Annals of Tourism Research 47 (2014) 77–95

Contents lists available at ScienceDirect

Annals of Tourism Research

journal homepage: www.elsevier.com/locate/atoures

Research Notes

Tourism subfield identification via journal clustering

Yulan Y. Yuan [a], Yuen-Hsien Tseng [b,*], Chun-Yen Chang [b]

[a] Jinwen University of Science and Technology, Taiwan, ROC
[b] National Taiwan Normal University, Taiwan, ROC

Recently there are growing interests in examining subject areas and research trends in the hospitality and tourism fields (Hall, 2011; Howey, Savage, Verbeeten, & Van Hoof, 1999; McKercher, Law, & Lam, 2006; Park, Philips, Canter, & Abbott, 2011). Each study must decide which specific set of journals to include or exclude for data analysis. Such decision affects how research fields could be perceived, and could also alter the conclusions drawn from the set (Tseng & Tsay, 2013). Yet, the question about inclusion of which set of journals remains unexamined. This research note reports an evidence-based investigation by proposing the use of bibliographic coupling as an objective way for journal selection. By doing so, we attempted to answer this question from an epistemological perspective.

Previous methods to choose journals for Tourism (or Hospitality) analysis relied on expert judgment. The experts made their judges either based on various journal ranking systems, such as peer assessment (McKercher et al., 2006), or based on journal titles and their mission statements (i.e., Cheng, Li, Petrick, & O'Leary, 2011; Kim, Savage, Howey, & Van Hoof, 2009). However, both methods are somewhat subjective, and therefore different sets of journals may be used in different studies.

The solution we proposed is to apply the bibliographic coupling (BC) method to delineate the possible subfields in a broader set of candidate journals. BC is a similarity concept that could be used to quantify the relatedness of knowledge base between two journals based on the common knowledge source they cite (Small & Koenig, 1977). Given that the set of references cited by the articles published in journal W (X) be denoted by W_s (X_s) and the common references cited in both journals be denoted by the intersection: $W_s \cap X_s$, the topical relatedness between the two journals can then be quantified by various similarity measures, such as the Dice coefficient (Salton, 1989): $2*|W_s \cap X_s|/(|W_s| + |X_s|)$, say $2*5/(10 + 15) = 0.4$, where a set enclosed by vertical bars denotes the number of references in the set. With the pairwise similarities, the journals in a given set can be clustered for subfield delineation. Taking Fig. 1 as an example, this method regards each journal as a singleton cluster at first. It then groups the most similar pair of clusters into a larger cluster. The same grouping rule applies repeatedly to the remaining clusters and newly created ones. The overall result is called a dendrogram if visualized in a tree structure. Another clustering method called multi-dimensional scaling (MDS) based on the same pairwise similarities could yield the result in a topic map for cross validation (Tseng & Tsay, 2013).

* Corresponding author. Tel.: +886 2 77345535; fax: +886 2 23222009.
E-mail addresses: yoyo@just.edu.tw (Y.Y. Yuan), samtseng@ntnu.edu.tw (Y.-H. Tseng), changcy@ntnu.edu.tw (C.-Y. Chang).
URL: http://web.ntnu.edu.tw/~samtseng/ (Y.-H. Tseng).

Ref. ID	Samples of Cited References from 37 Journals	Times Cited
1	FORNELL C, 1981, J MARKETING RES, V18, P39	302
4	Nunnally JC, 1978, PSYCHOMETRIC THEORY	192
7	Patton MQ, 2002, QUALITATIVE RES EVAL	171
10	Glaser B, 1967, DISCOVERY GROUNDED T	145
11	Lincoln Y, 1985, NATURALISTIC INQUIRY	135
19	Bagozzi R. P., 1988, J ACADEMY MARKETING, V16, P74	120
22	Zeithaml VA, 1996, J MARKETING, V60, P31	114
…	…	…
293096	Iroegbu H, 2002, TOURISM ANAL, V6, P155	1

