

MNP Inside Out: A Game Theory Assisted Machine Learning Model to Detect Subscriber Churn Behaviors under China's Mobile Number Portability Policy

Ye Ouyang, Ph.D
Telecom Artificial
Intelligence Lab, AsiaInfo
Technologies
Beijing, China
ye.ouyang@asiainfo.com

Aidong Yang, Ph.D
Telecom Artificial
Intelligence Lab, AsiaInfo
Technologies
Beijing, China
yangad@asiainfo.com

Shuming Zeng, Ph.D
Telecom Artificial
Intelligence Lab, AsiaInfo
Technologies
Beijing, China
zengsm3@asiainfo.com

Fanyu Meng
Electronical Engineering &
Computer Science Dept.,
University of California, Berkeley
Berkeley, United States
fy.meng@berkeley.edu

Abstract—Mobile number portability (MNP) policy highlights the problem of subscriber churn and thus enhance the liquidity and competition of the telecommunication market. China will implement MNP policy on November 1st, 2019 after 9 years of trials. This paper timely proposes, validates, and productizes a game theory assisted machine learning scheme to help the mobile network operator (MNO) in China make a strategy to proactively cope with their competitors in the same MNP market. The scheme further develops a set of MNP tactics for the MNO to detect user churn behaviors and to remedy the users with appropriate treatments. Experimental results demonstrate that the scheme can guide the MNOs to make targeted MNP strategy and precisely identify the “abnormal” subscribers who tend to churn out and potential new subscribers who may churn in. The scheme has been successfully implemented in production to greatly improve the marketing efficiency and user satisfaction in terms of an approximately 50% reduction of user churn for a tier-1 MNO in China.

Keywords—Subscriber Churn, Mobile Number Portability, Game Theory, Machine Learning;

I. INTRODUCTION

Mobile number is naturally considered to be one of the most important personal assets in today's information society. Conventionally, even if the service is unsatisfactory, mobile subscribers are reluctant to change operators due to the inconvenience to change their phone numbers [1-3]. To reduce the switching costs and gain competitions in the telecom markets, an increasing number of countries/regions including Japan [4], United States [5] and Europe [6], have been released mobile number portability (MNP) or wireless local number portability (WLNP) policy which enables subscribers to retain their existing phone numbers when switching operators. WLNP is defined by Federal Communication Commissions (FCC) in [7], allowing wireless consumers to switch from one mobile network operator (MNO) to another one within the same general metropolitan area and keep their existing phone number. Occasionally, it also allows consumers to keep their phone number while switching from a wireline plan to a wireless plan. Generally, MNP enables mobile users to retain their mobile telephone numbers when changing from one MNO to another regardless of voluntary or involuntary decisions from the users. The availability of MNP enhances the liquidity of the telecom markets [5], which introduces a new challenge for MNOs to retain the existing users and attract new users.

Many studies have investigated the influence of MNP on both sides: subscribers and operators. MNP is expected to

promote competition among operators and bring benefits to subscribers. Additionally, younger subscribers have a higher acceptance of MNP [1]. Handling fee also heavily affects the willingness towards MNP: about 17% of the mobile users in Japan would change the operators and upgrade their handsets, but the percentage drops to 5% with a handling fee [4]. Before the introduction of the MNP, the operators generally prefer charging a higher price to the “locked-in” subscribers rather than appealing to new subscribers [8], due to the significant switching costs. MNP services remove such switching barriers and further increase the complexity of the competitive environment of the telecom market.

China plans to implement the MNP policy starting from November 1st, 2019, after 9 years of trials in 5 provinces. Unfortunately, China's three tier-1 MNOs all reported declined revenue growth in their interim financial report recently (China Mobile declined by 0.6%, China Telecom declined by 0.5%, and China Unicom declined 1.1%) [9-11]. In the recession, MNP introduces another uncertain factor to either deteriorate or overturn the entire telecom sector in China. Three mobile operators would act more strategically under the MNP policy rather than just fighting a price war endlessly in the past. A clear sign of such change is that all of the three Chinese mobile operators have announced the end of unlimited plans starting from August 2019 [12].

Under such situation, the MNOs cannot adapt to the ever-changing environment only with a fixed marketing strategy in the MNP era. Even after the MNOs applied some flexible strategies that have taken into the account of the existence of MNP, the behaviors of the individual consumer are still very difficult to identify and predict. MNP-churning may also be affected by a numerous factor including rate plan, device quality to network experience, etc., and thus it's difficult for the operators to precisely lockdown and remedy the targeted users with an early sign to churn. Therefore, in this paper, we utilize a game theory assisted machine learning methodology to help MNOs in two folds:

- To create targeted MNP marketing strategy for the MNO to cope with the strategy portfolio from their competitors in the same MNP market;
- To develop a set of precise MNP tactics for the MNOs: (1) to identify and detect user churn behaviors, (2) to remedy the subscribers who present an early sign of MNP churn, and (3) to lockdown potential users who may port in from competitors.

In summary, the main contribution of this work including:

- For the upcoming MNP policy in China, we, for the first time, propose, develop, and utilize game theory to address the problem of MNOs' optimal MNP marketing strategy in an open and free competition market to maximize their interests.
- The interest expectations of 27 possible marketing game combinations of the three MNOs each with three possible market decisions are comprehensively analyzed to help MNOs dynamically achieving optimal decision-making in the volatile market.
- We reduce the operation strategy of MNOs into the prediction of subscriber transfer behavior, which is precisely a three-classification problem. By using our game theory assisted machine learning model and combining the big data of the three major MNOs in China, we can construct a complete probability matrix for any subscriber, and thus predict the transfer tendency of such subscriber.
- After comparing several prevalent shallow statistical models and deep neural networks, we find that the stacking neural networks (SNN) model presents a superior performance. On the extremely unbalanced real data set (the ratio of positive and negative sample is 1:1550), the F1-score of the SNN model reaches approximately 37%. This model can improve marketing efficiency and reduce nearly 50% customer churn for a tier-1 MNO in China.
- Through the analysis of the SNN model, we find that subscribers with transfer tendency have the characteristics of longer average call time and higher call frequency. They may have higher requirements for service quality for data and voice calls. Therefore, improving the data and voice service quality may effectively retain those subscribers.

This paper is organized as follows: Section II summarizes related works in user churning and behavior analysis with a MNP setting. Section III describes the proposed scheme of the game theory of MNP for China's telecom sector. Section IV presents the algorithm of user churn prediction based upon the MNO's marketing strategy deducted from the game theory of MNP. Section V applies computer simulations to evaluate the performance of the proposed methodologies. Finally, conclusions and potential future works are discussed in Section VI.

II. RELATED WORK

Churn prediction is an active field of research that has attracted attentions. Since the cost of attracting new customers is higher about five to six times expensive than that of retaining existing customers [13], detecting the potential churners before they leave is crucial to the MNOs, especially under the policy of MNP. In this section, we provide a brief overview of the relevant research progress.

