entropy

MDPI

Article

Cryptography in Hierarchical Coded Caching: System Model and Cost Analysis

Behrouz Zolfaghari [1,2], Vikrant Singh [3], Brijesh Kumar Rai [3], Khodakhast Bibak [4,*] and Takeshi Koshiba [5]

[1] CSE Department, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India; behrouz@cybersciencelab.org
[2] Cyber Science Lab, School of Computer Science, University of Guelph, Guelph, ON N1G 2W1, Canada
[3] EEE Department, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India; vs13@iitg.ac.in (V.S.); brijesh.rai@gmail.com (B.K.R.)
[4] Department of Computer Science and Software Engineering, Miami University, Oxford, OH 45056, USA
[5] Department of Mathematics, Faculty of Education and Integrated Arts and Sciences, Waseda University, Tokyo 169-8050, Japan; tkoshiba@waseda.jp
* Correspondence: bibakk@miamioh.edu

Abstract: The idea behind network caching is to reduce network traffic during peak hours via transmitting frequently-requested content items to end users during off-peak hours. However, due to limited cache sizes and unpredictable access patterns, this might not totally eliminate the need for data transmission during peak hours. Coded caching was introduced to further reduce the peak hour traffic. The idea of coded caching is based on sending coded content which can be decoded in different ways by different users. This allows the server to service multiple requests by transmitting a single content item. Research works regarding coded caching traditionally adopt a simple network topology consisting of a single server, a single hub, a shared link connecting the server to the hub, and private links which connect the users to the hub. Building on the results of Sengupta et al. (IEEE Trans. Inf. Forensics Secur., 2015), we propose and evaluate a yet more complex system model that takes into consideration both throughput and security via combining the mentioned ideas. It is demonstrated that the achievable rates in the proposed model are within a constant multiplicative and additive gap with the minimum secure rates.

Keywords: coded caching; secure delivery; hierarchical coded caching; cost analysis; system model

check for updates

Citation: Zolfaghari, B.; Singh, V.; Rai, B.K.; Bibak, K.; Koshiba, T. Cryptography in Hierarchical Coded Caching: System Model and Cost Analysis. *Entropy* **2021**, *23*, 1459. https://doi.org/10.3390/e23111459

Academic Editor: Boris Ryabko

Received: 8 October 2021
Accepted: 1 November 2021
Published: 3 November 2021

1. Introduction

Coded caching, proposed by Maddah-Ali and Niesen [1], refers to an augmented variant of caching. Coded caching follows two strategies during two transmission phases in order to avoid a traffic bottleneck in the network. The first transmission phase, referred to as the placement phase, takes place in off-peak hours. During this phase the system attempts at placing frequently-demanded content items in local memories of corresponding interested users in order to avoid unnecessary transmission during peak time. This helps deteriorate network bandwidth over utilization and underutilization problems during peak and off-peak intervals. An effective placement strategy should consider the statistical and probabilistic nature of the users' access patterns. The second phase, i.e., the delivery phase manages the transmission in peak hours. The ideal goal in the latter phase is to send only a single coded content item which is a function of the originally-requested content items. Each user—in the ideal case—should be able to calculate its own demanded item from the transmitted item. The more the system approaches this goal, the less amount of transmission during the delivery phase (rate) is required.

The authors in [1] made a lot of simplifying assumptions when establishing the first system model for a coded caching scheme. They assumed a simple network based on a star topology which provides a one-way content transmission from a single server storing

N files each of size F bits to K users and each user having cache size of MF bits. Each user requests a single file during the delivery phase. Every file sent by the server passes through a single shared link and arrives at the hub where it is duplicated and transmitted to each user through a private link.

This system model is obviously not realistic enough because it ignores many considerations of a real-world network among which we focus on scalability and security in this paper. There are various security-related issues, such as confidentiality, privacy, and distributed denial-of-service (DDoS) attack protection which need to be addressed in coded caching. Among the mentioned issues, confidentiality has received the most focus in recent years [2,3]. Previous works in this area have augmented the coded caching system model by adding an adversary with access to the shared link only during the delivery phase. The space required to store the cryptographic keys in the server memory and user caches, as well as the extra traffic caused by key exchange mechanisms should be considered as obvious costs of this variant of coded caching.

In order to address the scalability issue, some researchers have augmented the coded caching system model in another way via proposing hierarchical network topology [4]. In the proposed topology, the main server is mirrored in each cluster of users. This allows part of the traffic to be locally handled in user clusters which leads to improved scalability. This improvement is achieved at the cost of redundant servers and links.

Although scalability and security have been separately examined in previous research, the literature in this area has not come up with a study on the possibility or the costs of considering both issues at the same time. This paper addresses both of the mentioned issues via considering confidential content transmission over a hierarchical network. This goal is achieved by further augmenting the coded caching system model, as well as analyzing the related costs. In our proposed system model, the adversary can eavesdrop the shared links in each hierarchy level during the peak interval.

The costs of scalability have already been analyzed in previous research [4]. We compare the results of our mathematical cost evaluations with those obtained in [4] to analyze the extra cost posed by confidentiality considerations. The key contribution of the paper is the result that although the achievable rates are within a constant multiplicative and additive gap to the corresponding lower bounds in both schemes, confidentiality causes the constants to grow larger.

The rest of this paper is organized as follows. Section 2 defines the problem we are tackling in this paper. This section first studies relevant works and presents some preliminaries, and then the shortcomings of the previous works which motivate our work in this paper are discussed. Section 3 explains the secure hierarchical coded caching scheme and describes the system model and configuration. The fundamental limits as well as costs are analyzed in Section 4. In this section, the secure achievable rates, memory requirements and the lower bounds on the rates are calculated. A gap analysis between the secure achievable rates and the corresponding lower bounds is presented in Section 5. The last section of this paper is Section 6 which concludes the paper and suggests further research topics.

2. Problem Statement

In this section, we first present some preliminary discussions regarding coded caching and review the related literature and then highlight some shortcomings in the related works which motivate us to propose the secure hierarchical coded caching scheme.

2.1. Related Works

Caching is a solution to the problem of temporally-nonuniform access to contents stored in servers which may causes the network bandwidth to be underutilized in the off-peak interval while it can render a bottleneck in the peak interval. This technique helps achieve more uniform network traffic and deteriorate the bottleneck problem by allowing the system to store frequently-accessed content items in local caches during the off-peak time.

Coded caching has been a research focus during recent years [5–9]. Coded caching is finding its application in modern technologies and services, such as content delivery [10–12], mobile computing [13–15], and information-enteric networks [16]. Different aspects of coded caching have recently been studied among which we can refer to as [17], centralized [18,19] and decentralized [18,19] coded caching, placement [20] and delivery [21,22] schemes, as well as added pre-fetching phase [23], multi-casting [22,24–26], scheduling [27], error correction [28], clustering [29], heterogeneity [12,25,30], the impact of file size [31,32], dealing with non-uniform user demands [33] and peak-time traffic reduction [1]. Moreover, security in coded caching has been considered as a concern [20,34–36] and cryptography has been among the best-studied security mechanisms for use in coded caching [37,38].

Examining the above problems has led to different variants of caching schemes. In this paper, we focus on coded caching schemes which try to service as many user requests as possible by transmitting a single coded data item in the peak time. Coded caching schemes can be classified into the following categories with respect to their behaviors in the placement phase.

2.2. Centralized Schemes

In these schemes, the server decides the data items which are to be stored in user caches during the placement phase [1,39–43]. It has been shown that a multiplicative factor of $\frac{1}{1+KM/N}$ in size reduces the rate in centralized coded caching. This factor is referred to as *global caching gain*. As shown in [1], the centralized coded caching rate $R_C(M)$ is given by (1),

$$R_C(M) \triangleq K \cdot (1 - M/N) \cdot \min\left\{ \frac{1}{1+KM/N}, \frac{N}{K} \right\}. \tag{1}$$

2.3. Decentralized Schemes

In the latter schemes, users are allowed to store random data in their caches. It was shown in [44] that the rate [1] in decentralized coded caching can be obtained from (2),

$$R_D(M) \triangleq K \cdot (1 - M/N) \cdot \min\left\{ \frac{N}{KM}(1 - (1 - M/N)^K), \frac{N}{K} \right\}. \tag{2}$$

An important point to note here is that the term "decentralized" does not refer to the underlying network and the network topology adopted in [44,45] are the same as the one considered in [1].

2.4. Hierarchical Coded Caching Scheme

The scheme introduced in [4] proposes a hierarchical coded caching scheme in which the content stored in the main server can be mirrored by intermediate servers in different levels of hierarchy before being placed in end user caches. In this scheme, the requests issued by each end user are first forwarded to the closest intermediate server. If not serviced the request is then forwarded to the higher hierarchy level. This implies the existence of different peak and off-peak intervals in different hierarchy levels.

Two different caching schemes have been proposed in this paper. The first scheme referred to as Scheme A allows simultaneous coded multicasting in both hierarchy levels. Each mirror first downloads the content items requested by its corresponding users from the main servers. Then, the items are coded and forwarded to the users. In Scheme B, mirrors act as memory-less routers. They receive the items from the main server and forward them without being stored or coded. It has been demonstrated that both schemes can individually perform sub-optimally [4]. The authors in this paper argued that because of the disjunctive relation between Scheme A and Scheme B, the rate of each link is the sum of the individual rates induced by the two schemes. They proposed a hybrid scheme named as the *generalized coded caching scheme* that attempts to incorporates a proper combination of Scheme A and Scheme B in order to approximately minimize the overall rate.

2.5. Secure Coded Caching Scheme

The scheme presented in [3] argued that the shared link may be eavesdropped by an adversary since it is publicly accessible as a broadcast medium. Thus they proposed an on time pad (OTP) cryptosystem [46] to preserve the confidentiality of the data items exchanged through this link. In their proposed scheme, the keys are placed in user cache along with the data in the placement phase. These confidentiality measures can be applied in both centralized and decentralized coded caching systems.

It is demonstrated in this paper that the secure rates for the centralized scheme and the decentralized scheme can be obtained through replacing M/N by $(M-1)/(N-1)$ in Equations (1) and (2), respectively. The authors of [4] argued that the overall rate of the hierarchical network is the sum of the individual rates in different levels of the hierarchy. Thus, if the overall rate needs to be minimized, both levels should operate at their minimum rates.

The relationship between the goals followed by the mentioned schemes motivates our work in this paper. Moreover, we compare our results with the ones obtained in [4] as reference.

2.6. Motivations

The researchers who proposed the idea of coded caching made several simplifying assumptions regarding the system model [1]. These assumptions made the core idea more manageable in its early days. However, several aspects of the primary system model obviously need to be revisited in order for the scheme to be applicable to real-world networks. Scalability and security are two aspects considered by other researchers [3,4]. However, there are still several related issues which can motivate further research. For example, It should be considered that confidentiality (addressed in [3]) is not the only aspect of security. Moreover, the network topology (studied in [4]) is not the only factor affecting scalability. However, what motivates us for the work of this paper is the lack of a research on a system model which is both secure and confidential.

