

Optimal Pricing of Internet of Things: A Machine Learning Approach

Mohammad Abu Alsheikh, *Member, IEEE*, Dinh Thai Hoang, *Member, IEEE*, Dusit Niyato, *Fellow, IEEE*, Derek Leong, Ping Wang, *Senior Member, IEEE*, and Zhu Han, *Fellow, IEEE*

Abstract—Internet of things (IoT) produces massive data from devices embedded with sensors. The IoT data allows creating profitable services using machine learning. However, previous research does not address the problem of optimal pricing and bundling of machine learning-based IoT services. In this paper, we define the data value and service quality from a machine learning perspective. We present an IoT market model which consists of data vendors selling data to service providers, and service providers offering IoT services to customers. Then, we introduce optimal pricing schemes for the standalone and bundled selling of IoT services. In standalone service sales, the service provider optimizes the size of bought data and service subscription fee to maximize its profit. For service bundles, the subscription fee and data sizes of the grouped IoT services are optimized to maximize the total profit of cooperative service providers. We show that bundling IoT services maximizes the profit of service providers compared to the standalone selling. For profit sharing of bundled services, we apply the concepts of core and Shapley solutions from cooperative game theory as efficient and fair allocations of payoffs among the cooperative service providers in the bundling coalition.

Index Terms—Internet of Things (IoT), IoT pricing, IoT bundling, machine learning.

I. INTRODUCTION

Recent years have witnessed significant progress in using machine learning for solving challenging problems in the Internet of things (IoT) including data caching [1], [2], activity recognition [3], [4], channel estimation [5], and security [6], [7]. Such machine learning-based solutions require massive datasets for system training. Trading IoT data among firms allows data vendors to make profit by offering their data to service providers over the Internet. Service providers use the bought data in creating and training advanced machine learning-based IoT services¹, e.g., fraud detection, activity recognition, acoustic modeling, and medical diagnosis. However, economics of IoT services among service providers, data vendors, and customers is rarely studied in the literature. IoT market models and optimal pricing schemes are therefore

required to ensure maximum profits of firms by achieving optimal utilization of the IoT resources and optimal subscription fees for services.

IoT services can be either sold separately or bundled and sold as one service package. In particular, multiple service providers, e.g., fraud detection and recommender systems, can cooperate to sell a bundled service at a discounted rate while sharing the resulting profit². Several major questions related to this service bundling process arise. *Firstly, how is service quality defined and what are the optimal sizes of data that should be bought from data vendors? Secondly, should service providers cooperate to offer a bundled service instead of the standalone sales of services? Thirdly, once a cooperation is formed, how do cooperative service providers divide the resulting profit of a bundled service among themselves?*

This paper provides answers for the aforementioned questions by proposing IoT market models and optimal pricing schemes of selling machine learning-based IoT services separately or as bundled packages. The main objective is maximizing the profits of IoT service providers while providing bundled services to customers at discounted price. The key contributions of this paper can be summarized as follows:

- From the service provider's perspectives, we define the service quality as a mapping between the bought data size and the resulting accuracy of machine learning algorithms.
- We develop an IoT market model for the standalone selling of services. We then propose a non-linear optimization problem with the objective of maximizing the profit of service providers. Specifically, a service provider will make decisions on the optimal data size to buy from data vendors and the optimal subscription fee that should be charged to service customers.
- We formulate and solve an IoT service bundling optimization to obtain the optimal bundle subscription fee and data sizes for each service in the bundle. Here, the optimization objective is to maximize the total profit of the cooperative providers inside the bundling coalition. We show that a bundled service increases the profit of service providers compared to the standalone sales of services. Moreover, bundled IoT services are favored by customers due to their discounted subscription fees compared to the standalone sales.

M. Abu Alsheikh is with the Faculty of Science & Technology, University of Canberra, ACT, Australia 2617 (email: mabualsh@ieee.org). D. T. Hoang is with the School of Electrical and Data Engineering, University of Technology Sydney, NSW, Australia 2007 (email: hoang.dinh@uts.edu.au). D. Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (email: dniyato@ntu.edu.sg). P. Wang is with the Department of Electrical Engineering & Computer Science, York University, Canada M3J 1P3 (email: pingw@yorku.ca). Z. Han is with the University of Houston, TX, USA 77004 (email: zhan2@uh.edu), and also with the Department of Computer Science and Engineering, Kyung Hee University, South Korea 446-701.

¹For the rest of this paper, we use "machine learning-based IoT services" and "services" interchangeably.

²Product bundling is a marketing strategy which is widely used by economists, e.g., a fast food meal consisting of a sandwich and a soft drink.

- We present a profit sharing scheme to divide the generated profit of a bundled IoT service among the cooperative providers based on their contributions to the bundle profit. We apply the concepts of core solution and Shapley value to find the payoff allocations to providers.

Additionally, we provide closed-form solutions for each optimization problem in this paper. Unlike iterative solutions of nonlinear problems with inequality constraints, e.g., interior point methods [8], closed-form solutions can be evaluated in a finite time and are computationally efficient.

The rest of this paper is organized as follows. Section II presents the related work. Section III discusses the IoT market model and assumptions, and machine learning-based model is presented as a measure of service qualities. Then, optimization problems for profit maximizing are derived in Section IV for the standalone services and in Section V for the bundled IoT services. Section VI presents a model for payoff allocation and sharing among service providers in service bundling. Section VII discusses the experimental evaluation results. The paper is finally concluded in Section VIII.

II. RELATED WORK

A survey of 3,000 employees from 100 countries reports that successful firms apply machine learning five times more than low-performing ones [9]. This clearly shows the importance of adopting the recent advances in machine learning for generating revenues and meeting the market demands on intelligent IoT systems. Moreover, market models and pricing strategies are integral for maximizing the profit of selling products and services including wired and wireless network access [10]–[12], cloud computing [13], and mobile crowdsensing [14], [15], just to name a few.

A. Economics of Information Goods

Pricing information goods, e.g., software products and movies, is a well-studied problem in the literature. In [16], three pricing schemes of information goods were presented, namely, connection time-based, search-based, and subscription-based pricing. It has been argued that connection time-based pricing is less profitable than the other schemes for highly skilled users. The authors of [17] presented a nonlinear pricing scheme of information goods based on the customer usage, i.e., defining a fixed price to each usage level. In [18], the authors discussed the pay-per-use and unlimited subscription of information goods. It is shown that a maximum profit is achieved through competitive pricing. The authors in [19] discussed the benefits of bundling information goods. It is shown that an accurate estimation of the user behavior can be achieved for bundled goods compared to the individual sales of products.

Pricing and bundling of IoT services is more challenging than those of classical information goods, as the quality of information goods can be easily determined, e.g., the quality of a software is defined based on its supported features, while the casting and genre of a movie define its price. On the contrary, the quality of IoT services cannot be directly measured.

B. Economics of Query-based Data Services

Query-based data services extract data from structured databases and visualize them according to customer requests. Simple query pricing models can be preferable by buyers, but applying a complex pricing model increases the revenues of the content provider [20]. The authors in [21] presented a query-based pricing scheme of data. The price of a query, i.e., a question, is defined based on the number of data views required to provide the answer. In [22], the authors presented a pricing method in privacy-preserving query systems. This differential privacy pricing aims at maximizing the accuracy of queries while maintaining the privacy of the data owners. The privacy level is defined by the query buyers. The authors in [23] proposed a query pricing scheme with arbitrage-prevention. This prevents buyers from combining simple and cheap queries to achieve the same results of a complex and expensive query. Such existing query-based schemes are restricted by design to structured and relational data, and they do not address unstructured data, e.g., IoT data, which is the dominant form of data in real-world settings.

C. Incentive Mechanisms for Crowdsensing

Incentive mechanisms are required for maintaining high user participation and providing fair reward allocation. The authors in [24] presented a user-centric incentive mechanism formulated as a reverse auction, and a platform-centric incentive scheme modeled as a Stackelberg game. The user-centric mechanism enables the participants to compete for a higher reward. In the platform-centric, mechanism, the service acts as the game leader and announces the reward of completing the data collection task. The participants are the game followers and set their sensing time in order to maximize the received reward. Likewise, the authors in [25] modeled the crowdsensing problem as an online auction, while assuming that the arrival of participants has a stochastic distribution. In [26], the reward allocation problem was modeled as an all-pay auction such that the participant with the highest contribution receives the full reward.

This paper is fundamentally different from all past works. First, the problem of defining the quality of service in IoT has not been previously addressed. Second, existing works have not considered the data demand and pricing in IoT services when machine learning is heavily utilized. Third, none of the existing works has considered the problem of cooperation among service providers to form a bundled service which is an effective strategy for profit maximization. This paper addresses these limitations and presents a market model, bundling strategies, and optimal pricing schemes of IoT services. The novel data-driven optimizations of this paper allow achieving the maximum profits from offering IoT services to customers.

III. SYSTEM MODEL

In this section, we first briefly introduce the concept of machine learning-based IoT services and give real world examples of IoT services where optimal pricing and bundling are required. Then, we present a method for defining the value of data from a machine learning perspective.

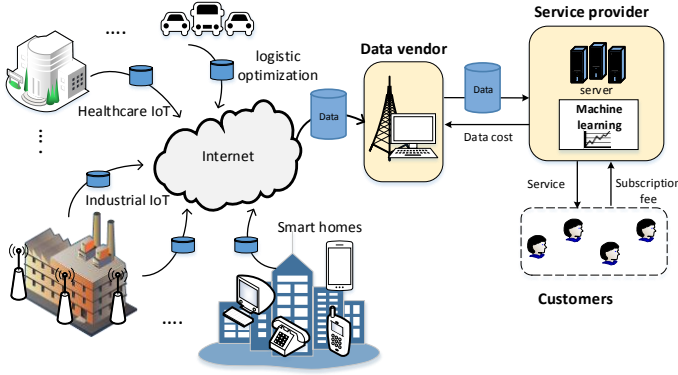


Fig. 1: System model of machine learning-based IoT services.

A. Machine Learning-Based IoT Services

We consider the IoT service architecture shown in Figure 1. An IoT service is typically composed of a data vendor, service provider, and service customers.

- *Data vendors*: The IoT data is firstly collected, stored and filtered by a data vendor. The data can be generated by different devices and technologies, e.g., sensor nodes, IoT gadgets, and smart phones. In addition to the deployment cost, the process of data gathering also requires costly human intervention for data annotation, validation, and preprocessing, e.g., anomaly detection and missing data imputation. Accordingly, the data vendor charges data buyers with a price. The cost of one data unit³ is denoted by c .
- *Service provider*: The raw data owned by data vendors is unprofitable unless suitable analytics tools are applied. A service provider is a business entity which buys data from one or more data vendors, uses machine learning tools, and offers a service to customers willing to pay a subscription fee. For profit maximization, a rational service provider decides the data size n (in data units) to buy from data sources and the subscription fee p_s to charge for his service.
- *Service customers*: We assume that there are a total of M customers. Each customer is an independent entity who decides whether to subscribe to a service provider based on his willingness-to-pay θ for the service and the subscription fee p_s set by the provider. θ defines the maximum subscription fee that a customer can pay for a service based on his evaluation and need for the service.

This IoT market model is useful in many service-oriented architectures.

- A *data marketplace* is an online store where entities can trade data. machine learning, e.g., machine learning algorithms, can be applied on the data to generate prediction models. Datasets of various types are offered and exchanged as assets. Examples of data markets include Azure Marketplace⁴, Qlik DataMarket⁵, and Infochimps⁶.

TABLE I: List of frequently used notations and symbols throughout the paper.

SYMBOL	DEFINITION
\mathcal{D}	Dataset
n	Requested data size
c	Cost of one data unit
M	Number of customers
p_s	Subscription fee
θ	Willingness-to-pay (reservation price) for a service by customers
$F(\cdot)$	Profit function of a service provider
\mathbf{H}	Hessian matrix of $F(\cdot)$
Δ_i	i th-order principal minor of \mathbf{H}
$\mathcal{L}(\cdot)$	Lagrangian dual problem
λ_i	Lagrangian multiplier
p_b	Bundle subscription fee
$q(\cdot)$	Service quality function
α	Fitting parameters of $q(\cdot)$
\mathcal{C}	The core solution set
\mathcal{K}	Bundling coalition
$F_{\mathcal{K}}(\cdot)$	Profit function of a bundling coalition \mathcal{K}
$\mathbf{H}_{\mathcal{K}}$	Hessian matrix of $F_{\mathcal{K}}(\cdot)$
$\mathcal{L}_{\mathcal{K}}(\cdot)$	Lagrangian dual function of a bundling coalition \mathcal{K}
φ_k	Payoff allocation for provider $k \in \mathcal{K}$ in the core solution
η_k	Payoff allocation for provider $k \in \mathcal{K}$ by the Shapley solution

- *Crowdsourcing services* also require optimized market models for data exchange. *Placemeter*⁷, for example, provides real-time information of pedestrian and vehicular movement in cities and urban areas. Placemeter allows users to upload videos of streets and public areas and they are paid back based on the video quality. Computer vision algorithms are used by Placemeter to extract information from real-time data.
- *IoT services* is a new trend of IoT platforms which are designated to visualize and trade IoT data. *Thingful*⁸ is an example of these platforms which allows IoT vendors and owners to visualize geolocations of connected devices around the world. Likewise, health care systems, such as *PatientsLikeMe*⁹, sell rich medical data which can be collected with IoT gadgets and other smart technologies.