A. Factors that influence the subscribers churn

In a highly saturated competitive telecommunications market, subscribers will migrate from one service operator to another based on any factors they deem important. That is to say, the cause to such churning may greatly varies. Authors in [14] used a longitudinal survey in Spain to analyze the factors that drive consumers to switch operators. They found that the satisfied consumers and those with mobile services bundled

with fixed services have low intension to leave the existing operator, and the more complex the bundle, the lower the switching rate. Authors in [5] defined three operational factors to investigate the motivations of subscribers' switching behaviors with the introduction of the MNP in the US. These factors are service, including call quality, value-added services and customer support; switching barriers, including switching cost and opportunity cost; prices, including subscriber lock-in, pricing structure, pricing scheme and additional service fee. The result of their studies indicate that call quality may not a significant factor because it has reached a level that customers cannot feel any differences. The finding is consistent with the study on German's telecommunication market [15]. However, in the northern region of China, service quality has a significant impact on subscribers' churn behavior [16]. Another study performed in US telecommunication market shows that age and education level have a direct influence on subscribers' switching attention, while gender does not [17]. In European Union countries, the switching time and charges are believed to have a significant impact on switching behavior [18].

Except for the customer attributes and the service attributes, the social network information is another key factor that influences subscribers' churning. Customers are not isolated; their behavior will be affected by the people around them [19]. Based on this idea, authors in [20] explored the role of social networks in customer retention and found an increase of 80% in churning rate after a neighbor churn. Expanded upon this finding, authors in [19] added several social network features, including the number of churn neighbors, seconds calling churners, the number of non-churn neighbors, seconds calling non-churners, out-of-network neighbor and seconds calling out-of-network to the customer information to build the churn prediction model, and achieved an improved accuracy. Additionally, a discrete choice experiment performed in Portuguese confirmed the importance of network effects in telecommunication markets [21]. Subscribers prefer to choose the same provider with their family and friends, which is named as the calling club network effects or the local network effects.

B. Customer churn prediction models

In a saturated telecommunication market, attracting new customers is more difficult than retaining current customers. It has been estimated that the average churn rate was 2.2% per month for mobile customers [13]. In order to reduce the churn rate, MNOs need to find out the determinants of customer churn and identify the potential churners. These two points can be achieved by building churn prediction models.

The churn prediction model is a data mining technique that uses historical data to classify whether the current customers are potential churners or not. To establish such churn prediction model, one should firstly collect a dataset that contains suitable customer's features and labels (churner or non-churner), and then select an algorithm to map the features to the labels. After evaluating the performance of the model in an independent dataset, we can use the model to predict the churn probability of a customer.

In the domain of telecommunications, many algorithms are used to construct such churn prediction models, including decision tree (DT), rule-based (RB) algorithms, Naïve Bayes (NB), support vector machine (SVM) and neural networks (NN) [13]. Authors in [13] argued that customer attributes are incomplete for prepaid customers. Thus, they incorporated

social network information in the prediction model and performed a social network analysis. This method has improved the accuracy, timeliness, and profitability of the model. Authors in [22] treated customers' churn behavior as a sequence of probable transitions. They used Markov chain to predict churning and retention rate in Nigeria's mobile telecommunication market. To obtain a better and more accurate prediction, authors in [23] proposed a hybrid learning model that combined tree method and genetic programming to derive the classification rules. Besides, they also analyzed the community effect of churn using game theory techniques.

However, most of the previous studies focused on predicting customers' churning behaviors without MNP policy. The introduction of the MNP policy not only makes the competition between operators more complex but also influences customers' switching tendency. In this paper, we use game theory to first analyze and derive the MNP marketing strategies for the MNOs and then propose a game theory assisted stacking neural networks model to classify customers as churners or not for the MNO to remedy under China's MNP policy.

III. THE MNP GAME THEORY FOR CHINA'S MNOs

When more than one company offer competitive products to the same market, their marketing behaviors will influence each other. Game theory provides a natural analytical mathematical tool to analyze the strategic interaction between rational decision-makers, and it applies to a wide range of fields such as economics, politics, logic and computer science [24]. In an oligopoly market, there are several classical game models according to different decision-making conditions. Cournot model assumes that two companies provide the same product, and each company makes its own supply decision without observing the output of other companies. The Stackelberg leadership model considers that at least one company in the market can choose a specific level of production before others, while Bertrand model describes the competition between companies that supply differentiated products. In China, China mobile (CM), China Unicom (CU) and China Telecom (CT) are the three companies that have dominated the telecom markets. The existence of MNP will promote the free flow of subscribers between them. To avoid disclosing customers' privacy, we use A, B and C below to refer to the three major operators in China.

A. Game Theory-based Algorithm

Game theory has achieved successful applications in sales price war[25], transaction decision-making[26], fraud punishment[27] and so on[28]. This section presents a game theory-based strategy evaluation method for MNOs' attack-defense marketing behaviors and applies it to the marketing decision-making of MNOs' under MNP in China.

Figure 1 presents the workflow of the proposed method of game theory-based attack-defense marketing strategy, where it consists of the parts of feature extraction of core business data, the comprehensive scoring of business value, the market game with competitors, and the output of optimal decision strategy. To maximize an MNO's benefits, a game theory-based decision-making algorithm is designed in the following table of Algorithm 1 for the MNP in China. The proposed algorithm dynamically outputs the current optimal marketing decision, *i.e.* the maximal switch-in (Max_{IN}), the minimal

switch-out (Min_{OUT}), or both the maximal switch-in and the minimal switch-out ($Max_{IN/OUT}$), depended on the current marketing objectives and market conditions.

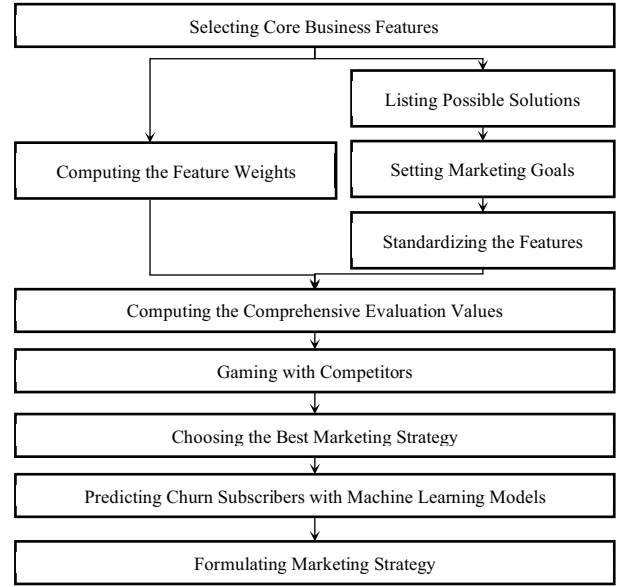


Fig. 1. The workflow of the game theory-based strategy evaluation method for attack-defense marketing decision-making.

