



Transforming Set-top Box Data into Audience Measurement Data

Set On/Set Off Editing Guidelines

Andrew Roberts
Antonio Carvalho

September 2009

Introduction

From our experience at TNS of processing Set-top box (STB) data over several years, it is clear that raw data straight off the STB cannot be used without addressing one fundamental issue:

Data extracted from digital STBs track events that occur on the STB itself but not on the connected television set. On occasion, viewers may turn the television set off whilst leaving the STB on or in standby mode. In this situation, the STB will report long viewing sessions which are not genuine and the processing of STB needs to account for such occurrences.

TNS has designed, developed and implemented Set On/Set Off (SOSO) editing guidelines to resolve the situation described above. Since initial implementation five years ago, the TNS SOSO editing guidelines have undergone a series of enhancements to further refine the process so that it more accurately reflects the true viewing status.

The latest version of the TNS SOSO guidelines incorporates four key characteristics which determine whether an STB tuning session should be edited and if so by how much. These key characteristics have been identified through analyses of STB and TV set (on/off) data. They are:

- the STB tuning session length
- the time of day
- the station tuned
- exceptional events

The TNS SOSO guidelines deal with both Type I and Type II errors. Type I errors occur when the session is too long (as a result of continuous STB feedback while the TV is off) and should be edited. Type II errors occur when the session is long but the TV is on and should not be edited.

STB tuning session length

The longer the STB is tuned to the same channel the greater is the chance that the TV set will have been turned off before the end of the tuning session. The probability that the STB tuning session includes time when the TV set was off progressively increases with the length of the tuning session.

While STB tuning session length is a key factor, the time of day also influences the probability of STB on-TV set off. Not surprisingly, breakfast and late night are the times when it is most likely that the TV set has been turned off as viewers are either going to work or going to bed.

Station tuned

STB tuning lengths also depend on the type of channel being viewed and this factor has to be considered before determining whether a particular STB session should be edited. Longer tuning sessions tend to be associated with the most popular channels, while news and music channels typically have much shorter sessions. The most extreme example is the EPG where real viewing sessions are very short indeed.

Special programming events

Occasionally programs are broadcast which attract large audiences for an extended period of time. Major programming events such as the Super Bowl or Academy Awards are prime examples. In these cases the normal editing procedures would result in erroneous editing very long tuning sessions where genuine (TV set on) viewing was taking place. The TNS SOSO editing guidelines consider the level of channel tuning versus each station's norm in order to detect such extremes in viewing behavior and thereby override the normal editing rules.

The process operates automatically by calculating for each channel the average number of tune-away events occurring in a given interval of time (the station's norm). If this norm is exceeded by a statistically significant margin, normal editing is suspended and viewing is deemed to continue until the actual tune-away event occurred.

TYPE I ERRORS

A Type I error occurs when:

- 1) The STB is left on generating long viewing sessions
- 2) The TV is turned off
- 3) The viewing statement is not edited

The presence of a large number of long viewing session lengths when the TV set is off can significantly influence TV ratings. The best way to analyse the distribution of the viewing session's length is to determine their frequency over the entire sample and display the results in histogram form. Below is an example of this type of plot, showing a significantly large number of short viewing sessions, with the frequency decreasing as the session length increases.

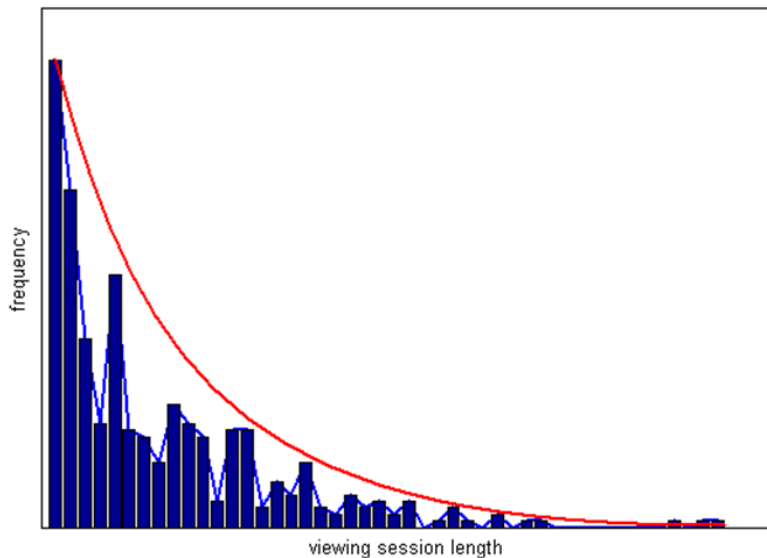


Figure 1 – Example of a viewing session length histogram.

Comparing the session lengths of set top box and People Meter data shows the presence of a significantly higher number of very long viewing sessions, which need to be adjusted via set on/set off editing.

