Practical Methods for Predicting Customer Retention

Cherkashin, Alexander and Sakhadzhi, Vladislav and Guliev, Ruslan and Bolshunova, Elena

T2 Mobile LLC

15 October 2024

Practical Methods for Predicting Customer Retention

Cherkashin Aleksandr
T2 Mobile LLC, Moscow, Russia
E-mail: a.cherkashin@t2.ru

Sakhadzhi Vladislav
T2 Mobile LLC, Moscow, Russia
E-mail: vladislav.sakhadzhi@t2.ru

Guliev Ruslan
T2 Mobile LLC, Moscow, Russia
E-mail: ruslan.guliev@t2.ru

Elena Bolshunova
T2 Mobile LLC, Moscow, Russia
E-mail: elena.bolshunova@t2.ru

Abstract. This study examines methods for analyzing and forecasting the retention of active subscribers in the telecommunications industry using various criteria for subscriber activity. The results demonstrate that the retention dynamics of an active subscriber base can be effectively modeled using a decreasing power function. This allows for medium-term forecasting based on truncated subscriber activity data. However, it is important to note the potential limitations in the effectiveness of the proposed approach for long-term forecasting, associated with changes in subscriber churn dynamics over time.
Keywords: subscriber base, customer retention, customer churn, power law, power function, telecommunications, LTV.
JEL Classification: C53, L96, D12, M31

Introduction

Existing research on predicting customer retention and churn primarily focuses on machine learning methods and big data analysis (Verhelst et al., 2021; Verbeke et al., 2012). These approaches often require significant computational resources and large datasets. Moreover, they are generally aimed at identifying factors influencing customer churn propensity (Ribeiro et al., 2023; Jain et al., 2021) or predicting individual customer behavior to enable targeted interventions. However, for economic modeling purposes, it is crucial to have accurate yet easily applicable methods for forecasting the behavior of an entire customer base or its large segments. For instance, this is essential for estimating the expected average total revenue per customer, known as Lifetime Value (LTV) (Gupta & Zeithaml, 2006). This study focuses on analyzing the behavior of large customer groups and the inverse of churn—retention of the active customer base, or, borrowing a term from biology, customer base survival. The research demonstrates that active customer base retention can be accurately modeled using power functions, and their extrapolation enables precise survival forecasting. However, before proceeding with modeling active customer base retention, it is necessary to discuss practical considerations regarding the selection of customer activity criteria, as these directly affect the metric itself and its interpretation.

1. Customer Activity Criteria

The analysis and forecasting of active customer base retention largely depend on the selection of criteria used to define customer activity. This section examines various approaches to determining customer activity and their impact on assessing customer base survival. The following classification of activity criteria is proposed:

1. Legal criteria;
2. Financial criteria;
3. Service usage criteria.

Each of these categories has its own advantages and limitations, which will be discussed in detail in the following sections.

1.1. Legal Criteria of Activity

Legal criteria of activity are based on the formal status of contractual relationships between a company and a subscriber. The primary indicator in this category is the presence of an active service agreement. However, in the context of mass-market services, subscriber activity is generally weakly correlated with the existence of a valid contract. There is often a significant lag between the actual cessation of service usage and the formal termination of the contract. For example, many mobile network operators, including T2, implement automatic contract termination after an extended period of inactivity. Under T2's terms, a contract is terminated if a subscriber does not perform any chargeable actions for 180 days while maintaining a zero or negative balance. In some industries, such as retail or one-time services, contractual relationships exist only within the scope of individual transactions, rendering this criterion unsuitable for long-term survival analysis. Consequently, the application of legal criteria to analyze subscriber base retention has significant limitations.

Despite its limited applicability in the telecommunications sector, legal criteria may be relevant in other economic sectors and legal frameworks where the formal status of a contract is more closely linked to actual service usage.

1.2. Financial Criteria of Activity

Financial criteria of activity are based on various aspects of financial interaction between a subscriber and a telecom operator, including revenue generation, subscriber payments, account balance status, and service-related costs. Using financial indicators as activity criteria has several clear advantages. Primarily, they directly reflect the economic value of a subscriber to the company, which is particularly important for economic modeling.

However, despite this advantage, financial criteria for assessing subscriber base retention come with significant limitations. One key factor affecting the reliability of financial criteria is their dependence on the company's accounting policies. Revenue recognition timing may vary significantly depending on the organization's established rules. For example, in the case of an upfront payment for an annual service plan, revenue can either be recorded as a lump sum or distributed over the entire paid period, leading to vastly different interpretations of subscriber activity. Such advance payments can either create an illusion of continuous activity even if the subscriber stops using the services after a few months or, conversely, indicate a complete lack of activity beyond the initial prepaid month. Neither scenario provides an accurate representation of actual subscriber behavior, making these data unsuitable for forecasting. This issue is particularly relevant in the context of marketing campaigns and special offers that incentivize long-term prepayments. A separate concern is the issue of remaining account balances for inactive subscribers. If a subscriber ceases using the service but still has funds in their account, the gradual deduction of these funds as revenue may create a misleading impression of ongoing activity. Such phenomena can lead to systematic distortions in the analysis of subscriber base retention.