The list of symbols used in this paper are summarized in Table I.

B. Machine Learning in IoT

As shown in Figure 2, consider a training data $\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^B$ which includes B tuples of connected feature set $\mathbf{x}_i \in \mathbb{R}^R$ and a class label $y_i \in \mathbb{R}$. y_i is only present in supervised learning, e.g., classification and regression problems. Based on the problem at hand, a feature set \mathbf{x}_i contains sensing data of R features such as images in vision problems, audio in acoustic modeling, text in document classification, and temperature values in weather forecasting. Generally, a machine learning optimization problem adheres to the following general formula:

$$\underset{\Psi}{\text{minimize}} \quad e(\mathcal{D}; \Psi) + \gamma g(\Psi), \quad (1)$$

³We assume that a data unit includes one percent of the full dataset samples.

⁴<https://www.datamarket.azure.com>

⁵<https://www.qlik.com>

⁶<http://www.infochimps.com>

⁷<https://www.placemeter.com>

⁸<https://www.thingful.net>

⁹<https://www.patientslikeme.com>

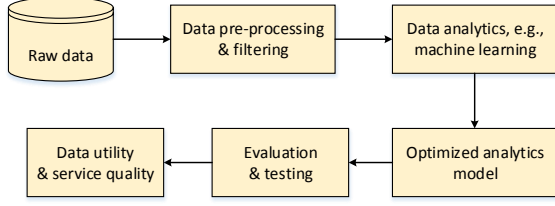


Fig. 2: Applying machine learning in IoT services.

where $e(\cdot)$ is a learning objective function designed to fit a model Ψ to historical data D , $g(\cdot)$ is a regularization term, γ is a weighted hyper-parameter. For example, a deep network [27] has a learning function which is defined as follows:

$$\underbrace{\frac{1}{B} \sum_{i=1}^B \left(\frac{1}{2} \|h_{\Psi}(\mathbf{x}_i) - y_i\|^2 \right)}_{e(D; \Psi)} + \underbrace{\frac{\gamma}{2} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} \left(\Psi_{ji}^{(l)} \right)^2}_{g(\Psi)}, \quad (2)$$

where L is the number of layers in the deep model, s_l is the number of neurons at layer l , $h_{\Psi}(\mathbf{x}_i)$ is the deep model prediction for input \mathbf{x}_i . The first term is an average sum-of-squared errors, the second term is a weight decay regularization which includes the summation modeling parameters. After the model is trained using the training data, it is tested to define the accuracy using the unseen testing data.

C. Service Quality

One of the key barriers in the development of data marketplaces is defining data quality and value to the potential service providers. In this section, we define the data quality based on the performance of machine learning models trained using the data.

Customers do not always subscribe to the cheapest IoT services. Instead they infer the subscription fee and service quality. More data is important to increase the accuracy of IoT services [28], [29]. However, this accuracy gain increases the cost of bought data. From the service provider's perspectives, we define the quality function $q(\cdot)$ of a service as a mapping between the bought data size to the resulting accuracy of the machine learning. The quality of the service is equal to the utility of bought data from the service provider's perspectives. For example, machine learning in intrusion detection systems by energy-constrained sensor networks is aimed for locating and identifying intruders [30]. The cost of detecting an intruder is a direct reflection of the value of data.

We introduce the service quality function $q(\cdot)$ to meet the following empirical assumptions:

- $q(\cdot)$ is an increasing function such that $q'(\cdot) > 0$. This assumption is intuitive as more data improves the accuracy of the analytics and quality of the service.
- $q(\cdot)$ has a decreasing marginal utility such that $q''(\cdot) \leq 0$. This assumption reflects the empirical accuracy of machine learning models.

In our data market and pricing framework, $q(\cdot)$ is defined as follows:

$$q(n; \alpha) = \alpha_1 - \alpha_2 \exp(-\alpha_3 n), \quad (3)$$

where n is the data size and $\alpha = [\alpha_1 \ \alpha_2 \ \alpha_3]$ is a tuple of three fitting parameters. To find the fitting parameters α , we vary the size of raw data D used to fit the model Ψ . Specifically, a series of O experimentation points $(n^{(1)}, \varepsilon^{(1)}), \dots, (n^{(j)}, \varepsilon^{(j)}), \dots, (n^{(O)}, \varepsilon^{(O)})$ is performed, where $\varepsilon^{(j)}$ is the service accuracy resulting from a data size of $n^{(j)}$ with $n^{(j)} > n^{(j+1)}$. $q(n; \alpha)$ is then found by applying nonlinear least squares for minimizing the sum of squared errors as follows:

$$\underset{\alpha}{\text{minimize}} \quad \frac{1}{O} \sum_{j=1}^O \left\| \varepsilon^{(j)} - q(n^{(j)}; \alpha) \right\|^2. \quad (4)$$

The parameter fitting problem in (4) can be solved iteratively to find the fitting parameter α [31]. O can be estimated before transmitting the data from the data vendor to the service provider through a third party broker which charges a broker fee. In real-world IoT services, the broker can also manage the financial transactions, service performance, and service delivery, e.g., measuring the compliance with the service-level agreement (SLA).

In Section VII-B, we present extensive experiments that show the validity of (3) in fitting the real accuracy of machine learning-based IoT services. For simplified notations, we use q instead of $q(n; \alpha)$ in the rest of this paper.

IV. IoT MARKET MODEL AND OPTIMAL PRICING OF STANDALONE SERVICES

In this section, we first develop an IoT market model of selling standalone IoT services, and we formulate an optimization problem to maximize the profit of a service provider. Then, the closed-form solutions for the optimal subscription fee and requested data size are provided.

We consider the separate sales of a monopolist service provider providing a service to a set of M potential customers as shown in Figure 1. Each customer has a different willingness-to-pay θ for the data service. If the willingness-to-pay of a particular customer is higher than or equal to the subscription fee p_s weighted by the service quality q such that $\theta q \geq p_s$, that customer will subscribe to the service. This indicates that the willingness-to-pay value depends on the customer evaluation of the service as well as its quality, i.e., high quality services will attract more customers. Based on our real-world customer surveys (Figure 6), we show that θ follows a uniform distribution. Then, the profit of the service provider is found as follows:

$$F(p_s, n) = \underbrace{Mp_s \mathbb{P}(\theta q \geq p_s)}_{\text{revenue}} - \underbrace{nc}_{\text{data cost}} \quad (5)$$

$$= Mp_s \left(1 - \frac{p_s}{q} \right) - nc \quad (6)$$

where $q = \alpha_1 - \alpha_2 \exp(-\alpha_3 n)$ is the service quality function. The first term of (5) is the revenue of offering the service to the customers with the subscription fee p_s . As θ is assumed

to follow a uniform distribution, the expression $\mathbb{P}(\theta q \geq p_s) = 1 - \frac{p_s}{q}$ defines the probability of subscribing to the service by the customers. The second term of (5) is the total data price paid to the data vendor which is equal to the requested data size multiplied by the data unit cost. The marginal cost of running the service is negligible.

We next formulate a profit maximization optimization based on the profit function in (5) for the service provider to decide their optimal requested data size and subscription fee.

A. Profit Maximization of Service Providers

The service provider is assumed to be rational and is interested in maximizing his individual profit. A non-linear optimization problem can be formulated to choose the optimal data size n^* that should be bought from the data vendor and optimal subscription fee p_s^* to be charged to the customers as follows:

$$\begin{aligned} \underset{n, p_s}{\text{maximize}} \quad & F(p_s, n) = Mp_s \left(1 - \frac{p_s}{q}\right) - nc \\ \text{subject to } & C1: p_s \geq 0, \\ & C2: n \geq 0. \end{aligned} \quad (7)$$

The objective function of (7) maximizes the profit of the service provider. The two constraints $C1$ and $C2$ are required to ensure non-negative solutions for the subscription fee and requested data size.

Proposition 1. $F(p_s, n)$ is concave, and hence the closed-form solution of the optimization problem $\underset{p_s, n}{\text{maximize}} F(p_s, n)$ is globally optimal.

Proof: We use Sylvester's criterion of twice differentiable functions to prove that the Hessian matrix \mathbf{H} of $F(p_s, n)$ is negative semidefinite, and hence the concavity of $F(p_s, n)$ [33]. Suppose that the principal minors of \mathbf{H} are denoted as $\{\Delta_i\}_{i=1,2}$, where i is the order of the principal minors. Based on Sylvesters criterion, \mathbf{H} is negative semidefinite if and only if the condition $(-1)^i \Delta_i \geq 0$ holds, i.e., every odd-order principal minor is non-positive and every even-order principal minor is non-negative. The Hessian matrix \mathbf{H} of $F(p_s, n)$ is obtained as shown in (11) at the top of the next page. The first-order principal minors of \mathbf{H} are derived as follows:

$$\Delta_{1,1} = \frac{-2M}{q} \leq 0, \quad (8)$$

$$\begin{aligned} \Delta_{1,2} = & \frac{-2M\alpha_2^2\alpha_3^2p_s^2 \exp(-2\alpha_3n)}{q^3} \\ & - \frac{M\alpha_2\alpha_3^2p_s^2 \exp(-\alpha_3n)}{q^2} \leq 0. \end{aligned} \quad (9)$$

The principal minor of order two is derived as follows:

$$\Delta_{2,1} = \frac{2\alpha_2(M\alpha_3p_s)^2}{q^3} \geq 0. \quad (10)$$

Accordingly, Sylvester's criterion is satisfied. \mathbf{H} is negative semidefinite and $F(p_s, n)$ is concave. ■

We next apply the Karush–Kuhn–Tucker (KKT) conditions to find the optimal solutions of (7). The KKT conditions are necessary and sufficient conditions for optimality when applied to concave functions [8].

B. Optimal Subscription Fee and Requested Data Size

Converting the constrained profit maximization problem (7) into a minimization problem and formulating the Lagrangian dual problem results in the following unconstrained Lagrange dual function:

$$\mathcal{L}(\{p_s, n\}, \lambda_1, \lambda_2) = -Mp_s \left(1 - \frac{p_s}{q}\right) + nc - \lambda_1 p_s - \lambda_2 n, \quad (12)$$

where λ_1 and λ_2 are Lagrange multipliers for the constraints $C1$ and $C2$, respectively. The first derivatives of (12) with respect to p_s and n are as follows:

$$\frac{\partial \mathcal{L}(\cdot)}{\partial p_s} = M \left(\frac{p_s}{q} - 1\right) + \frac{Mp_s}{q} - \lambda_1, \quad (13)$$

$$\frac{\partial \mathcal{L}(\cdot)}{\partial n} = c - \lambda_2 - \frac{M\alpha_2\alpha_3p_s^2 \exp(-\alpha_3n)}{q^2}, \quad (14)$$

where $\lambda_1 = 0$ and $\lambda_2 = 0$ (no active constraints). Setting both derivatives to zero, the optimal closed-form solutions are found as follows:

$$p_s^* = \frac{M\alpha_1\alpha_3 - 4c}{2M\alpha_3}, \quad (15)$$

$$n^* = \frac{1}{\alpha_3} \log \left(\frac{M\alpha_2\alpha_3}{4c} \right), \quad (16)$$

where the condition $M\alpha_1\alpha_3 \geq 4c$ must hold to ensure positive values of n^* and p_s^* . Accordingly, the maximum profit $F(p_s^*, n^*)$ of the provider is found by substituting the optimal values of n^* and p_s^* in (5).

V. IOT MARKET MODEL AND OPTIMAL PRICING OF BUNDLED SERVICES

In this section, we first present a market model where bundling is used by two cooperative service providers (denoted as Service 1 and Service 2). We denote the bundling coalition as \mathcal{K} . The two services are grouped together in one package and sold at a discounted subscription fee, i.e., subscribing to the two services separately costs more than the bundled service. It is important to note that customers can still subscribe to services separately. Then, we develop an optimization problem to select the optimal bundle subscription fee and requested data sizes by both providers.

