Wireless Networks (2020) 26:667–683
https://doi.org/10.1007/s11276-019-02176-3

A Game-theoretic analysis on the economic viability of mobile content pre-staging

Zhen Li[1] · Qi Liao[2] · Aaron D. Striegel[3]

Published online: 1 November 2019

Abstract
The rapid growth of demand for data in wireless communications has driven the mobile service carriers and the research community to seek both effective technical and alternative solutions to the data demand problem. One particular solution, content pre-staging, tries to push content as close to the mobile device as possible in order to lower demand at peak times. Assuming the interesting case that mobile device storage could be made available as part of the mobile carrier's system capacity either directly by the end user or indirectly by the carrier, this paper investigates the potential economic impacts on the mobile service business and various stakeholders of content pre-staging. We explore the economic implications of content pre-staging by modeling the interplay among the mobile carrier, end users, and the content provider in a game theoretic framework. The carrier designs pricing mechanisms to affect the behaviors of the content provider and end users for the purpose of profit maximization. In particular, two prices are introduced, the price charged to the content provider to pre-stage content on mobile device storage, and the monetary reward to compensate users for the usage of their mobile device storage. Although the individual incentive of the carrier is not necessarily aligned with social incentives, the welfare analysis of content pre-staging shows that the practice improves social welfare by increasing network efficiency. Localizing content increases the overall profitability of mobile service business which is positively related to the relevance of the pre-staged content. The carrier's pricing mechanisms determine the manner in which the increased profitability of the business is shared by various interested parties. While the carrier may design prices strategically to retain a larger share of the increased profitability, content pre-staging can benefit all the three parties in the game, i.e., the carrier gains in saved capacity and new revenue, users gain QoE, content, and financial rewards for sharing mobile device storage, and the content provider gains in increased revenue from increased content access.

Keywords Wireless mobile network · Content pre-staging · Operator-accessible storage · Economics · Game theory · Smart data pricing · Optimization · Social welfare

1 Introduction

Recent years have seen explosive demand for wireless data across a wide range of mobile devices. Conventionally, meeting the rising demand requires large amount of investment in mobile capacity. Mobile service carriers may be constrained by financial resources and/or motivation to increase capacity as much as demand. Furthermore, given the limited availability of wireless spectrum and the slow process by which new spectrum is re-purposed [1], the growth in the supply of wireless capacity is unlikely to keep pace with the massive increase in access demand. Mobile carriers are faced with the reality of both trying to improve capacity via efficiency or acquisition whilst simultaneously considering the economics of allocating limited resources.

To cope with the increased demand for cellular data, there have been significant research efforts to develop techniques for minimizing the cellular network traffic and

✉ Qi Liao
liao1q@cmich.edu

1 Department of Economics and Management, Albion College, Albion, USA

2 Department of Computer Science, Central Michigan University, Mount Pleasant, USA

3 Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, USA

improving user Quality of Experience (QoE) [2–5]. From the perspective of mobile carriers, techniques that increase network efficiency with no requirement on significant infrastructure changes can be quite appealing. The notion of caching as a fundamental principle of Internet structure has been a key driver of network efficiency since its initial inception [6]. Content caching can be widely used throughout the network to improve user QoE by reducing the pressure on service providers. Unlike wired networks whereby an in-network cache might reduce link bandwidth needs, for the mobile Internet sector, the bottleneck is at the access link where bandwidth is shared by all mobile terminals. Mobile devices tend to operate best when they all participate in the caching process due to changing link quality and mobile dynamics. Unfortunately, caching tends to be reactive, saving link capacity primarily against future access needs but offering little benefit when the content is first downloaded.

In contrast, content pre-staging is the process by which content is proactively pushed to the mobile device itself during off-peak times or via secondary access mechanisms (e.g., device-to-device (D2D) caching [7]). It is a promising approach to improve access to web content in face of exponential growth in cellular traffic with the pervasiveness of mobile devices and rich media services. When done properly, pushing content to local cache in mobile terminals can be quite effective [8].

This work is motivated by the interesting promise of pre-staging and explores the economic impact on the stakeholders, i.e., users, mobile carriers and content providers. Given the technical approaches the research community has proposed to deal with wireless traffic, there has been a distinct lack of economic explorations of the techniques regarding incentives and welfare implications from both individual and social perspectives. In this paper, we posit an interesting twist on content pre-staging and look at the notion of content pre-staging from an economic point of view. We ask the question about how the economics of pre-staging may change if there exists some mechanism whereby spare mobile device storage could be made accessible to the mobile carrier via a well-defined API for time shifting demand [9].

In the context of content pre-staging on mobile devices, the carrier implements a content localization policy to access the unused storage on mobile devices owned by end users and manage the storage as part of its system capacity. In the meantime, the carrier forms a contractual relationship with content providers to pre-stage selected content on mobile devices. The carrier may charge a content provider to pre-stage content on mobile devices and reward end users for the usage of their mobile device storage. Traditionally, Internet service providers (ISPs) and mobile carriers have used simple flat-rate broadband data plans for both wired and wireless network access. With the popularity of mobile devices and rapid increases in data demand, carriers around the world have started to explore alternative broadband access pricing, penalties, and accounting mechanisms to manage their networks [10]. Notably, pre-staging content on mobile devices could introduce new revenue streams for the carrier (e.g., content providers may have the incentive to pay carriers to localize their content to guarantee advertising income from smooth delivery of advertisement), but creates at least two more pricing questions for the carrier: how much to charge the content provider for localizing content and how much to reward end users for the usage of their mobile device storage.

To address the aforementioned research questions, we build a game theoretical model to study the interdependent decision-making by the three major types of economic agents involved in the content pre-staging practice in mobile service. The three interested parties are the mobile service carrier, the content provider, and the end users of mobile services who are also the consumers of content. The carrier sets the data access rate charged to the end users. The carrier also needs to decide on the content pre-staging rate charged to the content provider and the reward paid to end users for the use of their mobile device storage, if any, in case of content pre-staging. The model is formulated as a mechanism design. End users decide on data demand and the quantity of mobile device storage shared with the carrier. The content provider decides on the volume of content to pre-stage. Actual data consumption of users and the mobile storage users are willing to share and the quantity of the content the provider is willing to pre-stage are the messages they send as a function of their private information. The carrier designs optimal pricing mechanisms to induce the decision-making by end users and the content provider to maximize profit. By modeling the interactions of the three interested parties in the content pre-staging game, we study the economic impacts of operator-accessible mobile device storage on the mobile service business as a whole and on individual stakeholders respectively. We study how the pricing strategy of the mobile carrier may affect social welfare and the sharing of economic gains among stakeholders.

Specifically, our contributions include:

- Formulating the interactions among the mobile carrier, the content provider, and end users in case of pre-staging content on mobile devices as a mechanism design in a game theoretic framework.
- Deriving the subgame perfect Nash equilibrium solutions of the game.
- Evaluating the insights from model implications, such as the social welfare implications of content pre-

staging, financial incentives of individual parties, and the effects of the mobile carrier's pricing mechanisms, etc.

The social welfare implications of content pre-staging show that pushing content to local mobile device storage is feasible by providing appropriate financial incentives to the various stakeholders in the game. The subgame perfect Nash equilibrium of the three-party game shows that content localization increases social welfare, i.e., the *combined* economic well-being of the carrier, the content provider, and end users, by increasing the overall efficiency and hence profitability of the mobile service business. The extent of the efficiency improvement depends on the relevance of the pre-staged content and the initial severity of network congestion. Pre-staging can also benefit each individual party. First, end users receive improved QoE by reducing delay in data delivering via scarce bottleneck link since the content is readily available locally on their devices. Users may additionally receive a reward in return for renting their relatively cheap and unused storage. Users also gain in consumer surplus while enjoying more content access. Second, the carrier saves operating cost by time-shifting demand. With proper pricing, the carrier can maximize its profitability through new revenue from the content provider to allow the content provider to optimize its content delivery. Third, the content provider gains from satisfied users and increased advertising revenue from more content access by users. We note that the sharing of the increased profitability of the business from localizing content among stakeholders is arbitrary. The carrier has a range of pricing choices to charge the content provider and/or to reward end users to make them willing to cooperate and manipulate the allocation of economic gains.

The rest of the paper is organized as follows. Section 2 provides a discussion on the related literature seeking various solutions to reduce the backbone pressure on mobile networks, and how our research fits in the literature. Section 3 introduces a base model of a two-player game that involves only the carrier and end users in the mobile service business with no content pre-staging. The best response of end users is studied and the optimal pricing mechanism of the carrier is derived to induce profit maximization for the carrier. The base model serves as the basis of comparison with the content pre-staging case. In Sect. 4, an extended three-player game is developed by including the content provider in the game in the scenario of pushing content to mobile device storage. In Sect. 5, we search for the feasible and optimal pricing strategies of the mobile carrier and analyze the economic welfare implications of content localization and the carrier's pricing choice. We also study the financial incentives of the three interested parties. Section 6 provides numerical examples and graphical illustrations to further demonstrate the model solutions and implications. Finally, Sect. 7 concludes the paper.

2 Related work

As we enter the big data era, there has been a constant back and forth between the increasing demand for data and the limited growth of bandwidth and capacity in mobile communication technologies. In the mobile broadband domain, mobile service providers have to maintain capacity sufficient to satisfy the peak time demand for bandwidth to guarantee user QoE. Demand uncertainties can lead to under-supply of capacity during the peak time with the consequence of capacity rationing among users, causing unsatisfying user QoE while the same capacity would be considered an over-supply during off peak periods. In addition, the popularity of bandwidth-intensive mobile applications and mobile devices also poses a huge burden on the carriers to support the increasing wireless data traffic to invest in developing advanced networks or to expand the capacity of the current networks in an appropriate pace.