Balance top-ups, as a potential activity criterion, have the advantage of being independent of the company's accounting policies. However, they are not necessarily synchronized with actual service usage: a subscriber may deposit a large sum and then not use the services for an extended period, which would also lead to misinterpretations of their activity.

Using cost-related indicators as an activity criterion presents the same challenges as using revenue, further complicated by the intricacies of cost accounting and allocation. Moreover, costs are typically associated with service usage, which is more effectively measured directly (this will be further discussed in the next chapter on actual service usage criteria).

Given these considerations, it can be concluded that while financial indicators are essential for the economic evaluation of a subscriber base, their use as activity criteria requires an extremely cautious approach. These limitations may result in significant distortions in the analysis and forecasting of subscriber retention dynamics. To mitigate such issues, a thorough analysis of the company's accounting policies should be conducted before applying financial criteria. Additionally, when interpreting financial activity indicators, it is crucial to consider the business model and service delivery specifics. Although financial criteria are valuable for assessing the overall economic efficiency of a subscriber base, they cannot be regarded as a universal or sufficient tool for analyzing subscriber retention. Nonetheless, financial criteria can be effective for specific, narrowly focused tasks (e.g., analyzing the payment compliance of subscribers under fixed monthly fee plans).

1.3. Criteria for Actual Service Usage

Criteria for actual service usage represent the most objective and direct approach to assessing subscriber activity in the telecommunications industry. Unlike the legal and financial criteria discussed earlier, these indicators are based on real service consumption, ensuring high accuracy and relevance of subscriber behavior data. The key advantage of these criteria lies in their independence from the company's accounting policies and other managerial factors. The use of telecom services creates a direct load on the network infrastructure, reflecting real physical processes that are recorded at the moment they occur. This allows for obtaining up-to-date subscriber behavior data without the time delays inherent in financial indicators.

In the context of a mobile network operator, criteria for actual service usage encompass a wide range of services, including voice calls, data transmission, and text messaging. However, not all types of network activity are appropriate for evaluating subscriber retention. For example, technical traffic generated by a device to maintain network connectivity cannot be considered an indicator of subscriber activity (although it does indicate device activity). When developing a system of criteria for actual service usage, it is essential to consider the specifics of the provided services and their consumption patterns.

Despite their obvious advantages, actual service usage criteria also have certain limitations. The primary drawback is the potential discrepancy between actual service usage and its economic value for the company. A subscriber may actively use services included in a bundle without generating additional revenue. Conversely, a subscriber classified as inactive under these criteria may still be making regular payments, which is significant from a financial standpoint. For example, a subscriber may maintain a positive balance and regularly pay for services but not use them for various reasons. This situation can arise when a subscriber owns multiple SIM cards from different operators or retains a number solely for registration on online platforms. To address this limitation, one approach is to analyze the actual difference between "paying" and "service-using" subscribers and incorporate it into modeling. Another approach is to adopt a combined criteria that integrates both actual service usage and financial data.

The choice of a specific criterion or combination of criteria should be made considering the business model, analysis objectives, and characteristics of the subscriber base under study. There is no universal solution suitable for all cases; each company must develop its own approach to defining and measuring subscriber retention. In this study, we use two specially designed subscriber activity criteria:

1. Criterion A – considers only certain actual service consumption by the subscriber.
2. Criterion B – a combined approach that includes both actual service consumption and financial indicators of subscriber activity.

2. Conformity of Retention Dynamics to Power Functions

In the context of the telecommunications business, subscriber base retention, or survival, refers to the continued use of a company's services by a subscriber. The retention metric represents the proportion of active subscribers, according to the selected criterion, out of the initially connected cohort. Mathematically, retention for each period can be expressed using the following formula:

$$R_i = \frac{A_i(T,c)}{G_T} \tag{1}$$

where:

- R_i – retention in the i-th period;
- $A_i(T,c)$ – the number of active subscribers (according to the activity criterion c) in period i from the cohort onboarded in period T;
- G_T – the number of activations in period T;
- $i \in [1; +\infty]$ – the period for retention calculation;
- T – a fixed initial period (usually the connection period);
- c – the subscriber activity criterion (see Chapter 1).

The calculation and analysis of retention dynamics constitute cohort analysis (Fader & Hardie, 2009; Zhang & Chang, 2021) and serve two primary objectives: diagnosing the current state and providing a foundation for building predictive models. To identify patterns in subscriber behavior, it is advisable to examine retention dynamics over extended time intervals. In this study, we analyzed the retention dynamics of subscribers who joined the T2 network in November 2015, tracking their activity based on Criterion B over a period of 105 months (8.5 years) until July 2024 (Figure 1).

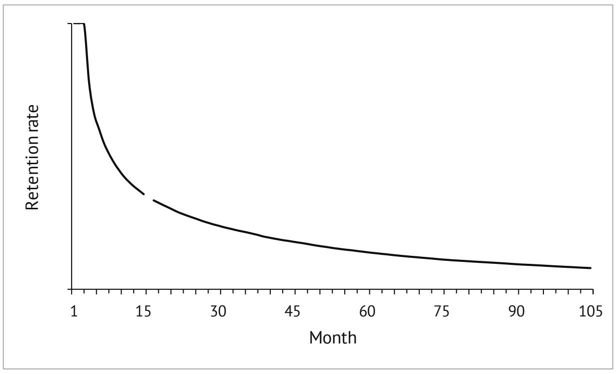

Figure 1. B-Retention of Subscribers Who Joined T2 in November 2015

Source: Developed by the authors based on T2 subscriber activity data.

