

An In-depth Study of Commercial MVNO: Measurement and Optimization

Ao Xiao^{1,2*}, Yunhao Liu^{1,3*}, Yang Li¹, Feng Qian⁴, Zhenhua Li^{1✉}
 Sen Bai¹, Yao Liu⁵, Tianyin Xu⁶, Xianlong Xin²

¹Tsinghua University ²Xiaomi Technology Co. LTD ³Michigan State University

⁴University of Minnesota, Twin Cities ⁵SUNY Binghamton ⁶University of Illinois Urbana-Champaign

ABSTRACT

Recent years have witnessed the rapid growth of mobile virtual network operators (MVNOs), which operate on top of the existing cellular infrastructures of base carriers, while offering cheaper or more flexible data plans compared to those of the base carriers. In this paper, we present a nearly two-year measurement study towards understanding various key aspects of today's MVNO ecosystem, including its architecture, performance, economics, customers, and the complex interplay with the base carrier. Our study focuses on a large commercial MVNO with about 1 million customers, operating atop a nation-wide base carrier. Our measurements clarify several key concerns raised by MVNO customers, such as inaccurate billing and potential performance discrimination with the base carrier. We also leverage big data analytics and machine learning to optimize an MVNO's key businesses such as data plan reselling and customer churn mitigation. Our proposed techniques can help achieve higher revenues and improved services for commercial MVNOs.

CCS CONCEPTS

- Networks → Network performance analysis; Network measurement; Mobile networks.

ACM Reference Format:

Ao Xiao, Yunhao Liu, Yang Li, Feng Qian, Zhenhua Li, Sen Bai, Yao Liu, Tianyin Xu, and Xianlong Xin. 2019. An In-depth Study of Commercial MVNO: Measurement and Optimization. In *The 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '19)*, June 17–21, 2019, Seoul, Republic of Korea. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3307334.3326070>

1 INTRODUCTION

As propelled by the increasing market demand, mobile virtual network operators (MVNOs) have quickly gained popularity and commercial success in recent years [35]. The global MVNO market revenue has reached about \$60.5 billion in 2018 [34], and is predicted to grow to \$103 billion in 2023 [19]. Also, the number of

* Co-primary authors. Zhenhua Li is the corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiSys '19, June 17–21, 2019, Seoul, Republic of Korea

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6661-8/19/06...\$15.00

<https://doi.org/10.1145/3307334.3326070>

MVNOs worldwide has exceeded 1000 [23]. MVNOs operate on top of the existing cellular infrastructures of *base carriers*, while offering cheaper or more flexible data and voice/SMS plans compared to those of the base carriers. In addition, MVNOs can enhance the utilization of base carriers' infrastructures. They may also promote the competition in the telecommunication market and prevent potential monopoly to some extent. Furthermore, MVNOs are expected to play a critical role in the 5G era [24]. MVNOs that obtain sliced radio access network resources from base carriers would act as a forerunner in mobile network slicing, which is a crucial technology in the revolution towards 5G cellular wireless networks.

Despite the unprecedented growth of MVNOs, end customers are still subject to manifold concerns when switching to MVNOs. For example, severe overcharge problems have been reported with regard to MVNO users [17]. Also, some base carriers may deliver MVNO users' data with lower priorities, leading to inferior performance experienced [25, 32, 36]. All above concerns may impede the reputation of MVNOs and hinder their development.

Due to the very few studies in the public literature, the research community lacks a thorough understanding of the MVNO ecosystem, including its architecture, performance, economics, customers, and the complex interplay between an MVNO and its base carrier. In this paper, we present a nearly two-year measurement study towards understanding the above aspects. Our study focuses on Xiaomi Mobile¹ (abbreviated as V-Mobile), a large commercial MVNO in China. It has about 1 million users and fully operates on top of China Telecom² (abbreviated as B-Mobile), a nationwide base carrier in China which has over 300 million users. V-Mobile is a representative *light* MVNO (§2), the most popular type of MVNOs that fully rely on the base carrier's cellular infrastructure [35], while having the capability of designing their own data plans independently of the base carriers. In other words, V-Mobile resells data plans purchased from B-Mobile to its users. In order to attract customers while gaining profits, V-Mobile has to (1) strategically design its own data plans, and (2) judiciously purchase data plans from B-Mobile to fulfill the data plans selected by V-Mobile customers, based on estimating their monthly data usage.

Studying the MVNO ecosystem is challenging, as it involves multiple stakeholders (MVNO customers, the MVNO carrier, and the base carrier) that incur complex interplay, as well as components that are not present in traditional MNOs (Mobile Network Operators) such as data plan reselling. To enable our large-scale measurement, we collaborate with V-Mobile and collect five datasets (§3): (1) BBSDataset, which consists of about 12,000 posts

¹Xiaomi Mobile Virtual Network Operator. <https://www.mi.com/mimobile/>.

²China Telecom Mobile Network Operator. <http://www.189.cn/>.

from the Bulletin Board System (BBS) maintained by V-Mobile; (2) UserDataset, which includes all customers' basic information, such as the gender, birthday, and the date of joining/leaving V-Mobile; (3) MonthlyDataset, which refers to each user's monthly data plan, payment, data usage, and account balance information; (4) PerfDataset, which comprises the network performance statistics of 342 V-Mobile users and 250 B-Mobile users who opted in to help us collect detailed network performance data; (5) SURDataset, which is the set of two types of state update reports (SURs) that are collected from V-Mobile's accounting center (AC). These datasets include customers' demographic information, selected data plans, monthly data usage, traffic characteristics, network performance, and billing events. Jointly examining these datasets helps reveal a comprehensive landscape of the MVNO ecosystem. Specifically, we perform detailed studies of the following aspects in this paper:

- **Data Plan and Usage Characterization (§4).** We find that V-Mobile customers are sensitive to prices in that almost half of the customers prefer the data plan with the lowest cost per GB (\$2.84/GB). Also, the customers' demographics are different from those of a regular MNO: there are significantly more female users (58.3%) than male users (41.7%), and V-Mobile customers are dominated by users younger than 30. Regarding the actual data usage, we find that V-Mobile users typically considerably under-utilize their data plans. Besides, users' actual data usage is weakly correlated with their subscribed data plans. The above findings suggest that V-Mobile users' data plan selection is oftentimes economically suboptimal, while such suboptimality is a key reason why operating an MVNO can be profitable.
- **Network Performance Characterization (§5).** By analyzing PerfDataset, we find no noticeable network performance difference between B-Mobile and V-Mobile. This is distinct from previous reports where MVNO users may suffer from performance degradations compared to the base carrier users [25, 32, 36]. Our results indicate that the performance degradation reported by prior studies is very likely a (man-manipulated) policy-level result by the base barrier rather than a technical necessity.
- **Monthly Data Usage Prediction (§6.1)** plays a critical role in the operation of an MVNO. We demonstrate that by strategically applying off-the-shelf machine learning algorithms enhanced with domain-specific data preprocessing, we can achieve an average prediction accuracy of 86.75%. In addition to performing monthly data usage prediction on a per-user basis, we also holistically examine the entire user base by conducting *uncertainty modeling*, which establishes a global statistical distribution for the prediction accuracy. Doing so facilitates the prediction for new V-Mobile users with insufficient training samples.
- **Data Reselling Optimization (§6.2).** We develop a framework that leverages our data usage prediction technique to optimize V-Mobile's reselling profit, *i.e.*, intelligently selecting the data plans to be purchased from B-Mobile in order to maximize the overall profit for V-Mobile. When applied to the 0.7 million active users in our dataset, our machine learning technique increases V-Mobile's profit rate by 17.1% compared to V-Mobile's current approach. Applying the uncertainty modeling further improves the profit rate by 34.1% – an overall improvement of 57.1%.

• **Customer Churn Profiling and Mitigation (§7).** Customer churn refers to when a customer cancels or “drops out” her subscribed cellular service. We mine our dataset to reveal key factors that indicate a user's forthcoming service cancellation. We then use them to proactively predict customers' churn, and manage to achieve both high precision (95.52%) and recall (94.45%). We further collaborate with V-Mobile to conduct a field trial. By sending to 1% of the customers with high churn probabilities (based on our prediction) a small gift card of \$1.44, V-Mobile can reduce the customers' churn rate from 0.83% to 0.32%.

• **Understanding Inaccurate Billing Issues (§8).** We systematically investigate V-Mobile users' complaints on billing issues, such as their unexpectedly high monthly charges. We find that most billing issues are caused by the excessive propagation delay of the control-plane billing state update reports (SUR). The overall delay often reaches several minutes, causing billing inconsistency between B-Mobile and V-Mobile.

To our knowledge, this is so far the most comprehensive study of a large commercial MVNO. Our high-level contributions are multifold. From an end user's perspective, our study demystifies the MVNO ecosystem, and clarifies several key concerns raised by customers. From an MVNO's perspective, we leverage statistical modeling, big data analytics, and machine learning to optimize an MVNO's key businesses such as data usage prediction, data plan reselling, and customer churn mitigation. Our study also exposes to the community new research topics, with intra-disciplinary nature, in the context of MVNO and its interplay with the base carrier.

