A Defense Mechanism Against Timing Attacks on User Privacy in ICN

Vignesh Sivaraman and Biplab Sikdar, *Senior Member, IEEE*

***Abstract*—While in-network caching is an essential feature of Information Centric Networks (ICN) for improved content dissemination and reducing the bandwidth consumption at the core of the network, it is prone to many privacy threats. For example, an adversary can passively breach the privacy of a consumer by simply analyzing the different retrieval times for the same content. This paper aims to address this problem of timing analysis attacks by developing privacy-enhancing caching strategies. The proposed caching strategies use two privacy metrics, namely mutual information from information theory and differential privacy, and formulates a privacy enhancing distributed optimization problem with the objective of optimizing the network cost incurred. We efficiently solve the optimization problem by considering it as a n-player, non-cooperative game. We show that Nash equilibrium exists for this game and compute it using an iterative best response algorithm. We compare and validate the performance of our approach on realistic network topologies by comparing it with the existing approaches in literature and the global optimal solutions.**

***Index Terms*—Information centric networks, timing attacks, privacy**

I. Introduction

Cisco visual networking index predicts that the global multimedia traffic would account for 82% of the total Internet traffic by 2022 [1], thus, rendering the content distribution as an important challenge to be addressed. The focus of the Internet has continually shifted from *where* is the content to *what* is the content, but the communication architecture still focuses on *where* the content is. Information Centric Networks (ICN) have been proposed to address this challenge of efficient distribution and retrieval of content. ICN are communication networks built with content as the primary entity of the network where the content names are used as identifiers (or addresses) and not the content locations. ICN has the following unique characteristics which are fundamental for the efficient dissemination of content: (i) Content's identity and its location are decoupled; (ii) Content is cached at the content stores of the intermediate nodes in the network (in-network caching); (iii) The content traverses the reverse path of its corresponding request (also referred to as interest).

In contrast to routers in the traditional Internet, additional storage is deployed at the ICN routers where the routers cache the forwarded content. When an interest arrives for a content, the router checks whether the content is present in its cache and if so, the content is served from the router's cache (instead of the interest being sent to the content producer). Thus, in-network caching enables faster content retrieval and improves the bandwidth efficiency of the network.

V. Sivaraman and B. Sikdar are associated with the National University of Singapore, Singapore. (E-mail: vignesh.sivaraman@u.nus.edu and bsikdar@nus.edu.sg)

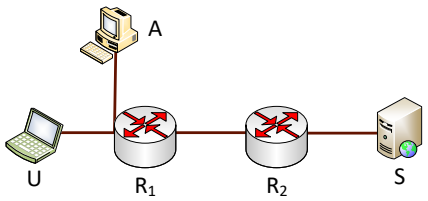


Fig. 1. Toy example illustrating timing attacks in ICN.

In the default caching approach of ICN, the routers cache all the content forwarded by them. This results in the content leaving a copy everywhere along the path from the content producer to the consumer. On the one hand this approach improves the network performance, but on the other hand it makes the network vulnerable to privacy attacks [2], [3]. The content in ICN traverses the same path as that of the corresponding interest (unlike TCP/IP networks). If the content is cached at a router, then it implies that the corresponding interest was forwarded by the router. Thus, when the router is on the edge of the network, it can be inferred that the requesting consumer is connected to the router. Timing analysis attacks leverage this feature of ICN along with its in-network caching to breach the privacy of the consumers [4].

A. An Example of Timing Analysis Attack in ICN

Consider the toy network in Figure 1. Here, U is a honest user, A is the adversary, R_1 and R_2 are ICN routers, and S is the content producer. U and A are connected to the same router R_1. Let a_r and a_s be the RTT between A and R_1, and, A and S, respectively. First, A sends a request for some unpopular content c_u and by measuring the RTT for c_u, A estimates a_s. Next, A requests for c_u again and as c_u is cached at R_1, R_1 responds back with c_u. By measuring the RTT for this request, A estimates a_r. Once A estimates the values of a_s and a_r, A can probe the cache of R_1 for various contents, measure the RTT t, and compare it with a_r and a_s. If $t \simeq a_r$ then A can infer that the content has been requested by U. Using this attack, A can identify the sensitive contents being accessed by U. In order to carry out this attack, A only needs to perform time measurements and does not need any other special resources or privileges. From the network's perspective, the behavior of A is similar to that of any other honest consumer.

The timing analysis attack in ICN posses a serious threat to the privacy of users. First, individual consumers' private browsing data, Internet usage and behavioral patterns can get leaked through such attacks. Furthermore, in scenarios where the adversary and the target share the same Internet Service Provider and are co-located (e.g. same building or business park), these attacks can lead to a business intelligence attack where the adversary is is interested to know contents requested by a competing company. Second, the adversary behaves exactly like a honest consumer and generates legitimate requests for legitimate contents. Therefore, the timing analysis attack is difficult to detect. The potential threat to users' privacy and the undetectability of the adversary makes the timing analysis attack in ICN an important challenge to be addressed.

B. Threat Model

We consider an adversary who is capable of inferring the contents of the cache of the routers in the network. The adversary can measure the round trip time of any content it requests and the adversary can use these time measurements to carry out a timing analysis attack to infer the source of the content. We note that such a timing analysis attack is a passive attack and the adversary does not modify any of the content. To this end, it is difficult to distinguish between an adversary and an honest user. The objective of the adversary is to, first, infer the contents of the cache of the routers in the network. Next, using this observation, the adversary tries to identify the user who had requested for the content, thereby breaching the privacy of the user. Organizational networks like university campus networks or office networks are easily susceptible to such timing attacks as the adversary may have sufficient side information about the users in the network and the caching routers. We assume that the network operator stores the caching strategies of all the users. Breaching the storage of the network operator requires a very strong adversary. Any adversary capable of breaching the network operator's infrastructure can easily breach directly into the privacy of users. Such an adversary does not need to infer the contents a router caches in order to breach the users' privacy. The assumption that the adversary cannot breach into the network operator's infrastructure is a valid assumption here because we consider timing analysis attacks in ICN. To perform this attack, an adversary only needs to precisely measure the content delivery time. There is a significant difference in the amount of resources required for an adversary to perform a (passive) timing analysis attack and to breach into the network operator's infrastructure. Thus, an adversary with limited capabilities (that are sufficient for timing analysis attacks) is unlikely to be able to breach the network operator's hardware infrastructure.

Timing analysis attacks can be deployed to determine the location of the consumers as well. Consider that the adversary has the side information about the specific content the consumer is interested in. Then, the adversary can request for the same content and measure the delay incurred. Based on the delay measured, the adversary can infer which router has cached the content and hence infer the location of the consumer.

While our approach addresses both the above mentioned threat models, we primarily focus our discussions using the first model where the adversary aims to figure out details about the content accessed by a consumer.

C. Contributions

While caching of content at the routers along the path (as discussed above) is easy to implement, it leads to data redundancy. Different off-path and co-operative caching mechanisms have been proposed in the literature to address this problem [5]–[9]. We exploit the off-path caching in ICN to address the timing analysis attack problem. The proposed solution formulates a non-cooperative game among the consumers (also henceforth referred to as users), where each consumer probabilistically chooses a router for caching a given content such that the cost incurred (in terms of network resources consumed and delay incurred) is minimized and its privacy requirements (in terms of information theoretic mutual information) are satisfied. We show that a Nash equilibrium exists in such a game and it can be computed using an iterative best response algorithm. Moreover, we formulate another game which incorporates differential privacy along with the information theoretic mutual information bound. This approach provides a framework for trade-off between the amount of information gained by an adversary (using a timing analysis attack) and the cost incurred by the caching strategy to mitigate it. The **contributions of this paper** can be summarized as:

- We propose a distributed content caching scheme to mitigate the impact of timing analysis attacks on ICN routers.
- The proposed privacy framework is based on a non-cooperative game formulation among the consumers. The privacy requirements are quantified using information theoretic mutual information and differential privacy. The Nash equilibrium of the game is computed using an iterative best response algorithm.
- The performance of the proposed scheme is compared with the existing approaches in the literature as well as with the globally optimal solution and the exponential mechanism (for differential privacy). The proposed approach, in addition to being distributed, also achieves near optimal cost and better cache utilization compared to the global solutions.

The rest of the paper is organized as follows. The related literature is presented in Section II and Section III describes the system model and the required preliminaries. The proposed game model is presented in Section IV and Section V presents the game with differential privacy constraints. The Nash equilibrium of the game, an iterative best-response algorithm to solve it, and its convergence are discussed in Section VI. The performance evaluation of the proposed approach is discussed in Section VII and Section VIII concludes the paper.

II. Related Work

As any other Internet architecture, ICN too is vulnerable to security attacks. The surveys in [2], [3], [10]–[12] highlight the various privacy and security threats present in ICN. The classical denial of service attack in ICN has been addressed using various approaches like rate limiting [13], [14], statistical

modeling [15], and PIT modification [16]. Different routing related issues like secure content naming [17], secure forwarding [18], and secure routing [19] have been discussed in the literature. The authors in [20] address the content poisoning attack where the adversary fills a router's cache with invalid contents. The authors in [21], [22] discuss the cache pollution attack where the adversary skews the content popularity by requesting less popular content frequently.

The timing analysis attack in ICN is an important problem that is being addressed by the research community. Different approaches in the literature that aim to address the timing analysis attack and enhance the cache privacy in ICN can be classified as follows:

1) Content placement based approaches: The authors in [23] decide whether to cache the content at a router or not based on the betweenness-centrality of the router. Sensitive contents are cached at routers with high betweenness-centrality in order to increase the consumers' anonymity set (the number of consumers using a given router). Using this mechanism, it is possible that the content is cached only at the upstream routers to achieve the required size of the anonymity set. Our caching strategy aims to achieve the same level of privacy by strategically caching the content at the edge routers of the network. Authors in [24] propose a caching scheme where the routers are clustered into groups. Routers in every group collaborate among each other and cache the content at one of the routers in the group. This scheme assumes a global privacy requirement of the system and does not take into account different privacy needs of the users. Moreover, the caching scheme does not take into account the effect of the privacy on the network performance. As our proposed approach too tries to enhance privacy by intelligent placement of content in the router caches, we compare our caching strategies with these strategies in our performance study.
2) Delay based approaches: The authors in [25] propose a mechanism to add delay to the content being served from the router's cache. The amount of delay added is such that it mimics the delay incurred when the content is served by one of the upstream routers or the server. In [4], the authors propose to introduce cache misses at the routers when the number of requests for the content is less than a threshold. The threshold chosen is random and is assumed to be a secret. The authors discuss the privacy corresponding to the value of this threshold. In [26], authors propose a detection based approach where detection techniques are deployed at the router and a delay is introduced only if an attack is detected. While introducing delays and cache misses may guarantee privacy for the consumers, they diminish the efficiency of the caches as the content is redundantly present in all the on-path routers and yet, consumers still experience significant delays.
3) Access control based approaches: Access control based approaches are proposed in [27]–[29]. Access control approaches do not directly aim to address cache privacy in ICN. Instead, the main aim here is to ensure that only authorized users are able to access the content in the network. While such methods partially enhance the cache privacy in ICN, they cannot ensure that users' privacy is maintained. For example, in scenarios like organizational networks, all the users including the adversary can have access to the content.

The problem of user privacy and timing analysis attack for the current Internet in general is a well explored problem. For example, approaches in [30], [31] enable an adversary to learn about the different content a user has recently accessed by reading the cache and cookies of the browser. The authors of [32] develop a timing analysis attack to gain access to the browser's cache (and hence breach the user's privacy). Similarly, [33] proposes a scheme to learn the users' call history in VoIP networks.