Note: The figure represents a graphical depiction of a time series of discrete retention values for each observation month. For better visualization, the discrete time series is displayed as a continuous line in this and subsequent figures. The vertical axis ranges from 0 to 1 (0% to 100%). To prevent disclosure of the company's actual data, numerical labels have been removed from this and all following figures. Data on active subscribers for February 2017 is missing, which appears as a gap in the graph.

A visual analysis of the graph reveals a monotonous and smooth retention trend, suggesting the hypothesis that it can be approximated using a relatively simple mathematical function.

To test this hypothesis, four types of functions were considered:

- Power function: ax^b;
- Exponential function: ae^{bx};
- Logarithmic function: $aln(x) + b$;

- Linear function: $ax + b$.

For simplicity, the parameters of the nonlinear functions were estimated after linearization using logarithmic transformation. Thus, the parameters were determined using the least squares method (LSM) for the following transformed functions:

- $\ln(y) = \ln(a) + b\ln(x) + \varepsilon$;
- $\ln(y) = \ln(a) + bx + \varepsilon$;
- $y = aln(x) + b + \varepsilon$;
- $y = ax + b + \varepsilon$.

This approach does not provide entirely accurate parameter estimates for the power and exponential approximations, as it assumes multiplicative residuals for the original expressions ax^b and ae^{bx}. Nevertheless, we consider it an acceptable method for an initial hypothesis assessment and selection of the most suitable function.

The approximation results are presented in Figure 2 and Table 1.

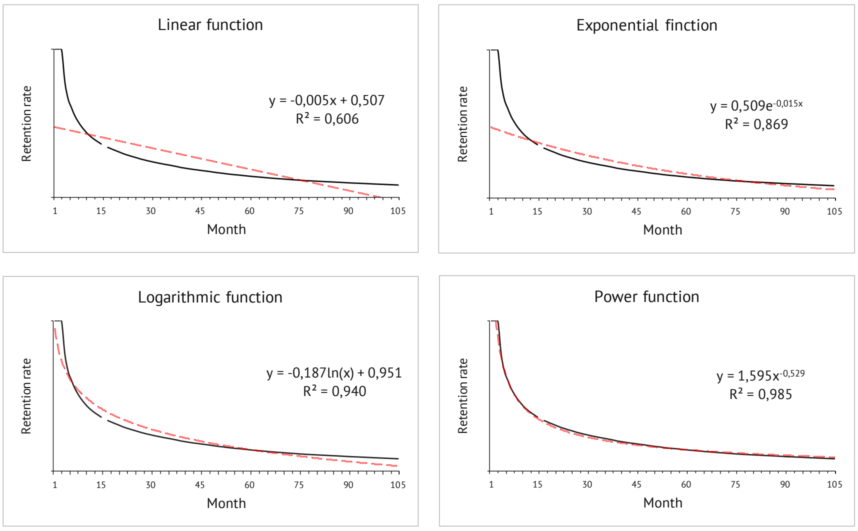

Figure 2. Approximation of the B-Retention Curve Using Different Functions

Source: Developed by the authors based on T2 subscriber activity data.

Analysis of the results shows that the power function provides the best approximation of the retention curve, with a coefficient of determination (R^2) = 0.985.

Table 1. Approximation of Retention Using Different Functions

Function type	Function	R^2
Power	$y = 1{,}595x^{-0.529}$	0,985
Logarithmic	$y = -0{,}187\ln(x) + 0{,}951$	0,940
Exponential	$y = 0{,}509e^{-0.015x}$	0,869
Linear	$y = -0{,}005x + 0{,}507$	0,606

Source: Compiled by the authors based on T2 subscriber activity data.

To verify the obtained results, we performed an approximation of actual retention based on Criterion A for various subscriber groups who joined the company's network in different periods and through different sales channels. The results are presented in Figure 3 and Table 2.

