The Institute for Perception

For technical reports from The Institute for Perception or for information about short courses, please visit www.ifpress.com or email us at mail@ifpress.com

Reprinted from IFPress (2016) 19(2) 3,4

eTURF 2.0: From Astronomical Numbers of Portfolios to a Single Optimum

John M. Ennis and William J. Russ

Background: TURF (Total Unduplicated Reach and Frequency) is a commonly used market research tool based on the media concepts of *reach* and *frequency*. In the original application of TURF, media schedulers wanted to maximize the number of people reached by and/or the frequency of individual exposures in a media campaign[1,2]. In this application, TURF was used to select the optimal set of media elements. When used for market research applications, however, TURF was typically used to find optimal combinations of items within a portfolio, with the goal of reaching as many consumers as possible with at least one item in the portfolio. While TURF was very popular in the late 1990s and early 2000s[3,4], its popularity faded as it became apparent that the original TURF algorithms had difficulty running on large datasets in reasonable time. In this report, we review how these computational limitations have been removed with recent advances in discrete mathematics, and demonstrate how these advances help to determine the optimal size and composition of a portfolio.

Scenario: You work for a flavor supplier and you have been assigned to create a family of fruit juice flavors from which your primary client will draw. Future applications include the development of a new line of sparkling fruit juice drinks. For this assignment, you have a set of 50 possible fruit juice flavors, and you would like to find a portfolio optimized to reach as many respondents as possible. In order to proceed with your project you decide to use TURF analysis.

Attempts at Direct Enumeration: Starting with the basic idea of treating your task as a Venn diagram problem, you consider the reach of various flavor combinations using direct enumeration of these combinations. While this approach is feasible for smaller numbers of items, such as sets of five items, you find that combinations with a larger number of items yield too many possible combinations for enumeration to be practical. This fact is clearly illustrated in Table 1 where even to consider twelve flavors out of fifty would require evaluation of over 121 billion combinations. And worse, you compute that the total number of portfolios, 2^{50} - 1, is over 1.1 quintillion. Using direct enumeration would require a minimum of one year of computation to find a solution to your problem.

Goal	Combinations
8 out of 50	536,878,650
9 out of 50	2,505,433,700
10 out of 50	10,272,278,170
11 out of 50	37,353,738,800
12 out of 50	121,399,651,100

Table 1. Number of possible combinations as a function of combination size.

eTURF 2.0: Difficulties encountered in handling large numbers of items using complete enumeration were first handled using heuristic approaches that provided useful solutions[5]. One insight, discovered in 2012, that reduced the number of item combinations to be considered, focused only on those combinations necessary to compute using the principle of "Non-Synergy."[5] According to this principle, a combination can never reach more respondents than the sum of the *reaches* of any pair of sub-combinations comprising it. Although this approach, called eTURF[6], enormously expanded the number of items in combinations that could be considered, an even more dramatic improvement closely followed. This improvement, due to Daniel Serra, used techniques from binary linear programming[7,8] to solve TURF-type problems. We have developed a new staged approach using these techniques to allow astronomically large problem sizes to be solved relatively quickly. In what follows, we refer to this staged approach as **eTURF 2.0** to distinguish it from previous versions of TURF analysis.

eTURF 2.0 Application to the Flavor Portfolio Project: In order to begin your project, you conduct a large-scale, nationwide internet-based survey. In this survey, 1800 respondents are asked to rate each of the 50 flavors on a 7-point purchase interest scale. A preliminary analysis reveals that 1601 respondents out of 1800 gave at least one flavor a "6" or a "7". In what follows, we restrict attention to these 1601 respondents who gave a top-2 box to at least one product, but all analyses could be duplicated with a stricter or more lenient definition of *reach*.