2 BACKGROUND

A base carrier typically owns a legal license to exclusively use certain frequencies of the radio spectrum in a country. This can often lead to market monopoly, service degradation, or under-utilization of radio resources. To address these problems, numerous MVNOs have emerged in recent years. An MVNO fully or partially leverages one or multiple base carriers' licensed radio spectrum and facilities.

According to their degrees of dependence on base carriers, today's MVNOs can be classified into three categories [10, 35]: *skinny MVNOs*, *light MVNOs*, and *thick MVNOs*. Skinny and light MVNOs do not have their own radio infrastructures; the former are mainly devoted to marketing and sales, and are thus also known as “branded resellers,” while the latter further have the ability to design specialized data plans independently of the base carriers. In some countries like China, MVNOs have emerged and flourished for only several years, and they have not been allowed to set up their own infrastructures so far. In fact, light MVNOs act more as a transitional paradigm between the skinny MVNOs and the thick MVNOs. In contrast, thick MVNOs have their own infrastructures to exert more control over their offerings, which however are not permitted in many countries. Among the three categories, light MVNOs are the most common and are permitted in most countries [35]. In this paper, our studied V-Mobile is a typical light MVNO.

Data Plan Reselling. As a light MVNO, V-Mobile resells data plans purchased from B-Mobile to its users. However, V-Mobile does not have a wholesale (discounted) price from B-Mobile, nor is it allowed to buy a single data plan from B-Mobile to serve multiple

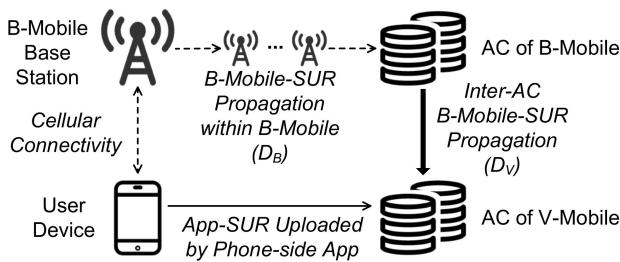


Figure 1: Interactions among V-Mobile, its customers, and B-Mobile.

users. Therefore, in order to attract customers while gaining profits, V-Mobile has to strategically design its own data plans. Specifically, let P_V be a data plan that V-Mobile offers its customers; let P_B be the same plan offered by B-Mobile; let P'_V be the corresponding plan that V-Mobile purchases from B-Mobile to fulfill P_V . Ideally, P_V needs to be cheaper than P_B , otherwise customers have no incentives to switch to V-Mobile; P'_V should be inferior to (and thus cheaper than) P_V , otherwise operating V-Mobile is not profitable.

Consider a concrete example. B-Mobile has a monthly plan of “\$7.23 for 2 GB plus \$0.01/MB overdraft” (P_B). To attract customers, V-Mobile lowers the price by offering “\$7.09 for 2 GB plus \$0.01/MB overdraft” (P_V). To fulfill P_V , V-Mobile can purchase from B-Mobile an inferior plan of “\$5.78 for 1 GB plus \$0.01/MB overdraft” (P'_V). The assumption here is that, despite having a 2 GB data plan, many users’ actual monthly usages are much lower. In the above example, the reselling is profitable if a V-Mobile user’s monthly usage is lower than $1000 + (7.09 - 5.78) / 0.01 = 1131$ MB. This example also illustrates the importance for V-Mobile to *accurately predict a user’s monthly data usage*.

Billing. As the base carrier of V-Mobile, B-Mobile has tens of millions of customers and over a million base stations across the country. Since V-Mobile is a light MVNO, all its SIM cards are issued by B-Mobile while possessing special phone numbers (MSISDN) that can be identified as MVNO numbers. Therefore, all the data traffic, voice calls, and SMS messages of V-Mobile users are delivered by the radio infrastructure of B-Mobile. Moreover, since V-Mobile establishes its own data plans, it needs to handle customers’ billing, which is processed at the AC, by itself. In order for V-Mobile to generate bills in an accurate and timely fashion, the SURs, which contain users’ up-to-date data usage information, need to be properly delivered to V-Mobile’s AC. This can be achieved in two ways. The major way is shown in the upper branch in Figure 1. Per the agreement between V-Mobile and B-Mobile, they establish a control-plane channel between their ACs, allowing B-Mobile to forward V-Mobile users’ SURs, which we call B-Mobile-SURs, to V-Mobile. In B-Mobile, B-Mobile-SURs are gathered and propagated by its base stations. Another way of delivering SURs is illustrated in the bottom branch in Figure 1. V-Mobile customers can optionally install an app developed by V-Mobile. When the app is running, it keeps monitoring the client device’s data usage and uploading SURs, which we call App-SURs, directly to V-Mobile’s AC.

3 MEASUREMENT DATA COLLECTION

To enable our large-scale measurement of a typical MVNO, we collect multiple datasets from various sources. Correlating these datasets focusing on different aspects helps reveal a complete landscape of the MVNO ecosystem. We next describe our datasets in details. Note that all users from whom we collected the data were presented with informed consent, and all customers’ identities were fully anonymized before any analysis was conducted.

(1) **BBS Dataset.** V-Mobile maintains a BBS (Bulletin Board System) website from which we gathered about 12,000 posts from the BBS users, who belonged to either the current or prospective users of V-Mobile. We then manually examine all the posts to understand the users’ concerns. The content of the posts can roughly be classified into four categories. 25.4% of the posts concern with the network performance such as whether V-Mobile can provide performance that is as good as that offered by B-Mobile; 5.2% of the posts express users’ doubts on their received bills and request for auditing; 37.9% of the posts involve specific complaints with regard to billing state update delay, e.g., notifications of data overdraft oftentimes arrive late, causing unexpected overdraft expense. We will revisit its root cause in §8. The remaining 31.5% of the posts are non-technical concerns, e.g., the appearance of V-Mobile’s SIM cards, and the purchase, repair and refund instructions.

(2) **UserDataset.** We obtain from V-Mobile a database containing each customer’s basic information, including the gender, birthday, the date of joining V-Mobile, and the date of canceling the service (for dropout users). Our dataset was captured in Oct. 2017, involving a total of 908,548 users since Jan. 2016. In this “snapshot”, 168,853 users terminated their contracts with V-Mobile and thus are excluded. We use this dataset for various types of analysis such as data usage characterization (§4), data usage prediction (§6.1), and customer churn analysis (§7). Hereafter, unless otherwise noted, we apply our analysis to the 0.7 million active customers.

(3) **MonthlyDataset.** The AC of V-Mobile archives every user’s monthly data plan, data usage, payment, and account balance information. They provide an essential data source for profiling, modeling, and predicting each user’s monthly data usage (§4 and §6.1). The data is also used for optimizing the data reselling profit of V-Mobile (§6.2), as well as customer churn profiling and mitigation (§7). In this paper, our studied **MonthlyDataset** covers 22 billed months, i.e., from January 2016 to October 2017.

(4) **PerfDataset.** A challenge we face is to monitor V-Mobile users’ network performance. We take a crowd-sourcing-based approach where we invite users of both V-Mobile and B-Mobile to voluntarily participate in a data collection campaign by installing an app developed by V-Mobile and turning on the “sampling performance” option for a whole month. The app we leverage for the campaign only works on Android (not iOS), so it is likely to introduce some bias to the measurement results. The app passively measures key performance metrics such as RTT and throughput of users’ traffic (we detail the measurement methodologies in §5). The collected statistics are securely uploaded to a remote server when users’ devices are idle, and no actual content payload was collected or uploaded. Overall, 342 V-Mobile users and 250 B-Mobile users from 269 cities in the country voluntarily participated in our study for 30 days, forming a decently large dataset consisting of 1 TB TCP

Monthly Data	Subscription Fee	Overdraft Fee	Percentage of Users
1 GB	5.64	\$0.01/MB	35.79%
2 GB	7.09	\$0.01/MB	10.52%
3 GB	8.53	\$0.01/MB	47.37%
4 GB	12.87	\$0.01/MB	6.32%

Table 1: V-Mobile’s data plans and their percentage of users.

Monthly Data	Subscription Fee	Overdraft Fee
500 MB	\$3.62	\$0.01/MB
1 GB	\$5.78	\$0.01/MB
2 GB	\$7.23	\$0.01/MB
3 GB	\$8.68	\$0.01/MB
4 GB	\$13.01	\$0.01/MB
6 GB	\$27.47	\$0.01/MB

Table 2: B-Mobile’s data plans (selected).

flows’ statistics. We leverage this unique dataset to characterize the performance of B-Mobile and V-Mobile in §5.

(5) SURDataset. Recall from §2 (Figure 1) that there are two types of SURs: B-Mobile-SURs and App-SURs. We collect both types of SURs at V-Mobile’s AC, and use this data to investigate the billing issues in §8. An SUR consists of a timestamp, a device ID, and the user’s data usage in bytes. The two types of SURs can be correlated by counting the data bytes consumed by a customer.

4 DATA USAGE CHARACTERIZATION

We characterize V-Mobile customers in three aspects: data plan selection, actual data usage, and the relationship between them. Whenever possible, we also compare V-Mobile with B-Mobile to unravel the differences between an MVNO and its base carrier.

