MATCHMAKING IN CYBERSPACE: AN APPLICATION OF WEB-BASED BRAND IMAGE MEASUREMENT

Ivar Vermeulen, VU University Amsterdam,The Netherlands

ABSTRACT

Brand images reside in consumers' minds. Measuring brand images therefore typically involves surveying (many) consumers. This paper shows how brand images can be measured without surveying real-life respondents, but by extracting consumer-based brand knowledge from the Web. The paper subsequently shows how the resulting brand images can be compared and further analyzed in order to support strategic marketing decisions. As a case in point, a study is presented exploring co-brand opportunities among the top 50 world brands.

INTRODUCTION

Since Keller's conceptualization of consumer-based brand equity (Keller, 1993) the dominant paradigm in marketing research has it that brand meaning is in the eye of the beholder (Oakenfull, Blair, Gelb, & Dacin, 2000). In other words, consumer perceptions, not corporate intentions, determine brand images. In order to monitor brands – their own, as well as those of their competitors – companies spend billions of dollars a year (Lee, 1999) on picking consumer brains. Today, the Web provides new opportunities to carry out valid assessments of brand – and other – images at low cost and in a time-effective manner, without making use of real-life respondents. This paper combines techniques from linguistics, social psychology, informatics, graph theory, and marketing research into a versatile, valid, straightforward, and above all, cheap Web-based method to study brand images and their interactions. The presented method is complementary, and in some cases an alterative, to traditional brand image measurement methods. As a demonstration of its use, the paper will present a study on co-brand opportunities among the top 50 world brands, based on consumer perceptions of brand image similarities that were extracted from the World Wide Web.

In co-branding, a new product is created using two brand names (Leuthesser, Kohli, & Suri, 2003). Well-known examples of co-brands are the Smart car (by Swatch and Mercedes), the Senseo coffee-machine (by Douwe Egberts and Philips), and the Beertender home tap (by Heineken and Krups). Co-brands are usually the result of extensive research, not only by R&D divisions, but also by marketers. The latter are aware of the notion that co-brands can only be successful if consumers come to accept the alliance between the two parent brands. In their assessments of co-brands, consumers thus build on their perceptions about both parent brands, as well as on their perception about the differences and similarities between both brands (Simonin & Ruth, 1998).

This paper proposes to measure such consumer perceptions by means of information extraction from the Web. As a tool for information extraction, it proposes to simply use search engine Google, albeit in an advanced manner. Google has rapidly become the largest mediator of information on the Web: Most of us use Google in search of (popular) knowledge about a wide range of subjects on a daily basis. In addition, Google has become an important source of information in itself. For example, a high Google-ranking of a company's Web site bears information as to the status and reputation of the company. Another important component of this "Google image" consists of associations: e.g., former sports teams, student bodies, and employers. Such associations may have a strong impact on how we judge someone. In fact, connectionist models of meaning representation claim that we in general attach meaning to notions – e.g., people, or brands – by associating them to other notions

(Anderson, 1983). In connectionist models, notions are therefore represented as association networks, consisting of a focal node with links to associated nodes. Links may vary in strength, representing that some associations are stronger – i.e., more representative for, or more accessible from, the focal notion – than others. When a node in the network is activated, or primed – for example, when we see a car – associated notions, such as comfort, speed, danger, family, work, driving, roads, traffic, and possibly a few car brands, will be prompted by spreading activation – i.e., propagation through the association network (Collins & Loftus, 1975).

Building on these connectionist models, current marketing theory explicitly assumes brand images to be represented in consumer memory as association networks (e.g., Keller, 1993). Brand images are by definition individual mental representations, but due to collective experiences, such as exposure to advertising campaigns, they are often relatively homogeneous among consumers (Elliott, 1994).

In computational linguistics, the degree of association between two notions is given by their co-occurrence in applicable text corpora (Church & Hanks, 1990). To extract brand associations from the Web, we will likewise pursue patterns of co-occurrences between generic brand associations and individual brand names. For each brand, the extracted brand associations constitute a brand association network, or brand image. Each brand image thus comprises a unique profile based on the brand knowledge of millions of Web-authors.

Once extracted, brand images may be further analyzed and evaluated. In the current study, images are compared in order to identify potential co-branding candidates among the top 50 world brands. In the co-branding literature, co-brand success is generally linked to the degree of congruency – or fit, overlap, similarity – between both parent brands (Leuthesser et al., 2003). Simonin & Ruth (1998) distinguish two types of congruency: product fit – overlap of (functional) product-related attributes – and brand fit – overlap of their (more symbolic) brand-related attributes. Their findings suggest that consumer evaluations of co-brands increase when both types of fit increase.

