_i, i = 1, \ldots, \ell$ that are labeled as discussed above. Let $\alpha = \sum_{i=1}^{\ell} |\mathbb{D}(v_i)|$ and $\beta = \sum_{i=1}^{\ell} |\mathbb{Z}(v_i)|$. Suppose that the lower bound computed by Algorithm 1 equals L. We define the associated problem instance as $P(\mathcal{T}, \alpha, \beta, L, N, K)$. We also define $\hat{\alpha} = |\cup_{i=1}^{\ell} \mathbb{D}(v_i)|$ and $\hat{\beta} = |\cup_{i=1}^{\ell} \mathbb{Z}(v_i)|$. A problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$ is optimal if all instances of the form $P'(\mathcal{T}', \alpha, \beta, L', N, K)$ are such that $L' \leq L$.

It is not too hard to see that it suffices to consider directed trees whose internal nodes have an in-degree at least two. In particular, if u has in-degree equal to 1, it is evident that $W_{new}(u) = \emptyset$ and thus, $y_{(u,out(u))} = 0$. In addition, we show in the Appendix (Claim 5) that w.l.o.g. it suffices to consider trees where internal nodes have in-degree at most two. Therefore, we will assume that all internal nodes have degree equal to two.

We can also conclude that each leaf v in an instance P is such that either $|\mathbb{Z}(v)| = 1$ or $|\mathbb{D}(v)| = 1$ but not both. If $|\mathbb{Z}(v)| = 1$, we call v a cache node; if $|\mathbb{D}(v)| = 1$ we call it a delivery phase node. In the subsequent discussion we will assume the delivery phase nodes are labeled in an arbitrary order $v_1, \dots, v_\alpha$ and the cache nodes from $v_{\alpha+1}, \dots, v_{\alpha+\beta}$, where we note that $\alpha + \beta = \ell$. Moreover, we let $\mathcal{D} = \{v_1, \dots, v_\alpha\}$ and $\mathcal{C} = \{v_{\alpha+1}, \dots, v_{\alpha+\beta}\}$.

We now explore some characteristics of optimal problem instances. In the tree $\mathcal{T}$ corresponding to problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$, consider an internal node u and the edge $e = (u, v)$. The incoming edges into u, denoted (u_l, u) and (u_r, u) are the last edges of the disjoint left and right subtrees denoted $\mathcal{T}_{u(l)}$ and $\mathcal{T}_{u(r)}$ respectively. Each of these subtrees defines a problem instance $P_l = P(\mathcal{T}_{u(l)}, \alpha_l, \beta_l, L_l, N, K)$ and $P_r = P(\mathcal{T}_{u(r)}, \alpha_r, \beta_r, L_r, N, K)$. We define $\mathcal{D}_{u(r)} = \{v \in \mathcal{D} : v \in \mathcal{T}_{u(r)}\}$ and $\mathcal{C}_{u(r)} = \{v \in \mathcal{C} : v \in \mathcal{T}_{u(r)}\}$ with similar definitions for $\mathcal{D}_{u(l)}$ and $\mathcal{C}_{u(l)}$. We also let $\mathcal{D}_u = \mathcal{D}_{u(l)} \cup \mathcal{D}_{u(r)}$ and $\mathcal{C}_u = \mathcal{C}_{u(l)} \cup \mathcal{C}_{u(r)}$.

Let $\Gamma_l = \cup_{v \in \mathcal{T}_{u(l)}} W_{new}(v)$ and $\Gamma_r = \cup_{v \in \mathcal{T}_{u(r)}} W_{new}(v)$, i.e., Γ_l and Γ_r are the subsets of $\{W_1, \dots, W_N\}$ that are used up in the problem instances P_l and P_r respectively. We shall often need to reason about the files recovered at the node u. Accordingly, we have the following definitions.

$$\Delta_{rl}(u) = Rec(\mathbb{Z}(u_r) \setminus \mathbb{Z}(u_l), \mathbb{D}(u_l)),$$
$$\Delta_{lr}(u) = Rec(\mathbb{Z}(u_l) \setminus \mathbb{Z}(u_r), \mathbb{D}(u_r)).$$

It can be observed that we have

$$\begin{aligned} W_{new}(u) &= \Delta(u,u) \setminus \mathbb{W}(u) \\ &= \Delta_{rl}(u) \cup \Delta_{lr}(u) \setminus \mathbb{W}(u). \end{aligned} \tag{2}$$

The second equality holds since $Rec(\mathbb{Z}(u_l), \mathbb{D}(u_l)) \cup Rec(\mathbb{Z}(u_r), \mathbb{D}(u_r)) \subseteq \mathbb{W}(u)$. Note that based on Algorithm 1, we can conclude that

$$\begin{aligned} \mathbb{W}(u) &= \cup_{v \in \{u_r, u_l\}} \mathbb{W}(v) \cup W_{new}(v) \\ &= \cup_{v \succ u} W_{new}(v). \end{aligned} \tag{3}$$

Often, we will need to refer to the singleton file subset that is recovered from $v \in \mathcal{D}$ and $v' \in \mathcal{C}$ where v and v' meet at node u. In this case, we denote

$$\Delta_{(v,v')}(u) = Rec(\mathbb{Z}(v'), \mathbb{D}(v)).$$

For any pair of leaf nodes v_i and v_j where $i, j \in \{1, \dots, \alpha+\beta\}$ we say that v_i and v_j meet at node u if there exist $path(v_i, u)$ and $path(v_j, u)$ in $\mathcal{T}$ such that $path(v_i, u) \cap path(v_j, u) = \emptyset$. In our subsequent discussion, we will often modify a given problem instance P to arrive at a different problem instance P'. In this situation we will use the superscript P or P' to refer to the appropriate instance, e.g., $W_{new}^P(u)$ will refer to $W_{new}(u)$ in the instance P.

For a problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$, it may be possible that $\hat{\beta} < \beta$. However, given such an instance, we can convert it into another instance where $\hat{\beta} = \beta$ without reducing the value of L. In fact the following stronger statement holds (see Appendix A for a proof).

Claim 1: For a problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$ consider an internal node u^* with associated problem instances $P_l = P(\mathcal{T}_{u^*(l)}, \alpha_l, \beta_l, L_l, N, K)$ and $P_r = P(\mathcal{T}_{u^*(r)}, \alpha_r, \beta_r, L_r, N, K)$ such that at least one of conditions (i) – (iii) below is true.

(i) $\hat{\beta}_l < \min(\beta_l, K)$.

(ii) $\hat{\beta}_r < \min(\beta_r, K)$.

(iii) $\hat{\beta} < \min(\beta, K)$.

Then, there exists another problem instance $P'(\mathcal{T}', \alpha, \beta, L', N, K)$ where $L' \geq L$ and none of the conditions (i) – (iii) hold.

Henceforth we will assume w.l.o.g. that $\hat{\beta} = \beta$ and that Claim 1 holds. Our next lemma shows a structural property of problem instances. Namely for a instance where $L < \alpha \min(\beta, K)$, increasing the number of files allows us to increase the value of L. This lemma is a key ingredient in our proof of the main theorem.

Lemma 1: Let $P = P(\mathcal{T}, \alpha, \beta, L, K, N)$ be an instance where $L < \alpha \min(\beta, K)$. Then, we can construct a new instance $P' = P(\mathcal{T}', \alpha, \beta, L', K, N+1)$, where $L' = L + 1$.

Informally, another property of optimal problem instances is that the same file is recovered as many times as possible at the same level of the tree. For instance, in Fig. 1, W_1 is recovered in both $\mathcal{T}_{u^*(l)}$ and $\mathcal{T}_{u^*(r)}$. In fact, intuitively it is clear that the same set of files can be reused in any subtrees of an internal node. Our next claim formalizes this intuition.

Claim 2: Consider an instance $P = P(\mathcal{T}, \alpha, \beta, L, K, N)$. At any node $u \in \mathcal{T}$, suppose w.l.o.g. that $|\Gamma_l| \geq |\Gamma_r|$. If there exists a node u such that $\Gamma_r \nsubseteq \Gamma_l$, then there exists another instance $P'(\mathcal{T}', \alpha, \beta, L', N', K)$ such that $N' < N$ and $L' \geq L$.

Definition 5: Saturation number. Consider an instance $P^*(\mathcal{T}^*, \alpha, \beta, L^*, N^*, K)$, where $L^* = \alpha \min(\beta, K)$, such that for all problem instances of the form $P(\mathcal{T}, \alpha, \beta, L^*, N, K)$, we have $N^* \leq N$. We call N^* the saturation number of instances with parameters (α, β, K) and denote it by $N_{sat}(\alpha, \beta, K)$.

In essence, for given α, β and K, saturated instances are most efficient in using the number of available files. It is easy to see that $N_{sat}(\alpha, \beta, K) \leq \alpha \min(\beta, K)$ since one can construct an instance with lower bound $\alpha \min(\beta, K)$ when $\alpha \min(\beta, K) \leq N$ (see remark 1 in the proof of Lemma 1).

Definition 6: Atomic problem instance. For a given optimal problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$ it is possible that there exist other optimal problem instances $P_i(\alpha_i, \beta_i, L_i, N, K), i = 1, \dots, m$ with $m \geq 2$ such that $\sum_{i=1}^m \alpha_i = \alpha, \sum_{i=1}^m \beta_i = \beta$

and $\sum_{i=1}^{m} L_i = L$, i.e., the value of L follows from appropriately combining smaller problems. In this case we call the instance P as non-atomic. Conversely, if such smaller problem instances do not exist, we call P an atomic problem instance.