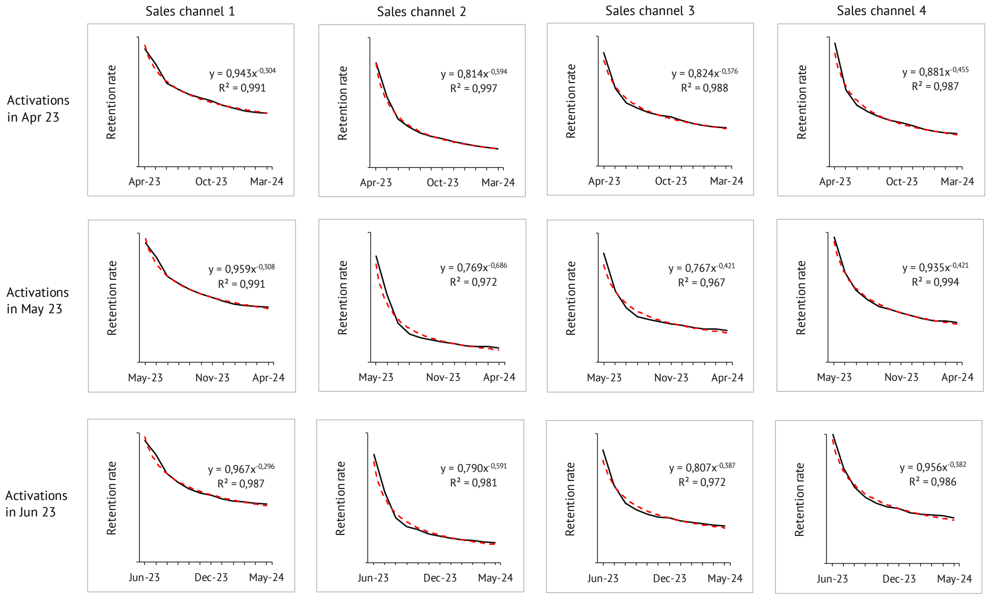

Figure 3. Approximation of A-Retention Curves Using a Power Function for Different Sales Channels and Activation Periods

Source: Developed by the authors based on T2 subscriber activity data.

Analysis of Table 2 confirms the high stability of power function approximation results. In all examined cases R^2 exceeds 0.96, indicating a strong fit between the model and empirical data, regardless of the sales channel or subscriber activation period.

Table 2. Power Function Approximation for Different Sales Channels and Activation Periods

Sales channel	Activation month	Function	R^2
	Apr’23	$y = 0{,}943\ x^{-0.304}$	0,991
1	May’23	$y = 0{,}959\ x^{-0.308}$	0,991
	Jun’23	$y = 0{,}967\ x^{-0.296}$	0,987
	Apr’23	$y = 0{,}814\ x^{-0.594}$	0,997
2	May’23	$y = 0{,}769\ x^{-0.686}$	0,972
	Jun’23	$y = 0{,}790\ x^{-0.591}$	0,981
	Apr’23	$y = 0{,}824\ x^{-0.376}$	0,988
3	May’23	$y = 0{,}767\ x^{-0.421}$	0,967
	Jun’23	$y = 0{,}807\ x^{-0.387}$	0,972
	Apr’23	$y = 0{,}881\ x^{-0.455}$	0,987
4	May’23	$y = 0{,}935\ x^{-0.421}$	0,994
	Jun’23	$y = 0{,}956\ x^{-0.382}$	0,986

Source: Compiled by the authors based on T2 subscriber activity data.

The obtained results allow us to formulate the following observations and hypotheses:

1. The high accuracy of retention approximation using a power function across different activity criteria, subscriber base segments, and time periods may indicate the existence of a fundamental pattern in subscriber retention that follows a power law.

2. Variations in the parameters of the approximating power function for different subscriber groups may exclusively reflect the influence of initial conditions (sales channel, activation period) on the long-term retention dynamics.
3. The preservation of retention curve shapes, despite the influence of numerous varying factors (service quality, retention activities, competitive acquisition efforts), suggests the possibility of a dynamic equilibrium within the competitive market. This equilibrium may exist between companies' efforts to retain subscribers and the factors contributing to churn.
4. Power functions are widely observed in describing patterns across multiple fields, including physics, astronomy, biology, geology, sociology, linguistics, psychology, and economics (Andriani & McKelvey, 2007). In M. Newman's research (Newman, 2005), power laws are examined in disciplines such as physics, biology, and economics, including their manifestations in Pareto distribution and Zipf's law. In economics and finance, power functions describe key phenomena such as income and wealth distribution, company sizes, stock market returns, trading volumes, and international trade metrics (Gabaix, 2009). The predominance of power-law dependencies in subscriber retention raises fundamental questions about the underlying processes governing subscriber behavior. For instance, J. Staddon's study (Staddon, 1978) attempts to justify the emergence of power-law dependencies in the behavior of living organisms in response to various stimuli.

The high effectiveness of power function approximation for subscriber retention opens up opportunities for forecasting subscriber base dynamics. The following chapters will explore the practical applications of these findings for building predictive models and discuss the limitations of the proposed approach.

3. Extrapolation of Retention and Its Use for Forecasting

The high accuracy of retention curve approximation using power functions, as demonstrated in Chapter 2, presents opportunities for their use in forecasting. Extrapolation allows for the estimation of subscriber base retention over long time intervals without waiting for actual data. This is particularly important for economic modeling and evaluating the effectiveness of subscriber acquisition strategies. Subscriber retention forecasting plays a key role in calculating critical economic indicators such as Customer Lifetime (expected subscriber lifespan) and Customer Lifetime Value (LTV) (Kumar, 2014). These metrics are essential for comparing against Subscriber Acquisition Cost (SAC) and, consequently, for making informed decisions about launching, continuing, or discontinuing campaigns and mechanisms for expanding the subscriber base (Krstevski & Mancheski, 2016).

To evaluate the effectiveness of power function extrapolation in retention forecasting, the following analysis was conducted:

1. Data on subscriber retention according to Criterion A was used for subscribers who joined in April 2023, covering a 12-month period.
2. Power function approximations were built using nonlinear regression with additive residuals (see Appendix A), applying an increasing number of actual retention values from the first months (ranging from 2 to 11 months).