In another study of co-brandings the authors find that consumer evaluations increase if consumers perceive the parent brands as mutually complementary. High overlap between the attributes of both parent brands implies little brand complementarity, and therefore poor consumer evaluations (Park, Jun, & Shocker, 1996). It should be noted that in their assessment of brand attribute overlap Park et al. consider only functional, not symbolic, attributes. Therefore, their analysis does not make clear whether, in order to obtain positive consumer evaluations of co-brands, the symbolic attributes of both parent brands should be either complementary or overlapping.

Yet another study has provided proof for a process assimilation of parent brand evaluations as a result of co-branding (Levin, 2002). For a co-branding effort to be beneficial to both parent brands, prior consumer evaluations of both parent brands should thus be similar. In sum, co-brands are potentially successful if (1) overlap of functional attributes is moderate (in order to ensure functional complementarity) (2) overlap of symbolic attributes is high (although the potential benefits of symbolic brand complementarity have never been explicitly tested), and (3) prior consumer evaluations of both brands are similar.

STUDY

The aim of this study is to demonstrate the use and application of extracting consumer brand knowledge from the Web by (1) extracting brand association networks for the top 50 world brands from the Web, (2) determining, for each pair of brands, similarity of their functional and symbolic associations, as well as in brand evaluation, and (3) determining suitable candidate brands to form a co-brand.

An association between two notions is indicated by their co-occurrence. Therefore, to establish step (1), we wrote a small computer program in programming language Perl that measured occurrences, pairwise co-occurrences, and co-occurrences with generic brand associations, for the top 50 world brands of the year 2005[1], on the Web. The program uses Google APIs, a free service made available to scholars by Google. Registered users may send up to 1000 automated queries a day to Google.

In our measurements, we distinguish three types of associations:

1. *Functional associations.* Product associations obviously are heterogeneous over product categories. The top 50 world brands represent many different product categories, and thus are associated with many product attributes, most of which do not apply to the majority of the other brands. In order to build our analysis on more generic functional associations, we make use of the notion that the degree to which the Web associates brand names to other brands to indicates functional substitutability (Vermeulen & Bruggeman, 2006). By using brands' associations to other top 50 brands as a proxy for functional associations, we abstract away from tangible brand-associations, but at the same time make the set of potential functional associations generic and exponentially smaller.
2. *Symbolic associations*. The symbolic meaning that consumers attach to a brand is often operationalized as brand personality. Prior research has distinguished five dimensions in brand personality – *sincerity, excitement, competence, sophistication, and ruggedness* – each of which is indicated by four more specific traits (Aaker, 1997). In addition to brand personality, we will use the non-evaluative dimensions of Osgood's semantic differential, i.e., *active/passive*, and *strong/weak*, to indicate symbolic meaning.
3. *Affective associations.* We use the degree to which the Web associates brand names to evaluative terms – such as *good* and *bad*, or *excellent* and *poor* – to indicate Web authors' affective associations with, the top 50 brands. Similar techniques for extracting affect from text corpora are common usage in computational linguistics (Turney, 2002).

Google allows searching for co-occurrences by interpreting the query "x y" as: "return all the pages that contain both the terms x and y". So, by typing "Amsterdam Hotels" in the Google search window, you are actually asking Google for pages that contain both the terms "Amsterdam" and "Hotels". Degree of co-occurrence can be computed in various ways. In the current study, we apply the *Jaccard association coefficient* $J(x,y)$ which denotes the overlap between sets relative to the size of both sets (Jackson, Somers, & Harvey, 1989). *J's* values are between 0 (denoting that x and y do not co-occur) and 1 (denoting that wherever x and y occur, they co-occur. So, if most Web pages about Amsterdam mention hotels, *and* if most pages about hotels mention Amsterdam, the co-occurrence between both terms is high.

To realize step (2) we need a measure for the similarity of two association networks. Graph theory provides such a measure by means of *structural equivalence* (Henderson, Iacobucci, & Calder, 2002). Two networks are structurally equivalent if they have ties of similar strength to similar nodes (Wasserman & Faust, 1995); structural equivalence thus in essence is a correlation measure, with values between –1 and 1. On the basis of structural equivalence (*SE*) we define a measure of structural complementarity (*SC*). Like *SE*, *SC* ranges between –1 (denoting that *SE* is three standard deviations removed from the mean of the population of *SE's*) and 1 (denoting that *SE* is equal to the average *SE*).