Claim 3: Let $P(\mathcal{T}, \alpha, \beta, L, K, N)$ be an instance where $\beta \le K$ and $L = \alpha\beta$. Then, $\hat{\alpha} = \alpha$.

Let $\rho(u) = \hat{\alpha}_l[\min(\hat{\beta}_r, K - \hat{\beta}_l)]^+ + \hat{\alpha}_r[\min(\hat{\beta}_l, K - \hat{\beta}_r)]^+$.

Claim 4: In instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$, consider an internal node u. We have

$$|W_{new}(u)| \le \min\left(\rho(u), N - |\Gamma_l \cup \Gamma_r|\right).$$

Proof: From eq. (2) it follows that

$$|W_{new}(u)| \le |\Delta_{rl}(u) \setminus \mathbb{W}(u)| + |\Delta_{lr}(u) \setminus \mathbb{W}(u)|.$$

Next, we observe that

$$\begin{aligned} &|\Delta_{rl}(u) \setminus \mathbb{W}(u)| = |Rec(\mathbb{Z}(u_r) \setminus \mathbb{Z}(u_l), \mathbb{D}(u_l)) \setminus \mathbb{W}(u)| \\ &\le |\mathbb{Z}(u_r) \setminus \mathbb{Z}(u_l)| \times |\mathbb{D}(u_l)| \\ &\overset{(a)}{\le} \hat{\alpha}_l \times [\min(\hat{\beta}_r, K - \hat{\beta}_l)]^+, \end{aligned}$$

where inequality (a) holds, since $|\mathbb{D}(u_l)| = \hat{\alpha}_l$ and $|\mathbb{Z}(u_r) \setminus \mathbb{Z}(u_l)| \le \min(\hat{\beta}_r, K - \hat{\beta}_l)]^+$. We can bound $|\Delta_{lr}(u) \setminus \mathbb{W}(u)|$ in a similar manner to obtain the first inequality. To see the second inequality we note that instances P_l and P_r recover a total of $|\Gamma_l \cup \Gamma_r|$ sources. As the total number of sources is N, $|W_{new}(u)| \le N - |\Gamma_l \cup \Gamma_r|$. ■

The following theorem and its corollary are the main results of our paper and can be used to identify optimal problem instances.

Theorem 1: Suppose that there exists an optimal and atomic problem instance $P_o(\mathcal{T} = (V, A), \alpha, \beta, L_o, N, K)$. Then, there exists optimal and atomic problem instance $P^*(\mathcal{T}^* = (V^*, A^*), \alpha, \beta, L^*, N, K)$ where $L^* = L_o$ with the following properties. Let us denote the last edge in P^* with (u^*, v^*). Let $P_l^* = P(\mathcal{T}^*_{u^*(l)}, \alpha_l, \beta_l, L_l^*, N_l, K)$ and $P_r^* = P(\mathcal{T}^*_{u^*(r)}, \alpha_r, \beta_r, L_r^*, N_r, K)$. Then, we have

$$\begin{aligned} L_l^* &= \alpha_l \min(\beta_l, K), \\ L_r^* &= \alpha_r \min(\beta_r, K), \text{ and} \\ L^* &= \min\left(\alpha \min(\beta, K), L_l^* + L_r^* + N - N_0\right), \end{aligned} \tag{4}$$

where $N_0 = \max(N_{sat}(\alpha_l, \beta_l, K), N_{sat}(\alpha_r, \beta_r, K))$[1]. Furthermore, at least one of β_l or β_r is strictly smaller than K.

Proof: Note that we assume that the problem instance P_o is atomic. This implies that $W_{new}^{P_o}(u^*) \neq \emptyset$. Using Claim 1 we can assert that $\hat{\beta}_l = \beta_l$ and $\hat{\beta}_r = \beta_r$.

Suppose that $L_l^* < \alpha_l \min(\beta_l, K)$. We apply the result of Lemma 1, by noting that $N_l < N$, and conclude that there exists another instance $P_l^{**} = P(\mathcal{T}^{**}_{u^*(l)}, \alpha_l, \beta_l, L_l^*+1, N_l+1, K)$ that can replace P_l^*, where the new file is denoted W^*. We also note that in P_o, $W^* \in W_{new}^{P_o}(u^*)$. Let us denote the new instance P_o'. We emphasize that the nature of the modification in Lemma 1 is such that $\Delta_{rl}^{P_o'}(u^*) = \Delta_{rl}^{P_o}(u^*)$ and $\Delta_{lr}^{P_o'}(u^*) = \Delta_{lr}^{P_o}(u^*)$. Moreover, we note that $\mathbb{W}^{P_o'}(u^*) = \mathbb{W}^{P_o}(u^*) \cup \{W^*\}$. We have,

$$\begin{aligned} &W_{new}^{P_o'}(u^*) = \Delta_{rl}^{P_o'}(u^*) \cup \Delta_{lr}^{P_o'}(u^*) \setminus \mathbb{W}^{P_o'}(u^*) \\ &= \Delta_{rl}^{P_o}(u^*) \cup \Delta_{lr}^{P_o}(u^*) \setminus \mathbb{W}^{P_o}(u^*) \cup \{W^*\} \\ &= W_{new}^{P_o} \setminus \{W^*\} \text{ (since } W^* \in \Delta_{rl}^{P_o}(u^*) \cup \Delta_{rl}^{P_o}(u^*)). \end{aligned}$$

Based on this argument, we can immediately conclude that we cannot have $L_l^* < \alpha_l \min(\beta_l, K)$ and $L_r^* < \alpha_r \min(\beta_r, K)$ as the file W^* can be used to simultaneously modify the instance P_r^*. Upon this modification, we can conclude that L^* can be increased by one, which contradicts the optimality of the instance P_o. Thus we assume that $L_r^* = \alpha_r \min(\beta_r, K)$. We can repeatedly apply the operation of moving files from $W_{new}^{P_o}(u^*)$ to P_l^* until we have $L_l^* = \alpha_l \min(\beta_l, K)$. It has to be the case that $|W_{new}^{P_o}(u^*)| > \alpha_l \min(\beta_l, K) - N_l$ so that we can repeatedly apply the operation of moving the files, for if this were not true, the instance P_o is not atomic.

We will denote the instance that we arrive at after completing these modification by P^*. We can also observe at this point that if we have $\beta_l \ge K$ and $\beta_r \ge K$, then $W_{new}^{P^*}(u^*) = \emptyset$ (by Claim 4) which implies that the original instance P_o is not atomic. Thus, either β_l or β_r or both have to be strictly smaller than K. In the discussion below we assume w.l.o.g. that $\beta_r < K$. It is easy to see that

$$L^* = L_l^* + L_r^* + |W_{new}^{P^*}(u^*)|.$$

By Claim 4, we have that

$$|W_{new}^{P^*}(u^*)| \le \min\left(\rho(u^*), N - \max(|\Gamma_l^*|, |\Gamma_r^*|)\right),$$

For an optimal instance, we claim that the above inequality is met with equality. This is because by Claim 2, we have either $\Gamma_l^* \subseteq \Gamma_r^*$ or $\Gamma_r^* \subseteq \Gamma_l^*$, so that there are $N - \max(|\Gamma_l^*|, |\Gamma_r^*|)$ new files that are available to be recovered at u^*.

Moreover, as $\beta_r < K$ and $L_r^* = \alpha_r \min(\beta_r, K)$, we can conclude that $\hat{\alpha}_r = \alpha_r$ by Claim 3. Next, we observe that if $\beta_l < K$, we can again use Claim 3 to conclude $\hat{\alpha}_l = \alpha_l$. On the other hand if $\beta_l > K$, then $[\min(\beta_r, K - \beta_l)]^+ = 0$. In both cases, we can conclude that

$$|W_{new}^{P^*}(u^*)| = \min\left(\tilde{\rho}(u^*), N - \max(|\Gamma_l^*|, |\Gamma_r^*|)\right),$$

where $\tilde{\rho}(u^*) = \alpha_l \times [\min(\beta_r, K - \beta_l)]^+ + \alpha_r \times [\min(\beta_l, K - \beta_r)]^+$. It is easy to verify that,

$$\alpha \min(\beta, K) = \alpha_l \min(\beta_l, K) + \alpha_r \min(\beta_r, K) + \tilde{\rho}(u^*).$$