Data Plan Selection. Recall from §2 that V-Mobile makes its data plans cheaper than those (counterparts) of B-Mobile to attract customers. Table 1 and Table 2 compare the specific data plans offered by both carriers, from which we find that V-Mobile customers can actually save a small portion (1.1%–2.4%) of money. Although the economic incentives seem weak, V-Mobile also provides “soft” value-added services to its customers in terms of billing, state update, customer service, and so on. Furthermore, V-Mobile can operate on more than one base carrier and is expected to support the cross-carrier portability of users’ phone numbers in the near future. Table 1 also shows the percentage of users subscribing to each of the V-Mobile’s data plans, as obtained from the UserDataset. As shown in Table 1, V-Mobile users are price conscious: the 3 GB data plan attracts most (47.37%) of the users because it has the lowest cost per GB (\$2.84/GB), followed by the 1 GB plan that has the lowest subscription cost. In contrast, the 4 GB plan has the fewest (6.32%) users, very likely because of its highest subscription cost as well as a large subscription cost increase from 3 GB to 4 GB.

V-Mobile’s data plans are designed according to its current commercial contract with its base carrier. Using one B-Mobile plan to serve multiple V-Mobile users is not allowed; this restriction stems from business strategies rather than technical reasons. This restriction is generalizable to the other MVNOs in China

such as Snail Mobile (<http://mobile.snail.com/>) and Suning Mobile (<http://10035.suning.com/>), but not the MVNOs in the US or Europe. As time goes by, it is likely to be cancelled or mitigated through business negotiation or even lawsuits in court. On the other hand, V-Mobile is allowed to adjust the threshold and price of each data plan without asking for B-Mobile’s permission.

We also make several interesting observations regarding V-Mobile users’ demographic distributions. As shown in Figure 2, there are significantly more female users (58.3%) than male users (41.7%) in particular for large (3 GB and 4 GB) plans. Women tend to prefer V-Mobile to B-Mobile, only 45% of whose customers are female. Furthermore, by analyzing UserDataset we notice that the ratio of female users increased by 1% in one year (from 2016 to 2017). Regarding the age distribution, as shown in Figure 3, most of the V-Mobile customers are younger than 30, with the exception where the 4 GB data plan holders are dominated by people between 30 and 39 who are presumably more financially sound. From 2016 to 2017, the ratio of users younger than 30 increased by 1.7%, and the ratio of users older than 50 increased by 1%; on the contrary, the other users between 30 and 49 decreased by 2.7%. In other words, V-Mobile’s penetrations have gradually increased as for children and seniors. The above observations can be leveraged by MVNOs for developing demographic-targeted products and promotions.

Figure 4 plots the distribution of V-Mobile users’ lifetime, which is defined as the time span from when the user joined V-Mobile to T_{end} . T_{end} is either the time when the user dropped out (canceled the V-Mobile service), or October 2017 which is the time when the data was collected if the user was still active at that time. We plot the lifetime for dropout and active users separately. As shown, active users’ lifetime ranges from 0 to 22 months and averages at nearly 16 months. Note that the 22-month cap exists because V-Mobile started its business 22 months before Oct. 2017. For dropout users, their lifetime is statistically even longer. Overall, although the absolute lifetime of V-Mobile users is significantly shorter than that of typical MNO users (ranging from 0 to 138 months and averaging at 52 months [18]), most V-Mobile users tend to have a relatively long lifetime compared to the lifetime of V-Mobile itself. This implies that most users appear to be satisfactory with or are willing to keep using V-Mobile’s services.

Actual Data Usage. We next study the data usage characteristics of V-Mobile users. Figure 5 plots the distribution of their monthly data usage across all of their billed months. It ranges between 0 and 6,018 MB with a mean (median) being 1,743 MB (1,950 MB). We make two observations from Figure 5. First, the tail indicates that overdraft may indeed occur in a non-trivial fraction of billed months, as to be quantified shortly. This brings potential obstacles for MVNOs towards making profits (§2). Second, more surprisingly, we see the other extreme for about 20% of the billed months in which a user consumes little data – we call this the *intermittent presences* of some users. In particular, some new users have just been enrolled for less than one month, which obviously have no monthly data usage. Suppose a user has been a V-Mobile customer for n full months while she barely used her data plan in m out of n months. We empirically define a “bare usage” as consuming no more than 50 MB of data in a month with only full months being considered. We then compute the *presence ratio* of the customer as $(n - m)/n$ to

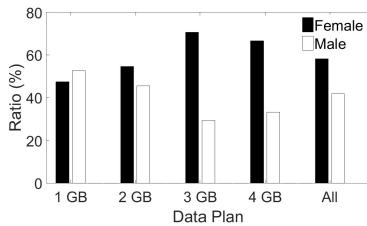


Figure 2: Gender distribution for each of V-Mobile’s data plans.

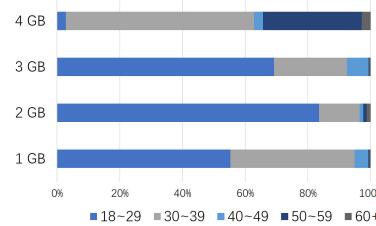


Figure 3: Age distribution for each of V-Mobile’s data plans.

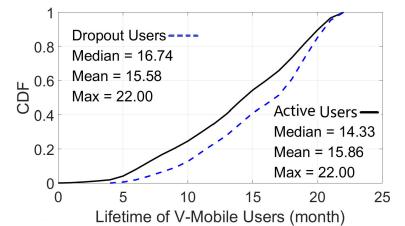


Figure 4: Lifetime of V-Mobile users (active and dropout, as of 10/2017).

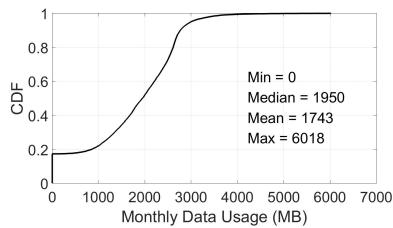


Figure 5: Distribution of V-Mobile active users’ monthly data usage.

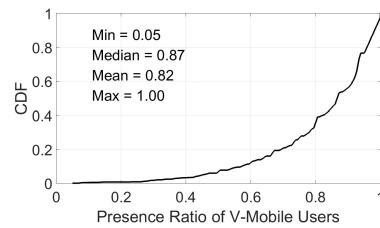


Figure 6: Distribution of V-Mobile active users’ presence ratios.

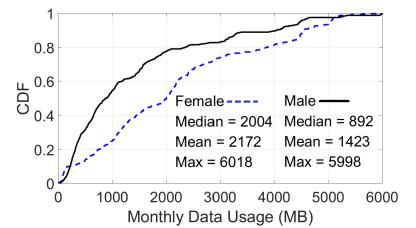


Figure 7: Distributions of male and female active users’ monthly data usage.

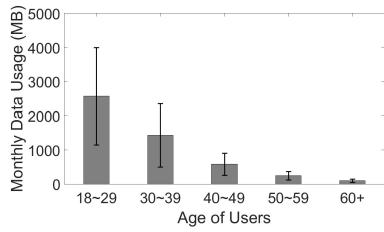


Figure 8: Relationship between users’ age and their monthly data usage.

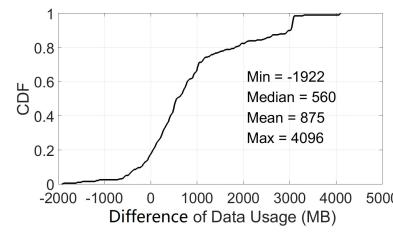


Figure 9: Difference between users’ selected data plans and actual data usage.

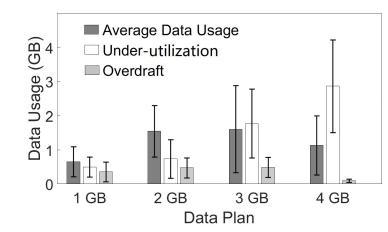


Figure 10: Avg. data usage, under-utilization & overdraft for each plan.

quantify how often she uses the V-Mobile service. Figure 6 plots the presence ratio distribution across users whose lifetimes are at least one full month. As shown, the presence ratio ranges from 0.05 to 1 and averages at 0.82. About 23% of the users have a presence ratio lower than 0.75. For the users with low presence ratios, V-Mobile made a survey with them randomly via telephone interviews, which reveal the reason – such low presence ratios are oftentimes from users who hold more than one SIM cards and use V-Mobile as a “backup” one that is used infrequently. These users can bring high profits if V-Mobile can accurately predict their non-presence.

We next investigate how users’ demographics affect the data usage. Figure 7 plots the distributions of the males’ and females’ monthly data usage across all billed months. It shows the monthly data usages of users whose lifetimes are at least one month. As shown, compared to males whose median (mean) monthly usage is 892 MB (1,423 MB), females tend to use much more data, with their median (mean) usage being 2004 MB (2,172 MB). The results echo our finding about the relationship between gender and data plan selection in Figure 2. They are also in line with a recent study showing females tend to be heavier smartphone users compared to males [1]. Figure 8 depicts the statistics (mean and stdev) of each

age group’s monthly data usage. We observe a clear trend where young people consume considerably more cellular data than older people: on average teenagers and the twenties consume 80% more data compared to people in their thirties, and 20× more compared to senior citizens (60+).

