

**Research Assessment Exercise 2020**  
**Impact Case Study**

**University:** City University of Hong Kong  
**Unit of Assessment (UoA):** 29

**Title of case study:** Improving Audience Measurement Methods for the TV and Online Media Industries

### 1. Summary of the impact

Professor Zhu's research on the generation and application of TV audience measurement (TAM) data has been used by at least three media research firms, 50+ TV stations and six online media websites. As illustrated by the six papers in Section 3 below, the research deals with how to draw random samples from online media (Zhu et al., 2011), how to integrate multi-source data (Lu & Zhu, 2014), how to merge big and small data (Zhu, 2016), how to harvest social media data (Liang & Zhu, 2017), how to measure mobile data (Zhu et al., 2018), and how to perform user analytics (Zhu et al., 2019). These findings have enabled TV and online media companies to better tailor their content to their viewers' requirements; advertising agencies and their clients to obtain optimal returns from their TV/online advertising spending; and several hundred million online and offline TV/media viewers around the world to watch more of the programmes they want to watch, when they want.

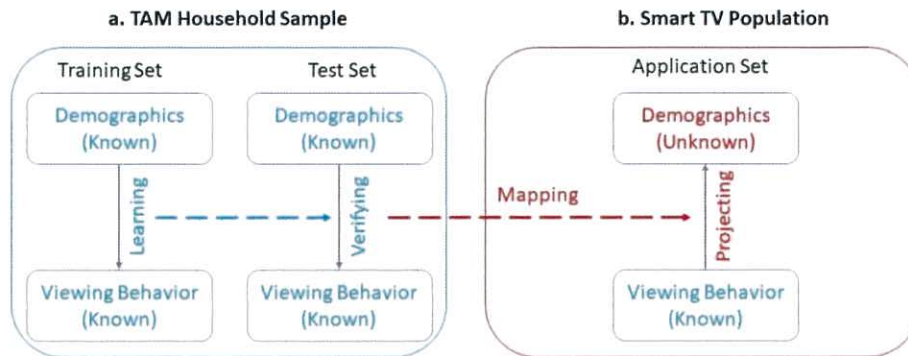
### 2. Underpinning research

TAM data register who watches what programmes at what times and for how long. As such, TAM data are the 'hard currency' in which the TV/online media industries operate. For example, TV stations use TAM data to schedule programmes and maximise audience size, whereas advertisers and their representatives (i.e., advertising agencies) use the data to design relevant TV commercials and select the optimal time slots for showing them. Members of the public benefit from these TAM-based operations by having access to the desired programmes at the right time. In short, all parties in the TV/online media industries require TAM data that are i) accurate, ii) readily available and iii) affordable to collect. Prof. Zhu's research has focused on improving the provision of these data.

As a seasoned media scholar, Prof. Zhu has been conducting basic and applied audience research for nearly four decades. For the first three decades, he worked on traditional TAM data in [REDACTED] which enabled him to develop first-hand insights into the traditional TAM methodology and establish a variety of connections with TV and online media companies. In 2008, Prof. Zhu expanded his research to online audience measurement through the Web Mining Lab (<http://weblab.com.cityu.edu.hk>), an interdisciplinary collaboration unit he created to conduct computational social science projects, including the integration of small and big TV/online audience data.

In non-technical terms, Prof. Zhu and his team use a variety of machine learning algorithms to *learn* and *verify* the relationships between the demographic characteristics (e.g., age, sex, etc.) of individual users and their viewing/browsing behaviour, both of which are included in the small data (see Fig. 1a), and then *map* the learned and verified relationships with the big data (Fig. 1b).

**Figure 1. Illustration of Multisource Data Integration**



Although this approach is intuitive and familiar to TAM researchers in the TV/online media industries, there are a series of technical challenges in determining the exact models, metrics and algorithms for deriving accurate results rapidly and at scale (i.e., involving billions or even trillions of records). Prof Zhu and his interdisciplinary team (consisting of media scholars, social psychologists, computer scientists, data visualisation experts, etc.) have combined their diverse expertise to find relevant solutions to these data integration problems.