A. Profit Maximization of Bundled Services

Consider two service providers that form a coalition \mathcal{K} to provide a bundled service to a base of M customers as shown in Figure 3. The subscription fee of the bundle is denoted as p_b . The qualities of Services 1 and 2 are denoted as q_1 and q_2 , respectively. The data unit costs paid by Services 1 and 2 are c_1 and c_2 , respectively. We define the reservation prices θ_1 and θ_2 as the willingness of the customers to pay for Services 1

$$\mathbf{H} = \begin{bmatrix} \frac{-2M}{q} & \frac{2M\alpha_2\alpha_3p_s \exp(-\alpha_3n)}{q^2} \\ \frac{2M\alpha_2\alpha_3p_s \exp(-\alpha_3n)}{q^2} & \frac{-2M\alpha_2^2\alpha_3^2p_s^2 \exp(-2\alpha_3n)}{q^3} - \frac{M\alpha_2\alpha_3^2p_s^2 \exp(-\alpha_3n)}{q^2} \end{bmatrix}. \quad (11)$$

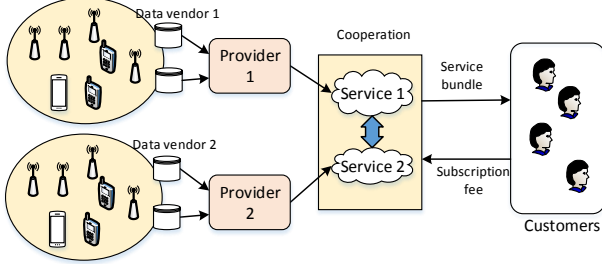


Fig. 3: System model of offering IoT services as one bundle.

and 2 in the bundle. Specifically, a customer will decide to subscribe to a bundled service if the following relation hold

$$\theta_1 q_1 + \theta_2 q_2 \geq p_b \quad (17)$$

which indicates that the customer evaluation of the bundled service is higher than its subscription fee. The profit optimization problem of a bundled service is then defined as follows:

$$\begin{aligned} & \underset{p_b, n_1, n_2}{\text{maximize}} F_K(p_b, n_1, n_2) = \\ & \quad \underbrace{Mp_b \mathbb{P}(\theta_1 q_1 + \theta_2 q_2 \geq p_b)}_{\text{revenue}} - \underbrace{n_1 c_1 - n_2 c_2}_{\text{data cost}} \\ & \text{subject to } C1, C2, \\ & \quad C3: p_b \geq 0, \\ & \quad C4: n_1 \geq 0, \\ & \quad C5: n_2 \geq 0, \end{aligned} \quad (18)$$

where $\mathbb{P}(\theta_1 q_1 + \theta_2 q_2 \geq p_b)$ is the demand on the bundle service by customers, i.e., the probability that a customer will decide to subscribe to a bundled service. The objective function of (18) maximizes the total bundling profit of cooperative service providers which is equal to the sale revenue minus the paid data cost. $C3$, $C4$, and $C5$ are required to ensure positive solution values for p_b , n_1 , and n_2 , respectively. The constraints $C1$ and $C2$ are the optimization constraints which depend on the bundle price and service quality. In particular, there are 4 possible demand patterns (defining the optimization constraints $C1$ and $C2$) where (17) holds as demonstrated by the shaded areas in Figure 4. Specifically, $\mathbb{P}(\theta_1 q_1 + \theta_2 q_2 \geq p_b)$ can be found for each case as follows:

$$\text{Case 1: } 1 - 0.5 \frac{p_b}{q_1} \frac{p_b}{q_2}, \quad (19)$$

$$\text{Case 2: } 0.5Mp_b \left(1 - \frac{p_b}{q_1} + 1 - \frac{(p_b - q_2)}{q_1} \right), \quad (20)$$

$$\text{Case 3: } 0.5Mp_b \left(1 - \frac{p_b}{q_2} + 1 - \frac{(p_b - q_1)}{q_2} \right), \quad (21)$$

$$\text{Case 4: } 0.5Mp_b \left(1 - \frac{(p_b - q_1)}{q_2} \right) \left(1 - \frac{(p_b - q_2)}{q_1} \right). \quad (22)$$

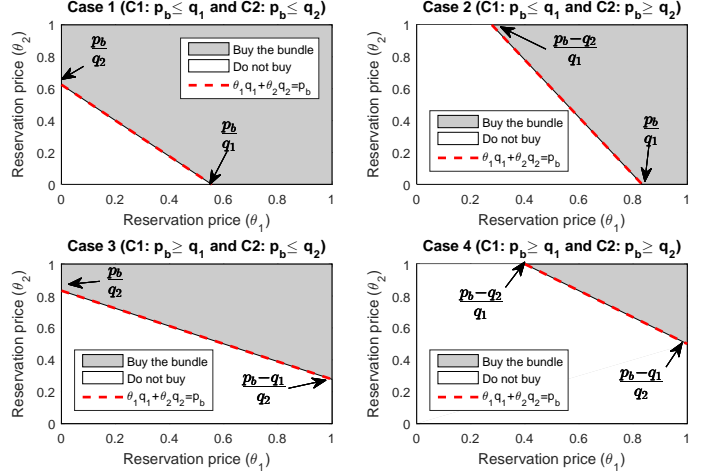


Fig. 4: The four demand cases on a bundled IoT service.

(19)-(22) correspond to Case 1 through Case 4 of Figure 4, respectively. $q_1 = \alpha_{11} - \alpha_{21} \exp(-\alpha_{31}n_1)$ and $q_2 = \alpha_{12} - \alpha_{22} \exp(-\alpha_{32}n_2)$ are the quality of IoT services 1 and 2, respectively.

Proposition 2. $F_K(p_b, n_1, n_2)$ is concave, and hence the closed-form solution of the optimization problem $\underset{p_b, n_1, n_2}{\text{maximize}} F_K(p_b, n_1, n_2)$ is globally optimal.

Proof: We will next prove the concavity of the profit function for Case 1 ($C1: p_b \leq q_1$ and $C2: p_b \leq q_2$) using Sylvester's criterion which provides a necessary and sufficient condition for the concavity of $F_K(\cdot)$ [33]. The other demand cases (Cases 2-4) can be analyzed similarly and are omitted due to the space limit. The Hessian matrix \mathbf{H}_K of $F_K(\cdot)$ for Case 1 is defined as in (23) shown at the top of the next page. $\Delta_{1,1}$, $\Delta_{1,2}$, and $\Delta_{1,3}$ are the three first-order principal minors of \mathbf{H}_K which can be derived as follows:

$$\Delta_{1,1} = -\frac{3Mp_b}{q_1 q_2} \leq 0, \quad (24)$$

$$\Delta_{1,2} = -\frac{M\alpha_{21}^2\alpha_{31}^2p_b^3 \exp(-2\alpha_{31}n_1)}{q_1^3 q_2} - \frac{0.5M\alpha_{21}\alpha_{31}^2p_b^3 \exp(-\alpha_{31}n_1)}{q_1^2 q_2} \leq 0, \quad (25)$$

$$\Delta_{1,3} = -\frac{M\alpha_{22}^2\alpha_{32}^2p_b^3 \exp(-2\alpha_{32}n_2)}{q_1 q_2^3} - \frac{0.5M\alpha_{22}\alpha_{32}^2p_b^3 \exp(-\alpha_{32}n_2)}{q_1 q_2^2} \leq 0. \quad (26)$$

$$\mathbf{H}_{\mathcal{K}} = \begin{bmatrix} \Delta_{1,1} & \frac{1.5M\alpha_{21}\alpha_{31}p_b^2 \exp(-\alpha_{31}n_1)}{q_1^2 q_2} & \frac{1.5M\alpha_{22}\alpha_{32}p_b^2 \exp(-\alpha_{32}n_2)}{q_1 q_2^2} \\ \frac{1.5M\alpha_{21}\alpha_{31}p_b^2 \exp(-\alpha_{31}n_1)}{q_1^2 q_2} & \Delta_{1,2} & \frac{-0.5M\alpha_{21}\alpha_{22}\alpha_{31}\alpha_{32}p_b^3}{q_1^2 q_2^2 \exp(\alpha_{31}n_1) \exp(\alpha_{32}n_2)} \\ \frac{1.5M\alpha_{22}\alpha_{32}p_b^2 \exp(-\alpha_{32}n_2)}{q_1 q_2^2} & \frac{-0.5M\alpha_{21}\alpha_{22}\alpha_{31}\alpha_{32}p_b^3}{q_1^2 q_2^2 \exp(\alpha_{31}n_1) \exp(\alpha_{32}n_2)} & \Delta_{1,3} \end{bmatrix}. \quad (23)$$

There are also three second-order principal minors for $\mathbf{H}_{\mathcal{K}}$ which are obtained as follows:

$$\begin{aligned} \Delta_{2,1} &= \frac{0.75M^2\alpha_{21}\alpha_{31}^2p_b^4 \exp(-\alpha_{31}n_1)(\alpha_{11}+q_1)}{q_1^4 q_2^2} \succcurlyeq \mathbf{0}, \\ \Delta_{2,2} &= \frac{0.75M^2\alpha_{22}\alpha_{32}^2p_b^4 \exp(-\alpha_{32}n_2)(\alpha_{12}+q_2)}{q_1^2 q_2^4} \succcurlyeq \mathbf{0}, \\ \Delta_{2,3} &= \frac{0.25M^2\alpha_{21}\alpha_{22}\alpha_{31}^2\alpha_{32}^2p_b^6 A_1}{q_1^4 q_2^4} \geq 0, \end{aligned} \quad (29)$$

where

$$A_1 = \exp(\alpha_{31}n_1 + \alpha_{32}n_2) \left(\alpha_{11}\alpha_{12} + \alpha_{11}\alpha_{22} \exp(-\alpha_{32}n_2) + \alpha_{12}\alpha_{21} \exp(-\alpha_{31}n_1) \right). \quad (30)$$

Finally, $\mathbf{H}_{\mathcal{K}}$ has one third-order principle minor defined as follows:

$$\Delta_{3,1} = \frac{-0.375M^3\alpha_{21}\alpha_{22}\alpha_{31}^2\alpha_{32}^2p_b^7 A_2}{q_1^5 q_2^5} \leq 0, \quad (31)$$

where

$$A_2 = \exp(-\alpha_{31}n_1 - \alpha_{32}n_2) (\alpha_{11}q_2 + \alpha_{12}q_1). \quad (32)$$

Accordingly, Sylvester's criterion is satisfied. This proves that $\mathbf{H}_{\mathcal{K}}$ is negative semidefinite and $F_{\mathcal{K}}(\cdot)$ is concave. Then, the solution of the optimization problem maximize $F_{\mathcal{K}}(\cdot)$ is globally optimal. \blacksquare

We next provide the closed-form solutions for the profit maximization problem of bundled services by applying the KKT conditions.