To realize step (3) we compute a measure for co-brand potential for each brand pair by averaging (a) structural complementarity of functional association networks, and structural equivalence of (b) symbolic, and (c) affective association networks.

[1] According to Business Week (BusinessWeekOnline, 2005).

Results

To measure the functional brand associations for each brand, we assessed brand-brand co-occurrences, amounting to (50·49)/2=1,225 measurements. To measure brands' symbolic associations, we conducted (50·24)=1,200 measurements. To measure the affective networks we conducted (50·6)=300 measurements. Finally, we assessed the individual occurrences of all 80 search terms used, amounting to a total of 2,805 measurements. To improve reliability, all measurements were conducted twice, rendering a total of 5,610 measurements. Test-retest reliability was r=0.996 (p<0.01). A small comparison (N=100) of the Google results by those of alternative search engine Altavista yielded a high correlation as well (r=0.987, p<0.01). Collection of data took place in the period from December 24, 2006 to January 1, 2007, at the rate of 750-1000 queries a day.

Row (b) in Table 1 shows the number of hits for the most visible 50 top world brands on the Web. *Google* turned out to be the brand with the highest visibility on the Web, followed – at significant distance – by *Microsoft* and *Sony*. There was no correlation between brand value as reported by Business Week 2005 and brand visibility on the Web (r=0.13, *ns*). Row (d) shows the ten brands with the highest scores on "Web-attitude", that is, the highest co-occurrences with positive affective associations, combined with the lowest co-occurrences with negative affective associations. Somewhat surprisingly, *Microsoft* has the highest score, followed by *IBM* and *Apple*. Row (e) shows the brand pairs with the highest and the lowest pairwise associations. Clearly, the strongest associations exist between brands operating in the same industry, which supports our choice for brand-brand associations as a generic proxy for functional brand associations. Row (f) shows the brands with the most and least similar functional associations. Again, high structural equivalence occurs within industries, with *Hewlett-Packard* and *Cisco* as the brand pair evoking the most similar associations. Note that the brand *MTV* is apparently functionally very different from many other brands in the top 50. Row (g) shows the brands with the most and least similar symbolic associations. The high similarity between *Toyota* and *BMW* is somewhat unexpected; both brands would seem to evoke other symbolic references in consumers. As was the case for functional networks, high degrees of symbolic structural equivalence were found within industries. Symbolically, *Google* turns out to be one-of-a-kind. With regard to affective associations (row (h)), *Pfitzer* and *Kellogg's* have the most similar scores, and *Nokia* has the most unique profile.

Our measure for co-brand potential – crediting high symbolic and affective similarity, and functional complementarity – renders *BMW* and *Nike* as the most promising pair of brands (row (i)). As it turns out, they fit the profile of symbolically and affectively similar brands with complementary functionality best. *Nike* shows co-brand potential with some other top 50 brands as well. *Google*, probably because it has distinct symbolic associations, shows generally poor co-brand potential. Finally, if we explore co-brands with symbolic complementarity, computations put *Disney* and *Siemens* forward as the most promising candidates (row (j)). They fit the profile of being complementary both in terms of functionality and symbolic meaning.

As an example of a Web-extracted brand image, Figure 1(a) shows the (combined) symbolic and affective image of *Coca-Cola*. Figure 1(b) shows the image similarities between the top 10 world brands.

Finally, *within* pairs of brands, structural equivalence between functional association networks strongly correlates with structural equivalence between affective association networks (r=0.62, p<0.01). Correlation with symbolic structural equivalence is much lower (r=0.18, p<0.01). Apparently, symbolic and functional equivalence are indeed complementary measures of image similarity.

DISCUSSION

This paper showed that images of 50 international brands could be fruitfully extracted from the Web. Subsequently, the paper showed how the extracted brand images could be compared and used for brand management purposes. The 5660 measurements on which the study was based, representing brand associations made on billions of international Web pages, were collected in less than a week by means of an fully automated and free data collection tool, and involved little human intervention, which demonstrates that Web-based research might be a feasible complement, and sometimes even alternative, to traditional – e.g., survey-based – research. Although the current study focuses on a marketing application, other research fields in which the interaction between public images plays a crucial role in designing strategies, campaigns or interventions, may benefit from Web based research methods as well. Obvious candidate fields are health communication, inter-group communication, and political communication.