Caching has been a commonly used method to alleviate network traffic. The store of recently-accessed content allows faster access to the content in the future. Recently, there has been a growing interest in more proactive techniques to improve content caching called pre-fetching, a method that caches selected content before it is actually needed [11, 12]. Content caching and pre-fetching can significantly reduce the mobile bottleneck link pressure and improve the QoE of end users. Research shows that users are highly likely to view the data in pre-fetched videos, indicating a promising opportunity to reduce network load [13]. Technical efforts searching for effective caching and pre-fetching have ranged from characterizing data and energy consumption by smartphone applications [14, 15] to actively examining the efficiency of the data transfers themselves [7, 8, 16].

Traditionally, cache pre-fetching happens at network edges. For mobile networks, the bottleneck tends to be the access link shared by multiple mobile terminals. To that end, researchers have proposed to push the caching paradigm further to implement a local cache on mobile devices [8], i.e., content pre-staging. Since this practice requires the availability of storage, an even more radical argument has been made to allow mobile carriers to access the unused mobile device storage for better time shifting of demand [9]. Content pre-staging affects both the demand and supply side of the mobile network capacity. While on one hand it time shifts demand to pre-fetch content during off-peak time for users to access at peak time, it also makes content providers to share the cost of capacity consumption along with end users. This paper focuses on an economic analysis of carrier-accessible mobile device storage for content pre-staging and illustrates the welfare impacts on mobile carriers, end users and content providers in a game theoretical setting.

Notably, there are no magical solutions to the data demand problem. Any technical or non-technical efforts must work either on the supply side or the demand side or both of the mobile system. For data demand management for example, there has been research on time-dependent pricing or peak-load pricing to smooth out data demand by charging more on peak-time traffic [17]. By charging users dynamically over time, time dependent pricing may flatten the temporal fluctuations of demand by motivating users to shift their usage to off-peak hours with lower price [18, 19].

For the supply of network capacity, one way to keep up financing capacity investment is to transfer the carriers' costs to other parties such as content providers. One recent proposal is content sponsoring that charges content providers instead of users for resources consumed in accessing the content [20–26]. The additional revenue streams will allow carriers to have the financial resources to build capacity. Content sponsoring however, does not necessarily alleviate the problem of matching limited network capacity with network demand because users may choose to consume more content when their content consumption is sponsored (at lease partially) by content providers. Both time dependent pricing and content sponsoring are affiliated with smart data pricing (SDP) concept, a broad set of ideas that goes beyond the traditional flat-rate or byte-counting models [10]. SDP is an umbrella term for a variety of pricing practices that have been proposed in recent years as opposed to the traditional flat-rate model.

Human mobility has been considered for caching and content delivery. Along with the high capabilities of modern mobile devices, local caching and sharing is possible by exploiting the correlations of user behavioral patterns to enable mobile networking in proximity [27]. Human mobility patterns are explored to develop a content delivery system that leverages the timing of content broadcasting to avoid transmitting duplicate copies of content to mobile users [28]. D2D communications may be leveraged to share cached information in cellular networks [29, 30]. Human mobility patterns and social tie for caching content in mobile devices for distribution via D2D communications is also studied [31].

Both content caching and pre-fetching have been studied in terms of economic incentive designed, including game theoretical modeling and analysis [32, 33]. This work extends the social welfare analysis devised in [33], formulates the problem as a mechanism design, and supplements the theoretical analysis with numerical evaluations and case studies. A common theme in recent cache economics research is to seek for service carriers' optimal strategies to monetize caching, from optimal caching and pricing policies of carriers [34] to the design of caching contracts between carriers and content providers [35]. Content caching can also be combined with dynamic pricing. For example, users may take advantage of D2D communications to cache content during off-peak time and trade cached content during peak time to save payments [36].

3 The base-case model

In this section, we first model the key interactions between the mobile service carrier and end users with no content pre-staging. The framework of the game is laid out and the subgame perfect equilibrium is derived. The equilibrium solutions for the base case serve as the basis of comparison with the solutions for the content pre-staging case which is analyzed in the following section. For reference, Table 1 lists the symbols and definitions of the variables, parameters, and functions included in the model setup.

3.1 Formulating the carrier-user game

Before content pre-staging is adopted, the interplay is mainly between the mobile service carrier and end users. The carrier is of unity, i.e., there is one carrier that can be understood as the representative of all carriers. The carrier sets a unit price to charge end users for their data usage. End users choose how much data to consume. The interplay between the carrier and end users can be formulated as a mechanism design problem in which the carrier designs optimal pricing mechanisms to induce the behaviors of users to reach the profit maximization outcome.

Consider one mobile carrier with a set of end users. The carrier sets the data access rate (denoted by p) to charge end users, and users choose how much data to consume given the price set by the carrier. By predicting the effect of data access price on user behaviors, the carrier designs an optimal pricing strategy for the purpose of profit maximization. End users own mobile devices and use them for web browsing, video streaming, reading e-books, playing games, taking photos, and running applications of all kinds. Users save photos, applications, music etc. on mobile devices so that they have both allocated/used storage (downloaded music, applications, saved photos, etc.) and free/unused storage on their mobile devices. A_j denotes a representative user j's allocated storage on mobile device.

We assume all network data demand by users comes from accessing content provided by the content provider of

Table 1 Definitions of variables, parameters and functions

Variable/parameter/function	Definition
p	Unit data access rate set by the carrier on end users
q	Unit reward received by end users for the carrier to access their mobile device storage
r	Unit content pre-staging rate the carrier charges the content provider
a	Content provider's per-content advertising revenue
θ	Average size of web content
g	Units of content the content provider chooses to pre-stage on mobile devices
n_j	Quantity of web content requested by representative user j
x_j	Realized data consumption by user j
U_j	User j's user surplus function
V_j	User j's total utility received from content access and usage of mobile device storage
s_j	Total storage capacity of user j's mobile device
A_j	User j's allocated/used storage on mobile device
$D(p)$	Realized demand for link capacity by all end users
X	Link capacity of the carrier
$C(X)$	Carrier's total cost function to build a link capacity of X
b	Carrier's operating cost per unit of capacity
β	Carrier's per-unit marginal cost in case of network congestion
W	Net social welfare
π	Carrier's profit
ES	End users' surplus
CS	Content provider's surplus
α	Parameter in isoelastic utility function for risk averse agents
γ_j	Scaling factor to differentiate user j's self valuation of content consumption and mobile storage usage
δ	Success rate of users' web content request
ξ	Percentage of pre-staged content that is relevant
λ_j	Lagrange multiplier associated with user j's mobile storage cap constraint

unity. Due to possible network congestion, not all content requests can be successfully served. The user may give up after certain trials, and an arbitrary content request is supposed to have a success probability of $\delta \in (0, 1]$. We assume user j generates n_j (a random number) network content accesses and the average size of heterogenous content is θ, thus the realized data consumption by user j for accessing network content is $x_j = \theta n_j$. The user's optimal/desired demand for network content is correspondingly $\frac{x_j}{\delta}$ since out of all content requests, a fraction of δ is successfully served. δ depends on the overall congestion on the network, and it essentially measures user QoE: the higher δ is, the better is the content access quality experienced by users.

Users value both consumption of network content and data saved on mobile devices. User j decides on used storage and network content access to maximize the user surplus function denoted by $U_j(x_j, p, A_j)$ for a given data access rate p. We assume that end users' preferences are additively separable on content access and usage of mobile device storage, i.e., $U_j = v_j(x_j, p) + v_j(A_j)$ where user j's total surplus is the sum of the user surplus from network content access and the usage of storage on mobile devices. In addition, we assume the two components of the user surplus function $v_j(x_j, p)$ and $v_j(A_j)$ have the same functional form. They are both concave and second-order differentiable with $v' > 0$ and $v'' < 0$, consistent with the economic principles of increasing total utility and diminishing marginal utility. As shown in the second term of the user surplus function $v_j(A_j)$, the size of allocated mobile device storage chosen by user j does not depend on the data access price p set by the carrier. When the carrier has no access to mobile device storage, the storage allocation of each user's mobile device plays no role in the carrier-user game in the base case.

We now consider the carrier's choices of a link capacity (X) and a data access rate (p) so as to maximize profit. Profit is revenue minus cost. The total revenue of the carrier is $p \sum_j x_j(p)$ where $\sum_j x_j(p)$ is the total data consumption by all end users, depending on the data access rate set by the carrier. The carrier's total cost is composed of two parts: the link capacity cost and the bottleneck congestion cost (e.g., rationing cost of having to allocate limited capacity or the cost of temporarily acquiring additional capacity), both are assumed to be linear and exogenous. Specifically, b is the unit cost of the link capacity and β is the unit network congestion cost. Thus, the total cost function $C(X)$ of the carrier is as follows,

$$C(X) = bX + \beta(D(p) - X) \tag{1}$$

where $D(p) = \sum_j x_j(p)$ is the realized demand for link capacity by all end users at a given price p.

According to the total cost function, the carrier chooses whether to maintain sufficient capacity in its network by comparing the two cost parameters, b and β,

$$X = \begin{cases} D(p) & \text{if } b \le \beta \\ 0 & \text{if } b\beta \end{cases} \tag{2}$$

We assume the cost of system failure always dominates the operating cost of the capacity, i.e., $b \ll \beta$. Thus the carrier uses the pricing strategy to manage data requests by end users to prevent network failure from happening. The cost function of the carrier is reduced to $C(X) = bX$.