B. Optimal Subscription Fee and Requested Data Sizes

1) *Case 1* ($C1: p_b \leq q_1$ and $C2: p_b \leq q_2$): Substituting the customer demand of (19) in (18), the Lagrangian dual function is defined as follows:

$$\mathcal{L}_{\mathcal{K}}(\{p_b, n_1, n_2\}, \lambda_1, \dots, \lambda_5) = -Mp_b \left(1 - \frac{0.5p_b^2}{q_1 q_2} \right) + n_1 c_1 + n_2 c_2 + \lambda_1 (p_b - q_1) + \lambda_2 (p_b - q_2) - \lambda_3 p_b - \lambda_4 n_1 - \lambda_5 n_2, \quad (33)$$

where $\lambda_1, \dots, \lambda_5$ are the Lagrange multipliers of the constraints C_1, \dots, C_5 in (18). The first derivatives of (33) with respect to p_b, n_1 , and n_2 are found as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathcal{K}}(\cdot)}{\partial p_b} &= \frac{Mp_b}{q_1 q_2} - \lambda_2 - \lambda_3 - \lambda_1 \\ &\quad - \frac{Mq_2(-0.5p_b + (\alpha_{11} - \alpha_{21} \exp(-\alpha_{31}n_1)))}{q_1 q_2}, \end{aligned} \quad (34)$$

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathcal{K}}(\cdot)}{\partial n_1} &= c_1 - \lambda_4 + \alpha_{21}\alpha_{31}\lambda_1 \exp(-\alpha_{31}n_1) \\ &\quad - \frac{M\alpha_{21}\alpha_{31}p_b \exp(-\alpha_{31}n_1)}{q_1} \\ &\quad + \frac{M\alpha_{21}\alpha_{31}p_b \exp(-\alpha_{31}n_1)(-0.5p_b + q_2 q_1)}{q_1 q_2}, \end{aligned} \quad (35)$$

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathcal{K}}(\cdot)}{\partial n_2} &= c_2 - \lambda_5 + \alpha_{22}\alpha_{32}\lambda_2 \exp(-\alpha_{32}n_2) \\ &\quad - \frac{M\alpha_{22}\alpha_{32}p_b \exp(-\alpha_{32}n_2)}{q_2} \\ &\quad + \frac{M\alpha_{22}\alpha_{32}p_b \exp(-\alpha_{32}n_2)(-0.5p_b + q_1 q_2)}{q_1 (q_2)^2}. \end{aligned} \quad (36)$$

Applying the KKT necessary and sufficient conditions with no active constraints ($\lambda_1, \dots, \lambda_5 = 0$) by setting the derivatives in (34)-(36) to zero, the closed-form solution can be deduced as follows:

$$n_1^* = \frac{1}{\alpha_{31}} \log \left(\frac{\alpha_{21}}{\alpha_{11}} - \frac{\frac{1}{6}\alpha_{21}A_3}{\alpha_{11}\alpha_{32}c_1} \right), \quad (37)$$

$$n_2^* = \frac{1}{\alpha_{32}} \log \left(\frac{\alpha_{22}}{\alpha_{12}} - \frac{\frac{1}{6}\alpha_{22}A_3}{\alpha_{12}\alpha_{31}c_2} \right), \quad (38)$$

$$p_b^* = -\frac{0.5A_3}{m\alpha_{31}\alpha_{32}}, \quad (39)$$

$$\lambda_1, \dots, \lambda_5 = 0, \quad (40)$$

where

$$A_3 = 3\alpha_{31}c_2 + 3\alpha_{32}c_1 - \left(\frac{8}{2}\alpha_{11}\alpha_{12}M^2\alpha_{31}^2\alpha_{32}^2 + 9\alpha_{31}^2c_2^2 - 18\alpha_{31}\alpha_{32}c_1c_2 + 9\alpha_{32}^2c_1^2 \right)^{0.5}. \quad (41)$$

The global solution is achieved (Proposition 2), and these closed-form expressions satisfy all constraints $C1, \dots, C5$ of the optimization problem in (18).

2) *Case 2* $\left(C1 : p_b \leq q_1 \text{ and } C2 : p_b \geq q_2\right)$: When the customer demand is defined as in (20) and the optimization constraints are $C1 : p_b \leq q_1$ and $C2 : p_b \geq q_2$, the Lagrangian dual problem of (18) is defined as follows:

$$\begin{aligned} \mathcal{L}_{\mathcal{K}}(\{p_b, n_1, n_2\}, \lambda_1, \dots, \lambda_5) = \\ -0.5Mp_b \left(1 - \frac{p_b}{q_1} + 1 - \frac{(p_b - q_2)}{q_1}\right) + n_1c_1 + n_2c_2 \\ + \lambda_1(p_b - q_1) - \lambda_2(p_b - q_2) - \lambda_3p_b - \lambda_4n - \lambda_5n_2, \end{aligned} \quad (42)$$

where $\lambda_1, \dots, \lambda_5$ are the Lagrange multipliers. Taking the derivatives of (42) with respect to p_b, n_1 , and n_2 , the closed-form solution can be found by solving the resulting derivatives as follows:

$$n_1^* = \frac{1}{\alpha_{31}} \log \left(\frac{0.25\alpha_{21}A_4}{\alpha_{32}c_1(\alpha_{11} - \alpha_{12})} \right), \quad (43)$$

$$n_2^* = \frac{1}{\alpha_{32}} \log \left(\frac{0.5}{c_2} \left(M\alpha_{22}\alpha_{32} - \frac{0.5\alpha_{22}A_4}{\alpha_{31}(\alpha_{11} - \alpha_{12})} \right) \right), \quad (44)$$

$$p_b^* = \frac{0.5A_4}{M\alpha_{31}\alpha_{32}} - \frac{2\alpha_{31}c_2 + 2\alpha_{32}c_1 - M\alpha_{12}\alpha_{31}\alpha_{32}}{M\alpha_{31}\alpha_{32}}, \quad (45)$$

$$\lambda_1^* = \frac{0.25A_4}{\alpha_{31}\alpha_{32}(\alpha_{11} - \alpha_{12})} - 0.5M, \quad (46)$$

$$\lambda_2^* = \frac{0.25A_4}{\alpha_{31}\alpha_{32}(\alpha_{11} - \alpha_{12})}, \quad (47)$$

$$\lambda_3^* = \lambda_4^* = \lambda_5^* = 0, \quad (48)$$

where

$$\begin{aligned} A_4 = 2\alpha_{31}c_2 + 2\alpha_{32}c_1 + M\alpha_{11}\alpha_{31}\alpha_{32} - M\alpha_{12}\alpha_{31}\alpha_{32} \\ - 2 \left(0.25M^2\alpha_{11}^2\alpha_{31}^2\alpha_{32}^2 - 0.5M^2\alpha_{11}\alpha_{12}\alpha_{31}^2\alpha_{32}^2 \right. \\ + 0.25M^2\alpha_{12}^2\alpha_{31}^2\alpha_{32}^2 + M\alpha_{11}\alpha_{31}^2\alpha_{32}c_2 - M\alpha_{11}\alpha_{31}\alpha_{32}^2c_1 \\ - M\alpha_{12}\alpha_{31}^2\alpha_{32}c_2 + M\alpha_{12}\alpha_{31}\alpha_{32}^2c_1 + \alpha_{31}^2c_2^2 \\ \left. + 2\alpha_{31}\alpha_{32}c_1c_2 + \alpha_{32}^2c_1^2 \right)^{0.5}. \end{aligned} \quad (49)$$

3) *Case 3* $\left(p_b \geq q_1 \text{ and } p_b \leq q_2\right)$: In this case, the Lagrangian dual function of (18) and (21) can be written as:

$$\begin{aligned} \mathcal{L}_{\mathcal{K}}(\{p_b, n_1, n_2\}, \lambda_1, \dots, \lambda_5) = \\ -0.5Mp_b \left(1 - \frac{p_b - q_1}{q_2} + 1 - \frac{p_b}{q_2}\right) + n_1c_1 + n_2c_2 \\ - \lambda_1(p_b - q_1) + \lambda_2(p_b - q_2) - \lambda_3p_b - \lambda_4n - \lambda_5n_2. \end{aligned} \quad (50)$$

Taking the derivatives of (50) with respect to p_b, n_1 , and n_2 , the closed-form solution can be expressed as follows:

$$n_1^* = \frac{1}{\alpha_{31}} \log \left(\frac{0.25\alpha_{21}A_5}{\alpha_{32}c_1(\alpha_{11} - \alpha_{12})} \right), \quad (51)$$

$$n_2^* = \frac{1}{\alpha_{32}} \log \left(\frac{0.5}{c_2} \left(M\alpha_{22}\alpha_{32} - \frac{0.5\alpha_{22}A_5}{\alpha_{31}(\alpha_{11} - \alpha_{12})} \right) \right), \quad (52)$$

$$p_b^* = \frac{0.5A_5}{M\alpha_{31}\alpha_{32}} - \frac{2\alpha_{31}c_2 + 2\alpha_{32}c_1 - M\alpha_{12}\alpha_{31}\alpha_{32}}{M\alpha_{31}\alpha_{32}}, \quad (53)$$

$$\lambda_1^* = 0.5M - \frac{0.25A_5}{\alpha_{31}\alpha_{32}(\alpha_{11} - \alpha_{12})}, \quad (54)$$

$$\lambda_2^* = -\frac{0.25A_5}{\alpha_{31}\alpha_{32}(\alpha_{11} - \alpha_{12})}, \quad (55)$$

$$\lambda_3^* = \lambda_4^* = \lambda_5^* = 0, \quad (56)$$

where

$$\begin{aligned} A_5 = 2\alpha_{31}c_2 + 2\alpha_{32}c_1 + M\alpha_{11}\alpha_{31}\alpha_{32} - M\alpha_{12}\alpha_{31}\alpha_{32} \\ - 2 \left(0.25M^2\alpha_{11}^2\alpha_{31}^2\alpha_{32}^2 - 0.5M^2\alpha_{11}\alpha_{12}\alpha_{31}^2\alpha_{32}^2 \right. \\ + 0.25M^2\alpha_{12}^2\alpha_{31}^2\alpha_{32}^2 + M\alpha_{11}\alpha_{31}^2\alpha_{32}c_2 - M\alpha_{11}\alpha_{31}\alpha_{32}^2c_1 \\ - M\alpha_{12}\alpha_{31}^2\alpha_{32}c_2 + M\alpha_{12}\alpha_{31}\alpha_{32}^2c_1 + \alpha_{31}^2c_2^2 \\ \left. + 2\alpha_{31}\alpha_{32}c_1c_2 + \alpha_{32}^2c_1^2 \right)^{0.5}. \end{aligned} \quad (57)$$

4) *Case 4* $\left(p_b \geq q_1 \text{ and } p_b \geq q_2\right)$: With (18) and (22), the Lagrangian dual function can be formulated as follows:

$$\begin{aligned} \mathcal{L}_{\mathcal{K}}(\{p_b, n_1, n_2\}, \lambda_1, \dots, \lambda_5) = \\ -0.5Mp_b \left(1 - \frac{(p_b - q_1)}{q_2}\right) \left(1 - \frac{(p_b - q_2)}{q_1}\right) + n_1c_1 + n_2c_2 \\ - \lambda_1(p_b - q_1) - \lambda_2(p_b - q_2) - \lambda_3p_b - \lambda_4n - \lambda_5n_2. \end{aligned} \quad (58)$$

Taking the derivatives of (58) with respect to p_b, n_1 , and n_2 , the closed-form solution can be deduced as follows:

$$n_1^* = \frac{1}{\alpha_{31}} \log \left(\frac{0.25\alpha_{21}\alpha_{31}A_6}{c_1(\alpha_{11}\alpha_{31}\alpha_{32} - \alpha_{12}\alpha_{31}\alpha_{32})} \right), \quad (59)$$

$$n_2^* = \frac{1}{\alpha_{32}} \log \left(\frac{0.5}{c_2} \left(m\alpha_{22}\alpha_{32} - \frac{0.5\alpha_{22}\alpha_{32}A_6}{\alpha_{11}\alpha_{31}\alpha_{32} - \alpha_{12}\alpha_{31}\alpha_{32}} \right) \right), \quad (60)$$

$$p_b^* = \alpha_{11} - \frac{\alpha_{21}c_1(\alpha_{11}\alpha_{31}\alpha_{32} - 1.0\alpha_{12}\alpha_{31}\alpha_{32})}{0.25\alpha_{21}\alpha_{31}A_6}, \quad (61)$$

$$\lambda_1^* = \frac{0.25A_6}{\alpha_{11}\alpha_{31}\alpha_{32} - \alpha_{12}\alpha_{31}\alpha_{32}} - 0.5m, \quad (62)$$

$$\lambda_2^* = -\frac{0.25A_6}{\alpha_{11}\alpha_{31}\alpha_{32} - \alpha_{12}\alpha_{31}\alpha_{32}}, \quad (63)$$

$$\lambda_3^* = \lambda_4^* = \lambda_5^* = 0, \quad (64)$$

where

$$\begin{aligned}
A_6 = & 2\alpha_{31}c_2 + 2\alpha_{32}c_1 + M\alpha_{11}\alpha_{31}\alpha_{32} - M\alpha_{12}\alpha_{31}\alpha_{32} \\
& - \left(M^2\alpha_{11}^2\alpha_{31}^2\alpha_{32}^2 - 2M^2\alpha_{11}\alpha_{12}\alpha_{31}^2\alpha_{32}^2 + M^2\alpha_{12}^2\alpha_{31}^2\alpha_{32}^2 \right. \\
& + 4M\alpha_{11}\alpha_{31}^2\alpha_{32}c_2 - 4M\alpha_{11}\alpha_{31}\alpha_{32}^2c_1 - 4M\alpha_{12}\alpha_{31}^2\alpha_{32}c_2 \\
& \left. + 4M\alpha_{12}\alpha_{31}\alpha_{32}^2c_1 + 4\alpha_{31}^2c_2^2 + 8\alpha_{31}\alpha_{32}c_1c_2 + 4\alpha_{32}^2c_1^2 \right)^{0.5}.
\end{aligned} \tag{65}$$

The cooperative providers should run the optimization of the four cases derived in this section. The case with the maximum resulting profit should be used in the bundled service.