Achieving the confidentiality promised in [3], as well as the scalability of the network studied in [4] by combining both ideas looks an enticing natural idea. However, the important issue to consider here is that combining these ideas can bring about new problems. In fact, the traffic and memory space overhead caused by the secure coded caching is against the scalability aimed by the hierarchical network. The key transmission occupies the bandwidth of the network which adversely affects the scalability. This problem will look more prominent when we consider the fact that OTP requires a new key for each single transmission. On the other hand, storing the keys in user caches prevents some frequently-requested data items to be stored during the placement phase because of the limited cache sizes. This will affect the peak time rate and may, consequently, overshadow the scalability of the underlying hierarchical network. Thus, every research focusing on simultaneous confidentiality and scalability should consider the trade-off between the two parameters. This trade-off will appear as an extra cost induced by the security-related constraints which should be tolerated by the hierarchical coded caching scheme.

In this paper, we first present an extended coded caching scheme which incorporates OTP confidentiality provisions and hierarchical network topology in the system model. Then, we analyze the extra cost induced by confidentiality via comparing the rate bounds to the case of non-secure hierarchical coded caching.

3. Secure Hierarchical Coded Caching

In this section, we present our secure hierarchical coded caching scheme and the related system model. Our system model needs to be defined in two aspects. We first introduce the topology and resources of the underlying network and then discuss the caching scheme which describes the transmissions in placement and delivery phases. Next, we discuss the security-related considerations.

In terms of topology, we adopt the 2-level hierarchical topology, described in [4]. In the top level of the hierarchy, the main server is connected to a hub via a shared link and then to mirror servers via separate links. In each cluster in the next hierarchy level, a shared link connects the mirror server to the hub while end users are connected to the same hub using separate links. We assume the number of the clusters to be equal to K_1 each of which connects K_2 end users.

As for the resources, the main server is assumed to store N files represented by W_1 through W_N each of size F bits. We assumed that the bits in a file are independent and uniformly distributed. The cache sizes in the mirrors and the end users are assumed to be M_1F and M_2F bits, respectively. The main and mirror servers are assumed to have unlimited processing power.

With respect to the caching scheme, we will assume the generalized caching scheme presented in [4] which is a combination of Scheme A and Scheme B. We follow the procedure to find the most efficient combination of the schemes.

During the delivery phase, each user makes exactly one demand. The local demands in each cluster are collected by the corresponding mirror server and then forwarded to the main server. The demand issued by $U_{(i,j)}$ is represented by the element $d_{i,j}$ in the demand matrix $\mathcal{D}$. According to the demands, the main server encodes the proper content along with with the orthogonal keys and transmits them within a file $X^{\mathcal{D}}$ of size $R_{S_1}F$ bits to all mirrors. Then, each mirror re-encodes (Scheme A) or forwards (Scheme B) the data requested by its corresponding users along with the related keys and transmits them within a file $Y^{\mathcal{D}}$ of size $R_{S_2}F$ bits.

Security-related constraints are considered in order to keep the transferred contents confidential from an external adversary assumed to have access to every shared link. In order to add confidentiality to our caching scheme, we adopt the security constraints proposed in [2,3]. Adopting the orthogonal key scheme proposed in [3], user caches, as well as mirror server memories are considered to be partitioned into Data and Key regions in order to keep space for storing the keys in the placement phase. Figure 1 shows the access model of adversary as well as the security-related configuration.

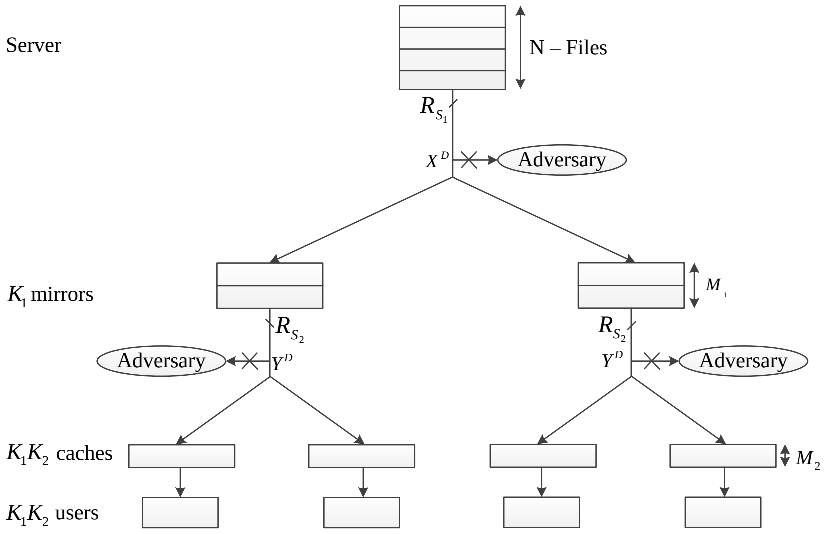

Figure 1. A hierarchical caching system with external adversaries acting overall shared links.

The mentioned security constraints guarantees that $I(X^{\mathcal{D}}; W_1, W_2, \ldots, W_N) = \epsilon_1$ and $I(Y^{\mathcal{D}}; W_1, W_2, \ldots, W_N) = \epsilon_2$, where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ which states that the external adversary cannot reveal any information regarding the files $W_1, W_2, \ldots, W_N$ by eavesdropping the shared links without access to users' and mirrors' caches. It is to be noted that $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ are for the case when file size is sufficiently large, i.e., when file size $\to \infty$. The minimum number of mirrors or users needed to be compromised in order to break the security was discussed in [3].

Adopting the security constraints from [3] requires some extra assumption regarding the placement phase in Scheme A. Since the users cannot establish an immediate communication with the main server, we assume a delegation between the main server and mirror servers in the placement phase in Scheme A. It means that the mirror servers are trusted and granted access to the keys because they need to decrypt and encrypt the contents again before and after re-encoding them.

Another assumption adopted from [3] in our system model is that every user is interested in no more than one file in the delivery phase and the demanded files are mutually different. The system cannot allocate resources, such as private links, network bandwidth, and cache space to a user with no demands in the delivery phase. Therefore, we suppose every user makes exactly one request in this phase. The latter assumptions obviously result in $N \geq K_1K_2$ as a criterion for the server to be able to answer all user requests. Again, we note that it is not reasonable to store files which will never be demanded. Thus, we assume that $N = K_1K_2$. Throughout, we assume that the placement phase is secure and links are error-free.

Let us represent the secure rate in the top hierarchy level by R_{S_1} and the second level secure rate by R_{S_2}. For a demand matrix $\mathcal{D}$ and for a large-enough file size F, a tuple $(M_1, M_2, R_{S_1}, R_{S_2})$ is said to be *feasible for* $\mathcal{D}$ if each user $U_{(i,j)}$ is able to recover its requested file $d_{i,j}$ securely with a probability arbitrarily close to unity. Moreover, $(M_1, M_2, R_{S_1}, R_{S_2})$ is *feasible* if it is feasible for all possible request matrices $\mathcal{D}$. Throughout, we assume feasible rate region in our analysis.

4. Fundamental Limits and Cost Analysis

The procedure we follow in our evaluations in this section can be described as follows. In order to maximize the secure achievable rate in the generalized scheme, we try to find the most effective combination of the Schemes A and B. To do this, we first parameterize the combination. We assume that a fraction equal to α of each file residing in the server (as well as transmissions in the top hierarchy level) are ruled by Scheme A and the rest $(1-\alpha)$ are transmitted on the basis of Scheme B. The corresponding fractions in the user cache (as well as transmissions in the second hierarchy level) are assumed to be equal to β and $1-\beta$, respectively. Then we try to find the best possible values for α and β which will result in the most effective combination. We denote the latter values by α^* and β^*. In the next step, we calculate the secure achievable rate for the generalized scheme via calculating the rates for both Schemes A and B and then combining the results together assigning the values α^* and β^* to α and β. We calculate the lower bounds of the rates through a similar procedure and then analyze the gap between the achievable rates and the rates specified by the lower bounds. A comparison between our results and those obtained in [4] highlights the cost of security in hierarchical network caching.

4.1. Preliminary Discussions

While analyzing the rates in each scheme, we separately consider each of the three regimes proposed in [4]. This makes it plausible to compare our results to those obtained in [4]. The mentioned regimes are characterized as shown in Equation (3) in terms of M_1 and M_2,

$$\begin{aligned}
1 M_1 + M_2K_2 &\geq N \text{ and } 0 \leq M_1 \leq N/4, \\
2 M_1 + M_2K_2 &< N, \\
3 M_1 + M_2K_2 &\geq N \text{ and } N/4 < M_1 \leq N. \qquad (3)
\end{aligned}$$

The results in [4], to which we compare our own results are as follows. The optimum values of α and β for the mentioned regimes in the non-secure hierarchical coded caching scheme are [4],

$$(\alpha^*, \beta^*) \triangleq \begin{cases} \left(\frac{M_1}{N}, \frac{M_1}{N}\right) & \text{in regime 1,} \\ \left(\frac{M_1}{M_1 + M_2K_2}, 0\right) & \text{in regime 2,} \\ \left(\frac{M_1}{N}, \frac{1}{4}\right) & \text{in regime 3.} \end{cases} \tag{4}$$

Moreover, the corresponding non-secure achievable rates for Scheme A and Scheme B have been calculated as functions of α^* and β^* in [4],

$$R_1(\alpha^*, \beta^*) \approx \begin{cases} \min\left\{\frac{K_1K_2}{1}, \frac{N}{M_2}\right\} & \text{in regime 1,} \\ \min\left\{K_1K_2, \frac{M_1}{M_1 + M_2K_2} \cdot \frac{(N - M_1)K_2}{M_1 + M_2K_2}, + \frac{M_1}{M_1 + M_2K_2} \cdot \frac{NK_2 - M_1}{M_1 + M_2K_2}\right\} & \text{in regime 2,} \\ \frac{(N - M_1)^2}{NM_2} & \text{in regime 3,} \end{cases} \tag{5}$$

$$R_2(\alpha^*, \beta^*) \approx \min\left\{K_2, \frac{N}{M_2}\right\}. \tag{6}$$

See Figure 2 for different regimes of M_1, M_2 for α^* and β^*. In (4) and (5), the approximation is within a constant additive and multiplicative as given by (7) and (8),

$$R_1 \quad \geq \quad R_1^{lb}(M_1, M_2) \geq \frac{1}{60} R_1(\alpha^*, \beta^*) - 4, \tag{7}$$

$$R_2 \quad \geq \quad R_2^{lb}(M_1, M_2) \geq \frac{1}{36} R_1(\alpha^*, \beta^*) - 16. \tag{8}$$