Various countermeasures exist to mitigate the cache privacy and timing attacks in the Internet. For example, [34] proposes an approach where the URLs are customized for every user so that the adversary cannot identify the content accessed by the user. In [35], the authors introduce a delay based approach to mitigate timing attacks (similar to [4], [25], [26]). Anonymous networks like ToR have also been widely studied in the context of user privacy and timing attacks [36]–[38]. We note that the cache privacy and timing analysis attack in ICN is different from that of the current Internet. Unlike the current Internet, the cache privacy and timing analysis attack in ICN completely focuses on the in-network caches at the routers.

The existing approaches for cache privacy and timing attacks in ICN preserve privacy by leveraging the uncertainty arising from the presence of other users using the on-path routers. In this work, we aim to preserve privacy by using the uncertainty provided by the other users in the network (on-path as well as off-path). As there are usually more users in the network than routers, our approach achieves better privacy for the same cost. Moreover, our approach inherently leaves only one copy of the content in the network, thereby improving the caching efficiency, without sacrificing the content access delay (although we show that we can easily extend our model to cache multiple copies of the content as required).

There are mechanisms to deploy off-path caching within a domain (like an autonomous system, enterprise network, and a campus network). In [5]–[7], the authors design a distributed off-path caching mechanism for autonomous system level networks. The authors in [8] propose techniques to discover off-path cached content using the trails left by the content. A survey of caching strategies can be found in [9]. We assume that an off-path caching mechanism is already available in the network. Proposing an architecture to achieve off-path caching is not a goal of this paper. Rather, we focus on making privacy enhancing caching decisions that can be deployed using the available mechanisms.

III. Preliminaries

A. Overview

To protect the users from adversaries performing timing analysis attacks, we aim to design an obscuring caching policy $\mathcal{A}$ for each user, following which the content accessed by the user is cached at different routers corresponding to the probability distribution output by $\mathcal{A}$. As we are concerned

about the quality of service and the incurred network resource costs, we consider delay and storage resource consumption in our cost model. While requesting for privacy, the user needs to decide on the appropriate mechanism $\mathcal{A}$. As the privacy of a user depends not only on her own mechanism but also on the mechanisms chosen by other users as well, we formulate the interactions between the users as a game. Game theory is an apt analytical tool to capture these strategic interactions between the users, including the users' strategies and payoffs.

Solution Approach: We propose two solution approaches to address the problem of timing analysis attack. The **first** approach formulates the problem as a non-cooperative game among the users. Every user solves an optimization problem with the objective to minimize the total cost incurred while achieving a minimum privacy level (different users may have different privacy levels). We quantify the privacy level of a user using the Shannon entropy of the system (to be precise, mutual information). Given the strategies of other users, a user evaluates her best response by solving the optimization problem. We show that the Nash equilibrium exists for our game and that it can be reached using a greedy iterative approach. While the **second** approach is similar to the first one, here, along with the Shannon entropy of the system, we also consider differential privacy of the users. Therefore, the optimization problem in this approach minimizes the total cost incurred subject to differential privacy and system entropy constraints.

B. Mutual Information

Let X and Y be a discrete random variables defined on sample spaces $\mathcal{X}$ and $\mathcal{Y}$ with probability mass functions $p_X(\cdot)$ and $p_Y(\cdot)$, respectively (for simplicity, the subscripts are dropped whenever they are clear from the context). The Shannon entropy, $H(X)$, of X is defined as:

$$H(X) = -\sum_{x\in\mathcal{X}} p_X(x)\log p_X(x). \tag{1}$$

The conditional entropy of X given Y is:

$$H(X|Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P(x,y)\log P(x|y). \tag{2}$$

Note that $0 \leq H(X|Y) \leq H(X)$. The mutual information between X and Y, $I(X,Y)$, is defined as:

$$I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{3}$$

C. Differential Privacy

Let us consider two databases D and D' which differ only on a single row (i.e., D and D' are adjacent databases). A randomized algorithm $\mathcal{A}$ is said to be ϵ-differentially private if $\mathcal{A}$ satisfies the following for all adjacent databases D and D' and $O \subseteq \text{Range}(\mathcal{A})$:

$$Pr[\mathcal{A}(D') \in O] \leq e^{\epsilon} \times Pr[\mathcal{A}(D) \in O].$$

We consider the caching strategy of all the users as one database. Therefore, the adjacent databases in our case vary just in the strategy of one user.

Symbol	Description
$\mathcal{R}$	Set of routers in the system
$\mathcal{U}$	Set of consumers in the system
p_u^r	Probability that consumer u chooses router r for caching
$\mathbf{p}_u$	Vector of probabilities with which consumer u chooses the routers in $\mathcal{R}$
q_u	Probability that the content is requested by consumer u
f_u	Cost function of consumer u
c_u^r	Cost incurred by consumer u when using router r for caching
$\mathbf{c}_u$	Vector of cost incurred by consumer u
U	Random variable indicating the requesting user as observed by the adversary
R	Random variable indicating the caching router as observed by the adversary
H	Shanon entropy function of the system
I	Mutual information function of the system
$\mathcal{B}_u$	Set of best response of consumer u
t_u	Upper bound on the mutual information of the system required by consumer u
A	Coefficient matrix of the constraints of optimization problem
t_{min}	Minimum mutual information of the system
t_{max}	Maximum mutual information of the system
$\boldsymbol{\beta}, \lambda$	Lagrangian multipliers of the optimization problem $P0$
$\mathbf{p}_u^*, \boldsymbol{\beta}, \lambda$	Optimal points of optimization problem $P0$ and its dual problem
$Q(\boldsymbol{\beta}, \lambda)$	System of equations to be solved to obtain Lagrangian multipliers of optimization problem $P0$
ϵ_u	Differential privacy parameter of consumer u
f^*	Conjugate function of function f
f^+	Positive homogeneous extension of function f
$\boldsymbol{\alpha}, \boldsymbol{\gamma}, \lambda$	Conjugate dual variables of optimization problem $DP1$
$\mathbf{x}, \mathbf{z}$	Auxiliary variables required to construct optimization problem $DP1$ from optimization problem $DP0$
$\mathbf{P}$	Caching probability matrix of all consumers
$\mathbf{P}_{prev}$	Caching probability matrix of the previous iteration

TABLE I
LIST OF VARIABLES AND SYMBOLS USED

Exponential Mechanism. The exponential mechanism [39] allows us to use differential privacy on arbitrary domains and ranges. Let the arbitrary domain and range for a randomized algorithm $\mathcal{A}$ be $\mathcal{D}$ and $\mathcal{R}$, respectively. Let $\mathcal{A}$ follow the exponential mechanism. Then, for an input $d \in \mathcal{D}$, $\mathcal{A}$ outputs $r \in \mathcal{R}$ with a probability proportional to $exp(\epsilon\frac{u(d,r)}{2\Delta u})$. Here $u : \{\mathcal{D} \times \mathcal{R}\} \rightarrow \mathbb{R}$ is the utility function and Δu is the sensitivity of u, i.e., is the largest possible difference in the utility when two inputs differ only on a single user's value, for any r.

The list of variables and symbols used in this paper are presented in Table I.

IV. Caching Strategy Games

A. System Model

We consider a network consisting of a set of routers $\mathcal{R}$ and a set of consumers $\mathcal{U}$. To provide high quality of service, we assume that all the routers in the network deploy storage and cache the contents forwarded by them. Under the default scenario, the content is cached along the path of travel and hence, a copy of the content is cached at the first hop router of the requesting user. The adversary can target the first hop router to gain information about the user's content preferences. Therefore, in our approach the content can be cached at any of the routers (of the edge network).

The proposed privacy aware caching approach does not require any architectural changes in ICN. We only assume that a manager is deployed in the edge network and a mechanism like [7] exists to perform off path caching (for the users requesting privacy). Here, the authors propose an automatic cache management system architecture where the network

operator specifies the requirements of the caching system and the architecture ensures that the requirements are satisfied. We assume that the network operator specifies the caching strategies of all users to this architecture. The architecture, in turn, manages the caching at the routers such that the privacy requirements of all users are satisfied. Initially, we assume that all the users are aware of the network topology and the other users' preferences (this assumption is relaxed subsequently). A user leverages these information and decides on the level of privacy required.

Let p_u^r be the probability with which consumer u chooses router r for caching. Let $\mathbf{p}_u$ be the vector of probabilities of consumer u corresponding to every router in the edge network, $\mathbf{p}_u = [p_u^1, p_u^2, \cdots, p_u^{|\mathcal{R}|}]^T$. The consumer decides on $\mathbf{p}_u$ based upon the level of privacy required and the corresponding cost incurred for it.

Remark. Our proposed approach leverages off-path caching which allows the content to be cached at any router of the network. In the absence of off-path caching, our approach can easily be restricted only to the on-path routers between a consumer-producer pair.

Let q_u be the content request generation rate of user u. For simplicity, we (i) focus our discussion on a single content, (ii) assume that all of the user's content is cached (therefore, content caching rate and the request generation rate is the same), and (iii) assume that q_us are normalized, i.e., $\sum_{u\in\mathcal{U}} q_u = 1$. Therefore, q_u can also be viewed as the probability with which the content is generated by u. We note that the content caching rate for user u at router r can be evaluated as $q_u^r = p_u^r q_u$, the total caching rate at a router is $q^r = \sum_{u\in\mathcal{U}} q_u^r$ and we also define $q_{-u}^r = \sum_{u'\neq u} q_{u'}^r$ as the cumulative content caching rate of all the users except user u at router r. We also define the vector of content generation rates as $\mathbf{q} = [q_1, q_2, \cdots, q_{|\mathcal{U}|}]^T$.

B. Game Model

We define a game $\mathcal{G}$ as a three tuple $\{\mathcal{P}, \mathcal{S}, \mathcal{F}\}$ where $\mathcal{P}$ is the set of players, $\mathcal{S}$ is the set of strategies, and $\mathcal{F}$ is the pay-off functions.

Players: The set of players correspond to $\mathcal{U}$, the set of all the users in the network.

Strategies: The set of strategies for every user is the different possible values of the probability vector $\mathbf{p}_u$. Hence, an instance of the strategies of all the users can be viewed as a strategy matrix $\mathbf{P}$ ($\mathbf{P}$ is a $|\mathcal{R}| \times |\mathcal{U}|$ matrix). Also, we denote the strategy of all users except user u by $\mathbf{P}_{-u}$.

Cost function: Let $f_u(\mathbf{p}_u)$ be the cost incurred for user u when its strategy is $\mathbf{p}_u$. Then, we have $f_u(\mathbf{p}_u) = \sum_{r\in\mathcal{R}} f_u^r(p_u^r)$, where f_u^r is the cost function of u corresponding to caching content at router r. The cost function takes two factors into consideration: first, the network resource consumed by choosing the router r which intuitively is a function of the caching capacity of r and second, the delay incurred due to choosing r. This factor depends on the location of u and r and is independent of the capacity of r. We assume that f_u is an non-decreasing convex function. We solve the game for the case when the cost function is affine, i.e., $f_u^r(p_u^r) = (\hat{c}_r q_{-u}^r + d_u^r) q_u p_u^r$. Here, $\hat{c}_r$ corresponds to the total caching capacity of r and q_{-u}^r corresponds to the current traffic at router r. Therefore, $\hat{c}_r q_{-u}^r$ represents the cost of caching at router r given the strategy $\mathbf{P}_{-u}$ of other users (expect u). The factor d_u^r corresponds to the delay cost incurred for user u while caching at router r. Let $c_u^r = (\hat{c}_r q_{-u} + d_u^r) q_u$ and $\mathbf{c}_u = [c_u^1, c_u^2, \cdots, c_u^{|\mathcal{R}|}]$. Then, the cost function can be written in vector form as $\mathbf{c}_u^T \mathbf{p}$. We consider the negative cost to be the *pay-off function* of the game.