3. Retention forecasts were generated by extrapolating each approximating function up to the 12th month. This resulted in model-based retention estimates for the entire 12-month period, with each model using a different amount of actual data.
4. The coefficient of determination (R^2) was calculated for each model based on 12 months of data, serving as both a measure of fit to actual retention and an indirect indicator of forecast quality.
5. A graph was constructed to illustrate the relationship between R^2 values and the amount of actual data used for model construction.

The analysis results are presented in Figures 4 and 5 and Table 3.

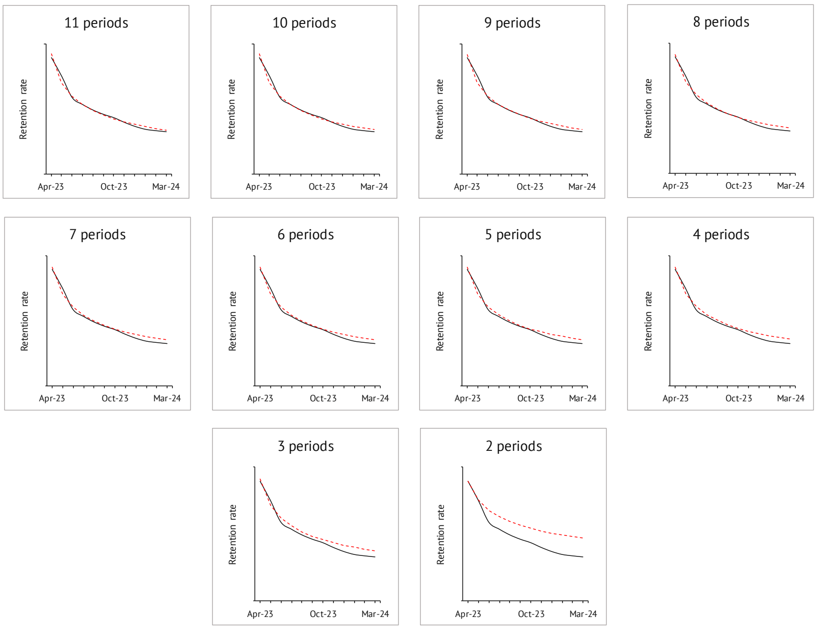

Figure 4. Comparison of Actual A-Retention with Different Forecasts Based on Power Function Extrapolation Using Varying Amounts of Actual Data

Source: Developed by the authors based on T2 subscriber activity data.

A sharp increase in forecast quality is observed when the first three actual retention values are used. Beyond this point, further improvements are minimal.

Table 3. Parameters of Approximating Functions and Forecast Errors

Actual periods	Forecast periods	Function	R^2
11	1	$y = 0{,}927x^{-0.370}$	0,988
10	2	$y = 0{,}925x^{-0.367}$	0,987
9	3	$y = 0{,}922x^{-0.361}$	0,986
8	4	$y = 0{,}919x^{-0.356}$	0,984
7	5	$y = 0{,}917x^{-0.351}$	0,982
6	6	$y = 0{,}917x^{-0.351}$	0,981
5	7	$y = 0{,}916x^{-0.348}$	0,980
4	8	$y = 0{,}915x^{-0.343}$	0,975
3	9	$y = 0{,}911x^{-0.327}$	0,954
2	10	$y = 0{,}900x^{-0.241}$	0,575

Source: Compiled by the authors based on T2 subscriber activity data.

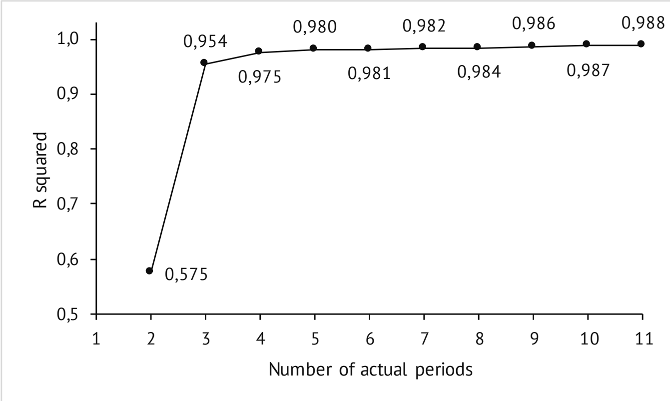

Figure 5. Dependence of R^2 on the Number of Actual Data Points Used in Models

Source: Developed by the authors based on T2 subscriber activity data.

The objective of this analysis was to propose a general approach for evaluating forecast quality, demonstrate the applicability of forecasting based on a limited set of actual data, and suggest a method for determining the optimal amount of historical data needed for building a forecast. The fact that using just a few actual values results in a sufficiently high forecast accuracy, with only marginal improvements thereafter, holds practical significance. For instance, it is possible to use only the first three retention values to achieve a reasonably accurate forecast of future retention.

It is important to note that extrapolation does not necessarily require the first few actual values. This method allows for assessing a "hypothetical past", meaning it can be used for retrospective analysis and evaluating alternative scenarios. This can be valuable when analyzing the effectiveness of past subscriber retention campaigns.

In practical applications, it is often reasonable to construct forecasts using all available retention data at a given moment. These could include retention values of different subscriber cohorts activated in different months. For example, if we consider a full year of activations, we could use: the first-month retention for subscribers who joined in January, February, ..., December (12 values), the second-month retention for subscribers who joined in January, February, ..., November (11 values), and so forth. This approach enhances forecast stability by reducing the impact of seasonality and random fluctuations.