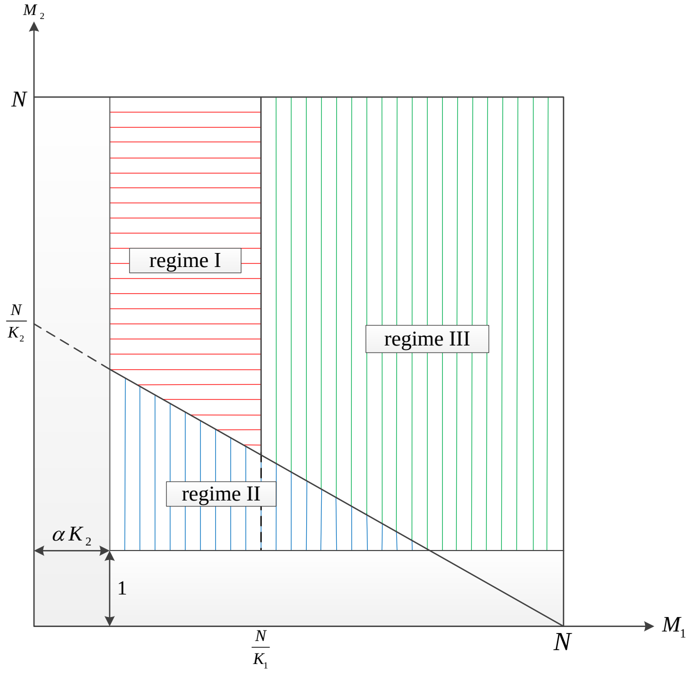

Figure 2. Different regimes of M_1, M_2 for α^* and β^*.

4.2. Secure Achievable Rates

Before beginning to develop our mathematical modelings, let us introduce some notations we will use in our equations. We will refer to the jth user ($j \in [1, 2, \ldots, K_2]$) in the ith cluster where $i \in [1, 2, \ldots, K_1]$ as $U_{(i,j)}$ and refer to the corresponding cache as $C_{(i,j)}$. Let us represent the coded content items transmitted in the first and second levels of the hierarchy by $X^{\mathcal{D}}$ and $Y^{\mathcal{D}}$, respectively, where $\mathcal{D}$ is the request matrix. Furthermore, let us represent the secure rate in the top hierarchy level by R_{S_1} and the second level secure rate by R_{S_2}.

Now, let us begin the derivation of our model by calculating R_{S_1} and R_{S_2} for scheme A. For N files and K_1 mirrors each with a cache size of $M_1 \in \frac{N-K_2}{K_1} \cdot t_1 + K_2$, where $t_1 \in \{0, 1, 2, \cdots, K_1\}$, R_{S_1} for scheme A is given by

$$R_{S_1}^{A} = K_2 \cdot r\left(\frac{M_1 - K_2}{N - K_2}, K_1\right), \tag{9}$$

where $r(.,.)$ is defined as:

$$r\left(\frac{M}{N}, K\right) \triangleq \left[K \cdot \left(1 - \frac{M}{N}\right) \cdot \frac{N}{KM}\left(1 - \left(1 - \frac{M}{N}\right)^{K}\right)\right]^{+}, \tag{10}$$

with $[x]^{+} \triangleq \max\{x, 0\}$. Moreover, R_{S_2} for Scheme A considering K_2 users each with a cache size of $M_2 \in \frac{N-1}{K_2} \cdot t_2 + 1$, where $t_2 \in \{0, 1, \ldots, K_2\}$, can be obtained from

$$R_{S_2}^{A} = r\left(\frac{M_2 - 1}{N - 1}, K_2\right). \tag{11}$$

Similarly, R_{S_1} and R_{S_2} for the scheme B can be calculated as

$$R_{S_1}^{B} = r\left(\frac{M_2 - 1}{N - 1}, K_1 K_2\right), \tag{12}$$

$$R_{S_2}^{B} = r\left(\frac{M_2 - 1}{N - 1}, K_2\right). \tag{13}$$

Let us normalize the total file size, mirror memory size, and user cache size involved by scheme A as shown in (14) and (15), respectively,

$$F' \triangleq \alpha F,$$

$$M_1' \triangleq \frac{M_1 F}{F'} = \frac{M_1}{\alpha}, \tag{14}$$

$$M_2' \triangleq \frac{\beta M_2 F}{F'} = \frac{\beta M_2}{\alpha}. \tag{15}$$

Moreover, let us normalize user cache size involved by scheme B as shown in (16),

$$F'' \triangleq (1 - \alpha) F,$$

$$M_2'' \triangleq \frac{(1 - \beta) M_2 F}{F''} = \frac{(1 - \beta) M_2}{1 - \alpha}. \tag{16}$$

Thus, the secure rates induced by scheme A and scheme B can be normalized with respect to the file size as given by

$$R_{S_1}^{A'} = \alpha K_2 \cdot r\left(\frac{M_1' - K_2}{(N-K_2)}, K_1\right) = \alpha K_2 \cdot r\left(\frac{M_1 - \alpha K_2}{\alpha(N-K_2)}, K_1\right), \tag{17a}$$

$$R_{S_2}^{A'} = \alpha \cdot r\left(\frac{M_2' - 1}{(N-1)}, K_2\right) = \alpha \cdot r\left(\frac{\beta M_2 - \alpha}{\alpha(N-1)}, K_2\right), \tag{17b}$$

and

$$R_{S_1}^{B'} = (1-\alpha) \cdot r\left(\frac{M_2'' - 1}{N-1}, K_1 K_2\right) = (1-\alpha) \cdot r\left(\frac{(1-\beta)M_2 - (1-\alpha)}{(1-\alpha)(N-1)}, K_1 K_2\right), \tag{18a}$$

$$R_{S_2}^{B'} = (1-\alpha) \cdot r\left(\frac{M_2'' - 1}{N-1}, K_2\right) = (1-\alpha) \cdot r\left(\frac{(1-\beta)M_2 - (1-\alpha)}{(1-\alpha)(N-1)}, K_2\right). \tag{18b}$$

In the next step, we will calculate α^* and β^* for each of the regimes in a way that both $R_{S_1}(\alpha,\beta)$ and $R_{S_2}(\alpha,\beta)$ can be minimized. Let us begin with regime **1**. According to (17b) and (18b), for $\alpha = \beta$ it holds that $R_{S_2}(\alpha,\alpha) = r((M_2-1)/(N-1), K_2)$. It can be verified that $\alpha = M/N$ results in a near-optimal value for $R_{S_1}(\alpha,\alpha)$ in Regime A. Thus, we chose $\alpha^* = M_1/N$ and $\beta^* = M_1/N$ in this regime. Choosing $\alpha^* = M_1/N$ allows each mirror to store the first part of each of the N files in the first transmission. Thus, there will be no need for further transmission between the server and the mirrors in the placement phase or key memory space in the mirrors.

Now let us proceed with regime **2**. In this regime, it can be verified from Equation (20b) that $M_2 < N/K_2$ which means that the M_2 cache area is very small. Thus, $R_{S_2}(\alpha,\beta)$ will be of order K_2 for any choice of α and β. Therefore, we only need to choose α and β in a way that $R_{S_1}(\alpha,\beta)$ is minimized. In this regime, the optimized values for α and β can be obtained as $\alpha^* = M_1/(M_1 + M_2K_2)$ and $\beta^* = M_1/(M_2(M_1 + M_2K_2))$.

In regime **3** (like in regime **1**), a choice of $\alpha = \beta = M_1/N$ is preferable. However, it should be considered that a large value of β leads to an unacceptably-large value of $R_{S_1}(\alpha,\beta)$. Thus, a minimum threshold of $\beta^* = M_1/M_2N$ should be considered. Similar to the case of regime **1**, no extra transmission between the server and the mirrors in the placement phase or key area in the cache is required in this regime.

After deciding the proper choice of α^*, β^*, let us calculate $R_{S_1}(\alpha^*,\beta^*)$ and $R_{S_2}(\alpha^*,\beta^*)$ for the generalized scheme as a combination of the secure rates in the two schemes A and B.

Theorem 1. *We have the following conditions on* $R_{S_1}(\alpha^*,\beta^*)$ *and* $R_{S_2}(\alpha^*,\beta^*)$

$$R_{S_1}(\alpha^*,\beta^*) \approx \begin{cases} \min\left\{K_1K_2, \dfrac{N-1}{M_2-1}\right\} & \text{in regime } \mathbf{1}, \\ \min\left\{K_1K_2, \dfrac{M_1}{M_1+M_2K_2} \cdot \dfrac{K_2(N-M_1)}{M_1+(M_2-1)K_2} \right. & \\ \qquad \left. + \dfrac{M_2K_2}{M_1+M_2K_2} \cdot \dfrac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}\right\} & \text{in regime } \mathbf{2}, \\ \dfrac{(N-M_1)^2}{N(M_2-1)} & \text{in regime } \mathbf{3}, \end{cases} \tag{19a}$$

and

$$R_{S_2}(\alpha^*,\beta^*) \leq K_1 \cdot \min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \tag{19b}$$

Proof. The normalized achievable secure rates for the generalized scheme can be calculated in the form of functions of α and β,

$$\begin{aligned} R_{S_1}(\alpha,\beta) &\triangleq R_{S_1}^{A'} + R_{S_1}^{B'}, \\ &= \alpha K_2 \cdot r\left(\frac{M_1 - \alpha K_2}{\alpha(N-K_2)}, K_1\right) + (1-\alpha)\cdot r\left(\frac{(1-\beta)M_2 - (1-\alpha)}{(1-\alpha)(N-1)}, K_1K_2\right), \end{aligned} \tag{20a}$$

and

$$\begin{aligned} R_{S_2}(\alpha,\beta) &\triangleq R_{S_2}^{A'} + R_{S_2}^{B'}, \\ &= \alpha \cdot r\left(\frac{\beta M_2 - \alpha}{\alpha(N-1)}, K_2\right) + (1-\alpha)\cdot r\left(\frac{(1-\beta)M_2 - (1-\alpha)}{(1-\alpha)(N-1)}, K_2\right). \end{aligned} \tag{20b}$$