It follows that

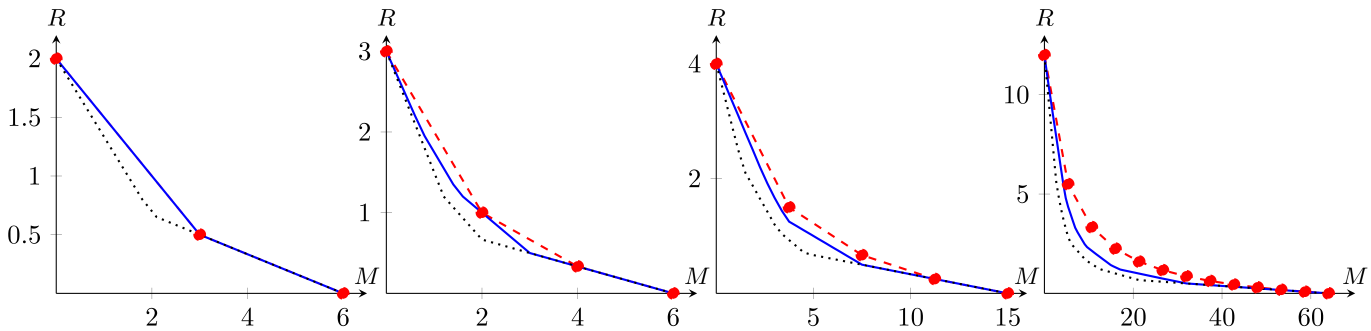
$$L^* = \min\left(\alpha \min(\beta, K), L_l^* + L_r^* + N - \max(|\Gamma_l^*|, |\Gamma_r^*|)\right).$$

Note that if $L^* < \alpha \min(\beta, K)$ we have

$$\begin{aligned} |W_{new}^{P^*}(u^*)| &= N - \max(|\Gamma_l^*|, |\Gamma_r^*|) \\ &\le N - \max(N_{sat}(\alpha_l, \beta_l, K), N_{sat}(\alpha_r, \beta_r, K)). \end{aligned} \tag{5}$$

We claim that P^* to be optimal, P_l^* and P_r^* have to be such that $N_l = N_{sat}(\alpha_l, \beta_l, K)$ and $N_r = N_{sat}(\alpha_r, \beta_r, K)$. To see, by the definition of saturation number problem instances, $P_l'(\mathcal{T}_l', \alpha_l, \beta_l, L_l', N_l', K)$ and $P_r'(\mathcal{T}_r', \alpha_r, \beta_r, L_r', N_r', K)$ exist such that $L_l' = L_l^*$, $L_r' = L_r^*$, $N_l' = N_{sat}(\alpha_l, \beta_l, K)$ and $N_r' = N_{sat}(\alpha_r, \beta_r, K)$. W.l.o.g let assume $N_l' \ge N_r'$. By the Claims

[1]As the instance is atomic, we have $N \ge N_0$.



(a) Case I: $N = 6$, $K = 2$ (b) Case II: $N = 6$, $K = 3$ (c) Case III: $N = 15$, $K = 4$ (d) Case IV: $N = 64$, $K = 12$

Fig. 2: Comparison of lower bounds. Blue curve: proposed lower bound, Red curve: achievable rate, Black curve: cutset lower bound.

1 and 2 problem instances P'_l and P'_r can be modified in such a way that $\hat{\beta}'_l = \min(\beta_l, K)$, $\hat{\beta}'_r = \min(\beta_r, K)$ and $\Gamma'_l \subseteq \Gamma'_r$. Also, $\cup_{v \in \mathcal{C}'_l} \mathbb{Z}(v)$ and $\cup_{v \in \mathcal{C}'_r} \mathbb{Z}(v)$ have minimum intersection. Now, consider problem instance $P'(\mathcal{T}', \alpha, \beta, L', N, K)$ with last edge (u', v') so that P'_l and P'_r are instances of u'_l and u'_r respectively. By setting $\Delta_{(v,v')}(u') \in \{W_1, \dots, W_N\} \setminus \Gamma'_l$, we can modify P' such that $|W_{new}(u')| = N - \max(N'_l, N'_r)$. Thus, we have $L = L^*_l + L^*_r + N - \max(N'_l, N'_r)$ and since $L^* \geq L$ therefore equality in eq. 5 holds. ■

Corollary 1: Suppose that there exists an optimal and atomic problem instance $P_o(\mathcal{T} = (V, A), \alpha, \beta, L_o, N, K)$. Consider problem instances $P'_l(\alpha'_l, \beta'_l, L'_l, N, K)$ and $P'_r(\alpha'_r, \beta'_r, L'_r, N, K)$ such that $\alpha'_l + \alpha'_r = \alpha$ and $\beta'_l + \beta'_r = \beta$ such that $N \geq N'_0 = \max(N_{sat}(\alpha'_l, \beta'_l, K), N_{sat}(\alpha'_r, \beta'_r, K))$. Then we have

$$L_o \geq \min\left(\alpha \min(\beta, K), L'_l + L'_r + N - N'_0\right).$$

Proof: The result follows by applying the arguments in the proof of Theorem 1, to the problem instance where P^*_l and P^*_r are replaced by P'_l and P'_r respectively. ■

IV. Discussion

Our first observation is that the cutset bound in [2] is a special case of the bound in eq. (4). In particular, suppose that $\alpha = \lfloor N/s \rfloor$, $\beta = s$ for $s = 0, 1, \dots, \min(N, K)$. Note that $\alpha\beta \leq N$. Thus, it is easy to construct a problem instance where $L = \alpha\beta$ (see remark 1 in the Proof of 1). This also follows from observing that $N_{sat}(\alpha, \beta, K) \leq \alpha\beta$.

Suppose that for a coded caching system with N files and K users, we first apply the cutset bound with certain α_1 and β_1 such that $\alpha_1\beta_1 < N$. This in turn implies that $N_{sat}(\alpha_1, \beta_1, K) < N$. Using Corollary 1 we can instead attempt to lower bound $2\alpha_1 R^\star + 2\beta_1 M$ and obtain the following inequality.

$$\begin{aligned} &2\alpha_1 R^\star + 2\beta_1 M \geq 2\alpha_1\beta_1 + N - N_{sat}(\alpha_1, \beta_1, K) \\ &\implies \alpha_1 R^\star + \beta_1 M \geq \alpha_1\beta_1 + (N - N_{sat}(\alpha_1, \beta_1, K))/2, \end{aligned}$$

which is strictly better than the cutset bound.

Example 2: Consider a system containing a server with four files and three users, $N = 4$ and $K = 3$. The cutset bounds corresponding to the given system are $4R^\star + M \geq 4$, $2R^\star + 2M \geq 4$ and $R^\star + 3M \geq 3$. Consider the second bound, $2R^\star + 2M \geq 4$ and instead attempt to obtain a lower bound on $4R^\star + 4M$.

In this case by exhaustive enumeration, it can be verified that $N_{sat}(2, 2, 3) = 3 < N$. Using Corollary 1, this results in the lower bound $L^* \geq \min(4 \times 3, 2 \times 4 + 4 - N_{sat}(2, 2, 3)) = 9$. Thus we can conclude $R^\star + M \geq 2.25$ which is better than the cutset bound $R^\star + M \geq 2$.

Theorem 1 can be leveraged effectively if it can also yield the optimal values of α_l, β_l and α_r, β_r. However, currently we do not have an algorithm for picking them in an optimal manner. Moreover, we also do not have an algorithm for finding $N_{sat}(\alpha, \beta, K)$. Thus, we have to use Corollary 1 with an appropriate upper bound on $N_{sat}(\alpha, \beta, K)$ in general.

Our proposed algorithm for upper bounding $N_{sat}(\alpha, \beta, K)$ is discussed in the Appendix (Algorithm 3). Setting $\alpha_l = \lceil \alpha/2 \rceil$, $\beta_l = \lfloor \beta/2 \rfloor$ in Theorem 1 and applying this approach to upper bound saturation number, we can obtain the results plotted in Fig. 2.

Example 3: Consider a system with $N = 64$, $K = 12$ and cache size $M = 16/3$. In this case using the cutset bound provides a lower bound $R^\star(M) \geq 77/27 = 2.852$. Now, using the approach of Theorem 1 for $\alpha = 12$, $\beta = 8$, $(\alpha_l, \beta_l) = (\alpha_r, \beta_r) = (6, 4)$ yields $12R^\star + 8M \geq \min(12 \times 8, 24 + 24 + 64 - N_{sat}(6, 4, 12))$. Using Algorithm 3 to upper bound N_{sat} we have $N_{sat}(6, 4, 12) \leq 17$ therefore $R^\star(M) \geq 157/36 = 4.361$. This is significantly closer to the achievable rate of 5.5 (from [2]).

References

[1] D. Wessels, *Web Caching*. O' Reilly, 2001.

[2] M. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. on Info. Th.*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[3] ——, "Decentralized coded caching attains order-optimal memory-rate tradeoff," *Networking, IEEE/ACM Transactions on*, 2014 (to appear).

[4] U. Niesen and M. Maddah-Ali, "Coded caching with nonuniform demands," in *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*, April 2014, pp. 221–226.

[5] N. Karamchandani, U. Niesen, M. Maddah-Ali, and S. Diggavi, "Hierarchical coded caching," in *Information Theory (ISIT), 2014 IEEE International Symposium on*, June 2014, pp. 2142–2146.

APPENDIX

Lemma 2: Algorithm 1 always provides a valid lower bound on $\alpha R^\star + \beta M$ where $\alpha = \sum_{i=1}^{\ell} |\mathbb{D}(v_i)|$ and $\beta = \sum_{i=1}^{\ell} |\mathbb{Z}(v_i)|$.