VI. PROFIT SHARING AMONG IOT PROVIDERS

After forming a coalition \mathcal{K} to sell IoT services as a bundle, the cooperative providers share the resulting profit. This section presents a profit sharing model using the core solution and Shapley value from cooperative game theory [34] to define the payoff allocations for the cooperative providers in \mathcal{K} .

A. The Core Solution

Let φ_k indicate the profit share of service provider $k \in \mathcal{K}$. The core solution set is calculated as follows [34]:

$$\mathcal{C} = \left\{ \varphi \mid \underbrace{\sum_{k \in \mathcal{K}} \varphi_k = F_{\mathcal{K}}^*}_{\text{group rationality}} \wedge \underbrace{\sum_{k \in \mathcal{S}} \varphi_k \geq F_{\mathcal{S}}^*}_{\text{individual rationality}}, \mathcal{S} \subseteq \mathcal{K} \right\} \tag{66}$$

where $\varphi \in \mathbb{R}^2$ is a vector of profit allocations φ_k , $F_{\mathcal{K}}^*$ is the profit of bundling, and $F_{\mathcal{S}}^*$ is the profit of separate selling when $|\mathcal{S}| = 1$. \mathcal{C} includes a set of profit allocations which guarantee no service provider will reject the payoff allocation, i.e., no incentive of leaving the coalition to sell services separately.

The core solution can be empty, containing a large number of possible solutions, or unfair to a service provider based on the individual contributions to the bundle formulation. Therefore, we next present the Shapley solution which provides a single and fair solution of the profit sharing problem.

B. The Shapley Solution

The Shapley solution provides a fair allocation of the bundling profit among the service providers forming a bundling coalition \mathcal{K} . For each cooperative provider $k \in \mathcal{K}$, the Shapley value η assigns a payoff η_k found as [34]:

$$\eta_k = \sum_{\mathcal{S} \subseteq \mathcal{K} \setminus \{k\}} \underbrace{\frac{|\mathcal{S}|! (|\mathcal{K}| - |\mathcal{S}| - 1)!}{|\mathcal{K}|!}}_{\text{probability of ordering}} \underbrace{\left(F_{\mathcal{S} \cup \{k\}}^* - F_{\mathcal{S}}^* \right)}_{\text{contribution}}. \tag{67}$$

The first term defines the probability of ordering providers to form the bundle coalition. The second term defines the marginal contribution of each provider $k \in \mathcal{K}$ to the bundle. The Shapley solution is efficient such that $\sum_{k \in \mathcal{K}} \eta_k = F_{\mathcal{K}}^*$.

VII. NUMERICAL RESULTS AND DISCUSSION

This section presents extensive numerical experiments to evaluate the performance of the proposed market model and pricing schemes for selling IoT services separately and as a bundle.

A. Datasets, machine learning, and Market Setup

We run experiments on three IoT services trained on real-world datasets. Today there are a few standard machine learning models that are widely used across tasks in IoT. To build the IoT service, we use three popular machine learning algorithms, namely, deep learning [27], support vector machines [35], and multinomial naive Bayes [36].

- *Service 1 (text categorization)*: IoT requires context awareness and consistent understanding of the user data [37]. We build Service 1 that predicts the user context from text data. Service 1 is trained on the 20 newsgroups dataset [38]. The dataset includes 18,828 data samples classified into 20 topics, e.g., sports, computer hardware, politics, etc. We use 13,180 samples for training the machine learning models and 5,648 samples for testing and accuracy prediction. A data unit contains 188 samples, i.e., one percent of the full dataset samples. We assume that the data vendor offers this dataset with a data unit cost of $c_1 = 0.1^{10}$.
- *Service 2 (sentiment analysis)*: IoT wearables provide a robust tool for determining the sentiment aspects, e.g., opinion, feedback, and emotion, of users [39]. Sentiment predicting IoT has many applications in smart homes, recommender systems, and content-based filtering. We build a data service to predict the sentiment (either positive or negative polarities) by training the machine learning models on the Sentiment140 dataset [40]. The dataset contains 629,146 tweet samples for model training and 419,431 samples for testing. Each data unit includes 10,482 samples. We assume that a data unit has a cost value of $c_2 = 0.05$.
- *Service 3 (Optical character recognition)*: IoT devices generally contain high-resolution cameras for capturing image and video data. The recent advances in machine learning enable extracting useful information from visual data. Service 3 is an IoT service that extracts text from handwritten images, e.g., this is needed for monitoring the expiration date of food in a smart fridge [41]. We train the machine learning models using the MNIST dataset [42]. The dataset contains 50,000 training image samples and 10,000 testing images. A data unit includes 600 samples.

Text classification (Services 1 and 2) typically requires transforming the text into numerical features suitable for different classifiers. In our experiments, we use a popular feature set known as term frequency-inverse document frequency (TF-IDF) [43] which counts the number of times that a word appears in a text. Unless otherwise stated, we assume a base of $M = 50$ customers.

¹⁰It is in monetary unit which is currency independent.

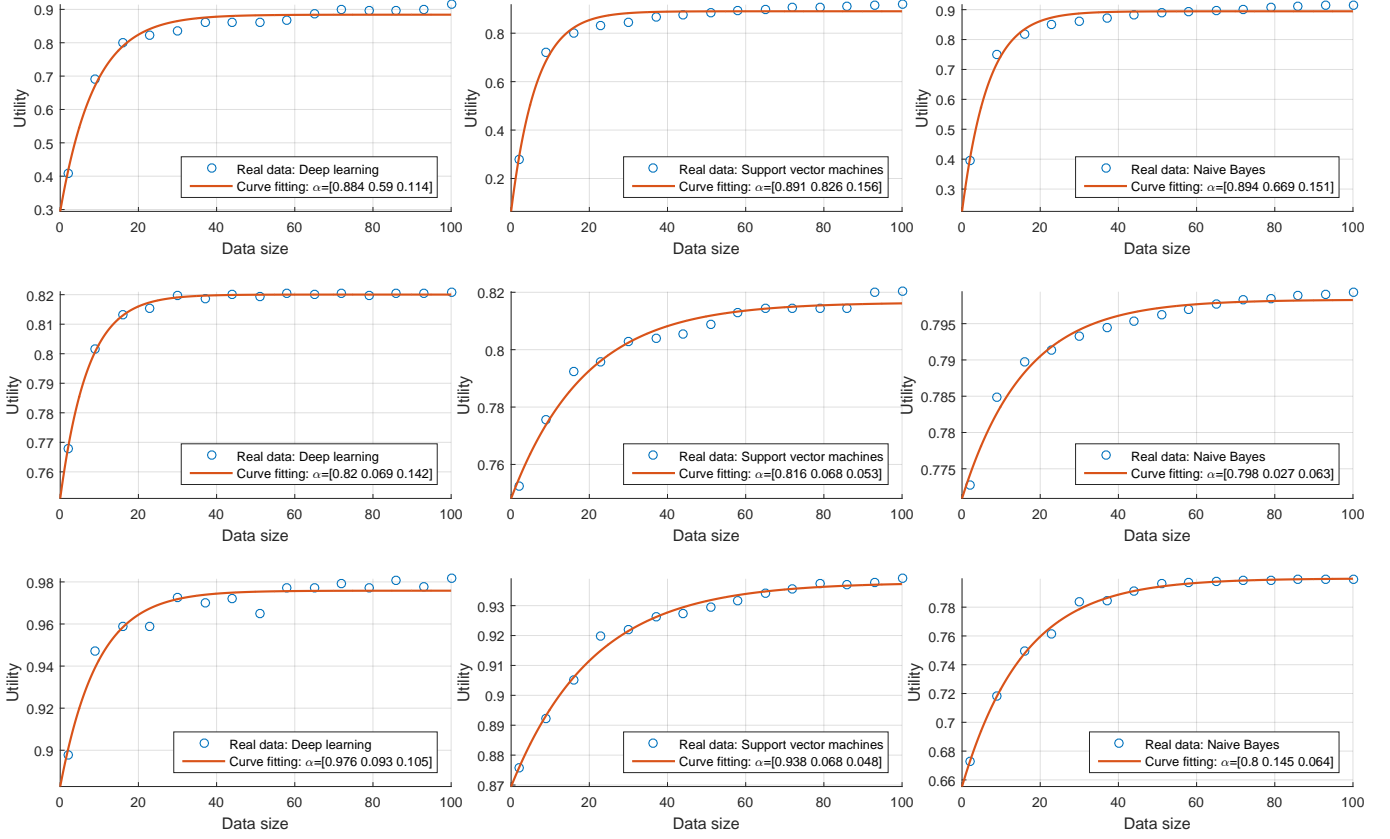


Fig. 5: Estimating the utility of data $q(n; \alpha)$ of three popular classifiers. The first row is for Service 1, the second row is for Service 2, and the third row is for Service 3, respectively.

Figure 6 shows the willingness-to-pay responses to our real-world customer surveys conducted using Google Surveys¹¹ for estimating the willingness-to-pay θ distribution in IoT. In each survey, we asked 100 users open-ended questions that elicit a direct expression of value. For example, to find θ of Service 2, we used the question of “how much would you be willing to pay for a wearable device that plays music and suggests activities (e.g., sports and readings) based on your sentiment and mood?”. Figure 6 shows that θ follows a uniform distribution. For model selection, we use the Akaike information criterion (AIC) [44]. We tested many relevant data distributions including the negative binomial, poisson, and geometric distributions. The uniform distribution has the minimum AIC values of 397.5 and 274.5 for Services 1 and 2, respectively.

B. Service Quality

Figure 5 shows our extensive experiments on finding the quality of Services 1-3. The recognition accuracy increases as the requested data size increases and vice versa. When the requested data size is increased, the increase in the accuracy diminishes. Besides, we note that the utility function defined in (3) fits the series of real data points well. We also observe

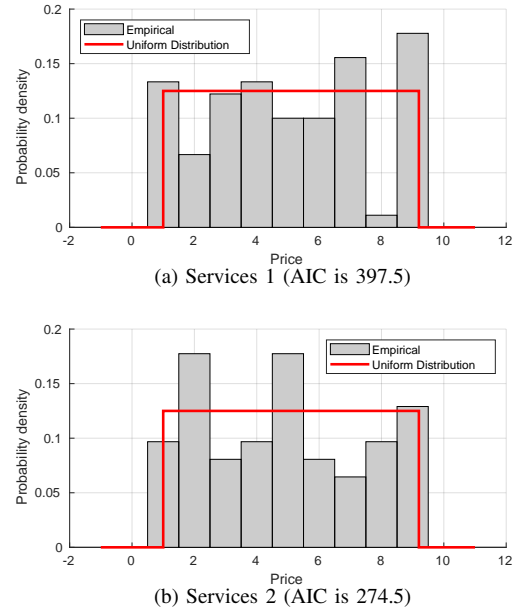


Fig. 6: Modeling the willingness-to-pay distribution based on real-world (empirical) customer surveys. θ of Services 1 and 2 follows a uniform distribution.

¹¹<https://surveys.google.com>

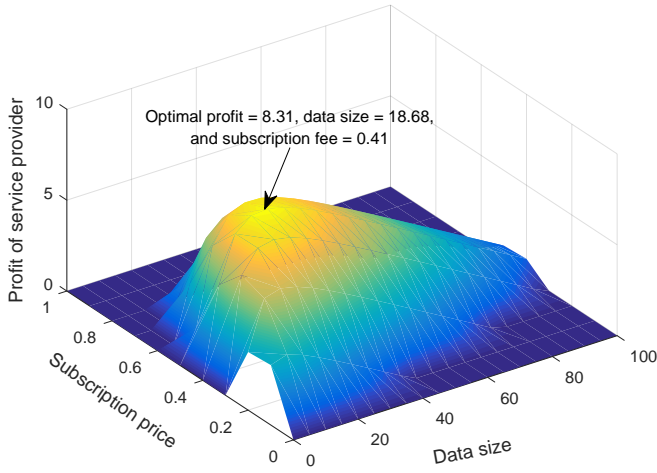


Fig. 7: Profit $F(p_s, n)$ of a monopolist service provider under varied data sizes and subscription fees.

a similar diminishing increase of accuracy for other machine learning algorithms such as k-nearest neighbors, logistic regression, random forests, and passive aggressive classifiers. However, due to the space limit, we omit them from this paper.

