

大数据时代的广播数据挖掘和应用

薛 未¹

我国广播广告从无到有，从1982年的2000万元到2012年250亿元的广告花费，可以说实现了从量到质的飞跃。回溯历史，新中国第一支商业广播广告诞生于1979年，广告主为上海家化的“春蕾”发乳。广告播出后，春蕾发乳从滞销转为畅销，风靡一时，成为众人心中的经典国货，中国广播广告在经历了10年的蛰伏之后迎来了全新开端，并在随后相当长时间里独领风骚。但随着改革开放和现代化建设的进一步纵深，电视、报纸、杂志分化了广播的大部分听众，而新媒体的强势崛起也让传统媒体受到了较大的冲击，广播的光芒被掩盖了，广播广告进入了低迷期。到了21世纪，广播由于国家政策的强力扶持，同时也因为城市化进程的加快，更主要得益于私家车保有量的高速增长，广播成为了“车轮上的媒体”，广播广告也在滚滚车流中得到了复兴。远誉与中国新广播事业一起成长与壮大，一起经历传统媒体时代、新媒体时代以及融媒体时代，2013年我们将进入一个全新时代，那就是大数据时代。

大数据带来的信息风暴正在变革我们的生活、工作和思维，大数据开启了一次重大的时代转型。说到大数据，就不得不说到最近很火的一部美剧《纸牌屋》，其从诞生之初就深深刻上了“大数据”烙印，从内容、导演到主演都是大量数据收集及演算的产物，甚至在剧集播出中，观众快进或回放的时间点都会回传至数据中心，用来统计哪些情节是乏味的，哪些情节是精彩的，以便在后期的剧本撰写中迎合观众口味的考量。《纸牌屋》让很多人认识到了大数据的力量，但大数据到底是什么呢？这个新兴词汇拥有明确的学术化解释，但过于专业和繁复，远誉认为，只要满足以下4大条件就是大数据：第一是**高容量（Volume）**，第二是数据的多样性（Variety），第三是实时性（Velocity），第四是数据的价值（Value）。

从这个层面来说，CSM数据完全可以满足这四个V，也就是说，CSM的数据就是大数据。其实，我们无需太在乎身处哪个时代，我们更在乎如何最佳地使用数据，如何让数据产生可观的价值回报并可被验证和改进，数据再大，最终还是属于知道正确运用它的人。下文将分享，远誉作为业内专业的广播广告专家，如何利用中国最权威的媒介数据研究专家CSM提供的数据，迎接这个庞大的时代。

大数据时代，数据之崛起

BIG DATA

2013年是大数据元年，大数据带来的信息风暴正在变革我们的生活、工作和思维，大数据开启了一次重大的时代转型

Volume. Variety. Velocity. Value
高容量/多样/实时/高价值



CSM数据符合4V要素，从这个层面来说，CSM数据就是大数据

我们无需太在乎我们身处哪个时代，我们更在乎如何最佳地使用数据，如何让数据产生可观的价值回报并可被验证和改进，**数据再大，最终还是属于知道正确运用它的人。**



一、市场：多项指标综合考量

制定一个完整的策划方案，第一步是要确定市场预算划分。每个城市经济水平不同，消费水平不同，人口基数不同，媒体的价格也不同，所以不能将预算进行平均分配，而是分为重要的和次要的市场。可如果按照一线、二线及三、四线城市进行分配，也存在较大的偏差。

在对城市的选择上，远誉会从几方面考虑，如人口数，收入，汽车保有量、广播到达率及CPM等。人口数越多，市

¹ 作者来自远誉广播广告公司。

场潜在的目标消费者就越多；收入高则可以表明该城市消费能力强；汽车的保有量也从一方面体现广播收听人群的含金量；广播到达率则是该城市广播渗透率的明证。

再引入CPM的考量，核算广播投放于该城市的千人成本。通过整理上述多项指标的数据数值，再加权计算以得到具体数值，通过这个数值来确定哪些城市需要重点投放，而哪些城市只需维持基本的投放量。

如表1中所示，分值在70分以上的城市有城市A到城市F，以上城市可以作为重点投放市场。通过远誉市场划分计算模型，可能有些市场保持了一线城市的实力，而有些城市的排名却跌出第一梯队，还有一些二线城市却显示出强大的潜力，成为急需突破和重投的市场。远誉市场划分计算模型是建立在充分利用CSM数据基础之上将“城市实力”量化的过程，且在实际操作中远誉会植入更多具有实际参考价值的数据，如竞品投放量、品牌渠道渗透率、客户在该城市年度销售目标等。