Data Plan vs. Actual Usage. To better understand the relationship between users’ selected data plan and users’ monthly actual usage, we plot the distribution of their difference in Figure 9 across all billed months, where a positive value indicates under-utilizing the data plan, and a negative value corresponds to overdraft. Figure 10 further shows the statistics of actually consumed, under-utilized, and overdraft bytes for all billed months of each type of data plan. As shown, customers typically under-utilize their data plans (in 82.59% of all billed months). An average user under-utilizes her data plan by as much as 893 MB. This indicates that most users are conservative with their data usage, in particular given that cellular data usage monitoring and limiting tools have been incorporated into today’s mobile OSes [20].

We also make a surprising observation from Figure 10 that users’ actual data usage is weakly correlated with their subscribed data plans: the average usage of the 2 GB plan is similar to that of the

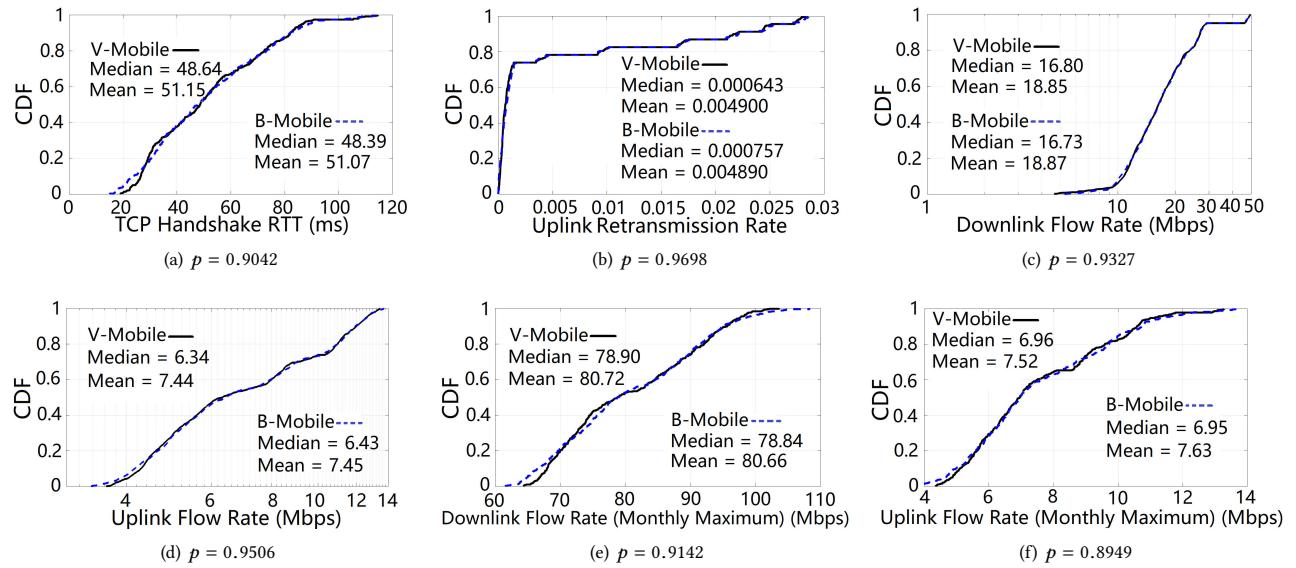


Figure 11: Network performance characterization of crowd-sourced V-Mobile and B-Mobile users.

3 GB plan, and the usage of the 4 GB plan is even less than 2 GB and 3 GB plans. This implies that *MVNO customers' data plan selection is oftentimes economically suboptimal* (especially for older customers who prefer larger data plans), while such suboptimality is a key reason why operating an MVNO can be profitable.

5 NETWORK PERFORMANCE

We now shift our focus to the network performance characterization by analyzing the PerfDataset, which, recall from §3, was collected from about 600 V-Mobile and B-Mobile users from an Android app. Once the “sampling performance” option is turned on, the app leverages libpcap to passively monitor the client device’s traffic (only the TCP/IP headers) and compute several key performance metrics: (1) TCP handshake RTT, (2) TCP flow size, *i.e.*, the total number of payload bytes within a flow, (3) TCP flow duration, *i.e.*, the time span between the first and last packet of a flow, (4) TCP flow rate, *i.e.*, the ratio between flow size and duration, with uplink and downlink measured separately, and (5) TCP uplink packet retransmission rate, *i.e.*, the ratio between retransmitted uplink data packets and the total number of uplink data packets within a flow (note that we are unable to measure the downlink retransmission rate). In our subsequent analysis, for flow rate calculation, following similar methodologies of prior studies [14, 29], we only consider flows that are at least 100 ms and at least 1MB; for uplink retransmission rate, we only examine flows containing at least 10 uplink packets, to make the metric meaningful.

The above are known to be the key metrics quantifying the network performance [29]. Compared to prior characterizations of commercial cellular network performance (*e.g.*, 3GTest in 2009 [15], 4GTest in 2011 [13], SpeedTest in 2011 [33] and an LTE study in 2012 [14]), our study focuses on both MVNO and MNO, and provides more up-to-date statistics that can benefit other work on, for example, synthetic cellular traffic generation and cellular performance modeling. Our measurement results are shown in Figure 11’s

six subplots. We plot the statistics for B-Mobile and V-Mobile separately. Each CDF in subplots (a) to (d) is across all eligible TCP flows; subplots (e) and (f) show the monthly maximum flow rates across all users’ billed months, to reveal the maximum capacity that today’s cellular networks can offer.

From the BBSDataset, we learn that many V-Mobile users have concerns that B-Mobile may potentially deliver their traffic at a lower priority. Surprisingly, our results in all plots of Figure 11 disprove this by showing that statistically, there is no noticeable performance difference between V-Mobile and B-Mobile in all six metrics. Since we collect the samples from a small number of users, we further use the paired-samples T-test [12] to compare their performances. The T-test result is a confidence (p -value) between 0 and 1. We accept the hypothesis that their mean values are equal if p is larger than α , a threshold typically set to 0.05. As listed in Figure 11, all the values of p are much bigger than 0.05. Therefore, even considering the impact of small samples, we still find no significant difference between their performances. We ascribe this to the non-discriminatory packet routing policy of B-Mobile in delivering V-Mobile users’ data, recalling that V-Mobile is a light MVNO that fully operates on B-Mobile’s communication infrastructure (§2).

In contrast, several previous reports [25, 32, 36] noted that the users of certain MVNOs had suffered from performance degradations compared to the base carriers’ users. For a light MVNO, our measurement results indicate that such a performance degradation is very likely a (man-manipulated) policy-level result rather than a technical necessity. The policy-level result means that it is the commercial policies of the base carrier that degrade the light MVNO users’ network performance, rather than technical difficulties. Consequently, when a light MVNO can have desirable business negotiations with its base carrier, its users would not suffer from inferior network performance. On the contrary, for a full MVNO that has its own wireless infrastructure, such a performance degradation might be (partially) owing to technical reasons.

6 DATA USAGE PREDICTION AND DATA RESELLING OPTIMIZATION

As described in §2, monthly data usage prediction plays a pivotal role in the operation of an MVNO. It also enables numerous other use cases such as data plan recommendation, cellular infrastructure planning, and cost-aware content prefetching [30]. For the MVNO customers, it can help them choose more appropriate data plans for saving money since most customers do not use up their monthly data plans (see Figure 9). For the base carrier, data usage prediction can help it design more appropriate data plans to fit more users' appetites. Additionally, our customer churn prediction can also benefit from data usage prediction to better retain customers (§7). In this section, we first describe the challenges and solutions for large-scale monthly data usage prediction (§6.1), and then detail how an MVNO can leverage it to optimize the reselling profit (§6.2).

6.1 Data Usage Prediction and Modeling

Monthly Data usage prediction is a typical *time series forecasting* problem, which utilizes a model to predict a variable's future value(s) based on its previously observed values. In our scenario, for each user we have a time series $X = \{X_1, X_2, \dots, X_n\}$ where X_i is the user's data usage in the i -th month. Our objective is to predict X_{n+1} , the data usage of the $(n+1)$ -th month. The prediction of each user is independently performed. For a given user, we define the prediction accuracy as:

$$\text{Accuracy} = 1 - \frac{|\text{predicted_usage} - \text{actual_usage}|}{\max(\text{predicted_usage}, \text{actual_usage})} \quad (1)$$

From a carrier's perspective, performing data usage prediction faces several challenges. First, customers' usage randomness makes an accurate prediction of *every* user's data usage impossible. We thus aim at achieving statistically good prediction results for the whole user base. Second, as the prediction is conducted on a per-month basis, each user only has a limited amount of historical data. Recall from §4 that an average V-Mobile user's lifetime is only 16 months. Third, as the carrier needs to perform the prediction for millions of users, we prefer an algorithm that is as lightweight as possible. In the literature, although there exist many studies on (typically short-term) traffic volume prediction, few studies, if any, specifically focus on long-term, large-scale monthly cellular data usage prediction with limited historical data. This is partially due to a lack of real cellular usage datasets.