With the proper choice of α^*, β^*, we proceed to calculate secure achievable rates $R_{S_1}(\alpha^*,\beta^*)$, $R_{S_2}(\alpha^*,\beta^*)$. As we observe, the secure achievable rates for the generalized caching scheme is a function of $r(.,.)$, as mentioned in the Equation (1). We observe the following,

$$r\left(\frac{M}{N}, K\right) \leq \begin{cases} \min\left\{K, \frac{N}{M} - 1\right\} & M \leq N, \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Now let us proceed with calculating the secure achievable rates for each of the regimes of M_1 and M_2 beginning with regime **1**. According to (5), (20), and (21), the secure achievable rates in this regime can be upper bounded as shown in inequalities (22a) and (22b),

$$\begin{aligned} R_{S_1}(\alpha^*,\beta^*) &= \frac{M_1}{N} K_2 \cdot r(1, K_1) + \left(1 - \frac{M_1}{N}\right) \cdot r\left(\frac{M_2 - 1}{N-1}, K_1K_2\right) \\ &\leq 0 + \min\left\{K_1K_2 \frac{N-1}{M_2 - 1}\right\} \\ &= \min\left\{K_1K_2, \frac{N-1}{M_2-1}\right\}, \end{aligned} \tag{22a}$$

and

$$\begin{aligned} R_{S_2}(\alpha^*,\beta^*) &= \frac{M_1}{N} \cdot \left(\frac{M_2 - 1}{N-1}\right) + \left(1 - \frac{M_1}{N}\right) \cdot r\left(\frac{M_2 - 1}{N-1}, K_2\right) \\ &= r\left(\frac{M_2 - 1}{N-1}\right) \\ &\leq \min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \end{aligned} \tag{22b}$$

Through a similar reasoning, the upper bounds to the secure achievable rate $R_{S_1}(\alpha^*,\beta^*)$ in regime **2** can be obtained from the inequality (23a) (the form of equations are different from regime **1**).

$$\begin{aligned}
R_{S_1}(\alpha^*,\beta^*) &= \frac{M_1K_2}{M_1+M_2K_2}\cdot r\left(\frac{M_1+M_2K_2-K_2}{N-K_2}, K_1\right) \\
&\quad +\frac{M_2K_2}{M_1+M_2K_2}\cdot r\left(\frac{(M_2-1)(M_1+M_2K_2)}{(N-1)M_2K_2}, K_1K_2\right) \\
&\leq \frac{M_1K_2}{M_1+K_2M_2}\cdot \min\left\{K_1, \frac{N-K_2}{M_1+M_2K_2-K_2}-1\right\} \\
&\quad +\frac{M_2K_2}{M_1+M_2K_2}\cdot \min\left\{K_1K_2, \frac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}-1\right\} \\
&\leq \frac{M_1}{M_1+M_2K_2}\cdot \min\left\{K_1K_2, \frac{K_2(N-M_1)}{M_1+(M_2-1)K_2}\right\} \\
&\quad +\frac{M_2K_2}{M_1+M_2K_2}\cdot \min\left\{K_1K_2, \frac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}\right\} \\
&\leq \min\left\{K_1K_2, \frac{M_1}{M_1+M_2K_2}\cdot\frac{K_2(N-M_1)}{M_1+(M_2-1)K_2}\right. \\
&\quad \left.+\frac{M_2K_2}{M_1+M_2K_2}\cdot\frac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}\right\}. \qquad (23a)
\end{aligned}$$

Furthermore, according to (20) and (21), the reader can easily verify that $R_{S_2}(\alpha^*,\beta^*)$ in regime **2** will be upper bounded by

$$R_{S_2}(\alpha^*,\beta^*) \leq K_2 = \min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \qquad (23b)$$

Additionally, for regime **3** we have,

$$\begin{aligned}
R_{S_1}(\alpha^*,\beta^*) &= \left(1-\frac{M_1}{N}\right)\cdot r\left(1-\frac{M_1}{N}\right) + r\left(\frac{(1-1/K_1)M_2-(1-M_1/N)}{(1-M_1/N)(N-1)}, K_2\right) \\
&\leq 0+\left(1-\frac{M_1}{N}\right)\cdot \min\left\{K_1K_2, \frac{(N-M_1)(N-1)K_1}{NM_2(K_1-1)-K_1(N-M_1)}\right\} \\
&\leq \left(\frac{N-M_1}{N}\right)\cdot\frac{K_1(N-M_1)(N-1)}{(K_1-1)N(M_2-1)} \\
&\leq \frac{K_1}{K_1-1}\cdot\frac{(N-M_1)^2}{N(M_2-1)}, \qquad (24a)
\end{aligned}$$

and

$$\begin{aligned}
R_{S_2}(\alpha^*,\beta^*) &= \frac{M_1}{N}\cdot r\left(\frac{M_2}{K_1M_1}, K_2\right) + \left(1-\frac{M_1}{N}\right)\cdot r\left(\frac{(1-1/K_1)M_2-(1-M_1/N)}{(1-M_1/N)(N-1)}, K_2\right) \\
&\leq \frac{M_1}{N}\cdot \min\left\{K_2, \frac{K_1(N-1)}{M_2-1}\right\} + \left(1-\frac{M_1}{N}\right)\min\left\{K_2, \frac{K_1(N-1)}{M_2-1}\right\} \\
&\leq \min\left\{K_2, \frac{K_1(N-1)}{(M_2-1)}\right\} \\
&\leq K_1\cdot\min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \qquad (24b)
\end{aligned}$$

Summarizing our results for the generalized scheme from the discussions above we have

$$R_{S_1}(\alpha^*,\beta^*) \approx \begin{cases}
\min\left\{K_1K_2, \dfrac{N-1}{M_2-1}\right\} & \text{in regime } \mathbf{1}, \\
\min\left\{K_1K_2, \dfrac{M_1}{M_1+M_2K_2}\cdot\dfrac{K_2(N-M_1)}{M_1+(M_2-1)K_2}\right. & \\
\qquad \left. +\dfrac{M_2K_2}{M_1+M_2K_2}\cdot\dfrac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}\right\} & \text{in regime } \mathbf{2}, \\
\dfrac{(N-M_1)^2}{N(M_2-1)} & \text{in regime } \mathbf{3},
\end{cases} \qquad (25a)$$

and

$$R_{S_2}(\alpha^*, \beta^*) \leq K_1 \cdot \min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \tag{25b}$$

The proof is now complete. □

4.3. Memory Requirements

An information-theoretical analysis will reveal some minimum requirements regarding the cache size in the mirrors and uses cache. Consider a caching system in which two files denoted by A and B are residing in the main server $N = 2$. Consider $K_1 = 2$ mirrors denoted by m_1 and m_2 with each of mirrors cache size M_1. Cache contents denoted by Z_1^m, Z_2^m cached by users in the placement phase. Let us assume that mirror m_1 demands a content item A_α in the delivery phase which is part of file A and mirror m_2 demands part of file B denoted by B_α. Both demanded items are assumed to be of size αF which can be considered a fraction equal to α of the size of file A or B. The mentioned demands can be represented by a demand vector $(d_1, d_2) = (A_\alpha, B_\alpha)$.

In this setup, the main server will transmit $X_{(A_\alpha, B_\alpha)}$ to the mirrors which should be capable of regenerating the items A_α and B_α when combined with Z_1^m and Z_2^m. From an information theoretical point of view, the criterion stated by inequality (26) should hold to make it possible to achieve this goal,

$$H(A_\alpha, B_\alpha | X_{(A_\alpha, B_\alpha)}, Z_1^m, Z_2^m) \leq \epsilon. \tag{26}$$

For the security constraint between the server and the mirrors, inequality (27) should hold in order to keep the delivery phase transmissions confidential,

$$I(A_\alpha, B_\alpha; X_{(A_\alpha, B_\alpha)}) \leq \delta. \tag{27}$$

Using (26) and (27) we have

$$\begin{aligned}
2\alpha F &\leq H(A_\alpha, B_\alpha), \\
&= I(A_\alpha, B_\alpha; X_{(A_\alpha, B_\alpha)}, Z_1^m, Z_2^m) + H(A_\alpha, B_\alpha | X_{(A_\alpha, B_\alpha)}, Z_1^m, Z_2^m), \\
&\leq I(A_\alpha, B_\alpha; X_{(A_\alpha, B_\alpha)}, Z_1^m, Z_2^m) + \epsilon, \\
&= I(A_\alpha, B_\alpha; X_{(A_\alpha, B_\alpha)}) + I(A_\alpha, B_\alpha; Z_1^m, Z_2^m | X_{(A_\alpha, B_\alpha)}) + \epsilon, \\
&\leq I(A_\alpha, B_\alpha; Z_1^m, Z_2^m | X_{(A_\alpha, B_\alpha)}) + \delta + \epsilon, \\
&\leq H(Z_1^m, Z_2^m | X_{(A_\alpha, B_\alpha)}) + \delta + \epsilon, \\
&\leq 2M_1 F + \delta + \epsilon.
\end{aligned} \tag{28}$$

From (28) we immediately obtain

$$M_1 \geq \alpha - \frac{\delta}{F} - \frac{\epsilon}{F}. \tag{29}$$

When δ and ϵ approach zero, inequality (29) will be converted to $M_1 \geq \alpha$. Again, we note that users should be able to recover both file A and file B from a single cached item Z_1^m along with two items within $X_{(A_\alpha, B_\alpha)}$ and $X_{(B_\alpha, A_\alpha)}$ transmitted in response to the demand vectors $(d_1, d_2) = (A_\alpha, B_\alpha)$ and $(d_1, d_2) = (B_\alpha, A_\alpha)$, respectively. The latter requirement leads to the following inequalities,

$$H(A_\alpha, B_\alpha | X_{(A_\alpha, B_\alpha)}, X_{(B_\alpha, A_\alpha)}, Z_1^m) \leq \epsilon, \tag{30}$$

$$I(A_\alpha, B_\alpha; X_{(A_\alpha, B_\alpha)}) \leq \delta. \tag{31}$$

Through similar reasoning the latter two inequalities will lead to $R_s^* + M_1 \geq 2\alpha$ where R_s^* is minimum rate between the server and the mirrors.

4.4. Lower Bounds

Now let us discuss the lower bounds on secure rates R_{S_1} and R_{S_2} for different values of M_1, M_2 given the feasibility of $(M_1, M_2, R_{S_1}, R_{S_2})$. To do this, we follow the approach taken in [3] for the secure non-hierarchical scheme and extend the discussions and results to the case of the secure hierarchical network.