Let U be a random variable corresponding to the event that user u has requested for the content, i.e., $P[U = u] = q_u$. Let R be a random variable corresponding to the event that the adversary observes the content cached at router r. Note that $P[R = r|U = u] = p_u^r$ and thus we have:

$$
\begin{aligned}
P[U = u, R = r] &= q_u p_u^r \\
P[R = r] &= \sum_{u'\in\mathcal{U}} P[R = r|U = u'] P[U = u'] \\
&= \sum_{u'\in\mathcal{U}} p_{u'}^r q_{u'} \\
P[U = u|R = r] &= \frac{P[R = r|U = u] P[U = u]}{P[R = r]} \\
&= \frac{q_u p_u^r}{\sum_{u'\in\mathcal{U}} p_{u'}^r q_{u'}}
\end{aligned}
$$

The mutual information between U and R is then given by:

$$
\begin{aligned}
I(U, R) &= H(U) - H(U|R) \\
&= -\sum_{u\in\mathcal{U}} q_u \log(q_u) \\
&\quad + \sum_{u\in\mathcal{U}} \sum_{r\in\mathcal{R}} q_u p_u^r \log\left(\frac{q_u p_u^r}{\sum_{u'\in\mathcal{U}} q_{u'} p_{u'}^r}\right)
\end{aligned}
$$

The user decides the level of privacy required by constraining the mutual information of the system to an upper bound. Since different users may have different privacy requirements, each user u decides on $\mathbf{p}_u$ accordingly.

C. Best Response

For any given strategy $\mathbf{P}_{-u}$ of other users, the set $\mathcal{B}_u(\mathbf{P}_{-u})$ of the best response of user u is given by

$$
\mathcal{B}_u(\mathbf{P}_{-u}) = \underset{\mathbf{p}_u\in\mathcal{S}_u}{\arg\min} f_u(\mathbf{p}_u, \mathbf{P}_{-u}). \tag{4}
$$

The best response $\overline{\mathbf{p}}_u$ of u in response to $\mathbf{P}_{-u}$ minimizes the total cost incurred for u while subjecting the mutual information of the system to an upper bound t_u. For a given $\mathbf{P}_{-u}$ and $\mathbf{q}$, $I(U, R)$ reduces to a function of $\mathbf{p}_u$ and for notational simplicity, we denote $I(U, R)$ as $I(\mathbf{p}_u)$. The best response of u can be formulated as the following optimization problem:

$$
\begin{aligned}
P0: \quad \underset{\mathbf{p}_u}{\text{minimize}} \quad & \mathbf{c}_u^T \mathbf{p}_u && (5) \\
\text{subject to} \quad & I(\mathbf{p}_u) \leq t_u, && (6) \\
& \mathbf{A}\mathbf{p}_u = \mathbf{b}, && (7) \\
& \mathbf{p}_u \geq \mathbf{0}. && (8)
\end{aligned}
$$

Here, constraint (6) is the upper bound on the leaked mutual information. Constraint (7) includes the constraint $\mathbf{1}^T\mathbf{p}_u = 1$ which along with constraint (8) ensures that $\mathbf{p}_u$ is a probability distribution. Moreover, constraint (7) allows us to accommodate for any additional equality constraints. Here, $\mathbf{A}$ is an $m \times |\mathcal{R}|$ coefficient matrix and $\mathbf{b}$ is a $m-$dimensional vector. We assume that $m \leq |\mathcal{R}|$ and that rank$(\mathbf{A}) = m$, i.e., $P1$ is feasible and has at least one positive feasible solution. Moreover, we assume that $\mathbf{c}_u$ does not belong to the column space of $\mathbf{A}^T$, i.e., $\mathbf{c}_u \notin C(\mathbf{A}^T)$ (if $\mathbf{c}_u = \mathbf{A}^T y$, it would imply that $\mathbf{c}_u^T x = \mathbf{y}^T\mathbf{A}\mathbf{p}_u = \mathbf{y}^T\mathbf{b} =$ constant).

For problem $P0$, the objective function (5) is affine. Also, since the function $I(U,R)$ is convex in $\mathbf{p}_u$ for a given $\mathbf{q}$ and $\mathbf{P}_{-u}$, constraint (6) is convex. Finally, constraints (7) and (8) are affine. Therefore, $P0$ is a convex optimization problem. For completeness, the proof for the convexity of $I(U,R)$ is presented in Appendix A.

Consider a user u with low privacy requirements and t_u chosen by it is large enough that the constraint (6) is inactive for all the feasible points. In this case, optimization problem $P1$ reduces to the following linear program:

$$P1: \quad \underset{\mathbf{p}_u}{\text{minimize}} \quad \mathbf{c}_u^T\mathbf{p}_u \tag{9}$$
$$\text{subject to} \quad \mathbf{A}\mathbf{p_u} = \mathbf{b}, \tag{10}$$
$$\mathbf{p}_u \geq \mathbf{0}. \tag{11}$$

Now, let us decrease the value of t_u and at some $t_u^{\max}$, constraint (6) becomes active, and remains active until a point $t_u^{\min}$. For any t_u less than $t_u^{\min}$, $P1$ becomes infeasible. Hence, when $t_u \in (t_u^{\min}, t_u^{\max})$, constraint (6) is active and this case is of interest to us. $t_u^{\max}$ is defined as $t_u^{\max} = I(\mathbf{p}_u^{LP}) = \min\{I(\mathbf{p}_u) | \mathbf{p}_u \text{ is an optimal solution to } P1\}$.

Now, consider the following optimization problem:

$$P2: \quad \underset{\mathbf{p}_u}{\text{minimize}} \quad I(\mathbf{p}_u) \tag{12}$$
$$\text{subject to} \quad \mathbf{A}\mathbf{p_u} = \mathbf{b}, \tag{13}$$
$$\mathbf{p}_u \geq \mathbf{0}. \tag{14}$$

Let $\mathbf{p}_u^{\min}$ be an optimal solution of $P2$. Then $I(\mathbf{p}_u^{\min})$ is the minimum possible mutual information. Hence, any t_u less than $I(\mathbf{p}_u^{\min})$ would make $P0$ infeasible, i.e., $t_u^{\min} = I(\mathbf{p}_u^{\min})$. Using the same arguments as used for $P0$, we note that $P2$ is a convex optimization problem.

Lemma 1. *For $t_u^{min} < t_u < t_u^{max}$, the optimal solution is strictly positive, i.e., $\mathbf{p}_u^* > \mathbf{0}$.*

Proof. Please refer to Appendix C. □

Theorem 1. *The optimal solution of $P1$ is given by:*

(i) $t_u \geq t_u^{max}$: the optimal solution to the linear programming, $P1$, $\mathbf{p}_u^{LP}$,

(ii) $t_u^{min} < t_u < t_u^{max}$: the following solution of $P0$,

$$p^{*r}{}_u = \frac{1}{q_u}\left(\frac{\sum_{v\neq u} q_v p_v^r}{\exp((c_u^r + \boldsymbol{\beta}^T\mathbf{a}_r)/q_u\lambda) - 1}\right), \ \forall r \in \mathcal{R}, \tag{15}$$

(iii) $t_u = t_u^{min}$: the optimal solution to convex problem $P2$, $\mathbf{p}_u^{min}$,

(iv) $t_u < t_u^{min}$: no solution,

where $\boldsymbol{\beta} \in \mathbb{R}^m, \lambda \geq 0$, and $\mathbf{a}_r$ is the r^{th} column of $\mathbf{A}$.

Proof. Please refer to Appendix D. □

D. Computing the Lagrangian Multipliers

Theorem 1 gives the form of the optimal solution, $\mathbf{p}_u^*$, for $P0$ in terms of the Lagrangian multipliers. This section focuses on case (ii) of Theorem 1 and computes the optimal values of the multipliers. Let $\mathbf{p}_u^*$ and $(\boldsymbol{\beta}^*, \lambda^*)$ be the primal and dual optimal points. Given that constraint (6) is active, we have the following from Karush-Kuhn-Tucker conditions:

$$c_u^r + \boldsymbol{\beta}^{*T} a_r + \lambda^* q_u \left[\log\left(\frac{q_u p_u^{r\,*}}{q_u p_u^{r\,*} + \sum_{v\neq u} q_v p_v^r}\right)\right] = 0, \ \forall r \in \mathcal{R}, \tag{16}$$
$$I(\mathbf{p}_u^*) - t_u = 0, \tag{17}$$
$$\mathbf{A}\mathbf{p}_u^* - \mathbf{b} = 0, \tag{18}$$
$$\boldsymbol{\beta}^* \in \mathbb{R}^m, \ \lambda^* > 0. \tag{19}$$

Solving Equation (16) for $p_u^{r\,*}$, we get

$$p^{*r}{}_u = \frac{1}{q_u}\left(\frac{\sum_{v\neq u} q_v p_v^r}{\exp((c_u^r + \boldsymbol{\beta}^{*T}\mathbf{a}_r)/q_u\lambda^*) - 1}\right), \ \forall r \in \mathcal{R}. \tag{20}$$

We have the primal optimal solution $\mathbf{p}_u^*$ as a function of the optimal dual solution $(\boldsymbol{\beta}^*, \lambda^*)$. As the optimal solution needs to satisfy the primal feasibility, i.e., Equation (17) and Equation (18), we determine the optimal dual solution $(\boldsymbol{\beta}^*, \lambda^*)$ by solving the primal feasibility equations. Here, we have $(m+1)$ equations and $(m+1)$ unknown multipliers.

Let us define the mapping $Q: \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$ as follows:

$$Q(\boldsymbol{\beta}, \lambda) = \begin{bmatrix} \mathbf{A}\mathbf{p}_u - \mathbf{b} \\ I(\mathbf{p}_u) - t_u \end{bmatrix}, \tag{21}$$

where $\mathbf{p}_u$ is given by Equation (20). By solving the system $Q(\boldsymbol{\beta}, \lambda) = 0$, we obtain the optimal dual solution. We solve $Q(\boldsymbol{\beta}, \lambda) = 0$ using the Newton-Kantorovich method. To guarantee local convergence of this method, the Jacobian Q' needs to be continuous and non-singular. Let $P(\boldsymbol{\beta}, \lambda) = \mathbf{A}\mathbf{p}_u - \mathbf{b}$ and $I(\boldsymbol{\beta}, \lambda) = I(\mathbf{p}_u)$. Then, Q' is given by:

$$Q' = \begin{bmatrix} \frac{\partial P(\boldsymbol{\beta},\lambda)}{\partial \boldsymbol{\beta}} & \frac{\partial P(\boldsymbol{\beta},\lambda)}{\partial \lambda} \\ \frac{\partial I(\boldsymbol{\beta},\lambda)}{\partial \boldsymbol{\beta}} & \frac{\partial I(\boldsymbol{\beta},\lambda)}{\partial \lambda} \end{bmatrix}. \tag{22}$$

Now, we evaluate the partial derivatives as follows:

(i) The i^{th} row of $\frac{\partial P(\boldsymbol{\beta},\lambda)}{\partial \boldsymbol{\beta}}$ can be written as

$$\left(\frac{-1}{q_u\lambda}\right)\sum_r a_{ir} w_r a_r^T,$$

where w_r is given by:

$$w_r = \frac{1}{q_u \sum_{v\neq u} q_v p_v^r}\left(q_u p_u^r (q_u p_u^r + \sum_{v\neq u} q_v p_v^r)\right).$$

Therefore, we get

$$\frac{\partial P(\boldsymbol{\beta}, \lambda)}{\partial \boldsymbol{\beta}} = \left(\frac{-1}{q_u\lambda}\right)\mathbf{A}\mathbf{W}\mathbf{A}^T \tag{23}$$

where $W = \text{diag}([w_1, w_2, \cdots, w_{|\mathcal{R}|}])$.

