



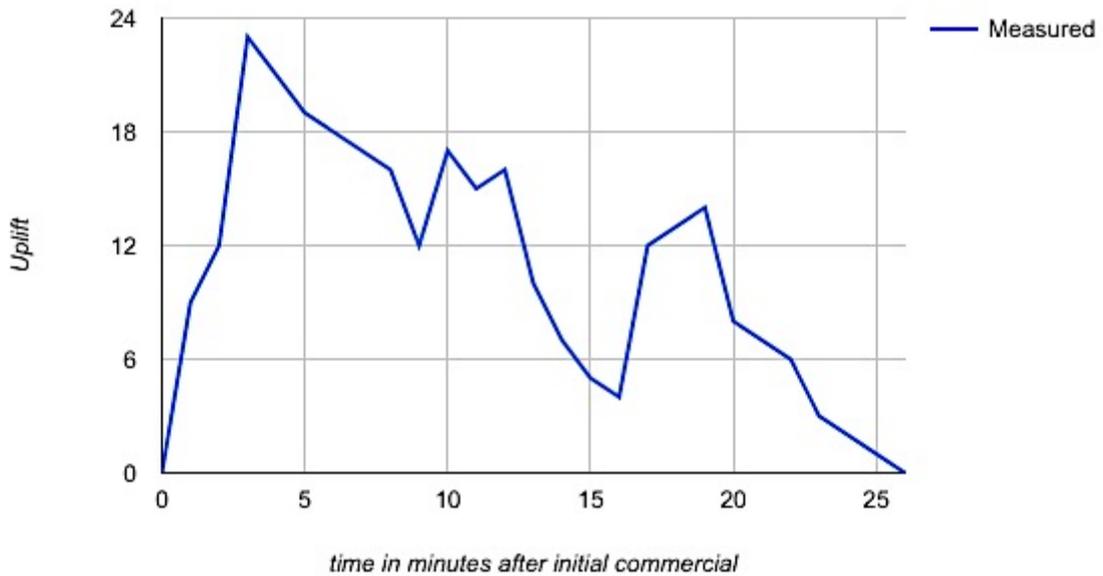
Blog post series: Overlapping TV commercials

This article about overlapping TV commercials belongs to a special series of blogposts, written by our own data wizards. It will offer you a glimpse into the engine room of Mediasynced. In these informative blogposts, we shed a light on the complexity of TV performance measurement in realtime and our robust statistical solutions.

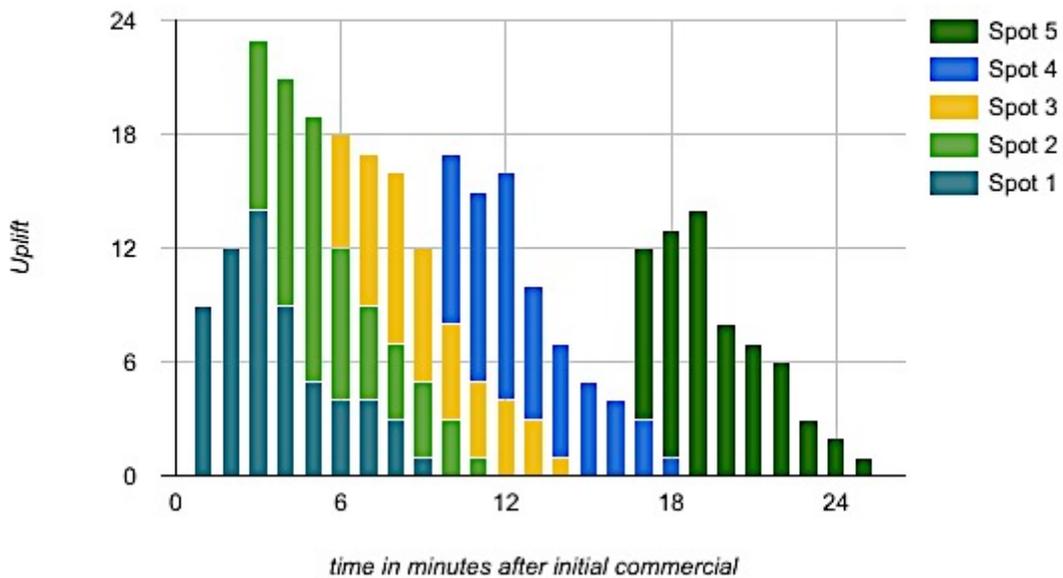
” The measured uplift of the overlapping part needs to be divided.

Reality is always more complicated than theory. From a data analysis standpoint, in an ideal campaign the commercials would be evenly divided between the time. A commercial would only start when the effect of the previous commercial has faded. This way the attribution model for each commercial could be kept very simple, everything above the baseline would be attributed to the commercial. Yet reality is quite different; it is common to see two or more commercials start at the same time or within minutes of each other. There are even situations where more than five different commercials start close to each other. In these kinds of situations, how would you attribute the total uplift to each commercial?

What we see.



What is actually happening



The simplest solution is to ignore these datapoints. You consider them to be outliers and hope that their removal does not negatively affect your analysis. In our opinion this is not an option. These situations are too common to ignore. In some campaigns almost 50% of all commercials overlap with at least one other commercial. This may be due to a deliberate 'TV-roadblock' strategy. Ignoring these data points would mean that you could not evaluate the effect of this option.

The solution that most people would offer when confronted with this problem is to attribute the uplift based upon a linear, time based model. While this is better than ignoring these points, it still has its own share of problems. It does not take the multiple factors into account that influence uplift. Some commercials have a significantly higher audience, or are being aired on better performing channels. Or have a significantly different response curve compared to the others. These and other factors are ignored in this model which could result in misleading conclusions.

We tackle this problem using a combination of models. When two commercials overlap, they create unclean data (the overlap). The measured uplift of the overlapping part needs to be divided. In most cases the overlap is partial, meaning there is a clean beginning of the first spot and a clean end of the second spot. We use the clean bits of data to predict the expected proportion from the unclean data. The prediction also takes the response curve into account, as well as the amount of clean data. We then validate the outcome with a set of statistical models, including machine learning models that are continuously being trained using historical data to calculate the most accurate attribution per spot.

Given the importance and complexity of attribution in the case of spot collisions, we have gone to great lengths to create the most robust attribution model for spot collisions.