Rank	*(a) Web visibility*	*(b) Hits (x1000)*	*(c) Value (xM$)*	*(d) Web attitude*	*(e) Strength of brand/brand associations*		*(f) Functional structural equivalence*		*(g) Symbolic structural equivalence*		*(h) Affective structural equivalence*		*(i) Co-brand potential*		*(j) Co-brand Potential w/ symbolic complementarity*	
1	Google	42800	8461	Microsoft	IBM	Intel	Hewl-P.	Cisco	Toyota	BMW	Pfizer	Kellogg's	BMW	Nike	Disney	Siemens
2	Microsoft	24500	59941	IBM	IBM	Oracle	Dell	UPS	JPMorgan	Goldm-S	Morgan-S	JPMorgan	Ford	UPS	Ford	Google
3	Sony	23600	10754	Apple	IBM	Apple	UPS	Google	Merrill-L	JPMorgan	UPS	Apple	HSBC	Ikea	BMW	Harley-D
4	Dell	21500	13231	Google	Merrill-L	Morgan-S	Citi	UBS	IBM	Intel	JPMorgan	Kellogg's	Louis-Vui	Gucci	Dell	MTV
5	Samsung	17500	14956	Intel	Intel	Apple	HSBC	UBS	Nokia	Samsung	Pfizer	JPMorgan	Pepsi	Nike	Ford	Nintendo
6	UPS	16750	9923	Oracle	Oracle	SAP	Canon	Nintendo	BMW	Nintendo	Pfizer	UBS	CocaCola	Ikea	Ford	MTV
7	Ford	13850	13159	UPS	IBM	Cisco	Hewlett-P	Oracle	Merrill-L	Morgan-S	Merrill-L	Goldm-S	Ameri-Ex	Nike	BMW	Louis-Vui
8	Apple	13750	7985	Sony	Merrill-L	Goldm-S	Kellogg's	UBS	McDonald'	Pepsi	Morgan-S	Kellogg's	Gillette	Nike	McDonald'	Gucci
9	Canon	12550	9044	Nokia	IBM	SAP	Citi	Goldm-S	Toyota	Honda	Pfizer	Morgan-S	CocaCola	Nike	Nike	Kellogg's
10	Nokia	11750	26452	GE	Merrill-L	JPMorgan	Intel	Dell	Coca-Cola	Pepsi	Merrill-L	Morgan-S	BMW	Ikea	BMW	Budweiser
...					...	...	...	...	...	...	...	...	...	...	...	...
1216					Nike	Google	Morgan-S	MTV	Coca-Cola	Google	Disney	Novartis	Goldm-S	Google	Sony	Goldm-S
1217					Heinz	Nintendo	JPMorgan	MTV	Microsoft	Citi	Nokia	Goldm-S	Sony	Goldm-S	Sony	Kellogg's
1218					Samsung	Gap	Merrill-L	MTV	Morgan-St	Google	Nokia	UBS	Sony	Morgan-S	Marlboro	Google
1219					GE	Dell	UBS	MTV	Goldm-S	Google	Nokia	Kellogg's	Sony	Kellogg's	Goldm-S	Google
1220					Sony	Gap	Citi	MTV	Microsoft	Kellogg's	Disney	JPMorgan	CocaCola	Google	Coca-Cola	Google
1221					Heinz	Gucci	CocaCola	Nokia	Google	Harley-D	Nokia	Citi	Merrill-L	Google	Merrill-L	Google
1222					Budweiser	SAP	Novartis	MTV	Microsoft	Marlboro	Disney	Morgan-S	Morgan-S	Google	Morgan-S	Google
1223					Samsung	Heinz	Goldm-S	MTV	Google	Kellogg's	Nokia	JPMorgan	Google	Kellogg's	Google	Kellogg's
1224					Oracle	Gucci	CocaCola	Sony	JPMorgan	Google	Disney	Merrill-L	JPMorgan	Google	JPMorgan	Google
1225					Ameri-Ex	Google	Pfizer	MTV	Citi	Google	Nokia	Morgan-S	Citi	Google	Citi	Google

Table 1: Top 50 world brands or brand pairs ordered by (a) Web visibility, (d) Web attitude, (e) mutual association, (f) functional, (g) symbolic, and (h) affective structural equivalence, and (i; j) two measures for co-brand potential .