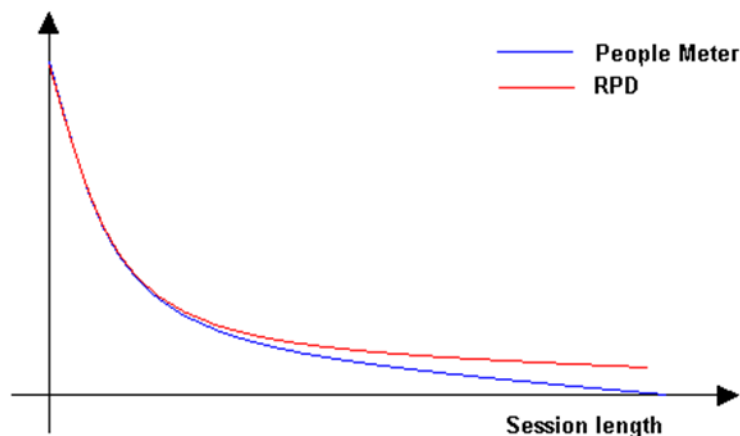


Figure 2 – Comparison between viewing session length for PeopleMeter and RPD data

TNS set on/set off editing guidelines have been developed to deal with these situations, and have gone through a series of improvements as described below.

- **Simple Editing**

The most basic implementation is to edit the set-top box data in such a way that very long sessions are capped after a given interval. In its simplest form, the TV is assumed to be off if the channel remains unchanged for a long period of time. The length of this time period is set by the capping level.

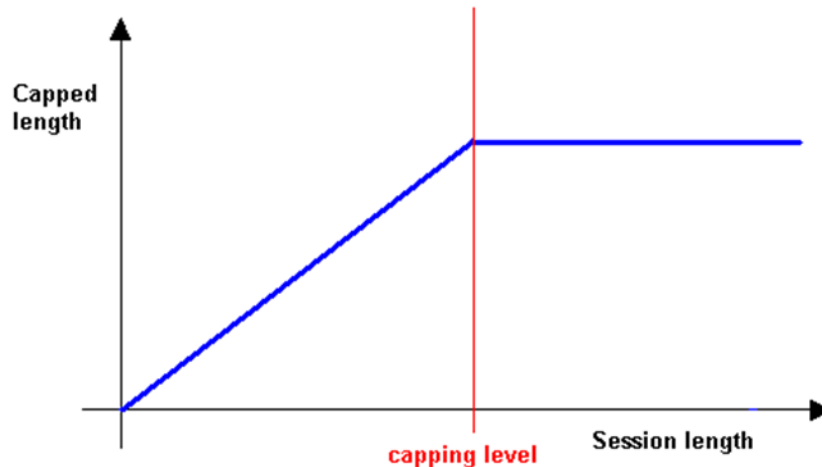


Figure 3 – Example of applying simple editing to longer viewing sessions. The editing level determines the maximum length allowed for viewing sessions.

The editing is done by daypart and also by channel or channel group. The editing level is adjusted with the aim of matching the average number of hours of tuning per home for each combination of daypart and channel group to known and accepted levels of overall TV viewing.

The first implementation of TNS’ simple editing guidelines resulted in viewing sessions that on average were no longer than 45 minutes. This approach strikes a balance between average viewing sessions and very long viewing sessions.

The two major drawbacks are having a fixed maximum session length, with the undesired effect of creating ripples in the data, and generally not allowing viewing for at least 60 minutes (the length of a large number of programs).

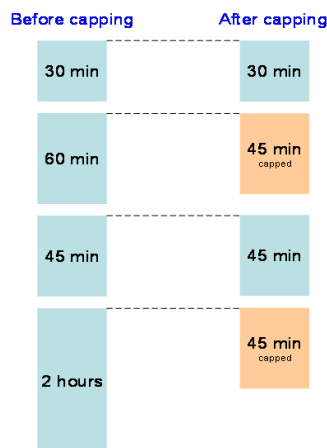


Figure 4 – Example of how the simple editing works, with the editing level set to 45 minutes.

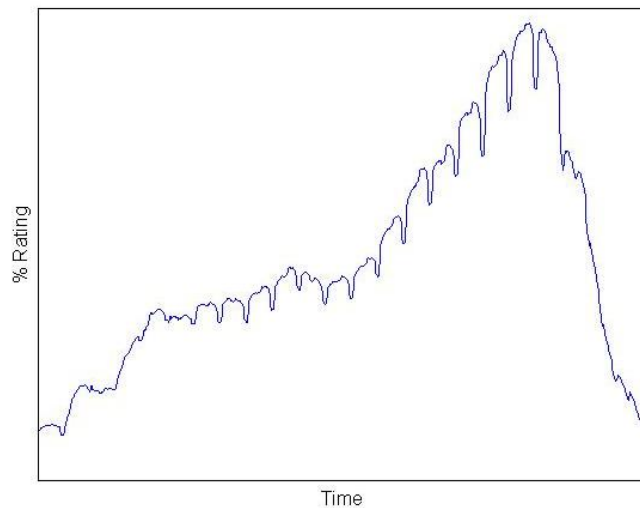


Figure 5 – Illustration of the possible “ripple effect” on the viewing data during the day with the introduction of simple editing.