(ii) Similarly, $\frac{\partial P(\boldsymbol{\beta},\lambda)}{\partial \lambda}$ is given as follows:

$$\frac{\partial P(\boldsymbol{\beta},\lambda)}{\partial \lambda} = \left(\frac{-1}{q_u \lambda}\right) \mathbf{A}\mathbf{y} \tag{24}$$

where y_r is given by

$$y_r = \frac{1}{\sum\limits_{v \neq u} q_v p_v^r} \left[q_u p_u^r (q_u p_u^r + \sum_{v \neq u} q_v p_v^r)\right] \log\left[\frac{q_u p_u^r}{q_u p_u^r + \sum\limits_{v \neq u} q_v p_v^r}\right].$$

(iii) Along the same lines, we get:

$$\frac{\partial I(\boldsymbol{\beta},\lambda)}{\partial \boldsymbol{\beta}} = \left(\frac{-1}{q_u \lambda}\right) \mathbf{y}^T \mathbf{A}^T, \tag{25}$$

$$\frac{\partial I(\boldsymbol{\beta},\lambda)}{\partial \lambda} = \left(\frac{-1}{q_u \lambda}\right) \mathbf{y}^T \mathbf{W}^{-1} \mathbf{y}. \tag{26}$$

Now, we can write Q' as follows:

$$Q'(\boldsymbol{\beta},\lambda) = \frac{-1}{q_u \lambda} \begin{bmatrix} \mathbf{A}\mathbf{W}\mathbf{A}^T & \mathbf{A}\mathbf{y} \\ \mathbf{y}^T\mathbf{A}^T & \mathbf{y}^T\mathbf{W}^{-1}\mathbf{y} \end{bmatrix} \tag{27}$$

Theorem 2. *If* $\mathbf{c}$ *not in the column space of* $\mathbf{A}^T$*, i.e.,* $\mathbf{c}_u \notin C(\mathbf{A}^T)$ *then* $Q'(\boldsymbol{\beta}, \boldsymbol{\lambda})$ *is a non-singular matrix.*

Proof. Please refer to Appendix E. □

Theorem 3. *Let* $(\boldsymbol{\beta}^*, \lambda^*)$ *be a solution to* $Q(\boldsymbol{\beta}, \lambda)$ *such that* $\lambda > 0$ *and* $\mathbf{c}_u \notin C(\mathbf{A}^T$*. Then* $(\boldsymbol{\beta}^*, \lambda^*)$ *is a point of attraction of the following Newton Kantorovich method:*

$$\begin{bmatrix} \boldsymbol{\beta}^{i+1} - \boldsymbol{\beta}^{i+1} \\ \lambda^{i+1} - \lambda^i \end{bmatrix} = \frac{-1}{q_u \lambda^i} \begin{bmatrix} \mathbf{A}\mathbf{W}^i\mathbf{A}^T & \mathbf{A}\mathbf{y}^i \\ (\mathbf{A}\mathbf{y}^i)^T & (\mathbf{y}^i)^T[\mathbf{W}^i]^{-1}\mathbf{y}^i \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{A}\mathbf{p_u}^i - b \\ I(\mathbf{p}_u^i) - t_u \end{bmatrix} \tag{28}$$

where

$$w_r^i = \frac{1}{q_u \sum\limits_{v \neq u} q_v p_v^r} \left(q_u (p_u^r)^i (q_u (p_u^r)^i + \sum_{v \neq u} q_v p_v^r)\right) \tag{29}$$

$$y_r^i = \frac{1}{\sum\limits_{v \neq u} q_v p_v^r} \left(q_u (p_u^r)^i (q_u (p_u^r)^i + \sum_{v \neq u} q_v p_v^r)\right) . \log\left(\frac{q_u (p_u^r)^i}{q_u (p_u^r)^i + \sum\limits_{v \neq u} q_v p_v^r}\right) \tag{30}$$

$$(p_u^r)^i = \frac{1}{q_u} \left(\frac{\sum_{v \neq u} q_v p_v^r}{\exp((c_u^r + (\boldsymbol{\beta}^i)^T \mathbf{a}_r)/q_u \lambda^i) - 1}\right) \tag{31}$$

and i *represents the* i^{th} *iteration of the method.*

Proof. From Lemma 1 we have $\mathbf{p}_u > 0$. We also have $\lambda > 0$ and we assume $\mathbf{c}_u \notin C(\mathbf{A}^T)$. Hence, from Theorem 2, the Jacobian Q' is continuous and non-singular. $\lambda > 0$ is in a neighborhood of λ^* as $\lambda* > 0$. Then $(\boldsymbol{\beta}^*, \lambda^*)$ is a point of attraction [40]. □

V. Differentially Private Caching Games

This section enhances the caching game discussed in Section IV. We add differential privacy constraints to the existing game model so that for user u, the probability of choosing two routers with similar cost is similar. As in Section IV-C, we formulate the differentially private best response of u as an optimization problem with the objective to minimize the cost incurred subject to the mutual information **and** differential privacy constraints as follows:

$$\begin{aligned} DP0: \quad \underset{\mathbf{p}_u}{\text{minimize}} \quad & \mathbf{c}_u^T \mathbf{p}_u && (32) \\ \text{subject to} \quad & I(\mathbf{p}_u) \leq t_u, && (33) \\ & \frac{p_u^r}{p_u^{r'}} \leq e^{\epsilon_u |c_r - c_{r'}|} \quad \forall r, r' \in \mathcal{R}, && (34) \\ & \mathbf{1}^T \mathbf{p_u} = 1, && (35) \\ & \mathbf{p}_u \geq \mathbf{0}. && (36) \end{aligned}$$

Here, constraint (34) ensures that differential privacy is satisfied for the choice of routers for user u. ϵ_u is the differential privacy parameter for user u and the difference of the cost factors $|c_r - c'_r|$ is taken as the distinguishing metric between routers r and r'. From constraint (34), we can see that ϵ_u plays an important role in determining $\mathbf{p}_u$. Therefore, assuming ϵ_u as a given input parameter is not ideal. Rather, ϵ_u needs to be chosen such that the constraints of $DP0$ are satisfied. One approach to address this problem is to consider both $\mathbf{p}_u$ and ϵ_u as the decision variables of $DP0$. In this case, we can observe that $DP0$ is no longer a convex optimization problem and is difficult to solve.

We take the following **two step approach** to determine $\mathbf{p}_u$: **step 1:** determine a feasible value of ϵ_u by solving the equation $I(\mathbf{p}_u) = t_u$ and **step 2:** solve $DP0$ to obtain the best response of u using the feasible ϵ_u.

A. Determining ϵ_u

As the differential privacy constraint (34) holds for any mechanism that satisfies ϵ_u-differential privacy, we use the exponential mechanism [39] to solve $I(\mathbf{p}_u) = t_u$ and obtain a feasible ϵ_u. According to the exponential mechanism, $p_u^r \propto e^{-\epsilon_u c^r / 2\Delta u}$, i.e., we consider the negative cost coefficient c^r as the utility. Therefore, using constraint (35), we get

$$p_u^r = \frac{e^{-\epsilon_u c^r / 2\Delta u}}{\sum\limits_{r' \in \mathcal{R}} e^{-\epsilon_u c^{r'} / 2\Delta u}}. \tag{37}$$

Using $\mathbf{p}_u$ from Equation (37), we solve $I(\mathbf{p}_u) = t_u$ for ϵ_u, e.g., by using the Netwon-Raphson method. It can be verified that the corresponding Jacobian is non-zero.

Remark. A vice versa relation also exists, i.e., for a given ϵ_u, we can identify a feasible value of t_u. To maintain consistency with the model in Section IV-C, we evaluate ϵ_u for a given t_u. We also note that it is possible that a user has an independent choice of the differential privacy parameter irrespective of the mutual information parameter. Let us denote that by ϵ'_u. Then we can use $\hat{\epsilon}_s$ as the differential privacy parameter in DP0 where $\hat{\epsilon}_u = \min(\epsilon_u, \epsilon'_u)$. For notational simplicity, we denote the differential privacy parameter as ϵ_u.

B. Best Response

Next, we solve $DP0$ using dual conjugate theory to determine the best response of u. Similar to Section IV-C, we focus on the case where $t_u^{\min} < t_u < t_u^{\max}$ (since a non-linear constraint exists in this case and the optimization problem is linearly constrained for other cases). First, we define the following:

Definition 1. *Conjugate function.* Let $f : \mathbb{R}^n \to \mathbb{R}$ and dom f be its domain. Then the conjugate function $f^* : \mathbb{R}^n \to \mathbb{R}$ is defined as follows:

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}\in \text{dom } f} (\mathbf{y}^T\mathbf{x} - f(\mathbf{x})) \tag{38}$$

$$\text{dom } f^* = \{\mathbf{y} \mid \sup_{\mathbf{x}\in \text{dom } f} (\mathbf{y}^T\mathbf{x} - f(\mathbf{x})) < \infty,\ \mathbf{y} \in \mathbb{R}^n\} \tag{39}$$

Note that $f(\mathbf{x}) + f^*(\mathbf{y}) \geq \mathbf{y}^T\mathbf{x},\ \forall \mathbf{x} \in \text{dom } f,\ \forall \mathbf{y} \in \text{dom } f^*$. Moreover, the equality occurs for the sub-gradient sets, i.e.,

$$f(\mathbf{x}) + f^*(\mathbf{y}) = \mathbf{y}^T\mathbf{x},\ \text{if } \mathbf{y} \in G(\mathbf{x}) \text{ or } \mathbf{x} \in G^*(\mathbf{y}) \tag{40}$$

where $G(\mathbf{x})$ and $G^*(\mathbf{y})$ are the sub-gradient of f and f^* at $\mathbf{x}$ and $\mathbf{y}$, respectively. Next, we define the positive homogeneous extension of f^* as follow:

Definition 2. *Positive homogeneous extension.* Let $f^* : \mathbb{R}^n \to \mathbb{R}$ and dom f^* be its domain. Then the positive homogeneous extension $f^+ : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is defined as

$$f^+(\mathbf{y}, \lambda) = \begin{cases} \lambda f^*(\mathbf{y}/\lambda),\ \lambda > 0 \\ \sup\limits_{\mathbf{x}\in \text{dom } f^*} \mathbf{y}^T\mathbf{x},\ \lambda = 0 \end{cases} \tag{41}$$

$$\begin{aligned}\text{dom } f^+ = \{(\mathbf{y}, \lambda) | \lambda = 0,\ \sup_{\mathbf{x}\in \text{dom } f^*} (\mathbf{y}^T\mathbf{x}),\ \mathbf{y} \in \mathbb{R}^n\} \\ \cup \{(\mathbf{y}, \lambda) | \lambda > 0,\ \mathbf{y}/\lambda \in \text{dom } f^*\}.\end{aligned} \tag{42}$$

In order to compute the dual of $DP0$, we first formulate the following equivalent optimization problem:

$$\begin{aligned} DP1: \quad \underset{\mathbf{p}_u, \mathbf{x}, \mathbf{z}}{\text{minimize}} \quad & \mathbf{c}_u^T\mathbf{p}_u && (43)\\ \text{subject to} \quad & I(\mathbf{x}) \leq t_u, && (44)\\ & \mathbf{A}\mathbf{p}_u - \mathbf{z} \geq 0, && (45)\\ & \mathbf{z} = \mathbf{b}, && (46)\\ & \mathbf{p}_u - \mathbf{x} = 0, && (47)\\ & \mathbf{p}_u \geq \mathbf{0}, && (48)\end{aligned}$$

where $\mathbf{A}\mathbf{p}_u - \mathbf{b} \geq 0$ contains constraints (34) and (35) (i.e., the differential privacy and probability distribution constraint, respectively).