Methodology. We apply mainstream machine learning techniques for a user's monthly data usage prediction, including SVR (Support Vector Regression [37]), RBFNN (Radial Basis Function Neural Network [4]), BPNN (Back Propagation Neural Network [31]), and ULR (Unary Linear Regression [3]). Each user is trained and tested separately. We empirically select the algorithms' parameters as follows. As for SVR, it uses the RBF (Radial Basis Function) kernel with the kernel parameter of 0.01 and regularization parameter of 100. As for RBFNN, we set the spread of radial basis function to 0.1 and the number of neurons which are added between displays to 100. Meanwhile, we limit the maximum number of neurons to 400. BPNN uses "tansig" and "purelin" for hidden layers and output layers. The maximum number of iterations is set to 100. Parameter settings are actually complex in our prediction algorithms because

different datasets have their specific characteristics. However, the algorithms are fixed and thus only the parameters need to be tuned to fit the datasets.

Another decision we need to make is to determine n , the window length of the time series X . It poses a tradeoff: a large n reduces the number of valid training samples and may cause potential overfitting, while a small n makes the training data less expressive. Specifically, our model is trained in the form of sliding windows on each user's historical data usage, where each user's historical data usage constitutes a time series. We test different values of n with comprehensive experiments on our collected big data and empirically choose $n=3$ to balance the above tradeoff.

Furthermore, in order to filter out outliers and to avoid potential overfitting, we perform preprocessing on each user's entire historical data usage using k -means clustering in both training and testing. Specifically, we set $k=2$ and only retain the larger cluster by removing the smaller cluster and concatenating the remaining samples in a chronological order. We empirically observe that such a simple preprocessing step reasonably handles the intermittent presence problem of most users (§4) as the infrequent low-usage months can typically be filtered out. As a matter of fact, at the beginning (after data cleaning) we did not filter out outliers and the results turned out not good enough. Thus, we explored various possible ways to improve the accuracy of data usage prediction, and we found that filtering out outliers was the most effective to achieve our goal. In detail, quite a few "abnormal" factors such as users' travel and death may degrade the prediction accuracy of users' data usage. Nonetheless, we need to focus on the general trends of users' data usage by eliminating those extreme or accidental situations, so as to figure out a proper prediction model with a sufficiently high accuracy. Our methodology inevitably introduces some bias, which however is acceptable according to our experiences.

Results. For each user whose lifetime is l months, since we pick $n=3$, we use $\{s_1, s_2, s_3 \rightarrow s_4\}, \dots, \{s_{l-3}, s_{l-2}, s_{l-1} \rightarrow s_l\}$ to train a model, where s_i is the data usage of the i -th month obtained from the *MonthlyDataset*, using each of the aforementioned machine learning algorithms ("→" separates the features and label). In the prediction phase, we employ the model to predict s_{l+1} based on $\{s_{l-2}, s_{l-1}, s_l\}$. To evaluate our approach, we compute the prediction accuracy, defined in Equation (1), for each user, and compare the average accuracy of the five algorithms in Figure 12. In fact, V-Mobile itself adopts a quite simple method to predict each user's monthly data usage by calculating the average value in previous months. However, it does not work well since the prediction accuracy is merely 68.67%. Among all these methods, our method achieves an overall prediction accuracy of up to 86.75% across users (using SVR with RBF kernel with outlier filtering). The outlier filtering effectively improves the best accuracy across all algorithms from 80.35% to 86.75%.

We observe that users with a longer lifetime tend to exhibit a higher prediction accuracy because more historical data is available. This is illustrated in Figure 13, which plots the average prediction accuracy for users with different lifetime (using SVR with RBF kernel and outlier filtering). Moreover, SVR with RBF kernel, SVR with linear kernel, and ULR are very computationally efficient in that the total running time for all the 0.7 million active users is less

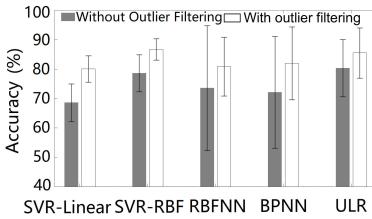


Figure 12: Average prediction accuracy of different ML techniques.

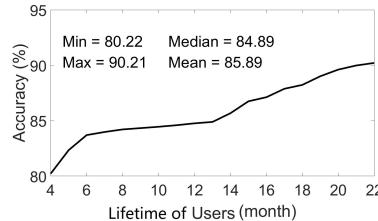


Figure 13: Avg. prediction accuracy increases with the lifetime of users.

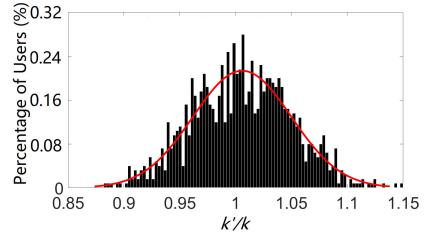


Figure 14: The random variable k'/k follows a normal distribution.

than 20 minutes including both training and testing on a server with Intel E5-Xeon Broadwell (V4) CPU and 16 GB memory. RBFNN and BPNN take a longer yet still acceptable time of 11 hours to train and test the whole dataset. We had also tried to use more attributes for data usage prediction, such as user information, plan type, and phone model. Unfortunately, this did not essentially improve the prediction accuracy, and using some attributes even degraded the accuracy. However, it is an interesting direction to explore how to further improve the accuracy, probably by considering other key attributes which we have not noticed yet.

Statistical Uncertainty Modeling. Despite the overall good results, Figure 12 indicates the high variation of the prediction accuracy across individual users. In particular, for users with a short lifetime, their data usage prediction is inherently difficult due to a lack of training samples. This motivates us to further holistically examine the entire user base by performing *uncertainty modeling*, which establishes statistical distributions for the prediction accuracy. As to be shown in §6.2, the uncertainty modeling helps users with insufficient training data; it also provides a building block for deriving a generic cost-aware optimization framework.

In metrology, all measurements are subject to uncertainty – a measurement result is complete only when it is accompanied by a statement of the associated uncertainty [11]. In our case, the way to reveal uncertainty is to understand the probability distribution of the prediction accuracy in the face of a large user base. We exemplify our modeling approach using the prediction results produced by SVR-RBF with outlier filtering, which yields the highest accuracy as shown in Figure 12. Other prediction methods’ results can be modeled in a conceptually similar manner (the actual distribution may differ though). For SVR-RBF, we make a key observation that the ratio between a user’s actual monthly data usage and predicted usage follows a normal distribution, as illustrated in Figure 14. Quantitatively, let k be the predicted data usage of a V-Mobile user, and k' be the actual usage in a billed month. Then the random variable $\frac{k'}{k}$ follows the normal distribution $N(1, \sigma^2)$, and its probability density function is:

$$h\left(\frac{k'}{k}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(k'-1)^2}{2\sigma^2}} \quad (2)$$

Thereby, the probability density function of the actual data usage (k') can be derived as:

$$g(x) = \frac{1}{\sqrt{2\pi}k\sigma} e^{-\frac{(x-1)^2}{2k\sigma^2}} \quad (3)$$

We next perform a hypothesis test to confirm our observation. We adopt the *Kolmogorov-Smirnov (K-S) test* [21], which is a non-parametric test that compares an empirical distribution of a set of samples with a reference probability distribution (in our case, the normal distribution). Suppose we have independently and identically distributed samples $\{X_1, \dots, X_n\}$ with some unknown distribution \mathbb{P} , and we want to decide between the following hypotheses:

$$H_0 : \mathbb{P} = N(1, \sigma^2), \quad H_1 : \mathbb{P} \neq N(1, \sigma^2). \quad (4)$$

For the random variable $\frac{k'}{k}$, we run the K-S test on our data. The test result is a confidence (p -value) between 0 and 1. we accept H_0 if the p -value is larger than α , a threshold typically set to 0.05. The numerical results indicate this is indeed the case, as the p -value is calculated to be $0.7949 \gg 0.05 = \alpha$. We note that for several other prediction algorithms such as ULR and RBFNN, their $\frac{k'}{k}$ also follows a normal distribution with $p > 0.05$ (e.g., $p = 0.7655$ for ULR). However, this may not hold for algorithms with a lower prediction accuracy such as BPNN and SVR-Linear, whose prediction results may conform to other types of statistical distributions.

6.2 Data Reselling Optimization

We now consider an important application of our data usage prediction: data reselling optimization. Recall that a V-Mobile user oftentimes does not use up her subscribed data plan P_V , so V-Mobile can purchase an inferior plan, P'_V , from B-Mobile to fulfill P_V , and utilize the price difference to make profits. Our goal here is to make the selection of P'_V more intelligent, in order to increase the profit for V-Mobile. This process is transparent to V-Mobile users.

Since our data usage prediction may not be accurate for a specific user, we are unable to guarantee that every individual user’s reselling profit is maximized. Instead, we aim at maximizing the *expected* data reselling profit across all users. We now describe our formulation and solution. Let there be a group of m customers with their predicted data usage $\{k_1, k_2, \dots, k_m\}$, and a collection of the base carrier’s data plans $\{P_1, P_2, \dots, P_n\}$. Consider any given user, say user j whose predicted usage is k_j . Let $p_i(k_j)$ be the *expected* expense, which includes the monthly subscription fee and the overdraft cost, that V-Mobile pays B-Mobile if V-Mobile selects P_i as the underlying plan to fulfill this user’s subscription. Since each user is handled independently by V-Mobile, the minimum total expense that V-Mobile pays B-Mobile is:

$$\sum_{j=1}^m \min_{1 \leq i \leq n} \{p_i(k_j)\}. \quad (5)$$

To calculate Equation (5), we need to first calculate $p_i(k_j)$:

$$p_i(k_j) = \int_0^{+\infty} f(x)g(x)dx, \quad (6)$$

where x is the independent variable of the customer's data usage, $g(x)$ is the probability density function of the user's actual data usage derived through uncertainty modeling (e.g., Equation (3) for SVR-RBF), and $f(x)$ is the payment function of x under data plan P_i . For example, $f(x)$ for “\$3.62 for 500 MB plus \$0.01/MB overdraft” is defined as $f(x) = 3.62$ for $0 \leq x \leq 500$, and $f(x) = \frac{x}{100} - 1.38$ for $x > 500$.