Theorem 2. *We have*

$$\begin{aligned} R_{S_1} &\geq \max_{\substack{s_1\in\{1,2,\ldots,K_1\} \\ s_2\in\{1,2,\ldots,K_2\}}} max\left\{ s_1 s_2\left[1 - \frac{1}{N - 2s_1 s_2}\Big(s_1 M_1 + s_1 s_2 (M_2 - 1)\Big)\right], \frac{s_1 s_2 (N - s_1 M_1 - s_1 s_2 M_2)}{N} \right\} \\ &\triangleq R^{lb}_{S_1}(M_1, M_2), \end{aligned}$$

and

$$R_{S_2} \geq \max_{t\in\{1,2,\ldots,K_2\}} \frac{t(N - tM_2)}{N}.$$

Proof. Let us begin with the lower bound on R_{S_1}. For $s_1 \in \{1,2,\ldots,K_1\}$ and $s_2 \in \{1,2,\ldots,K_2\}$, suppose the first s_1 mirrors store $Z^m_1, Z^m_2, \ldots, Z^m_{s_1}$. Furthermore, assume that for $i \in \{1,2,\ldots,s_1\}$ and $j \in \{1,2,\ldots,s_2\}$, every user $C_{(i,j)}$ caches $Z^u_{i,1}, Z^u_{i,2}, \ldots, Z^u_{i,s_2}$. Suppose the mentioned users issue the demand matrix $\mathcal{D}^1$ defined as $d^1_{i,j} = (i-1)s_2 + j$ which includes requests for the first $s_1 s_2$ files residing in the main server. The items transmitted by the main server within $X_1 = X_{(d_{1,1},\ldots,d_{s_1 s_2})}$, along with the mirrored items $Z^m_1, Z^m_2, \ldots, Z^m_{s_1}$ and cached items $Z^u_{i,1}, Z^u_{i,2}, \ldots, Z^u_{i,s_2}$ must able to decode the files $W_1, W_2, \ldots, W_{s_1 s_2}$.

Similarly, for the different request matrix $\mathcal{D}$, where user $U_{(i,j)}$ demands $d_{i,j} = s_1 s_2 + (i-1)s_2 + j$, i.e., requesting next $s_1 s_2$ files from the server. The transmission X_2, along with mirrors $Z^m_{i,1}, Z^m_{i,2}, \ldots, Z^m_{i,s_2}$ and users cache $Z^u_{i,1}, Z^u_{i,2}, \ldots, Z^u_{i,s_2}$ must be able to decode the files $W_{s_1 s_2+1}, W_{s_1 s_2+2}, \ldots, W_{2 s_1 s_2}$. Likewise, considering all $\lfloor N/s_1 s_2 \rfloor$ request matrices, multicast transmission $X_1, \ldots, X_{\lfloor N/s_1 s_2 \rfloor}$ along with mirrors $Z^m_1, Z^m_2, \ldots Z^m_{s_1}$ and users cache $Z^u_{i,1}, Z^u_{i,2}, \ldots, Z^u_{i,s_2}$, must be able to recover the files $W_1, \ldots, W_{s_1 s_2 \lfloor N/s_1 s_2 \rfloor}$. Let

$$\begin{aligned} \widetilde{W} &= \{W_1, \ldots, W_{s_1 s_2 \lfloor N/s_1 s_2 \rfloor}\} \\ \widetilde{X} &= \{X_1, \ldots, X_{\lfloor N/s_1 s_2 \rfloor}\} \\ \widetilde{X}\backslash\{l\} &= \{X_1, \ldots, X_{l-1}, X_{l+1}, \ldots, X_{\lfloor N/s_1 s_2 \rfloor}\} \\ \widetilde{Z}^m &= \{Z^m_1, \ldots, Z^m_{s_1}\} = \widetilde{Z}^m_i \\ \widetilde{Z}^u &= \{Z^u_{1,1}, \ldots, Z^u_{1,s_2}, Z^u_{2,1}, \ldots, Z^u_{s_1 s_2}\} = \{Z^u_{i,j}\}. \end{aligned}$$

Another point implied by the feasibility of $(M_1, M_2, R_{s_1}, R_{s_2})$ in our system model is that the external adversary should not be able to retrieve any information regarding the contents being transmitted in the delivery phase. This criterion is formally described by inequalities (32) and (33),

$$H(\widetilde{W}|\widetilde{X}, \widetilde{Z}^m, \widetilde{Z}^u) \leq \epsilon_1, \tag{32}$$

and

$$I(\widetilde{W}; X_l) \leq \epsilon_2,\ l = 1, \ldots, q. \tag{33}$$

Consider the information flow consisting of multicast transmission $X_1, \ldots, X_{\lfloor N/s_1 s_2 \rfloor}$, mirrors $Z_1, Z_2, \ldots, Z_{s_1}$ and users cache $Z_{i,1} \ldots, Z_{i,s_2}$ for decoding file $W_1, \ldots, W_{s_1 s_2 \lfloor N/s_1 s_2 \rfloor}$. We have

$$\begin{aligned}
s_1s_2\lfloor N/s_1s_2\rfloor F &\leq H(\widetilde{W})\\
&= I(\widetilde{W};\widetilde{X},\widetilde{Z}^u,\widetilde{Z}^m)+H(\widetilde{W}|\widetilde{X},\widetilde{Z}^u,\widetilde{Z}^m)\\
&\leq I(\widetilde{W};\widetilde{X},\widetilde{Z}^u,\widetilde{Z}^m)+\epsilon_1\\
&= I(\widetilde{W};X_l)+I(\widetilde{W};\widetilde{X}\backslash\{l\},\widetilde{Z}^m,\widetilde{Z}^u|X_l)+\epsilon_1\\
&\leq I(\widetilde{W};\widetilde{X}\backslash\{l\},\widetilde{Z}^m,\widetilde{Z}^u|X_l)+\epsilon_1+\epsilon_2\\
&\leq H(\widetilde{X}\backslash\{l\},\widetilde{Z}^m,Z^u)+\epsilon\\
&\leq \sum_{k=1,k\neq l}^{\lfloor N/s_1s_2\rfloor} H(X_k)+\sum_{i=1}^{s_1} H(Z_i^m)+\sum_{i=1}^{s_1}\sum_{j=1}^{s_2} H(Z_{i,j}^u)+\epsilon\\
&\leq (\lfloor N/s_1s_2\rfloor-1)R_{S_1}F+s_1M_1F+s_1s_2M_2F+\epsilon. \qquad (34)
\end{aligned}$$

So,

$$s_1s_2\lfloor N/s_1s_2\rfloor \leq (\lfloor N/s_1s_2\rfloor-1)R_{S_1}+s_1M_1+s_1s_2M_2+\frac{\epsilon}{F}. \qquad (35)$$

Solving and optimizing for all possible values of s_1 and s_2 we obtain

$$\begin{aligned}
R_{S_1} &\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} \lim_{\epsilon\to 0}\frac{1}{\lfloor N/s_1s_2\rfloor-1}\left\{s_1s_2\lfloor N/s_1s_2\rfloor-s_1M_1-s_1s_2M_2-\frac{\epsilon}{F}\right\}\\
&\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} s_1s_2-\frac{s_1M_1+s_1s_2(M_2-1)}{N/s_1s_2-2}\\
&\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} s_1s_2\left(1-\frac{s_1M_1+s_1s_2(M_2-1)}{N-2s_1s_2}\right). \qquad (36)
\end{aligned}$$

We can obtain an alternate lower bound by using $\lceil N/s_1s_2\rceil$ transmissions instead of $\lfloor N/s_1s_2\rfloor$ in (35),

$$\begin{aligned}
R_{S_1} &\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} \frac{1}{\lceil N/s_1s_2\rceil-1}(N-s_1M_1-s_1s_2M_2)\\
&\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} \frac{1}{N/s_1s_2}(N-s_1M_1-s_1s_2M_2)\\
&\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} \frac{1}{N}(s_1s_2(N-s_1M_1-s_1s_2M_2)). \qquad (37)
\end{aligned}$$

The inequalities (36) and (37) hold for any value of $s_1\in\{1,2,\dots,K_1\}$ and $s_2\in\{1,2,\dots,K_2\}$. So, we obtain the following lower bound on R_{S_1} for the tuple $(M_1,M_2,R_{S_1},R_{S_2})$ to be feasible,

$$\begin{aligned}
R_{S_1} &\geq \max_{\substack{s_1\in\{1,2,\dots,K_1\}\\ s_2\in\{1,2,\dots,K_2\}}} \max\left\{s_1s_2\left[1-\frac{1}{N-2s_1s_2}\Big(s_1M_1+s_1s_2(M_2-1)\Big)\right],\frac{s_1s_2(N-s_1M_1-s_1s_2M_2)}{N}\right\}\\
&\triangleq R_{S_1}^{lb}(M_1,M_2). \qquad (38)
\end{aligned}$$

After calculating the lower bound for R_{S_1}, let us proceed with that of R_{S_2} assuming the feasibility of $(M_1,M_2,R_{S_1},R_{S_2})$. Let $t\in\{1,2,\dots,K_2\}$. Consider the t users cache $U_{(1,j)}$ as $Z_{1,1}^u,Z_{1,2}^u,\dots,Z_{1,t}^u$ with $j\in\{1,2,\dots,t\}$. Consider the request matrix $\mathcal{D}$ with demands $d_{1,j}=j$, i.e., requesting t files from the server. The transmission $Y_1=Y_{(d_{1,1},\dots,d_{1,t})}$, along with the users cache $U_{(1,j)}$ $Z_{1,1}^u,Z_{1,2}^u,\dots,Z_{1,t}^u$ must be able to decode the files $W_1,\dots,W_t$. Similarly, for the different request matrix $\mathcal{D}$, where the user demands $d_{i,j}=t+j$, i.e., requesting another t files from the server. The transmission Y_2 along with the users cache $U_{1,j}$, must be able to decode the files $W_{t+1},\dots,W_{2t}$. Likewise, considering the all $\lceil N/t \rceil$ request matrices, multicast transmission, $Y_1, \ldots, Y_{\lceil N/t \rceil}$ along with the users cache $Z^u_{1,1}, Z^u_{1,2}, \ldots, Z^u_{1,t}$ must be able to recover the files $W_1, \ldots, W_N$. Assuming Y_l be the information leaked to the external adversary through the link connecting between the mirror and its corresponding users.

$$\begin{aligned}
\widetilde{W} &= \{W_1, \ldots, W_N\} \\
\widetilde{Y} &= \{Y_1, \ldots, Y_{\lceil N/t \rceil}\} \\
\widetilde{Y}\backslash\{l\} &= \{Y_1, \ldots, Y_{l-1}, Y_{l+1}, \ldots, Y_{\lceil N/t \rceil}\} \\
\widetilde{Z}^u &= \{Z^u_{1,1}, \ldots, Z^u_{1,t}\} = \{Z^u_{1,j}\}, \text{ where } j \in \{1, 2, \ldots, t\}.
\end{aligned}$$

The file recovery and security constraints can be stated as

$$\begin{aligned}
H(\widetilde{W}|\widetilde{Y}, \widetilde{Z}^u) &\leq \epsilon_1, && (39) \\
I(\widetilde{W}; Y_l) &\leq \epsilon_2, \quad l = 1, \ldots, Y_{\lceil N/t \rceil} && (40)
\end{aligned}$$