Based on these results, for the rest of our experiments we use $q_1 = 0.884 - 0.59 \exp(-0.114n_1)$ and $q_2 = 0.82 - 0.069 \exp(-0.142n_2)$ for service qualities which correspond to using deep learning on the 20 newsgroups (Service 1) and Sentiment140 (Service 2) datasets.

C. Separate Selling of Services

In this section, we study the separate sales of Service 1 under the model and pricing schemes of Section IV. Assuming a deep learning implementation of Service 1, the service quality parameters are $\alpha_1 = 0.884$, $\alpha_2 = 0.59$, and $\alpha_3 = 0.114$ as specified in Figure 5.

1) *Profit Maximization:* In Figure 7, we analyze the profit $F(p_s, n)$ of Service 1 under varied requested data sizes n and subscription fees p_s . When the subscription fee is high, fewer customers will be interested in buying the service which will accordingly decrease the resulting profit. A low subscription fee has a similar negative effect on the profit even though more customers will decide to buy the service. Likewise, a poor spending on buying data from the vendor results in poor profit for the provider as fewer customers will buy the service due to its low accuracy. When the requested data size is high, the service accuracy will increase, but the profit will be negatively affected due to the high data cost. Recall that the profit is the revenue value that remains after accounting for the data cost as in (5).

Using (16) and (15) to determine the optimal service setting, we find that the optimal data size $n^* = 18.68$ and subscription fee $p_s^* = 0.41$ resulting in an optimal profit of $F(n^*, p_s^*) = 8.31$.

As data is the main commodity in IoT services, we next study the impact of data price (Section VII-C2) and data quality (Section VII-C3) in the economics of a monopolist service provider.

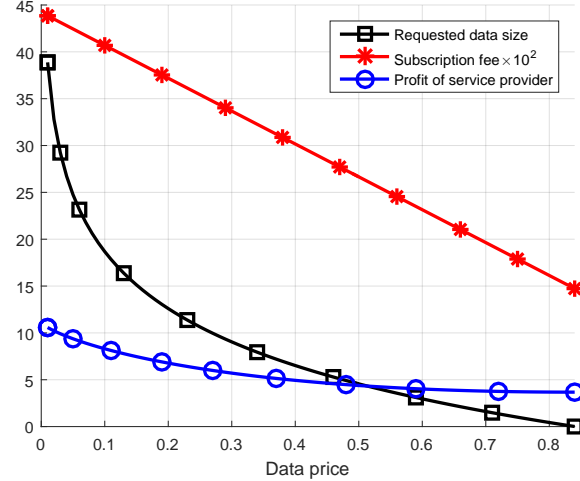


Fig. 8: Optimal profit $F(p_s^*, n^*)$ of the service provider, subscription fee p_s^* , and requested data size n^* under varied data prices.

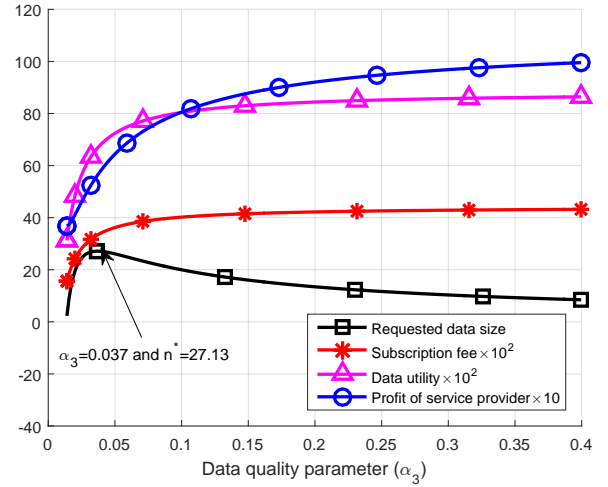


Fig. 9: Optimal profit $F(p_s^*, n^*)$ of the service provider, subscription fee p_s^* , requested data size n^* , and data utility $q(n^*, \alpha)$ under varied data quality parameter α_3 .

2) *The Impact of Data Cost:* Figure 8 shows the impact of the data cost c on the optimal solutions, i.e., optimal requested data size n^* and subscription fee p_s^* , and the resulting profit $F(n^*, p_s^*)$ from the separate sales of Service 1. Firstly, we observe that the service provider will make a lower profit out of the service if the data price is increased. Secondly, when the data price is high, the service provider will try to decrease the cost by requesting less data from the data vendor. Then, as requesting less data decreases the service quality q , the service provider should also decrease the subscription fee in order to attract more customer subscriptions. This is intuitive as the customers infer both the subscription fee and the service quality when making their subscription decision as defined in (5). Increasing the data price beyond $c = 0.84$ produces infeasible solutions, i.e., negative profit, as the constraint $M\alpha_2\alpha_3 \geq 4c$ will not be satisfied.

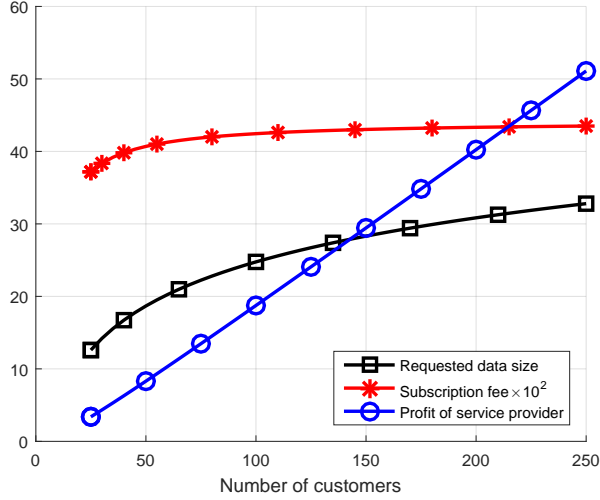


Fig. 10: Optimal profit $F(p_s^*, n^*)$ of the service provider, subscription fee p_s^* , and requested data size n^* under varied customer bases M .

3) *The Impact of Data Quality:* We consider the impact of buying data of different qualities q by changing the data quality parameter α_3 . Specifically, higher values of α_3 result in higher data qualities q as specified in (3). We observe two important results from Figure 9. Firstly, the service provider should always look for high quality data as it helps them to maximize their profit as less data is required for service fitting. If the provider buys more quality data up to a level ($n^* = 27.13$ at $\alpha_3 = 0.037$), then less data will be requested as enhancing the service quality beyond this level will not be essential for most customers. This is because the customers infer both the service quality and the subscription fee. Secondly, when the data quality is high, less data is required to achieve a satisfactory quality. The low spending in buying data maximizes the provider's profit.

4) *The Impact of Customer Base Size:* Figure 10 shows the optimal solutions, i.e., optimal requested data size n^* and subscription fee p_s^* , and the generated profit $F(n^*, p_s^*)$ under varied sizes of the customer base M . We observe that the service provider will increase his investment in buying more data as the customer size increases (more customer demand). Likewise, the subscription fee p_s^* of the service will be increased as a result of the increased demand. This increase in demand positively impacts the resulting profit of the service provider.

D. Service Bundling

In the following, we consider the marketing strategy of grouping Services 1 and 2 into one bundled service forming a coalition \mathcal{K} . Deep learning is used for machine learning in both services resulting in quality functions as shown in Figure 5. Specifically, the service quality parameters are $\alpha_{11} = 0.884$, $\alpha_{21} = 0.59$, and $\alpha_{31} = 0.114$ for Services 1, while the quality parameters are $\alpha_{12} = 0.82$, $\alpha_{22} = 0.069$, and $\alpha_{32} = 0.142$

for Service 2. We use the analysis of Section V for service bundling and pricing.

1) *Profit Maximization:* The profits $F_{\mathcal{K}}(p_b, n_1, n_2)$ of the bundled service is plotted in Figure 11 which correspond to the four demand cases of (19)-(22). The maximum profit is achieved by using Case 1 ($C1 : p_b \leq q_1$ and $C2 : p_b \leq q_2$) with the closed-form solutions defined in (37)-(40). The optimal solutions are $n_1^* = 19.29$, $n_2^* = 7.01$, $p_b^* = 0.658$ which generate an optimal profit of $F_{\mathcal{K}}(p_b^*, n_1^*, n_2^*) = 19.67$.

Figure 12 illustrates the market equilibrium of selling Services 1 and 2 separately or as one bundle under coalition \mathcal{K} . Using (16), (15), and (5) to analyze the separate sales, we find that the optimal subscription fee and profit values of Services 1 and 2 are $p_{s1}^* = 0.41$, $F_1(n_1^*, p_{s1}^*) = 8.31$, $p_{s2}^* = 0.39$, and $F_2(n_2^*, p_{s2}^*) = 9.58$, respectively.

- For separate sales (Figure 12a), the optimal market equilibrium is shown as the vertical $\theta_1 q_1 = p_{s1}^* = 0.41$ and horizontal $\theta_2 q_2 = p_{s2}^* = 0.39$ lines for Services 1 and 2, respectively. There are four decision areas centered by the point $\mathcal{H} \equiv \left(\frac{p_{s1}^*}{q_1}, \frac{p_{s2}^*}{q_2} \right)$. Firstly, the customers with reservation prices that lie to the northeast of \mathcal{H} will subscribe to both services. Secondly, the customers with reservation prices that lie to the southeast of \mathcal{H} will subscribe to Service 1 only, Thirdly, the customers with reservation prices that lie to the northwest of \mathcal{H} will subscribe to Service 2 only, Finally, the customers with reservation prices that lie to the southwest of \mathcal{H} will not subscribe to any service as their evaluations of both service are below the market equilibrium values.
- Figure 12b shows that customers with reservation prices that lie above the line $\theta_1 q_1 + \theta_2 q_2 = p_b^*$ will decide to subscribe to the bundled service. On the other hand, the customers with reservation prices that lie below the line $\theta_1 q_1 + \theta_2 q_2 = p_b^*$ will not subscribe to the bundle.

We can also observe that the customers will be interested in buying the two services as one bundled service at the discounted subscription price $p_b^* < p_{s1}^* + p_{s2}^*$. From the provider's perspectives, the bundled service helps in maximizing their profits $F_{\mathcal{K}}(p_b^*, n_1^*, n_2^*) > F_1(n_1^*, p_{s1}^*) + F_2(n_2^*, p_{s2}^*)$. This profit is shared among the two service providers as will be discussed later in Section VII-E.

2) *The Impact of Data Cost:* While fixing the data price of Service 2 ($c_2 = 0.05$), we have varied the data price of Service 1 and observed the market parameters of the bundled service as in Figure 13. There are several important observations that can be highlighted. The optimal requested data sizes n_1^* and n_2^* , bundle subscription fee p_b^* , and bundling profit $F_{\mathcal{K}}(p_b^*, n_1^*, n_2^*)$ decrease as the data price c_1 increases. Specifically, when the data price c_1 is high, the cooperative providers will try to minimize the total cost by buying less data for Service 1 n_1^* . Likewise, this will decrease the ability of buying data for Service 2 n_2^* which will be slightly decreased. Finally, to maintain the market equilibrium, the subscription fee p_b^* of the service bundle should be also decreased for high values of data price c_1 . This inverse correlation can be deduced from (17) where the customers consider the qualities of Services 1 and 2 as well as the bundle subscription fee