Algorithm 1 shows the flowchart of game theory of MNP for the operators in China. Firstly, the important features of MNO's business published and/or private data are selected as inputs $X_i, i = 1, 2, \dots, n$, and then its feature importance vector β_i between the input X_i and the input $X_j, j = 1, 2, \dots, n$, is calculated by

$$\beta_{i,j} = \begin{cases} 1, & \text{if } X_i \text{ is more important than } X_j, \\ 0, & \text{if } X_i \text{ is less important than } X_j, \end{cases} \quad (1)$$

where $\beta_{i,j}$ is the j^{th} element of importance vector β_i . The final score α_i and its weights ω_i can be obtained by

$$\alpha_i = \sum_{j=1}^n \beta_{i,j}, \quad (2)$$

$$\omega_i = \frac{\alpha_i}{\sum_{i=1}^n \alpha_i}. \quad (3)$$

Algorithm 1: The proposed Game Theory of MNP for MNOs in China

Select core features $X_i, i = 1, 2, \dots, n$
 Compute parameters $\beta_{i,j}, \alpha_i$, and ω_i via equation (1), (2), (3), respectively
For all $v \in \{Max_{IN}, Min_{OUT}, Max_{IN/OUT}\}$ **do**
 Set the operation objective values \tilde{x}_i^v for the i^{th} feature and v^{th} marketing strategy, $i = 1, 2, \dots, n$
 Compute the score $\tilde{s}_i^v, i = 1, 2, \dots, n$, via table I
 Compute comprehensive evaluation value $Z_v = \sum_{i=1}^n \omega_i \tilde{s}_i^v$
End
For all $v \in \{Max_{IN}, Min_{OUT}, Max_{IN/OUT}\}$ **do**
 For all $o \in \{A, B, C\}$ **do**
 Compute the game score $S_{v(o)}^{o(m)}$ via (5)
 End
End
 Compute the mathematical expectation $E\{S_{v(o)}^{o(m)}\}$ for all $v \in \{Max_{IN}, Min_{OUT}, Max_{IN/OUT}\}$ and all $o \in \{A, B, C\}$
 Output the optimal i of $\max [E\{S_{v(o)}^{o(m)}\}]$ for all $o \in \{A, B, C\}$.

Assuming that the marketing operation objectives for MNP in the i^{th} feature and v^{th} marketing strategy as \tilde{x}_i^v , where v denotes Max_{IN} , Min_{OUT} or $\text{Max}_{\text{IN/OUT}}$, and $i = 1, 2, 3, \dots, n$, and the scoring benchmark is defined by Table I. By substituting the \tilde{x}_i^v to the scoring benchmark of Table I, the score \tilde{s}_i^v of \tilde{x}_i^v can be obtained. Finally, the comprehensive value of marketing strategy v is evaluated by

$$\mathbf{Z}_v = \sum_{i=1}^n \omega_i \tilde{s}_i^v. \quad (4)$$

TABLE I. THE SCORING BENCHMARK

Feature	Scoring									
	5		4		3		2		1	
	\geq	$<$	\geq	$<$	\geq	$<$	\geq	$<$	\geq	$<$
\tilde{x}_1	S ₁₀	S ₁₁	S ₁₂	S ₁₃	S ₁₄	S ₁₅	S ₁₆	S ₁₇	S ₁₈	S ₁₉
\tilde{x}_2	S ₂₀	S ₂₁	S ₂₂	S ₂₃	S ₂₄	S ₂₅	S ₂₆	S ₂₇	S ₂₈	S ₂₉
\tilde{x}_3	S ₃₀	S ₃₁	S ₃₂	S ₃₃	S ₃₄	S ₃₅	S ₃₆	S ₃₇	S ₃₈	S ₃₉
\vdots										
\tilde{x}_n	S _{n0}	S _{n1}	S _{n2}	S _{n3}	S _{n4}	S _{n5}	S _{n6}	S _{n7}	S _{n8}	S _{n9}

Since each of three telecom operators in China, *i.e.* A, B and C, has three possible marketing strategies, *i.e.* Max_{IN} , Min_{OUT} , and $\text{Max}_{\text{IN/OUT}}$, in the game among the economic entities in actual society, and any two of three operators will game each other, we will have 27 possible permutations and combinations for the game theory-based algorithm. To get the optimal solution for each game player, all the possible game theory results are calculated by

$$\mathbf{S}_{v(i)}^{o(m)} = 2 \cdot \mathbf{Z}_{v(i)}^{o(m)} - \mathbf{Z}_{v(j)}^{o(n)} - \mathbf{Z}_{v(k)}^{o(l)}, \quad (5)$$

where $\mathbf{S}_{v(i)}^{o(m)}$ is the final score of the operator $o(m)$ selecting marketing strategy $v(i)$ after game with the other players, \mathbf{Z}_v^o denotes the comprehensive value in Eq.(4) of the o^{th} operator selecting the v^{th} marketing strategy, and $o \in \{A, B, C\}$, $v \in \{\text{Max}_{\text{IN}}, \text{Min}_{\text{OUT}}, \text{Max}_{\text{IN/OUT}}\}$, $i, j, k, m, n, l = 1, 2, 3$, and $m \neq n \neq l$.

By computing the expectation value of $\mathbf{E}\{\mathbf{S}_{v(i)}^{o(m)}\}$ for all the v and o , where $\mathbf{E}\{\cdot\}$ is the mathematical expectation, the optimal marketing strategy $v(i)$ of the operator $o(m)$ is obtained by

$$v\{i\} \leftarrow \max[\mathbf{E}\{\mathbf{S}_{v(i)}^{o(m)}\}], \quad (6)$$

where all $v \in \{\text{Max}_{\text{IN}}, \text{Min}_{\text{OUT}}, \text{Max}_{\text{IN/OUT}}\}$ and all $o \in \{A, B, C\}$.

IV. PREDICTING SUBSCRIBER CHURN ALGORITHMS

In China's telecommunications market, subscribers can only select one of the three major telecom operators, *i.e.* A, B, and C. Once the MNP policy is fully implemented, subscribers are allowed to switch to any one of the telecom operators for free. To maximize the benefit of a telecom operator, the subscriber churn rate must be predicted in advance.

To predict the churning of a subscriber, we need to compute the transition probability. Assuming that a subscriber

i is to switch from a provider m to a provider n , the transition decision p_{mn}^i can be expressed as

$$p_{m,n}^i = \sum_k \delta_{m,k} p_{k,n}^i, \quad (7)$$

where $\delta_{m,k}$ is the Dirac delta function. Eq.(7) means that the subscriber i can only switch from m to the other providers. According to Eq.(7), we can construct the transition matrix. The whole transition matrix M is defined as Table II, where operators A, B and C are denoted as 1, 2 and 3, respectively. For any subscriber, his/her choice is determined and unique, hence the transition decision $P_{m,n}$ for any subscriber can be written by

$$P_{m,n} = \begin{cases} 1, & \text{if the subscriber switches from } m \text{ to } n, \\ 0, & \text{others.} \end{cases} \quad (8)$$

From Eq.(8) and Table II, one can see that there is only one element in the transition matrix which is 1 and all the other values are zeros.