The conducted research demonstrates the strong potential of using power function extrapolation for forecasting subscriber retention. However, despite its effectiveness, this approach has certain limitations that require careful examination. The next chapter will analyze these limitations, specifically discussing the long-term applicability of power functions for retention modeling.

4. Limitations of Applicability

As demonstrated in Chapter 2, power functions of the form ax^b provide a highly accurate approximation of actual subscriber base retention. Chapter 3 explored the potential use of these functions for forecasting. However, despite their effectiveness, it is evident that the high quality of forecasts based on power function extrapolation has certain limits of applicability. From a mathematical standpoint, as $x \to +\infty$, a function of the form ax^b (where $a > 0$ and $b < 0$) asymptotically approaches zero without crossing the x-axis:

$$\lim_{x \to +\infty} a\,x^b = 0, \quad \text{for } a > 0, b < 0 \tag{2}$$

This implies that forecasts based on such functions predict that retention will never reach zero. However, from a practical perspective, this assumption does not reflect reality. The retention of a subscriber group is an aggregated measure of individual subscriber states, where each subscriber at any given moment can only be in one of two states: 1 (active subscriber) or 0 (inactive subscriber). Logically, the number of active subscribers will continuously decrease over time until only one active subscriber remains. Once this last subscriber leaves, the group's overall retention will inevitably drop to zero—an event that a function of the form ax^b can never predict.

Moreover, subscriber retention is influenced not only by factors related to customers' willingness to use the company's services but also by objective constraints, such as the physical ability to use telecom services. Subscribers are living individuals (except for the segment of devices with SIM cards, which also have a limited lifespan). Inevitably, a point will be reached when every subscriber physically can no longer continue using the service.

These considerations highlight a potential limitation in the long-term applicability of power function extrapolation for retention forecasting. For a more detailed analysis of this limitation, we turn to the B-retention curve and its power function approximation, presented in Figures 1 and 2. Closer examination of the graph's later periods reveals that, despite the high overall accuracy of the approximation, as time progresses, there is a systematic overestimation of actual retention by the approximating curve.

To better illustrate this overestimation effect, we will first construct a more precise approximation (where Figure 2 previously showed an approximation assuming multiplicative residuals), using nonlinear regression with additive residuals (see Appendix A). This will be done by using only the first three actual retention values, after which we will extrapolate retention over the entire period.

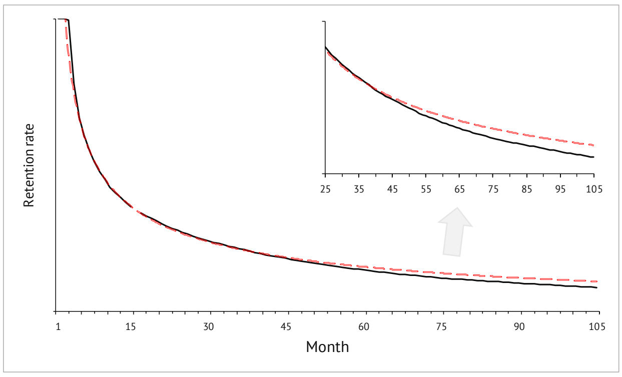

Figure 6. Extrapolation of B-Retention Based on Initial Actual Values

Source: Developed by the authors based on T2 subscriber activity data.

Note: Due to the methodological specifics of Criterion B, the first four values provide a distorted representation of actual subscriber activity. Therefore, for extrapolation, the actual values from months 5, 6, and 7 were selected. Deviations in the first few months are also explained by the specific nature of the chosen activity criterion.

The results (Figure 6) show a growing discrepancy between the forecasted and actual retention as the subscriber cohort ages. This observation suggests the hypothesis that additional factors emerge over time, influencing subscriber base retention. As previously discussed, one such factor may be the physical age of subscribers, whose effect on telecom service usage presumably increases with the duration of their presence in the subscriber base.

Based on the survival curves of large mammals, including humans (Type I curve in Figure 7) [Deevey, 1947; Pearl & Miner, 1935], we can hypothesize the general shape of the subscriber retention curve while considering the age factor.

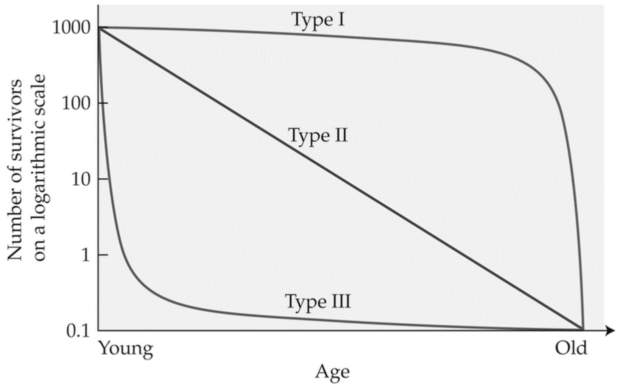

Figure 7. Three Main Types of Survival Curves for Animal Populations in the Wild According to E. Deevey

Source: [Hill et al., 2021]

If we consider the subscriber retention curve as the probability of maintaining activity, then the physical age factor should gradually modify it in accordance with a Type I survival curve (Figure 7). As a result, the general shape of the subscriber retention curve may take on an S-shaped form, consisting of two distinct phases separated by an interval of uncertain duration (Figure 8). The first phase reflects retention dynamics influenced by service quality and customer loyalty. The second phase demonstrates the impact of other objective physical factors.