Proof: Consider any internal node $v \in \mathcal{T}$. We have

$$\begin{aligned} &\sum_{u \in in(v)} H(\mathbb{Z}(u) \cup \mathbb{D}(u) | \mathbb{W}(u) \cup W_{new}(u)), \\ &\overset{(a)}{\geq} \sum_{u \in in(v)} H(\mathbb{Z}(u) \cup \mathbb{D}(u) | \mathbb{W}(v)), \\ &\overset{(b)}{\geq} H(\mathbb{Z}(v) \cup \mathbb{D}(v) | \mathbb{W}(v)), \\ &\overset{(c)}{=} I(W_{new}(v); \mathbb{Z}(v) \cup \mathbb{D}(v) | \mathbb{W}(v)) \\ &+ H(\mathbb{Z}(v) \cup \mathbb{D}(v) | \mathbb{W}(v) \cup W_{new}(v), \end{aligned}$$

where inequality in (a) holds since $\mathbb{W}(u) \cup W_{new}(u) \subseteq \mathbb{W}(v)$ and conditioning decreases entropy, (b) holds since $\cup_{u \in in(v)} \mathbb{Z}(u) = \mathbb{Z}(v)$ and $\cup_{u \in in(v)} \mathbb{D}(u) = \mathbb{D}(v)$ and (c) holds by the definition of mutual information. Let V_{int} denote the set of internal nodes in $\mathcal{T}$. Let v^* denote the root and (u^*, v^*) denote its incoming edge. Then,

$$\begin{aligned} &\sum_{v \in V_{int}} \sum_{u \in in(v)} H(\mathbb{Z}(u) \cup \mathbb{D}(u) | \mathbb{W}(u) \cup W_{new}(u)) \geq \\ &\sum_{v \in V_{int}} y_{(v,out(v))} + \sum_{v \in V_{int}} H(\mathbb{Z}(v) \cup \mathbb{D}(v) | \mathbb{W}(v) \cup W_{new}(v)), \end{aligned}$$

where we have ignored the infinitesimal terms introduced due to Fano's inequality (for convenience of presentation). Note that the RHS of the inequality above contains entropy label of all nodes $v \in V_{int}$ (including u^*). On the other hand the LHS contains the entropy label of all nodes including the leaf nodes but excluding the node u^*. Canceling the common terms, we obtain,

$$\begin{aligned} &\sum_{i=1}^{\ell} H(\mathbb{Z}(v_i) \cup \mathbb{D}(v_i) | W_{new}(v_i)) \geq \\ &\sum_{v \in V_i} y_{(v,out(v))} + H(Z \cup \mathbb{D}(u^*) | \mathbb{W}(u^*), W_{new}(u^*)), \end{aligned}$$

since $\mathbb{W}(v_i) = \phi$ for $i = 1, \ldots, \ell$. We can therefore conclude that

$$\sum_{i=1}^{\ell} H(\mathbb{Z}(v_i), \mathbb{D}(v_i)) \geq \sum_{v \in V} y_{(v,out(v))} \quad (6)$$

$$\implies \sum_{i=1}^{\ell} H(\mathbb{Z}(v_i)) + \sum_{i=1}^{\ell} H(\mathbb{D}(v_i)) \geq \sum_{v \in V} y_{(v,out(v))} \quad (7)$$

Noting that $M \geq H(\mathbb{Z}(v_i))$ and $R^\star \geq H(\mathbb{D}(v_i))$ we have the required result. ∎

Claim 5: Consider a problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$ such that there exists a node $u \in \mathcal{T}$ with $|in(u)| \geq 3$. Then, there exists another instance $P'(\mathcal{T}', \alpha, \beta, L', N, K)$ where $L' \geq L$ and $|in(u)| \leq 2$ for all nodes $u \in \mathcal{T}'$.

Proof: We iteratively modify P to arrive at an instance where every node has in-degree at most two. Towards this end, we first identify a node u with in-degree $\delta \geq 3$ such that no other node of degree at least 3 is topologically higher than it.

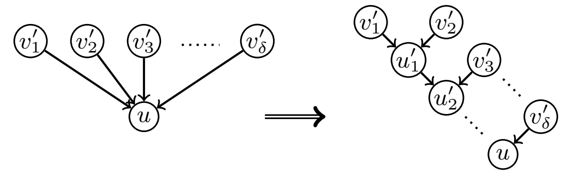


Fig. 3: Tree modification example

We modify the instance P by replacing u with a directed in-tree where each node has in-degree exactly two. Specifically, arbitrarily number the nodes in $in(u)$ from $v_1', \ldots, v_\delta'$. We replace the node u with a directed in-tree $\mathcal{T}_u$ with leaves $v_1', \ldots, v_\delta'$ and root u. $\mathcal{T}_u$ has $\delta - 2$ internal nodes numbered $u_1', \ldots, u_{\delta-2}'$ such that $in(u_i') = \{u_{i-1}', v_{i+1}'\}$ where $u_0' = v_1'$ (see Fig. 3). Let us denote the new instance by $P_o = P_o(\mathcal{T}_o, \alpha, \beta, L_o, N, K)$. We claim that $L_o \geq L$. To see this, suppose that $W^* \in W_{new}^P(u)$. We show that $W^* \in \cup_{u' \in \mathcal{T}_u} W_{new}^{P_o}(u')$. This ensures that $L_o \geq L$. To see this we note that

$$\begin{aligned} \mathbb{Z}^P(u) &= \mathbb{Z}^{P_o}(u) \\ \mathbb{D}^P(u) &= \mathbb{D}^{P_o}(u), \text{ and, thus} \\ \Delta^P(u,u) &= \Delta^{P_o}(u,u). \end{aligned}$$

Thus, if $W^* \in W_{new}^P(u)$, there exists an internal node $u_i' \in \mathcal{T}_u$ with the smallest index $i \in \{1, \ldots, \delta - 2\}$ such that $W^* \in \Delta^{P_o}(u_i', u_i')$. Note that if $i > 1$, we have $W^* \in W_{new}^{P_o}(u_i')$ since $W^* \notin \Delta^{P_o}(u_{i-1}', u_{i-1}')$ which in turn implies that $W^* \notin \mathbb{W}^{P_o}(u_i')$. On the other hand if $i = 1$, then a similar argument holds since it is easy to see that $W^* \notin \mathbb{W}^{P_o}(u_1')$.

Note that modification in instance P can only affect nodes that are downstream of u. Now consider u' such that $u \in in(u')$. It is evident that $\mathbb{Z}^{P_o}(u') = \mathbb{Z}^P(u')$ and $\mathbb{D}^{P_o}(u') = \mathbb{D}^P(u')$. Moreover $\mathbb{W}^{P_o}(u') = \cup_{v \in in(u')} \mathbb{W}^{P_o}(v) \cup W_{new}^{P_o}(v)$. Now for $v \neq u$, $\mathbb{W}^{P_o}(v) = \mathbb{W}^P(v)$ and $W_{new}^{P_o}(v) = W_{new}^P(v)$ as there are no changes in the corresponding subtrees. Moreover, as $\Delta^P(u,u) = \Delta^{P_o}(u,u)$, we have that $\mathbb{W}^{P_o}(u) \cup W_{new}^{P_o}(u) = \mathbb{W}^P(u) \cup W_{new}^P(u)$. This implies that $\mathbb{W}^{P_o}(u') = \mathbb{W}^P(u')$. Thus, we can conclude that $W_{new}^{P_o}(u') = W_{new}^P(u')$. Applying an inductive argument we can conclude that the $W_{new}^{P_o}(u') = W_{new}^P(u')$ for all u' such that $u \succ u'$.

The above process can iteratively be applied to every node in the instance that is of degree at least three. Thus, we have the required result. ∎

We define the function $\psi : \mathcal{D} \times \mathcal{C} \to \{0,1\}$ that allows us to express L in another way. For nodes $v_i \in \mathcal{D}, v' \in \mathcal{C}$ we can define their meeting point $u \in \mathcal{T}$. The function $\psi(v_i, v')$ is determined by means of Algorithm 2, where the sequence in which we pick the nodes $v_1, \ldots, v_\alpha$ is fixed. Each element of $W_{new}(u)$ can be recovered from multiple pairs of nodes that meet there. The array $\Omega(u, \delta_u)$ keeps track of the first time the file δ_u is encountered. The function $\psi(v_i, v')$ takes the value 1 if the file W^* recovered from the pair $(\mathbb{Z}(v'), \mathbb{D}(v_i))$ at u belongs to $W_{new}(u)$ and has not been encountered before and 0 otherwise. A formal description is given in Algorithm 2.

Algorithm 2 Computing ψ

Input: $P(\mathcal{T},\alpha,\beta,L,N,K)$, Array $\Omega(u,\delta_u)$, where $u\in\mathcal{T}$, $\delta_u\subseteq W_{new}(u), |\delta_u|=1$.