TABLE II. THE TRANSITION MATRIX OF OPERATORS A, B, C

Operator	A	B	C
A	P ₁₁	P ₁₂	P ₁₃
B	P ₂₁	P ₂₂	P ₂₃
C	P ₃₁	P ₃₂	P ₃₃

Suppose that there are N subscribers in the telecom market. At time $t+1$ (unit: month), the total number of subscribers who switch into the provider m is given by

$$\text{In}(m, t+1) = \sum_{i=1}^N \sum_{n \neq m} p_{mn}^i(t), \quad (9)$$

while the total number of subscribers who switch out from the provider m is

$$\text{Out}(m, t+1) = \sum_{i=1}^N \sum_{n \neq m} p_{mn}^i(t). \quad (10)$$

To attract subscribers to switch-in or reduce switch-out, we must know the transition matrix M^i ex-ante for each subscriber i . According to the previous studies [1-4, 8, 21, 24], The subscriber churn decision is influenced by lots of factors, including education, the bundle status, gender, value-added services, switching cost, call quality, monthly fee and call time. Hence the transition decision is related to such attributes,

$$p_{mn}^i(t) = p_{mn}^i(t, D^i), \quad (11)$$

where D^i is the features of subscriber i .

One challenging factor is to find a mapping from (t, D^i) to p_{mn}^i in Eq. (11). Recently, several techniques have been developed to predict p_{mn}^i , such as Markov Chain modeling [22], Logistic Regression (LR) [5], social network analysis [19], and some other statistical analysis methods [29,14]. However, these studies either lack of interpretability or do not focus on the MNP policy, especially China's MNP policy.

Also, subscribers' behaviors depend on both the current state and the history. To avoid dealing with timing problems, we add the historical information to the current state and create some new features including the average value of the call time

in the last three months, changes in call time and changes in call frequency.

Considering that subscribers have three choices, retain, switch to the other two operators, the predicting problem is a three-class problem. In the paper, the stacking model is used for superior performance and interpretability. Besides, several other models, such as Random Forest (RF) model, Recurrent Neural Network (RNN) [30], DT, Gradient Boosting Decision Tree (GBDT), XGBoost [32], ExtraTree and NN, are also constructed for comparison. The performance of these models are characterized by accuracy, recall, F1-score and the area (AUC) under the receiver operating characteristic (ROC) curve. Since the number of subscribers that switch out is relatively small, resampling and de-sampling techniques are used to solve the problem of data imbalance.

After constructing the machine learning-based churn prediction model, we can compute the expected gain of each strategy as follows,

$$\max\{\ln(m, t+1)\} = \max\left\{\sum_{i=1}^N \sum_{n \neq m} p_{n,m}^{i*}(t, D^i)\right\}, \quad (12)$$

$$\min\{\text{Out}(m, t+1)\} = \min\left\{\sum_{i=1}^N \sum_{n \neq m} p_{m,n}^{i*}(t, D^i)\right\}, \quad (13)$$

$$\max\left\{\frac{\ln(m, t+1)}{\text{Out}(m, t+1)}\right\} = \max\left\{\frac{\sum_{i=1}^N \sum_{n \neq m} p_{n,m}^{i*}(t, D^i)}{\sum_{i=1}^N \sum_{n \neq m} p_{m,n}^{i*}(t, D^i)}\right\}, \quad (14)$$

where p^* is the transition decision predicted by the ML model.

In this method, whatever strategy the game theory obtained, the first step is to get the transition decision matrix through ML model. For strategy Max_{IN} , Min_{OUT} , and $\text{Max}_{IN/OUT}$, we can use the same ML model to predict churn tendency. After recognizing the churners, the providers need to perform corresponding marketing activities according to the strategy obtained by the game theory. In turn, the results of the ML model can also affect the process decision-making based on game theory.

V. SIMULATIONS AND DISCUSSIONS

To thoroughly investigate the performance of the game-theory-based machine learning methods and validate the effectiveness of the theoretical analyses given in the previous section, we employ Monte Carlo simulations to analyze the game theory for dynamic marketing decision of operator A in an actual market environment as an example, and then evaluate the algorithm performance of the proposed subscriber churn prediction models based on the marketing decision.

A. The Analysis of Game Theory

In this section, we use the MNP case as an example to show how operator A achieve the optimal marketing strategy using.

As shown in Table III, the core feature, $X_i, i = 1, 2, 3 \dots, 8$, of monthly telecom business operation data is extracted and input to the game theory Algorithm 1. The features can be changed resonantly at any time according to the needs of business and marketing operation. From Eq. (1), (2) and (3),

TABLE III. THE FEATURE EXTRACTION OF TELECOM CORE MONTHLY BUSINESS DATA

Operator	Feature							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
A	-5.88%	-70.4	58.50%	-29.7	68.20%	4701.9	3.093	162.5
B	-0.79%	-14.8	13.90%	1.3	5.60%	389.5	0.467	38.8
C	-3.91%	-15.3	27.60%	1.3	26.20%	1807.2	1.391	76.6

* X_1 - The Ratio of Monthly Income Change; X_2 - The Change of Monthly New Subscribers (in ten thousand); X_3 - The Change Ratio of Monthly New Subscribers; X_4 - The Change of Monthly Retained Subscribers (in ten thousand); X_5 - The Change Ratio of Monthly Retained Subscribers; X_6 - The Number of Monthly Retained Subscribers (in ten thousand); X_7 - Monthly Income (in billions); X_8 - The Number of New Subscribers (in ten thousands);

the parameters of β_i, α_i , and ω_i for the feature $X_i, i = 1, 2, 3, \dots, 5$ are calculated in Table IV, where $X_1 \sim X_5$ are the selected important features for the MNP marketing of operator A in current month, and X_A is an extensible feature which can be adjusted by marketing operator.

TABLE IV. THE FEATURE PAIRWISE COMPARISON MATRIX

Feature	Feature						Score	Weight
	X_1	X_2	X_3	X_4	X_5	X_A	α_i	ω_i
X_1	0	1	1	0	0	$1 - \beta_{A1}$	2	0.2
X_2	0	0	1	0	0	$1 - \beta_{A2}$	1	0.1
X_3	0	0	0	1	0	$1 - \beta_{A3}$	1	0.1
X_4	1	1	0	0	0	$1 - \beta_{A4}$	2	0.2
X_5	1	1	1	1	0	$1 - \beta_{A5}$	4	0.4
* X_A	β_{A1}	β_{A2}	β_{A3}	β_{A4}	β_{A5}	0	$\sum \beta_{Aj}$	$\frac{\alpha_A}{\sum_{i=1}^5 \alpha_i}$

* X_A - the extensible feature.

TABLE V. THE MARKETING OBJECTIVES OF THE SELECTED CORE FEATURES IN DIFFERENT MARKETING STRATEGIES FOR OPERATOR A

Marketing Strategy v	Feature					
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4	\tilde{x}_5	\tilde{x}_A
Max_{IN}	2%	11	57%	-5	66%	NA
Min_{OUT}	-4%	20	61%	-8	64%	NA
$\text{Max}_{IN/OUT}$	5%	-14	54%	-1	69%	NA

To achieve the marketing objectives while gaming with the other competitors, the marketing operators need to set their marketing objective values \tilde{x}_i^v for the three marketing operation strategies $v(i), i = 1, 2, 3$ as shown in Table V, in which \tilde{x}_A are the optional features that can be used as marketing needed. By substituting the marketing values \tilde{x}_i^v in Table V to the scoring benchmark Table VI, the comprehensive evaluation values Z_v in (4) for the marketing strategy Max_{IN} , Min_{OUT} , and $\text{Max}_{IN/OUT}$ are obtained by Table VII.