3.2 Equilibrium analysis in the base case

We derive the model solution by studying first the end users' optimal strategy. A representative user j takes the data access rate set by the carrier as given and chooses data access request so as to maximize user surplus U_j. We denote this optimal demand for data as $x_j^*(p)/\delta$. Realized or served user demand is thus x_j^*.

We take an isoelastic utility function, a commonly used functional form of utility in economics that satisfies both increasing and concave properties of utility. Specifically, it is the only class of utility functions with constant relative risk aversion. The parameter (denoted by α) in Eq. (3) is positive for risk averse agents. It is isoelastic because the optimal decision at given level of wealth is also optimal for another level of wealth. In the context of mobile storage allocation, this type of utility function leads to a constant proportionate allocation of mobile storage between used and unused storage by individual end users regardless of the actual size of mobile storage [37, 38].

Users benefit from accessing web content and using mobile device storage. Users also have to pay for data usage. The difference between user benefit and cost is defined as user surplus. User j's optimal choices of web content consumption (x_j) and mobile storage usage (A_j) satisfy the following user surplus maximization problem:

$$\begin{aligned} \max_{x_j, A_j} \ & U_j = \frac{(x_j/\delta)^{1-\alpha_j}}{1-\alpha_j} + \gamma_j \frac{A_j^{1-\alpha_j}}{1-\alpha_j} - p x_j \\ \text{s.t. } & A_j \le s_j \end{aligned} \tag{3}$$

where s_j is the total storage capacity of user j's mobile device. γ_j is a scaling factor to differentiate user j's self valuation of web content consumption and allocated mobile device storage. If the user prefers web content to data saved on allocated storage, $\gamma_j < 1$. If the user assigns a higher value on data stored on mobile device than accessing online content, $\gamma_j > 1$. Users incur a data cost increasing in data access rate and the volume of data consumption. By setting the data access rate, the carrier may use its pricing strategy to affect end users' choice of web content consumption. Note the user's choices of x_j and A_j are not interdependent in the base case. They would become interdependent in the extended pre-staging model in Sect. 4.

We use the underline to denote the optimal solutions to the base model. The first order condition of Eq. (3) with respect to x_j provides us with the realized web data demand by user j,

$$\underline{x}_j^* = p^{-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}} \tag{4}$$

The quantity of content access by the user is accordingly $\underline{n}_j^* = \frac{\underline{x}_j^*}{\theta} = \frac{p^{-\frac{1}{\alpha_j}} \delta^{(1-\frac{1}{\alpha_j})}}{\theta}$.

In the base case, how much storage users occupy is independent of the carrier's decisions. The optimal choice of A_j can be found from the first order condition of Eq. (3) with respect to A_j,

$$\underline{A}_j^* = \left(\frac{\lambda_j}{\gamma_j}\right)^{-\frac{1}{\alpha_j}} \tag{5}$$

where λ_j is the Lagrange multiplier associated with the mobile storage cap constraint.

Next we consider the carrier's profit (π) maximization strategy by choosing data access rate p and link capacity X so that the price limits end users' actual Internet content access to no higher than the carrier's chosen capacity,

$$\begin{aligned} \max_{p, X} \ & \pi = p \sum_j x_j^* - bX \\ \text{s.t. } & \sum_j x_j^* \le X \end{aligned} \tag{6}$$

Given a price p, the carrier can find $X^*(p)$, the optimal link capacity as a function of the price so as to maximize profit. We use $S(p) = X^*(p)$ to denote this supply side function. When end users and the carrier are at a market equilibrium, supply equals demand, i.e., $S(p) = D(p)$. At such an equilibrium price $\underline{p}^*$, each end user maximizes user surplus by consuming $\underline{x}_j^*(\underline{p}^*)$ link capacity, and the carrier maximizes profit by providing just enough capacity $\underline{X}^*(\underline{p}^*) = \sum_j \underline{x}_j^*(\underline{p}^*) = \sum_j \underline{p}^{*-\frac{1}{\alpha_j}}\delta^{1-\frac{1}{\alpha_j}}$ from Eq. (4). The carrier's optimal profit is $\underline{\pi}^* = (\underline{p}^* - b)\sum_j \underline{p}^{*-\frac{1}{\alpha_j}}\delta^{1-\frac{1}{\alpha_j}}$ where $(\underline{p}^* - b)$ is the per-unit profit.

Suppose initially there is no contractual relationship between the carrier and the content provider to pre-stage content. The content provider thus plays no part in the game between the carrier and its end users. Assuming the pre-staged content includes both non-advertising and advertising content, which generates revenue for the content provider, and the cost function of the content provider is held fixed, therefore maximizing profit is equivalent to generating the maximum advertising revenue to the content provider. We suppose the content provider earns an average revenue a for a content of average size θ. The advertising revenue received by the content provider in the base case is hence $a\sum_j \underline{n}_j^*(p) = \frac{a}{\theta}\sum_j \underline{p}^{*-\frac{1}{\alpha_j}}\delta^{1-\frac{1}{\alpha_j}}$ where $\frac{a}{\theta}$ is the per-unit advertising revenue.

4 Three-player pre-staging game

This section expands the base model by accommodating the practice of content pre-staging. In particular, the carrier implements the content localization policy to have access to unused storage on end users' mobile devices. In the meantime, the carrier forms a contractual relationship with the content provider to pre-stage selected content on mobile devices. To facilitate the practice of pushing content to local mobile device storage, the carrier provides proper financial incentives for the content provider and end users to cooperate.

4.1 Formulating mobile device storage access game

In principle, pushing content to mobile device storage may benefit the mobile carrier by saving link capacity and generating content pre-staging revenue. Content pre-staging may also improve user QoE and guarantee stable advertising income to the content provider. Realizing said potential gains requires properly chosen pricing schemes. To formulate pre-staging, we add a third party, the content provider to the game (Fig. 1).

In the extended three-party game, the carrier sets three prices, the data access rate (denoted by p), the reward rate to pay end users for the usage of their mobile device storage (denoted by q), and the content pre-staging rate

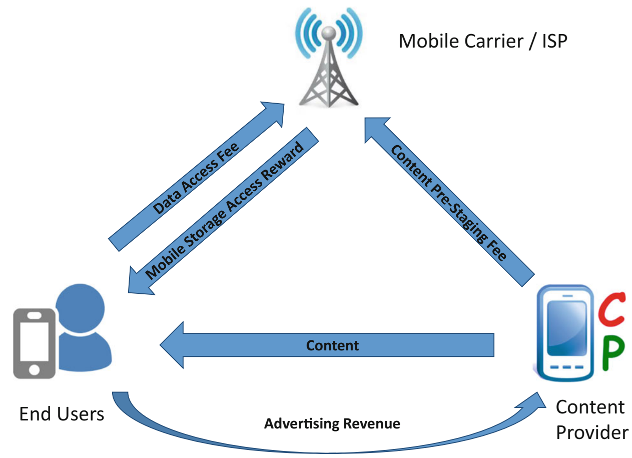


Fig. 1 The economic relationships between the mobile carrier, the content provider, and users in the content pre-staging game

charged to the content provider (denoted by r). All the three prices are on a per-unit basis. To focus the analysis on the two new prices q and r associated with content localization, we suppose the carrier keeps the same data access rate as in the base case, i.e., $p^* = \underline{p}^*$, and chooses optimal q and r. The content provider responds by determining the maximum amount of content it is willing to pay to store locally. End users respond by choosing how much content to consume and the distribution of mobile device storage between allocated space and unused space. The unused space is the part of mobile device storage available for the carrier to access.

An essential challenge for the success of content pre-staging is the relevance of pre-staged content. Even with the increased storage capabilities, it is not possible to completely anticipate the content needed by a user. It is also unlikely that all the pre-staged content is useless. Thus, we assume that a fraction of $\xi \in (0,1)$ content loaded onto mobile devices is relevant, and assume the advertising revenue is linearly proportional to the volume of advertising content being viewed. These are simplifying assumptions for the purpose of tractability and they do not fundamentally affect the model conclusions.

4.2 Content provider's best response

In contrast to the base case, maximizing the content provider's profit is no longer equivalent to maximizing its advertising revenue when the content provider can pay a fee to pre-stage content. The content provider's advertising revenue depends on end users' access to content. The content provider's decision-making is to choose the quantity of content to pre-stage on mobile devices, up to the maximum available unoccupied storage.

The content provider would be willing to pre-stage content on mobile devices if the increased advertising revenue exceeded the cost of pre-staging. If the content is not pre-staged, it is successfully accessed with a probability of δ, as in the base case. Staging content locally increases the chance of the content being accessed by $\xi(1-\delta)$. The per-unit gain in advertising revenue of pre-staged content is $\frac{a\xi(1-\delta)}{\theta}$.

Let g be the units of content that the content provider chooses to pre-stage on mobile devices. Given the pre-staging rate set by the carrier, the content provider's best response g^* is straightforward as follows.

$$g^* = \begin{cases} \sum_j (s_j - A_j)/\theta & \text{if } r \leq a\xi(1-\delta)/\theta \\ 0 & \text{if } r > a\xi(1-\delta)/\theta \end{cases} \tag{7}$$

That is, the content provider would choose to pre-stage the maximum possible content as long as the gain of pre-staging exceeds the cost. The content provider's advertising revenue along with profit would increase in this case.