Initialization
  for all $u\in\mathcal{T}$, $\delta_u\subseteq W_{new}(u)$ where $|\delta_u|=1$ **do**
    $\Omega(u,\delta_u)\leftarrow 0$,
  end for
end Initialization
for $i\leftarrow 1$ to α **do**
  for all $v'\in\mathcal{C}$ **do**
    $\delta_u=\Delta_{(v_i,v')}(u)$.
    if $\delta_u\in W_{new}(u)$ and $\Omega(u,\delta_u)==0$ **then**
      $\psi(v_i,v')\leftarrow 1$, and $\Omega(u,\delta_u)\leftarrow 1$.
    else
      $\psi(v_i,v')\leftarrow 0$.
    end if
  end for
end for

Claim 6: For an instance $P(\mathcal{T},\alpha,\beta,L,N,K)$ the following equality holds

$$L=\sum_{i=1}^{\alpha}\sum_{v'\in\mathcal{C}}\psi(v_i,v'). \tag{8}$$

Proof: We first note that at the end of the algorithm above $\Omega(u,\delta_u)=1$ for all $u\in\mathcal{T}$ and all $\delta_u\subseteq W_{new}(u), |\delta_u|=1$. To see this suppose that there is a $u_1\in\mathcal{T}$ and a singleton subset δ_{u_1} of $W_{new}(u_1)$ such that $\Omega(u_1,\delta_{u_1})=0$. Now δ_{u_1} is recovered from some delivery phase node and cache node, otherwise it would not be a subset of $W_{new}(u_1)$. As our algorithm considers all pairs of delivery phase nodes and cache nodes, at the end of the algorithm it has to be the case that $\Omega(u_1,\delta_{u_1})=1$.

Next, we note that for each pair (u_1,δ_{u_1}) where $u_1\in\mathcal{T}$ and δ_{u_1} is singleton subset of $W_{new}(u_1)$, we can identify a unique pair of nodes (v_i,v') where $v_i\in\mathcal{D}$ and $v'\in\mathcal{C}$ such that $\psi(v_i,v')$ and $\Omega(u_1,\delta_{u_1})$ are set to 1 at the same step of the algorithm. The remaining pairs (v_i,v') that cannot be put in one to one correspondence with a pair (u_1,δ_{u_1}) are such that $\psi(v_i,v')$ are set to 0. Moreover as $\sum_{u\in\mathcal{T}}\sum_{\delta_u\subseteq W_{new}(u),|\delta_u|=1}\Omega(u,\delta_u)=\sum_{u\in\mathcal{T}}|W_{new}(u)|=L$, it follows that $L=\sum_{i=1}^{\alpha}\sum_{v'\in\mathcal{C}}\psi(v_i,v')$. ■

A. Proof of Claim 1

Proof: W.l.o.g let assume that condition (a), namely $\hat{\beta}_l<\min(\beta_l,K)$ holds. This implies that there is a set of leaves in $\mathcal{T}_{u^*(l)}$ denoted $\{v_{i_1},\ldots,v_{i_m}\}$ such that $\mathbb{Z}(v_{i_1})=\cdots=\mathbb{Z}(v_{i_m})=\{Z_j\}$. Let $\Lambda=\{u\in\mathcal{T}_{u^*(l)}:(v_{i_a},v_{i_b})$ meet at u, for all distinct $v_{i_a},v_{i_b}\in\{v_{i_1},\ldots,v_{i_m}\}\}$. We identify $u_0\in\Lambda$ such that no element of Λ is topologically higher than u_0 (note that u_0 may not be unique) and let $v^*_{i_a}$ and $v^*_{i_b}$ be the corresponding nodes in $\{v_{i_1},\ldots,v_{i_m}\}$ that meet at u_0. W.l.o.g we assume that $v^*_{i_b}\in\mathcal{T}_{u_0(r)}$ and $v^*_{i_a}\in\mathcal{T}_{u_0(l)}$.

We construct instance P' as follows. Choose a member of $\{Z_1,\ldots,Z_K\}\setminus\{\mathbb{Z}(v'):v'\in\mathcal{C}_{u^*(l)}\}$ and denote it by Z_k. We set $\mathbb{Z}^{P'}(v^*_{i_b})=\{Z_k\}$. Also, for any $u\in\mathcal{D}_{u_0(r)}$ and $\mathbb{D}^P(u)=X_{d_1,\ldots,d_K}$ we set $\mathbb{D}^{P'}(u)=X_{d'_1,\ldots,d'_K}$ such that $d'_j=d_k$ and $d'_k=d_j$ and $d'_i=d_i$ for $i\notin\{j,k\}$.

We now show that $L'\geq L$. In particular, for $u\in\mathcal{T}_{u_0(l)}$, we have $W^{P'}_{new}(u)=W^{P}_{new}(u)$. Also we claim that $W^{P'}_{new}(u)=W^{P}_{new}(u)$ for $u\in\mathcal{T}_{u_0(r)}$. To see this, note that for $v\in\mathcal{D}_{u_0(r)}$ and $v'\in\mathcal{C}_{u_0(r)}$ we have $\Delta^{P'}(v',v)=\Delta^{P}(v',v)$ if $\mathbb{Z}(v')\notin\{Z_j,Z_k\}$. If $\mathbb{Z}^{P'}(v')=\{Z_k\}$ and $\mathbb{D}^{P'}(v)=X_{d'_1,\ldots,d'_K}$ then,

$$\begin{aligned}\Delta^{P'}(v',v)&=Rec(\{Z_k\},X_{d'_1,\ldots,d'_K})\\&=\{W_{d'_k}\}=\{W_{d_j}\}\\&=Rec(\{Z_j\},X_{d_1,\ldots,d_K})\\&=\Delta^{P}(v',v).\end{aligned}$$

Furthermore note that there not exist any $v'\in\mathcal{C}_{u_0(r)}$ such that $\mathbb{Z}(v')=\{Z_j\}$ since we picked u_0 such that no element of Λ is topologically higher than u_0. It is not hard to see that this in turn implies that $W^{P'}_{new}(u)=W^{P}_{new}(u)$ for $u\in\mathcal{T}_{u_0(r)}$.

It follows therefore that $\mathbb{W}^{P'}(u_0)=\mathbb{W}^{P}(u_0)$ (see eq. (3)). Let us now consider the other nodes. As the changes are applied only to $\mathcal{T}_{u_0(r)}$ so $label(u)$ changes only for nodes u such that $u_0\succ u$. Consider the subset of internal nodes $U=\{u_0,u_1,\ldots,u_t\}$ such that $u_i\succ u_{i+1}$, i.e., the set of internal nodes including u_0 and all nodes downstream of u_0 such that u_t is root node. W.l.o.g we assume that $u_{i-1}\in\mathcal{T}_{u_i(l)}$ for $i\geq 1$. We now show that $\cup_{u\in U}W^{P}_{new}(u)\subseteq\cup_{u\in U}W^{P'}_{new}(u)$. Towards this end we have the following observations for $u\in U$.

$$\begin{aligned}&\mathbb{Z}^{P'}(u)=\mathbb{Z}^{P}(u)\cup\{Z_k\}\ \text{(from construction of } P')\\&\Delta^{P'}(u,u)=\cup_{v\in\mathcal{D}_u}\Delta^{P'}(u,v).\end{aligned}$$

Now, for $v\notin\mathcal{D}_{u_0(r)}$ we have $\mathbb{D}^{P'}(v)=\mathbb{D}^{P}(v)$ so that

$$\begin{aligned}\Delta^{P'}(u,v)&=Rec(\mathbb{Z}^{P'}(u),\mathbb{D}^{P'}(v))\\&=Rec(\mathbb{Z}^{P'}(u),\mathbb{D}^{P}(v))\\&\supseteq\Delta^{P}(u,v)\ \text{since}\ \mathbb{Z}^{P'}(u)\supseteq\mathbb{Z}^{P}(u).\end{aligned}$$

Conversely for $v\in\mathcal{D}_{u_0(r)}$ we have

$$Rec\left(\{Z_j,Z_k\},\mathbb{D}^{P'}(v)\right)=Rec\left(\{Z_j,Z_k\},\mathbb{D}^{P}(v)\right),$$

and

$$Rec\left(\{Z_i\},\mathbb{D}^{P'}(v)\right)=Rec\left(\{Z_i\},\mathbb{D}^{P}(v)\right),$$

for $Z_i\notin\{Z_j,Z_k\}$. Now note that $\{Z_k,Z_j\}\subset\mathbb{Z}^{P'}(u)$ so that

$$\begin{aligned}\Delta^{P'}(u,v)&=Rec\left(\mathbb{Z}^{P'}(u),\mathbb{D}^{P'}(v)\right)\\&=Rec\left(\mathbb{Z}^{P'}(u),\mathbb{D}^{P}(v)\right),\\&\supseteq Rec\left(\mathbb{Z}^{P}(u),\mathbb{D}^{P}(v)\right)=\Delta^{P}(u,v),\end{aligned}$$

since $\mathbb{Z}^{P'}(u)\supset\mathbb{Z}^{P}(u)$. We can therefore conclude that

$$\Delta^{P}(u,u)=\cup_{v\in\mathcal{D}_u}\Delta^{P}(u,v)\subseteq\cup_{v\in\mathcal{D}_u}\Delta^{P'}(u,v)=\Delta^{P'}(u).$$

Now we consider a $W^*\in W^{P}_{new}(u_i)$ so that $W^*\in\Delta^{P}(u_i,u_i)$ which by above condition means that $W^*\in\Delta^{P'}(u_i,u_i)$. Thus either $W^*\in W^{P'}_{new}(u_i)$ or $W^*\in\mathbb{W}^{P'}(u_i)$. In the latter case there exists a node $u_{i'}$ where $0\leq i'<i$ such that $W^*\in$

$W_{new}^{P'}(u_{i'})$ since we have shown that $\mathbb{W}^{P'}(u_0) = \mathbb{W}^P(u_0)$. Thus, we observe that

$$\begin{aligned} L' &= \sum_{u \in \mathcal{T}', u \notin U} |W_{new}^{P'}(u)| + |\cup_{u \in U} W_{new}^{P'}(u)|, \\ &\geq \sum_{u \in \mathcal{T}, u \notin U} |W_{new}^{P}(u)| + |\cup_{u \in U} W_{new}^{P}(u)|, \\ &= L, \end{aligned}$$

where the second inequality holds since $\sum_{u \in \mathcal{T}', u \notin U} |W_{new}^{P'}(u)| = \sum_{u \in \mathcal{T}, u \notin U} |W_{new}^{P}(u)|$ and $|\cup_{u \in U} W_{new}^{P'}(u)| \geq |\cup_{u \in U} W_{new}^{P}(u)|$.