TABLE VI. THE SCORING BENCHMARK FOR OPERATOR A

Feature	Scoring							
	5 =>	4 =>	3 <	2 =>	1 <	0 =>	-1 <	-2 <
\tilde{x}_1	5	3	5	0	3	-3	0	-3
\tilde{x}_2	15	10	15	5	10	-10	0	-10
\tilde{x}_3	58	57	58	56	57	55	56	55
\tilde{x}_4	5	3	5	0	3	-5	0	-5
\tilde{x}_5	70	68	70	66	68	64	66	64

TABLE VII. COMPREHENSIVE EVALUATION VALUE OF OPERATOR A IN DIFFERENT MARKETING STRATEGIES

Marketing Strategy v	X_1	X_2	X_3	X_4	X_5	X_A	CEV Z_v
	0.20	0.10	0.10	0.20	0.40	0.00	
Max_{IN}	3	4	4	2	3	NA	3.00
Min_{OUT}	1	5	5	1	2	NA	2.20
$\text{Max}_{IN/OUT}$	5	1	1	2	4	NA	3.20

* CEV - Comprehensive Evaluation Value

B. Nash Equilibrium Marketing Strategy

When a telecom operator is to compose a marketing strategy, *i.e.* Max_{IN} , Min_{OUT} , and $Max_{IN/OUT}$, the other two telecom operators' choices also have a great impact on its strategy decision. Usually, the strategies chosen by other operators are unknown before marketing. Therefore, the operator needs to make the optimal decision without knowing competitor information to maximize its own interests.

The Nash Equilibrium (NE) can be used to solve this problem. NE is one of game theory-based methods that provide an exact analytical method to simulate the strategy interaction among all the rational decision-makers. Through game simulation with the other operators for each of the three marketing strategies, we can obtain the expected marketing score once a certain strategy is chosen, and then select the best strategy that benefits the operator. Since each operator has three marketing strategies, *i.e.* Max_{IN} , Min_{OUT} , and $Max_{IN/OUT}$, we have 27 competing outcomes as shown in Table IX in appendix.

The Table IX is the Nash equilibrium simulation results for all the three operators when they choose different three marketing strategies. Similarly, based on the prior information of the other market competitors, *i.e.* known, partly known, or completely unknown, the expected benefit returns in different marketing strategies for each operator can be obtained. As an example, the Table VIII shows the expected benefit returns for each operator without knowing any competitor's strategy v .

TABLE VIII. BENEFIT RETURNS FOR EACH OPERATOR IN DIFFERENT MARKETING STRATEGIES

Marketing Strategy v	A			B			C		
	$*Y_1$	$*Y_2$	$*Y_3$	$*Y_1$	$*Y_2$	$*Y_3$	$*Y_1$	$*Y_2$	$*Y_3$
Max_{IN}	0.17	44	56	-0.1	33	56	0.27	44	44
Min_{OUT}	1.37	0	100	-2.5	100	0	-1.33	100	0
$Max_{IN/OUT}$	0.77	22	78	0.7	0	89	0.67	11	78

*Y₁- Expected Revenue (in billion); *Y₂- Negative Income Probability (in %); *Y₃- Positive Income Probability (in %);

According to the results of the table above, the operators would select the marketing strategies which should be adopted in the current market environment to maximize their own interests. For example, operator A should choose Min_{OUT} marketing strategy in current market environment and is expected to get a 1.37 billion revenue with 100% of positive income probability. It is important to note that when the competitor's marketing strategy is known or partially known, the statistical data in Table V would change, and thus the optimal marketing strategy will be affected as well.

C. Subscriber Churn Prediction

According to the analysis of game theory, the best strategy for operator A is $Min(Out(A, t))$. For any subscriber, we only need to classify whether he/she will leave or retain, and there is no need to determine which operator he/she will switch to. Therefore, the three-classification problem degenerates into a binary-classification problem.

The dataset is collected from different moths, and we divide it into two independent groups, group I and group II. The attributes of subscriber in group I is collected from moth t_1-2 to t_1 , and the label is obtained at t_1+1 . While the attributes in group II is collected from moth t_2-2 to t_2 , and the label is obtained at t_2+1 . Each group contains 31 million subscribers with 216 attributes. The subscriber is labeled 1 if he switched

out at moth $t+1$ ($t = t_1, t_2$), else labeled 0. The data in group I is used to train and to validate the ML model, while group II is used for testing. Since there are only 20,000 positive samples in group I, we use all of them and randomly select 60,000 negative samples to form the training and validation dataset with 80,000 samples.

The stacking neural network (SNN) we use in this work is sketched in Fig.2. This model contains two layers. The first layer contains five sub-models, RF, ExtraTree, GBDT, XGBoost, and NN, while the second layer only contains a NN sub-model. During the training process, the sub-models in the first layer are trained and validated using five-fold cross-validation, and the output of these model are used as inputs of the second layer. Before training, the positive samples in the training set are resampled so that the ratio of the positive and negative samples is 1:1, while the ratio is kept as 1:3 in the validation set. The output of the NN in the second layer will be the output of SNN. The performance of the model is characterized by AUC and shown in Fig.3. For comparison, the results of DT, RNN, RF, ExtraTree, GBDT, XGBoost and NN models are also presented.

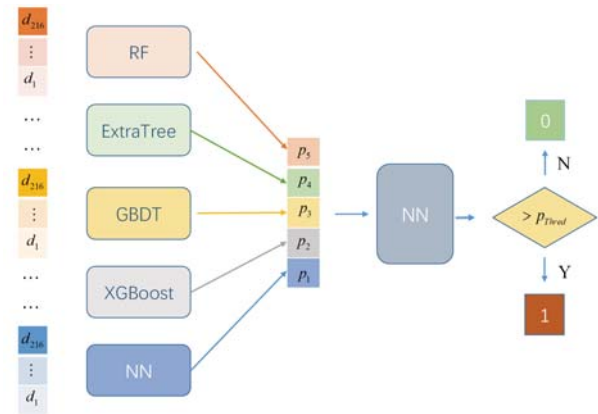


Fig. 2. The structure of the stacking neural networks model

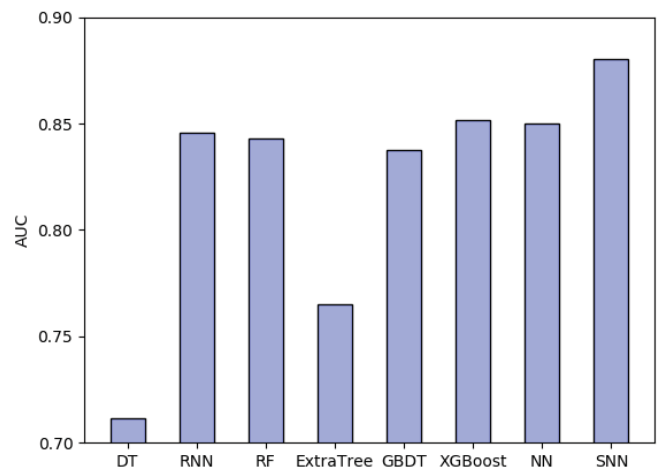


Fig. 3. AUC of different models on test set

As shown in Fig.3, The AUC of the SNN model is 0.8801, which is comparably larger than any other models. Since the AUC of the DT model is considerably smaller, we do not include it as a sub-model for the SNN. We also explored the possibility of using RNN to predict churn behaviors. We construct a long short-term memory (LSTM) [31] binary-classification model for comparison. As shown in Fig.3, the