The objective function of the dual program is computed as the sum of the conjugate function of the primal objective, Equation (43), subject to constraint (46) and the positive homogeneous extension of constraint (44). The former is given as

$$\sup_{\substack{\mathbf{p}_u \\ \mathbf{z}=\mathbf{b}}} \mathbf{s}^T\mathbf{p}_u + \mathbf{z'}^T\mathbf{z} - \mathbf{c}_u^T\mathbf{p}_u = \mathbf{b}^T\mathbf{z'}. \tag{49}$$

Here $\mathbf{s}$ and $\mathbf{z'}$ are conjugate dual variables of $\mathbf{p}_u$ and $\mathbf{z}$, respectively. The above holds as the supremum occurs at $\mathbf{s} = \mathbf{c}_u$.

Let $\mathbf{y}$ be the conjugate dual variable of $\mathbf{x}$. Then, the conjugate transform of $I(\mathbf{x}) - t_u$ is given by

$$\begin{aligned}-\sum_{r\in\mathcal{R}} \left(\sum_{\substack{v\in\mathcal{U} \\ v\neq u}} q_v p_v^r \left[\log(1 - e^{y_r/q_u}) + \log\left(\frac{q_v p_v^r}{\sum\limits_{v'\neq u} q'_v p'^r_v} \right) \right] \right) \\ + \sum_{u\in\mathcal{U}} q_u \log(q_u) + t_u\end{aligned} \tag{50}$$

which has the positive homogeneous extension as follows:

$$-\lambda \sum_{r\in\mathcal{R}} \left(\sum_{\substack{v\in\mathcal{U} \\ v\neq u}} q_v p_v^r \left[\log(1 - e^{y_r/\lambda q_u}) + \log\left(\frac{q_v p_v^r}{\sum\limits_{v'\neq u} q'_v p'^r_v} \right) \right] + t_u \right) \tag{51}$$

for $\lambda > 0,\ \lambda \in \mathbb{R}$ (for completeness, we provide the derivation in Appendix B). The case that $\lambda = 0$ renders constraint (44) inactive. Further, from Equation (40), the primal and dual variables are related as

$$x_r = \frac{\sum_{v\neq u} q_v p_v^r}{\lambda q_u (1 - e^{y_r/\lambda q_u})} \tag{52}$$

and the objective of the dual program is

$$\begin{aligned}f(\mathbf{z'}, \mathbf{y}, \lambda) = \mathbf{b}^T\mathbf{z'} \\ - \lambda \sum_{\substack{r\in\mathcal{R} \\ v\in\mathcal{U} \\ v\neq u}} \left(q_v p_v^r \left[\log\left(\frac{q_v p_v^r (1 - e^{y_r/\lambda q_u})}{\sum\limits_{v'\neq u} q'_v p'^r_v} \right) \right] + t_u \right)\end{aligned} \tag{53}$$

subject to $\mathbf{s} = \mathbf{c}_u,\ \mathbf{z} = \mathbf{b},\ \lambda \geq 0$.

Finally, the feasible set of the dual program is the dual cone of the primal cone generated by constraints (45), (47), and (48). Let $\boldsymbol{\alpha}, \boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ be the Lagrangian multipliers for constraints (45), (47), and (48), respectively. Then, using Definition 1, we have $\mathbf{s} = \mathbf{A}^T\boldsymbol{\alpha} + \boldsymbol{\delta} + \boldsymbol{\gamma}$, $\mathbf{y} = -\boldsymbol{\gamma}$, and $\mathbf{z'} = -\boldsymbol{\alpha}$. Here $\boldsymbol{\alpha} \geq \mathbf{0}$, $\boldsymbol{\gamma} \in \mathbb{R}^{|\mathcal{R}|}$, and $\boldsymbol{\delta} \geq \mathbf{0}$.

Now, we can form the dual program of $DP1$ as follows:

$$\begin{aligned} DP2: \quad \underset{\boldsymbol{\alpha}, \boldsymbol{\gamma}, \lambda}{\text{minimize}} \quad & f(-\boldsymbol{\alpha}, -\boldsymbol{\gamma}, \lambda) && (54)\\ \text{subject to} \quad & \mathbf{A}^T\boldsymbol{\alpha} - \mathbf{c}_u + \boldsymbol{\gamma} \leq 0, && (55)\\ & \boldsymbol{\alpha} \geq 0, && (56)\\ & \lambda \geq 0, && (57)\\ & \boldsymbol{\gamma} \in \mathbb{R}^{|\mathcal{R}|}. && (58)\end{aligned}$$

$DP2$ is a linearly constrained convex optimization problem and as it satisfies Slater's condition. Thus, strong duality holds and $\mathbf{p}_u^*$ can be computed using (52).

Remark. In Equation (7) and Equation (35), we can modify the constraint $\mathbf{1}^T\mathbf{p}_u = 1$ to $\mathbf{1}^T\mathbf{p}_u = n$. This ensures that the expected number of copies of the content in the network is n. The caching probabilities at the routers are computed such that expected number of copies in the network will be n. The computational complexity for a system with n copies of the content is the same as the computational complexity for scenarios with only one copy. Therefore, our proposed caching

Algorithm 1: Iterative Best Response (IBR) algorithm

Iterative best (IBR) response **Input:** $\mathcal{G}$
Output: $\mathbf{p}_u^{NE}$ (Nash Equilibrium solution)
Randomly choose a feasible $\mathbf{p}_u^0$
$Update(\mathbf{P})$
do
  $Get(\mathbf{P}_{-u})$
  $\mathbf{P}_{prev} \leftarrow \mathbf{P}$
  $\mathbf{p}_u \leftarrow OptimalResponse(\mathbf{P}_{-u})$
  $Update(\mathbf{P})$
while $\mathbf{P} \neq \mathbf{P}_{prev}$;

strategy can be easily extended to cache multiple copies of the content in the network.

VI. Nash Equilibrium

This section shows that Nash equilibrium exists for our games and presents an iterative best-response algorithm to compute it.

Definition 3. Nash Equilibrium: A caching probability matrix $\mathbf{P}$ is a Nash equilibrium if and only if

$$f_u(\mathbf{p}_u, \mathbf{P}_{-u}) \leq f_u(\mathbf{p}'_u, \mathbf{P}_{-u}),\ \forall \mathbf{p}'_u \in \mathcal{S}_u,\ \forall u \in \mathcal{U}. \quad (59)$$

Theorem 4. *For game $\mathcal{G}$, a pure strategy Nash equilibrium (PNE) exists.*

Proof. We can observe from $P2$ and $DP0$ that the strategy space of $\mathcal{G}$ is compact, convex, and non-empty and the cost function is convex. Hence, a pure strategy Nash equilibrium exists for $\mathcal{G}$ [41]. □

We use Algorithm 1 (IBR) to compute the pure strategy Nash equilibrium. Here, during the i^{th} update round, all consumers compute their optimal response based on the strategies of other consumers. This procedure is continued until the strategies converge to a Nash equilibrium. We observe that the optimal response of u does not depend directly on the strategy of the other individual consumers. Rather, it depends on a few aggregate values (e.g., $\sum q_v p_v^r$ and $\sum q_v p_v^r \log(q_v p_v^r)$).

Theorem 5. *The iterative best response algorithm (IBR) converges to a pure strategy Nash equilibrium (PNE).*

Proof. The optimal response of u is subject to the constraint $I(\mathbf{p}_u) < t_u$. Hence, $\mathcal{G}$ is a game of strategic substitutes (with convex strategy sets) because if a consumer with lower privacy requirement increases the mutual information of the system (to minimize its cost), then an another consumer with higher privacy requirement substitutes for it (to meet its own privacy constraint). Moreover, the computation of best response in each round can be performed simultaneously in IBR. Then, under the assumption that all best response correspondences are single valued, IBR converges to a PNE [42]. □

Remark. A user using our privacy aware caching strategy does not need to know the privacy needs and strategies of every other user in the system. Rather, the user only requires to know a few aggregate values like $\sum q_v p_v^r$ and $\sum q_v p_v^r \log(q_v p_v^r)$ (Equations (29)-(31) and Equations (51)-(53)). These aggregate values do not leak the individual privacy requirements and strategies of other users. We assume that all the users send their individual strategies to the network operator and the network operator broadcasts the aggregate values to all the users.

As an alternative, all the users may provide their privacy requirements to the network operator. The network operator then functions as a neutral entity and computes the strategies of all the users in the network. For both of the above mentioned approaches, a Nash Equilibrium exists and the communication overhead is $O(n)$ messages, where n is number of users in the network.

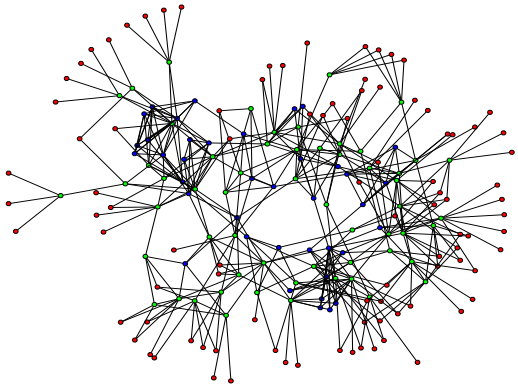
For the case when a user does not have complete information about the other users' strategies, the user considers a probability distribution over the possible strategies of other users. Then, the user aims to minimize her expected cost (instead of the exact cost). Then, the non-cooperative game in Section VI becomes a non-cooperative Bayesian game and a Bayesian Nash Equilibrium exists for such a game.

VII. Performance Evaluation

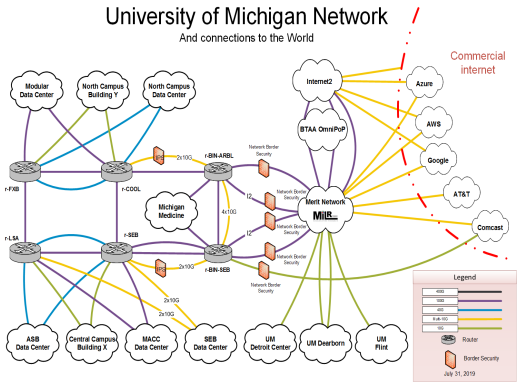
This section evaluates the performance of the proposed caching strategies. First we compare the performance of our proposed caching strategies with the approaches proposed in [23], [24]. The approach in [23] preserves privacy by caching content only at the routers with a minimum betweenness centrality. In [24], the routers ensure privacy by collaborating with each other to increase the uncertainty of the adversary. For every approach, we plot the cost incurred as a function of the privacy. The privacy is measured in terms of the anonymity set size of the router. Anonymity set size is a well studied privacy metric [43]–[46]. It is a natural choice of privacy metric for comparing different approaches as the adversary's certainty to breach users' privacy decreases with the increasing size of the anonymity set.