The first step in your eTURF 2.0 analysis is to determine the minimum number of flavors in a portfolio that will reach all of the reachable respondents. You find that the smallest portfolio that reaches all of these respondents contains 32 flavors. This stage of the analysis completes in only 0.04 seconds, and sets an upper bound on the size of the future searches. Informed by your first stage, you proceed with the next stages to determine 1) the maximum overall *reach* of portfolios for each portfolio size between 1 and 32, 2) the maximum *reach frequency*, and 3) the maximum *reach penetration* to find a single optimal portfolio. The entire analysis takes just over one minute to complete, with the time to complete the search for each size portfolio and for each stage of the search shown in Figure 1.

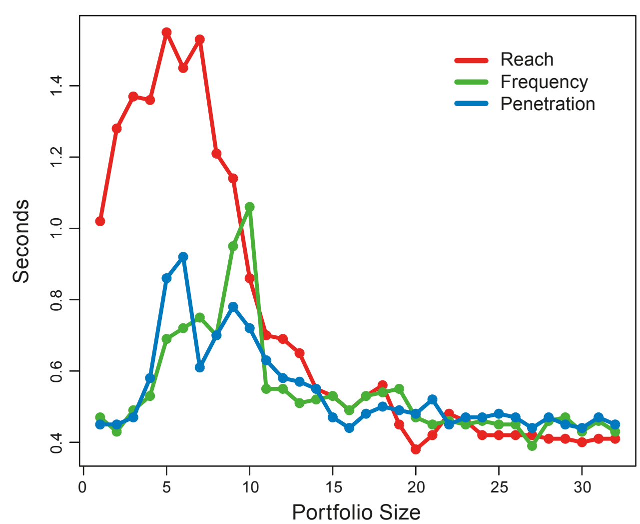

Figure 1. Time to complete the main stages of the eTURF 2.0 analysis as a function of portfolio size.

The *reach* values of these optimal portfolios vary from a maximal *reach* of 1116 respondents for a single flavor by itself to the maximum overall *reach* of 1601 respondents for optimal portfolios of size 32. These *reach* values are shown in Figure 2 together with total *reach frequencies* for each portfolio size. Similarly, Figure 3 depicts the mean *penetration* of the flavors in the optimal portfolios for each portfolio size expressed as the harmonic mean percent of consumers reached per product.

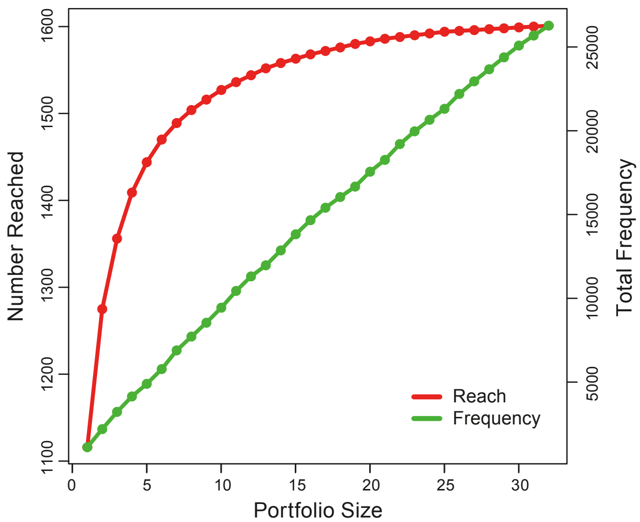

Figure 2. *Reach* percentages and total *reach frequencies* of optimal portfolios as a function of portfolio size.

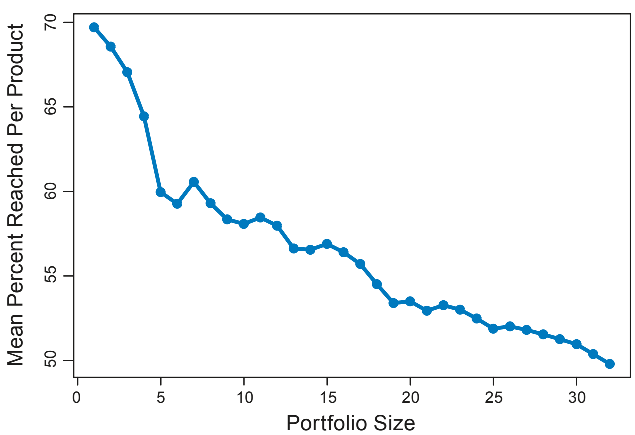

Figure 3. Harmonic mean product *penetration* within optimal portfolios, as a function of portfolio size.