Based on our modeling results in §6.1, $\frac{k_j'}{k_j}$ follows a certain distribution. Assume we use SVR-RBF or ULR whose prediction results follow a normal distribution $N(1, \sigma^2)$. In this case, the maximum likelihood estimation $\hat{\sigma}^2$ is:

$$\hat{\sigma}^2 = \frac{1}{m} \sum_{j=1}^m \left(\frac{k_j'}{k_j} \right)^2 - 1. \quad (7)$$

From the historical data of V-Mobile, we calculate $\hat{\sigma}$ to be 0.04 for SVR-RBF and 0.22 for ULR.

To calculate Equation (6), we first rewrite it as a sum of two integrations $\int_0^S f(x)g(x)dx + \int_S^\infty f(x)g(x)dx$, where S is the data plan size (e.g., 500 MB in the above example), due to the piecewise nature of $f(x)$. We then apply the rectangle method [6] to numerically calculate each integration.

Evaluation. To assess the effectiveness of the above method, we apply it to the historical usage data of V-Mobile users between January 2016 and October 2017 (obtained from the *MonthlyDataset*), to optimize the reselling profit for the billing cycle of November 2017. The key evaluation metric, *profit rate (PR)*, is defined as $PR = (m_1 - m_2)/m_2$ where m_1 is the total expense that customers pay V-Mobile, and m_2 is the cost that V-Mobile pays B-Mobile. A higher *PR* is always preferred by V-Mobile. Note that a *PR* may be negative. We compare four optimization methods as follows:

- **The Current Approach** employed by V-Mobile, learned by us based on our communication with the company, works as follows: in each month, V-Mobile estimates each user's usage in the next month simply by calculating the arithmetic mean of all historical months' usage. V-Mobile then purchases the cheapest data plan (with the overdraft cost taken into account) from B-Mobile based on this rough usage estimation. We apply this method to our data and calculate the overall *PR* to be 3.5%.
- **Machine Learning Only.** For each customer, V-Mobile employs machine learning (SVR-RBF with outlier filtering, see §6.1) to predict next month's data usage, and then determines the data plan to be purchased from B-Mobile accordingly. This approach yields an overall *PR* of 4.1%.
- **Machine Learning with Uncertainty Estimation.** We apply the full method in §6.1 and §6.2 that combines the machine learning (SVR-RBF) and statistical uncertainty estimation. This approach leads to an overall *PR* of 5.5%, which is 57.1% higher than V-Mobile's current approach and 34.1% higher than the ML-only approach. We find that the significant increase of *PR* is attributed to better selections of B-Mobile's data plans for

customers with a short lifetime. For such customers with insufficient training samples, per-user machine learning is oftentimes ineffective. Instead, the cross-user prediction accuracy modeling (§6.1) can provide a more reasonable estimation of the expected expense ($p_i(k_j)$) based on the “big data.”

- **The Optimum.** To estimate the upper bound of the profit that V-Mobile can make, we use the ground-truth data, i.e., customers' actual data usage in Nov. 2017, to compute the optimal data plan selection. This leads to a *PR* of 6.2%, only 12.7% higher than our achieved *PR* of 5.5%.

7 CUSTOMER CHURN PROFILING AND MITIGATION

Customer churn refers to when a customer ceases her relationship with a company (in our case, canceling her cellular service or dropout). Its mitigation [2] is extremely important for both base carriers and MVNOs because of the high expense that a carrier needs to pay to recruit new customers.

Correlation Analysis. We seek answers to the following key question: which properties (features) of customers are good indicators of their forthcoming service cancellation? Understanding this is a key prerequisite of performing effective churn management in order to retain customers. We take a data-driven approach to address this problem. Specifically, we first compile a set of 12 features that cover a wide range of properties of users' personal, data usage, and performance. Obtained from *UserDataset* and *MonthlyDataset*, the 12 features are listed as follows: users' gender, age, lifetime, data plan, monthly data usage, monthly expense, monthly uplink bytes, monthly downlink bytes, roaming events, account balance, state update performance (§8), and device type. For each of a user's billed months, we generate a vector containing the 12 features, as well as obtain a binary label representing the churn state (whether the user has canceled her service) in the next month. To further enrich the feature set, we extend the features with numerical values by computing basic statistical functions such as *mean()*, *median()*, *stdev()*, and *mean_diff()* (mean of the first order difference) over the user's past months.

We next study the correlation between each of the features and the label (the churn state of the next month) across all customers' billed months. We compute the Spearman's rank correlation coefficient [8], a robust, nonparametric measure of the dependency between the rankings of two statistical variables. The coefficient ranges between -1.0 (negative correlation) and +1.0 (positive correlation). We include both the active and dropout users (about 1 million in total) in this study. Figure 15 lists the top 8 features that have the strongest correlations with the customer churn. We make several observations as detailed below. To help quantify our findings, we define a metric called *Dropout Contribution (DC)*, which is the percentage of dropout users with certain properties (e.g., having a specific lifetime range) among all dropout users. *DC* is calculated by examining each user's features associated with the month in which she dropped out.

- Figure 16 plots the relationship between users' lifetime and their *DC*. As shown, the *DC* increases as users stay longer. For example, 26.6% of all V-Mobile users have a lifetime between 18 and

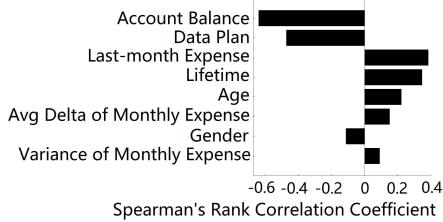


Figure 15: Top features ranked by the Spearman’s rank correlation coefficient.

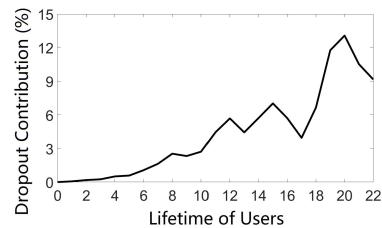


Figure 16: V-Mobile users’ lifetime vs. Dropout Contribution (DC).

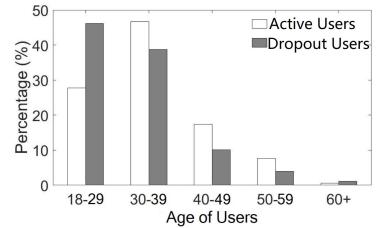


Figure 17: V-Mobile users’ age, for active and dropout users.

ML Model	Precision	Recall	F1 score
Naive Bayes	76.72%	50.54%	0.61
RBF Neural Network	81.52%	91.85%	0.86
SVM	78.79%	96.89%	0.87
Logistic Regression	92.38%	85.02%	0.89
Decision Tree	94.45%	94.10%	0.94
Random Forest	95.52%	94.45%	0.95

Table 3: Customer churn prediction results.

22 months, but their *DC* is as high as 51%. In contrast, for the 8% users who have joined V-Mobile within half a year, their *DC* is only 2%. In today’s MVNO industry, vendors usually pay much attention to attracting new users while oftentimes ignoring existing, and in particular, long-time loyal customers. This practice appears to be unwise according to our observation.

- Females account for a larger fraction (58%) in all V-Mobile customers, but are less likely ($DC=0.45$) to drop out compared to males ($DC=0.55$). Besides, young people are more likely to cancel their services compared to older customers. As shown in Figure 17, 28.4% users are between 18 and 29, while their *DC* is as high as 46.7%. These findings suggest the potential effectiveness of demographic-aware strategies for customer churn mitigation.
- A good indicator of a user’s imminent dropout is her monthly expense being high and her account balance staying low. For example, if a user spends more than \$57.8 in a month while the account balance is lower than \$1.5, then in 72.3% of such cases, she will drop out in the next month. This is likely due to the user’s intent of using up her account balance, which will not be refunded when she drops out.

Customer Churn Prediction. By jointly leveraging our identified features, we consider proactively predicting the customer churn. We apply off-the-shelf machine learning algorithms (Naive Bayes, SVM, Logistic Regression, RBF Neural Network, Decision Tree, and Random Forest) to the 8-dimensional feature vectors (Figure 15). Details of configuring the machine learning algorithms are omitted for brevity. Unlike the data usage prediction where we build a per-user model, here we construct a single model for all users’ monthly records, to capture the common behaviors regarding service cancellation among V-Mobile customers.

We evaluate the prediction accuracy using 10-fold cross validation based on *users*. In each fold, we train a model using the monthly records from 90% of the users, and use the trained model to predict the records of the remaining 10% users. The results shown in Table 3

indicate that the random forest model achieves the best precision (95.52%) and F1 score (0.95), as well a high recall rate (94.45%).