This is similar case of single layer secure scheme. Consider the information flow consisting of multicast transmission $Y_1, \ldots, Y_{\lceil N/t \rceil}$ and users cache $Z^u_{1,1}, Z^u_{1,2}, \ldots, Z^u_{1,t}$ for decoding the files $W_1, W_2, \ldots, W_N$. We have

$$\begin{aligned}
NF &= H(\widetilde{W}) \\
&= I(\widetilde{W}; \widetilde{Y}, \widetilde{Z}^u) + (\widetilde{W}|\widetilde{Y}, \widetilde{Z}^u) \\
&\leq I(\widetilde{W}; \widetilde{Y}, \widetilde{Z}^u) + \epsilon_1 \\
&= I(\widetilde{W}; Y_l) + I(\widetilde{W}; \widetilde{Y}\backslash\{l\}, \widetilde{Z}^u|Y_l) + \epsilon_1 \\
&\leq I(\widetilde{W}; \widetilde{Y}\backslash\{l\}, \widetilde{Z}^u|Y_l) + \epsilon_1 + \epsilon_2 \\
&\leq H(\widetilde{Y}\backslash\{l\}, \widetilde{Z}^u) + \epsilon, \text{ where } \epsilon_1 + \epsilon_2 = \epsilon \\
&\leq \sum_{i=1, i\neq l}^{\lceil N/t \rceil} H(Y_i) + \sum_{j=1}^{t} H(Z^u_{1,j}) + \epsilon \\
&\leq (\lceil N/t \rceil - 1) R_{S_2} F + t M_2 F + \epsilon. && (41)
\end{aligned}$$

Therefore,

$$N = (\lceil N/t \rceil - 1) R_{S_2} + t M_2 + \frac{\epsilon}{F}. \qquad (42)$$

Solving and optimizing for all value of t, we obtain the following lower bound

$$\begin{aligned}
R_{S_2} &\geq \max_{t \in \{1, 2, \ldots, K_2\}} \lim_{\epsilon \to 0} \frac{N - t M_2 - \frac{2\epsilon}{F}}{\lceil N/t \rceil - 1} \\
&\geq \max_{t \in \{1, 2, \ldots, K_2\}} \frac{N - t M_2}{N/t} \\
&= \max_{t \in \{1, 2, \ldots, K_2\}} \frac{t(N - t M_2)}{N} && (43) \\
&\triangleq R^{lb}_{S_2}(M_1, M_2).
\end{aligned}$$

□

5. Gap Analysis

In this section, we analyze the gap between the secure achievable rates and the corresponding lower bounds.

5.1. $R_{S_1}(\alpha^*, \beta^*)$ against $R^{lb}_{S_1}(M_1, M_2)$

Theorem 3. *$R_{S_1}(\alpha^*, \beta^*)$ will vary between a constant multiplicative and a constant additive gap with $R^{lb}_{S_1}(M_1, M_2)$. Specifically,*

$$R_{S_1} \geq R^{lb}_{S_1}(M_1, M_2) \geq \frac{1}{96} R_{S_1}(\alpha^*, \beta^*) - 4.$$

Proof. The values of α^* and β^*, and consequently $R_{s_1}(\alpha^*, \beta^*)$ obviously depend on the regime characterized by M_1 and M_2. This makes it necessary to examine each of the regimes. We begin with regime 1 assuming that $N \geq K_1K_2$, $K_1 \geq 4$ and $K_2 \geq 4$.

Regime 1: $M_1 + M_2K_2 \geq N$ and $0 \leq M_1 \leq \frac{N}{K_1}$ where $M_1 \geq \frac{M_1K_2}{N}$, $M_2 \geq 1$. Inequalities (25a) and (38) give the achievable secure rate $R_{S_1}(\alpha^*, \beta^*)$, as well as the lower bound on $R_{S_1}(M_1, M_2)$ for regime 1.

In order to make the margin of the gap more manageable, we further divide our discussions regarding this regime into three sub-regimes specified as follows:

$$\begin{array}{ll}
\text{1.A)} & \dfrac{M_1K_2}{N} \leq M_1 \leq \dfrac{N}{2K_1}, \quad \dfrac{3N}{4K_2} \leq M_2 \leq \dfrac{N}{4}, \\
\text{1.B)} & \dfrac{N}{2K_1} \leq M_1 \leq \dfrac{N}{K_1}, \quad \dfrac{3N}{4K_2} \leq M_2 \leq \dfrac{N}{4}, \\
\text{1.C)} & \dfrac{M_1K_2}{N} \leq M_1 \leq \dfrac{N}{K_1}, \quad \dfrac{N}{4} \leq M_2 \leq N.
\end{array}$$

For sub-regime 1.A), let us feed $s_1 = 1$ and $s_2 = \left\lfloor \frac{N}{2M_2} \right\rfloor$ (which is a valid choice because $\lfloor z \rfloor \geq z/2$ for any $z \geq 1$) into (38) which gives

$$\begin{aligned}
R^{lb}_{S_1}(M_1, M_2) &\geq \frac{\left\lfloor \frac{N}{2M_2} \right\rfloor \left(N - M_1 - \left\lfloor \frac{N}{2M_2} \right\rfloor M_2 \right)}{N} \\
&\overset{(a)}{\geq} \frac{1}{N} \left[\frac{N}{4M_2} \left(N - \frac{N}{2K_1} - \frac{N}{2M_2} \cdot M_2 \right) \right] \\
&\geq \frac{N}{4M_2} \left(1 - \frac{1}{2K_1} - \frac{1}{2} \right) \\
&\overset{(b)}{\geq} \frac{N}{4M_2} \left(\frac{1}{2} - \frac{1}{8} \right) \\
&\overset{(c)}{\geq} \frac{3N}{32M_2} \geq \frac{3}{32} \cdot \frac{4}{5} \cdot \frac{N-1}{M_2 - 1}, \\
&\geq \frac{3}{40} \min \left\{ K_1K_2, \frac{N-1}{M_2 - 1} \right\}.
\end{aligned} \tag{44}$$

In deriving (44), we have used $(a) : \lfloor z \rfloor \geq z/2 \; \forall z \geq 1$, $(b) : K_1 \geq 4$ and $(c) : N \geq K_1K_2$. Combining (44) and (22a), we obtain

$$R^{lb}_{S_1}(M_1, M_2) \geq \frac{3}{40} R_{S_1}(\alpha^*, \beta^*). \tag{45}$$

For sub-regime 1.B, let

$$(s_1, s_2) = \begin{cases} \left(\left\lfloor \frac{N}{K_1M_1} \right\rfloor, \left\lfloor \frac{M_1}{M_2} \right\rfloor \right) & \text{for} \quad M_1 \geq M_2, \\ \left(\left\lfloor \frac{N}{K_1M_1} \right\rfloor, 1 \right) & \text{otherwise}. \end{cases}$$

Note that for $M_1 \geq M_2$, we have

$$\begin{aligned} 1 &= \left\lfloor \frac{N}{K_1.N/K_1} \right\rfloor \leq \left\lfloor \frac{N}{K_1 M_1} \right\rfloor \leq \frac{N}{K_1 M_1} \leq 2, \\ 1 &\leq \left\lfloor \frac{M_1}{M_2} \right\rfloor \leq \frac{M_1}{M_2} \leq \frac{N/K_1}{3N/(4K_2)} = \frac{4K_2}{3K_1}, \end{aligned}$$

and for $M_1 < M_2$, we have

$$1 = \left\lfloor \frac{N}{4N/4} \right\rfloor \leq \left\lfloor \frac{N}{4M_2} \right\rfloor \leq \left\lfloor \frac{N}{K_1 M_1} \right\rfloor \leq \frac{N}{K_1 M_1} \leq 2.$$

Finally, feeding the chosen values of s_1, s_2 into (38) we obtain

$$\begin{aligned} R^{lb}_{S_1}(M_1, M_2) &\geq \frac{\frac{N}{4K_1 M_2}\left(N - \frac{N}{K_1 M_1} \cdot M_1 - \frac{N}{4M_2} \cdot M_2\right)}{N} \\ &\geq \frac{N}{4K_1 M_2}\left(1 - \frac{1}{K_1} - \frac{1}{4}\right) \\ &\geq \frac{N}{32 M_2} \geq \frac{1}{32} \cdot \frac{4}{5} \cdot \frac{N-1}{M_2 - 1} \\ &\geq \frac{1}{40} \min\left\{K_1 K_2, \frac{N-1}{M_2-1}\right\}. \end{aligned} \tag{46}$$

Combining (46) and (22a), we obtain

$$R^{lb}_{S_1}(M_1, M_2) \geq \frac{1}{40} R_{S_1}(\alpha^*, \beta^*). \tag{47}$$

Similarly, in sub-regime 1.C, we have

$$\begin{aligned} R^{lb}_{S_1}(M_1, M_2) &\geq \frac{N}{M_2} - 4 \geq \frac{N-1}{M_2-1} - 4 \\ &\geq \min\left\{K_1 K_2, \frac{N-1}{M_2-1}\right\} - 4. \end{aligned} \tag{48}$$

Combining (48) and (22a), we obtain

$$R^{lb}_{S_1}(M_1, M_2) \geq R_{S_1}(\alpha^*, \beta^*) - 4. \tag{49}$$

Our analysis for sub-regimes 1.A, 1.B and 1.C demonstrate that the secure achievable rate $R^{lb}_{S_1}(M_1, M_2)$ is within a constant multiplicative and additive gap for regime 1.