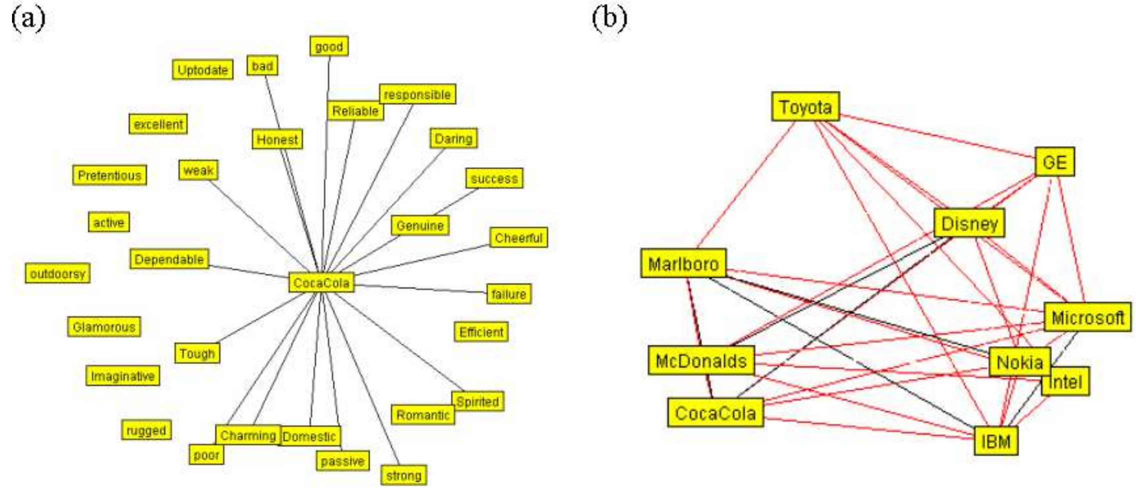

Figure 1: (a) Combined symbolic and affective network of Coca-Cola; (b) structural equivalence for top 10 world brands.

REFERENCES

Aaker, J. (1997), "Dimensions of Brand Personality", *Journal of Marketing Research, 34*(3), 347-356.

BusinessWeekOnline. (2005), *Top 100 Global Brands Scoreboard*. Retrieved December 1, 2006, from http://bwnt.businessweek.com/brand/2005/

Church, K., & Hanks, P. (1990), "Word association norms, mutual information, and lexicography", *Computational Linguistics, 16*(1), 22-29.

Elliott, R. (1994), "Exploring the symbolic meaning of brands," *British Journal of Management, 5*(2), 13-19.

Henderson, G. R., Iacobucci, D., & Calder, B. J. (2002), "Using network analysis to understand brands," *Advances in Consumer Research, Volume Xxix, 29*, 397-405.

Jackson, D., Somers, K., & Harvey, H. (1989), "Similarity Coefficients: Measures of Co-Occurrence and Association or Simply Measures of Occurrence?," *The American Naturalist, 133*(3), 436-453.

Keller, K. L. (1993), "Conceptualizing, Measuring and Managing Customer-Based Brand Equity," *Journal of Marketing, 57*(1), 1-22.

Lee, J. (1999), "Research Means Business," *Campaign, 10*(37).

Leuthesser, L., Kohli, C., & Suri, R. (2003), "2+ 2= 5? A framework for using co-branding to leverage a brand," *Journal of Brand Management, 11*, 35-47.

Levin, A. M. (2002), "Contrast and assimilation processes in consumers' evaluations of dual brands," *Journal of Business and Psychology, 17*(1), 145-154.

Oakenfull, G., Blair, E., Gelb, B., & Dacin, P. (2000), "Measuring brand meaning," *Journal of Advertising Research, 40*(5), 43-53.

Park, C., Jun, S., & Shocker, A. (1996), "Composite Branding Alliances: An Investigation of Extension and Feedback Effects," *Journal of Marketing Research, 33*(4), 453-466.

Simonin, B. L., & Ruth, J. A. (1998), "Is a company known by the company it keeps? Assessing the spillover effects of brand alliances on consumer brand attitudes." *Journal of Marketing Research, 35*(1), 30-42.

Turney, P. (2002), "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 417–424.

Vermeulen, I., & Bruggeman, J. (2006), ""Automated Longitudinal Data Collecton on the Web: Competition between Search Engines, 1993-2000," *ASSR Working Paper, 06*(03).

Wasserman, S., & Faust, K. (1995). *Social network analysis*: Cambridge Univ. Press.