Real-world Churn Mitigation. We collaborate with V-Mobile to conduct a field trial for churn mitigation in late October 2017. Specifically, we first apply the random forest model (trained using the data from 2/2016 to 9/2017) to predict whether each active user (as in 10/2017) will cancel her service in 11/2017. For each test sample, the random forest model outputs a probability indicating the likelihood of dropout. As shown in Figure 18, the dropout probability exhibits a bimodal distribution where nearly 98% (1%) of the users have a churn probability near zero (one). Next, we select the top 1% of the users with the highest dropout probability, and recommend V-Mobile send to each of them via an SMS message a digital gift card containing a small amount (\$1.44) of top-up fee (the gift card can only be activated next month). After conducting this churn mitigation campaign, V-Mobile observed that the churn rate, defined as the fraction of dropout users among all active users in the previous month, decreased significantly from 0.83% in October 2017 to 0.32% in November 2017. This indicates the potential effectiveness of our churn prediction approach ³.

8 INACCURATE BILLING IN MVNO

Recall from §3 that based on our analysis of the BBSDataset, we find that more than 40% of the posts in the V-Mobile BBS concern inaccurate billing or delayed billing state notifications. These problems, if occur, are highly undesired because handling a billing issue requires considerable manual efforts from the customer service team, and problematic billing may also endanger the reputation of an MVNO. We manually examine the posts involving billing complaints in the BBSDataset, and identify two issues:

Issue 1. *Termination of service despite a positive account balance.* A user suddenly loses her access to the cellular data service. Meanwhile, however, she still observes a positive account balance from the V-Mobile’s app or website.

Issue 2. *Unexpectedly high monthly charge experienced by customers.* A user receives a monthly bill that substantially exceeds her expected cost. Typically, the user believes she has not yet used up her monthly data plan but the received bill charges more than the data plan subscription cost.

³To be more rigorous, ideally V-Mobile should also add a control group by distributing gift cards to 1% of users who are randomly chosen, in a different month. Unfortunately V-Mobile was not able to conduct this due to non-technical reasons. Nevertheless, we feel it is still worthwhile to report our experiment and its results.

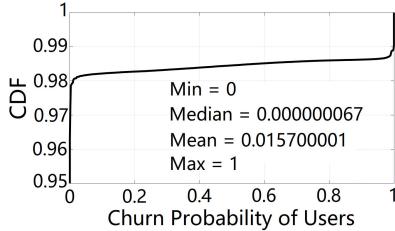


Figure 18: Churn probability distribution of V-Mobile users.

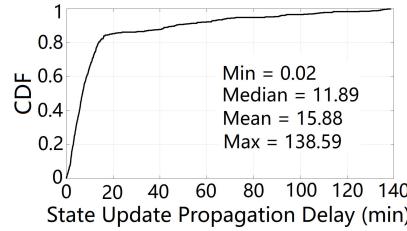


Figure 19: B-Mobile-SUR propagation delay in B-Mobile (D_B).

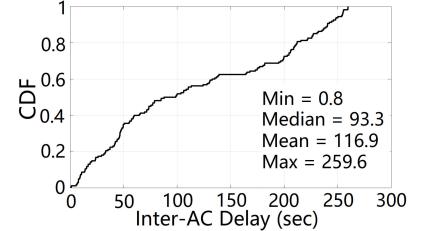


Figure 20: Inter-AC B-Mobile-SUR propagation delay (D_V).

In addition, we learn from V-Mobile that their internal billing sometimes also experiences inconsistencies:

Issue 3. *Unexpectedly high charge experienced by V-Mobile.* V-Mobile notices that for certain customers, the bill that it receives from B-Mobile appears to be higher than what it should be, based on its own data usage records of these customers. This does not affect V-Mobile customers' bills, but causes V-Mobile to lose profit.

To understand the root causes of the above issues, we carefully analyze the SURDataset. Recall from §3 that it captures SURs delivered to V-Mobile's AC from two sources: users' apps (App-SURs) and B-Mobile's AC (B-Mobile-SURs). Our basic analysis methodology is to *correlate* the two types of SURs. For App-SURs, their delivery takes place in almost real time (typically less than 1 second), so their reception time can be largely regarded as a “ground-truth” of the time when the data is actually consumed. However, App-SURs are only used for troubleshooting purposes due to their low user coverage and the possibility of being tampered/spoofed. In contrast, B-Mobile-SURs are used for the actual billing. However, their delivery may experience high delays. To quantify that, as shown in Figure 1, we use D_B to denote the B-Mobile-SUR propagation delay within B-Mobile, and use D_V to denote the delivery latency from B-Mobile to V-Mobile. We can first calculate $D_B + D_V$ by taking the reception difference between a B-Mobile-SUR and its corresponding App-SUR, and then calculate D_V as a delta between the timestamp field in the B-Mobile-SUR (denoting the time when it leaves B-Mobile's AC) and its reception time at V-Mobile's AC. D_B can thus be derived as $(D_B + D_V) - D_V$.

Figure 19 and Figure 20 plot the distributions of D_B and D_V respectively. As shown, D_B ranges from several seconds to 2.3 hours, and averages at nearly 16 minutes. D_V ranges from 1 second to 4.3 minutes, with an average of around 2 minutes.

Given the above finding, we can explain the root causes of the three aforementioned issues (Issue 1 to 3). For Issue 1, we find that the user-perceived account balance, despite being positive, was typically quite low when the user could not access the cellular service. In fact, at the time when the cellular service was terminated by the B-Mobile, the user's account balance had indeed decreased to zero or negative at B-Mobile's AC. However, the B-Mobile-SUR had not yet been propagated to V-Mobile's AC due to the high delay of D_V . Such a billing inconsistency is a unique issue in an MVNO.

For Issue 2, we find that in most cases the user actually had not used up her data plan of the current month. However, at the end of the *previous* month, due to the high propagation delay ($D_B + D_V$), a small portion of the B-Mobile-SURs arrived late at V-Mobile's AC,

so V-Mobile was not able to add them to the previous month's bill. Instead, they were then added to the current month's bill, making it look unreasonable.

For Issue 3, we owe it to the excessive inter-AC delay (D_V). Specifically, after a V-Mobile user switches to a different data plan, the subscription is not always delivered from V-Mobile's AC to B-Mobile's AC in real time when the inter-AC channel is congested. As a result, before the new data plan actually takes effect at B-Mobile, the user's subscription will experience inconsistency between V-Mobile and B-Mobile. This inconsistency will be reflected on the next bill that B-Mobile sends to V-Mobile.

9 RELATED WORK

MVNO Measurements. There exist only a limited number of studies on commercial MVNOs. In 2014, Zarinni *et al.* [36] conducted a first study (based on their claim) of the MVNO performance using controlled experiments, and noticed performance-wise differential treatments between some MVNOs and their base carriers. Also in 2014, Vallina *et al.* [35] studied the relationship between many MVNOs and their base carriers, and pinpointed several potential issues for light and thick MVNOs. In 2016, Schmitt *et al.* noted that a packet may traverse a longer network path in an MVNO network than in the base carrier's network [32]. Recently, Oshiba studied how MVNOs affect bandwidth estimation algorithms [25]. Our study differs from theirs in several aspects: the study scale, the examined topics, and the findings.

Cellular Network Performance has been extensively studied in the literature. Several large-scale measurements include [13–15, 33]. Our study in §5 focus on an MVNO and its base carrier, and provide more up-to-date performance statistics. We also compare our results with those in [14].

Traffic Prediction and MVNO Economics. There exist a plethora of work on (typically short-term) network traffic prediction, such as [9, 27, 37], to name a few. However, it is unclear whether they are suitable for our long-term, large-scale data usage prediction with limited historical data. Researchers have also conducted analytical modeling and formulation on MVNO economics, such as market sharing between MVNOs and MNOs [7], leveraging users' feedback for pricing [22], and QoS-aware scheduling [38]. Compared to these theoretical studies, our reselling optimization takes a more practical and intuitive approach based on statistical modeling and machine learning. We also evaluate its effectiveness using real MVNO customers' data.

Customer Churn Mitigation has been widely studied in industries such as news media [5] and banking [28]. To predict dropout users, various techniques have been put forward [5, 16, 26]. We instead consider churn mitigation in the context of MVNO, which has a unique business model and customer-churn characteristics.

10 CONCLUDING REMARKS

We conduct an in-depth investigation of V-Mobile, a large and representative light MVNO with about 1 million customers. Our findings shed light on various topics that are of the interests of the stakeholders: MVNO customers, the MVNO carrier, and the base carrier. At a high level, our study delivers several take-away messages. First, MVNOs can leverage the unique demographics and usage patterns of their customers to promote and improve their services. Second, AI and big data can significantly boost the revenue and service quality for MVNOs. Third, the cross-layer and cross-entity interactions, in particular those between an MVNO and its base carrier, need to be better handled. Finally, we cannot quantify how generalizable our results are as to other MVNOs because we do not have their data. Qualitatively, we feel that the generalizability of our results depends mostly on the operating model of a different MVNO, rather than which country/area the different MVNO is operated in. We plan to perform deeper explorations in all above aspects in our future work.

ACKNOWLEDGMENTS

We sincerely thank our shepherd Dr. Yunxin Liu and the anonymous reviewers for their valuable feedback. We also appreciate Xiaoyu Piao for her help in data collection. This work is supported in part by the National Key R&D Program of China under grant 2018YFB1004800, and the National Natural Science Foundation of China (NSFC) under grants 61822205, 61432002 and 61632020.