表1 远誉市场划分计算模型



远誉市场划分计算模型

计算方法：

●以目标市场的人均收入、媒体千人成本及广播收听情况和市场人口数、及汽车保有量的指数加权所得

Rank	城市	人均收入 20%	千人成本(CPM) 25%	到达率 25%	人口数(万) 15%	汽车保有量(万) 15%	综合指数					
1	城市A	89	40288	90	30	91	69	100	2280	30	262	80
2	城市B	350	40542	93	19	73	52	44	1055	44	230	75
3	城市C	90	38488	82	56	76	54	87	2069	100	820	73
4	城市D	80	32570	100	18	93	66	29	681	27	141	73
5	城市E	71	29626	64	28	66	71	16	1413	45	214	71
6	城市F	88	35704	60	30	100	71	27	880	44	227	70
7	城市G	85	34456	30	60	71	51	54	1275	46	241	57
8	城市H	67	27284	26	71	76	94	60	1418	59	304	57
9	城市I	89	36322	39	46	77	55	54	816	30	159	56
10	城市J	72	26400	80	36	63	44	33	786	24	125	51

二、频率：选台复合化指标

频率的选择是制定广播媒体策略中最重要的一环。传统的选台指标往往依赖于收听到达表现，以收听率或到达率作为选台的单一指标。假设在某一市场，收听表现最好的为a频率，该频率的到达率在最高的时候可以是第二位b频率的2倍，按照常规的选台方针，肯定“非a莫属”；但这种“肯定”一定是正确的吗？a频率拥有强势的收听表现，其刊例价也同样“一骑绝尘”，满档加收更成为常态，在此溢价基础

上，以CPRP而论，a频率不再具备绝对优势；另外，a频率电台主力收听群众与投放品牌的受众重合度较低，而b频率无论从听众集中度还是电台属性上均具有较高的契合度，更重要的是b频率是省台，a频率为市台，从辐射范围上来说，b频率远胜a频率。综合以上因素，频率选择归属不言而喻。远誉选台复合化指标，包含到达率、收听点成本（CPRP）、电台属性、覆盖区域、受众集中等，通过对CSM多样、实时、海量及高价值的广播数据的筛选、收集及综合考量，将选台精准度进一步提升（图1）。

图1 远誉选台复合化指标



三、时点：选点精细化指标

点位的选择是广播媒体策略中容易被忽视的环节。在传统印象里，早晚高峰为黄金点位，而套播广告则是以逸待劳的绝佳选择。但随着广播广告客户的含金量不断提升、广播节目品牌化的深入、受众收听习惯的改变等诸多趋势，“传统”需要改革。远誉是这一改革的先行者，制定选点精细化指标，该指标包含了分时收听率、节目属性、广告环境、受众行为、CPRP及预算等诸多因素（图2）。

以黑龙江都市女性广播频率为例，首先该频率的常年收听最高峰不是出现在上下班时间，而是出现在下午4点半至6点，造成“黄金点位偏移”的原因来自于台柱式节目《叶文有话要说》，该档节目主打“感情牌”，以探讨情感、婚恋话题为主，高居黑龙江所有频率，所有时段，所有节目的第一位，缔造了收听神话，拥有一大批的女性忠实听众，所以，如

果对时点进行深度营销的话，该节目的广告开口可作为以女性为购买主导的产品的绝佳选择。

除考虑到以上所涉及的分时收听率及节目属性之外，远誉还会把广告环境作为重点考量因素，其一为广告播出时段整体广告环境，若某一时段已经被医药、卖场促销类广告所充斥，品牌继续入驻该时段，势必会稀释品牌的格调和好感度；其二为广告播出时段的竞品情况，若某一时段已经被竞争品牌率先并强势占领，品牌选择和竞品在声量“火拼”并非明智之举，大量的预算损耗未必换来声势上的压制，可采用“田忌赛马”战术，避其优势，攻其软肋。

另外，远誉在选台策略中更创新引入了“受众行为”这一数据。远誉在与电商类客户合作过程中，收集并整理大量数据发现，电商类受众的消费高峰点出现在早10点和晚9点；整合以上数据结合广播收听数据，可以得出早9点至10点，晚7点至9点是电商类广告于广播平台传播的黄金时点，可以有效提升网络购物这一即时消费行为的转换。从这一结果不难发现，相对传统的早晚高峰，黄金时点出现了小小的偏移。但也正是这小小的偏移，既节省了预算又提升了点位的有效性。

每个时点的收听率、CPRP、刊例有高低，每个时点的节目风格、广告环境、触及受众有差异，每个时点的广告版本、传播内容、预算分配有不同，每个时点都具备深度经营的价值。基于此，远誉的选点精细化指标是在传播碎片化大背景之下，深度经营“每个时点”的数据指标，也是充分重视CSM数据的指标，更是在海量数据中发现“珍珠”的指标。

图2 远誉选点精细化指标

远誉选点精细化指标



四、验证：精益求精式效果验证

每个策划制作完成后，远誉还会利用CSM数据进行效果预测以验证策略的正确性。

我们身处大数据时代，宏观大局，微观数字，让广告方案策略做的更好更精准，使效果得到进一步的提升。大数据，抑或小数据，归根结底不过是时代的一个名字。数据让我们减少失败，但创造让我们走向成功。不管是最好的时代，还是最坏的时代，大时代还是小时代，唯有勇于前瞻、敢于创造、不断超越并真正实行才能立于行业的前端。让我们共同挖掘，共同发现，共同探讨，携手开拓，共创辉煌。