We abbreviate the solution of the caching game and the differentially private caching game as NC and NDC, respectively. The betweenness centrality approach and the collaborative caching approach are abbreviated as BC and CC, respectively. Then, to further validate our caching strategies, we compare them with the global optimal strategy for caching game, the differentially private caching game, and the exponential mechanism in Equation (37) (abbreviated as GC, GDC and EDC, respectively). The objective of the global optimal solution is to minimize the sum of costs of all the consumers subject to the privacy constraints.

We consider two different network topologies for our evaluation. The first topology is obtained from Rocketfuel network topology traces for ISP Exodus [47] and the second one is a campus network topology obtained from the University of Michigan [48]. The topologies are depicted in Figure 2a and Figure 2b, respectively. For the Exodus topology in Figure 2a, the total delay incurred between two backbone routers is assumed to be 20 ms, a backbone and a gateway router is 10 ms, and two gateway routers is 10 ms. In Figure 2b, the



(a) Exodus topology



(b) University of Michigan campus topology

Fig. 2. Network topologies used for performance study. (a) The blue, green and red nodes are backbone, gateway, and access routers, respectively. (b) The six routers (grey in color) form the core routers of the network and gateway routers in different parts of the campus network connect to the core routers.

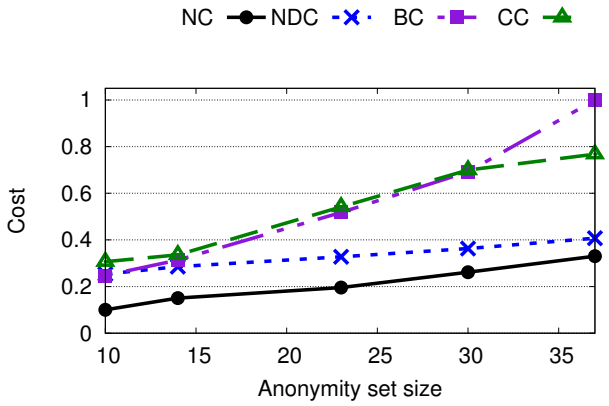


Fig. 3. Effect of anonymity set size on the cost for Rocketfuel topology.

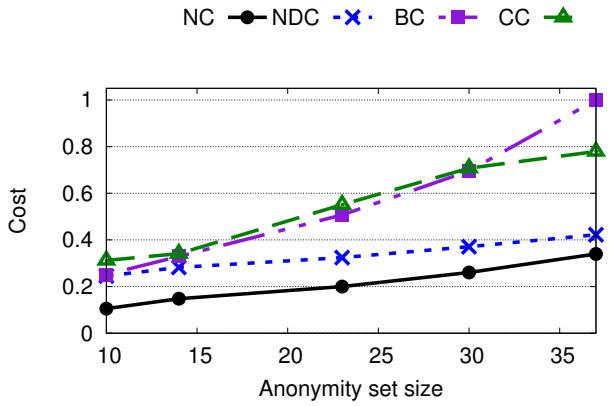


Fig. 4. Effect of anonymity set size on the cost for campus topology.

delay between the core routers is assumed to be 6 ms and the delay between the core and gateway routers is assumed to be 3 ms. We assume q_u to be uniformly distributed. The cost values presented in our evaluations are the average cost per consumer and are normalized in the range $[0, 1]$.

The performance evaluations are performed using Named Data Networks (NDN). We assume a payload size of 1024 bytes, the cache size of each router to be 1 GB, and the number of requests generated by the consumers is randomly selected from the range $[2, 20]$ requests per second. We would like to note that our caching strategy can be used for any ICN architecture which deploys caches at its routers. The aim of the adversary is to breach the privacy of the users. To do so, the adversary observes the cache of the routers and then tries to identify the user who requested the content. To enhance the user's privacy, our proposed caching strategy aims to increase the uncertainty of the adversary's inference.

For a uniform and fair comparison, we consider the anonymity set size as the privacy metric. Here we consider a network with 40 consumers. To evaluate the impact of number of consumers in the network, we vary the number of consumers in the network in our further evaluations. Figure 3 and Figure 4 depict the cost incurred by NC, NDC, BC and CC as the anonymity set size increases from 10 to 37. In general, the cost for all the approaches increases with the increasing anonymity set size. This is because all the approaches cache content at routers farther away from the consumers to achieve a bigger anonymity set size. Moreover, as more consumers start caching at the same router for increasing the anonymity set size, the cost of caching at that router increases, thereby increasing the cost for all the consumers. Comparing NC and NDC with BC and CC, we observe that for the same anonymity set size, BC and CC incur higher cost than NC and NDC. For example, in the Exodus case with the anonymity set size of 14, the cost incurred for NC, NDC, BC and CC are 0.15, 0.28, 0.31 and 0.33, respectively. For the same topology with a set size of 37, the cost incurred for NC, NDC, BC and CC are 0.33, 0.40, 1 and 0.76, respectively. We can observe that the cost difference between our proposed approaches (NC and NDC) and the existing approaches (BC and CC) increases significantly as the privacy (anonymity set size) increases. As BC is a betweenness centrality based approach, BC users the more centrally placed routers which are expensive. CC is a collaborative approach which randomly caches at any one of the router in the network. While such an approach ensures privacy, it is oblivious to the caching cost. In contrast, NC and NDC achieve the same level of privacy by leveraging off-path caching where less expensive and closer routers can be used to increase the set size.

Here, to increase the anonymity set size of NC and NDC, we

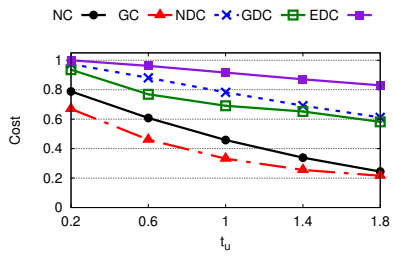


Fig. 5. Effect of t_u on the cost for Rocketfuel topology.

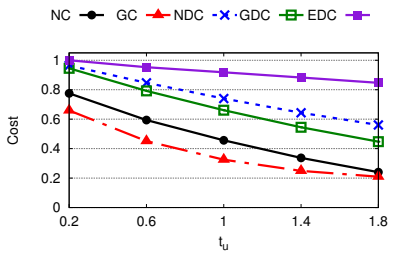


Fig. 6. Effect of t_u on the cost for campus topology.

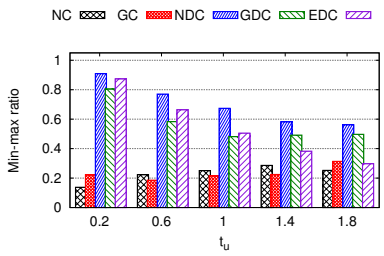


Fig. 7. Effect of t_u on cache utilization for Rocketfuel topology.

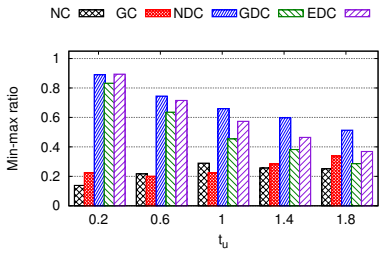


Fig. 8. Effect of t_u on cache utilization for campus topology.

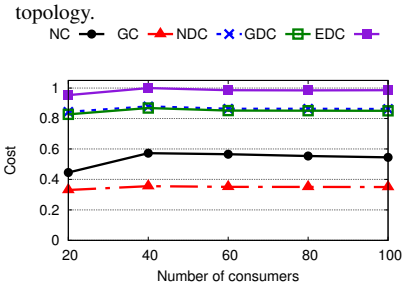


Fig. 9. Effect of number of consumers on the cost for Rocketfuel topology.

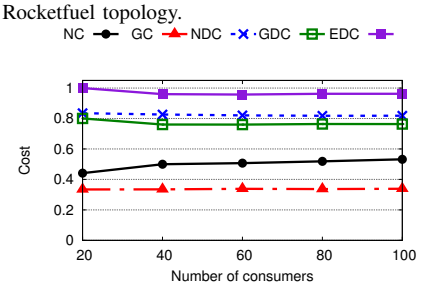


Fig. 10. Effect of number of consumers on the cost for campus topology.

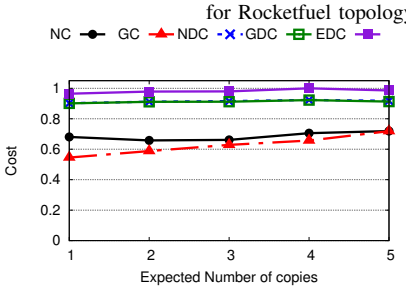


Fig. 11. Effect of number of content copies on the cost for Rocketfuel topology.

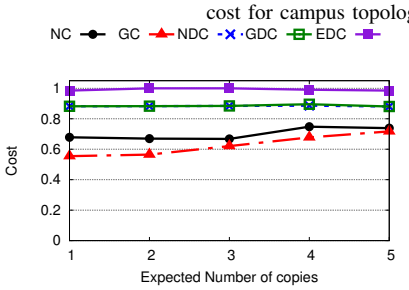


Fig. 12. Effect of number of content copies on the cost for campus topology.

decrease the t_u value from 1.8 to 0.2. Now, we compare NC and NDC with GC, GDC and EDC as the value of t_u varies from 0.2 to 1.8. Figure 5 and Figure 6 show the cost incurred for both the topologies. We observe that for all the approaches the cost incurred decreases as the t_u values increases (i.e. the privacy decreases). For larger values of t_u, the consumers can choose less expensive routers with higher probability, thereby reducing the average cost incurred. Due to the distributed and selfish nature of NC and NDC, they incur more cost as compared to GC and GDC. This is expected as GC and GDC are global optimal solutions. The difference in cost for NDC and GDC is smaller compared to the difference between NC and GC because of the differential privacy constraints which renders the strategies of NDC and GDC quite similar to each other. For example, in the Exodus topology with $t_u = 0.2$, NC and GDC incur a cost of 0.76 and 0.67, respectively while NDC and GDC incur a cost of 0.97 and 0.93, respectively. Moreover, as the privacy requirements (i.e., increasing t_u) are relaxed, both NC and NDC tend to converge towards GC and GDC, respectively. We also observe that NDC performs significantly better than EDC, especially as t_u increases. For example, NDC, GDC and EDC incur a cost of 0.69, 0.68 and 0.87, respectively, when $t_u = 1.4$.

Efficient usage of network resources is an important objective of any network operator. Therefore, we study how balanced are our caching approaches. We plot the ratio of the minimum (caching) traffic to the maximum (caching) traffic at the routers. Figure 7 and Figure 8 depict the ratios as the value of t_u increases. We observe that in general, NDC, GDC and EDC are more balanced when compared to NC and GC. This is due to the differential privacy constraints where the traffic at different routers are not independent but are dependent on each other and on the value of ϵ_u. As t_u increases, this constraint relaxes and therefore, the traffic balance in general decreases for NDC, GDC and EDC. We also see that among NDC, GDC and EDC, NDC generally achieves a better balance and this difference is more significant as t_u increases. For NC and GC, it is difficult to observe a pattern because the traffic at different routers are only required to satisfy the mutual information constraint and the traffic is independent otherwise.

Next we study impact of the number of consumers on the performance of our approaches. Figure 9 and Figure 10 depict the cost incurred by all the approaches as the number of consumers increase as 20, 40, 60, 80, and 100. Here, the t_u values for the consumers lie in the range $[0.9, 1]$. We again see that NC incurs slightly more cost compared to GC owing to its distributed and selfish approach while NDC incurs cost very similar to GDC. Interestingly, we observe that the variation in the costs is small as the number of consumers increases. This implies that the cost incurred for a consumer depends on the

privacy requirements and not on the number of consumers in the network.