4.3 End users' best response

The surplus function of end users depends on the costs and benefits associated with user access to Internet content and mobile device storage. In contrast to the base case, users now face a tradeoff between their own mobile device storage usage and the forgone possible reward for unused storage to be incorporated into the carrier's system capacity. The amount of mobile device storage that users would be willing to share to the carrier depends on the reward rate. Therefore, the current form of the user surplus function is $U_j(x_j, p, A_j(q))$ where user j's choice of allocated mobile device storage depends on the reward rate q set by the carrier.

Users do not have to be aware of whether a particular content is pre-staged or not. User j consumes x_j amount of data towards the goal of maximizing user surplus, of which $\xi(s_j - A_j)$ is served locally with pre-staged content. The rest is delivered directly from the content provider via the bottleneck link with the size of $x_j - \xi(s_j - A_j)$. User j's total demand for data is thus $\frac{x_j - \xi(s_j - A_j)}{\delta} + \xi(s_j - A_j)$, and the user's net data cost is $px_j - q(s_j - A_j)$, data consumption fee net of rewards received from sharing mobile device storage with the carrier.

Taking an isoelastic utility function, user j's maximization problem is

$$\begin{aligned} \max_{x_j, A_j} \quad U_j = & \frac{\left\{\frac{x_j - \xi(s_j - A_j)}{\delta} + \xi(s_j - A_j)\right\}^{1-\alpha_j}}{1-\alpha_j} \\ & + \gamma_j \frac{A_j^{1-\alpha_j}}{1-\alpha_j} - (px_j - q(s_j - A_j)) \\ \text{s.t.} \quad & A_j \leq s_j \end{aligned} \tag{8}$$

Given the data access rate p, user j's optimal choice of data consumption x_j^* is provided by the first order condition of Eq. (8) as

$$x_j^* = p^{-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}} + \xi(1-\delta)(s_j - A_j) \tag{9}$$

At this content access level, the user generates $\frac{a}{\theta}\{p^{-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}} + \xi(1-\delta)(s_j - A_j)\}$ advertising revenue to the content provider. Pre-staging a certain content increases the probability of successful access by $\xi(1-\delta)$. When the content provider chooses to take full unused storage on mobile devices, user j's actual content consumption increases by $\xi(1-\delta)(s_j - A_j)$ (from Eq. (4) to Eq. (9)) at the same data access rate.

From the first order condition of Eq. (8) with respect to A_j and replacing x_j with its optimal solution in Eq. (9), we derive the optimal allocated storage usage by user j as $\left(\frac{q - p\xi(1-\delta)}{\gamma_j}\right)^{-\frac{1}{\alpha_j}}$. Considering the mobile storage cap constraint, the user's optimal choice of allocated mobile device storage A_j^* is

$$A_j^* = min\left\{\left(\frac{q - p\xi(1-\delta)}{\gamma_j}\right)^{-\frac{1}{\alpha_j}}, s_j\right\} \tag{10}$$

4.4 The carrier's decision

When unused mobile device storage is manageable by the carrier, it becomes part of the system capacity of the carrier, and may be allocated as seen fit by the carrier to pre-stage selected content. The carrier's own link capacity is used to serve users' online content access requests that cannot be fulfilled locally.

The carrier has two revenue sources: data consumption payment by end users and content pre-staging payment by the content provider. The carrier's costs include the cost of link capacity and the rewards to end users for the usage of their mobile device storage. The carrier decides on three rates (p, q, r) and chooses a link capacity (X) with the goal of maximizing profit.

$$\max_{p,q,r,X} \ \pi = p\sum_j x_j + r\theta g - q\sum_j (s_j - A_j) - bX$$
$$\text{s.t.} \ \sum_j \{x_j - \xi(s_j - A_j)\} \le X \tag{11}$$

In optimum, the actual link capacity consumed by end users must be equal to the link capacity chosen by the carrier, i.e., $\sum_j \{x_j^* - \xi(s_j - A_j^*)\} = X$. The choice of optimal capacity X^* and the optimal data access price p^* have the following relationship.

$$X^* = \sum_j \left\{ p^{*-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}} - \delta\xi(s_j - A_j^*) \right\} \tag{12}$$

We call the triple (p^*, q^*, r^*) the optimal pricing mechanism of the carrier. To solve for the mechanism, we turn to social welfare analysis of how content pre-staging may affect the combined economic well-being of the mobile service business as a whole and individual interested parties respectively.

5 Social welfare and optimal pricing

In this section, we illustrate the economic impacts of content pre-staging on social welfare, i.e., the combined welfare of the mobile carrier, end users, and the content provider. We also discuss how to introduce proper financial incentives to facilitate pre-staging in practice, based on which the carrier's pricing mechanism is derived.

5.1 Pre-staging increases social welfare

Combining the economic well-being of the carrier, end users, and the content provider, the social welfare function is of the following form:

$$W = \pi + ES + CS \tag{13}$$

where W = net social welfare, π = profit of the carrier, ES = end users' surplus, and CS = the content provider's surplus. Since the payments made by one party to another (such as content pre-staging fee paid to the carrier by the content provider or the data access payment end users make to the carrier) do not affect social welfare, the net economic benefit of mobile service business depends on users' valuation of mobile services and the usage of mobile devices, the carrier's costs of providing mobile services, and the advertising revenue generated from users' content access, i.e.,

$$W = \sum_j V_j - bX + a\sum_j \frac{x_j}{\theta} \tag{14}$$

where V_j is user j's total utility received from content access and the usage of mobile device storage.

Compared to the base case, the change in social welfare is

$$\Delta W = \Delta \sum_j V_j + \frac{a}{\theta}\Delta \sum_j x_j - \Delta bX \tag{15}$$

In particular, the change in the carrier's link capacity cost is

$$\begin{aligned} \Delta bX &= b(X^* - \underline{X}^*) \\ &= -b\left\{ \sum_j \delta^{1-\frac{1}{\alpha_j}} (\underline{p}^{*-\frac{1}{\alpha_j}} - p^{*-\frac{1}{\alpha_j}}) + \delta\xi \sum_j (s_j - A_j^*) \right\} \end{aligned} \tag{16}$$

The change in the content provider's advertising revenue is

$$\begin{aligned} \frac{a}{\theta}\Delta \sum_j x_j = \frac{a}{\theta}\Bigg\{ &\sum_j \delta^{1-\frac{1}{\alpha_j}} (p^{*-\frac{1}{\alpha_j}} - \underline{p}^{*-\frac{1}{\alpha_j}}) \\ &+ (1-\delta)\xi \sum_j (s_j - A_j^*) \Bigg\} \end{aligned} \tag{17}$$

The change in end users' total utility depends on the change in utility from accessing content and allocated mobile device storage where

$$\begin{aligned} \Delta \sum_j V(x_j) = \sum_j \Bigg\{ &\frac{\left(p^{*-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}} + (1-\delta)\xi(s_j - A_j^*) \right)^{1-\alpha_j}}{1-\alpha_j} \\ &- \frac{(\underline{p}^{*-\frac{1}{\alpha_j}} \delta^{1-\frac{1}{\alpha_j}})^{1-\alpha_j}}{1-\alpha_j} \Bigg\} \end{aligned} \tag{18}$$

$$\begin{aligned} \Delta \sum_j V(A_j) = \gamma_j \sum_j \Bigg\{ &\frac{(min\{\frac{q^* - p^*\xi(1-\delta)}{\gamma_j})^{-\frac{1}{\alpha_j}}, s_j\})^{1-\alpha_j}}{1-\alpha_j} \\ &- \frac{\left(\frac{\lambda_j}{\gamma_j}\right)^{-\frac{1-\alpha_j}{\alpha_j}}}{1-\alpha_j} \Bigg\} \end{aligned} \tag{19}$$

For simplicity, let $q^* - p^*(1-\delta) = \lambda$ for all end users so that users do not change the way they use mobile device storage, thus $A^* = \underline{A}^*$ and $\Delta \sum_j V_j = \Delta \sum_j V(x_j)$. Alternatively, suppose at the initial optimal mobile storage consumption level $\underline{A}^*$, the marginal cost of giving up one unit of storage is equal to the mobile storage access fee q, then the user is neutral to the financial reward of q regarding whether to give up own usage of mobile storage to earn rewards.

If the carrier charges the same data access rate with or without pre-staging, i.e., $p^* = \underline{p}^*$, then all of the three components of social welfare improves.

$$\Delta \sum_j V_j = \sum_j \left\{ \frac{\left(p^{*^{-\frac{1}{\alpha_j}}} \delta^{1-\frac{1}{\alpha_j}} + (1-\delta)\xi(s_j - A_j^*)\right)^{1-\alpha_j}}{1-\alpha_j} - \frac{\left(p^{*^{-\frac{1}{\alpha_j}}} \delta^{1-\frac{1}{\alpha_j}}\right)^{1-\alpha_j}}{1-\alpha_j} \right\} > 0 \tag{20}$$

$$\frac{a}{\theta} \Delta \sum_j x_j = \frac{a}{\theta}(1-\delta)\xi \sum_j (s_j - A_j^*) > 0 \tag{21}$$

$$\Delta bX = -b\delta\xi \sum_j (s_j - A_j^*) < 0 \tag{22}$$

That is, user utility increases as they consume more content; the advertising revenue increases from increased content viewing by users; link capacity cost is saved as part of the content demand is met locally, hence $\Delta W > 0$. Pre-staging content on mobile devices increases the combined economic well-being of the three interested parties.