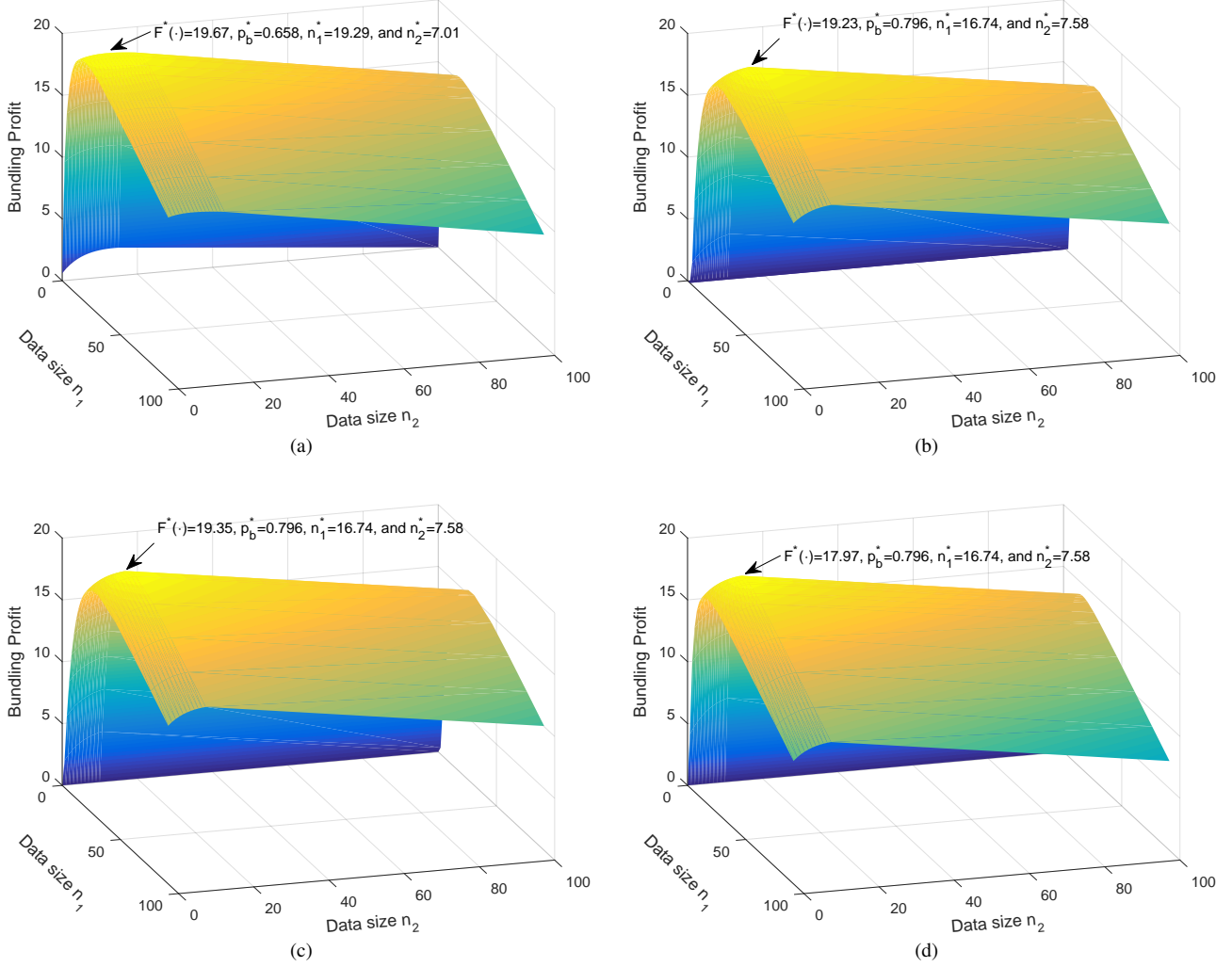


Fig. 11: Profit from the bundled service under varied data sizes (n_1 and n_2). (a)-(d) are for Cases 1-4, respectively.

when making their subscription decisions.

3) *The Impact of Customer Base Size*: Figure 14 shows that the optimal requested data sizes n_1^* and n_2^* , subscription fee p_b^* , and bundling profit $F_K(p_b^*, n_1^*, n_2^*)$ are proportional to the number of customers M . In particular, an increased number of customers M results in higher values of n_1^* , n_2^* , p_b^* , and $F_K(p_b^*, n_1^*, n_2^*)$. As any other demand relationship in economics, the subscription fee and profit increase for high demands by customers. Similarly, the service providers will buy more data to provide better quality services as more customers would be interested in buying the service.

E. Profit Sharing of the Service Bundle

We next analyze the profit sharing among the two service providers forming the bundled service.

1) *Payoff Allocation*: Recall from Section VII-D1 that the bundling profit is $F_K(p_b^*, n_1^*, n_2^*) = 19.67$, while the profits of separate selling are $F_1(n_1^*, p_{s1}^*) = 8.31$ and $F_2(n_2^*, p_{s2}^*) = 9.58$ of Services 1 and 2, respectively. Figure 15 shows four important solution concepts. The first solution (gray shaded

area) gives the feasible payoffs. The second solution gives “efficient payoffs” which satisfy the constraint $\varphi_1 + \varphi_2 = F_K(p_b^*, n_1^*, n_2^*)$. However, this solution does not guarantee that the individual payoff of each provider resulting from the bundling strategy is higher than that of the separate service selling. Therefore, a provider may reject his payoff and decide to sell his service separately. The third solution is the core which is found as in (66). This solution guarantees that the individual payoffs of providers are higher with bundling compared to the separate selling of services. The final solution is the Shapley value as in (67) which gives fair payoffs for the cooperative providers in K based on their individual contribution to the service bundle.

2) *The Impact of Individual Contributions on Payoffs*: In Figure 16, we further analyze the fair allocation of payoffs with the Shapley concept. In particular, we consider the case when the data price of Service 1 is varied while fixing the data price of Service 2 ($c_2 = 0.05$). The payoff allocations by the Shapley value are always inside the core solution area (the gray shaded area). Firstly, we compare the profit of selling Service 1 and 2 separately or as a bundled service. The payoff

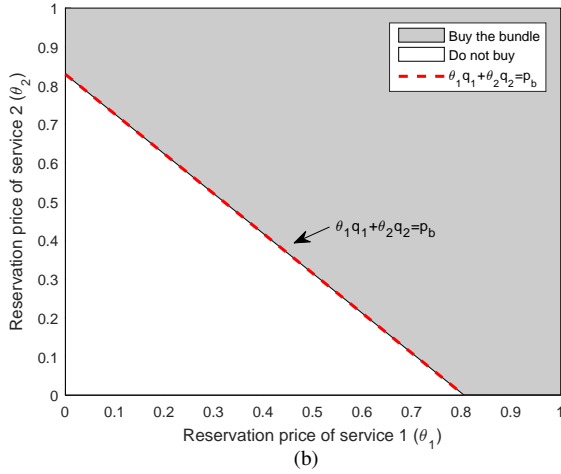
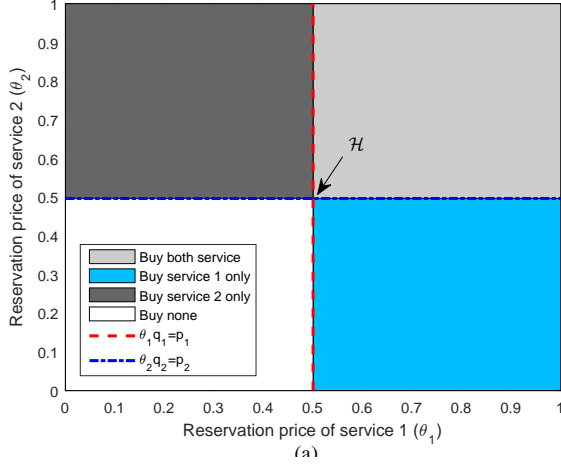


Fig. 12: Market equilibrium under different customer reservations. (a) Separate selling of the two services with subscription fees of $p_{s1}^* = 0.41$ and $p_{s2}^* = 0.39$. (b) Selling the two services as a bundle with bundle subscription fee of $p_b = 0.659$.

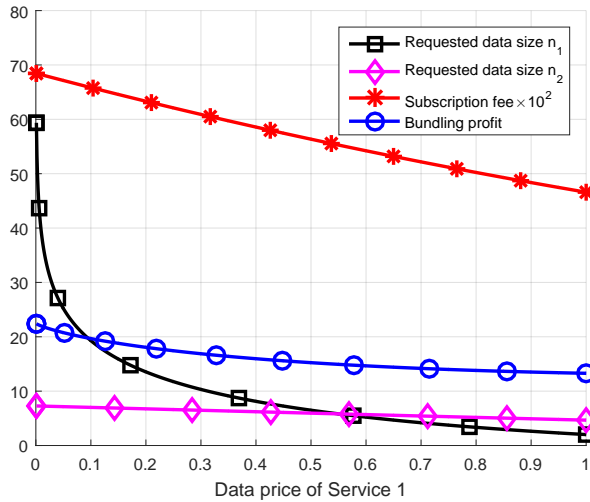


Fig. 13: Optimal requested data sizes n_1^* and n_2^* , bundle subscription fee p_b^* , and bundling profit $F_K^*(\cdot)$ of service providers under varied data prices of Service 1 c_1 .

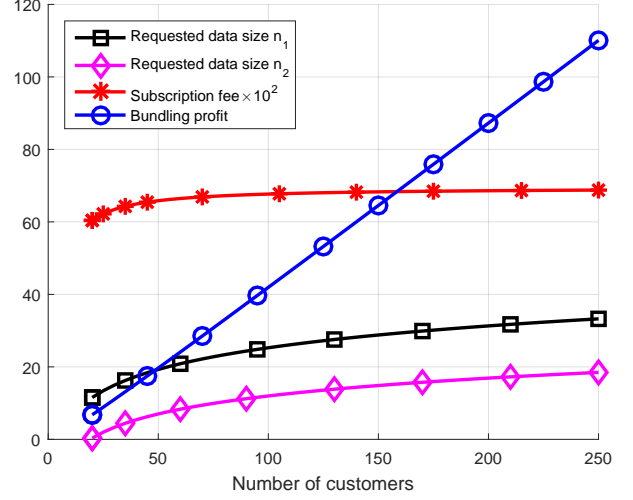


Fig. 14: Optimal requested data sizes n_1^* and n_2^* , bundle subscription fee p_b^* , and bundling profit $F_K^*(\cdot)$ of service providers under varied sizes of the customer base M .

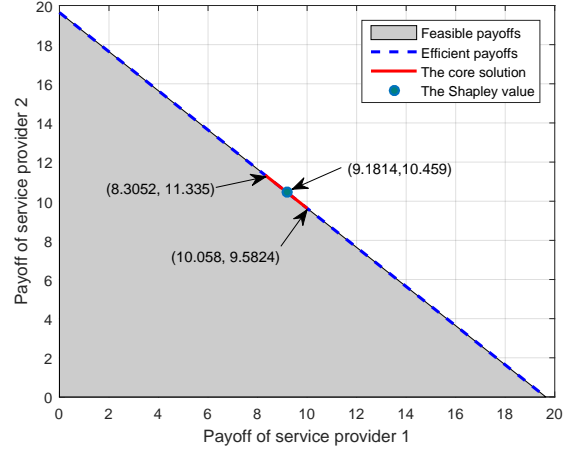


Fig. 15: Profit sharing of the bundled service.

allocations to each service by the Shapley value are higher than those of the separate sales. In the separate selling, changing the price of one service does not change the profit of the other service. Secondly, the payoff allocations in the bundled service reflect the contributions of an individual service to the profit of the bundled service. When the data price of a service is high, the payoff of that particular service will be fairly decreased due to the high cost introduced in building the bundle. This fair allocation of payoff in the Shapley concept provides an incentive for cooperative service providers to search for the best deals and vendors of data. Thirdly, it can be noted that the high data cost of Service 1 reduces the incentive of Service 2 in the bundle formation as the residual margin between the profit of separate and bundled sales diminishes and becomes zero at $c_1 = 1$.

Finally, Figure 17 shows that higher quality of one service increases the profit of the other cooperative service within

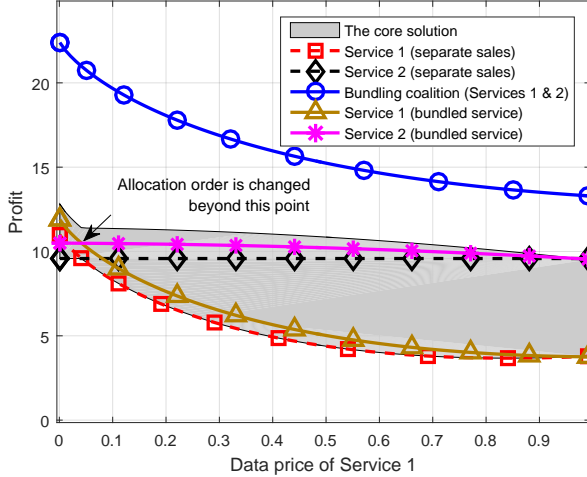


Fig. 16: Profit of Services 1 and 2 under varied data prices of Service 1 c_1 .

