Speedtest Intelligence Methodology

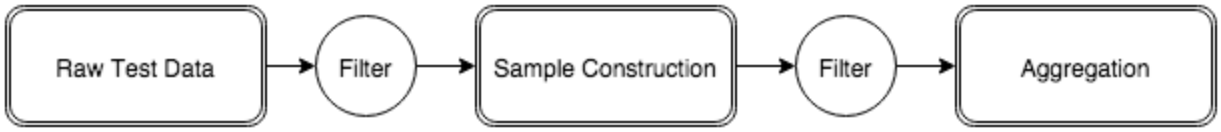
1. Introduction

With over 9 million tests each day, Speedtest has the most comprehensive view of worldwide internet performance. We have worked diligently to devise the most accurate method for determining the fastest ISPs and mobile networks around the world.

The purpose of this document is to outline the methodology used for our Speedtest Intelligence product. We believe we have constructed an unbiased and equitable approach that ensures stable results, controls for extraneous variables, and ensures every internet user gets a fair number of “votes”.

2. Methodology Overview

In general, our methodology across platforms (web, mobile, etc.) flows as follows :



A. Raw Test Data

These are raw test results taken by Speedtest users from individual platforms. Two platforms are currently represented in our raw test results: Broadband and Mobile.

Broadband
Broadband results from ISPs include any result taken on Speedtest.net, from any supported browsing technology and/or operating system. Each broadband result comes parceled with data covering several categories of interest for constructing samples for the Awards rankings and Speedtest Intelligence metrics, such as IP addresses, GeoIP Information, download speed, upload speed, and latency, among others.

We combine Google Maps, Maxmind, and an internal GeoIP system to select the most accurate location based on the IP address a test was taken from. GeoIP systems have wide use specifically for this purpose. Our use of three parallel systems ensures as accurate of a location fix as is currently possible.

Mobile
Mobile results include Android and iOS results taken on their respective Speedtest application. Each mobile result comes parceled with data covering several categories of interest for constructing samples for the Awards rankings and Speedtest Intelligence metrics, such as

MNC/MCC codes, GPS location, IP addresses, connection types, download speed, upload speed, and latency, among others.

We rely mostly on GPS coordinates to get a latitude and longitude coordinate for a device at the time a test was taken. However, when required, we fall back on GeoIP systems. An example of this would be when a user disables location services. To match with a location identifier (city, region, etc.), we do a reverse-lookup using great-circle distance as our measure of distance.

B. Sample Construction

"Sample Construction" enables us to create standardized data points that can be used for further statistical analysis. In general, the goal in creating samples is to reduce the effect of extraneous variables, and ensure every user gets fair representation in the final analysis, as well as to prevent gaming of our ranking systems. Without a sample process, for example, a power user that takes 100 times as many tests as a casual user would have their correlated speeds weighted higher in many of our final aggregates.

There are many ways to create samples given our data, however we employ a method we believe is the most representative of what speeds each user would experience on a day-to-day basis.

We apply several filters to ensure the raw data is clean, including filtering out "spoofed" web results or removing mobile results with different pre- and post-test connection types. Then we remove unwanted data dependencies, which limits the effect of extraneous variables, and ensures each user gets fair and relatively equal representation in the data.

A single "user-sample" is constructed straight from a database table containing raw test results by using a query language. Each sample is considered one group. In practice we are producing user-samples defined as "mean speed for User X per day using provider Y in location Z." The location could be a city, region, or country. That means we combine our data by several key identifiers and produce a few metrics that are aggregate values, such as the mean download speed, or mean upload speed for those key identifiers.

The main difference between a mobile user-sample, and a broadband user-sample is in how we determine the user identifier on each platform. Additionally, mobile platforms report their connection technology whereas this must be inferred or surveyed on for broadband results. All other sample fields are the same or very similar (e.g. mobile provider ID vs. broadband provider ID).

A mobile user's test would have the user identifier determined by the unique device identifier. A broadband user's test would utilize a combination of the IP-address and the session identifier. Session identifiers for users who log in to Speedtest.net are always the same.

C. Aggregation

The aggregation step occurs after the user-samples are constructed, and involves creating data subsets and computing various statistics on these subsets that can be used for tracking progress or comparing providers and/or locations to one another. The thresholding and statistics used are designed to capture some attribute or information that can be used for comparing these data subsets to one another in a user-friendly fashion.