REFERENCES

- [1] Ionut Andone, Konrad Blaszkiewicz, Mark Eibes, Boris Trendafilov, Christian Montag, and Alexander Markowetz. 2016. How Age and Gender Affect Smartphone Usage. In *Proceedings of ACM UbiComp*. Heidelberg, Germany.
- [2] Wenjie Bi, Meili Cai, Mengqi Liu, and Guo Li. 2016. A Big Data Clustering Algorithm for Mitigating the Risk of Customer Churn. *IEEE Transactions on Industrial Informatics* 12, 3 (June 2016), 1270–1281.
- [3] Hocine Bourennane, Dominique King, and Alain Couturier. 2000. Comparison of Kriging with External Drift and Simple Linear Regression for Predicting Soil Horizon Thickness with Different Sample Densities. *Geoderma* 97, 3-4 (September 2000), 255–271.
- [4] David S Broomhead and David Lowe. 1988. *Radial Basis Functions, Multi-variable Functional Interpolation and Adaptive Networks*. Technical Report. Royal Signals and Radar Establishment Malvern (United Kingdom). <https://apps.dtic.mil/dtic/tr/fulltext/u2/a196234.pdf>.
- [5] Kristof Coussement and Dirk Van den Poel. 2008. Churn Prediction in Subscription Services: An Application of Support Vector Machines while Comparing Two Parameter-selection Techniques. *Expert Systems with Applications* 34, 1 (January 2008), 313–327.
- [6] Philip J Davis and Philip Rabinowitz. 1984. *Methods of Numerical Integration*. Academic Press, 51–198.
- [7] Merouane Debbah, Loubna Echabbi, and Chahinez Hamlaoui. 2012. Market Share Analysis between MNO and MVNO under Brand Appeal Based Segmentation. In *Proceedings of IEEE NetGcoP*. Avignon, France.
- [8] Yadolah Dodge. 2008. *Spearman Rank Correlation Coefficient*. Springer New York, New York, NY, 502–505. https://doi.org/10.1007/978-0-387-32833-1_379.
- [9] Huifang Feng and Yantai Shu. 2005. Study on Network Traffic Prediction Techniques. In *Proceedings of IEEE WCNM*. Wuhan, China.
- [10] Trevor Fiatal. 2012. Mobile Virtual Network Operator. US Patent 8,107,921.
- [11] A Gustavo González and M Ángeles Herrador. 2007. A Practical Guide to Analytical Method Validation, Including Measurement Uncertainty and Accuracy Profiles. *Trends in Analytical Chemistry* 26, 3 (January 2007), 227–238.
- [12] Henry Hsu and Peter A Lachenbruch. 2007. Paired T-test. *Wiley Encyclopedia of Clinical Trials* (September 2007), 1–3.
- [13] Junxian Huang, Feng Qian, Alexandre Gerber, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2012. A Close Examination of Performance and Power Characteristics of 4G LTE Networks. In *Proceedings of ACM MobiSys*. Low Wood Bay, Lake District, UK.
- [14] Junxian Huang, Feng Qian, Yihua Guo, Yuanyuan Zhou, Qiang Xu, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2013. An In-depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance. *ACM SIGCOMM Computer Communication Review* 43, 4 (October 2013), 363–374.
- [15] Junxian Huang, Qiang Xu, Birjodh Tiwana, Z Morley Mao, Ming Zhang, and Paramvir Bahl. 2010. Anatomizing Application Performance Differences on Smartphones. In *Proceedings of ACM MobiSys*. San Francisco, CA, USA.
- [16] Shin-Yuan Hung, David C Yen, and Hsiu-Yu Wang. 2006. Applying Data Mining to Telecom Churn Management. *Expert Systems with Applications* 31, 3 (October 2006), 515–524.
- [17] Mathew Ingram. 2006. MVNOs: Phone Companies without the Equipment. <https://www.theglobeandmail.com/technology/mvnos-phone-companies-without-the-equipment/article729367/>.
- [18] Aditya Kapoor. 2017. Churn in the Telecom Industry - Identifying Customers Likely to Churn and How to Retain Them. <https://wp.nyu.edu/adityakapoor/2017/02/17/churn-in-the-telecom-industry-i-identifying-customers-likely-to-churn-and-how-to-retain-them/>.
- [19] Pankaj Lanjukdar and Seapee Bajaj. 2017. Mobile Virtual Network Operator (MVNO) Market-Global Opportunity and Forecasts, 2017-2023. <https://www.alliedmarketresearch.com/mobile-virtual-network-operator-market>.
- [20] Zhenhua Li, Weiwei Wang, Tianyin Xu, Xin Zhong, Xiang-Yang Li, Yunhao Liu, Christo Wilson, and Ben Y Zhao. 2016. Exploring Cross-Application Cellular Traffic Optimization with Baidu TrafficGuard. In *Proceedings of USENIX NSDI*. Santa Clara, CA, USA.
- [21] Raul HC Lopes. 2011. *Kolmogorov-Smirnov Test*. Springer Berlin Heidelberg, Berlin, Heidelberg, 718–720. https://doi.org/10.1007/978-3-642-04898-2_326.
- [22] Charalampos Meidanis, Ioannis Stiakogiannakis, and Maria Papadopoulou. 2014. Pricing for Mobile Virtual Network Operators: The Contribution of U-map. In *Proceedings of IEEE DySPAN*. McLean, VA, USA.
- [23] Anne Morris. 2015. Number of MVNOs Exceeds 1,000 Globally. <https://www.fiercewireless.com/europe/report-number-mvnos-exceeds-1-000-globally>.
- [24] Akihiro Nakao, Ping Du, Yoshiaki Kiriha, Fabrizio Granelli, Anteneh Gebremariam, Tarik Taleb, and Miloud Bagaa. 2017. End-to-end Network Slicing for 5G Mobile Networks. *Journal of Information Processing* 25 (January 2017), 153–163.
- [25] Takashi Oshiba. 2018. Accurate Available Bandwidth Estimation Robust Against Traffic Differentiation in Operational MVNO Networks. In *Proceedings of IEEE ISCC*. Hague, Netherlands.
- [26] Marcin Owczarek. 2010. Churn Models for Prepaid Customers in the Cellular Telecommunication Industry Using Large Data Marts. *Expert Systems with Applications* 37, 6 (June 2010), 4710–4712.
- [27] Hassan Peyravi and Rahul Sehgal. 2017. Link modeling and delay analysis in networks with disruptive links. *ACM Transactions on Sensor Networks* 13, 4 (December 2017), 31.
- [28] Džulijana Popović and Bojana Bašić. 2009. Churn Prediction Model in Retail Banking Using Fuzzy C-means Algorithm. *Informatica* 33, 2 (May 2009), 243–247.
- [29] Feng Qian, Alexandre Gerber, Zhuoqing Morley Mao, Subhabrata Sen, Oliver Spatscheck, and Walter Willinger. 2009. TCP Revisited: A Fresh Look at TCP in the Wild. In *Proceedings of ACM IMC*. Chicago, IL, USA.
- [30] Lenin Ravindranath, Sharad Agarwal, Jitendra Padhye, and Chris Riederer. 2014. Procrastinator: Pacing Mobile Apps' Usage of the Network. In *Proceedings of ACM MobiSys*. Bretton Woods, NH, USA.
- [31] David Rumelhart, Geoffrey Hinton, and Ronald Williams. 1986. Learning Representations by Back-propagating Errors. *Nature* 323, 6088 (October 1986), 533.
- [32] Paul Schmitt, Morgan Vigil, and Elizabeth Belding. 2016. A Study of MVNO Data Paths and Performance. In *Proceedings of PAM*. Heraklion, Crete, Greece.
- [33] Joel Sommers and Paul Barford. 2012. Cell vs. WiFi: On the Performance of Metro Area Mobile Connections. In *Proceedings of ACM IMC*. Boston, MA, USA.
- [34] Statista. 2018. Size of the Global MVNO Market from 2012 to 2022. <https://www.statista.com/statistics/671623/global-mvno-market-size/>.
- [35] Nareso Vallina-Rodriguez, Srikanth Sundaresan, Christian Kreibich, Nicholas Weaver, and Vern Paxson. 2015. Beyond the Radio: Illuminating the Higher Layers of Mobile Networks. In *Proceedings of ACM MobiSys*. Florence, Italy.
- [36] Fatima Zarinni, Ayon Chakraborty, Vyas Sekar, Samir R Das, and Phillipa Gill. 2014. A First Look at Performance in Mobile Virtual Network Operators. In *Proceedings of ACM IMC*. Vancouver, BC, Canada.
- [37] Lizong Zhang, Nawaf R Alharbe, Guangchun Luo, Zhiyuan Yao, and Ying Li. 2018. A hybrid forecasting framework based on support vector regression with a modified genetic algorithm and a random forest for traffic flow prediction. *Tsinghua Science and Technology* 23, 4 (August 2018), 479–492.
- [38] Tianxiao Zhang, Huasen Wu, Xin Liu, and Longbo Huang. 2016. Learning-aided Scheduling for Mobile Virtual Network Operators with QoS Constraints. In *Proceedings of IEEE WiOpt*. Tempe, AZ, USA.