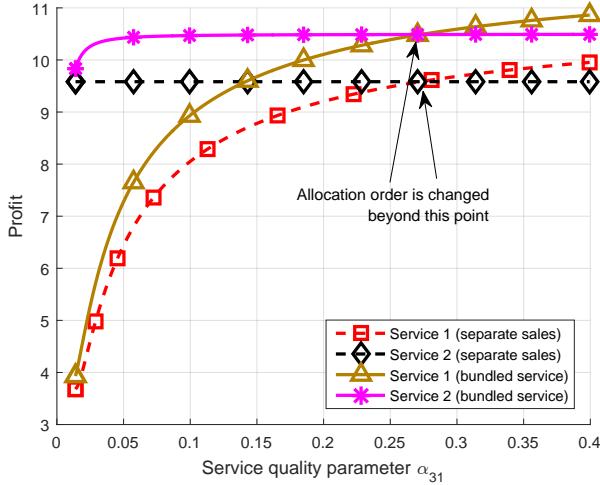


Fig. 17: Profit of Services 1 and 2 under varied quality of Service 1 α_{31} .

the bundle, and vice versa. When the quality of Service 1 is increased (high values of α_{31}) while fixing the quality of Service 2, the payoff allocations to Services 1 and 2 increase. However, it is observed that the profit increase of Service 1 is greater than that of Service 2. In the separate selling case, Service 2 will not benefit from enhancing the quality of Service 1. Therefore, there is a profit incentive for service providers to look for high quality services to cooperate with when forming bundled services.

VIII. CONCLUSION

In this paper, we have developed IoT market models and optimal pricing schemes for the standalone and bundled sales of machine learning-based IoT services. The proposed optimizations maximize the profit of service providers by deciding the optimal data sizes that should be bought from data vendors and the optimal subscription fees at which the services should be offered to customers. Bundling IoT services has been shown

as an effective marketing strategy for maximizing the profits of providers. We have observed that service providers have a profit incentive in searching for the best quality services to form a bundling coalition. Finally, we have presented a profit sharing models for allocating the profit payoffs among cooperative service providers.

In this paper, we have assumed that the service providers are economically rational, i.e., they take decisions that maximize their revenues. Future work can improve our optimizations by exploring the possible competition among service providers. Moreover, our profit maximization models can be extended to support complementary and substitute services.

ACKNOWLEDGMENT

This work was supported in part by the Australian Research Council (ARC) under grant DE200100863. It was also partially supported by Singapore NRF National Satellite of Excellence, Design Science and Technology for Secure Critical Infrastructure NSoE DeST-SCI2019-0007, A*STAR-NTU-SUTD Joint Research Grant Call on Artificial Intelligence for the Future of Manufacturing RGANS1906, WASP/NTU M4082187 (4080), Singapore MOE Tier 1 2017-T1-002-007 RG122/17, MOE Tier 2 MOE2014-T2-2-015 ARC4/15, Singapore NRF2015-NRF-ISF001-2277, Singapore EMA Energy Resilience NRF2017EWT-EP003-041, US MURI AFOSR MURI 18RT0073, NSF EARS-1839818, CNS1717454, CNS-1731424, CNS-1702850, and CNS-1646607.

REFERENCES

- [1] S. Müller, O. Atan, M. van der Schaar, and A. Klein, "Context-aware proactive content caching with service differentiation in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1024–1036, 2017.
- [2] S. O. Somuyiwa, A. György, and D. Gündüz, "A reinforcement-learning approach to proactive caching in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, pp. 1331–1344, June 2018.
- [3] J. Xu, L. Song, J. Y. Xu, G. J. Pottie, and M. Van Der Schaar, "Personalized active learning for activity classification using wireless wearable sensors," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, pp. 865–876, Aug 2016.
- [4] M. Abu Alsheikh, D. Niyato, S. Lin, H.-P. Tan, and Z. Han, "Mobile big data analytics using deep learning and Apache Spark," *IEEE Network*, vol. 30, pp. 22–29, May 2016.
- [5] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in ofdm systems using sparse bayesian learning," *IEEE Transactions on Signal Processing*, vol. 62, pp. 3591–3603, July 2014.
- [6] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, "A mobile offloading game against smart attacks," *IEEE Access*, vol. 4, pp. 2281–2291, June 2016.
- [7] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 10037–10047, December 2016.
- [8] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [9] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," *MIT Sloan Management Review*, vol. 52, p. 21, December 2011.
- [10] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "A survey of smart data pricing: Past proposals, current plans, and future trends," *ACM Computing Surveys*, vol. 46, p. 15, November 2013.
- [11] T. Seregina, O. Brun, R. El-Azouzi, and B. J. Prabhu, "On the design of a reward-based incentive mechanism for delay tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 16, pp. 453–465, February 2016.

- [12] M. Zhang, L. Gao, J. Huang, and M. L. Honig, "Hybrid pricing for mobile collaborative internet access," *IEEE/ACM Transactions on Networking*, June 2019.
- [13] M. Khodak, L. Zheng, A. S. Lan, C. Joe-Wong, and M. Chiang, "Learning cloud dynamics to optimize spot instance bidding strategies," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 2762–2770, IEEE, 2018.
- [14] C. Jiang, L. Gao, L. Duan, and J. Huang, "Scalable mobile crowdsensing via peer-to-peer data sharing," *IEEE Transactions on Mobile Computing*, vol. 17, pp. 898–912, April 2017.
- [15] S. Xu, X. Chen, X. Pi, C. Joe-Wong, P. Zhang, and H. Y. Noh, "iLOCuS: Incentivizing vehicle mobility to optimize sensing distribution in crowd sensing," *IEEE Transactions on Mobile Computing*, May 2019.
- [16] S. Jain and P. Kannan, "Pricing of information products on online servers: Issues, models, and analysis," *Management Science*, vol. 48, pp. 1123–1142, September 2002.
- [17] A. Sundararajan, "Nonlinear pricing of information goods," *Management Science*, vol. 50, pp. 1660–1673, December 2004.
- [18] S. Balasubramanian, S. Bhattacharya, and V. V. Krishnan, "Pricing information goods: A strategic analysis of the selling and pay-per-use mechanisms," *Marketing Science*, vol. 34, pp. 218–234, March 2015.
- [19] Y. Bakos and E. Brynjolfsson, "Bundling information goods: Pricing, profits, and efficiency," *Management science*, vol. 45, pp. 1613–1630, December 1999.
- [20] M. Balazinska, B. Howe, and D. Suciu, "Data markets in the cloud: An opportunity for the database community," *Proceedings of the VLDB Endowment*, vol. 4, no. 12, pp. 1482–1485, 2011.
- [21] P. Koutris, P. Upadhyaya, M. Balazinska, B. Howe, and D. Suciu, "Query-based data pricing," *Journal of the ACM*, vol. 62, p. 43, November 2015.
- [22] C. Li, D. Y. Li, G. Miklau, and D. Suciu, "A theory of pricing private data," *ACM Transactions on Database Systems*, vol. 39, p. 34, December 2014.
- [23] B.-R. Lin and D. Kifer, "On arbitrage-free pricing for general data queries," *Proceedings of the VLDB Endowment*, vol. 7, no. 9, pp. 757–768, 2014.
- [24] D. Yang, G. Xue, X. Fang, and J. Tang, "Incentive mechanisms for crowdsensing: Crowdsourcing with smartphones," *IEEE/ACM Transactions on Networking*, vol. 24, pp. 1732–1744, June 2016.
- [25] A. Singla and A. Krause, "Truthful incentives in crowdsourcing tasks using regret minimization mechanisms," in *Proceedings of the 22nd International Conference on World Wide Web*, pp. 1167–1178, ACM, 2013.
- [26] T. Luo, S. K. Das, H. P. Tan, and L. Xia, "Incentive mechanism design for crowdsourcing: An all-pay auction approach," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 7, no. 3, p. 35, 2016.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning." <http://www.deeplearningbook.org>, 2016. MIT Press.
- [28] A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE Intelligent Systems*, vol. 24, pp. 8–12, March 2009.
- [29] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, pp. 78–87, October 2012.
- [30] M. Abu Alsheikh, S. Lin, D. Niyato, and H.-P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 1996–2018, Fourthquarter 2014.
- [31] T. Strutz, *Data fitting and uncertainty: A practical introduction to weighted least squares and beyond*. Vieweg and Teubner, 2010.
- [32] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," in *NIPS Workshop on Private Multi-Party Machine Learning*, 2016.
- [33] E. K. Chong and S. H. Zak, *An introduction to optimization*, vol. 76. John Wiley & Sons, 2013.
- [34] R. B. Myerson, *Game theory*. Harvard university press, 2013.
- [35] I. Steinwart and A. Christmann, *Support vector machines*. Springer Science & Business Media, 2008.
- [36] A. M. Kibriya, E. Frank, B. Pfahringer, and G. Holmes, "Multinomial naive Bayes for text categorization revisited," in *Advances in Artificial Intelligence*, pp. 488–499, Springer, 2004.
- [37] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE communications surveys & tutorials*, vol. 16, pp. 414–454, First Quarter 2014.
- [38] J. Rennie, "The 20 newsgroups dataset." <http://qwone.com/~jason/20Newsgroups>, 2008. accessed 19-February-2016.
- [39] T. W. AlHanai and M. M. Ghassemi, "Predicting latent narrative mood using audio and physiologic data," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, 2017.
- [40] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision." <http://www.sentiment140.com>, 2009.
- [41] S. Luo, H. Xia, Y. Gao, J. S. Jin, and R. Athauda, "Smart fridges with multimedia capability for better nutrition and health," in *Proceedings of the International Symposium on Ubiquitous Multimedia Computing*, pp. 39–44, IEEE, 2008.
- [42] Y. LeCun and C. Cortes, "Mnist handwritten digit database." <http://yann.lecun.com/exdb/mnist>, 2010. Online; accessed 23-March-2016.
- [43] A. Rajaraman, J. D. Ullman, J. D. Ullman, and J. D. Ullman, *Mining of massive datasets*, vol. 1. Cambridge University Press, 2012.
- [44] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected papers of Hirotugu Akaike*, pp. 199–213, Springer, 1998.

BIOGRAPHIES



Mohammad Abu Alsheikh (mabualsh@ieee.org) [S'14–M'17] is an Assistant Professor and ARC DECRA Fellow at the University of Canberra, Australia. He was a Postdoctoral Researcher at Massachusetts Institute of Technology, USA. He designs and creates novel IoT systems that leverage both machine learning and convex optimization with applications in people-centric sensing, human activity recognition, and smart cities. His doctoral research at Nanyang Technological University, Singapore was focused on optimizing the data collection in wireless sensor networks. After graduating with a B.Eng. degree in computer systems from Birzeit University, Palestine, he worked as a software engineer at a digital advertising start-up and Cisco.



Dinh Thai Hoang (hoang.dinh@uts.edu.au) [M'16] is a faculty member with the School of Electrical and Data Engineering, the University of Technology Sydney, Australia. He received his Ph.D. in Computer Science and Engineering from the Nanyang Technological University, Singapore, in 2016. His research interests include emerging topics in wireless communications and networking such as ambient backscatter communications, cognitive radios, wireless energy harvesting, IoT, mobile edge and 5G networks. He has been serving as an active member of technical program committees and a reviewer for many IEEE Transactions, Journals, Magazines, and Conferences. He received the best reviewer award from IEEE Transactions of Wireless Communications and he is currently an editor of IEEE Wireless Communications Letters and IEEE Transactions on Cognitive Communications and Networking.



Dusit Niyato (dniyato@ntu.edu.sg) [M'09–SM'15–F'17] is currently a Professor in the School of Computer Science and Engineering, at Nanyang Technological University, Singapore. He received B.Eng. from King Mongkut's Institute of Technology Ladkrabang (KMUTL), Thailand in 1999 and Ph.D. in Electrical and Computer Engineering from the University of Manitoba, Canada in 2008. His research interests are in the area of energy harvesting for wireless communication, Internet of Things (IoT) and sensor networks.



Derek Leong (derekleong@alumni.cmu.edu) received the B.S. degree in electrical and computer engineering from Carnegie Mellon University in 2005, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 2008 and 2013, respectively. He was a Scientist with the Smart Energy and Environment cluster at the Institute for Infocomm Research (I²R), A*STAR, Singapore. His research and development interests include distributed systems, sensor networks, smart cities, and the Internet of Things.



Ping Wang (wangping@ntu.edu.sg) [M'08–SM'15] received the Ph.D. degree in electrical engineering from University of Waterloo, Canada, in 2008. Currently she is an Associate Professor in the Lassonde School of Engineering, York University, Canada. Her current research interests include resource allocation in multimedia wireless networks, cloud computing, and smart grid. She was a corecipient of the Best Paper Award from IEEE Wireless Communications and Networking Conference (WCNC) 2012 and IEEE International Conference on Communications (ICC)

2007.



Zhu Han (zhan2@uh.edu) [S'01–M'04–SM'09–F'14] received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor in Boise State University, Idaho. Currently, he is a Professor in

Electrical and Computer Engineering Department as well as Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, several best paper awards in IEEE conferences, and is currently an IEEE Communications Society Distinguished Lecturer.