A “host distance” filter is applied to the user-samples before computing most of our statistics. This is used in order to prevent sabotage of competing provider rankings, as well as to control for lower speeds achieved by users that tend to test frequently over longer distances.

From our mobile samples, we compute the 95th percentile “Average Distance to the Host” from our user-sample data, for each country, for the time period being analyzed. This is then used as a “host distance” threshold for user-samples from that country before aggregation can commence. In practice this removes ~5% of samples from the data set while giving an equitable boost to metrics for each provider. For broadband user-samples, the same methodology is followed with the exception that the 90th percentile “host distance” is used in the calculation.

Note that this threshold is applied to the samples using the “average distance to the host” field, rather than on individual tests. By design, it’s possible for users to test to far away hosts and still have their user-samples pass through the threshold so long as they don’t test to far away hosts too often in a single day.

The host-distance threshold is never less than 100 km, and the maximum allowable host-distance threshold is 2000 km. Otherwise, the 95th and 90th percentiles are used respectively so long as they fall between those bounds for the country and time period of interest.

We aggregate the samples by time frame, provider and location into statistics that best allow us to evaluate markets and providers.

Provider and Location
Provider information is available on a city, region and country level and allows the providers to compare their performance with their competitors in each location subset.

Additionally, we are able to analyze the performance information of a given city, region, and/or country, without consideration to individual provider performance. For these aggregates, we use location information without showing a specific provider and compute the aggregates based on that location. This is useful for comparing the speeds an average user achieves location-to-location.

Modern Devices
For mobile results, outside of technology types used to identify the mobile connection, we have created a special mobile market type called “Modern Devices” to aid in analysis of mobile user’s typical speeds--those achieved most commonly throughout a country. This market type includes all cellular tests, regardless of connection technology used, as long as they were taken on devices that are identified as being capable of achieving the fastest speeds commonly-available in the country in question.

The definition of “Modern-Market type” changes country to country. What is considered modern in some countries may not be available in less developed countries. We have an automated mechanism to determine modern-device classification requirements on a country by country basis. For highly-developed countries a “Modern-Device” is likely to be synonymous with “LTE-capable”, however in less developed countries it’s possible that 3G or even 2G are the most modern technology categories available to their consumers.

In practice, this market type is a subset of all mobile technology types. However, we do not include samples from device-versions that are considered out-of-date, or incapable of connecting to the fastest commonly-available networks within a country.

D. Awards

For awards, we review and categorize every ISP and carrier that appears in the Speedtest Intelligence data to ensure that they are still active in the marketplace, and offer residential services with access to internet for end users.

We require that a provider remain an individual entity during the entire award period. If two providers merge, or one is acquired, we will combine the test results of these providers moving forward from the merger/acquisition date. Our requirement here may disqualify any absorbed providers from receiving an award, however the parent provider would not have any such restrictions and in fact a merger/acquisition could push the parent provider upwards in ranking.

In addition, we require providers receiving awards to have “Top Provider” status, meaning they have a significant share of users (at least 3% of total samples) in their respective market. Finally, we also determine whether there are any ties between providers in a market to ensure that we are determining clear winners based on what would be a material difference in service to the end user.

Fastest ISP
It is common for ISPs to offer multiple subscription plans to consumers at different price points, each resulting in a different level of performance. The goal of the Speedtest Award is to determine the fastest ISP regardless of actual consumer subscription tier. To achieve that top-level view, the Speedtest Award determination is made by analyzing the fastest generally-available speeds by each provider. We use the 90th percentile download performance

for each ISP's Speedtest results to represent the experience of consumers who are subscribed to the fastest service tiers available from a given provider.

Fastest Mobile Network

Performance on mobile networks is generally not tied to a specific subscription plan as compared to fixed ISPs. Mobile networks are impacted by the devices that individual consumers use, with many older devices only being able to connect to older, slower technologies such as 2G and 3G. To determine the Speedtest Award for mobile networks, we use only results taken from "modern devices" capable of connecting to the market's fastest, generally available technology. As a result, Speedtest Awards for a mobile network are calculated using average download speeds achieved on modern devices. The results include all tests taken from modern devices, regardless of the actual connection technology used during a test. This methodology also provides a view of overall network performance across an entire award geography, not simply the performance in highly-developed areas.