However, there is still one question remaining: how are the improved efficiency and benefits of the mobile service business distributed among the parties?

5.2 Carrier's pricing mechanism

To provide financial incentives for each party, the distribution of the gains must satisfy $\Delta\pi \geq 0$, $\Delta ES \geq 0$, and $\Delta CS \geq 0$ when compared to the base case. That is, in case of content pre-staging, no party shall be worse off than the no-pre-staging scenario.

For the content provider, $\Delta CS \geq 0$ requires that the increased advertising revenue from content pre-staging must be no less than the pre-staging cost of the content, i.e.,

$$\frac{a}{\theta} \Delta \sum_j x_j \geq r \sum_j (s_j - A_j^*) \tag{23}$$

Combining Eqs. (21) and (23), we derive

$$r \in \left[0, \frac{a\xi(1-\delta)}{\theta}\right]$$

which is consistent with the best response of the content provider in 4.2.

For end users, their gain in user surplus is

$$\Delta ES = \Delta \sum_j V_j + q \sum_j (s_j - A_j) > 0$$

so that users gain at any level of reward rate q for content pre-staging on mobile devices. Therefore, the lower bound on q is zero, and the theoretical upper bound satisfies $\Delta\pi = 0$.

The change in the carrier's profit is

$$\Delta\pi = r \sum_j (s_j - A_j) - q \sum_j (s_j - A_j) - \Delta bX$$

Combined with Eq. (22) we get

$$\Delta\pi = (r - q + b\delta\xi) \sum_j (s_j - A_j) \tag{24}$$

When $\Delta\pi = 0$, $q = r + b\delta\xi$ thus the feasible range of pre-staging reward payment to end users is

$$q \in [0, r + b\delta\xi]$$

Accordingly, $\Delta\pi \geq 0$. It is reasonable to assume that in practice, the carrier must tend to charge a higher content pre-staging fee to the content provider than the reward paid to users, i.e., $r > q$, thus the profitability of the carrier increases in case of content pre-staging.

In summary, when the carrier initiates content localization mechanism with no change in service charge to end users, the carrier's optimal pricing mechanism is $(p^* = \underline{p}^*, q^* = 0, r^* = \frac{a\xi(1-\delta)}{\theta})$. In this case, users gain from improved QoE and more content access, the content provider is equally well between content pre-staging or not, and the carrier receives the maximum possible gain in profitability.

$$\Delta\pi^* = \left(\frac{a\xi(1-\delta)}{\theta} + b\delta\xi\right) \sum_j (s_j - A_j^*) \tag{25}$$

Note the socially optimal value of r or q is not unique. The actual values of q and r determine only the income distribution among the carrier, the content provider, and end users with no impact on social welfare as long as q and r fall in the specified ranges.

5.3 Further discussions

Charging the content provider a pre-staging rate of $\frac{a\xi(1-\delta)}{\theta}$ may not be feasible since the content provider does not gain from content pre-staging and thus has a lack of incentive to pre-stage. To motivate the content provider, the carrier may lower r to share profit with the content provider. The increase in profitability of the mobile service business ($\Delta\pi^*$) may be shared between the carrier and the content provider in an arbitrary manner.

In the carrier's optimal pricing solution, users receive no financial rewards for content localization but rather gain from improved QoE and increased content consumption at $q^* = 0$. For customer willingness to participate, q^* may have to be positive to overcome psychological hurdles such as users' unwillingness to yield part of the control over their mobile devices to the carrier out of privacy or ownership concerns.

End users may receive financial gains from content pre-staging in two alternative ways: being rewarded a fee for the carrier's usage of their mobile device storage, as in the model setup, or via reduced or discounted data access rate tailored to the level of pre-staging. Charge-then-refund practice has entered the pricing model of the mobile industry when it appeared as part of the pricing structure of Google Fi: Google will refund the customer at the end of the month for any data purchased but not used [39]. The refund idea in the content pre-staging context is to refund the customer for any mobile device storage occupied by the carrier to pre-stage. Direct reward and indirect refund are equivalent at appropriately chosen prices. The difference is whether the financial benefit to end users is perceived explicitly. We recommend the reward practice because an explicit monetary gain is more likely to induce end users to comply.

Besides prices, the most important factor determining how effective pre-staging can be is the relevance of pre-staged content. As shown in Eqs. (20)–(22), all three components of changes in social welfare are affected by the relevance of selected content, and social welfare is positively related to parameter ξ. The carrier and the content provider must have the answer to the question "what content to pre-stage in order to increase the probability of a content hit?" One possibility for targeted pre-staging is based on machine learning over information collected through online activities of users, e.g., locations, purchase history, social networks, similar to targeted advertising. While there has been research on intelligent content caching [40, 41], development of algorithms for intelligent content pre-staging is beyond the scope of this paper, which explores the economic implications of content pre-staging.

The content pre-staging policy requires mechanism design. The two new rates (the reward received by users for sharing unused mobile storage and the content pre-staging rate charged to the content provider) introduced by the pre-staging policy serve as two price mechanisms to change the incentives of end users and the content provider to reach the desired outcome.

Since this paper aims at studying the economic implications of content pre-staging, it considers a simplified pre-staging policy that is non-discriminatory. End users and content providers are both of one type with similar preferences. The prices charged to users and the content provider, or the reward to users are all flat rate. A formal mechanism design analysis would be a worthwhile topic of future research to focus on the mobile carrier's profit-maximizing pricing strategy and the design of pre-staging policy in a game of incomplete information with agents of different types. Various pre-staging and pricing mechanisms could be designed to induce the best outcome for the mobile carrier. Mechanisms could also be designed to constrain the behavior of the carrier as individual incentives are not always aligned with social incentives.

6 Case study and evaluation

In this section, we further study a few key issues discussed in previous sections via numerical examples and graphical illustrations. Throughout the numerical simulation, the values of parameters are estimated, and both the utility function and the cost function take the format as specified in the model setup. The analytics demonstrate the model solutions, especially the important relationship between the pricing strategy of the mobile carrier and the sharing of improved economic welfare among the three stakeholders involved in the content pre-staging game.

6.1 Determinants of welfare effects of content pre-staging

First, we study the range of feasible content pre-staging rate by estimating the additional advertising revenue that content pre-staging may generate. According to the Internet Advertising Revenue Report by PricewaterhouseCoopers [42], in the first half of 2018, digital video advertising revenues on mobile devices comprise about 60% of all digital video ad revenues, reaching \$4.2 billion, and 220 million users stream 47 billion digital videos in a typical month. Based on this data, we suppose 28 billion digital videos is mobile stream (about 60% of total streaming) per month, equivalent to 168 billion digital videos for a half year. Therefore, per-video advertising revenue generated is estimated to be $a = \$0.025$. Watching Netflix uses about 1 GB of data per hour for each stream of standard definition video, and up to 3 GB per hour for each stream of HD video. Suppose the average time of digital video is 3 minutes, then the per-GB advertising revenue received by the content provider would be about $\frac{a}{\theta} = \$0.50$.

The range of both the relevance rate of pre-staged content (ξ) and the success rate of users' web content request (δ) is between 0 and 1. Figure 2 shows how the theoretical upper bound on content pre-staging rate (r) depends on the two parameters. Note Fig. 2 also illustrates the increased advertising revenue due to content pre-staging when the shared storage on user mobile device is normalized to one (i.e., $s_j - A_j^* = 1$), from Eq. (21). As shown, the range of maximum possible r is between \$0 and \$0.50. The best case for content pre-staging is when all pre-staged contents are relevant, i,e., $\xi = 1$, and no online content request can be completed, i.e., $\delta = 0$, then the carrier may charge the content provider a content pre-

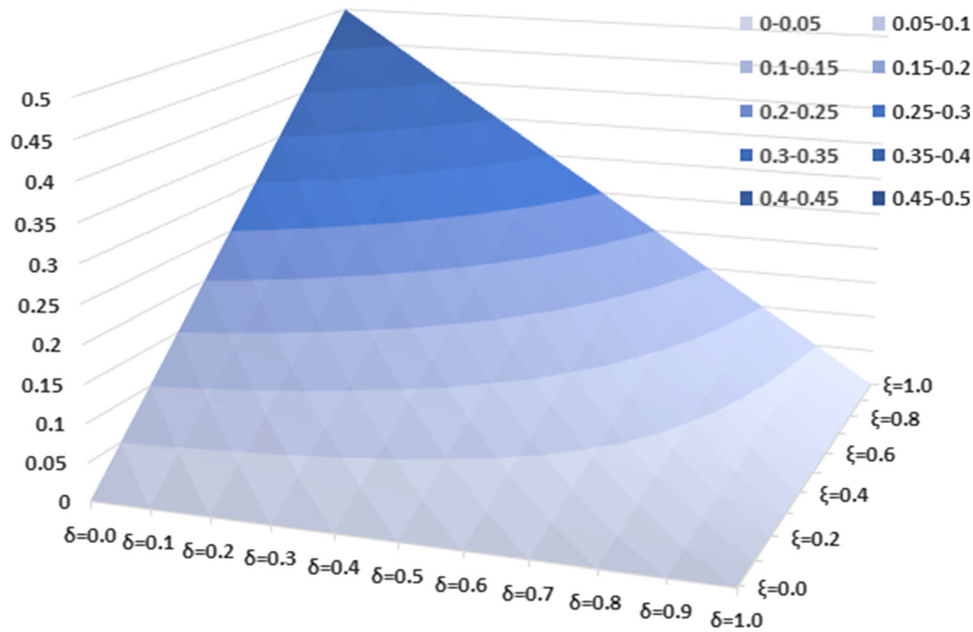


Fig. 2 Cap on content pre-staging rate (r): pre-staging rate is capped by the increased advertising revenue generated from content pre-staging that is increasing in content relevance rate and decreasing in web request success rate

staging rate up to per-GB advertising revenue of $r = \$0.50$. The worst case for content pre-staging is when none of the pre-staged content is relevant and all web content requests are fulfilled, then the carrier would not be able to charge content pre-staging fee on the content provider, i.e., $r = \$0$. The range of pre-staging rate is increasing in advertising revenue which sets the theoretical upper bound on the pre-staging rate. At any $\{\xi, \delta\}$ combination, as long as the actual content pre-staging rate is lower than that shown in Fig. 2, the welfare-improving condition for the content provider as in Eq. (23) is satisfied, i.e., the content provider is better off than the base case of no content pre-staging.

