



Assessment of NetForecast, Inc. Methodology

Prepared for NetForecast, Inc.

March 13th, 2018

Jonathan Falk

I. Introduction

We have been asked by NetForecast, Inc. (“NetForecast”) to perform an independent assessment from a statistical perspective of the methodology they use to audit the accuracy of monthly data usage by customers of Internet Service Providers (“ISPs”).

Our conclusion is that this methodology is capable of giving unbiased estimates of the quantities at issue: the divergence between measured usage and actual usage, the fraction of total customers out of specification and measured in excess of actual usage. In addition, this methodology can also give reliable measures of the uncertainty of these figures: the margin of error.

While this methodology is *capable* of reliably estimating both the required metrics and the margin of error for those metrics, whether or not it actually does so is a question of whether the sample of customers used is drawn in a statistically sensible fashion to be representative of the entire population. And even if the sample is statistically representative, whether it is *adequate* to give results within an acceptable margin of error will depend on the size of the sample. Both of these features (sampling method and sample size) are in the control of the particular ISP client, not NetForecast. NetForecast has prepared documents which state criteria required to make the sample representative and sufficiently accurate, but the implementation of these criteria is outside of NetForecast’s control. That said, an ISP which follows the criteria recommended by NetForecast can be confident that the results returned by NetForecast are as reliable as promised.

This report focuses on the *statistical* aspects of the NetForecast methodology. There are a host of technical considerations as to how divergences between bytes measured and bytes used are calculated about which we have been informed but have no opinion.

II. Background

NetForecast’s audit of data usage accuracy works by comparing an ISP’s hourly measurement of bytes sent and received by some customer against NetForecast’s own measurements of the same quantity in a router placed at the customer’s premises. ISPs measure usage on the network, not at the particular customer premise, which leads to a number of conceptual differences in these measures, particularly with respect to timing and to the overhead associated with frame bytes which are stripped off *en route* or in the router. There is a substantial amount of processing required to reconcile these differences whose accuracy is not a subject of this report. For our purposes, we assume that each hour there is a measurement by the ISP (Ibytes) and a measurement at the customer’s router (Rbytes) which in principle are expected to be identical, but in practice differ. In each hour h , therefore, for each sampled router r , there are three important quantities:

$$R_{\text{bytes}_{rh}}, I_{\text{bytes}_{rh}} \text{ and } E_{rh} = (I_{\text{bytes}_{rh}} - R_{\text{bytes}_{rh}}) / R_{\text{bytes}_{rh}}$$

Note that the error term for a particular router in a particular hour E_{rh} , is calculated from the perspective of the router measurement, so that a positive value indicates an overmeasurement by the ISP and the value is the percentage overmeasured.

While the underlying data operates on an hourly basis, the critical metrics work on the aggregate monthly accuracy, which is simply uses the sum of $Rbytes_{rh}$ and $Ibytes_{rh}$ across the month, and a calculation of an aggregate monthly error using these quantities.¹

The aggregation of errors at a monthly basis corrects most of the timing issues which create divergences between the time at which data is sent or received at the router and the time at which it is recorded by the ISP. However, the first hour of the month is generally discarded because of these issues. In addition, other hours are discarded when there are various mechanical issues, such as the rebooting of a router, power failures, etc. All of these issues are outside of the scope of this methodological review, which assumes that we have, however derived, a series of *valid* hourly observations by customer of $Rbytes$ and $Ibytes$. NERA has been taken through the process used to discard particular hours as nonvalid. These procedures certainly seem superficially valid but we have no particular expertise in evaluating these procedures. From a statistical perspective, they are an attempt to excluded data contaminated by processes other than those which are intended to be measured. To the extent they succeed, they should be used.

III. Metrics

After examining the spreadsheets that NetForecast compiles to make the measurements they report to their client ISPs. Fundamentally, they simply calculate the monthly error rate for each sampled customer. Additionally, some additional trimming is used for significant outlier hours which are deemed valid by the basic processing but nonetheless likely to be contaminated. These trimmed hours do not contribute materially to the monthly average errors.

The basic metrics reported to the customer are: overall error rate, fraction of ISP measurements which are high and fraction of ISP measurements which are either above or below an exceedance threshold say 1%. Overmeasurement by ISPs is deemed a worse problem than undermeasurement because data caps are employed by ISPs which only affect customer behavior and service when they are exceeded². Thus, only when ISPs overestimate usage *and* the estimated usage exceeds the cap, while *actual* usage is below the cap is there a potential billing issue. The current reporting does not distinguish between overmeasurement above data caps and

¹ This will be different than simply averaging the hourly errors E_{rh} , since the denominator weights $Rbytes_{rh}$ are not constant.

² While overmeasurement is deemed more serious, NetForecast collects and reports separately on exceedance levels above and below sampled customer usage.

overmeasurement below caps, or the frequency with which usage is measured over the cap while actual usage is under the cap, although the data exists to provide these calculations should they be requested.

IV. Analysis

A. Estimation of Metrics

The mean calculations performed by NetForecast are correctly performed. In essence, each metric is simply averaged across the sampled ISP customers. This provides an unbiased estimate of the population quantities so long as the sample can be taken to be random.

In fact, NetForecast has no idea whether the samples are taken at random or not. For most of their customers, the ISPs themselves solicit participation in the NetForecast process, which involves the willingness to have a NetForecast router (with measurement software) installed at the customer's premises. Some ISPs use the company's own employees for this purpose which raises numerous questions about representativeness.