### 3. References to the research

- R1** Zhu, J. J. H., Mo, Q., Wang, F., & Lu, H. (2011). A random digit search (RDS) method for sampling of blogs and other web content. *Social Science Computer Review*, 29(3), 327- 339.
- R2** Lu, H., & Zhu, J. J. H. (2014). Using single source data to better understand user-generated content (UGC) behavior. In *Advances in Social Networks Analysis and Mining (ASONAM)*, 2014 IEEE/ACM International Conference on (pp. 790-795). IEEE.
- R3** Zhu, J. J. H. (2016, in Chinese). Integrations of big data and sampled small data. *Rating China*, 12, 24-29.
- R4** Liang, H., & Zhu, J. J. H. (2017). Big data, collection of (social media, harvesting). In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods*, Wiley & Sons.
- R5** Zhu, J. J. H., Chen, H. X., Peng, T. Q., Liu, X. F., & Dai, H. X. (2018). How to measure sessions of mobile phone use. *Mobile Media and Communication*, 6(2), 215-232.
- R6** Zhu, J. J. H., Guan, L., Zhou, Y. X., Hou L., Shen, A. Q., & Lu, H. (2019). Applying user analytics to uses and effects of social media in China. *Asian Journal of Communication*, 29(3), 291-306.

### 4. Details of the impact

#### 4.1 Knowledge Dissemination

Prof. Zhu has communicated his research to interested users in the TV and online media industries through a number of channels such as workshops and trade publications. These include 12 workshops and talks attended by 2000 professionals from three industries in seven countries or regions. These events have enabled Prof. Zhu to demonstrate applications of classic

quantitative social science methods (e.g., surveys, experiments and statistical analysis) and emerging computational methods (e.g., log analytics, text mining and machine learning) to professionals in the TV, online media and online advertising industries [S1].

#### 4.2 Industry Adoption

The methods and expertise generated through Prof. Zhu's research are invaluable to a number of actors in the TV, online media and advertising industries. For example, Prof. Zhu's methods have been applied by the media research company [REDACTED] which is a primary provider of TV/online audience research [REDACTED] in a survey of more than 56,500 households, representing the habits of 1.3 billion TV viewers and 1.2 billion radio listeners [REDACTED] and 6.55 million viewers [REDACTED] (see S1). Prof. Zhu's methods have also been used by [REDACTED] the [REDACTED] branch of a world leading media research firm that studies consumer behaviour in more than 100 countries (S2), and [REDACTED] a global leader in interactive entertainment media employing over 700 people with offices in [REDACTED] many other cities around the world (S2). In a letter to Prof. Zhu, [REDACTED] stated that 'your expertise and insights on various ways to integrate TV and digital audience measurement and big data/machine learning have a wide range of impact on the perceptions and practices of media research industry' (S2). Similarly, [REDACTED] Business Intelligence Director [REDACTED] described [REDACTED] decade long collaboration with Prof. Zhu on audience measurement methods as 'helping us improve our services, content production quality and integration of traditional TV and online media to more than 6 million viewers [REDACTED]' (S3). In summary, as stated by [REDACTED] [REDACTED] in his letter about Prof. Zhu's contributions to [REDACTED] seminar on big data and communication, "his insights about big data has helped the audience a lot in understanding the rapid changes in communication studies" (S4).

#### 4.3 Industry Collaboration

Prof. Zhu has established deep relationships with a number of companies in the form of collaborative applied research projects and joint research labs. Two joint labs are currently in operation: one formed in collaboration with [REDACTED] and the other established with [REDACTED]. Both of these companies are industry leaders in social media data analytics. For example, Prof. Zhu and [REDACTED] collaborated to devise a sampling method (called Random Digit Search or RDS) for social media research (see R1). Based on the collaborative research, they are currently applying for a [REDACTED] Joint-Project to establish an artificial intelligence platform for creative and cultural industries (see an excerpt of the application in S5). Similarly, Prof. Zhu (along with other colleagues [REDACTED] [REDACTED] and [REDACTED]) are currently collaborating on a large-scale project to develop deep-learning based recommendation systems for commercial clients.

### 5. Sources to Corroborate the Impact

[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]