• Progressive Editing

To address the major drawbacks of Simple Editing TNS developed Progressing Editing, which is the approach used in the US today. Our aim was to develop set on/set off editing guidelines that did not require specific channel knowledge and could, solely on the basis of the available STB data, determine when a STB statement requires capping and by how much. Progressive editing was developed based on understanding gained from significant research done on the UK BARB data to understand when the data needs to be capped.

In Progressive Editing, editing levels are proportional to the length of the raw viewing statement, i.e., the longer a viewing session is, the greater the probability that it will be edited. A maximum statement length is defined, and for each statement a probability of “need for editing” is calculated.

Using a highly effective and verifiable probabilistic method, editing parameters in Progressive Editing are optimised so the resulting data achieves similar values of total TV Ratings to known and accepted levels of overall TV viewing. Once the editing parameters are determined, channel indices are applied that reflect the differences between viewing patterns to the different channels. These channel indices are self referencing and, unlike Simple Editing, do not need to be maintained as they are very stable over time.

TYPE II ERRORS

Although Progressive Editing works well, there are certain programmes that, by their nature, last longer and lead to longer viewing sessions.

A Type II error occurs when:

- 1) *The STB is left on generating long viewing sessions*
- 2) *The TV is turned on*
- 3) *The viewing statement should not be edited.*

Some programming examples of this would include regularly scheduled sporting event, special programming events like the Super Bowl, Academy Awards, etc. or major breaking live news. To address Type II errors, TNS developed Dynamic Editing.

One way to determine whether long tuning events are legitimate is to examine the number of statement ends during the day. Long viewing sessions that are legitimate will have an atypical high number of statement ends when they finish (Figure 6).

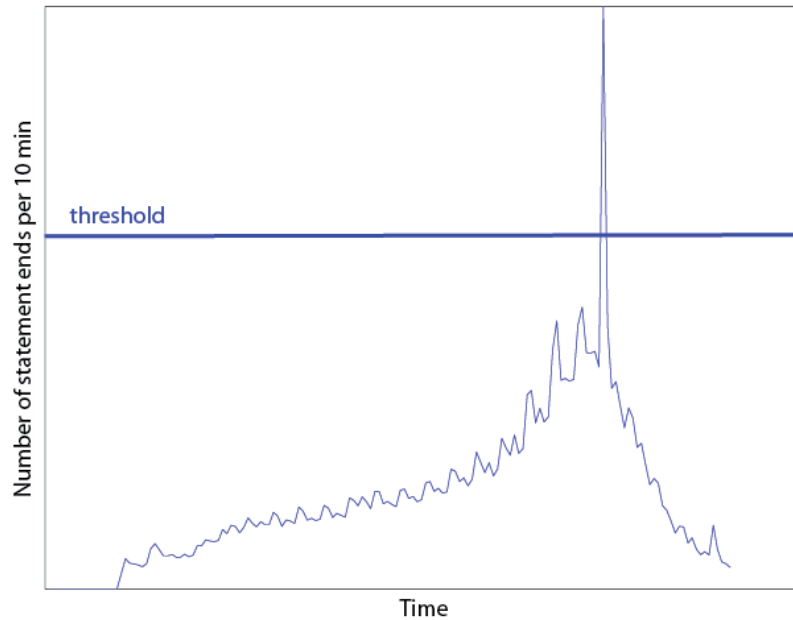


Figure 6 – Example of how to identify legitimate longer viewing statements, using the number of statement ends over the day period.

A statistical threshold is defined above which a certain program will have a longer duration, and therefore no set on/set off editing will be applied.

The need to deal with the issue of set on/set off has long been recognized by TNS and as shown here, our procedures have evolved through ongoing investment and research on the subject. Today we believe the state-of-the-art approach pioneered by TNS results in the most accurate estimate of viewing levels based on STB data in the industry. Our set on/set off editing guidelines not only deal with the well publicized Type I error, but importantly the lesser publicized Type II error. Further information on the subject is available on request.

© 2009 TNS Media Research.

About TNS Media Research

TNS Media Research is the leader in providing digital TV audience and advertising measurement tools and intelligence derived from second-by-second return path data. With a proven track record in audience research services across 32 countries, TNS Media Research is at the forefront of technological developments in audience measurement in the media industry, working with broadcasters and advertisers, media agencies and industry committees. TNS Media Research is a Kantar Group company. For more information on TNS Media Research and DIRECTView, visit www.tns-mr.com.

For more information, contact:
BUD BREHENY
 (212) 991-6080
bud.breheny@tns-mr.com