Next, we study the impact of increasing the number of copies of the content. Here we consider 40 consumers in the network with $t_u \in [0.9, 1]$. We vary the number of copies as $1, 2, \cdots, 5$. We achieve this by slightly modifying $P0$ and $DP0$ by including a constraint that the expected number of copies of the content is $i, i \in \{1, 2, \cdots, 5\}$. We plot the average cost incurred per consumer per copy of the content in Figure 11 and Figure 12. We observe that the cost incurred per copy increases slightly as the number of copies increases (evidently for NC and GC). For example, NC incurs a cost of 0.65, 0.66, and 0.70 when the number of copies increases as 2, 3, and 4, respectively. This is because the cost of caching depends on the traffic at the router and as the number of copies increases, the traffic increases too. While the slight increase encourages the network operator to cache multiple copies of the content, the network operator needs to trade-off between the quality of experience and redundancy occurring due to the presence of multiple copies of the same content.

VIII. Conclusion

This paper proposed a private off-path caching mechanism to mitigate the timing analysis attacks in ICN. The caching strategy is computed in a distributed manner by formulating a non-cooperative game among the users in the network. The Nash equilibrium of the game is computed using an iterative best response-algorithm. Our results show that the proposed approach significantly outperforms the existing approaches and achieves near optimal results and better cache utilization when compared to the global solutions.

Appendix

A. Convexity of $I(U, R)$

The first order derivative of $I(U, R)$ with respect to p_u^r is evaluated as

$$\begin{aligned}\frac{\partial I}{\partial p_u^r} &= \frac{\partial}{\partial p_u^r}\left[q_u p_u^r \log\left(\frac{q_u p_u^r}{\sum_{u'\in\mathcal{U}} q'_u p_{u'}^r}\right) + \sum_{u'\neq u} q_{u'} p_{u'}^r \log\left(\frac{q_{u'} p_{u'}^r}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r}\right)\right] \\ &= \frac{\partial}{\partial p_u^r}\left[q_u p_u^r\Big(\log(q_u p_u^r) - \log(q_u p_u^r + \sum_{v\neq u} q_v p_v^r)\Big) + \sum_{u'\neq u} q_{u'} p_{u'}^r\Big(\log(q_{u'} p_{u'}^r) - \log(q_u p_u^r + \sum_{v\neq u} q_v p_v^r)\Big)\right] \\ &= q_u\left[\log(q_u p_u^r) + 1 - \log(q_u p_u^r + \sum_{v\neq u} q_v p_v^r) - \frac{q_u p_u^r}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r} - \sum_{u'\neq u}\frac{q_{u'} p_{u'}}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r}\right] \\ &= q_u\left[\log(q_u p_u^r) - \log(q_u p_u^r + \sum_{v\neq u} q_v p_v^r)\right] \end{aligned} \tag{60}$$

From Equation (60), we have the following second order derivatives:

$$\frac{\partial^2 I(U,R)}{\partial {p_u^r}^2} = {q_u}^2\left(\frac{1}{q_u p_u^r} - \frac{1}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r}\right),\ \forall r \in \mathcal{R} \tag{61}$$

$$\frac{\partial^2 I(U,R)}{\partial p_u^r \partial p_u^{r'}} = 0, \qquad \forall r \neq r'; r, r' \in \mathcal{R} \tag{62}$$

As $q_u p_u$ and $\sum_{v\neq u} q_v p_v^r$ are non negative, from Equation (61), $\frac{\partial^2 I(U,R)}{\partial {p_u^r}^2}$ is non negative for all r. Therefore, the Hessian matrix of $I(U, R)$ has non-negative diagonal elements. From Equation (62), the non-diagonal elements of the Hessian matrix are 0. Therefore the Hessian matrix is positive semi-definite and $I(U, R)$ is a convex function of $\mathbf{p}_u$.

B. Conjugate function of $I(x) - t_u$

Let $I^*(y)$ be conjugate of $I(x) - t_u$. Then $I^*(y)$is given by:

$$\begin{aligned} I^*(\mathbf{y}) &= \sup_{\mathbf{x}}(\mathbf{y}^T\mathbf{x} - (I(x) - t_u)) \\ &= \sup_{\mathbf{x}}\Bigg[\mathbf{y}^T\mathbf{x} - \Bigg(-\sum_{u\in\mathcal{U}} q_u \log(q_u) + \sum_{u\in\mathcal{U}}\sum_{r\in\mathcal{R}} q_u x_r \log\left(\frac{q_u x_r}{\sum_{u'\in\mathcal{U}} q_{u'} p_{u'}^r}\right) - t_u\Bigg)\Bigg]. \end{aligned} \tag{63}$$

Using the first order derivative, the supremum occurs at

$$x_r = \frac{\sum_{u'\in\mathcal{U}} q_{u'} p_{u'}^r}{q_u(e^{-y_r/q_u} - 1)}.$$

Substituting the value of x_r in Equation (63), we get

$$I^*(\mathbf{y}) = -\sum_{r\in\mathcal{R}}\left(\sum_{\substack{v\in\mathcal{U}\\ v\neq u}} q_v p_v^r\Bigg[\log(1 - e^{y_r/q_u}) + \log\left(\frac{q_v p_v^r}{\sum_{v'\neq u} q'_v p_v'^r}\right)\Bigg]\right) + \sum_{u\in\mathcal{U}} q_u \log(q_u) + t_u.$$

Computing the positive homogeneous extension is straightforward by replacing y_r with y_r/λ for $\lambda > 0$.

C. Proof of Lemma 1

As $t_u < t_u^{\max}$, we have $\mathbf{c}_u^T\mathbf{p}_u^* > \mathbf{c}_u^T\mathbf{p}_u^{LP}$ by construction. Let us consider $\mathbf{y} > \mathbf{0}$ such that $\mathbf{1}^T\mathbf{y} = 1$. For some $k \in [0, 1]$ and $l \in [0, 1]$, we define

$$\mathbf{s} = (1 - l)\mathbf{p}_u^* + l[k\mathbf{p}_u^{LP} + (1 - k)y]. \tag{64}$$

We can easily see that for $k < 1$ and $l > 0$, $z > \mathbf{0}$. If k is sufficiently close to 1 and $l > 0$, then $\mathbf{c}_u^T\mathbf{z} < \mathbf{c}_u^T\mathbf{p}_u^*$. As the slope of $I(\mathbf{p}_u)$ is infinite at the boundary of the feasible domain, if $p_u^{r*} = 0$ we have $I(\mathbf{z}) < I(p_u^*)$ for $0 < k < 1$ and sufficiently small $t > 0$. Thus, we can construct a feasible solution $\mathbf{z}$ with $\mathbf{c}_u^T\mathbf{z} < \mathbf{c}_u^T\mathbf{p}_u^*$. This is a contradiction as $\mathbf{p}_u^*$ is an optimal solution. Hence $p_u^{r*} > 0, \forall r \in \mathcal{R}$.

D. Proof of Theorem 1

Case (i). For the case $t_u > t_u^{\max}$, the result is immediate as constraint (6) is inactive for all the feasible points. For the case $t_u = t_u^{\max}$, we have $I(\mathbf{p}_u^{LP}) = t_u$, i.e., $\mathbf{p}_u^{LP}$ is feasible as well as optimal (as $\mathbf{c}_u^T\mathbf{p}_u$ cannot be further minimized even if constraint (6) is removed) and case (i) follows.

Case (ii). For $t_u^{\min} < t_u < t_u^{\max}$, $\mathbf{p}_u^{\min}$ is a feasible solution to $P1$ such that $I(\mathbf{p}_u^{\min}) < t_u$. This is immediate as $I(\mathbf{p}_u^{\min}) = t_u^{\min} < t_u$. As $I(\mathbf{p}_u)$ is strictly convex, there exists at least one optimal solution $\mathbf{p}_u^*$. For $t_u^{\min} < t_u < t_u^{\max}$, constraint (6) is active for an optimal solution $\mathbf{p}_u^*$. To prove this statement, assume that $I(\mathbf{p}_u^*) < t_u$. Then, as the constraint is inactive, $\mathbf{p}_u^{LP}$ is an optimal solution. Hence, $I(\mathbf{p}_u^{LP}) \leq t_u$. However, according to this case's assumption, $I(\mathbf{p}_u^{LP}) = t_u^{\max} > t_u$. This is a contradiction. Therefore, we conclude that constraint (6) is active and therefore $I(\mathbf{p}_u^*) = t_u$.

Now, we prove (15). The Lagrangian L of $P0$ is given by:

$$L(\mathbf{p}_u, \boldsymbol{\beta}, \lambda) = \mathbf{c}_u^T\mathbf{p}_u + \lambda\left(I(\mathbf{p}_u) - t_u\right) + \boldsymbol{\beta}^T(\mathbf{A}\mathbf{p}_u - \mathbf{b}). \quad (65)$$

From Lemma 1, we note that constraint (8) is satisfied with strict inequality (i.e., constraint is inactive) and thus, it is not included in the Lagrangian. The Lagrange dual function is:

$$g(\boldsymbol{\beta}, \lambda) = \inf_{\mathbf{p}_u} \mathbf{c}_u^T\mathbf{p}_u + \lambda\left(I(\mathbf{p}_u) - t_u\right) + \boldsymbol{\beta}^T(\mathbf{A}\mathbf{p}_u - \mathbf{b}) \quad (66)$$

From Lemma 1, we observe that when $t_u^{\min} < t_u < t_u^{\max}$, there exists a point in the feasible domain that is strictly feasible. Therefore Slater's condition is satisfied. As $P0$ is convex and satisfies Slater's condition, strong duality holds and the optimal duality gap is zero. Now, using the Karush-Kuhn-Tucker conditions, the gradient of L vanishes at $\mathbf{p}_u^*$. Therefore,

$$c_u^r + \boldsymbol{\beta}^T a_r + \lambda q_u\left[\log(\frac{q_u p_u^r}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r})\right] = 0, \quad (67)$$

$\forall r \in \mathcal{R}$. Here $\boldsymbol{\beta} \in \mathbb{R}^m$ and $\lambda > 0$ since constraint (6) is active. Solving (67) for p_u^r, we get (15). The Lagrangian multipliers $\boldsymbol{\beta}$ and λ in (15) are computed in Section IV-D.

Case(iii) and Case(iv). The results of case (iii) and case (iv) are immediate.

E. Proof of Theorem 2

A necessary and sufficient condition for Q' to be singular is that any row is a linear combination of the other m rows (or the rank of $Q' < m+1$). Let us assume that Q' is singular. Then, as the rows of $\mathbf{A}$ are linearly independent, the $(m+1)^{\text{th}}$ row of Q' is a linear combination of the first m rows. Performing row operations on the $(m+1)^{\text{th}}$ row of Q', we get:

$$s = \mathbf{y}^T[\mathbf{A}^T(\mathbf{A}\mathbf{W}\mathbf{A}^T)^{-1}\mathbf{A} - \mathbf{W}^{-1}]\mathbf{y} = 0. \quad (68)$$

Let $\hat{\mathbf{A}} = \mathbf{A}W^{1/2}$ and $\hat{\mathbf{y}} = \mathbf{W}^{-1/2}\mathbf{y}$, i.e., $\hat{y}_r = \left(q_u(q_u p_u^r(q_u p_u^r + \sum_{v\neq u} q_v p_v^r))\right)^{1/2} \log\left(\frac{q_u p_u^r}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r}\right)$.