AUC of the best RNN model is 0.8456, on par with XGBoost and NN. However, due to the limited length of our available data, we only use 4 month sequences as in the input. Additionally, we can only model the problem as a binary classification instead of the traditional way of predicting time to next churn, which may harm the performance of the RNN model. The RNN exhibits a noticeable instability and a tendency towards converging to predicting only true or only false even with heavy normalization. Due to these factors, we also do not include RNN as a sub-model for the SNN.

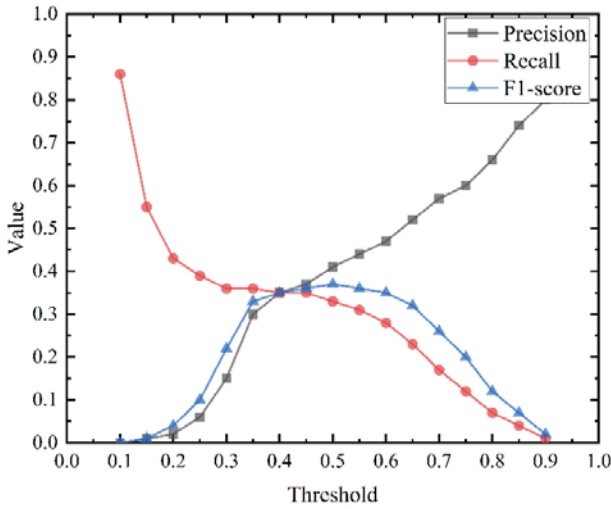


Fig. 4. The classification results of the SNN model at different thresholds on test set

In previous studies, the AUC of the logistic regression, DT, Bagging Tree and RF models are 0.8217, 0.8088, 0.8284 [33] and 0.7350 [34], respectively, which are all inferior compared to our proposed method. Since the SNN has the best performance, in the following discussions, we focus on the analysis of the results of the SNN. The test set (group II) contains 21,582 negative samples and nearly 31 million negative samples, the classification precision, recall and F1-score of the SNN at different thresholds in this data set are shown in Fig.4.

As the threshold decreases, the accuracy decreases and recall value increases. When the threshold is 0.55, the SNN achieves its best F1-score of 0.37. Compared to the F1-score of 0.68 on the validation set, the performance of the model in test set seems noticeably worse. This is due to the heavy data imbalance in the test set. In validation set, the ratio of the positive samples and negative samples is 1:3, while in test set, the ratio is 1:1550. Compared with traditional marketing strategies, SNN-based strategy would be more efficient. Besides, when threshold is between 0.35 and 0.65, the performance of SNN is nearly the same, indicating the robustness of the model. In an actual marketing setting, the threshold of the model is set to 0.15 according to the marketing strategy of operator A, so about 55% of the MNP churners can be correctly identified. The threshold could be further reduced to achieve a higher correct classification rate.

To further analyze the characteristics of the subscribers who tend to switch out, the top eight attributes to the model and the relative importance (RI) are shown in Fig.5. The RI is defined as

$$RI^i = \frac{F_1^i - F_0}{F_0}$$

where F_0 is the F1-score, F_1^i is the F1-score when the normalized attribute i is set to 0. In Fig.5, the RIs of the top eight attributes are scaled to 1. F_0 and F_1^i are calculated at threshold 0.55.

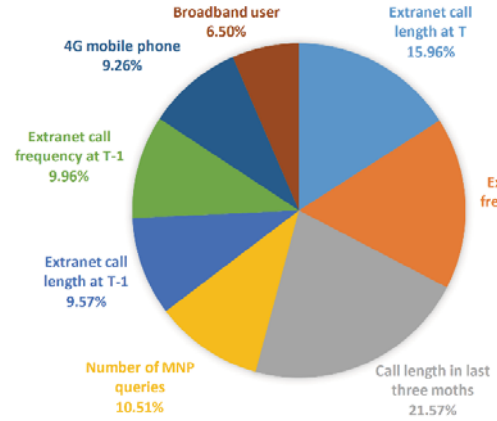


Fig. 5. The RI of top eight attributes that influencing the SNN model

As presented in Fig.5, the total call length, extranet call frequency and extranet call length are the most important factors that drive subscribers to switch out from operator A. Subscribers with long call time and high call frequency may be more sensitive to call cost and call quality [16]. In addition, subscribers who pay more attention to MNP also tend to leave operator A. Therefore, to minimize the number of churners, operator A should attach much weight to the above mentioned people that are labeled as churner by the SNN.

VI. PRODUCTION: MAP-MNP ANALYTICS PLATFORM

The total solution has been developed in production in a platform named MAP (MNP Analytics Platform) launched in multiple provinces in a tier-1MNO in China recently. The system structure of MAP is outlined in Fig.6. MAP is designed with 4 components. The foundation component is the MAP BigData Platform where the extranet and intranet source data is extracted, transformed, loaded, aggregated, and correlated in terms of MNP marketing data features. Typical MAP source data portfolio includes marketing data, subscribers' behaviors, profile, and property data, network and application performance data, and other competitors' strategy data. The second component is MAP Algorithm Engine, which contains three sub-components, MAP strategy making, MNP Churn Prediction and Treatment Matching. In the MAP strategy making sub-component, the real dynamic marketing strategy is output by the game theory-based algorithm with all possible marketing strategy options. Once the best marketing strategy is determined, our proposed machine learning model assisted by game theory in the MNP Churn Prediction sub-component will start to run and predict the MNP churn subscribers according to the corresponding market strategy. In the Treatment Matching sub-component, treatment recommendation algorithms are further used to match the churn users to her/his best treatment from the treatment library. In the Care Delivery component, the treatment will be delivered to the given subscriber along with the best marketing channel and timing through the customer care, with an ultimate goal to reducing his/her willingness to churn. Finally, the overall workflow and result is reactively assessed in the