Next, we evaluate how the gain in the profitability of the mobile carrier depends on its price setting. According to the released wireless pricing data for the Big Four mobile network providers (AT&T, Verizon, Spring, and T-Mobile) [43], the average price of the Big Four for 1 GB of data across the board is \$13.01. Considering the fact that intense competition among the Big Four has driven down prices for years [44], we set the data access rate at $p = \$10$ per GB in the numerical example, which is converted into an operating cost of $b = \$8$ per GB if the carrier's operating margin is 20% [45].

The profit-driven mobile carrier would want to maximize the content pre-staging rate charged to the content provider ($r = r^*$) and minimize the reward paid to end users for pre-staging content on their mobile device ($q = 0$). At $\frac{a}{\theta} = \$0.50$ and $b = \$8$, the carrier's per-unit maximum possible profit received from accessing end users' mobile device storage is $\triangle\pi^* = 0.5\xi + 7.5\delta\xi$ from Eq. (25), as illustrated in Fig. 3. It is also the overall possible welfare gain of content pre-staging that is increasing both in the content relevance rate and the web request success rate. This is the optimal case for the mobile carrier as it would receive all the gains of content pre-staging, leaving the content provider no difference as in the base case of no content pre-staging.

The increased profit of mobile service business comes from two sources: increased advertising revenue and saved operating cost. In particular, the carrier's saved operating cost is $8 \times \delta \times \xi$ per GB for each unit of the mobile device storage used (from Eq. (22)). The cost benefit the carrier may receive from content pre-staging is increasing in both the content relevance rate and the web request success rate.

How total utility for user j changes with the relevance rate of pre-staged content and the success rate of web content request is shown in Fig. 4. We normalize the shared storage on user mobile device in Eq. (20) (i.e., $s_j - A_j^* = 1$). At $p = \$10$ and $\alpha = 0.5$. Content pre-staging on local mobile device storage increases users' consumption of content, which leads to the increase in user utility, hence end users are better off regardless whether they receive additional payments for sharing mobile device storage with the carrier. The plausible reward the carrier may offer is a bonus.

6.2 Profit sharing between carrier and content provider

As shown earlier in Fig. 3, the maximum possible gain that content pre-staging may generate to the mobile service derives from both the increased advertising revenue and the saved operating cost. The saved cost part of the gain is

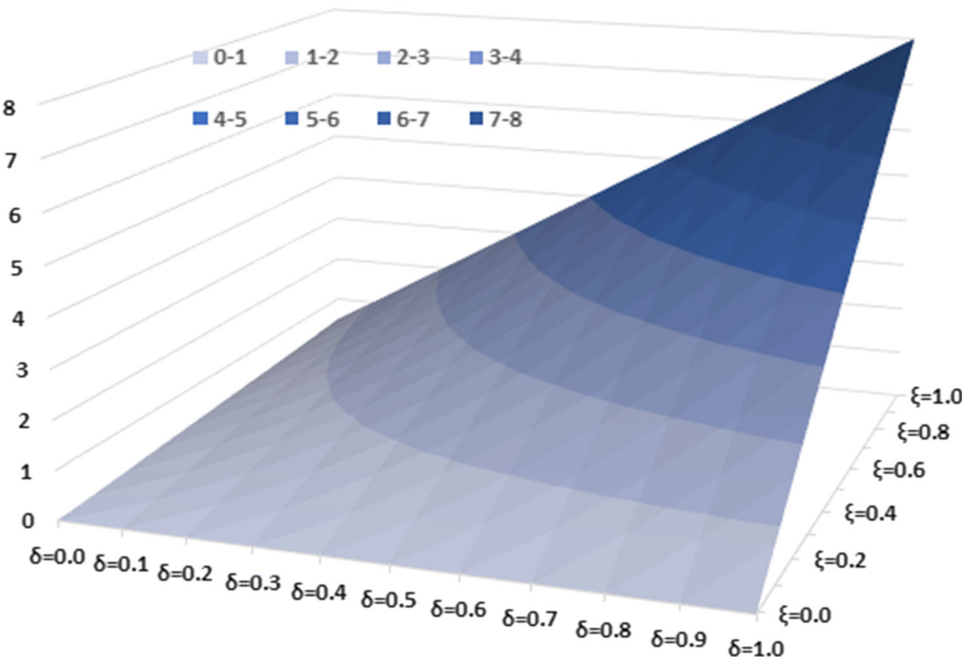


Fig. 3 Carrier's maximum possible pre-staging profit, which is also the overall monetary gain of content pre-staging that is increasing in both content relevance rate and web request success rate. In the carrier's optimal pricing case, the carrier receives all the increased profits, leaving the content provider equally well as in the base case

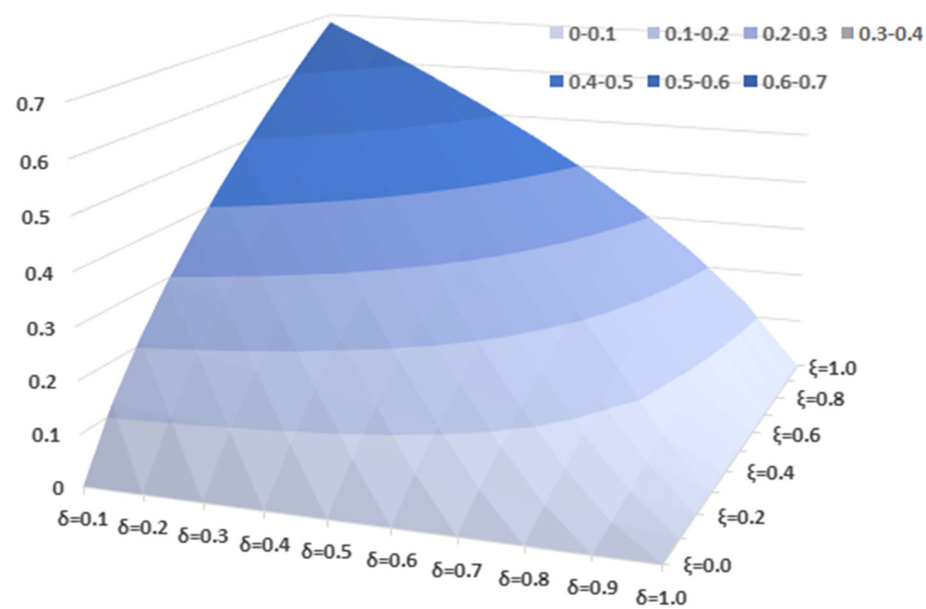
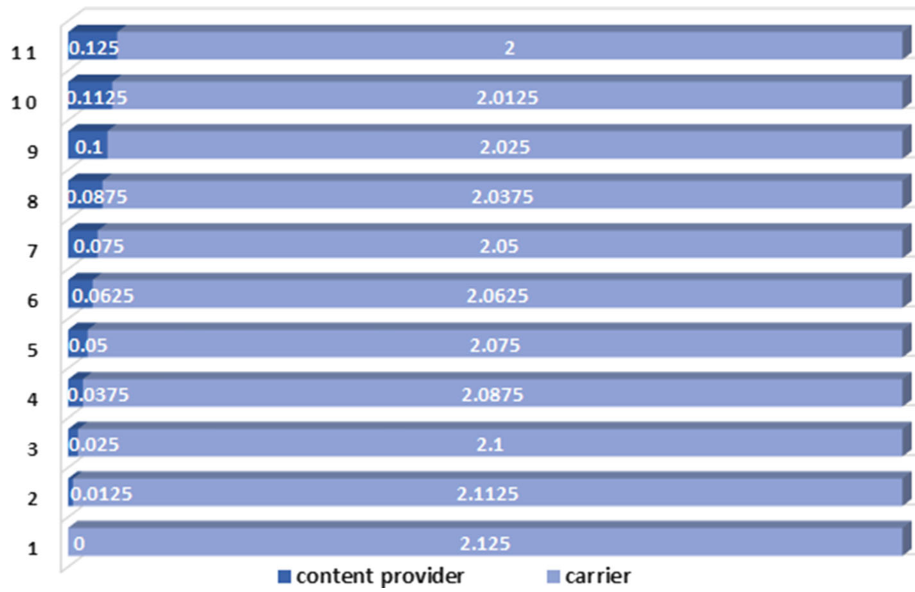


Fig. 4 Increased user utility per shared mobile storage: content pre-staging increases users' consumption of content. The increase in user utility is increasing in content relevance rate and decreasing in content request success rate

specific to the mobile carrier. Specifically, how is the gain from increased advertising revenue distributed between the carrier and the content provider depends on the content pre-staging rate the carrier charges the content provider. If the content provider is charged a pre-staging rate up to the cap, e.g., $r = \$0.50$ as in the numerical example, the content provider would be equally well as in the base case, thus lacking financial incentive to pre-stage content. As long as the pre-staging rate is below the cap, content pre-staging is mutually beneficial for the mobile carrier and the content provider. The two parties will share this part of the increased profit related to increased advertising revenue.