As for regime 2, we further divide it into the following sub-regimes.

$$\begin{aligned} &(2.\text{A}) \quad \frac{M_1 K_2}{M_1 + M_2 K_2} \leq M_1 < \frac{N}{K_1}, \quad 1 \leq M_2 < \frac{N}{K_1 K_2}, \\ &(2.\text{B}) \quad \frac{M_1 K_2}{M_1 + M_2 K_2} \leq M_1 < \frac{N}{K_1}, \quad \frac{N}{K_1 K_2} \leq M_2 < \frac{N}{3K_2}, \\ &(2.\text{C}) \quad \frac{M_1 K_2}{M_1 + M_2 K_2} \leq M_1 < \frac{N}{K_1}, \quad \frac{N}{3K_2} \leq M_2 < \frac{N}{4}, \\ &(2.\text{D}) \quad \frac{N}{K_1} \leq M_1 \leq N, \quad 1 \leq M_2 < \frac{N - M_1}{2K_2}, \\ &(2.\text{E}) \quad \frac{N}{K_1} \leq M_1 \leq N, \quad \frac{N - M_1}{2K_2} \leq M_2 < \frac{N - M_1}{K_2}. \end{aligned}$$

For sub-regime 2.A, we assume $s_1 = \lfloor \frac{K_1}{3} \rfloor$ and $s_2 = K_2$. Using $\lfloor z \rfloor \geq z/2$ for any $z \geq 1$, we see that it is a valid choice of s_1, s_2, since $K_1 \geq 4$ and thus $\lfloor K_1/3 \rfloor \geq 1$. Equating the values of s_1, s_2 in (38), we obtain

$$\begin{aligned}
R^{lb}_{S_1}(M_1,M_2) &\geq \frac{1}{N}\left[\left\lfloor\frac{K_1}{3}\right\rfloor K_2(N-\left\lfloor\frac{K_1}{3}\right\rfloor M_1-\left\lfloor\frac{K_1}{3}\right\rfloor K_2M_2)\right] \\
&\overset{(a)}{\geq} \frac{1}{N}\left[\frac{K_1K_2}{6}\left(N-\frac{M_1K_1}{3}-\frac{M_2K_1K_2}{3}\right)\right] \\
&\overset{(b)}{\geq} \frac{1}{N}\left[\frac{K_1K_2}{6}\left(N-\frac{N}{3}-\frac{N}{3}\right)\right] \\
&= \frac{K_1K_2}{18} \\
&\geq \frac{1}{18}\min\Bigg\{K_1K_2,\frac{M_1}{M_1+M_2K_2}\cdot\frac{K_2(N-M_1)}{M_1+(M_2-1)K_2} \\
&\quad +\frac{M_2K_2}{M_1+M_2K_2}\cdot\frac{(N-1)K_2M_2}{(M_2-1)(M_1+M_2K_2)}\Bigg\}, \qquad (50)
\end{aligned}$$

where (a) follows from $\lfloor z\rfloor \geq z/2$ for any $z \geq 1$; and (b) follows from $M_1 < N/K_1$, $M_2 < N/(K_1K_2)$. Combining the result with (23a), we obtain

$$R^{lb}_{S_1}(M_1,M_2) \geq \frac{1}{18}R_{S_1}(\alpha^*,\beta^*). \qquad (51)$$

The remaining sub-regimes of this regime can be analyzed in a similar manner, thus we present only the values chosen for s_1 and s_2, as well as the final inequality for each sub-regime. The values $(\lfloor\frac{N}{3M_2K_2}\rfloor, K_2)$, $(1, K_2)$, $(1, \lfloor\frac{N}{4M_2}\rfloor)$ and $(1, \lfloor\frac{N-M_1}{2M_2}\rfloor)$ are chosen for (s_1, s_2) in sub-regimes 2.B, 2.C, 2.D and 2.E, respectively. Moreover, inequalities (52) through (55) demonstrate the gaps for the same sub-regimes, respectively,

$$\begin{aligned}
R^{lb}_{S_1}(M_1,M_2) &\geq \frac{2}{135}R_{S_1(\alpha^*,\beta^*)}, && (52)\\
R^{lb}_{S_1}(M_1,M_2) &\geq \frac{3}{64}R_{S_1}(\alpha^*,\beta^*), && (53)\\
R^{lb}_{S_1}(M_1,M_2) &\geq \frac{1}{32}R^{lb}_{S_1}(\alpha^*,\beta^*), && (54)\\
R^{lb}_{S_1}(M_1,M_2) &\geq \frac{1}{96}R_{S_1}(\alpha^*,\beta^*). && (55)
\end{aligned}$$

We will also study regime 3 through dividing it into two sub-regimes as follows:

$$\begin{aligned}
&\text{3.A)} \quad \frac{N}{K_1} \leq M_1 \leq N, \qquad \frac{N-M_1}{K_2} \leq M_2 < \frac{N-M_1}{2},\\
&\text{3.B)} \quad \frac{N}{K_1} \leq M_1 \leq N, \qquad \frac{N-M_1}{2} \leq M_2 \leq N.
\end{aligned}$$

The reasoning method is similar to the case of sub-regimes 2.A through 2.E. Therefore, we briefly mention only the chosen values for (s_1, s_2) and the final inequality obtained for each sub-regime. For sub-regime 3.A, we chose $s_1 = 1$ and $s_2 = \lfloor\frac{N-M_1}{2M_2}\rfloor$ and derive

$$R^{lb}_{S_1}(M_1,M_2) \geq \frac{1}{16}R_{S_1}(\alpha^*,\beta^*). \qquad (56)$$

In sub-regime 3.B, we obtain

$$
\begin{aligned}
R^{lb}_{S_1}(M_1,M_2) &\geq 0=\frac{8}{3}-\frac{8}{3}\\
&\geq \frac{4}{3}\cdot 2\cdot 1-\frac{8}{3}\\
&\overset{(a)}{\geq} \frac{K_1}{K_1-1}\cdot\frac{N-M_1}{M_2}\cdot\frac{N-M_1}{N}-\frac{8}{3}\\
&\geq \frac{K_1}{K_1-1}\frac{(N-M_1)^2}{M_2N}-\frac{8}{3}\\
&= \frac{M_2-1}{M_2}\cdot\frac{K_1}{K_1-1}\frac{(N-M_1)^2}{N(M_2-1)}-\frac{8}{3}\\
&\overset{(b)}{\geq} \frac{5}{6}R_{S_1}(\alpha^*,\beta^*)-\frac{8}{3},
\end{aligned}
\tag{57}
$$

where (a) follows from $\frac{N-M_1}{M_2}\leq 2$ and (b) follows from $N\geq K_1K_2$, $K_1\geq 4$, and (24a).

The results obtained for sub-regimes 3.A and 3.B suggest that the gap analysis for regime 3 will be similar to the case of regime 1 and regime 2. On the other hand, we show that regimes 1, 2 and 3 cover the entire (M_1,M_2) plane. This helps us come into the conclusion that in each subregime $R_{S_1}(\alpha^*,\beta^*)$ and $R^{lb}_{S_1}(M_1,M_2)$ are within a constant multiplicative and additive gap. Therefore, the unified final result which we will obtain for all the studied regimes is

$$
R_{S_1}\geq R^{lb}_{S_1}(M_1,M_2)\geq\frac{1}{96}R_{S_1}(\alpha^*,\beta^*)-4. \tag{58}
$$

□

5.2. $R_{S_2}(\alpha^*,\beta^*)$ against $R^{lb}_{S_2}(M_1,M_2)$

Theorem 4. *$R_{S_2}(\alpha^*,\beta^*)$ is within a constant multiplicative and additive gap with $R^{lb}_{S_2}$ for every possible value of (M_1,M_2). Specifically,*

$$
R_{S_2}\geq R^{lb}_{S_2}(M_1,M_2)\geq\frac{1}{45}R_{S_2}(\alpha^*,\beta^*)-16.
$$

Proof. Let us focus on the case where $N\geq K_1K_2$, $K_1\geq 4$ and $K_2\geq 4$. Recall from (25b) that achievable secure rate $R_{S_2}(\alpha^*,\beta^*)$ is upper bounded as

$$
R_{S_2}(\alpha^*,\beta^*)\leq K_1\cdot\ \min\left\{K_2,\frac{N-1}{M_2-1}\right\}. \tag{59}
$$

Furthermore, the lower bound on $R_{S_2}(\alpha^*,\beta^*)$ can be obtained from (43) as

$$
R^{lb}_{S_2}(M_1,M_2)=\max_{t\in\{1,2,\ldots,K_2\}}\frac{t(N-tM_2)}{N}. \tag{60}
$$

In the rest of our discussion we will partition the (M_1,M_2) plane by distinguishing the following two cases:

$$
\begin{aligned}
&(1) \qquad 1\leq M_2\leq\frac{N}{4},\\
&(2) \qquad \frac{N}{4}\leq M_2\leq N.
\end{aligned}
$$

We will examine the mentioned cases in order to improve the margin of the gap.

(1) $1\leq M_2\leq\dfrac{N}{K_2}$, let $t=\left\lfloor\frac{1}{3}\min\left\{K_2,\frac{N}{M_2}\right\}\right\rfloor$ in (60). This is a valid choice since $K_2\geq 4$. Thus,

$$
1\leq\left\lfloor\frac{1}{3}\min\left\{K_2,\frac{N}{M_2}\right\}\right\rfloor\leq\frac{K_2}{3}.
$$

By feeding the value of t into (60), it follows that

$$R_{S_2}^{lb}(M_1, M_2) \geq \frac{1}{N}\left[\left\lfloor\frac{1}{3}\min\left\{K_2, \frac{N}{M_2}\right\}\right\rfloor \cdot \left(N - \left\lfloor\frac{1}{3}\min\left\{K_2, \frac{N}{M_2}\right\}\right\rfloor M_2\right)\right].$$

Since $\forall z \geq 1 : \lfloor z \rfloor \geq z/2$, we can continue as follows,

$$\begin{aligned} R_{S_2}^{lb}(M_1, M_2) &\geq \frac{1}{N}\left[\frac{1}{6}\min\left\{K_2, \frac{N}{M_2}\right\}\left(N - \frac{N}{3}\right)\right] \\ &= \frac{1}{9}\min\left\{K_2, \frac{N}{M_2}\right\}. \end{aligned}$$

Because $N \geq K_1K_2$ and $K_1 \geq 4$, we have

$$\begin{aligned} R_{S_2}^{lb}(M_1, M_2) &\geq \frac{1}{9} \cdot \frac{4}{5}\min\left\{K_2, \frac{N-1}{M_2-1}\right\} \\ &= \frac{4}{45}\min\left\{K_2, \frac{N-1}{M_2-1}\right\} \\ &\geq \frac{1}{45} \cdot K_1 \cdot \min\left\{K_2, \frac{N-1}{M_2-1}\right\}. \end{aligned} \tag{61}$$

From (61) and (25b) we obtain

$$R_{S_2}^{lb}(M_1, M_2) \geq \frac{1}{45}R_{S_2}(\alpha^*, \beta^*). \tag{62}$$

(2) For $\frac{N}{4} \leq M_2 \leq N$, it holds that

$$\begin{aligned} R_{S_2}^{lb}(M_1, M_2) &\geq 0 = K_1\frac{N}{M_2} - K_1\frac{N}{M_2} \\ &\geq K_1 \cdot \min\left\{K_2, \frac{N}{M_2}\right\} - K_1\frac{N}{M_2} \\ &\overset{(K_1 \geq 4)}{\geq} K_1 \cdot \frac{3}{4}\min\left\{K_2, \frac{N-1}{M_2-1}\right\} - 16. \end{aligned}$$

Therefore,

$$R_{S_2}^{lb}(M_1, M_2) \geq \frac{3}{4}R_{S_2}(\alpha^*, \beta^*) - 16. \tag{63}$$

The entire (M_1, M_2) plane is obviously covered by cases 1) and 2). Thus, $R_{S_2}(\alpha^*, \beta^*)$ and $R_{S_2}^{lb}(M_1, M_2)$ are shown to be embraced by constant additive and multiplicative curves as shown by (64) which is derived via combining (62) and (63),

$$R_{S_2} \geq R_{S_2}^{lb}(M_1, M_2) \geq \frac{1}{45}R_{S_2}(\alpha^*, \beta^*) - 16. \tag{64}$$

□

6. Conclusions and Further Works

In this paper, we further developed the system model of a coded caching scheme by simultaneously assuming a hierarchical network and adversaries tapping the shared links in peak time. We calculated the secure achievable rates of each link in the proposed scheme. The parameters considered in previously-proposed hierarchical scheme have been reconsidered here to obtain approximate minimum achievable rates. Furthermore, we calculated the lower bound on the feasible rates. We showed that the secure achievable rates are within a constant multiplicative and additive gap to the corresponding lower bounds. These results are similar to those obtained in the non-secure hierarchical scheme, but the cost of security appears in the form of larger constants. Our work can be continued by proposing and evaluating yet more complex system models. More complicated models can assume that the adversary has access to the shared links in the placement phase or allow the users to issue more than one request in the delivery phase.