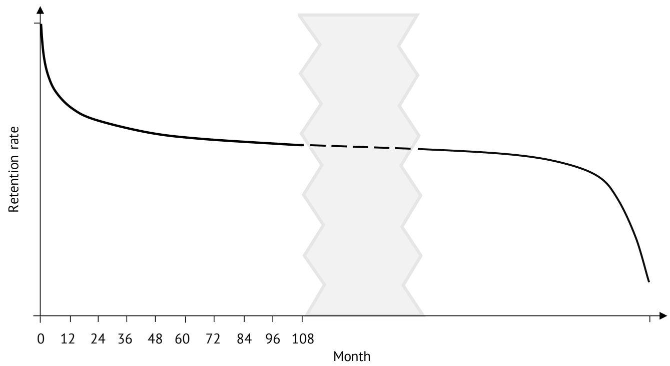

Figure 8. General Shape of the S-Shaped Subscriber Retention Curve

Source: Developed by the authors.

It is possible that this S-shaped curve, with its deceleration and acceleration phases, represents a universal survival pattern applicable not only to telecom subscribers but also to biological organisms. If comparable starting points of observation were aligned, all three types of survival curves might be variations of this generalized curve, with differences arising due to the positioning of deceleration and acceleration zones and the life stage most emphasized in analysis.

Studying the long-term effects of age and other demographic factors on subscriber base retention presents a promising direction for future research. Such an analysis could contribute to the development of more accurate long-term forecasting models and help define the overall shape of the subscriber retention curve throughout the entire subscriber lifecycle. However, as demonstrated earlier, for short- and medium-term forecasting, this factor can be reasonably neglected.

Conclusion

This study proposes a classification of subscriber activity criteria and discusses their practical applications. Combined indicators of actual service usage, sometimes in conjunction with financial indicators, appear to be the most universal approach for modeling subscriber base retention.

It has been demonstrated that subscriber retention dynamics follow a power law, specifically curves of the form ax^b ($where\ a > 0$ and $b < 0$). This finding can be utilized for forecasting subscriber retention through simple extrapolation, using only three to four actual data points in our case. The ability to construct simple and accurate forecasts based solely on initial retention values has significant applications in economic modeling and in assessing the total expected impact of subscriber acquisition, including key metrics such as average subscriber lifetime and LTV (Customer Lifetime Value).

At the same time, long-term forecasts based on power function extrapolation have limitations, which are presumably linked to the growing influence of additional factors over time, such as subscriber age. It can be hypothesized that the general shape of the retention function should be a combination of different patterns, where the initial part of the curve follows a power law, while the later stages deviate from this form.

This study leaves several key questions open, each of which requires further investigation:

1. Why does subscriber retention follow power functions so precisely?
2. Why does retention (and, correspondingly, churn) appear so uniform and seemingly dependent only on initial conditions? Could this be a manifestation of a dynamic equilibrium among competitive market forces?
3. What could be the general long-term shape of the subscriber retention curve?

References

1. Andriani, P. Beyond Gaussian averages: Redirecting international business and management research toward extreme events and power laws / P. Andriani, B. McKelvey // Journal of International Business Studies. – 2007. – №38(7) . – P. 1212-1230.
2. Deevey, E. S., Jr. Life tables for natural populations of animals // The Quarterly Review of Biology. – 1947. – №22. – P. 283-314.
3. Fader, P. S. Probability models for customer-base analysis / P. S. Fader, B. G. Hardie // Journal of interactive Marketing. – 2009. – №23(1). – P. 61-69.
4. Gabaix, X. Power Laws in Economics and Finance // Annual Review of Economics. – 2009. – №1. – P. 255-293.
5. Gupta, S. Customer metrics and their impact on financial performance / S. Gupta, V. Zeithaml // Marketing Science. – 2006. – №25(6). – P. 718-739.
6. Hill, R. W. Youth in the study of comparative physiology: insights from demography in the wild / R. W. Hill, D. Sleboda, J. J. Millar // Journal of Comparative Physiology B. – 2021. – №191. – P. 1-16.
7. Jain, H. Telecom churn prediction and used techniques, datasets and performance measures: a review / H. Jain, A. Khunteta, S. Srivastava // Telecommunication Systems. – 2021. – №76(4). – P. 613-630.
8. Krstevski, D. Managerial Accounting: Modeling Customer Lifetime Value – An Application in the Telecommunication Industry / D. Krstevski, G. Manchenski // European Journal of Business and Social Sciences. – 2016. – №5(1). – P. 64-77.
9. Kumar, V. Costumer lifetime value: The path to profitability. – Boston: Now Publishers, 2008. – 112 p.
10. Newman, M. E. J. Power laws, Pareto distributions and Zipf's law // Contemporary Physics. –2005. – №46(5). – P. 323-351.
11. Pearl, R. Experimental studies on the duration of life. XIV. The comparative mortality of certain lower organisms / R. Pearl, J. R. Miner // The Quarterly Review of Biology. – 1935. – №10. – P. 60-79.
12. Ribeiro, H. Determinants of churn in telecommunication services: a systematic literature review / H. Ribeiro, B. Barbosa, A. C. Moreira & R. G. Rodrigues // Management Review Quarterly. – 2023. – №74(3). – P. 1327-1364.
13. Staddon, J. Theory of behavioral power functions // Psychological Review. – 1978. – №85(4). – P. 305-320.
14. Verhelst, T. Understanding Telecom Customer Churn with Machine Learning: From Prediction to Casual Inference / T. Verhelst, O. Caelen, J. C. Dewitte, B. Lebichot, G. Bontempi // Artifival Intelligence and Machine Learning. – 2021. – №1996. – P. 182-200.
15. Verbeke, W. Building comprehensible customer churn prediction models with advanced rule induction techniques / W. Verbeke, D. Martens, C. Mues, B. Baesens // Expert Systems with Applications. –2011. – №38(3). – P. 2354-2364.
16. Zhang J. Z., Chang, C. W. Consumer dynamics: Theories, methods, and emerging directions / J. Z. Zhang, C. W. Chang // Journal of the Academy of Marketing Science. – 2021. – №49(1). – P. 166-196.