To conclude we note that the above modification of the original instance can be iteratively repeated until we have $\hat{\beta}_l' = \min(\beta_l, K)$. Following this we can repeat the process on the instance P_r' if $\hat{\beta}_r' < \min(\beta_r, K)$ to ensure that $\hat{\beta}_r' = \min(\beta_r, K)$.

It remains to show that $\hat{\beta} = \min(\beta, K)$. Towards this end, we consider a variation of the above argument. Let v^* be the root of $\mathcal{T}$ so that (u^*, v^*) is the last edge in $\mathcal{T}$. By applying the above arguments we have $\hat{\beta}_l = |\mathbb{Z}(u_l^*)| = \min(\beta_l, K)$ and $\hat{\beta}_r = |\mathbb{Z}(u_r^*)| = \min(\beta_r, K)$. Note that if either β_l or β_r be larger or equal to K then $\hat{\beta} = |\mathbb{Z}(u^*)| = |\mathbb{Z}(u_l^*) \cup \mathbb{Z}(u_r^*)| = K$ or equivalently $\hat{\beta} = \min(\beta, K)$. So we only consider the case that both β_l and β_r are smaller than K and $\hat{\beta} < \min(\beta, K)$. In this case it is easy to see that there exists a unique $v_{i_1}^* \in \mathcal{C}_{u^*(r)}$ and a unique $v_{i_2}^* \in \mathcal{C}_{u^*(l)}$ such that $\mathbb{Z}(v_{i_1}^*) = \mathbb{Z}(v_{i_2}^*) = \{Z_j\}$ (for some Z_j).

We pick $Z_k \in \{Z_1, \ldots, Z_K\} \setminus \mathbb{Z}(u^*)$ and construct $P'(\mathcal{T}', \alpha, \beta, L', N, K)$ from P by applying the following changes. We set $\mathbb{Z}^{P'}(v_{i_1}^*) = \{Z_k\}$ and for any $v \in \mathcal{D}_{u^*(r)}$ such that $\mathbb{D}^P(v) = X_{d_1, \ldots, d_K}$ we set $\mathbb{D}^{P'}(v) = X_{d_1', \ldots, d_K'}$ where $d_j' = d_k$, $d_k' = d_j$ and $d_i' = d_i$ for all $i \notin \{j, k\}$.

Since, the changes only affect $\mathcal{T}_{u^*(r)}$ therefore $W_{new}^{P'}(u) = W_{new}^P(u)$ for all $u \in \mathcal{T}_{u^*(l)}$. Also by arguments similar to the ones made earlier, we can show that $W_{new}^{P'}(u) = W_{new}^P(u)$ for $u \in \mathcal{T}_{u^*(r)}$ therefore $\mathbb{W}^{P'}(u^*) = \mathbb{W}^P(u^*)$. Furthermore, since $\mathbb{Z}^{P'}(u^*) = \mathbb{Z}^P(u^*) \cup \{Z_k\}$ we have $W_{new}^P(u^*) \subseteq W_{new}^{P'}(u^*)$. Thus, we conclude that $L' \geq L$ while $\hat{\beta}' = \hat{\beta} + 1$.

If still $\hat{\beta}' < \min(\beta, K)$ then we keep applying the above changes until we construct P' such that $\hat{\beta}' = \min(\beta, K)$. ■

Proof of Lemma 1

Proof: For a node v_i, where $1 \leq i \leq \alpha$, we have

$$\begin{aligned} \sum_{v' \in \mathcal{C}} \psi(v_i, v') &\leq |\cup_{v' \in \mathcal{C}} \mathbb{Z}(v')| \\ &= \hat{\beta}, \\ &= \min(\beta, K). \end{aligned} \tag{9}$$

From eq. (9) we can conclude that $L \leq \alpha \min(\beta, K)$.

Remark 1: If $N \geq \alpha \min(\beta, K)$, then it is easy to construct an instance such that $L = \alpha \min(\beta, K)$. Specifically, pick any directed tree on $\alpha + \beta$ leaves. Suppose that node $v \in \mathcal{D}, v' \in \mathcal{C}$ meet at node u. We label the leaves such that $|\cup_{(v,v') \in \mathcal{D} \times \mathcal{C}} \Delta_{v,v'}(u)| = \alpha \min(\beta, K)$.

Given the conditions of the theorem, it is evident that there exists an index $i^* \in \{1, \ldots, \alpha\}$ such that $\sum_{v' \in \mathcal{C}} \psi(v_{i^*}, v') < \min(\beta, K)$. We set i^* to be the smallest such index. Let $\Pi^1(v_{i^*}) = \{v' \in \mathcal{C} : \psi(v_{i^*}, v') = 1\}$ and $\Pi^0(v_{i^*}) = \{v' \in \mathcal{C} : \psi(v_{i^*}, v') = 0, \mathbb{Z}(v') \nsubseteq \cup_{v \in \Pi^1(v_{i^*})} \mathbb{Z}(v')\}$. Note that $\Pi^0(v_{i^*})$ is non-empty since $|\cup_{v' \in \mathcal{C}} \mathbb{Z}(v')| = \min(\beta, K)$ and $\sum_{v' \in \mathcal{C}} \psi(v_{i^*}, v') < \min(\beta, K)$.

Next, we determine the set of nodes where v_{i^*} and the nodes in $\Pi^0(v_{i^*})$ meet, i.e., we define $\Lambda^0(v_{i^*}) = \{u \in \mathcal{T} : \exists v' \in \Pi^0(v_{i^*})$ such that v_{i^*} and v' meet at $u.\}$. Note that there is a topological ordering on the nodes in $\Lambda^0(v_{i^*})$. Pick the node $u^* \in \Lambda^0(v_{i^*})$ such that no element of $\Lambda^0(v_{i^*})$ is topologically higher than u^*. Let the corresponding node in $\Pi^0(v_{i^*})$ be denoted by v_{j^*} where $j^* \in \{\alpha + 1, \ldots, \alpha + \beta\}$.

Suppose that $\mathbb{Z}(v_{j^*}) = \{Z_k\}$ and that $\mathbb{D}(v_{i^*}) = X_{d_1, \ldots, d_K}$. We modify the instance P as follows. Set $d_k = N+1$ (i.e., the index of the $N+1$ file). Thus, the only change is in $\mathbb{D}(v_{i^*})$. Let us denote the new instance by $P' = P(\mathcal{T}', \alpha, \beta, L', N+1, K)$.

We now analyze the value of L'. W.l.o.g. we assume that $v_{i^*} \in \mathcal{T}'_{u^*(l)}$ and $v_{j^*} \in \mathcal{T}'_{u^*(r)}$. Note that $W_{new}^{P'}(u) = W_{new}^P(u)$ for $u \in \mathcal{T}'_{u^*(r)}$ as the subtree $\mathcal{T}'_{u^*(r)}$ is identical to $\mathcal{T}_{u^*(r)}$. We also have

$$W_{new}^{P'}(u) = W_{new}^P(u) \text{ for } u \in \mathcal{T}'_{u^*(l)}.$$

To see this suppose that this is not true. This implies that the file W_{N+1} is recovered at some node in $\mathcal{T}'_{u^*(l)}$, i.e., there exists $v' \in \mathcal{C}$ such that $v' \in \mathcal{T}'_{u^*(l)}$ and $\mathbb{Z}(v') = Z_k$. However this is a contradiction, since this implies the existence of node that is topologically higher than u^* where v_{i^*} and v' meet. It follows that $\mathbb{W}^{P'}(u^*) = \mathbb{W}^P(u^*)$.