From Lemma 1, we have $p_u^r > 0\ \forall r$. Hence,

$$s = \hat{\mathbf{y}}^T[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{A}}^T)^{-1}\hat{\mathbf{A}} - \mathbf{I}]\hat{\mathbf{y}} = -\hat{\mathbf{y}}^T\mathbf{P}\hat{\mathbf{y}}. \quad (69)$$

We note that $\mathbf{P}$ is the orthogonal projection onto the nullspace of $\hat{\mathbf{A}}$. Therefore,

$$\begin{aligned} s = 0 &\Leftrightarrow \mathbf{P}\hat{\mathbf{y}} = 0 \\ &\Leftrightarrow \hat{\mathbf{y}} \in C(\hat{\mathbf{A}}^T) \\ &\Leftrightarrow \hat{\mathbf{y}} = \mathbf{X}^{1/2}\mathbf{A}^T\mathbf{k}, \text{ for some } \mathbf{k}. \end{aligned} \quad (70)$$

Simplifying Equation (16) and Equation (70), we get

$$\log\left(\frac{q_u p_u^r}{q_u p_u^r + \sum_{v\neq u} q_v p_v^r}\right) = \frac{-c_u^r - \boldsymbol{\beta}^T\mathbf{a}_r}{q_u\lambda} = \mathbf{a_r}^T\mathbf{k}'. \quad (71)$$

As $\boldsymbol{\beta}^T\mathbf{a}_r = \mathbf{a}_r^T\boldsymbol{\beta}$, we get

$$c_u^r = \mathbf{a_r}^T(-\boldsymbol{\beta} - q_u\lambda\mathbf{k}'), \text{ i.e.,} \quad (72)$$

$$\mathbf{c}_u = \mathbf{A}^T(-\boldsymbol{\beta} - q_u\lambda\mathbf{k}') \in C(\mathbf{A}^T). \quad (73)$$

Thus, $s = 0 \Leftrightarrow \mathbf{c}_u \in C(\mathbf{A}^T)$. However, we have assumed $\mathbf{c}_u \notin C(\mathbf{A}^T)$. Therefore, $s \neq 0$ and Q' is non-singular.

REFERENCES

[1] CISCO. (2020, March) Cisco Annual Internet Report (2018–2023) White Paper. [Online]. Available: "https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html"

[2] A. Chaabane, E. De Cristofaro, M. A. Kaafar, and E. Uzun, "Privacy in content-oriented networking: Threats and countermeasures," *ACM SIGCOMM Computer Communication Review*, 2013.

[3] E. G. AbdAllah, H. S. Hassanein, and M. Zulkernine, "A survey of security attacks in information-centric networking," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1441–1454, 2015.

[4] G. Acs, M. Conti, P. Gasti, C. Ghali, G. Tsudik, and C. A. Wood, "Privacy-aware caching in information-centric networking," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 2, p. 313–328, 2019.

[5] S. Saha, A. Lukyanenko, and A. Ylä-Jääski, "Cooperative caching through routing control in information-centric networks," in *Proc. of IEEE INFOCOM*, 2013, pp. 100–104.

[6] Z. Li and G. Simon, "Time-shifted tv in content centric networks: The case for cooperative in-network caching," in *Proc. of IEEE ICC*, 2011.

[7] V. Sourlas, L. Gkatzikis, P. Flegkas, and L. Tassiulas, "Distributed cache management in information-centric networks," *IEEE Transactions on Network and Service Management*, vol. 10, no. 3, pp. 286–299, 2013.

[8] O. Ascigil, V. Sourlas, I. Psaras, and G. Pavlou, "Opportunistic off-path content discovery in information-centric networks," in *Proc. of IEEE LANMAN)*, 2016.

[9] M. Zhang, H. Luo, and H. Zhang, "A survey of caching mechanisms in information-centric networking," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1473–1499, 2015.

[10] E. Bertino and M. Nabeel, "Securing named data networks: Challenges and the way forward," in *Proc. of SACMAT*, 2018, p. 51–59.

[11] R. Tourani, S. Misra, T. Mick, and G. Panwar, "Security, privacy, and access control in information-centric networking: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 566–600, 2018.

[12] T. Chatterjee, S. Ruj, and S. D. Bit, "Security issues in named data networks," *Computer*, vol. 51, no. 1, pp. 66–75, 2018.

[13] A. Afanasyev, P. Mahadevan, I. Moiseenko, E. Uzun, and L. Zhang, "Interest flooding attack and countermeasures in named data networking," in *Proc.of IFIP Networking Conference*, 2013, pp. 1–9.

[14] H. Dai, Y. Wang, J. Fan, and B. Liu, "Mitigate ddos attacks in ndn by interest traceback," in *Proc. of IEEE INFOCOM WKSHPS*, 2013, pp. 381–386.

[15] K. Wang, H. Zhou, Y. Qin, and H. Zhang, "Cooperative-filter: countering interest flooding attacks in named data networking," *soft computing*, vol. 18, no. 9, pp. 1803–1813, 2014.

[16] K. Wang, J. Chen, Z. Huachun, and Y. Qin, "Content-centric networking: Effect of content caching on mitigating dos attack," *International Journal of Computer Science Issues*, vol. 9, pp. 43–52, 12 2012.

[17] W. Wong and P. Nikander, "Secure naming in information-centric networks," in *Proc. of the Re-Architecting the Internet Workshop*, 2010.

[18] B. A. Alzahrani, V. G. Vassilakis, and M. J. Reed, "Securing the forwarding plane in information centric networks," in *Proc. of CEEC*, 2013, pp. 174–178.

[19] B. A. Alzahrani, M. J. Reed, and V. G. Vassilakis, "Enabling z-filter updates for self-routing denial-of-service resistant capabilities," in *Proc. of CEEC*, 2012, pp. 100–105.

[20] D. Kim, S. Nam, J. Bi, and I. Yeom, "Efficient content verification in named data networking," in *Proc. of ACM ICN*, 2015, p. 109–116.

[21] M. Conti, P. Gasti, and M. Teoli, "A lightweight mechanism for detection of cache pollution attacks in named data networking," *Comput. Netw.*, vol. 57, no. 16, p. 3178–3191, 2013.

[22] G. Mauri, R. Raspadori, M. Gerlay, and G. Verticale, "Exploiting information centric networking to build an attacker-controlled content delivery network," in *Proc. of MED-HOC-NET*, 2015.

[23] N. Abani, T. Braun, and M. Gerla, "Betweenness centrality and cache privacy in information-centric networks," in *Proc. of ACM ICN*, 2018.

[24] A. Jones and R. Simon, "A privacy-preserving collaborative caching approach in information-centric networking," in *Stabilization, Safety, and Security of Distributed Systems*, 2020, pp. 133–150.

[25] A. Mohaisen, H. Mekky, X. Zhang, H. Xie, and Y. Kim, "Timing attacks on access privacy in information centric networks and countermeasures," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 6, pp. 675–687, 2014.

[26] N. Kumar and S. Srivastava, "A triggered delay-based approach against cache privacy attack in ndn," in *Proc. of IEEE/ACIS ICIS*, 2018.

[27] Y. Jingtao, Z. Li, and Q. Junlei, "Design of an internet of brain things network security system based on icn," *IEEE Access*, vol. 7, pp. 1694–1705, 2019.

[28] C. Ghali, G. Tsudik, and C. A. Wood, "When encryption is not enough: Privacy attacks in content-centric networking," in *Proc. of ACM ICN*, 2017.

[29] N. Kumar, A. Aleem, A. K. Singh, and S. Srivastava, "Nbp: Namespace-based privacy to counter timing-based attack in named data networking," *Journal of Network and Computer Applications*, vol. 144, pp. 155 – 170, 2019.

[30] R. Focardi, R. Gorrieri, R. Lanotte, A. Maggiolo-Schettini, F. Martinelli, S. Tini, and E. Tronci, "Formal models of timing attacks on web privacy1 1research partially supported by murst progetto cofinanziato tosca." *Electronic Notes in Theoretical Computer Science*, vol. 62, pp. 229 – 243, 2002.

[31] R. Gorrieri, R. Lanotte, A. Maggiolo-Schettini, F. Martinelli, S. Tini, and E. Tronci, "Automated analysis of timed security: A case study on web privacy," *Int. J. Inf. Secur.*, vol. 2, no. 3, p. 168–186, 2004.

[32] A. Bortz and D. Boneh, "Exposing private information by timing web applications," in *Proc. of WWW*, 2007, p. 621–628.

[33] G. Zhang, S. Fischer-Huebner, L. A. Martucci, and S. Ehlert, "Revealing the calling history of sip voip systems by timing attacks," in *Proc. of ARES*, 2009, pp. 135–142.

[34] M. Jakobsson and S. Stamm, "Web camouflage: Protecting your clients from browser-sniffing attacks," *IEEE Security Privacy*, vol. 5, no. 6, pp. 16–24, 2007.

[35] S. Schinzel, "An efficient mitigation method for timing side channels on the web," in *Proc. of COSADE*, 2011.

[36] V. Shmatikov and M.-H. Wang, "Timing analysis in low-latency mix networks: Attacks and defenses," in *Proc. of ESORICS*. Springer, 2006, pp. 18–33.

[37] Y. Gilad and A. Herzberg, "Spying in the dark: Tcp and tor traffic analysis," in *Proc. of PETS*, 2012, pp. 100–119.

[38] S. J. Murdoch and G. Danezis, "Low-cost traffic analysis of tor," in *Proc. of IEEE SP*, 2005, pp. 183–195.

[39] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proc. of IEEE FOCS*, 2007, pp. 94–103.

[40] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*. Siam, 1970, vol. 30.

[41] S. Cachon, Gérard P.and Netessine, *Game Theory in Supply Chain Analysis*. Boston, MA: Springer US, 2004, pp. 13–65.

[42] P. Dubey, O. Haimanko, and A. Zapechelnyuk, "Strategic complements and substitutes, and potential games," *Games and Economic Behavior*, vol. 54, no. 1, pp. 77–94, 2006.

[43] A. Pfitzmann and M. Köhntopp, "Anonymity, unobservability, and pseudonymity—a proposal for terminology," in *Designing privacy enhancing technologies*. Springer, 2001, pp. 1–9.

[44] C. Diaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *Proc. of PETS*. Springer, 2002, pp. 54–68.

[45] C. Diaz, "Anonymity metrics revisited," in *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2006.

[46] P. Golle and K. Partridge, "On the anonymity of home/work location pairs," in *Proc. of PerCom*. Springer, 2009, pp. 390–397.

[47] N. Spring, R. Mahajan, and D. Wetherall, "Measuring isp topologies with rocketfuel," *ACM SIGCOMM Comput. Commun. Rev*, 2002.

[48] (2019, July) University of Michigan Campus Network Diagram. [Online]. Available: "https://its.umich.edu/enterprise/wifi-networks/campus-network-diagram-description"

Sivaraman Vignesh received the B.Tech. degree in computer science and engineering from National Institute of Technology, Tiruchirappalli India, in 2014, the M.Tech. degree in computer science and engineering from the Indian Institute of Technology, Guwahati, India, in 2016, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, in 2021. His research interests include Information Centric Networks, privacy in networks and data center networks.

Biplab Sikdar (S'98-M'02-SM'09) received the B.Tech. degree in electronics and communication engineering from North Eastern Hill University, Shillong, India, in 1996, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1998, and the Ph.D. degree in electrical engineering from the Rensselaer Polytechnic Institute, Troy, NY, USA, in 2001. He was on the faculty of Rensselaer Polytechnic Institute from 2001 to 2013, first as an Assistant and then as an Associate Professor. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. His research interests include computer networks, and security for IoT and cyber physical systems. Dr. Sikdar is a member of Eta Kappa Nu and Tau Beta Pi. He served as an Associate Editor for the IEEE Transactions on Communications from 2007 to 2012 and for the IEEE Transactions on Mobile Computing from 2014-2017.