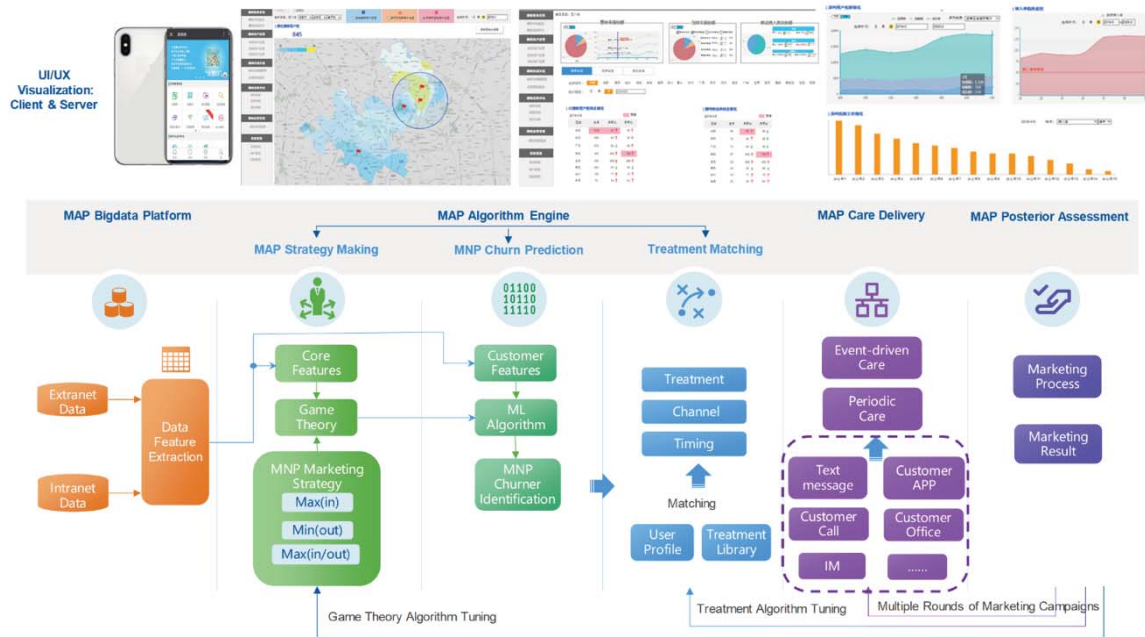


Fig. 6. The System Architecture of MAP (MNP Analytics Platform)

Posterior Assessment component in which the final feedback from the subscribers are looped back to the prior components for the purpose of incremental algorithm tuning. These four parts are equally important for the MAP platform. However, in this paper, we only focus on the Algorithm Engine part. Other parts of the MAP platform will be introduced in the future. After using the MAP platform, the average number of MNP churners in a particular province has dropped from about 20,000 to about 10,000 per month, reduced by nearly 50% (average value in six months).

VII. CONCLUSION

The introduction of the MNP enhances the liquidity and competition of the telecommunication market. To maximize the interest of the MNO, we have established a game theory-based scheme to capture the strategic interactions between MNOs. The strategies of MNO can be translated into detecting and predicting subscriber behaviors for further customer care to reduce the MNP driven churn. Driven by tens of millions of data, this problem can be solved by a game theory assisted machine learning algorithm as shown in this paper. As an example, we analyze the three their-1 MNOs in a given province in China and find that the best MNP strategy for MNO-A in that province is to minimize the number of switch out subscribers. To achieve this goal, we establish the game theory assisted stacking neural networks (SNN) model to predict the switching tendency, which can improve the marketing efficiency of MNO and reduce the customer churn by nearly 50% in a particular province and for a particular company. In addition, the SNN can demonstrate the top deciding factor for churning, which is the quality of data and the audio services in this case.

VIII. REFERENCES

- [1] J. P. Maicas, Y. Polo, and F. J. Sese, "Reducing the level of switching costs in mobile communications: The case of mobile number portability," *Telecommun. Policy*, vol. 33, no. 9, pp. 544–554, 2009.
- [2] Z. Yang and R. T. Peterson, "Customer perceived value, satisfaction, and loyalty: The role of switching costs," *Psychol. Market.*, vol. 21, no. 10, pp. 799–822, 2004.
- [3] J. Farrell and P. Klemperer, "Coordination and lock-in: Competition with switching costs and network effects," *Handbook of industrial organization*, vol. 3, pp. 1967–2072, 2007.
- [4] T. Otsuka and H. Mitomo, "User benefits and operator costs of mobilenumber portability in japan and impact on market competitiveness," *Telecommun. Policy*, vol. 37, no. 4-5, pp. 345–356, 2013.
- [5] D. H. Shin, "A study of mobile number portability effects in the united states," *Telemat. Inform.*, vol. 24, no. 1, pp. 1–14, 2007.
- [6] S. Bühler, R. Dewenter, and J. Haucap, "Mobile number portability in europe," *Telecomm. Policy*, vol. 30, no. 7, pp. 385–399, 2006.
- [7] FCC. "Wireless Local Number Portability (WLNP)". Internet: <https://www.fcc.gov/general/wireless-local-number-portability-wlnp>, May 18, 2016[Jul 26, 2019]
- [8] V. B. Viard, "Do switching costs make markets more or less competitive? the case of 800-number portability," *The RAND Journal of Economics*, vol. 38, no. 1, pp. 146–163, 2007.
- [9] China Mobile. "Announcements". Internet: <https://www.chinamobileltd.com/en/ir/announcements.php>, Apr. 29th, 2019 [Jul 26, 2019]
- [10] China Telecom. "Investor Calendar". Internet: <https://www.chinatelecom-h.com/en/ir/calendar.php>, Apr. 29th, 2019 [Aug 23, 2019]
- [11] China Unicom. "PRESENTATIONS & WEBCASTS". Internet: <https://www.chinaunicom.com.hk/en/ir/presentations.php>, Aug. 14th, 2019 [Aug 23, 2019]
- [12] SINA. "Operators will cancel the unlimited packages. Telecom will no longer reach the speed limit from next month". Internet: <https://news.sina.com.cn/s/2019-08-01/doc-ihyctitm6122548.shtml>, Aug 1th, 2019[Aug 25, 2019]
- [13] A. Berson, S. Smith, and K. Thearling, *Building data mining applications for CRM*. McGraw-Hill Professional, 1999.
- [14] A. Berson, S. Smith, and K. Thearling, *Building data mining applications for CRM*. McGraw-Hill Professional, 1999.
- [15] B. Garc'ia-Mari~ noso and D. Suárez, "Switching mobile operators: Evidence about consumers behavior from a longitudinal survey," *Telecommun. Policy*, vol. 43, no. 5, pp. 426–433, 2019.
- [16] T. J. Gerpott, W. Rams, and A. Schindler, "Customer retention, loyalty, and satisfaction in the german mobile cellular telecommunications market," *Telecommun. Policy*, vol. 25, no. 4, pp. 249–269, 2001.
- [17] D. Liang, Z. Ma, and L. Qi, "Service quality and customer switching behavior in china's mobile phone service sector," *J. Bus. Res.*, vol. 66, no. 8, pp. 1161–1167, 2013.
- [18] D.-H. Shin and W.-Y. Kim, "Forecasting customer switching intention in mobile service: An exploratory study of predictive factors in mobile

number portability,” Technol. Forecast. Soc. Change, vol. 75, no. 6, pp. 854–874, 2008.