Next, we claim that $W_{new}^{P'}(u^*) = W_{new}^P(u^*) \cup \{W_{N+1}\}$. To see this consider the following series of arguments. Let the singleton subset $\Delta^P(v_{i^*}, v_{j^*}) = \{W^*\}$. Note that $\psi^P(v_{i^*}, v_{j^*}) = 0$. This implies that there exist $v \in \mathcal{D}_{u^*}$ and $v' \in \mathcal{C}_{u^*}$ such that v and v' meet at u^* and recover the file W^* where $(v, v') \neq (v_{i^*}, v_{j^*})$. Thus, as $\mathbb{Z}^{P'}(u^*) = \mathbb{Z}^P(u^*)$, we can conclude that

$$\begin{aligned} \Delta^{P'}(u^*, u^*) &= Rec(\mathbb{Z}^P(u^*), \mathbb{D}^{P'}(u^*)) \\ &= \Delta^P(u^*, u^*) \cup \{W_{N+1}\}. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} &W_{new}^{P'}(u^*) = \Delta^{P'}(u^*, u^*) \setminus \mathbb{W}^{P'}(u^*) \\ &= \Delta^P(u^*, u^*) \cup \{W_{N+1}\} \setminus \mathbb{W}^P(u^*) \\ &= W_{new}^P(u^*) \cup \{W_{N+1}\}, \text{ (since } W_{N+1} \notin \mathbb{W}^P(u^*)). \end{aligned}$$

For u such that $u^* \succ u$ we inductively argue that $W_{new}^{P'}(u) = W_{new}^P(u)$. To see this suppose that $u^* = u_r$. It is evident that $\Delta_{rl}^{P'}(u) = \Delta_{rl}^P(u)$. Next, $\Delta_{lr}^{P'}(u) = \Delta_{lr}^P(u)$ since $Z_k \notin \mathbb{Z}(u_l) \setminus \mathbb{Z}(u_r)$. Thus, as $W_{N+1} \notin \Delta_{rl}^P(u) \cup \Delta_{lr}^P(u)$,

$$\begin{aligned} W_{new}^{P'}(u) &= \Delta_{rl}^{P'}(u) \cup \Delta_{lr}^{P'}(u) \setminus \mathbb{W}^{P'}(u) \\ &= \Delta_{rl}^P(u) \cup \Delta_{lr}^P(u) \setminus \mathbb{W}^{P'}(u) \\ &= \Delta_{rl}^P(u) \cup \Delta_{lr}^P(u) \setminus \mathbb{W}^P(u) \cup \{W_{N+1}\} \\ &= \Delta_{rl}^P(u) \cup \Delta_{lr}^P(u) \setminus \mathbb{W}^P(u) \\ &= W_{new}^P(u). \end{aligned}$$

Next, we note that $\mathbb{W}(u) = \mathbb{W}(u_r) \cup W_{new}(u_r) \cup \mathbb{W}(u_l) \cup W_{new}(u_l)$. It is evident that $\mathbb{W}^{P'}(u_l) = \mathbb{W}^P(u_l)$ and $W_{new}^{P'}(u_l) = W_{new}^P(u_l)$. Next, $\mathbb{W}^{P'}(u_r) = \mathbb{W}^{P'}(u^*) = \mathbb{W}^P(u^*)$ (from above) and $W_{new}^{P'}(u^*) = W_{new}^P(u^*) \cup \{W_{N+1}\}$, so that $\mathbb{W}^{P'}(u) = \mathbb{W}^P(u) \cup \{W_{N+1}\}$.

As the induction hypothesis we assume that for any node u downstream of u^*, we have $W_{new}^{P'}(u) = W_{new}^P(u)$ and $\mathbb{W}^{P'}(u) = \mathbb{W}^P(u) \cup \{W_{N+1}\}$. Consider a node u' such that $u'_r = u$. As before we have $\mathbb{W}^{P'}(u'_l) = \mathbb{W}^P(u'_l)$, $W_{new}^{P'}(u'_l) = W_{new}^P(u'_l)$. Moreover, we have $\mathbb{W}^{P'}(u'_r) = \mathbb{W}^P(u'_r) \cup \{W_{N+1}\}$ and $W_{new}^{P'}(u'_r) = W_{new}^P(u'_r)$, by the induction hypothesis, so that $\mathbb{W}^{P'}(u') = \mathbb{W}^P(u') \cup \{W_{N+1}\}$.

Next, we argue similarly as above that $\Delta_{rl}^{P'}(u') = \Delta_{rl}^P(u')$ and $\Delta_{lr}^{P'}(u') = \Delta_{lr}^P(u')$ and the sequence of equations above can be used to conclude to that $W_{new}^{P'}(u') = W_{new}^P(u')$.

We conclude that $L' = L + 1$. ■

B. Proof of Claim 2

Proof: We pick one of the nodes where $\Gamma_r \nsubseteq \Gamma_l$ and apply the following arguments. We denote this node by u^*. Note that since $|\Gamma_l| \geq |\Gamma_r|$, there exists an injective mapping $\phi : \Gamma_r \setminus \Gamma_l \to \Gamma_l \setminus \Gamma_r$. Let $\mathbb{Z}(u_r^*) = \{Z_{i_1}, \ldots, Z_{i_m}\}$.

We construct the instance P' as follows. For a $v \in \mathcal{D}_{u_r^*}$ suppose $\mathbb{D}(v) = \{X_{d_1,\ldots,d_K}\}$. For $j = 1, \ldots, m$, if $d_{i_j} \in \Gamma_r \setminus \Gamma_l$, we replace it by $\phi(d_{i_j})$; otherwise, we leave it unchanged. In other words, we modify the delivery phase signals so that the files that are recovered in $\mathcal{T}_{u^*(r)}$ are a subset of those recovered in $\mathcal{T}_{u^*(l)}$.

As our change amounts to a simple relabeling of the sources, for $u \in \mathcal{T}_{u^*(r)}$ we have $|W_{new}^{P'}(u)| = |W_{new}^P(u)|$. Furthermore, the relabeling of the sources only affects $u \in \mathcal{T}'$ such that $u^* \succ u$. Note that $\mathbb{W}^{P'}(u^*) \subset \mathbb{W}^P(u^*)$ (the inclusion is strict since there is at least one source in $\Gamma_r \setminus \Gamma_l$ is mapped to $\Gamma_l \setminus \Gamma_r$) since we have $\Gamma_r^{P'} \subseteq \Gamma_l^{P'}$ and $\Gamma_l^{P'} = \Gamma_l^P$.

Now, we note that

$$\Delta_{rl}^{P'}(u^*) = \Delta_{rl}^P(u^*), \text{ and}$$
$$\Delta_{lr}^{P'}(u^*) = \Delta_{lr}^P(u^*),$$

where the first equality holds since $\mathbb{Z}^P(u_r^*) = \mathbb{Z}^{P'}(u_r^*)$, $\mathbb{Z}^P(u_l^*) = \mathbb{Z}^{P'}(u_l^*)$ and $\mathbb{D}^P(u_l^*) = \mathbb{D}^{P'}(u_l^*)$. The second equality holds since our modification to the delivery phase signals in $\mathcal{T}_{u^*(r)}$ does not affect files that are recovered from $\mathbb{Z}^P(u_l^*) \setminus \mathbb{Z}^P(u_r^*)$. It follows therefore that $|W_{new}^{P'}(u^*)| \geq |W_{new}^P(u^*)|$.

We make an inductive argument for nodes u that are downstream of u^*; w.l.o.g. we assume that $u^* \in \mathcal{T}_{u(r)}$. Specifically, our inductive hypothesis is that for a node u that is downstream of u^*, we have $\mathbb{W}^{P'}(u) \subseteq \mathbb{W}^P(u)$, $\Delta_{rl}^{P'}(u) = \Delta_{rl}^P(u)$ and $\Delta_{lr}^{P'}(u) = \Delta_{lr}^P(u)$.

Now consider a node u' downstream of u such that $u'_r = u$. We have, $\mathbb{W}(u') = \mathbb{W}(u'_l) \cup W_{new}(u'_l) \cup \mathbb{W}(u) \cup W_{new}(u)$. Note that we can express $\mathbb{W}(u) \cup W_{new}(u) = \mathbb{W}(u) \cup \Delta_{rl}(u) \cup \Delta_{lr}(u)$. It is evident that $\mathbb{W}^{P'}(u'_l) = \mathbb{W}^P(u'_l)$ and $W_{new}^{P'}(u'_l) = W_{new}^P(u'_l)$. Moreover, by the induction hypothesis, $\mathbb{W}^{P'}(u) \subseteq \mathbb{W}^P(u)$ and $\Delta_{rl}^{P'}(u) \cup \Delta_{lr}^{P'}(u) = \Delta_{rl}^P(u) \cup \Delta_{lr}^P(u)$. Thus, the induction step is proved. We conclude therefore that $L' \geq L$ and that $N' < N$. ■