Finally, to determine which size portfolio to recommend, you conduct a series of statistical tests. You consider, for each portfolio size, whether the optimal portfolio of one size larger has statistically greater *reach*. For your statistical test, you conduct a McNemar's test on the counts of respondents reached by the smaller portfolio but not the larger, and vice versa. The results of these tests, together with the differences in the percentages of the respondents reached by the variously sized portfolios, are depicted in Figure 4. Based on this analysis you decide on a portfolio size of 12, which reaches 97% of the reachable respondents and is not significantly different from a portfolio size of 13.

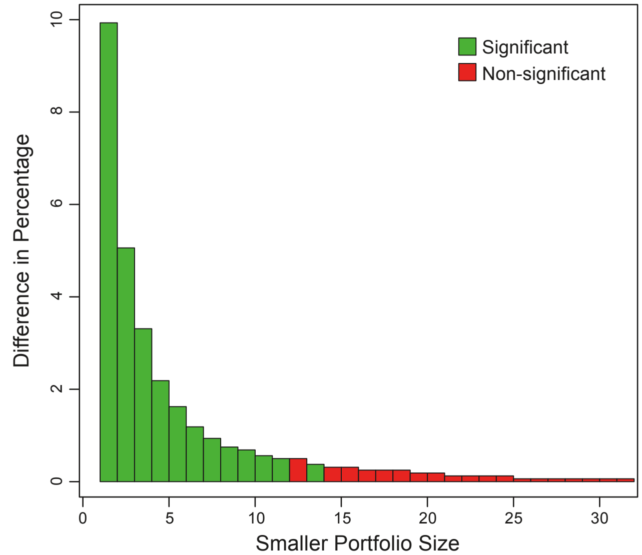

Figure 4. Differences in reach percentages between optimal portfolios of consecutive sizes.

Conclusion: TURF, once a fashionable market research instrument, fell out of favor because of its difficulty in handling large problem sizes. More recently, however, TURF has come back into favor as advances in algorithmic theory help to solve this difficulty. By applying a staged approach to create a more efficient TURF technique, such as eTURF 2.0, it is now possible to quickly determine the minimum number of products needed in a portfolio to reach as many respondents as possible and to determine high quality portfolios of smaller sizes. In addition, statistical tests can be applied to check whether solutions of smaller sizes are significantly different in quality from solutions of larger sizes. These developments remove a limitation on TURF analysis which provides value in the search for optimized portfolios.

References

1. Miaoulis, G., Parsons H., and Free, V. (1990). TURF: A new planning approach for product line extensions. *Marketing Research,* **2**(1), 28-40.
2. Cohen, E. (1993, June). TURF analysis. *Quirk's Marketing Research Review,* 10–13.
3. Krieger, A. and Green, P. (2000). TURF revisited: Enhancements to total unduplicated reach and frequency analysis. *Marketing Research,* **12**(4), 30–37.
4. Markowitz, L. (2005). *Going beyond TURF to complement and extend existing product lines* (Technical Report, November 2005). Ipsos Insight.
5. Adler, T., Smith, C., and Dumont, J. (2010). Optimizing product portfolios using discrete choice modeling and turf. In S. Hess and A. Daly (Eds.), *Choice modelling: The state-of-the-art and the state-of-practice: Proceedings from the inaugural international choice modelling conference* (pp. 485–497). Bingley, England: Emerald Publishing Group Limited.
6. Ennis, J. M., Fayle, C., and Ennis, D. M. (2012). eTURF: A competitive TURF algorithm for large datasets. *Food Quality and Preference,* **23**(1), 44–48.
7. Nemhauser, G. and Wolsey, L. (1988). *Integer and combinatorial optimization.* New York: Wiley.
8. Serra, D. (2013). Implementing TURF analysis through binary linear programming. *Food Quality and Preference,* **28**(1), 382–388.