Author Contributions: Formal analysis, B.Z., V.S., B.K.R., K.B. and T.K.; Writing—original draft, B.Z., V.S., B.K.R., K.B. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: TK was supported in part by JSPS Grant-in-Aids for Scientific Research (A) No.16H01705 and No.21H04879, for Scientific Research (B) No.17H01695, and for Challenging Exploratory Research No.19K22849.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the reviewers for helpful comments that improved the presentation of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Maddah-Ali, M.A.; Niesen, U. Fundamental limits of caching. *IEEE Trans. Inf. Theory* **2014**, *60*, 2856–2867. [CrossRef]
2. Sengupta, A.; Tandon, R.; Clancy, T.C. Decentralized caching with secure delivery. In Proceedings of the IEEE International Symposium on Information Theory, Honolulu, HI, USA, 29 June–4 July 2014.
3. Sengupta, A.; Tandon, R.; Clancy, T.C. Fundamental limits of caching with secure delivery. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 355–370. [CrossRef]
4. Karamchandani, N.; Niesen, U.; Maddah-Ali, M.A.; Diggavi, S.N. Hierarchical coded caching. *IEEE Trans. Inf. Theory* **2016**, *62*, 3212–3229. [CrossRef]
5. Bai, B.; Li, W.; Wang, L.; Zhang, G. Coded caching in fog-ran: b-matching approach. *IEEE Trans. Commun.* **2019**, *67*, 3753–3767. [CrossRef]
6. Cao, H.; Yan, Q.; Tang, X. Reducing search complexity of coded caching by shrinking search space. *IEEE Commun. Lett.* **2019**, *23*, 568–571. [CrossRef]
7. Kim, G.; Hong, B.; Choi, W.; Park, H. Mds-coded caching leveraged by coordinated multi-point transmission. *IEEE Commun. Lett.* **2018**, *22*, 1220–1223. [CrossRef]
8. Luo, T.; Peleato, B. The transfer load-i/o trade-off for coded caching. *IEEE Commun. Lett.* **2018**, *22*, 1524–1527. [CrossRef]
9. Zhang, J.; Lin, X.; Wang, X. Coded caching under arbitrary popularity distributions. *IEEE Trans. Inf. Theory* **2018**, *64*, 349–366. [CrossRef]
10. Cao, Y.; Tao, M. Treating content delivery in multi-antenna coded caching as general message sets transmission: A dof region perspective. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 3129–3141. [CrossRef]
11. Ngo, K.-H.; Yang, S.; Kobayashi, M. Scalable content delivery with coded caching in multi-antenna fading channels. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 548–562. [CrossRef]
12. Yang, Q.; Gündüz, D. Coded caching and content delivery with heterogeneous distortion requirements. *IEEE Trans. Inf. Theory* **2018**, *64*, 4347–4364. [CrossRef]
13. Pedersen, J.; Amat, A.G.; Andriyanova, I.; Brännström, F. Optimizing mds coded caching in wireless networks with device-to-device communication. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 286–295. [CrossRef]
14. Shariatpanahi, S.P.; Caire, G.; Khalaj, B.H. Physical-layer schemes for wireless coded caching. *IEEE Trans. Inf. Theory* **2019**, *65*, 2792–2807. [CrossRef]
15. Tang, A.; Roy, S.; Wang, X. Coded caching for wireless backhaul networks with unequal link rates. *IEEE Trans. Commun.* **2018**, *66*, 1–13. [CrossRef]
16. Panigrahi, B.; Shailendra, S.; Rath, H.K.; Simha, A. Universal caching model and markov-based cache analysis for information centric networks. *Photonic Netw. Commun.* **2015**, *30*, 428–438. [CrossRef]
17. Cheng, M.; Jiang, J.; Yan, Q.; Tang, X. Constructions of coded caching schemes with flexible memory size. *IEEE Trans. Commun.* **2019**, *67*, 4166–4176. [CrossRef]
18. Shangguan, C.; Zhang, Y.; Ge, G. Centralized coded caching schemes: A hypergraph theoretical approach. *IEEE Trans. Inf. Theory* **2018**, *64*, 5755–5766. [CrossRef]
19. Vilardebó, J.G. A novel centralized coded caching scheme with coded prefetching. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1165–1175. [CrossRef]

20. Wang, J.; Cheng, M.; Yan, X.; Tang, Q. Placement delivery array design for coded caching scheme in d2d networks. *IEEE Trans. Commun.* **2019**, *67*, 3388–3395. [CrossRef]
21. Asghari, S.M.; Ouyang, Y.; Nayyar, A.; Avestimehr, A.S. An approximation algorithm for optimal clique cover delivery in coded caching. *IEEE Trans. Commun.* **2019**, *67*, 4683–4695. [CrossRef]
22. Zheng, L.; Yan, Q.; Chen, Q.; Tang, X. Delivery design for coded caching over wireless multicast networks. *IEEE Access* **2019**, *7*, 72803–72817. [CrossRef]
23. Zhang, K.; Tian, C. Fundamental limits of coded caching: From uncoded prefetching to coded prefetching. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1153–1164. [CrossRef]
24. Bayat, M.; Mungara, R.K.; Caire, G. Achieving spatial scalability for coded caching via coded multipoint multicasting. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 227–240. [CrossRef]
25. Vettigli, G.; Ji, M.; Shanmugam, K.; Llorca, G.; Tulino, A.M.; Caire, G. Efficient algorithms for coded multicasting in heterogeneous caching networks. *Entropy* **2019**, *21*, 324. [CrossRef]
26. Zhong, S.; Wang, X. Joint multicast and unicast beamforming for coded caching. *IEEE Trans. Commun.* **2018**, *66*, 3354–3367. [CrossRef]
27. Combes, R.; Ghorbel, A.; Kobayashi, M.; Yang, S. Utility optimal scheduling for coded caching in general topologies. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1692–1705. [CrossRef]
28. Karat, N.S.; Thomas, A.; Rajan, B.S. Error correction in coded caching with symmetric batch prefetching. *IEEE Trans. Commun.* **2019**, *67*, 7264–7274. [CrossRef]
29. Pääkkönen, G.; Barreal, A.; Hollanti, C.; Tirkkonen, O. Coded caching clusters with device-to-device communications. *IEEE Trans. Mob. Comput.* **2019**, *18*, 264–275. [CrossRef]
30. Ibrahim, A.A.; Zewail, A.M.; Yener, A. Coded caching for heterogeneous systems: An optimization perspective. *IEEE Trans. Commun.* **2019**, *67*, 5321–5335. [CrossRef]
31. Zhang, J.; Lin, X.; Wang, C.-C.; Wang, X. Coded caching for files with distinct file sizes. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 1686–1690.
32. Lampiris, E.; Elia, P. Adding transmitters dramatically boosts coded-caching gains for finite file sizes. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 1176–1188. [CrossRef]
33. Niesen, U.; Maddah-Ali, M.A. Coded caching with nonuniform demands. *IEEE Trans. Inf. Theory* **2014**, *63*, 221–226.
34. Ding, Y.; Wang, L.; Wu, H.; Shen, H.V.; Poor, X. Tradeoff of content sharing efficiency and secure transmission in coded caching systems. In Proceedings of the IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20–24 May 2018.
35. Kiskani, M.K.; Sadjadpour, H.R. Secure coded caching in wireless ad hoc networks. In Proceedings of the International Conference on Computing, Networking and Communications (ICNC), Silicon Valley, CA, USA, 26–29 January 2018.
36. Zewail, A.A.; Yener, A. Coded caching for resolvable networks with security requirements. In Proceedings of the IEEE Conference on Communications and Network Security (CNS): The Workshop on Physical-Layer Methods for Wireless Security, Philadelphia, PA, USA, 17–19 October 2016.
37. Kamel, M.; Wigger, S.; Sarkiss, M. Decentralized coded caching for wiretap broadcast channels. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018.
38. Suthan, I.; Chugh, C.H.H.; Krishnan, P. An improved secretive coded caching scheme exploiting common demands. In Proceedings of the IEEE Information Theory Workshop (ITW), Kaohsiung, Taiwan, 6–10 November 2017.
39. Hachem, J.; Karamchandani, N.; Diggavi, S.N. Coded caching for multi-level popularity and access. *IEEE Trans. Inf. Theory* **2017**, *63*, 3108–3141. [CrossRef]
40. Lim, S.H.; Wang, C.; Gastpar, M. Information-theoretic caching: The multi-user case. *IEEE Trans. Inf. Theory* **2017**, *63*, 7018–7037. [CrossRef]
41. Sengupta, A.; Tandom, R.; Clancy, T.C. Improved approximation of storage-rate tradeoff for caching via new outer bounds. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015.
42. Vijit, K.K.P.; Rai, B.K.; Jacob, T. Towards the exact rate memory tradeoff in coded caching. In Proceedings of the National Conference on Communications (NCC), Bangalore, India, 20–23 February 2019.
43. Wei, Y.; Ulukus, S. Coded caching with multiple file requests. In Proceedings of the 55th Annual Allerton Conference on Communication, Control and Computing (Allerton), Monticello, IL, USA, 3–6 October 2017.
44. Maddah-Ali, M.A.; Niesen, U. Decentralized coded caching attains order-optimal memory-rate tradeoff. *IEEE/ACM Trans. Netw. (TON)* **2015**, *23*, 1029–1040. [CrossRef]
45. Wei, Y.; Ulukus, S. Novel decentralized coded caching through coded prefetching. In Proceedings of the 2017 IEEE Information Theory Workshop (ITW), Kaohsiung, Taiwan, 6–10 November 2017.
46. Shannon, C.E. Communication theory of secrecy systems. *Bell Syst. Tech. J.* **1949**, *28*, 656–715. [CrossRef]