Figure 5 provides an example of profit sharing by the carrier and the content provider at $\delta = \xi = 0.5$. In this case, the per-unit increase in profit with content pre-staging is \$2.125, comprising \$0.125 in advertising revenue and \$2.0 in saved operating cost. The saved cost is specific to the carrier. The content provider can share the increased profit up to \$0.125. The split of the gain in advertising revenue between the carrier and the content provider is arbitrary. Figure 5 is a simplified linear case where the per-unit advertising revenue generated from increased content consumption is constant, or constant marginal revenue (MR) in economic terminology. In this case, the content

Fig. 5 Profit sharing by carrier and content provider: content pre-staging increases the profitability of mobile services via increased advertising revenue and saved capacity cost. The latter is specific to the carrier. The carrier and the content provider may share the former in an arbitrary way



provider's choice of pre-staging content on users' mobile devices is binary: it will pre-stage the maximum units if the pre-staging rate is below the increased advertising revenue; it will not choose to pre-stage if the pre-staging rate is above the increased advertising revenue.

Nevertheless, the marginal revenue is likely to be diminishing, i.e., additional advertising revenue is increasing in users' content access at a decreasing rate, so that the content pre-staging rate imposed on the content provider will affect the quantity of pre-staging. Figure 6 illustrates how the pre-staging rate (r) may affect the quantity of content pre-staging choice (Q) by the content provider (and hence social welfare of content pre-staging), and how the increased advertising profit is shared by the carrier and the content provider in case of diminishing marginal revenue.

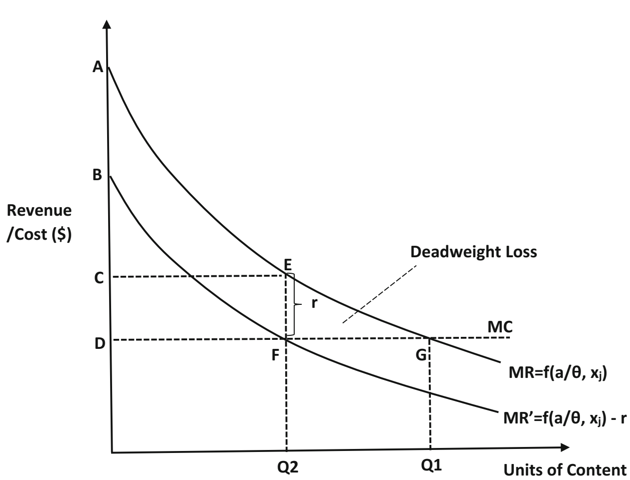


Fig. 6 When advertising revenue is increasing at a decreasing rate in pre-staged content, social welfare decreases as the content pre-staging rate increases, resulting in a deadweight loss

We suppose the marginal cost (MC) for the content provider to supply an additional unit of content is constant (assuming increasing marginal cost does not change any insights). In Fig. 6, the upper MR curve is the marginal revenue generated from an additional unit of content access by end users. In absence of content pre-staging fee (i.e., $r = 0$), the content provider chooses to pre-stage $Q1$ units of content. The profit (and social welfare) is represented by area ADG, which all goes to the content provider because there is no pre-staging fee. When the carrier charges a pre-staging rate of r, the content provider's marginal revenue curve shifts down and the vertical distance between the two MR curves is the pre-staging rate. At presence of the fee, the content provider chooses to pre-stage a quantity of $Q2 < Q1$. The total profit is area $ADFE$, of which the area of $CDFE$ goes to the carrier, and the area of ACE is received by the content provider. The area EFG is the foregone advertising revenue due to the fee, a net decrease in social welfare, i.e., the deadweight loss. Apparently, charging a content pre-staging fee benefits the carrier, harms the content provider, and the whole economy loses. Therefore, when marginal advertising revenue is decreasing, it is socially optimal to have free content pre-staging so that all the possible gains of pre-staging would be achieved. Nevertheless, since the carrier is a stakeholder rather than a benevolent social planner, it would charge a fee to earn profit. The larger the fee, the bigger the deadweight loss.

To sum, content pre-staging is socially beneficial. When the per-unit advertising revenue is unrelated to the total quantity of content consumption, the content pre-staging rate does not change the gain of pre-staging. It results in merely a transfer of profit between the content provider and the mobile carrier although the sharing of profit can be arbitrary. When the per-unit advertising revenue is decreasing in content consumption, an increase in pre-staging rate decreases quantity of content pre-staging chosen by the content provider, resulting in a deadweight loss. It would be socially optimal to set pre-staging rate at $r = 0$.

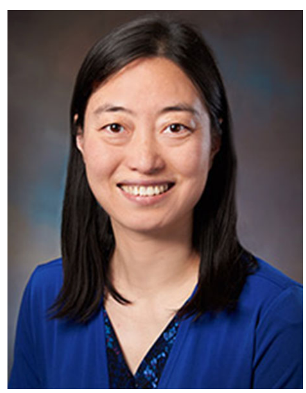
7 Conclusion

One of the largest challenges in wireless communications has been the dramatic growth in demands for data. In this paper, we studied the economic implications of pre-staging selected content on unused mobile device storage. The interactions among the mobile carrier, the content provider, and end users were formulated in a game theoretic framework as a mechanism design problem, in which the carrier designs the pre-staging policy and pricing mechanisms to maximize profit by predicting the best responses of end users and the content provider. Equilibrium solutions were derived and analyzed. We investigated the pricing strategies of the mobile carrier to reach the optimal operation point of the system by providing proper financial incentives to the content provider and end users. We showed that content localization via carrier-accessible mobile device storage is socially welfare enhancing, i.e., it increases the combined benefits of all interested parties and improves efficiency in capacity management of the mobile service business.

References

1. RajanishEmail, N., & Banakar, R. M. (2019). Cognitive radio: A conceptual future radio and spectrum sensing techniques—A survey. In S. Satapathy, K. Raju, K. Shyamala, D. Krishna, & M. Favorskaya (Eds.), *Advances in decision sciences, image processing, security and computer vision* (Vol. 4, pp. 155–165)., Learning and analytics in intelligent systems book series (LAIS) Cham: Springer.
2. Han, B., Hui, P., Kumar, V. A., Marathe, M. V., Pei, G., & Srinivasan, A. (2010). Cellular traffic offloading through opportunistic communications: A case study. In *Proceedings of the 5th ACM workshop on challenged networks*, Chicago, Illinois, September 24 (pp. 31–38).
3. Zhuo, X., Gao, W., Cao, G., & Hua, S. (2014). An incentive framework for cellular traffic offloading. *IEEE Transactions on Mobile Computing*, *13*(3), 541–555.
4. Chen, X., Wu, J., Cai, Y., Zhang, H., & Chen, T. (2015). Energy-efficiency oriented traffic offloading in wireless networks: A brief survey and a learning approach for heterogeneous cellular networks. *IEEE Journal on Selected Areas in Communications*, *33*(4), 627–640.
5. Zhou, H., Wang, H., Li, X., & Leung, V. C. M. (2018). A survey on mobile data offloading technologies. *IEEE Access*, *6*, 5101–5111.
6. Paschos, G. S., Iosifidis, G., Tao, M., Towsley, D., & Caire, G. (2018). The role of caching in future communication systems and networks. *IEEE Journal on Selected Areas in Communications*, *36*(6), 1111–1125.
7. Ji, M., Caire, G., & Molisch, A. F. (2015). Wireless device-to-device caching networks: Basic principles and system performance. *IEEE Journal on Selected Areas in Communications*, *PP*(99), 1.
8. Finamore, A., Mellia, M., Gilani, Z., Papagiannaki, K., Erramilli, V., & Grunenberger, Y. (2013). Is there a case for mobile phone content pre-staging? In *Proceedings of the 9th ACM conference on emerging networking experiments and technologies (CoNEXT'13)*, New York, NY (pp. 321–326).
9. Striegel, A., Hu, X., & Song, L. (2015). A case for making mobile device storage accessible by an operator. *IEEE COMSOC MMTC E-Letter*, *10*(1), 7–10.
10. Sen, S., Joe-Wong, C., Ha, S., & Chiang, M. (2013). A survey of smart data pricing: Past proposals, current plans, and future trends. *ACM Computing Surveys (CSUR)*, *46*(2), 1–37.
11. Aisamuddin, A., Shariff, M., Katuk, N., & Zakaria, N. H. (2017). An overview to pre-fetching techniques for content caching of mobile applications. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2–12), 37–43.
12. Cao, P. (2015). Opportunities and challenges for caching and prefetching on mobile devices. In *Proceedings of 3rd IEEE workshop on hot topics in web systems and technologies (HotWeb)*, Washington, DC, November 12–13.
13. Bao, X., Gowda, M., Mahajan, R., & Choudhury, R. R. (2013). The case for psychological computing. In *Proceedings of the 14th workshop on mobile computing systems and applications (HotMobile'13)*, no. Article No. 6, Jekyll Island, Georgia, February 26–27.
14. Ding, N., Wagner, D., Chen, X., Pathak, A., Hu, Y. C. & Rice, A. (2013). Characterizing and modeling the impact of wireless signal strength on smartphone battery drain. In *Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*, Pittsburgh, PA, June 17–21, 2013 (pp. 29–40).
15. Li, Y., Wang, Y., & Lan, T. (2018). Mobile ad prefetching and energy optimization via tail energy accounting. *IEEE Transactions on Mobile Computing*, *18*(9), 2117–2128.
16. Qian, F., Huang, J., Erman, J., Mao, Z. M., Sen, S. & Spatscheck, O. (2013). How to reduce smartphone traffic volume by 30%? In *Proceedings of the 14th international conference on passive and active measurement (PAM)*, Hong Kong, China, March 18–19, 2013 (pp. 42–52).
17. Zhang, L., Wu, W., & Wang, D. (2014). Time dependent pricing in wireless data networks: Flat-rate vs. usage-based schemes. In *Proceedings of IEEE INFOCOM*, Toronto, Canada, April 27–May 2, 2014.
18. Ha, S., Sen, S., Joe-Wong, C., Im, Y., & Chiang, M. (2012). Tube: Time-dependent pricing for mobile data. In *Proceedings of the ACM SIGCOMM conference on applications, technologies, architectures, and protocols for computer communication*, Helsinki, Finland, August 13–17, 2012 (pp. 247–258).
19. Joe-Wong, C., Ha, S., Sen, S., & Chiang, M. (2015). Do mobile data plans affect usage? Results from a pricing trial with ISP customers. In *Proceedings of the 16th international conference on passive and active network measurement (PAM)*, New York, NY, March 19–20, 2015 (pp. 96–108).
20. Andrews, M., Ozen, U., Reiman, M. I., & Wang, Q. (2013). Economic models of sponsored content in wireless networks with uncertain demand. In *Proceedings of IEEE INFOCOM workshop*