Appendix A. Estimation of Nonlinear Regression Parameters for a Power Function with Additive Residuals Using the Least Squares Method

A nonlinear regression model for a power function with additive residuals is given by:

$$\boxed{y_t = ax_t^b + \varepsilon_t,\ t = 1 \div T}$$

Considering the general equation of nonlinear regression: $y = ax^b + \varepsilon$

Taking the natural logarithm of both sides: $\ln y = \ln(ax^b + \varepsilon)$,

Let $ax^b = z$, then: $\ln y = \ln(z + \varepsilon)$

Using the Maclaurin series expansion for $f(\varepsilon) = f(0) + f'(0)\varepsilon + O(\varepsilon^2)$,
where $f(\varepsilon) = \ln(z + \varepsilon)$, we obtain:

$$f(\varepsilon) = \ln(z + \varepsilon) = \ln(z) + \frac{1}{z}\varepsilon + O(\varepsilon^2) \approx \ln(z) + \frac{\varepsilon}{z}$$

Since $z = ax^b$, we substitute:

$$f(\varepsilon) \approx \ln(ax^b) + \frac{\varepsilon}{ax^b} = \ln a + b\ln x + \frac{\varepsilon}{y - \varepsilon} \approx \ln a + b\ln x + \frac{\varepsilon}{y}$$

Thus, we obtain a linear regression model with heteroskedastic residuals:

$$\boxed{\ln y_t = \ln a + b\ln x_t + \frac{\varepsilon_t}{y_t} = \tilde{a} + b\ln x_t + \frac{\varepsilon_t}{y_t},\ t = 1 \div T, \tilde{a} = \ln a}$$

Regression Residuals: $\varepsilon_t = y_t(\ln y_t - \tilde{a} - b\ln x_t),\ t = 1 \div T$

Sum of Squared Regression Residuals:

$$\Sigma(\tilde{a}, b) = \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} y_t^2(\ln y_t - \tilde{a} - b\ln x_t)^2$$

To find the optimal parameters, we solve the normal equations system:

$$\frac{\partial\Sigma(\tilde{a}, b)}{\partial\tilde{a}} = -\sum_{t=1}^{T} 2y_t^2(\ln y_t - \tilde{a} - b\ln x_t) = 0$$

$$\frac{\partial\Sigma(\tilde{a}, b)}{\partial b} = -\sum_{t=1}^{T} 2y_t^2(\ln y_t - \tilde{a} - b\ln x_t)\ln x_t = 0$$

After simplification, we obtain a system of linear equations:

$$\tilde{a}\sum_{t=1}^{T} y_t^2 + b\sum_{t=1}^{T} y_t^2\ln x_t = \sum_{t=1}^{T} y_t^2\ln y_t$$

$$\tilde{a}\sum_{t=1}^{T} y_t^2\ln x_t + b\sum_{t=1}^{T} y_t^2(\ln x_t)^2 = \sum_{t=1}^{T} y_t^2\ln y_t\ln x_t$$

Define the summation terms:

$$S_1 = \sum_{t=1}^{T} y_t^2\ ,\ \ S_2 = \sum_{t=1}^{T} y_t^2\ln x_t\ ,\ \ S_3 = \sum_{t=1}^{T} y_t^2\,(\ln x_t)^2\ ,\ \ S_4 = \sum_{t=1}^{T} y_t^2\ln y_t\ ,\ \ S_5 = \sum_{t=1}^{T} y_t^2\ln y_t\ \ln x_t$$

Then, the system of equations can be rewritten as:

$$\tilde{a}S_1 + bS_2 = S_4$$
$$\tilde{a}S_2 + bS_3 = S_5$$

Solving the system of equations, we obtain the estimates for $\tilde{a}$ and b:

$$\boxed{\tilde{a} = \frac{S_4S_3 - S_5S_2}{S_1S_3 - S_2^2}\ ,\quad b = \frac{S_1S_5 - S_4S_2}{S_1S_3 - S_2^2}\ ,\quad a = e^{\tilde{a}}}$$