Algorithm 3 Upper Bound on $N_{sat}(\alpha, \beta, K)$

Input: α, β and K.
Initialization
  Let (u^*, v^*) be last edge and set $U_{new} = \{u^*\}$.
  Set $\mathbb{Z}(u^*)$ be any subset of $\{Z_1, \ldots, Z_K\}$ of size $\min(\beta, K)$ and $\beta(u^*) = \beta$, $\alpha(u^*) = \alpha$.
  $\mathcal{C} = \emptyset$ and $\mathcal{D} = \emptyset$.
end Initialization
procedure CACHE NODES LABELING
  while U_{new} is nonempty **do**
    Pick u from U_{new}, create nodes u_l and u_r, edges (u_l, u) and (u_r, u), add them to $\mathcal{T}$.
    Set $\alpha_l(u) = \lceil \alpha(u)/2 \rceil$, $\beta_l(u) = \lfloor \beta(u)/2 \rfloor$ and $\alpha_r(u) = \alpha(u) - \alpha_l(u)$, $\beta_r(u) = \beta(u) - \beta_l(u)$.
    Set $\mathbb{Z}(u_l)$ and $\mathbb{Z}(u_r)$ be subsets of $\mathbb{Z}(u)$ of sizes $\min(\beta_l(u), K)$ and $\min(\beta_r(u), K)$ respectively with minimum intersection.
    Remove u from U_{new}.
    if $\alpha_l(u) + \beta_l(u) \geq 2$ **then**
      Add u_l to U_{new}.
    else
      If $\beta_l(u) == 1$ add u_l to $\mathcal{D}$ otherwise to $\mathcal{C}$.
    end if
    if $\alpha_r(u) + \beta_r(u) \geq 2$ **then**
      Add u_r to U_{new}.
    else
      If $\beta_r(u) == 1$ add u_r to $\mathcal{D}$ otherwise to $\mathcal{C}$.
    end if
  end while
end procedure
Set $U_{unlab} = \{u \in \mathcal{T} : in(u) \subset \mathcal{C} \cup \mathcal{D}\}$ and $U_{lab} = \emptyset$.
Set $n_v = 0$ for all $v \in \mathcal{C} \cup \mathcal{D}$.
procedure DELIVERY NODES LABELING
  while U_{unlab} is not empty **do**
    Pick $u \in U_{unlab}$ such that $in(u) \subseteq U_{lab}$ and set $\mathcal{I}_l = \{i : Z_i \in \mathbb{Z}(u_l) \setminus \mathbb{Z}(u_r)\}$ and $\mathcal{I}_r = \{i : Z_i \in \mathbb{Z}(u_r) \setminus \mathbb{Z}(u_l)\}$.
    Set $\rho_{rl} = |\mathcal{I}_r| \times |\mathcal{D}_{u(l)}|$ and $\rho_{lr} = |\mathcal{I}_l| \times |\mathcal{D}_{u(r)}|$.
    Set $j = 1$, $m_u = \max(n_{u_l}, n_{u_r})$ and $n_u = \rho_{rl} + \rho_{lr} + m_u$.
    for all $v \in \mathcal{D}_{u(l)}$ **do**
      Let $\mathbb{D}(v) = X_{d_1,\ldots,d_K}$.
      for $i \in \mathcal{I}_r$ **do**
        Set $d_i = j + m_u$.
        $j \leftarrow j + 1$.
      end for
    end for
    for all $v \in \mathcal{D}_{u(r)}$ **do**
      Let $\mathbb{D}(v) = X_{d_1,\ldots,d_K}$.
      for $i \in \mathcal{I}_l$ **do**
        Set $d_i = j + m_u$.
        $j \leftarrow j + 1$.
      end for
    end for
    Remove u from U_{unlab} and add it to U_{lab}.
    If $out(u) \notin U_{lab}$ add it to U_{unlab}.
  end while
end procedure
Output: $\hat{N}_{sat}(\alpha, \beta, K) = n_{v^*}$.

C. Proof of Claim 3

Proof: Suppose that this is not the case. This means that there exist at least two delivery nodes $v_{i_1}, v_{i_2} \in \mathcal{D}$ such that $\mathbb{D}(v_{i_1}) = \mathbb{D}(v_{i_2})$. Let u^* be the node where v_{i_1} and v_{i_2} meet and w.l.o.g let $i_1 < i_2$ and $v_{i_2} \in \mathcal{T}_{u^*(r)}$. In eq. (9) we show that $\sum_{v' \in \mathcal{C}} \psi(v, v') \leq \min(\beta, K)$ for any $v \in \mathcal{D}$. As $L = \alpha\beta$, we conclude that $\sum_{v' \in \mathcal{C}} \psi(v, v') = \beta$ for any $v \in \mathcal{D}$. Next, as $i_1 < i_2$ and v_{i_1}, v_{i_2} have the same label (and any two leaves meet at some node) we have that $\psi(v_{i_2}, v') = 0$ for $v' \notin \mathcal{C}_{u^*(r)}$. From this, we can conclude that $\mathcal{C}_{u^*(r)} = \mathcal{C}$. However, this implies that $Rec(\mathbb{Z}(u^*), \mathbb{D}(v_{i_1}) \backslash \mathbb{W}(u^*) = \emptyset$ since all files in $Rec(\mathbb{Z}(u^*), \mathbb{D}(v_{i_1}))$ have already been recovered in subtree $\mathcal{T}_{u^*(r)}$ and $\hat{\beta} = \beta$. This in turn implies that $\psi(v_{i_1}, v') = 0$ for all $v' \in \mathcal{C}$ which contradicts the fact that $L = \alpha\beta$. ■

D. Upper bound on $N_{sat}(\alpha, \beta, K)$

Note that by the definition of saturation number any problem instance $P(\mathcal{T}, \alpha, \beta, L, N, K)$ where $L = \alpha \min(\beta, K)$ can be used to find upper bound for $N_{sat}(\alpha, \beta, K)$. We claim that the proposed problem instance in Algorithm 3 is such that $L = \alpha \min(\beta, K)$ hence $\hat{N}_{sat}(\alpha, \beta, K)$ is a valid upper bound. To see this, we have the following argument.

For fixed $u \in \mathcal{T}$, by the algorithm for $v \in \mathcal{D}_{u(l)}$ with $\mathbb{D}(v) = X_{d_1,\dots,d_K}$ and $i \in \mathcal{I}_r$ such that $\{Z_i\} = \mathbb{Z}(v')$ for some $v' \in \mathcal{C}_{u(r)}$, we have $\Delta_{(v,v')}(u) = \{W_{d_i}\}$ therefore $W_{d_i} \in \Delta(u, u)$. It is easy to verify that $W_{d_i} \in W_{new}(u)$ where $d_i = j + m_u$, $j \in \{1, \dots, \rho_{lr} + \rho_{rl}\}$. Note that $d_i = j + m_u > m_u \geq n_v$ for any $v \succ u$. Now assume that $W_{d_i} \notin W_{new}(u)$. Since $W_{d_i} \in \Delta(u, u)$ this means that there exists a node $v \succ u$ where $W_{d_i} \in W_{new}(v)$ so that $d_i = j' + m_v$. But this contradicts the assumption that $d_i > n_v$ since $d_i = j' + m_v \leq n_v$ therefore $W_{d_i} \in W_{new}(u)$.

By our algorithm for different choices of $i \in \mathcal{I}_r$ and $v \in \mathcal{D}_{u(l)}$ there exist a distinct $j \in \{1, \dots, \rho_{rl}\}$ such that $d_i = j + m_u$, therefore d_i takes all values in $m_u + 1, \dots, m_u + \rho_{rl}$. By the same argument for $\mathcal{D}_{u(r)}$, $\mathcal{I}_l$ and $j \in \{\rho_{rl} + 1, \dots, n_u\}$ the number of distinct values for d_i is $\rho_{rl} + \rho_{lr}$. Moreover, since $W_{d_i} \in W_{new}(u)$ it is easy to see that $|W_{new}(u)| \geq \rho_{rl} + \rho_{lr}$. Note that in the cache labeling phase of the algorithm we choose $\mathbb{Z}(u_l)$ and $\mathbb{Z}(u_r)$ to have the minimum possible intersection. Therefore, $|\mathcal{I}_r| = [\min(\beta_r(u), K - \beta_l(u))]^+$. A similar argument holds for $|\mathcal{I}_l|$ and we can conclude that

$$
\begin{aligned}
|W_{new}(u)| \geq \rho_{rl} + \rho_{lr}, \\
&= \alpha_l(u)[\min(\beta_r(u), K - \beta_l(u))]^+ \\
&+ \alpha_r(u)[\min(\beta_l(u), K - \beta_r(u))]^+. \qquad (10)
\end{aligned}
$$

Along with the matching upper bound in Claim 4, we can assert that equality holds in eq. (10).

Now we claim that $\sum_{v \succeq u} |W_{new}(v)| = \alpha(u) \min(\beta(u), K)$ by induction. It is easy to verify that this is true when $u \in \mathcal{C} \cap \mathcal{D}$ since for leaves either $\mathbb{Z}(u)$ or $\mathbb{D}(u)$ is empty therefore $W_{new}(u) = \emptyset$ and either $\alpha(u)$ or $\beta(u)$ is zero. Assume that the equation holds for u_l and u_r and we show that it also holds for u. To see this note that

$$
\begin{aligned}
&\sum_{v \succeq u} |W_{new}(v)| \\
&= \sum_{v \succeq u_l} |W_{new}(v)| + \sum_{v \succeq u_r} |W_{new}(v)| + |W_{new}(u)| \\
&= \alpha_l(u) \min(\beta_l(u), K) + \alpha_r(u) \min(\beta_r(u), K) + |W_{new}(u)| \\
&= \alpha_l(u)\{\min(\beta_l(u), K) + [\min(\beta_r(u), K - \beta_l(u))]^+\} \\
&+ \alpha_r(u)\{\min(\beta_r(u), K) + [\min(\beta_r(u), K - \beta_r(u))]^+\},
\end{aligned}
$$

where we used the size of $W_{new}(u)$ in eq. (10) and the fact that the induction hypothesis holds for u_l and u_r. It is easy to verify that $\min(\beta_l(u), K) + [\min(\beta_r(u), K - \beta_l(u))]^+ = \min(\beta, K)$ where $\beta_l(u) + \beta_r(u) = \beta(u)$. Therefore,

$$
\sum_{v \succeq u} |W_{new}(v)| = \alpha(u) \min(\beta(u), K).
$$

So the equation $\sum_{v \succeq u} |W_{new}(v)| = \alpha(u) \min(\beta(u), K)$ holds for any $u \in \mathcal{T}$ in the algorithm 3. Applying this to the node u^* where $\alpha(u^*) = \alpha$ and $\beta(u^*) = \beta$ completes our claim.