on smart data pricing (SDP), Turin, Italy, April 14–19, 2013 (pp. 345–350).
21. Joe-Wong, C., Ha, S., & Chiang, M. (2015). Sponsoring mobile data: An economic analysis of the impact on users and content providers. In *Proceedings of IEEE INFOCOM*, Kowloon, Hong Kong, April 26–May 1, 2015.
22. Joe-Wong, C., Sen, S., & Ha, S. (2018). Sponsoring mobile data: Analyzing the impact on internet stakeholders. *IEEE/ACM Transactions on Networking*, *26*(3), 1179–1192.
23. Zhang, L., Wu, W., & Wang, D. (2015). Sponsored data plan: A two-class service model in wireless data networks. In *Proceedings of the ACM SIGMETRICS international conference on measurement and modeling of computer systems*, Portland, OR, June 15–19, 2015 (pp. 85–96).
24. Hande, P., Chiang, M., Calderbank, R., & Rangan, S. (2009). Network pricing and rate allocation with content provider participation. In *Proceedings of IEEE INFOCOM*, Rio de Janeiro, Brazil, April 19–25, 2009 (pp. 990–998).
25. Andrews, M., Bruns, G., & Lee, H. (2014). Calculating the benefits of sponsored data for an individual content provider. In *Proceedings of the 48th annual conference on information sciences and systems (CISS)*, Princeton, NJ, March 19–21, 2014 (pp. 1–6).
26. Zhang, L., & Wang, D. (2014). Sponsoring content: Motivation and pitfalls for content service providers. In *Proceedings of IEEE conference on computer communications workshops (INFOCOM WKSHPS)*, Toronto, Canada, April 27–May 2, 2014 (pp. 577–582).
27. Thilakarathna, K., Jiang, F.-Z., Mrabet, S., Kaafar, M. A., Seneviratne, A., & Xie, G. (2017). Crowd-cache: Leveraging on spatio-temporal correlation in content popularity for mobile networking in proximity. *Computer Communications*, *100*, 104–117.
28. Lau, C. P., Alabbasi, A., & Shihada, B. (2018). An efficient content delivery system for 5G CRAN employing realistic human mobility. *IEEE Transactions on Mobile Computing*, *18*(4), 742–756.
29. Asadi, A., Wang, Q., & Mancuso, V. (2014). A survey on device-to-device communication in cellular networks. *IEEE Communications Surveys & Tutorials*, *16*(4), 1801–1819.
30. Malak, D., Al-Shalash, M., & Andrews, J. G. (2016). Optimizing content caching to maximize the density of successful receptions in device-to-device networking. *IEEE Transactions on Communications*, *64*(10), 4365–4380.
31. Wu, Y., Yao, S., Yang, Y., Zhou, T., Qian, H., Hu, H., et al. (2016). Challenges of mobile social device caching. *IEEE Access*, *4*, 8938–8947.
32. Chen, Z., Liu, Y., Zhou, B., & Tao, M. (2016). Caching incentive design in wireless d2d networks: A stackelberg game approach. In *IEEE international conference on communications (ICC)*, Kuala Lumpur, Malaysia, May 23–27, 2016.
33. Li, Z., Liao, Q., & Striegel, A. D. (2016). On the economics of mobile content pre-staging. In *The 5th IEEE INFOCOM workshop on smart data pricing (SDP)*, San Francisco, CA, April 11, 2016.
34. Tadrous, J., Eryilmaz, A., & Gamal, H. E. (2014). Joint pricing and proactive caching for data services: Global and user-centric approaches. In *Proceedings of IEEE INFOCOM workshop on smart data pricing (SDP)*, Toronto, ON, April 27–May 2, 2014 (pp. 616–621).
35. Ma, R. T. B., & Towsley, D. (2015). Cashing in on caching: On-demand contract design with linear pricing. In *Proceedings of the 11th ACM conference on emerging networking experiments and technologies (CoNEXT)*, Heidelberg, Germany, December 1–4, 2015.
36. Alotaibi, F., Hosny, S., Tadrous, J., Gamal, H. E., & Eryilmaz, A. (2015). Towards a marketplace for mobile content: Dynamic pricing and proactive caching. arXiv preprint arXiv: 1511.07573.
37. Arrow, K. J. (1971). The theory of risk aversion. In K. Arrow (Ed.), *Essays in the theory of risk-bearing* (pp. 90–120). Chicago: Markham.
38. Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*(1–2), 122–136.
39. Lyons, D. (2015). Google wireless and the evolution of usage-based pricing. *TechPolicyDaily.com*, May 6, 2015. Available: http://www.techpolicydaily.com/internet/google-usage-based-pricing/
40. Agarwal, S. (2006). Intelligent content caching for mobile devices. In *Proceedings of the 13th international conference on telecommunications*, May 9–12, 2006 (p. 4).
41. Quan, W., Liu, Y., Jiang, X., & Guan, J. (2016). Intelligent popularity-aware content caching and retrieving in highway vehicular networks. *EURASIP Journal on Wireless Communications and Networking*. https://doi.org/10.1186/s13638-016-0688-z.
42. PWC, IAB internet advertising revenue report: 2018 first six months results. *PricewaterhouseCoopers*, November 2018.
43. Louis, T. (2013). The real price of wireless data, *Forbes*, September 22, 2013.
44. Leubsdorf, B. (2017). How cell-phone plans with unlimited data limited inflation. *The Wall Street Journal*, May 19, 2017.
45. DeGrasse, M. (2017). How carriers are controlling network opex, *RCR Wireless News*, September 22, 2017.



Zhen Li is currently an E. Maynard Aris Endowed Professor of Economics in the Department of Economics and Management at Albion College. She received her Master's Degree and Ph.D. in Economics from Princeton University under the direction of Dr. Michael Woodford. She graduated with her Bachelor's Degree in International Economics from Peking University. Dr. Li conducted research on applied macroeconomics and international finance, in particular on international financial integrity and related policy issues. Dr. Li's recent research interests include inter-disciplinary research study on economics and game theory of computer networks and information security.

Qi Liao is currently a Professor of Computer Science at Central Michigan University (CMU). He received his M.S. and Ph.D. in Computer Science and Engineering (CSE) from the University of Notre Dame, and a B.S. and departmental distinction in Computer Science (minor in Mathematics) from Hartwick College, New York. Dr. Liao's research interests include computer security, machine learning, visual analytics, and economics/game theory at the intersection of network usage and cybersecurity. He received best paper awards at USENIX LISA, IEEE ICCCBDA, Emerald Literati Awards for Excellence for Information and Computer Security, IEEE VAST Challenge Award, winner of National Security Innovation Competition, Center for Research Computing Award for Computational Sciences and Visualization, and CMU College of Science & Engineering Award for Outstanding Research. Dr. Liao was a visiting research scientist at IBM Research, Argonne National Lab, and ASEE Fellow at U.S. Air Force Research Lab.

Aaron D. Striegel is currently a Professor of Computer Science and Engineering at the University of Notre Dame. He also serves on the Executive Committee of the Wireless Institute at the University of Notre Dame. Prof. Striegel received his Ph.D in 2002 in Computer Engineering at Iowa State University under the direction of Dr. G. Manimaran. Prof. Striegel's research interests focus on instrumenting the wireless networked ecosystem to gain insight with respect to user behavior and optimizing network performance. Flagship projects of Prof. Striegel include the NetSense, NetHealth, and Tesserae involving the instrumentation and analysis of data from hundreds of smartphones and wearables over a nearly seven-year period of continuous data streaming. Further research interests of Prof. Striegel include heterogeneous network optimization (cellular, WiFi), content distribution via edge device pre-staging, and network security dynamics.