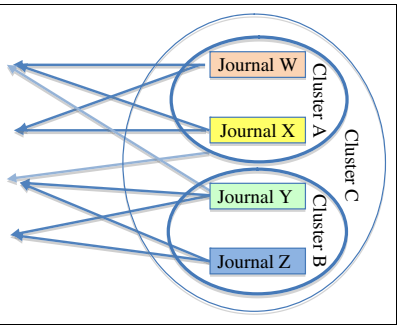

Fig. 1. A demonstration of how the journals are grouped together (clustered) based on journal bibliographic coupling.

In Journal Citation Reports (JCR), tourism journals have been grouped together with others to establish a subject category: 'Hospitality, Leisure, Sport & Tourism' (HLST). Journals indexed in HLST were the most frequently selected journals used in previous studies. Therefore, to address this question, the bibliographic data of 37 journals in the HLST category from 2008 to 2012 were used and downloaded from Web of Science on Feb. 20, 2013. They were analyzed and clustered by the freeware at: http://goo.gl/29co12. There were 7420 journal articles having 293,096 cited references in total (examples can be seen in Fig. 1). The journal clustering resulted in five major clusters. Their dendrograms are shown in Fig. 2 and their spatial distributions in terms of topical relationships are illustrated in Fig. 3. These five clusters are manually labeled as: Tourism (ToR); Hospitality (HoS); Sport Psychology (SP); Sociology of Leisure & Sport (SLS); and Sport Economics (SE). Three journals are not grouped with other clusters; they form singleton clusters due to their low similarity of knowledge base with others. These three journals are: *Journal of Hospitality Leisure Sport & Tourism Education* (JoHLSTE), *Journal of Tourism and Cultural Change* (JTCC), and *International Journal of the History of Sport* (IJHS).

From the dendrogram, it can be seen that most of the journals are clustered in a reasonable sense. Fig. 3 verifies the clustering, as most journals in the same clusters locate in vicinity. In addition, the topical relations among the five journal clusters are also illustrated. For examples, journals in HoS and ToR are closer, which means that hospitality and tourism journals share more knowledge bases

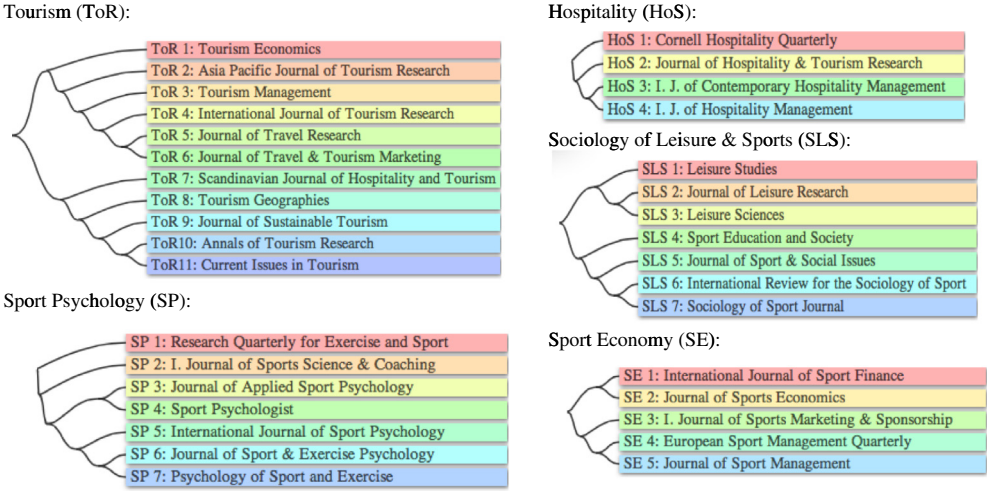

Fig. 2. Clustering results in dendrograms derived by journal bibliographic coupling.