- [19] M. E. Slade, “Vancouver’s gasoline-price wars: An empirical exercise in uncovering supergame strategies,” Rev. Econ. Stud., vol. 59, no. 2, pp. 257–276, 1992.
- [20] B. U. Sánchez and G. Asimakopoulos, “Regulation and competition in the european mobile communications industry: An examination of the implementation of mobile number portability,” Telecommun. Policy, vol. 36, no. 3, pp. 187–196, 2012.
- [21] J. Confraria, T. Ribeiro, and H. Vasconcelos, “Analysis of consumer preferences for mobile telecom plans using a discrete choice experiment,” Telecomm. Policy, vol. 41, no. 3, pp. 157–169, 2017.
- [22] S. O. Adebisi, E. O. Oyatoye, and J. N. Mojekwu, “Predicting customer churn and retention rates in nigerias mobile telecommunication industry using markov chain modelling,” Acta Universitatis Sapientiae, Economics and Business, vol. 3, no. 1, pp. 67–80, 2015.
- [23] V. Yeshwanth, V. V. Raj, and M. Saravanan, “Evolutionary churn prediction in mobile networks using hybrid learning,” in Twenty-fourth international FLAIRS conference, 2011.
- [24] R. Zhang and Q. Zhu, “A game-theoretic defense against data poisoning attacks in distributed support vector machines,” in 2017 IEEE 56th Annual Conference on Decision and Control (CDC). IEEE, 2017, pp.4582–4587.
- [25] M. E. Slade, “Vancouver’s gasoline-price wars: An empirical exercise in uncovering supergame strategies,” Rev. Econ. Stud., vol. 59, no. 2, pp. 257–276, 1992.
- [26] R. Ferrero, S. Shahidepour, and V. Ramesh, “Transaction analysis in deregulated power systems using game theory,” IEEE T. Power Syst., vol. 12, no. 3, pp. 1340–1347, 1997.
- [27] B. Magaloni, “The game of electoral fraud and the ousting of authoritarian rule,” Am. J. Polit. Sci., vol. 54, no. 3, pp. 751–765, 2010.
- [28] M. Casson, The economics of business culture: Game theory, transaction costs, and economic performance. Oxford University Press, 1994.
- [29] M. Czajkowski and M. Sobolewski, “How much do switching costs and local network effects contribute to consumer lock-in in mobile telephony?” Telecommun. Policy, vol. 40, no. 9, pp. 855–869, 2016.
- [30] G. Dorffner, “Neural networks for time series processing,” in Neural network world. Citeseer, 1996.
- [31] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [32] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM, 2016, pp. 785–794.
- [33] M. Ballings and D. Van den Poel, “Customer event history for churn prediction: How long is long enough?” Expert Syst. Appl., vol. 39, no. 18, pp. 13517–13522, 2012.
- [34] S. Mitrović, B. Baesens, W. Lemahieu, and J. De Weerd, “On the operational efficiency of different feature types for telco churn prediction,” Eur. J. Oper. Res., vol. 267, no. 3, pp. 1141–1155, 2018.

IX. APPENDIX

TABLE IX. GAME BENEFITS OF THREE OPERATORS IN DIFFERENT MARKETING STRATEGIES

Game Index		Game Players		
		A	B	C
1 st Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN/OUT}	Max _{IN/OUT}
	Benefit Score	0.3	-0.3	0
2 nd Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN/OUT}	Max _{IN}
	Benefit Score	1.5	0.9	-2.4
3 rd Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN/OUT}	Min _{OUT}
	Benefit Score	-0.1	-0.7	0.8
4 th Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN}	Max _{IN/OUT}
	Benefit Score	-0.3	0.9	-0.6

5 th Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN}	Max _{IN}
	Benefit Score	0.9	2.1	-3
6 th Round	Marketing Strategy	Max _{IN/OUT}	Max _{IN}	Min _{OUT}
	Benefit Score	-0.7	0.5	0.2
7 th Round	Marketing Strategy	Max _{IN/OUT}	Min _{OUT}	Max _{IN/OUT}
	Benefit Score	0	0.3	-0.3
8 th Round	Marketing Strategy	Max _{IN/OUT}	Min _{OUT}	Max _{IN}
	Benefit Score	1.2	1.5	-2.7
9 th Round	Marketing Strategy	Max _{IN/OUT}	Min _{OUT}	Min _{OUT}
	Benefit Score	-0.4	-0.1	0.5
10 th Round	Marketing Strategy	Max _{IN}	Max _{IN/OUT}	Max _{IN/OUT}
	Benefit Score	-1.3	0.5	0.8
11 th Round	Marketing Strategy	Max _{IN}	Max _{IN/OUT}	Max _{IN}
	Benefit Score	-0.1	1.7	-1.6
12 th Round	Marketing Strategy	Max _{IN}	Max _{IN/OUT}	Min _{OUT}
	Benefit Score	-1.7	0.1	1.6
13 th Round	Marketing Strategy	Max _{IN}	Max _{IN}	Max _{IN/OUT}
	Benefit Score	-1.9	1.7	0.2
14 th Round	Marketing Strategy	Max _{IN}	Max _{IN}	Max _{IN}
	Benefit Score	-0.7	2.9	-2.2
15 th Round	Marketing Strategy	Max _{IN}	Max _{IN}	Min _{OUT}
	Benefit Score	-2.3	1.3	1
16 th Round	Marketing Strategy	Max _{IN}	Min _{OUT}	Max _{IN/OUT}
	Benefit Score	-1.6	1.1	0.5
17 th Round	Marketing Strategy	Max _{IN}	Min _{OUT}	Max _{IN}
	Benefit Score	-0.4	2.3	-1.9
18 th Round	Marketing Strategy	Max _{IN}	Min _{OUT}	Min _{OUT}
	Benefit Score	-2	0.7	1.3
19 th Round	Marketing Strategy	Min _{OUT}	Max _{IN/OUT}	Max _{IN/OUT}
	Benefit Score	0.7	-0.5	-0.2
20 th Round	Marketing Strategy	Min _{OUT}	Max _{IN/OUT}	Max _{IN}
	Benefit Score	1.9	0.7	-2.6
21 st Round	Marketing Strategy	Min _{OUT}	Max _{IN/OUT}	Min _{OUT}
	Benefit Score	0.3	-0.9	0.6
22 nd Round	Marketing Strategy	Min _{OUT}	Max _{IN}	Max _{IN/OUT}
	Benefit Score	0.1	0.7	-0.8
23 rd Round	Marketing Strategy	Min _{OUT}	Max _{IN}	Max _{IN}
	Benefit Score	1.3	1.9	-3.2
24 th Round	Marketing Strategy	Min _{OUT}	Max _{IN}	Min _{OUT}
	Benefit Score	-0.3	0.3	0
25 th Round	Marketing Strategy	Min _{OUT}	Min _{OUT}	Max _{IN/OUT}
	Benefit Score	0.4	0.1	-0.5
26 th Round	Marketing Strategy	Min _{OUT}	Min _{OUT}	Max _{IN}
	Benefit Score	1.6	1.3	-2.9
27 th Round	Marketing Strategy	Min _{OUT}	Min _{OUT}	Min _{OUT}
	Benefit Score	0	-0.3	0.3