At our request, NetForecast has begun the process of providing customers with guidelines on creating statistically representative samples. A statistically representative sample has two features: first, every customer has some probability of being sampled and second, that probability of sampling can be estimated. If that sampling takes place at random, then simple averaging will give unbiased estimates of the underlying population probabilities. Should some ISP in the future produce sampling probabilities which are not uniform, weighted averages will have to be taken in which the weights reflect the probability of sampling: higher sampling probabilities get lower weights.

NetForecast has created several levels of potential response to the sample solicitation process. These responses are in two dimensions: representativeness and size. Representativeness deals with both geographic coverage and having sampling probabilities for each customer. Sample size deals with reliability, to which we turn next.

B. Reliability

While the current NetForecast methodology gives an unbiased estimate of the metrics they report, we are generally unsatisfied only to get an unbiased estimate; we would also like some measure of the reliability of that estimate; how likely is it that the true value is outside some defined threshold? In general, the more observations we have, the more reliable the estimate, but statisticians have made the concept of reliability more precise through the use of two concepts: the *confidence level* and the *margin of error*. Reliability estimates require both, as well as sample sizes and, in the case of population ratios (as all but the aggregate error rate are) base

rates for the metric. NetForecast has prepared and disseminated tools to enable its customers to choose sample sizes commensurate with particular confidence levels and margins of error. At the lowest levels, they warn clients that results can be considered for information purposes only. At the highest levels, they assert that the results will, over time, have statistical validity.

The “over time” statement above is important. In order to achieve statistically reliable measures, customers need to be measured for a number of months in order to achieve sample sizes large enough to achieve statistical validity. This assumes that the results for a particular household sampled are statistically independent from month to month. We have examined a sample of households and find the error rates to be largely uncorrelated from month-to-month, which suggests that this assumption is reasonable.

We can think of all these connected together: when we use an unbiased procedure to estimate a population quantity, the probability that the true population quantity is within some given threshold of the estimated quantity is a direct function of the base rate expected and the sample size.

To make things concrete, suppose we have 20 sampled customers (sampled at random from the entire population) and we estimate the mean error rate at -0.3% , that is the mean customer is slightly undermeasured³. What is the probability that the entire population has an average usage out of specification high; in other words, although the average observed is -0.3% , what is the probability that the true average, were we to measure everyone, would be over $+1.0\%$ (1.3% higher than the observed rate in the sample)?

The simple way to make this calculation is to assume that errors are normally distributed in the population. We can then calculate the standard deviation of the error rates for the 20 observed samples and calculate the probability that a sample of 20 will be off by more than $+1.3\%$ ⁴. Say this probability is 5% . Then we have 95% confidence that the margin of error is less than $+1.3\%$.

V. Conclusion

NetForecast has a methodology which produces statistically reliable unbiased estimates of the accuracy of ISP metering systems for its actual population on the assumption that the samples that are given to it are representative of the population and sampled with uniform probabilities.

³ In the data we have seen, the mean customer is usually undermeasured for technical reasons which are not important for this discussion.

⁴ This probability will depend on the sample size; the larger the sample size, the lower the probability.

Report qualifications/assumptions and limiting conditions

This report is for the exclusive use of the NERA Economic Consulting client named herein. There are no third-party beneficiaries with respect to this report, and NERA Economic Consulting does not accept any liability to any third party.

Information furnished by others, upon which all or portions of this report are based, is believed to be reliable but has not been independently verified, unless otherwise expressly indicated. Public information and industry and statistical data are from sources we deem to be reliable; however, we make no representation as to the accuracy or completeness of such information. The findings contained in this report may contain predictions based on current data and historical trends. Any such predictions are subject to inherent risks and uncertainties. NERA Economic Consulting accepts no responsibility for actual results or future events.

The opinions expressed in this report are valid only for the purpose stated herein and as of the date of this report. No obligation is assumed to revise this report to reflect changes, events or conditions, which occur subsequent to the date hereof.

All decisions in connection with the implementation or use of advice or recommendations contained in this report are the sole responsibility of the client. This report does not represent investment advice nor does it provide an opinion regarding the fairness of any transaction to any and all parties.

About NERA

NERA Economic Consulting (www.nera.com) is a global firm of experts dedicated to applying economic, finance, and quantitative principles to complex business and legal challenges. For over half a century, NERA's economists have been creating strategies, studies, reports, expert testimony, and policy recommendations for government authorities and the world's leading law firms and corporations. We bring academic rigor, objectivity, and real-world industry experience to bear on issues arising from competition, regulation, public policy, strategy, finance, and litigation.

NERA's clients value our ability to apply and communicate state-of-the-art approaches clearly and convincingly, our commitment to deliver unbiased findings, and our reputation for quality and independence. Our clients rely on the integrity and skills of our unparalleled team of economists and other experts backed by the resources and reliability of one of the world's largest economic consultancies. With its main office in New York City, NERA serves clients from more than 25 offices across North America, Europe, and Asia Pacific.

NERA

ECONOMIC CONSULTING

NERA Economic Consulting
1166 Avenue of the Americas
New York, New York 10036
Tel: 1 (212) 345-3000 Fax: 1 (212) 345-4650
www.nera.com